weather@home - development and validation of a very large ensemble modelling system for probabilistic event attribution

N. Massey[*ab], R. Jones[da], F. E. L. Otto[a], T. Aina[c], S. Wilson[d], J. M. Murphy[d], D. Hassell[f], Y. H. Yamazaki[eg], M. R. Allen[ae]

[a]*Environmental Change Institute, University of Oxford, Oxford, OX1 3QY*
[b]*Smith School for Enterprise and the Environment, University of Oxford, Oxford, OX1 2BQ, UK.*
[c]*Oxford e-Research Centre, University of Oxford, Oxford, OX1 3QG*
[d]*Met Office Hadley Centre, Fitzroy Road, Exeter, EX1 3PB, UK*
[e]*Atmospheric, Oceanic and Planetary Physics, University of Oxford, Oxford, OX1 3PU, UK.*
[f]*Department of Meteorology, University of Reading, Earley Gate, Reading, RG6 6BB, UK*
[g]*School of Geography, Politics and Sociology, Newcastle University, Newcastle upon Tyne, NE1 7RU, UK*
[*]*Correspondence to: Environmental Change Institute, University of Oxford, Oxford, OX1 3QY.*
Email: neil.massey@ouce.ox.ac.uk

Demonstrating the effect that climate change is having on regional weather is a subject which occupies climate scientists, government policy makers and the media. After an extreme weather event occurs, the question is often posed, "Was the event caused by anthropogenic climate change?" Recently, a new branch of climate science (known as attribution) has sought to quantify how much the risk of extreme events occurring has increased or decreased due to climate change. One method of attribution uses very large ensembles of climate models computed via volunteer distributed computing. A recent advancement is the ability to run both a global climate model and a higher resolution regional climate model on a volunteer's home computer. Such a setup allows the simulation of weather on a scale that is of most use to studies of the attribution of extreme events.

This paper introduces a global climate model that has been developed to simulate the climatology of all major land regions with reasonable accuracy. This then provides the boundary conditions to a regional climate model (which uses the same formulation but at higher resolution) to ensure that it can produce realistic climate and weather over any region of choice. The development process is documented and a comparison to previous coupled climate models and atmosphere only climate models is made. The system (known as weather@home) by which the global model is coupled to a regional climate model and run on volunteer's home computers is then detailed. Finally, a validation of the whole system is performed, with a particular emphasis on how accurately the distributions of daily mean temperature and daily mean precipitation are modelled in a particular application over Europe. This builds confidence in the applicability of the weather@home system for event attribution studies.
Copyright © 2014 Royal Meteorological Society

Copyright © 2014 Royal Meteorological Society
Prepared using qjrms4.cls

Q. J. R. Meteorol. Soc. **00**: 2–22 (2014)

1. Introduction

climate*prediction*.net is the largest ensemble climate modelling experiment to date. Using distributed volunteer computing (Anderson 2004), over 126 million years of coupled atmosphere-slab ocean (Stainforth *et al.* 2005), coupled atmosphere-ocean (Frame *et al.* 2009) and medium-resolution atmosphere only (Pall *et al.* 2011) models have been completed.

While very large ensembles of low-resolution climate models have increased our understanding of key feedbacks and large scale processes in the climate system (e.g. Stainforth *et al.* (2005)), the real power of very large ensembles lies in the potential to simulate extreme events which, by their definition, are rare. To obtain statistics on the magnitude and frequency of occurrence of extreme weather events, large numbers of simulations of possible weather under the same climate conditions need to be run to assess the odds of such events. However, as most extreme weather events occur on a relatively small spatial scale high-resolution global or regional models are required to realistically capture impact relevant extreme events.

Projects such as PRUDENCE (Jacob *et al.* 2007) and ENSEMBLES (van der Linden and Mitchell 2009) have produced multi-model ensembles of regional climate models (RCM) with boundary forcings sourced from difference global climate models (GCM). This has helped increase our understanding of the sensitivity of the climatology of the RCM to both its formulation and its boundary forcings. However, to truly understand the distribution of extreme weather events within these RCMs, and to determine whether they have shifted under anthropogenic climate change, requires an ensemble size in the tens of thousands, rather than in the hundreds.

weather@home utilises the climate*prediction*.net volunteer distributed computing network to compute very large ensembles of the HadRM3P regional climate model driven by the HadAM3P atmosphere–only global climate model (AGCM). HadAM3P is based on the coupled Hadley Centre GCM HadCM3 (Pope *et al.* (2000), Gordon *et al.* (2000)), which is, despite having had several successors, a remarkably good model in representing 20th century climatology (Solomon *et al.* (2007), Lei *et al.* (2013), Collins *et al.* (2001)).

In general, coupled–climate models show a westerly bias in the zonal winds across the Atlantic and corresponding blocking errors (Scaife *et al.* 2010). These errors are largely due to biases in the sea–surface temperatures (SST) in the North Atlantic (Scaife *et al.* 2011) and that in eliminating these biases the zonal winds and blocking errors are reduced. (Scaife *et al.* 2011) shows that these errors in SST can be removed by either increasing the resolution of the ocean in coupled–climate models or by forcing an atmosphere only model with observed SSTs. In this paper we show that, in comparison to HadCM3, forcing a higher resolution atmosphere only model with observed sea–surface temperatures (SSTs) and improving its physics produces a better represetnation of large scale weather events. The results of Scaife *et al.* (2010, 2011) thus suggest that the model will represent an improvement over coupled–climate models in general and HadAM3P's ability to simulate large scale events (Otto *et al.* 2012) indicate that it is a good model to use to force the boundaries of a higher resolution regional model.

Given the importance of the regional climatology of the GCM used by weather@home, the first part of this paper (Sections 2 to 3) describes the details of the development and validation of the model over a single historical run between 1961 and 1990. While previous studies have established that specific individual weather events are captured satisfactorily by the regional model (e.g. Massey *et al.* (2012), Sparrow *et al.* (2013), Rupp *et al.* (2012)), we show here that the improved resolution and physics of the driving model have improved the climatology of the global simulations. We demonstrate that with the increase of the resolution of the atmosphere component of HadCM3, HadAM3 simulates weather features that are not represented in the low-resolution version. This is especially the case for mid–latitudial climates where the representation of winter storm tracks is improved, which is important for the simulation of extreme precipitation events. However, the increase in the resolution leads to a loss of compensating errors and thus increases the temperature bias compared to the lower resolution version. We reformulate some of the sub–grid scale physics parameterisations and model chemistry which produces an improved representation of cloud cover which also, for example, results in a much improved simulation of temperature extremes.

These advancements in the representation of the immediate diagnostic variables such as temperature and precipitation, and also in the representation of the large-scale circulation, lead to confidence that the individual weather events that are realistically represented in HadAM3P are "right for the right reasons". That is, they are not a product of the cancellation of errors in the tuning process, but are instead the realistic representation of underlying physical processes. Thus, we conclude that HadAM3P is a good tool to analyse large-scale extreme weather events and provide boundary conditions to drive a regional climate model. In this study we use the HadRM3P regional model, which has essentially the same model formulation and vertical resolution as HadAM3P, but increases the horizontal resolution to either 50 or 25km.

The second part of this paper (Sections 4 to 5) details how the distributed computing infrastructure is used to run a large ensemble of both the GCM and RCM over the same historical period used in the validation of the GCM, 1961 to 1990. This ensemble is again validated with respect to observations, with the aim of determining the general suitability of using large ensembles of the model for studies of extreme weather events, without concentrating on a single event as in previous studies. It should be noted that this validation section is not only checking the suitability of the model, but the system as a whole, including the forcing files used, the initial condition perturbation and the method of using volunteer distributed computing to run the models.

This validation, analysis of the spread of the ensembles and consistency check between the GCM and RCM reveals that weather@home is indeed a good tool to investigate changes and drivers of extreme weather events especially in mid–latitudinal climates.

2. Model development

weather@home uses the HadAM3P atmosphere-only global climate model (AGCM) and a regional climate model (RCM) variant, HadRM3P, both from the UK Met Office Hadley Centre. These models are based upon the atmosphere component of HadCM3, a well documented and

widely used coupled ocean-atmosphere model described in Gordon *et al.* (2000). HadRM3P is the regional model used in the Providing Regional Climates for Impacts Studies project (PRECIS, Jones *et al.* (2004)), also originating from the UK Met Office.

HadAM3P and HadRM3P (henceforth HadAM3P/RM3P) share the same model formulation, with the only differences being the spatial resolution, timestep length and physical parameter values associated with length scales. HadAM3P is integrated with a 15 minute timestep, has 19 vertical levels and a horizontal resolution of $1.875°$ longitude and $1.25°$ latitude, which approximates to grid boxes of length ~ 150km at mid-latitudes and ~ 200km in the tropics. HadRM3P also has 19 vertical levels, but has a horizontal resolution of either 50km or 25km and a timestep of 5 minutes. HadAM3P's grid is defined as a regular latitude-longitude grid with regular poles, whereas HadRM3P employs a rotated grid, with artificial poles defined on a per-region basis so that the region of interest lies along the equator in the rotated grid. This ensures that each grid box in the region has approximately the same area. Hassell and Jones (1999) and Hudson and Jones (2002) show that the higher resolution of a RCM can produce more realistic weather features, such as tropical cyclones, that GCMs struggle to represent. Also, Denis *et al.* (2003) show, in a study where they derived coarse resolution boundary conditions from filtering high resolution simulations that, up to a resolution difference of 12 between an RCM and its driving data, the RCM can reproduce realistic high resolution weather features seen in the original simulation.

HadAM3P/RM3P is a grid-point model which solves equations of motion, radiative transfer and dynamics explicitly on the same scale as the grid. The atmospheric equations are a quasi-hydrostatic version of the primitive equations with full representation of the Coriolis force, as described in Cullen (1993). Other, mostly thermodynamic, processes which occur at the sub-grid scale are represented by physical parameterisations. The vertical resolution of HadAM3P/RM3P remains the same as HadCM3, although the horizontal resolution increases two-fold for HadAM3P and either 8-fold or 16-fold for HadRM3P.

2.1. Improvements to the HadCM3 model

The relatively low horizontal resolution of the atmospheric component of HadCM3 (denoted HadAM3) contributes to significant regional simulation biases which compromise inferences made about regional climate change from using the model. Increasing horizontal resolution substantially reduces some of these biases, notably in extra-tropical surface winds and temperatures during Northern Hemisphere winter. However, some other aspects of model performance are degraded because the increase in resolution upsets balances between compensating errors present at lower resolution. This can be reversed by making some significant changes and improvements to the physical parameterisations in HadAM3, primarily to improve the simulation of clouds and radiative fluxes while preserving the benefits of higher resolution. With the new model formulation, denoted HadAM3P, the primary surface variables of temperature, precipitation and surface pressure are simulated better than, or at least as well as, HadCM3 over all major continental regions.

Details of these improvements are shown in Figures 1–5 and outlined in Sections 2.3.1 and 2.3.2. Section 3 provides an overview of the potential effect of these improvements on driving regional models, in that it shows reductions in the root mean square error for a number of regions, when comparing the new model formulation (HadAM3P) to the previous formulation (HadCM3).

2.2. Description of the models and experiments

The models described here are based on the HadCM3 coupled model (Gordon *et al.* 2000). Its atmospheric component, HadAM3, forms the basis of the model development described below. HadAM3 is fully described by Pope *et al.* (2000) and is generally applied on a regular latitude/longitude grid of horizontal resolution $2.5\times3.75°$ with 19 vertical levels. This will be referred to as standard resolution. In an earlier version of the model, HadAM2b, Stratton (1999) found an improved simulation of the North Atlantic storm track in winter on increasing the horizontal resolution to $0.833\times1.25°$. In HadAM3, Pope and Stratton (2002) show that most of the improvement found at $0.833\times1.25°$ can be replicated at $1.25\times1.875°$ and so we choose this coarser resolution to limit the extra computing resources required. Pope and Stratton (2002) investigated the impact of increasing vertical resolution from 19 to 30 levels, finding mixed benefits. Moist and cold biases in the upper troposphere were reduced, but some aspects of the tropical climate were simulated less well due to a deterioration of the performance of the convection scheme. In view of this we decided not to increase vertical resolution in HadAM3P, although the potential improvements from doing so are being investigated as part of the strategy for developing new climate models in the Hadley centre (Johns *et al.* 2006).

In our $1.25\times1.875°$ version of HadAM3 (hereafter Hi-res HadAM3), the time step is 15 minutes for both dynamics and physics, c.f. 30 minutes at standard resolution. The physical parameterisations are identical to HadAM3, and the dynamical formulation is identical apart from resolution dependent adjustments required for the calculations of diffusion and gravity wave drag. HadAM3P consists of Hi-res HadAM3 augmented by a number of changes to the sub–grid scale physics and chemistry, listed below.

2.2.1. Calculation of cloud cover

In HadAM3 the cloud cover and cloud water content in a grid box are both calculated from a saturation variable q_c defined as the difference between total water (i.e. water vapour + liquid + ice) and the saturation vapour pressure (Smith 1990). When provided with observed grid box values of total water and temperature the Smith scheme reproduces observed cloud water contents quite well but underestimates cloud cover, based on data from stratocumulus regions and the upper troposphere (Wood and Field 2000). More generally, HadAM3 reproduces the effects of clouds on the global radiation budget quite well, but through a compensation of errors in which insufficient cloud cover tends to be offset by excessively high cloud optical thicknesses. One reason for this is that clouds are assumed to fill the entire volume of a model layer, neglecting the possibility of thin layers of cloud associated with sub-grid scale variations in cloud water in the vertical. We address this by introducing a modification in which

	HadCM3	HadAM3	Hi-res HadAM3	HadAM3P	HadRM3P
Model type	AOGCM	AGCM	AGCM	AGCM	RCM
Resolution	2.5x3.75°	2.5x3.75°	1.25x1.875°	1.25x1.875°	0.44x0.44°
Scale at mid–lat	300km	300km	150km	150km	50km
Vertical levels	19	19	19	19	19
Time step	30 mins	30 mins	15 mins	15 mins	5 mins
Improved physics	No	No	No	Yes	Yes

Table I. Summary of models discussed in this study. AOGCM = coupled atmosphere–ocean GCM, AGCM = atmosphere only GCM. Improved physics indicates that the model has the improvements to the sub–grid scale physics and chemistry detailed in Section 2. Scale at mid–lat is the width of a grid box at 45°N.

the cloud scheme is called for three sub-layers within each model grid box, calculating cloud cover from values of q_c for each sub-layer obtained by vertical interpolation. Areal cloud cover for the grid box is taken as the maximum of the values found in the three sub-layers, generally resulting in larger values (with corresponding reductions in optical thickness) compared with results from the standard parameterisation used in HadAM3.

