Using the mid-Holocene 'greening' of the Sahara to narrow acceptable ranges on climate model parameters

Peter O. Hopcroft[1], Paul Valdes[2], and William Ingram[3]

[1]University of Birmingham
[2]University of Bristol
[3]Met Office

November 23, 2022

Abstract

During the early to mid-Holocene vegetation expanded to cover much of the present-day Sahara. Although driven by a well-understood difference in the orbital configuration, general circulation models have generally failed to simulate the required rainfall increase. One possible explanation is the presence of systematic biases in the representations of atmospheric convection which might also impact future projections. We employ a Bayesian method to learn from an ensemble of present day and mid-Holocene simulations that vary parameters in the convection, boundary layer and cloud schemes. The model can reproduce the 'Green Sahara' rainfall if mixing between convective plumes and the environment is increased in the upper troposphere relative to lower down. This does not appreciably impact the present day simulation, meaning that the palaeoclimate reconstructions are able to narrow constraints on suitable parameter ranges. This suggests that other uncertain components of climate models could be targeted in this way.

Using the mid-Holocene 'greening' of the Sahara to narrow acceptable ranges on climate model parameters

Peter O. Hopcroft [1], Paul J. Valdes [2], and William Ingram [3]

[1]School of Geography, Earth & Environmental Sciences, University of Birmingham, Birmingham, B15 2TT, U.K.
[2]Bristol Research Initiative for the Dynamic Global Environment, School of Geographical Sciences, University of Bristol, University Road, Bristol, BS8 1SS, U.K.
[2]Met Office Hadley Centre, FitzRoy Road, Exeter, EX1 3PB, U.K.

Key Points:

- A wide range of proxy data agree in showing a mid-Holocene "green Sahara".
- General circulation models (GCMs) struggle to simulate this.
- Bayesian tuning of a GCM succeeds for the mid-Holocene, finding the improvement has little impact on the present-day simulation.dual

Corresponding author: Peter Hopcroft, `p.hopcroft@bham.ac.uk`

Abstract

During the early to mid-Holocene vegetation expanded to cover much of the present-day Sahara. Although driven by a well-understood difference in the orbital configuration, general circulation models have generally failed to simulate the required rainfall increase. One possible explanation is the presence of systematic biases in the representations of atmospheric convection which might also impact future projections. We employ a Bayesian method to learn from an ensemble of present day and mid-Holocene simulations that vary parameters in the convection, boundary layer and cloud schemes. The model can reproduce the 'Green Sahara' rainfall if mixing between convective plumes and the environment is increased in the upper troposphere relative to lower down. This does not appreciably impact the present day simulation, meaning that the palaeoclimate reconstructions are able to narrow constraints on suitable parameter ranges. This suggests that other uncertain components of climate models could be targeted in this way.

Plain Language Summary

General circulation models are complex computational representations of the Earth's climate system. Run on supercomputers, these can be used to predict future climate change. Past climate changes can also be used to test climate models. One example of this is the greening of the Sahara around 11,000 to 4,000 years ago. Almost all models fail to capture the amplitude of the so-called 'Green' Sahara. One possible reason for this is that small scale features such as clouds and storms in the atmosphere must be approximated using parametrizations. These parametrizations are poorly constrained by available climate observations and they thus potentially introduce errors in predictions of past or future climate changes. In this work we show that the greening can be simulated accurately when the parametrizations are tuned not only with present day observed climate fields, but additionally with the past green Sahara state. This suggests that climate model parametrizations may be significantly improved and uncertainties reduced if climate states from the past are used in developing climate models.

1 Introduction

The hydrological cycle response in a warming climate will be a major driver of future socio-economic impacts (Hoegh-Guldberg et al., 2013). Projections from general circulation models (GCMs: the most detailed and physically-based models of the global climate system) consistently predict a warmer future almost everywhere, but precipitation projections are much more divergent, especially in the tropics (Chadwick et al., 2016; Allen & Ingram, 2002; Kent et al., 2015; Rowell et al., 2016).

Much of the uncertainty in future precipitation change is related to processes associated with clouds and convection. In reality these physical processes occur over spatial scales up to 10s of km. In contrast, most GCMs have a resolution of around 150-200 km. Processes such as convection must therefore be parametrized, but all parametrizations are approximations, and the structure of many parametrizations is unavoidably far simpler than reality, leaving no way to choose a parameter value a priori or directly from observations (e.g. Stensrud, 2007)

Palaeoclimate changes can provide valuable 'out-of-sample' tests for modelling the climate system (Valdes, 2011; Schmidt et al., 2014; Harrison et al., 2015). This is because past climate states provide examples of both fast and gradual changes that are larger in amplitude than historical climate changes and are therefore more comparable in amplitude to expected future climate change. To be useful in this way, a palaeoclimate state or transition must be associated with a good understanding of both the underlying forcings (e.g. a change in greenhouse gas levels or a change in orbit) and the resultant impacts in the climate system.