2.2.2. Specification of relative humidity threshold for cloud formation, Rh_{crit}

It is assumed that the sub-grid scale distribution of q_c can be represented by a symmetric triangular function (Smith 1990) which depends on Rh_{crit}, the grid box mean relative humidity above which cloud begins to form. These assumptions allow cloud fraction (C) to be specified as a quadratic spline passing through the points (RH= Rh_{crit}, C=0), (RH=1,C=0.5), (RH=1+ Rh_{crit}, C=1). A fixed value of Rh_{crit} is specified for each model level in HadAM3: at standard resolution this ranges from 0.95 in the lowest layer to 0.7 for layers in the free atmosphere. However, Cusack *et al.* (1999) argued that the assumption that Rh_{crit} does not vary in time or with geographical location is unrealistic. Based on evidence from aircraft observations and high resolution analyses for numerical weather prediction Cusack *et al.* (1999) proposed that σ_{clim}, the standard deviation of q_c within a climate model grid box, can be parameterised in terms of σ_{3x3} the standard deviation of q_c over neighbouring grid boxes. Specifically

$$\sigma_{\text{clim}} = A(p) * \sigma_{3x3}$$

where $A(p)$ is a coefficient which varies with pressure (i.e. model level) but has no geographical or time dependence. Cusack *et al.* (1999) found that using this relation to predict σ_{clim} (and hence Rh_{crit}) led to reduced biases in cloud and relative humidity in the upper troposphere in the standard resolution version of the model. Here we assess this Rh_{crit} parameterisation in the higher resolution HadAM3, using values of $A(p)$ appropriate for a 150km grid (S. Cusack, pers. comm.).

2.2.3. Improved calculation of the radiative effects of convection

The parameterisation of convection in HadAM3 calculates a cloud fraction which is assumed constant between the diagnosed cloud base and cloud top. This approach takes no account of anvil clouds, leading to an underestimation

of high cloud of intermediate optical thickness and an overestimation of high, optically thick cloud (Ringer and Allan 2004). In order to rectify this, HadAM3P includes a set of empirical modifications developed by Gregory (1999), which are informed by basic observed properties of anvils as ice clouds which form in the presence of deep convection and tend to have their bases at the freezing level. When deep convection occurs, the modified scheme increases cloud fraction linearly with height from the freezing level to the cloud top to represent the anvil, and decreases cloud fraction to a constant value below the freezing level to represent the convective tower. Deep clouds are defined as those having their bases in the boundary layer and their tops above the freezing level. If convection is not diagnosed as deep then no change is made to the calculation of cloud fraction used in HadAM3.

In addition, simulated convective cloud water amounts are too large in HadAM3. The Gregory (1999) modifications address this by reducing the values of cloud water used in the calculation of radiative transfer. This is done partly by excluding convective precipitation from the water path (because rain drops are much less radiatively active than smaller cloud droplets), and also by introducing a scaling factor which accounts crudely for concentration of cloud water in a small fraction of the cloud associated with the convective updraught.

2.2.4. Additional minor changes

In HadAM3 the land surface is coupled to the underlying soil by a heat conduction term. This is appropriate for a bare soil surface but leads to an underestimation of the diurnal cycle for vegetated surfaces. In HadAM3P vegetated surfaces are assumed to be coupled radiatively to the underlying soil. This weakens the coupling between the ground and surface air temperatures (Best and Hopwood, 2001), leading to an improved diurnal cycle and the removal of unrealistic peaks in the frequency distribution of minimum temperatures associated with soil freezing in winter. HadAM3 occasionally simulates unrealistically high surface temperatures in arid regions. This occurs because the model only updates radiative fluxes every three hours, preventing rapid rises in surface temperature driven by strong solar heating from being simultaneously offset by increases in long wave cooling. This problem does not arise in non-arid areas where there is sufficient moisture to allow evaporation to limit the increase in temperature. In HadAM3P the upward surface long wave radiation flux is updated every model timestep (fifteen minutes), thus removing this unrealistic behaviour.

Prepared using qjrms4.cls

Precipitation is assumed to fall on a fraction of a model grid box in HadAM3. The specified fractions influence the land surface hydrology, including the partitioning of evaporation between surface evapotranspiration and free evaporation from a wet vegetated canopy (Dolman and Gregory, 1992). Values of 0.4 (for convective precipitation) and 0.5 (for large scale precipitation) were found to be appropriate for the spatial resolution of HadAM3P, based on results obtained by spatial aggregation of instantaneous precipitation fields from regional climate model simulations.

2.3. Experimental design of climate simulations and sensitivity tests

Three types of experiments are assessed in this Section. In order of increasing length these are: sensitivity tests of 18 to 42 months to check the effect of changing individual model components; short, decadal-scale, climatological tests for the period 1980-1990 to check the effect of multiple changes; and long climate simulations to provide a comprehensive assessment of model performance for the period 1961-1990. At the ocean surface, the lower boundary condition is provided by monthly values of global sea–surface temperature and sea-ice concentrations from the HadISST1 reconstruction of Rayner et al. (2003). The experiments presented below with the standard and high resolution versions of HadAM3 and the tests involving the changes to the large–scale cloud scheme and multiple physics changes all use this period. Only results from HadAM3P and HadCM3 use the standard 30 year climatological period. For the sensitivity tests, the first 6 months are used to spin the model up and is not considered in any subsequent analysis, and these are used to assess the seasonal or annual impact of individual physics changes such as changes in precipitation efficiency on cloud water.

2.3.1. Model response to changes in formulation

The initial change applied to HadAM3, a doubling of the horizontal resolution, significantly improves the realism of the winter storm tracks in both hemispheres with a pole–ward migration improving their position and reducing high polar pressure biases. In the case of the Northern Hemisphere (Figure 1), this reduces both a significant high latitude easterly bias and North Eurasian cold bias. With the full package of changes in HadAM3P (bottom panels of Figure 1) both of these biases are reduced further though with an increase in the area of warm bias over North–East Asia and North America. In the Southern Hemisphere summer Figure 1, the resolution increase has little impact but the other changes in HadAM3P warm Australia, either reducing cool or increasing warm biases, and Southern and equatorial Africa to reduce or remove cool biases.

Also apparent from Figure 1 is the impact of using observed SSTs and sea–ice. The second panel demonstrates that they are responsible for part of the improvements in the Northern Hemisphere winter–time circulation. This was confirmed in the results of a sensitivity experiment (not shown) where high resolution HadAM3 used SSTs and sea–ice from the parallel HadCM3 integration. In this experiment the improvements seen in Figure 1 panel 3 were not realised. This was due to the excessive sea–ice extents simulated by HadCM3 reducing temperatures in polar regions and leading to a high pressure bias. (Interestingly,

the opposite situation occurred in the previous Hadley Centre coupled model, HadCM2, where too little Artic sea–ice gave rise to the removal of a high pressure bias seen in HadAM2 and thus realistic storm tracks in HadCM2.)

In Northern Hemisphere summer the impact of resolution increase is not similarly neutral or beneficial. There are improvements in circulation (not shown) though the main impact is the worsening of warm biases over all of Eurasia and North America south of 50N (Figure 2). This develops because the increase in resolution upsets a balance of errors operating in HadAM3. The cloud scheme underestimates cloud cover (Figure 2), but this is partly compensated by other biases, notably excessive cloud water contents and an insufficiently vigorous hydrological cycle. Increasing horizontal resolution leads to an intensification of the hydrological cycle via stronger vertical motions resulting in reductions in atmospheric relative humidity (not shown) and more heavy precipitation events. This results in even lower cloud cover (Figure 2) leading to excessive surface solar heating and hence an increased surface warming which worsens warm biases seen in these regions in HadAM3. With the full package of changes in HadAM3P the negative bias in clouds is significantly reduced almost everywhere as are the Northern Hemisphere warm biases. However, due to the simulated clouds having more realistic (lower) optical thicknesses, those Northern Hemisphere regions with insufficient cloud still have warm biases. In other regions, tropical Africa and Australia in winter, these lower optical thicknesses result in reduced cool biases (Figure 2).

As a result of the increase in resolution, but more so from the other changes in HadAM3P, precipitation is reduced over most land areas. For most of the Northern Hemisphere summer this reduces wet biases (Figure 3) though also increases dry biases in west Asia and eastern North America which could also be contributors to the warm biases in these regions through lowering available soil moisture for evaporative cooling at the surface.

2.3.2. Effects of changing the model physics

The Rh_{crit} parameterisation introduces significant inhomogeneity into the the spatial distribution of Rh_{crit}. It gives higher values than the level–constant values of HadAM3 in most regions and significantly so in much of the stratosphere and tropical troposphere (not shown). This significantly changes cloud distributions in most regions, for example with large reductions in the lower tropical troposphere (Figure 4). It also reduces upper tropospheric moist biases which reduces a positive bias seen in the simulation amounts of the cirrus clouds. Another specific cloud change is a reduction in high-top optically thick midlatitude clouds, again reducing a positive bias. The impact of the change to the improved cloud fraction calculation is, as expected, mainly to increase cloud amounts (with the biggest signal in the lower troposphere at high latitudes) and to reduce cloud optical thicknesses. It has a large impact on low–top clouds, significantly increasing the medium thickness clouds in this category (and so removes nearly all of the HadAM3 error) as well as reducing the thick clouds which are also improvements (Figure 4). It also increases mid–top thin and intermediate thickness clouds which again are improvement though it does increase high–level thin clouds which is a degradation.

The effects of the Rh_{crit} parameterisation and the cloud fraction modification combined is to significantly
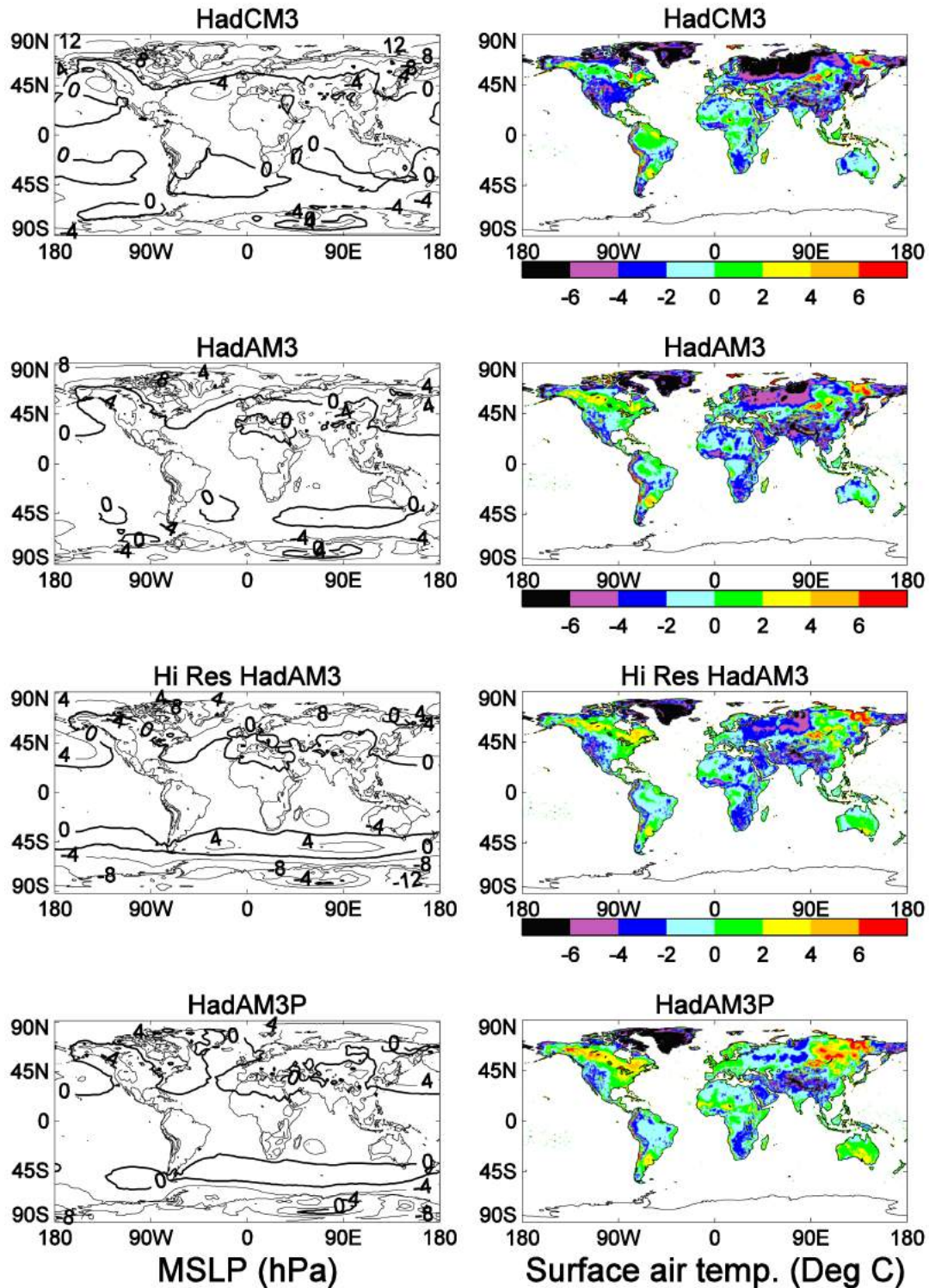
Figure 1. Errors in December–January–February (DJF) mean sea level pressure, left column, and surface air temperature, right column, for (from top to bottom) HadCM3, HadAM3, HadAM3 at double horizontal resolution and HadAM3P. The circulation biases are calculated with respect to the ERA15 reanalysis (Gibson *et al.* 1997) and the land-surface temperature biases with respect to the CRU-TS dataset (Mitchell and Jones 2005).

increase cloud cover outside the tropics to more realistic values. This provided a substantial cooling in summer, eliminating the Northern Hemisphere positive biases but introducing an overall cool bias and a negative top of the atmosphere radiation bias. The latter comes from a remaining positive bias in cloud optical thicknesses resulting from two factors, excessive cloud droplet number concentrations and cloud water/ice contents. The former resulted from assuming, in HadAM3, an unrealistically high

value for this quantity hence reducing the droplets effective radii and thus increasing the cloud brightness for given amounts of cloud water. This problem is overcome as a consequence of introducing a representation of the sulphur cycle, as one aim of this is to predict atmospheric aerosol concentrations. This then determines the concentration of cloud condensation nuclei (CCN) which control the size of cloud droplets (and thus cloud brightness, the first indirect effect of aerosols). With the sulphur cycle parameterisation
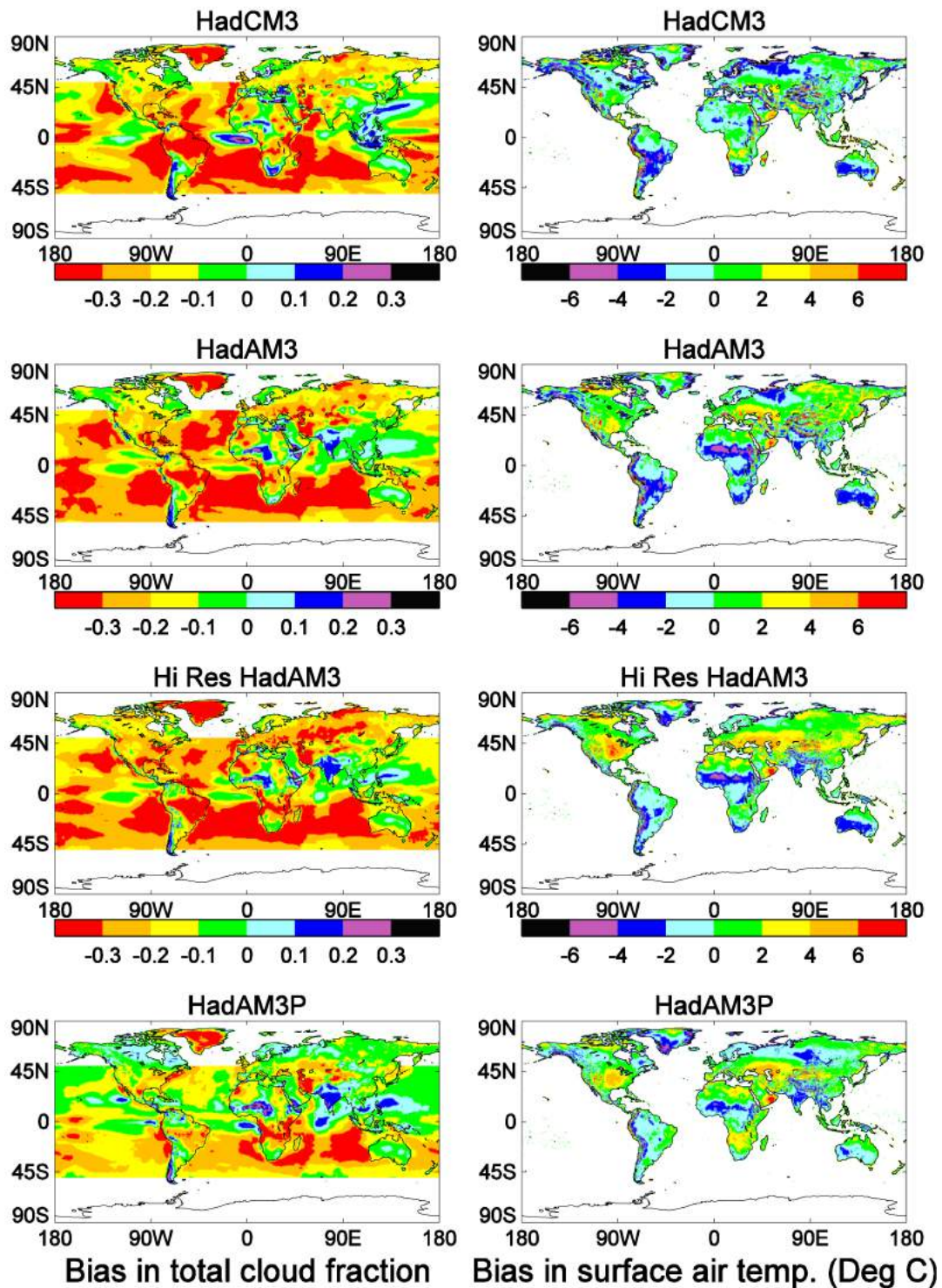
Figure 2. As Figure 1 but for June–July–August (JJA) cloud fraction, left column, and surface air temperature, right column. The cloud fraction biases are calculated with respect to data from the ISCCP (International Satellite Cloud Climatology Project) dataset (Rossow and Schiffer 1999) and the land-surface temperature biases with respect to the CRU-TS dataset (Mitchell and Jones 2005).

predicting realistic CCN concentrations (Jones *et al.* 2001), the brightness of clouds was immediately reduced to more realistic values. To reduce the excessive cloud water and ice concentrations, the rate at which cloud liquid water is converted to precipitation (C_t) was increased as was the fall speed of ice (V_{F1}), an example of the effect of the former in reducing medium and thick clouds is shown in Figure 4.

The above cloud changes refer to the calculation of large–scale (or stratiform) clouds, i.e. those due to large-scale dynamical processes. Similar problems of insufficient cloud

extents but excessive cloud brightness were also present in the HadAM3 representation of convectively generated clouds. The low cloud extents are due to a lack of vertical variation of cloud amount in HadAM3, i.e. there is no representation of the amount of deep convective, or anvil, clouds increasing with height. The excessive brightness of convective clouds in HadAM3 is due to unrealistic amounts of cloud water. The introduction of the convective anvils allowed an appropriate choice of the shape of deep convective clouds allowing improved cloud extents (not
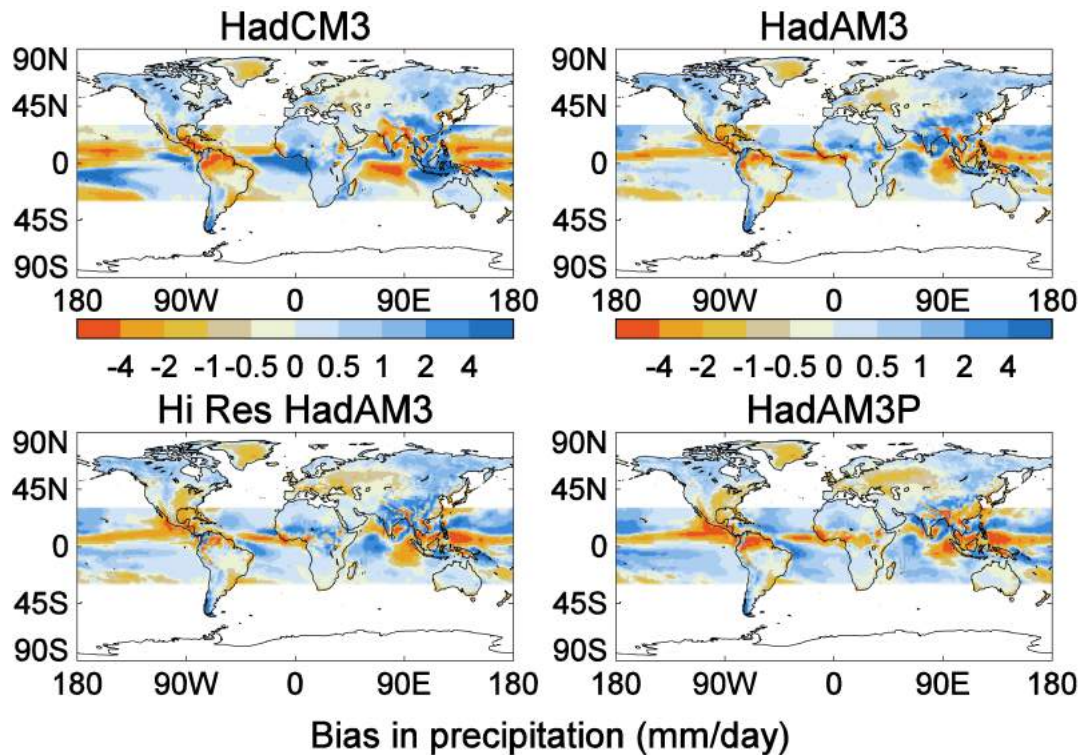
Figure 3. Errors in JJA precipitation for HadCM3 and HadAM3 (top row) and HadAM3 at double horizontal resolution and HadAM3P (bottom row) The land-surface precipitation biases are calculated with respect to the CRU-TS dataset (Mitchell and Jones 2005) and those over the tropical oceans with respect to the Climate Prediction Center (CPC) Merged Analysis of Precipitation (CMAP), (Xie and Arkin 1997).

shown) and also their impact on long–wave and short–wave cloud forcing in the tropics where anvil clouds dominate radiative balance. The anvils thus largely remove compensating excessive incoming shortwave and outgoing long–wave radiation fluxes (Figure 5). Reductions in convective cloud water were obtained by the improvements in the representation of convective updrafts.

The combined effects of the changes described above results in significant differences between HadAM3 and HadAM3P in clouds and their interaction with radiation. In order to generate a realistic global radiation balance (an important constraint on the system) and long–wave and short–wave cloud forcings (crucial factors in the models response to climate change) further fine–tuning of cloud parameters was performed. Thresholds for the conversion of cloud water to precipitation were increased over land and reduced over sea for both large–scale and convective precipitation. This reflects the significantly higher concentration of aerosols and thus CCNs over land and thus the higher total volume of water that can be supported within clouds (i.e. as smaller droplets) before forming precipitation. This tends to increase cloud lifetimes over land whilst decreasing them over sea. The combined effect of these and the other changes on cloud forcing are shown in Figure 5. Shortwave cloud forcing is improved over land compared to high-resolution HadAM3 (large positive biases in the Northern Hemisphere are reduced) and over sea compared to HadCM3 (large negative biases, especially in the North Pacific, are reduced). Long-wave cloud forcing is improved generally with respect to both models, with significant improvements in the tropics due to the representation of convective anvils in HadAM3P.

The improvements in the coupling between the soil and the land-surface and the treatment of surface radiation

fluxes have had little effect on the mean climatology of HadAM3P. However, they have substantially improved the simulation of temperature extrema. For example, in many areas significant biases in diurnal temperature range in HadAM3 (both at standard and high resolution) are removed in HadAM3P (not shown).

3. Results from model development

3.1. Comparison of model surface climatologies

In this Section we concentrate on comparing HadAM3P with the model it is designed to replace in terms of providing boundary conditions for regional climate modelling, i.e. HadCM3. Qualitatively, much of this comparison has already been described in the preceding section and so here we just provide some quantitative summary measures of the difference in performance of the two models. Given the intended focus of work with this model being on the regional implications of climate change we concentrate on assessing biases in the main surface variables of mean sea level pressure, temperature, precipitation and cloud cover. The first provides a measure of the realism of the large–scale circulation patterns in the models and thus acts as a check on the models ability to correctly simulate the drivers of regional weather phenomena. The second and third are the primary variables of interest when considering potential impacts of climate change. Another important variable in this context is also surface radiation for which reliable globally extensive observations are not available and so we use the closely related variable of cloud cover (which itself is often used to derive surface radiation changes in modelling impacts).

Table II compares the global skill of the two models in terms of their root mean square errors. Mean sea
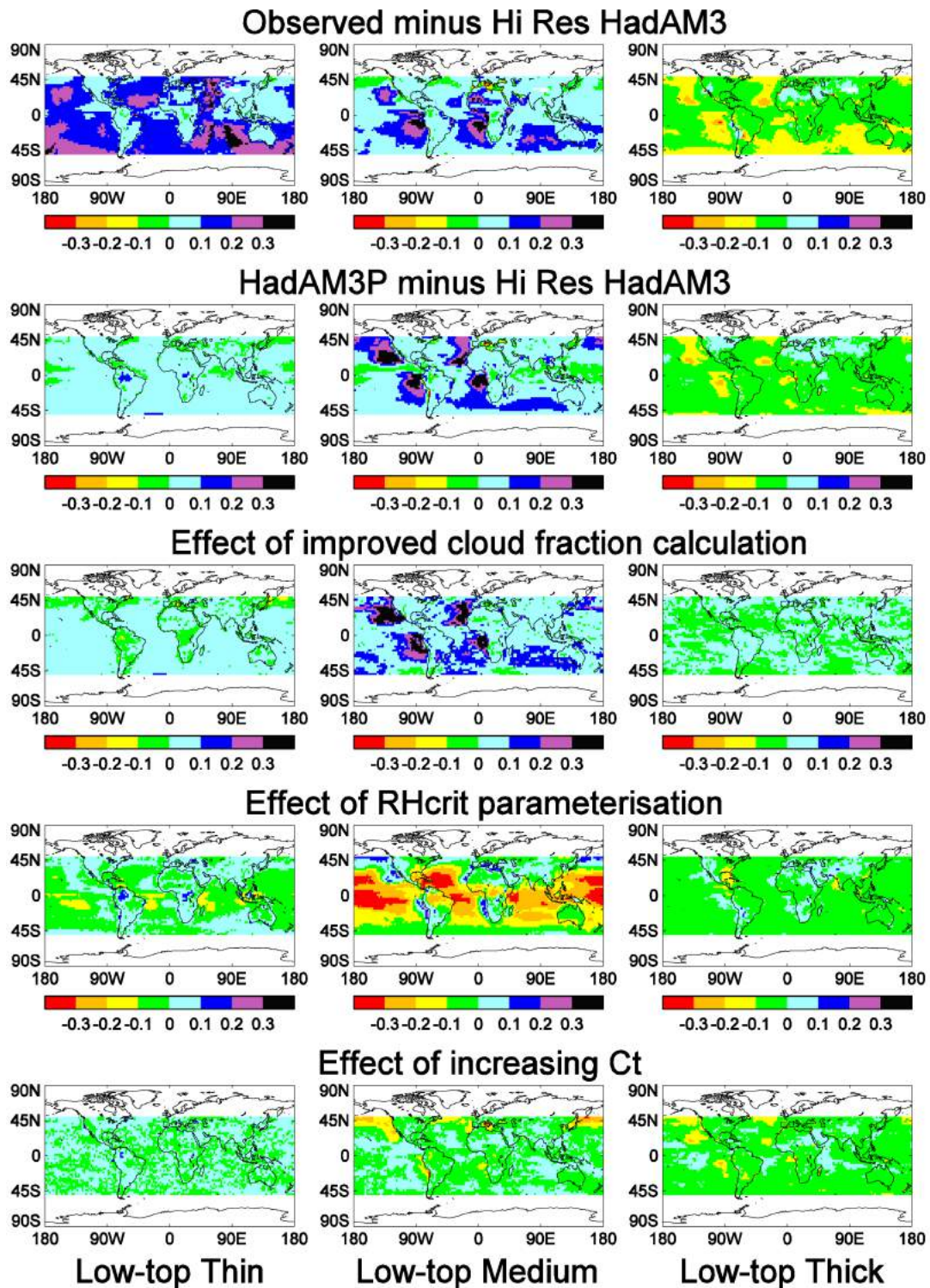
Figure 4. Differences in three optical thickness categories of ISCCP low cloud for JJA, comparing (from top to bottom) HadAM3 with observations and HadAM3P and then assessing the effect of the improved cloud fraction calculation, the Rhcrit scheme and the effect of increasing Ct. The cloud thickness biases are calculated with respect to data from the ISCCP dataset (Rossow and Schiffer 1999).

level pressure is compared with the global fields from the ERA15 reanalysis (Gibson *et al.* 1997) and the other variables are compared with land–only data of the CRU-TS dataset (Mitchell and Jones 2005). This shows that HadAM3P clearly performs as well as and mostly better than HadCM3. More specifically, HadAM3P significantly outperforms HadCM3 except for precipitation where the models have similar skill. The more similar behaviour for precipitation is mainly due to the effects noted in the previous section for boreal summer over the Eurasian and North American continental interiors where large temperature and precipitation biases at high resolution in HadAM3 are only partially compensated for by the improved physics in HadAM3P.

This picture is further confirmed when repeating the analysis over individual continental regions (Table III). For each region HadAM3P performs better overall. Across the eight season/variable combinations, HadCM3 only
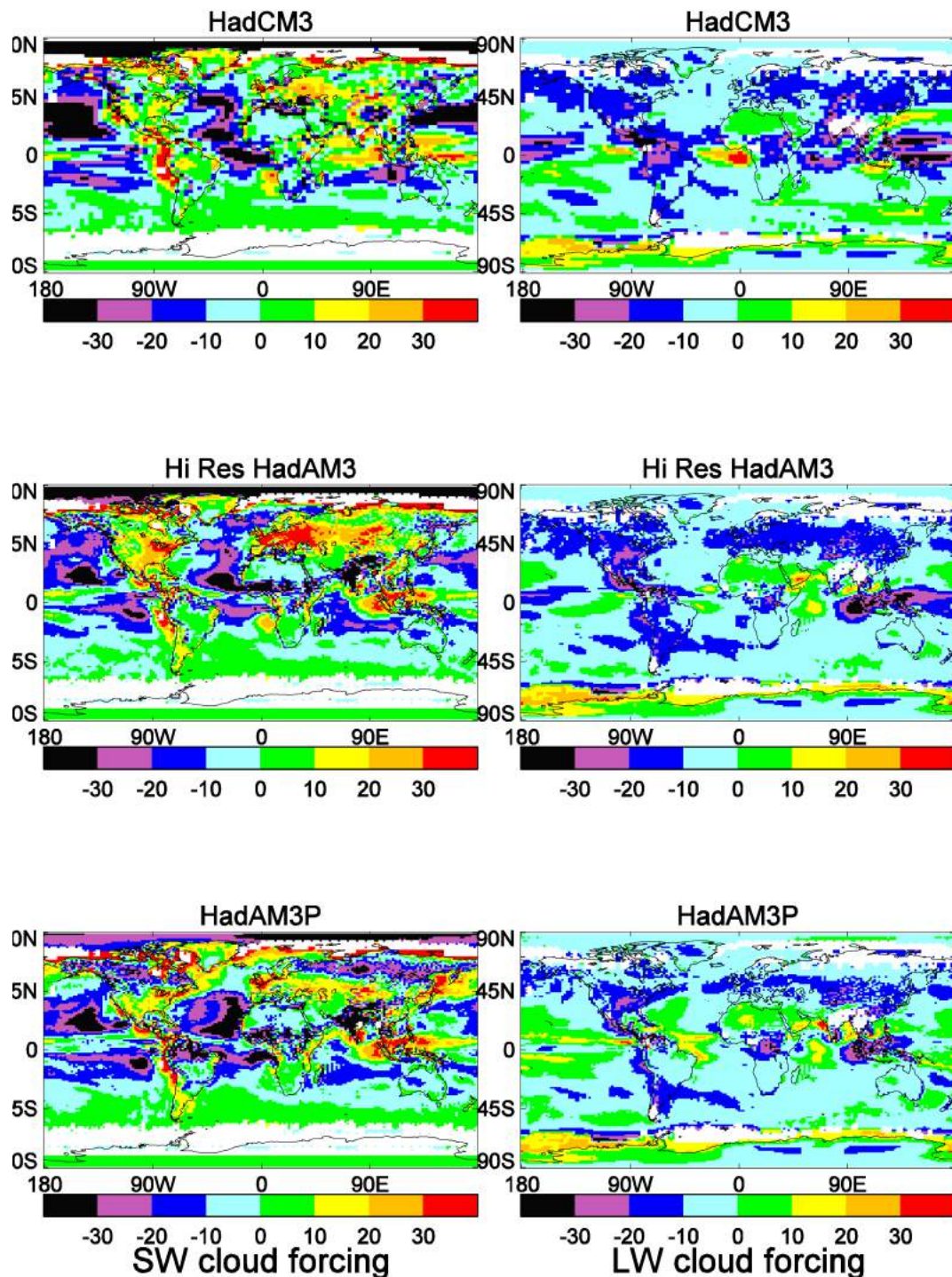
Figure 5. Errors in JJA shortwave (left) and longwave (right) cloud forcing in W/m2 in HadCM3 (top) and HadAM3 at double horizontal resolution (middle) and HadAM3P (bottom) The cloud forcing biases are calculated with respect to data from the ERBE (Earth Radiation Budget Experiment) dataset (Ramanathan *et al.* 1989).

performs better for one or two and HadAM3P for four or five. The only exception is North America where the models skill is more comparable, with HadCM3 better in two cases and HadAM3P in three. Finally, when comparing the models in terms of the different variables, HadAM3P is clearly superior in all but precipitation where again, on this regional analysis, their performance is comparable.

4.   Setup of the distributed computing experiment

The previous Sections (2 and 3) detail the development of the HadAM3P model, which is shown to produce a realistic climate over a number of regions. With the increase in computing power, it is now possible to run this model on a home PC, coupled via the lateral boundary conditions to a regional model variant, HadRM3P. Using the distributed computing network of climate*prediction*.net (CPDN) allows for many thousands of these models to be run on volunteer's home computers. This Section outlines the architecture of the distributed computing network and, by considering the GCM, RCM and distributed computing network as a single system, details an experiment to determine the system's suitability for use in probabilistic event attribution studies.

	DJF		JJA	
	HadCM3	HadAM3P	HadCM3	HadAM3P
Mean sea level pressure (hPa)	3.15	**2.96**	3.05	**2.73**
Surface air temperature (K)	3.55	**2.89**	2.41	**2.19**
Precipitation (mm/day)	1.47	1.49	2.33	2.28
Fractional cloud cover	0.19	**0.14**	0.16	**0.13**

Table II. Root mean square errors in HadCM3 and HadAM3P for the two extreme seasons of December, January, February (DJF) and June, July, August (JJA) for various surface fields. Errors are calculated at land points only for all but the first quantity. See text for the observational datasets. Figures are given in bold where the rms error for that model is 90% or less of the value for the other model.

	DJF		JJA	
	HadCM3	HadAM3P	HadCM3	HadAM3P
Europe=[-30,30,60,70]				
Mean sea level pressure (hPa)	5.40	**2.32**	3.12	3.20
Surface air temperature (K)	4.84	**2.67**	2.66	**2.54**
Precipitation (mm/day)	0.64	0.69	**0.58**	0.71
Fractional cloud cover	0.16	**0.14**	0.18	**0.16**
Asia=[60,10,175,70]				
Mean sea level pressure (hPa)	2.98	3.09	**4.78**	5.35
Surface air temperature (K)	4.40	**3.45**	2.46	**2.13**
Precipitation (mm/day)	**0.58**	0.70	2.18	2.05
Fractional cloud cover	0.13	**0.11**	0.16	**0.13**
Africa=[-15,-35,50,35]				
Mean sea level pressure (hPa)	2.13	2.29	2.91	2.73
Surface air temperature (K)	1.96	**1.73**	**1.88**	2.15
Precipitation (mm/day)	1.27	**1.06**	3.33	3.36
Fractional cloud cover	0.16	**0.17**	0.20	**0.20**
N America=[-160,30,75,60]				
Mean sea level pressure (hPa)	3.09	**2.25**	**2.49**	2.85
Surface air temperature (K)	3.43	**2.53**	2.31	2.53
Precipitation (mm/day)	0.87	0.88	**0.66**	0.74
Fractional cloud cover	0.13	0.12	0.18	**0.15**
S America=[-80,-55,-35,10]				
Mean sea level pressure (hPa)	2.88	3.09	**2.03**	2.24
Surface air temperature (K)	2.53	**2.09**	2.80	**1.86**
Precipitation (mm/day)	2.41	2.52	**2.42**	2.47
Fractional cloud cover	0.18	**0.11**	0.30	**0.18**
Australia=[110,-40,150,-10]				
Mean sea level pressure (hPa)	**1.84**	3.11	2.42	**1.56**
Surface air temperature (K)	1.50	1.60	1.96	**1.43**
Precipitation (mm/day)	1.48	**1.13**	0.51	**0.45**
Fractional cloud cover	**0.09**	0.11	0.11	**0.10**

Table III. As Table II but for individual regions. The region definitions are given as the [W,S,E,N] corners of a rectangular latitude–longitude box in degrees.

This experiment mirrors that used in Section 3 and so a direct comparison between those results and results from the HadAM3P/RM3P models running under the CPDN infrastructure can be made.

4.1. Required inputs to the models

While running under the distributed network, HadAM3P/RM3P requires a number of inputs, which must be supplied to the volunteer's computers. These

include the initial condition of the model and, as the model is atmosphere only, forcings are required at the sea–surface boundary, in the form of sea–surface temperatures (SST) and sea-ice fraction (SIF). Atmospheric concentrations of the well-mixed greenhouse gases are required, including carbon dioxide (CO_2), nitrous oxide (N_2O), methane (CH_4) and the halocarbons (CFC113, CFC11, CFC12, HCFC22, HFC124 and HFC134A). Ozone (O_3) concentrations are required as zonal averages at each model level and the inputs to the sulphur cycle are also required.

4.2. Distributed computing

To enable the computation of large ensembles of the GCM and RCM, volunteer distributed computing (VDC) is used. climate*prediction*.net (CPDN) uses VDC to great effect, generating very large ensembles of coupled slab layer ocean and atmosphere models (Stainforth *et al.* 2005), high resolution atmosphere only models (Pall *et al.* 2011) and coupled atmosphere–ocean models (Rowlands *et al.* 2012).

CPDN uses the Berkeley Open Infrastructure for Network Computing (BOINC; Anderson (2004)) to leverage the idle computing power of volunteers in a client / server model. Each volunteer signs up to the CPDN project via the BOINC client software, which then downloads the GCM and RCM to the volunteer's computer. CPDN scientists control the project's servers, which hand out workunits to volunteer's client computers. Each workunit contains all the information needed by the climate models to run an experiment for a certain period of model time, under a specified climate scenario. weather@home builds upon CPDN's success to use the same infrastructure to compute large ensemble simulations using the HadAM3P/RM3P models. The model integrations described in this paper are performed under a climate scenario designed to replicate the historical period of 1961 to 1990.

Unfortunately, not all workunits that are sent out by the server are completed by the clients. The failure to complete a workunit can be due to a number of factors, including unstable hardware, failure of hardware or termination of the workunit by the volunteer. The ratio of the number of sent out to completed workunits is called the attrition rate. Previous CPDN studies have shown that workunits with the lowest attrition rate take approximately one week to complete (Christensen *et al.* 2005), which equates to about one model-year of HadAM3P/RM3P integration and, therefore the workunit length in weather@home is set to be one model-year.

4.3. Model coupling

Under weather@home, HadAM3P/RM3P run on the same client computer in an interleaved manner. The GCM (HadAM3P) runs first for one full model day, providing the lateral boundary conditions (LBCs) to the RCM (HadRM3P) which also runs for one full model day. The coupling between the GCM and RCM is strictly one-way, in that the GCM feeds the RCM but there is no feedback from the RCM to the GCM. The RCM defines a four point buffer zone around the perimeter of the region. The main variables comprising the LBCs (atmospheric pressure at the surface along with horizontal winds, temperature and humidity for all atmospheric layers) are relaxed across the buffer zone to values temporarily interpolated from 6 hourly output from the GCM.

4.4. Model initial conditions

As mentioned above, the distributed system runs the HadAM3P/RM3P models for one year at a time, in a time-slice manner. In order to produce a timeseries of integrated models, a continuation system is used. Initially, the project creates a pool of workunits, each with the same generic starting conditions, at five-yearly intervals. These workunits are handed out to client computers, integrated over the model year under the specified climate scenario, and the results from the integration are returned, along with the final state of the model. This final state is then incorporated into a new workunit describing the next year of the climate scenario, using this final state as the starting condition. This process is repeated with the final states of the subsequent integrations, enabling a timeseries of climate model integrations to be built from the single-year runs.

As noted above, each initial pool of workunits, which are created at five-yearly intervals have the same, generic initial condition, which is the state of the model at the 1st December 1968, after 9 years of integration under an observed climate forcing scenario. Integrating the model to this time produces an initial condition which is close to the climatology of the 20th Century. In order to produce the full range of internal variability that is possible with the model, each workunit defines an initial condition perturbation which is applied to the generic starting conditions for the GCM. There is no perturbation applied to the RCM. The initial condition perturbation is drawn from a large set of possible perturbations defined as deltas in potential temperature and calculated from next day differences within a year long integration of the GCM. The perturbation is calculated as a fully three dimensional field, with a scaling function applied for all levels above a certain level in the models atmosphere. This is to ensure that there is no perturbation at the top of the atmosphere and so the top of atmosphere flux is not influenced too greatly. The scaling function has the form:

$$S = 1 - \frac{sin(Z_c - Z_0)}{N_z - Z_0}$$

where Z_c is the level at which the scaling is applied, Z_0 is the level above which no perturbation occurs and N_z is the number of levels in the model's atmosphere. The maximum amplitude of the perturbation is also limited to 5K to minimise the risk that a large perturbation in potential temperature could lead to an instability in the model. Finally, five global scaling factors are applied to the perturbations to generate a set of 1740 initial condition perturbations, from which the workunits can draw from. These initial condition perturbations are only specified for the original pool of workunits that have the generic starting condition and no perturbations are applied to the starting conditions used in the continuation process, so as to allow continuous integrations of models under a specific climate scenario.

Computing the climate models via VDC also adds the potential for perturbations to arise from the variety of computing platforms that the models will run on. weather@home supports Linux, Mac OS X and Microsoft Windows platforms and requires an Intel–compatible CPU. This allows for many permutations of operating system, CPU manufacturer and CPU model. A previous CPDN study using the HadCM3 coupled GCM (Knight *et al.* 2007)

found that this perturbation due to platform differences has approximately the same influence as an initial condition perturbation.

4.5. Forcings at the (sea-)surface

As both the GCM and RCM are atmosphere-only models, they both require forcings at the boundary between the atmosphere and ocean in the form of sea–surface temperatures (SSTs) and sea-ice fractions (SIFs). These quantities are defined per grid box for the GCM, with the RCM using the same field interpolated to the finer grid.

For the historical 1961 to 1990 climate scenario described in this paper, HadISST1 (Rayner *et al.* 2003) is used to provide both the SST and SIF fields. HadISST has been specifically designed to drive atmospheric climate models (Rayner *et al.* 2003) and is used in the ERA-40 reanalysis for the 1958 to 1981 period (Uppala *et al.* 2005). After 1981, the NOAA/NCEP 2D-Var dataset (Reynolds *et al.* 2002) is used in ERA-40 for the SST and SIF fields. HadISST also provides boundary layer forcings for the regional model in the PRECIS system (Jones *et al.* 2004), which uses the same HadRM3P model as weather@home. HadISST is provided as a global coverage dataset, for non-land points only, as monthly means with a spatial resolution of $1° \times 1°$. In order to use this data to drive the HadAM3P/RM3P model, it must be regridded to the GCM resolution of $1.875° \times 1.25°$. This is performed by an area–weighted averaging method. The discrepancy between the HadAM3P and HadISST land-sea masks (LSM) will cause missing data to be present in the regridded data. This is dealt with by assigning any grid box where missing data occurs with the mean of the surrounding grid boxes which themselves do not have missing data.

In addition to the spatial regridding, the HadISST data is also temporarily regridded. Woollings *et al.* (2010) show that the North Atlantic storm track in high resolution atmosphere only climate models is sensitive to the temporal resolution of the forcing SSTs, and that by increasing the temporal resolution from monthly mean values of SST to weekly mean values, an improvement is made to storm track density, matching ERA-40 more closely. As the SST values in HadISST are monthly, the actual weekly values of the SSTs cannot be recovered or reconstructed. However, the interpolation method of Sheng and Zwiers (1998) is used to recover some of the variability from the time series of monthly mean SSTs and SIFs.

HadAM3P/RM3P interpolates monthly values of SST and SIF to daily values using a simple linear interpolation scheme. Such a scheme has two disadvantages: the mean of the daily values will not equal the original monthly mean and the interpolated values are smoothed, losing inter-month variability (Sheng and Zwiers 1998). Applying this method overcomes these problems by adjusting the values of the monthly means, then linearly interpolating between the adjusted values, in the case of weather@home, to 5 day means. The adjusted mean values are derived by constraining the interpolated values over a month so that the mean of interpolated values for that month is equal to the original monthly mean. This temporal interpolation is performed offline, forming SST and SIF files for a single year, which then becomes part of the workunit describing the climate scenario for that year.

4.6. Atmospheric composition

In the distributed experiment, the concentrations of CO_2, CH_4 and N_2O follow the timeseries of the observed quantities in the IPCC 4th Assessment Report (AR4, Solomon *et al.* (2007)). The halocarbon gases (CFC113, CFC11, CFC12, HCFC22, HFC124 and HFC134A) are represented as a single value per timepoint in the timeseries, which produces the equivalent radiative forcing as if all 6 gases were modelled. Ozone (O_3) also follows the observations from AR4, including the appearance of the ozone hole in the 1980s.

4.7. Additional climate drivers

The standard inputs to the HadAM3P/RM3P sulphur cycle scheme are the surface anthropogenic emissions of SO_2, elevated emissions (e.g. from a chimney stack) of SO_2 which are released at a higher model layer than the surface, natural emissions of SO_2 from regularly erupting, but small scale, sources and natural emissions of DMS, primarily from phytoplankton in the ocean. An extra modification to HadAM3P/RM3P for weather@home is the addition of large scale volcanic eruptions emitting large quantities of SO_2. These large natural emissions are modelled in four latitude bands, with a value prescribed for each band for each model year. Input files are required for the five emission types above, as well as the 3D fields for the oxidisation variables. Figure 6 shows the timeseries of the modification to the optical depth due to volcanic activity.
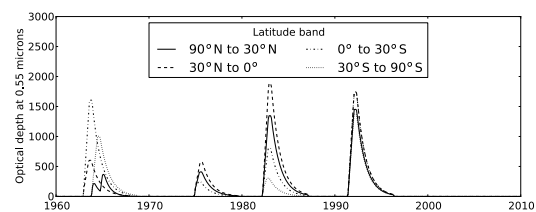

Figure 6. Increase in optical depth due to volcanic activity in four latitude bands

weather@home also applies a small modification to HadAM3P to account for variations in solar activity by allowing anomalies from the solar constant to be specified. These anomalies are taken from Krivova *et al.* (2007) which take into account the 11 year solar cycle as well as longer timescale solar processes. Figure 7 shows the timeseries of the anomaly to the solar constant.
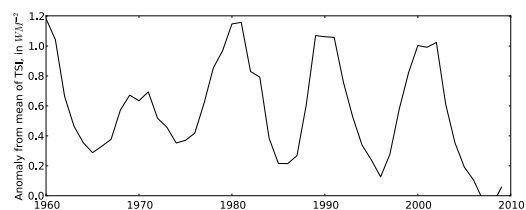

Figure 7. Anomaly to the solar constant, based on Krivova *et al.* (2007)

5. Results from the distributed computing experiment

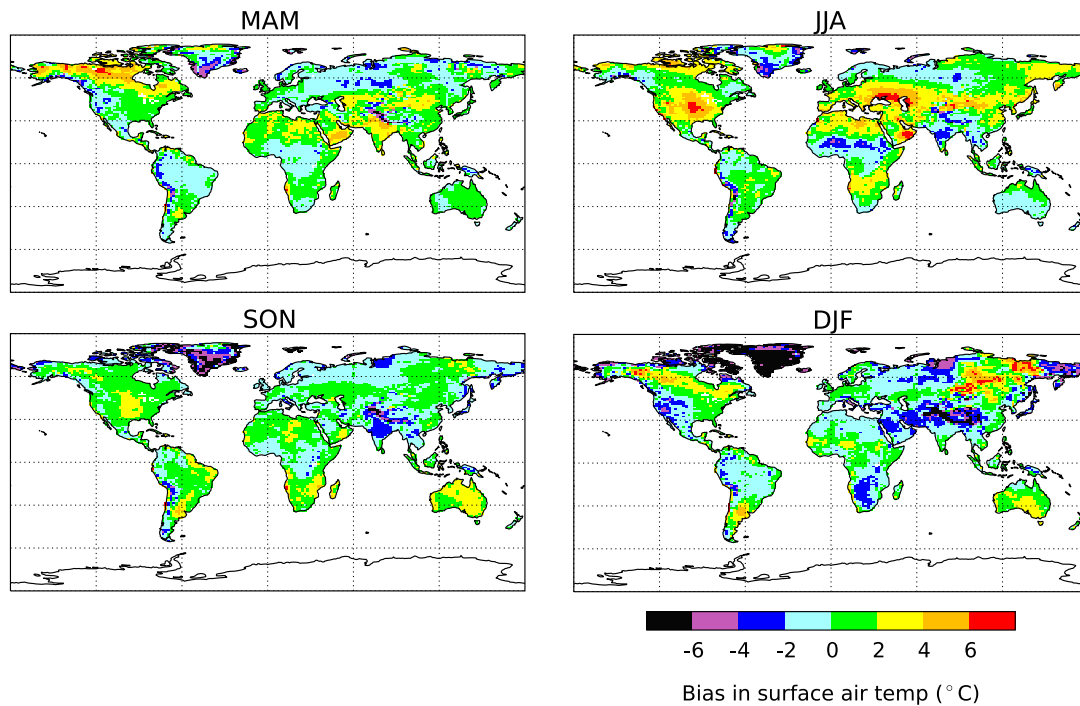Using the experimental setup detailed in Section 4, approximately 500 ensemble members per year have been

Figure 8. Ensemble mean of the temperature bias in the weather@home global HadAM3P model with respect to CRU-TS.

computed, for the historical period of 1961 to 1990. This Section analyses the results from those model runs, firstly in a manner similar to Section 2, examining the bias in the GCM and RCM seasonal mean temperature and precipitation variables. The distribution of the global and regional model's daily temperature and precipitation variables are then examined so as to determine their suitability for use in a probabilistic event attribution study. In this paper a subset of 25 ensemble members per year is used, for the period 1961 to 1990, giving a total ensemble size of 750 over the 30 year period. The ensemble members in the subset are chosen randomly, with each ensemble member having equal chance of being chosen to be part of the subset. The large ensemble capacity of weather@home is of most use when examining extreme weather events which, by their very nature, are rare and require a large ensemble to generate examples of the event. Conversely, to characterise the climatology and overall distribution of climate variables over a 30 year period a much smaller ensemble is required.

Between the results from the model development in Sections 2 and 3 and the launch of weather@home, there was a gap of around 5 years. During this period an error was found and corrected in the representation of one of the soil properties affecting the mobility of soil moisture. Thus, in addition to presenting results from the entire weather@home system, this Section demonstrates the impact of this change, along with running these models on different computing platforms compared to the supercomputer on which the initial configurations were tested.

5.1. Global model

Analysis of the GCM follows the analysis in Sections 2 and 3, in that the bias between the seasonal mean of the model and the seasonal mean of CRU-TS is calculated for each

ensemble member, and then the ensemble mean of these biases is produced. Figures 8 and 9 are directly comparable to Figures 1, 2 and 3 as the same colour schemes and observational datasets are used. Throughout this Section, the HadAM3P model running under weather@home will be denoted HadAM3P-W@H, whereas the model development version from Section 2 will be denoted HadAM3P-MD.

Figure 8 shows the ensemble mean of the bias in the near–surface temperature of HadAM3P-W@H, when driven with HadISST SSTs and historical atmospheric forcings over the period 1961 to 1990, as detailed in Section 4.

In the Northern Hemisphere winter (DJF), there are large temperature biases over Greenland, Eastern and Arctic Russia, China, South Asia and South Africa, as well as the Western edge of the Americas. Biases over Europe, Scandinavia and Northern Africa are much lower, with some errors over the Alps. This indicates that the HadAM3P GCM is suitable for driving the HadRM3P RCM over the European domain. Compared to the same season in Figure 1 HadAM3P-W@H shows the same pattern of bias as HadAM3P-MD, with some reduction in bias over Eastern Europe and an increase in bias over Arctic Russia. Table IV indicates that, globally, there is more bias in HadAM3P-W@H (with an RMSE score of 3.77) than HadAM3P-MD (2.19), whereas for the European domain, there is less bias in HadAM3P-W@H (2.17) than in HadAM3P-MD (2.67).

In the Northern Hemisphere summer (JJA), there are large warm biases over the United States of America, and the Caspian and Black Sea areas of Europe. Although the patterns are largely similar, these specific locations have biases that are greater in HadAM3P-W@H than in HadAM3P-MD. Comparing Tables II and IV shows that, globally, there is a similar error in HadAM3P-W@H (RMSE of 2.27) as in HadAM3P-MD (2.19) but that over the European domain HadAM3P-W@H (2.86) performs worse than HadAM3P-MD (2.54). When run as part of the weather@home system, it is expected that these biases will

be propagated to HadRM3P, impacting the performance of the RCM over Europe.

Biases in Spring (MAM) and Autumn (SON) are less than in DJF or JJA, with SON showing a cold bias over Greenland which, although large, is not as large as the cold bias over Greenland in DJF. MAM shows a warm bias over Northern Canada and some cold bias over Greenland. These smaller biases are confirmed by the RMSE which, for both seasons, are less than DJF and JJA over both the global domain and the European domain. This leads to the expectation that biases in MAM and SON in HadRM3P should also be low.

Figure 9 shows the ensemble mean of the bias in precipitation in HadAM3P-W@H over the 1961 to 1990 period. For the Northern Hemisphere Winter (DJF) there are large dry biases over the Amazon, Greenland, Indonesia and Madagascar. However, over Europe there is a mix of a small wet bias (up to 1mm/day) in Western Europe, a slight dry bias (up to -1mm/day) in some areas of Southern Europe and the UK and very little bias in Eastern Europe. Globally, in DJF, HadAM3P-W@H has a smaller RMSE (1.32) than HadAM3P-MD (1.49). However, over Europe the RMSE are much more similar (HadAM3P-W@H: 0.71, HadAM3P-MD: 0.69).

In the Northern Hemisphere Summer (JJA), HadAM3P-W@H has large dry biases over Columbia, Venezuela and Central America, a dry bias over the Eastern USA and dry biases over Western Africa, South East Asia and Indonesia. Europe shows a slight dry bias (up to -1mm/day) with a slight wet bias (up to 0.5mm/day) over Spain and Scandinavia. This is a very similar pattern of bias as in HadAM3P-MD, shown in Figure 3, with an improvement to the bias over Eastern Europe and the Alps. This improvement is quantified by comparing the RMSE of HadAM3P-MD in Tables II and III and HadAM3P-W@H in Table IV. Globally, HadAM3P-W@H (RMSE of 1.62) has less bias in precipitation than HadAM3P-MD (RMSE of 2.19). However, over the European domain the scores are much more similar (HadAM3P-W@H: 0.80, HadAM3P-MD: 0.71). This is due to the improvement in the bias over Eastern Europe in HadAM3P-W@H being outside the domain used to calculate the European RMSE. However, both models have a low RMSE.

Bias in Spring (MAM) and Autumn (SON) are similar to those in JJA and DJF, with MAM having dry biases in Southern America and Indonesia and a wet / dry bias in Africa. SON also has a dry bias in Indonesia, with further dry biases in Central America and South East Asia. Both seasons have a dry Greenland. Globally, MAM has an RMSE (1.53) less than JJA (1.62) and SON has an RMSE (1.32) equal to DJF. Over the European domain, MAM has the lowest RMSE (0.50) of all the seasons and SON has an RMSE (0.67) greater than this but less than DJF and JJA. Overall the error in precipitation over Europe is in line with both HadAM3P-MD and HadCM3, except for JJA whose larger error is likely to be due to the warm bias in temperature in that season.

5.2.   Regional and global model over European domain

Section 5.1 evaluated the bias in the GCM component of weather@home for the entire globe. This Section examines the biases in the RCM modelling a European domain, and compares these biases to the same spatial domain in the GCM. This evaluation serves two purposes. Firstly, to check

	MAM	JJA	SON	DJF
Global domain				
Air temperature (K)	1.91	2.27	2.17	*3.77*
Precipitation (mm/day)	1.53	**1.62**	1.32	**1.32**
Europe=[-30,30,60,70]				
Air temperature (K)	1.54	*2.86*	1.40	**2.17**
Precipitation (mm/day)	0.50	*0.80*	0.67	0.71

Table IV. Root Mean Square Errors (RMSE) in the HadAM3P GCM model running as part of the weather@home system, for all seasons, with respect to the CRU-TS dataset. Both the RMSE over the global domain and over Europe are given. Values in bold are 90% or less of the corresponding values in Tables II and III. Values in italic are 110% or more of the corresponding values.

that there is consistency between the GCM and RCM, in that the patterns of biases are largely the same and that the RCM does not introduce any new biases into the domain. Secondly, to determine whether the RCM has less or more bias overall than the GCM within the European domain.

Figure 10 shows the results of this comparison for seasonal mean temperature biases with respect to CRU-TS (Mitchell and Jones 2005). The GCM and RCM are largely consistent over the European domain in all seasons. In MAM, there is a reduction in the cold bias of the GCM in the RCM over Iceland and Western Norway. Some additional cold bias is present at the eastern edge of the domain in the RCM. In JJA there is an increase in the RCM of the warm bias over the Balkans (west of the Black Sea), over Northern Europe and over the alps. In SON there is again a reduction in the cold bias seen in the GCM over Western Norway and also a reversal of the cold bias over Italy. In DJF, the RCM shows a reduction in the cold bias over regions in Southern Europe that are close to the Mediterranean Sea and a reduction in the cold bias over Northwest Russia. However this is balanced by an increase in the warm bias over Norway and a slight increase in the warm bias over Central Europe.

Table V provides a more quantitative assessment of these findings. Computing the root mean square error (RMSE) for biases in both the GCM and RCM, for each season, it shows that there is less overall error in the RCM than in the GCM for every season except for JJA.

Figure 11 shows the comparison between the HadAM3P GCM and the HadRM3P RCM for biases in seasonal mean precipitation over the European domain, with respect to CRU-TS. As with the temperature biases, there is largely agreement between the GCM and RCM, with the location and magnitude of the precipitation biases remaining the same. Table V provides a quantitative assessment of these errors, by computing the RMSE for the biases in both the GCM and RCM. Unlike the temperature, there is actually an increase in the error in the precipitation when dynamically downscaling from the GCM to the RCM.

These results are consistent with the performance of the HadAM3H RCM run of the PRUDENCE project (Jacob *et al.* 2007), which compares a number of RCMs driven by the same boundary conditions. In particular, HadRM3P shows the same warm and dry bias in JJA in
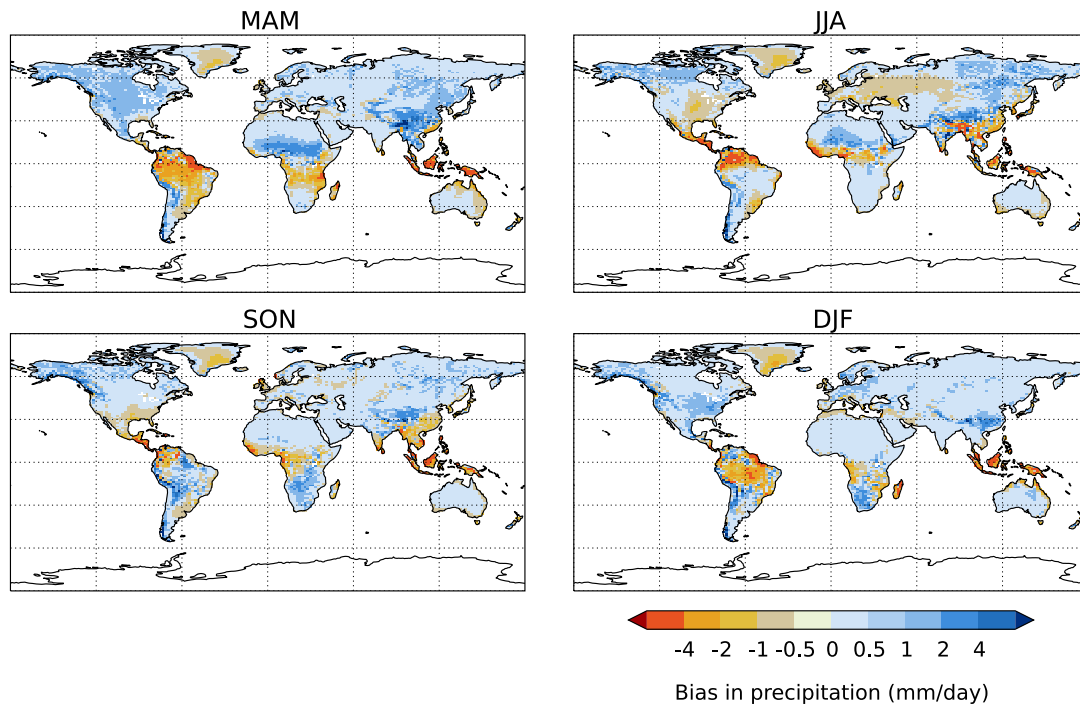
Prepared using qjrms4.cls

Q. J. R. Meteorol. Soc. **00**: 2–22 (2014)

Figure 9. Ensemble mean of the precipitation bias in the global HadAM3P model with respect to CRU-TS.
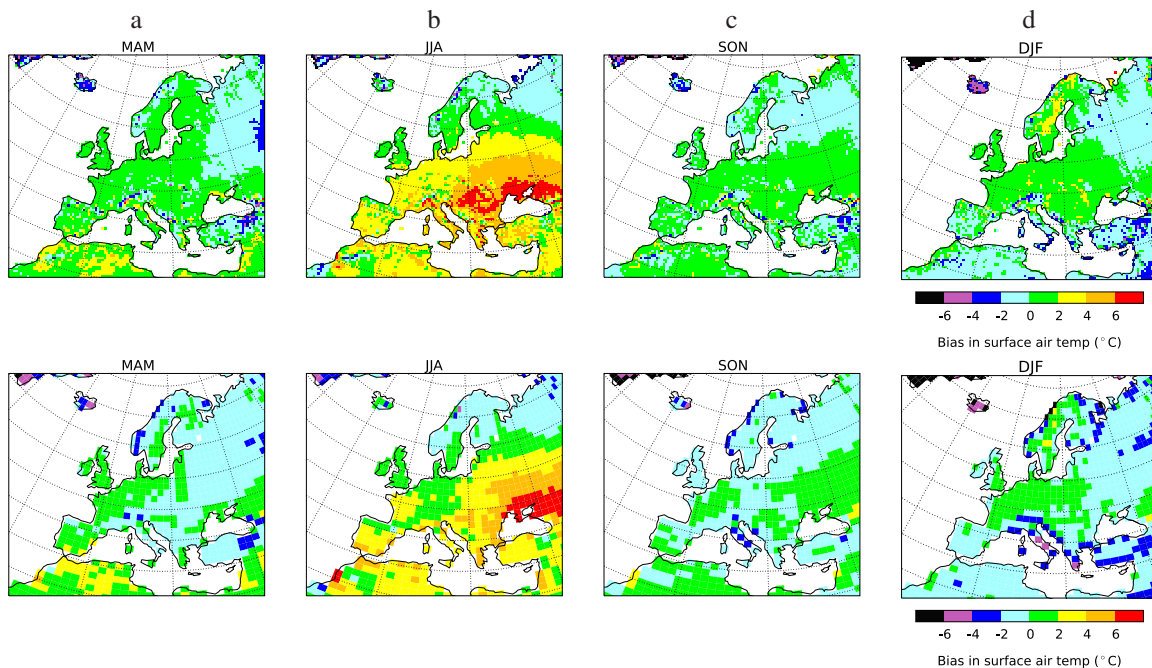

Figure 10. Ensemble mean of the temperature bias in the regional HadRM3P model (top row) and global HadAM3P model (bottom row) with respect to CRU-TS.

the Mediterranean and Eastern Europe as HadRM3H (Jacob *et al.* (2007), Figure 3).

5.3.   Distribution of daily variables in the regional and global model over the UK and Ireland

While Section 5.1 evaluated the weather@home system over a global scale for seasonal means of climate variables and Section 5.2 examined the biases in the RCM, also for seasonal means, the extreme events which probabilistic event attribution is interested in occur on smaller spatial

scales and shorter timescales, typically a single day to a week. In light of this, this Section evaluates the system's ability to represent the distribution of daily temperature and precipitation values over a region encompassing the land–points of the UK and Ireland.

As Section 5.2 shows, there is some small improvement to the temperature bias in the RCM, when compared to the GCM, and a small increase to the precipitation bias. In light of this, it would be reasonable to ask what the added value of the regional model is. In this Section we show that, by moving away from just considering climatological seasonal
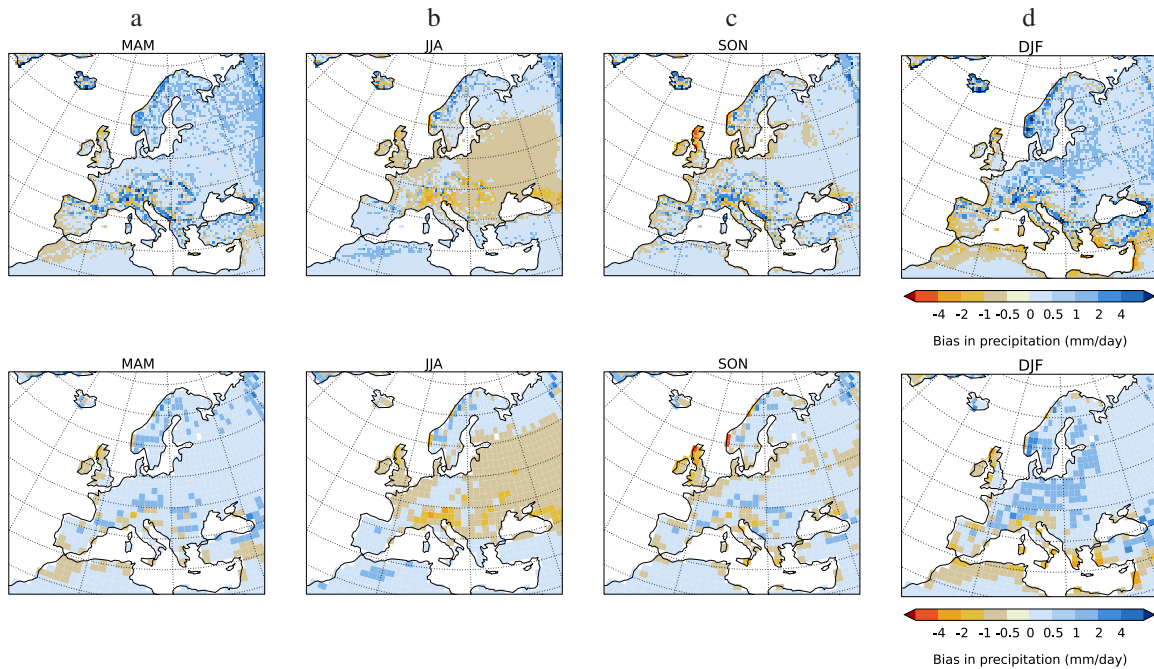
Figure 11. Ensemble mean of the precipitation bias in the regional HadRM3P model (top row) and global HadAM3P model (bottom row) with respect to CRU-TS.

	MAM		JJA	
	HadAM3P	HadRM3P	HadAM3P	HadRM3P
Air temperature (K)	1.60	1.41	2.73	2.97
Precipitation (mm/day)	0.48	0.87	0.75	0.84
	SON		DJF	
	HadAM3P	HadRM3P	HadAM3P	HadRM3P
Air temperature (K)	1.68	1.27	2.80	1.92
Precipitation (mm/day)	0.62	0.89	0.68	1.07

Table V. Root Mean Square Errors (RMSE) for the global HadAM3P and the regional HadRM3P models, over the European domain, as modelled in HadRM3P. The HadAM3P data has been remapped to the HadRM3P rotated grid and domain before the RMSE calculation.

means, the RCM improves the modelling of the distribution of the daily temperature and precipitation variables.

Figures 12 and 13 show quantile–quantile plots (Q–Q plots), comparing the distribution of the daily mean temperature and daily mean precipitation in the large ensemble of RCM and GCM runs to the distribution of these variables in the E-OBS dataset (Haylock *et al.* 2008), over the UK and Ireland. In this Section, the same ensemble as in the previous Sections is used, with the corresponding observations taken from 1961 to 1990 from E-OBS. In effect we are comparing two ensembles to the observations: an ensemble of GCMs and an ensemble of RCMS, which have been driven at the LBCs by the GCM. To construct the distribution of the variables from the large ensemble, daily values are recovered from every ensemble member at every grid box that is land in the UK and Ireland region. The models use a 360 day year and, therefore, across the 750 member ensemble, values at 270,000 instances of the UK and Ireland domain are produced. HadAM3P has 18 grid boxes in this domain, whereas HadRM3P has

136 grid boxes. For the ensemble, this produces 4,860,000 values from the GCM and 36,720,000 from the RCM. Such large sets of values ensures that the ensemble is truly representative of the behaviour of the model, allows the extreme values in low and high percentiles to be captured and also enables testing of statistical significance.

In order to compare this high temporal and spatial resolution model ensemble to the real world, an observational dataset with similar temporal and spatial resolution is needed. We have chosen the E-OBS dataset (Haylock *et al.* 2008) for this part of the analysis as it is available on the same rotated grid (50km, $0.44°$) as the RCM, meaning that no remapping of the observation to model grid is needed. For the GCM, E-OBS is also available on a regular latitude–longitude grid at $0.5°$ resolution. In order to compare this to the GCM, a remapping to the GCM grid is required. For daily mean temperature this is done by bilinear interpolation. For daily mean precipitation, this is done by a conservative remapping scheme that ensures the same total

amount of precipitation is in the remapped data as is in the original data.

Figures 12 and 13 feature a solid line and an envelope. The solid line contains the values at the percentiles for the whole ensemble – i.e. all 750 members. The envelope shows the 5th to 95th percentile range for the values at each percentile when considering each ensemble member. So, for example, the value at the 50th percentile will have 750 potential values (one from each ensemble member) and the 5th and 95th percentile of these 750 values is found. This allows both the uncertainty and internal variability of the model to be assessed.

In order to quantify the differences in how well the modelled distribution of the climate variable matches the observed distribution, the root mean square difference (RMSD) of the values at the 1st to 99th percentiles is used. This can be expressed by:

$$RMSD = \frac{1}{99}\sqrt{\left(\sum_{p=1}^{p=99}(o_p - e_p)^2\right)}$$

where o_p is the value at the pth percentile in the observations and e_p is the value at the same percentile in the ensemble. In Table VI this RMSD is calculated for the percentile values computed for all ensemble members, i.e. the solid line in Figures 12 and 13.

Figure 12 shows the quantile–quantile plots for daily mean temperature over the UK and Ireland in both the RCM and GCM, for all four seasons. In MAM there is a good correspondence between the observed and ensemble quantile values for both the RCM and GCM. This is confirmed in Table VI, with both the GCM and the RCM having an RMSD of 0.04, for all percentiles. In the higher percentiles, the GCM and RCM also have a very similar performance, with an RMSD of 0.31 for the RCM and 0.30 for the GCM. However, in the lower percentiles the RCM (0.22) does peform better than the GCM (0.37). There is a greater spread in the ensemble values for the higher percentiles than for the lower and middle percentiles, but this spread is similar between the RCM and GCM.

In JJA the RCM actually performs worse than the GCM, for higher and middle percentiles. This reduction in skill corresponds to the increase in bias in the RCM of the monthly mean temperatures, as seen in Table V. For all percentiles, the RMSD increases from 0.10 for the GCM to 0.12 for the RCM. However, in the upper percentiles, the difference is more apparent, with the RMSD increasing from 0.98 to 1.17. There is a much larger spread in the values at the higher percentiles than in the lower percentiles, although the spread over all percentiles remains much the same between the GCM and RCM. For the higher percentiles, the spread is asymmetrical, with the 99th percentile having ensemble members with a much lower value than the value derived from all ensemble members. This indicates that the model is capable of producing a wide range of mean temperatures in JJA, Therefore, to fully sample all of the weather patterns that could produce these mean temperatures, a large ensemble is necessary.

In SON, both the RCM and GCM perform similarly, with the RCM having the better performance in the lower percentiles. For all percentiles, the RMSD reduces from 0.17 to 0.13. For the upper percentiles ($90th+$), the RCM also has a slightly lower RMSD (0.29) than the GCM (0.31). Again, the upper percentiles have more of a spread in values

than the lower percentiles, with the GCM having a slighter wider spread of values in the lower percentiles than the RCM.

For DJF, both the RCM and GCM show good skill in modelling the percentile values, with the RCM having slightly more skill, especially in the lower percentiles. This is confirmed by the RMSD scores of 0.10 for all percentiles in the GCM and 0.03 in the RCM. For the higher percentiles, the regional model (0.12) has higher skill than the GCM (0.17). Most importantly, for a season where one extreme weather event of interest is very cold temperatures, the RMSD for the lower percentiles in the RCM (0.10) is much lower than in the GCM (0.82) indicating that the RCM has much more skill in representing very cold days at the correct frequency, when compared to the observations.

Figure 13 shows the quantile–quantile plots for precipitation in the RCM and GCM for all seasons. Unlike the daily mean temperature, which shows an improvement in modelling the percentile values for all seasons except JJA, the percentile values for precipitation show little improvement in the RCM compared to the GCM, except for JJA. As Table VI shows, the RMSD for all percentiles is much the same for the RCM and GCM in MAM and DJF. In JJA the RCM (0.13) performs better than the GCM (0.19), whereas in SON the RCM (0.18) performs worse than the GCM (0.14). In the higher percentiles (90th to 99th) the RCM performs better in MAM, JJA and DJF, but worse in SON.

The largest improvement to the modelling of precipitation in the RCM occurs during the JJA season. Despite there being a similar warm bias in the seasonal mean temperatures over much of the region, as show in Figure 10 and Table V and a dry bias in the seasonal mean precipitation (Figure 11), the RCM has a distribution of daily mean precipitation that matches the observed distribution much more closely than in the GCM, although there is still considerable bias. This is confirmed in Table VI where the RMSD for all percentiles in the GCM is 0.19 and the RCM is 0.13. In the higher percentiles this improvement is even more apparent, with a RMSD of 1.68 in the GCM and a RMSD of 1.02 in the RCM. This shows that the increase in resolution in the RCM has a positive effect on modelling the daily mean precipitation, even though the same biased surface temperature as in the GCM is being used as drivers, both at the sea–surface boundary and the lateral boundaries.

Although, as discussed above, the performance of the RCM in representing the distribution of daily mean precipitation is worse than the GCM in MAM, SON and DJF, Figure 13 shows that the bias in the RCM is much more consistent than the bias in the GCM. The RCM consistently under–predicts the precipitation (the model is too dry) across all percentiles, whereas the GCM under–predicts in lower percentiles and over–predicts (the model is too wet) in higher-percentiles in DJF and MAM. This consistency of the RCM is an advantage as it allows for the use of a straight–forward scaling and offset bias correction to be used (Massey *et al.* 2012), whereas the inconsistency of the GCM may require a more complicated bias correction method.

6. Discussion and conclusion

In the distributed computing project weather@home, the atmosphere only Hadley Centre model HadAM3P is coupled to a higher resolution regional equivalent, HadRM3P, which was produced as part of the UK Met
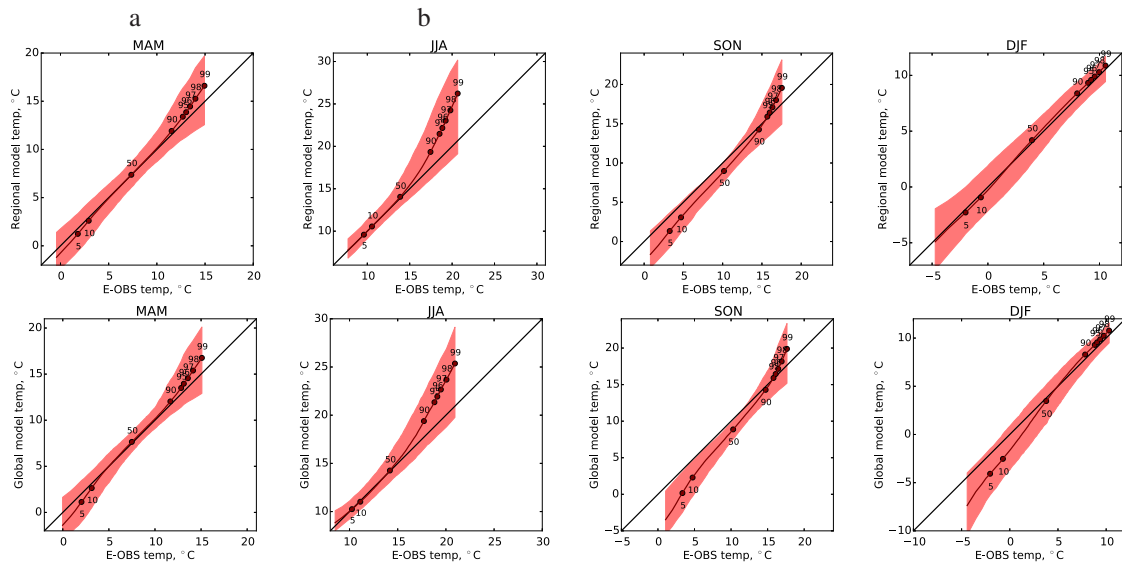
Figure 12. Quantile-quantile plot of the distribution of daily mean temperature in the regional model (top row) and the global model (bottom row) compared to the distribution of daily mean temperature in the E-OBS dataset over the UK and Ireland. The black line shows the quantile values for the entire ensemble. The red envelope shows the 5th to 95th percentile range of values for individual ensemble members.
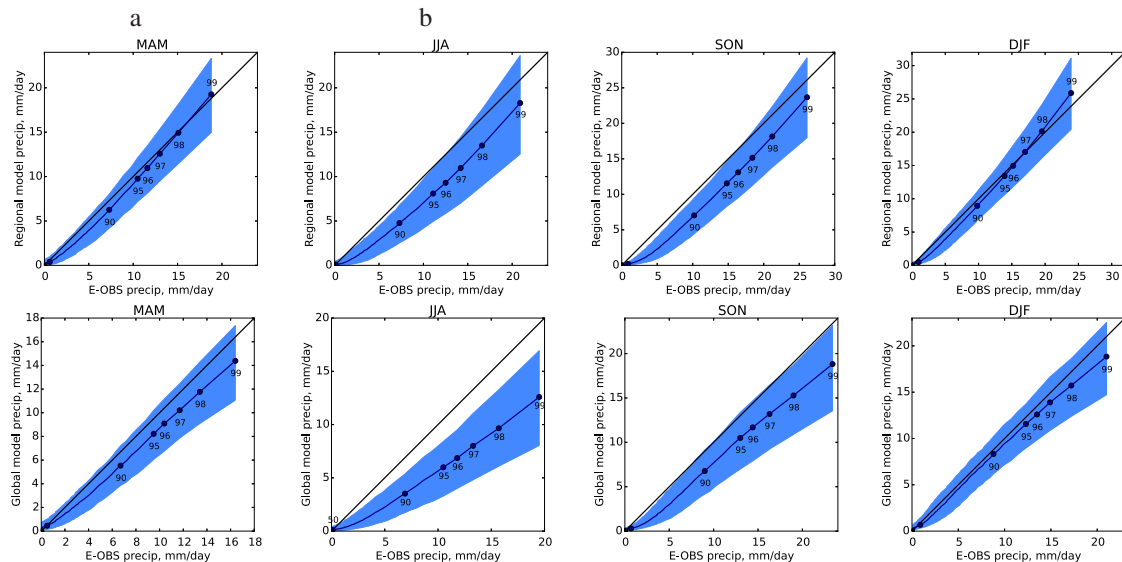

Figure 13. Quantile-quantile plot of the distribution of daily mean precipitation in the regional model (top row) and the global model (bottom row) compared to the distribution of daily mean precipitation in the E-OBS dataset over the UK and Ireland. The black line shows the quantile values for the entire ensemble. The blue envelope shows the 5th to 95th percentile range of values for individual ensemble members.

Office PRECIS project. The distributed computing used to simulate large ensembles with perturbed initial condition requires the use of a relatively old modelling base, the Hadley Centre HadCM3 model, which has the advantage of only requiring a small memory footprint when running on a typical home computer. The changes made in producing the HadAM3P model from this model base include improvements to the resolution and model physics (Sections 2 and 3). These changes result in improvements to the climatology compared to other models derived from the same modelling base. HadCM3 is still a very successful model, despite its age, and is included in both the CMIP5 multi-model ensemble and the IPCC AR5 report. HadCM3 is shown to have errors on a par with other members of the CMIP5 ensemble when comparing the global seasonal-cycle climatology with observations from 1980 to 2005 (Stocker *et al.* (2013), Figure 9.7). Furthermore, Sillmann

et al. (2013) show that HadCM3 is competitive with other CMIP5 models when evaluating climate extreme indices, with the exception of consecutive wet and dry days and variables related to the diurnal cycle. This performance of HadCM3, along with the improvements in HadAM3P related to the diurnal cycle and SST biases, gives confidence to the competitiveness of the weather@home modelling set-up.

The modelling system as a whole is a powerful tool to understand changes in weather events all over the world and also allows attribution studies to be conducted with respect to different external drivers of the climate system. Although this paper has concentrated on the European region, weather@home currently has six regions being modelled by HadRM3P. In order to make the most of these opportunities, the regional climate model data produced by the six regions are hosted and analysed by experts

	MAM		JJA	
	HadAM3P	HadRM3P	HadAM3P	HadRM3P
Mean air temperature (K) *all*	0.04	0.04	0.10	0.12
Mean air temperature (K) *1–10th*	0.37	0.22	0.05	0.01
Mean air temperature (K) *90th+*	0.30	0.31	0.98	1.17
Precipitation (mm/day) *all*	0.07	0.06	0.19	0.13
Precipitation (mm/day) *90th+*	0.50	0.27	1.68	1.02
	SON		DJF	
	HadAM3P	HadRM3P	HadAM3P	HadRM3P
Mean air temperature (K) *all*	0.17	0.13	0.10	0.03
Mean air temperature (K) *1–10th*	1.27	0.75	0.82	0.10
Mean air temperature (K) *90th+*	0.31	0.29	0.17	0.12
Precipitation (mm/day) *all*	0.14	0.18	0.05	0.07
Precipitation (mm/day) *90th+*	1.03	1.11	0.37	0.30

Table VI. Root Mean Square Difference (RMSD) for all percentiles from 1 to 99 for the daily mean air temperature and precipitation over the UK and Ireland (*all*), the 1st to 10th (*1–10th*) and for the 90th to 99th percentiles (*90th+*).

of the respective region: Europe (Oxford, UK), Western US (Oregon, USA), South Africa (Cape Town, South Africa), Australia and New Zealand (Melbourne, Tasmania, Wellington) and South Asia (Pune, India). It is an aim of the project to extend the network of regions to cover all parts of the land surface of the Earth. The collaborative network makes it not only feasible to store and analyse the large amount of data generated by this approach but also ensures that experts of different regional climates will be involved in future development of the modelling system.

weather@home uses just one GCM driving a single RCM. The CORDEX project (Giorgi *et al.* 2009) aims to understand some of the uncertainties in regional modelling by comparing many RCMs driven by both observations and output from multiple GCMs. Although weather@home is not part of CORDEX, it is aligning itself with the methodologies of CORDEX as closely as possible. For example, the European domain presented in this paper has the same rotated pole as the CORDEX domain and contains the agreed common interior. The Australia and New Zealand domain is the same as the Australasia domain in CORDEX and the South Asia domain in weather@home is identical to the CORDEX South Asia (SASIA) domain. This aligment with CORDEX will enable our collaborators to compare results with other models and to determine where the weather@home RCM output fits into the distribution of the multi-model ensemble members in CORDEX.

In conclusion, the modelling approach is an excellent tool to analyse statistics of regional extreme weather events. The driving model has biases in surface temperatures with larger biases in boreal Winter and Summer in the Northern Hemisphere but the representation of precipitation is exceptionally good with respect to the dynamics and physics represented in the model. This produces an accurate representation of the distribution of the daily mean precipitation values, and an accurate account of the high precipitation values which would be classed as an extreme event. The strength in the representation of precipitation in the model, despite the presence of relative high biases in surface temperatures in some regions, shows that the model is not sensitive to biases in the external drivers. The fact that the individual weather events which are realistically represented in HadAM3P are "right for the right reasons", and not a product of the cancellation of errors in the tuning process, indicates that the results of weather@home simulations are comparable to state of the art global climate model simulations. Furthermore the distributed computing approach allows for the simulation of very large ensembles of global and regional climate, thus statistics of weather events can be obtained. This allows for the analysis of the frequency of occurrence of extreme events, which would be impossible with ensembles smaller than of the order of 100. The analysis of the spread of the ensembles and consistency check between the global and regional model reveals, in addition, that weather@home is also a good tool to investigate changes and drivers of extreme weather events especially in mid–latitudinal climates, rendering the set-up ideal for probabilistic event attribution.

Acknowledgements

Copyright © 2014 Royal Meteorological Society
Prepared using qjrms4.cls

Q. J. R. Meteorol. Soc. **00**: 2–22 (2014)

project ENSEMBLES (http://ensembles-eu.metoffice.com) and the data providers in the ECA&D project (http://www.ecad.eu)

Finally, we would like to thank all of the volunteers who have donated their computing time to climate*prediction*.net and weather@home.

References

Anderson DP. 2004. BOINC: A system for public-resource computing and storage. *5th IEEE / ACM International Workshop on Grid Computing* .

Christensen C, Aina T, Stainforth D. 2005. The challenge of volunteer computing with lengthy climate modelling simulations. *First International Conference on e-Science and Grid Technologies (e-Science 2005), 5-8 December 2005, Melbourne, Australia* .

Collins M, Tett S, Cooper C. 2001. The internal climate variability of hadcm3, a version of the hadley centre coupled model without flux adjustments. *Climate Dynamics* **17**: 61–81.

Cullen MJP. 1993. The unified forecast/climate model. *Meteorological Magazine* **122**: 81–94.

Cusack S, Edwards J, Kershaw R. 1999. Estimating the subgrid variance of saturation, and its parametrization for use in a GCM cloud scheme. *Quarterly Journal of the Royal Meteorological Society* **125**(560): 3057–3076.

Denis B, Laprise R, Caya D. 2003. Sensitivity of a regional climate model to the resolution of the lateral boundary conditions. *Climate Dynamics* **20**: 107–126.

Frame DJ, Aina T, Christensen CM, Fall NE, Knight SHE, Piani C, Rosier SM, Yamazaki K, Yamazaki Y, Allen M. 2009. The climate*prediction*.net BBC climate change experiment: Design of the coupled model ensemble. *Philosophical Transactions of The Royal Society* **367**: 855–870.

Gibson JK, Kllberg P, Uppala S, Hernandez A, Nomura A, Serrano E. 1997. ERA description. *ECMWF Re-Analysis Project Report Series No. 1* : 71pp.

Giorgi F, Jones C, Asrar GR. 2009. Addressing climate information needs at the regional level: the CORDEX framework. *World Meteorological Organization (WMO) Bulletin* **58**(3): 175–181.

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. *Climate Dynamics* **16**: 147–168.

Gregory JM. 1999. Representation of the radiative effect of convective anvils. *Hadley Centre Technical Note 7 (HCTN7)* : 20pp.

Hassell D, Jones RG. 1999. Simulating climatic change of the Southern Asian monsoon using a nested regional climate model (HadRM2). *Hadley Centre Technical Note 8 (HCTN8)* : 34pp.

Haylock M, Hofstra N, Tank AK, Klok E, Jones P, New M. 2008. A European daily high-resolution gridded dataset of surface temperature and precipitation for 1950–2006. *J. Geophys. Res (Atmospheres)* **113**(D20), doi:10.1029/2008JD010201.

Hudson DA, Jones RG. 2002. Regional climate model simulations of present-day and future climates of Southern Africa. *Hadley Centre Technical Note 39 (HCTN39)* : 40pp.

Jacob D, Brring L, Christensen O, Christensen J, de Castro M, Dqu M, Giorgi F, Hagemann S, Hirschi M, Jones R, Kjellstrm E, Lenderink G, Rockel B, Snchez E, Schr C, Seneviratne S, Somot S, van Ulden A, van den Hurk B. 2007. An inter-comparison of regional climate models for Europe: model performance in present-day climate. *Climatic Change* **81**(1): 31–52, doi:10.1007/s10584-006-9213-4, URL http://dx.doi.org/10.1007/s10584-006-9213-4.

Johns T, Durman C, Banks H, Roberts M, McLaren A, Ridley J, Senior C, Williams K, Jones A, Rickard G, *et al.* 2006. The new Hadley Centre climate model (HadGEM1): Evaluation of coupled simulations. *Journal of Climate* **19**(7): 1327–1353.

Jones A, Roberts DL, Woodage MJ, Johnson CE. 2001. Indirect sulphate aerosol forcing in a climate model with an interactive sulphur cycle. *Journal of Geophysical Research: Atmospheres (1984–2012)* **106**(D17): 20 293–20 310.

Jones R, Noguer M, Hassell D, Hudson D, Wilson S, Jenkins G, Mitchell J. 2004. *Generating high resolution climate change scenarios using PRECIS*. Met Office Hadley Centre, Exeter, UK.

Knight CG, Knight SHE, Massey N, Aina T, Christensen C, Frame DJ, Kettleborough JA, Martin A, Pascoe S, Sanderson B, Stainforth DA, , Allen MR. 2007. Association of parameter, software and hardware variation with large scale behavior across 57,000 climate models. *PNAS* **104**(30): 12 259–12 264.

Krivova N, Balmaceda L, Solanki S. 2007. Reconstruction of solar total irradiance since 1700 from the surface magnetic flux. *Astronomy & Astrophysics* **467**: 335–346.

Lei Y, Hoskins B, Slingo J. 2013. Natural variability of summer rainfall over China in HadCM3. *Climate Dynamics* **42**(1-2): 1–16.

Massey N, Aina T, Rye C, Otto FEL, Wilson S, Jones RG, Allen MR. 2012. Have the odds of warm November temperatures and of cold December temperatures in Central England changed? *Bulletin of the American Meteorological Society* **93**: 1041–1067.

Mitchell TD, Jones PD. 2005. An improved method of constructing a database of monthly climate observations and associated high-resolution grids. *Int. J. Climatology* **25**: 693–712.

Otto FEL, Massey N, van Oldenborgh GJ, Jones RG, Allen MR. 2012. Reconciling two approaches to attribution of the 2010 Russian heat wave. *Journal of Geophysical Research* **39**.

Pall P, Aina T, Stone D, Stott PA, Nozawa T, Hilberts AGJ, Lohmann D, Allen MR. 2011. Anthropogenic greenhouse gas contribution to flood risk in England and Wales in Autumn 2000. *Nature* : 382–386.

Pope V, Gallani M, Rowntree P, Stratton R. 2000. The impact of new physical parameterizations in the Hadley Centre climate model: HadAM3. *Climate Dynamics* **16**: 12–146.

Pope V, Stratton R. 2002. The processes governing horizontal resolution sensitivity in a climate model. *Climate Dynamics* **19**: 211–236.

Ramanathan V, Cess R, Harrison E, Minnis P, Barkstrom B, Ahmad E, Hartmann D. 1989. Cloud-radiative forcing and climate: Results from the earth radiation budget experiment. *Science* **243**(4887): 57–63.

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

Reynolds RW, Rayner NA, Smith TM, Stokes DC, Wang W. 2002. An improved in–situ and satellite SST analysis for climate. *Journal of climate* **15**(13): 1609–1625.

Ringer MA, Allan RP. 2004. Evaluating climate model simulations of tropical cloud. *Tellus A* **56**(4): 308–327.

Rossow WB, Schiffer RA. 1999. Advances in understanding clouds from ISCCP. *Bulletin of the American Meteorological Society* **80**(11): 2261–2287.

Rowlands D, Frame D, Ackerley D, Aina T, Booth B, Christensen C, Collins M, Faull N, Forest CE, Grandey B, Gryspeerdt E, Highwood E, Ingram W, Knight S, Lopez A, Massey N, McNamara F, Meinshausen N, Piani C, Rosier S, Sanderson B, Smith L, Stone D, Thurston M, Yamazaki K, Yamazaki H, Allen M. 2012. Broad range of 2050 warming from an observationally constrained large climate model ensemble. *Nature Geoscience* **5**(4): 256–260.

Rupp D, Mote P, Massey N, Rye C, Jones R, Allen M. 2012. Did human influence on climate make the 2011 Texas drought more probable? *Bulletin of the American Meteorology Society* **93**: 1052–1054.

Scaife AA, Copsey D, Gordon C, Harris C, Hinton T, Keeley S, O'Neill A, Roberts M, Williams K. 2011. Improved atlantic winter blocking in a climate model. *Geophysical Research Letters* **38**(23): 6pp.

Scaife AA, Woollings T, Knight J, Martin G, Hinton T. 2010. Atmospheric blocking and mean biases in climate models. *Journal of Climate* **23**(23): 6143–6152.

Sheng J, Zwiers F. 1998. An improved scheme for time-dependent boundary conditions in atmospheric general circulation models. *Climate Dynamics* **14**: 609–613.

Sillmann J, Kharin VV, Zhang X, Zwiers FW, Bronaugh D. 2013. Climate extremes indices in the cmip5 multimodel ensemble: Part 1. model evaluation in the present climate. *Journal of Geophysical Research: Atmospheres* **118**(4): 1716–1733, doi:10.1002/jgrd.50203, URL http://dx.doi.org/10.1002/jgrd.50203.

Smith R. 1990. A scheme for predicting layer clouds and their water content in a general circulation model. *Quarterly Journal of the Royal Meteorological Society* **116**(492): 435–460.

Solomon S, Qin D, Manning M, Chen Z, Marquis M, Averyt K, Tignor M, Miller H (eds). 2007. *IPCC, 2007: Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Sparrow S, Huntingford C, Massey N, Allen M. 2013. The use of very large atmospheric model ensemble to assess potential anthropogenic influence on the UK Summer 2012 high rainfall totals. *Bulletin of the American Meteorology Society* **94**: S36–S41.

Stainforth DA, Aina T, Christensen C, Collins M, Faull N, Frame DJ, Kettleborough JA, Knight S, Martin A, Murphy JM, 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 T, Qin D, Plattner G, Tignor M, Allen S, Boschung J, Nauels A, Xia Y, Bex V, Midgley P (eds). 2013. *IPCC, 2013: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Stratton RA. 1999. A high resolution AMIP integration using the Hadley Centre model HadAM2b. *Climate Dynamics* **15**: 9–28.

Uppala SM, Kållberg PW, Simmons AJ, Andrae U, Bechtold VDC, Fiorino M, Gibson JK, Haseler J, Hernandez A, Kelly GA, Li X, Onogi K, Saarinen S, Sokka N, Allan RP, Andersson E, Arpe K, Balmaseda MA, Beljaars ACM, Berg LVD, Bidlot J, Bormann N, Caires S, Chevallier F, Dethof A, Dragosavac M, Fisher M, Fuentes M, Hagemann S, Hólm E, Hoskins BJ, Isaksen L, Janssen PAEM, Jenne R, McNally AP, Mahfouf J, Morcrette J, Rayner NA, Saunders RW, Simon P, Sterl A, Trenberth KE, Untch A, Vasiljevic D, Viterbo P, Woollen J. 2005. The ERA-40 re-analysis. *Quarterly Journal of the Royal Meteorological Society* **131**: 2961–3012, doi: 10.1256/qj.04.176.

van der Linden P, Mitchell JFB (eds). 2009. *ENSEMBLES: Climate Change and its Impacts: Summary of research and results from the ENSEMBLES project*. Met Office Hadley Centre, FitzRoy Road, Exeter EX1 3PB, UK.

Wood R, Field PR. 2000. Relationships between total water, condensed water, and cloud fraction in stratiform clouds examined using aircraft data. *Journal of the atmospheric sciences* **57**(12): 1888–1905.

Woollings T, Hoskins B, Blackburn M, Hassell D, Hodges K. 2010. Storm track sensitivity to sea surface temperature resolution in a regional atmosphere model. *Climate Dynamics* **35**: 341–354.

Xie P, Arkin PA. 1997. Global precipitation: A 17-year monthly analysis based on gauge observations, satellite estimates, and numerical model outputs. *Bulletin of the American Meteorological Society* **78**(11): 2539–2558.