The early to mid-Holocene (around 11,000-4,000 years before present) is frequently highlighted as such a period (Braconnot et al., 2012; Harrison et al., 2015; Biasutti et al., 2018). This is because the ultimate forcing during this time was a very well-understood change in the configuration of Earth's orbit (Berger & Loutre, 1991). The resulting increase in northern hemisphere summer insolation drove enhanced monsoon circulation and precipitation (Kutzbach & Street-Perrott, 1985). This led to a so-called 'Greening' of the Sahara (e.g. Claussen et al., 2017). This is evidenced by the development of savanna or steppe-like vegetation (Hély et al., 2014), expansion of lakes and rivers (Kohfeld & Harrison, 2000; Skonieczny et al., 2015), a reduction in dust deposited over the Atlantic (de Menocal et al., 2000; Williams et al., 2016) and the presence of neolithic settlements and domesticated animals (Manning & Timpson, 2014).

Pollen and macro-fossil evidence suggests that annual mean precipitation increased by 1.1 mm day$^{-1}\pm0.1$ (Bartlein et al., 2011) relative to the present day (mean$\pm$standard error for 15-30°N). There is some uncertainty on the spatial pattern of change, with earlier compilations of pollen suggesting a relatively uniform vegetation change (Hoelzmann et al., 1998), while more recent datasets suggest greater changes in vegetation in the South compared with further North (Hély et al., 2014). The pollen samples have been integrated with a vegetation model to infer climate, showing the same overall precipitation increase (Wu et al., 2007). A larger rainfall increase of around 1.5 (0.9-2.8) mm day^{-1} has been inferred from marine core leaf-wax hydrogen isotope ratios (Tierney et al., 2017). Despite these uncertainties, all lines of evidence agree on a minimum increase in precipitation of at least 0.7 mm day^{-1}(Joussaume et al., 1999) that enabled vegetation to grow across much of the present-day Sahara (Ritchie & Haynes, 1987; Street-Perrott et al., 1990; Pachur & Holzmann, 1991; Jolly et al., 1998; Peyron et al., 2006).

All GCM simulations driven with the orbital configuration for 6,000 years before present (6ka BP), simulate an increase in precipitation, but almost always much smaller than these pollen observations imply over the Sahara itself (Joussaume et al., 1999; Braconnot et al., 2007, 2012; Brierley et al., 2020). This is also true when dynamic vegetation and/or dynamic dust processes are enabled (Perez-Sanz et al., 2014; Harrison et al., 2015; Hopcroft & Valdes, 2019). In contrast, when significant changes in the land surface albedo and/or significant reductions in dust aerosols are specified, sufficient rainfall can be simulated (e.g. Levis et al., 2004; Skinner & Poulsen, 2016; Pausata et al., 2016).

Saharan dust particles are less absorbing than is prescribed in most climate models which tend to use outdated radiative properties (Hopcroft & Valdes, 2019; Albani & Mahowald, 2019). A significant dust reduction during the mid-Holocene probably did not appreciably enhance convective precipitation (Hopcroft & Valdes, 2019). Moreover, the reduction in dust loading would have altered cloud formation through dust-cloud interactions, and this has been shown to reduce stratiform precipitation (Thompson et al., 2019). Land-surface feedbacks can efficiently drive the monsoon northwards (Texier et al., 2000; Levis et al., 2004; Skinner & Poulsen, 2016) but there is little agreement about how the 'greening' of the Sahara should be configured in models (Street-Perrott et al., 1990; Texier et al., 2000; Hopcroft et al., 2017; Lu et al., 2018; Chen et al., 2020). It is thus not trivial to judge whether or not a sufficient precipitation enhancement is achieved for the right reasons in model simulations of the mid-Holocene. The model-data disparity over North Africa has persisted for several decades across multiple GCMs (Biasutti et al., 2018). This suggests systematic biases that either require more detailed physical representations or a different approach to parameter choices.

In this work we use the atmospheric component of the coupled GCM HadCM3 (Pope et al., 2000; Gordon et al., 2000; Valdes et al., 2017) which also does not simulate a 'greening' of the Sahara under mid-Holocene boundary conditions (Braconnot et al., 2007). We use this GCM to evaluate what the model failure may reveal about the representation of precipitation in GCMs and to compare the parametric constraints from present-day and mid-Holocene climatic conditions.

2 Methods

2.1 General circulation model and boundary conditions

We use the Met Office Hadley Centre atmosphere model 3 (HadAM3) with the MOSES 2.1 land surface scheme and prescribed vegetation cover (Pope et al., 2000; Essery et al., 2003; Valdes et al., 2017), specifically HadAM3B-M2.1aN. This GCM has horizontal resolution of $3.75\times2.5°$ (longitude-latitude) with 19 vertical levels. HadAM3 uses the mass-flux convection scheme by Gregory & Rowntree (1990) which is comparable in complexity to schemes used in several other GCMs (Stensrud, 2007; Maraun & Widmann, 2018). Relative to the published configuration of the model (here labelled ORIG), a revised (REV) model version was developed here that includes a humidity-dependence of the mixing and forced detrainment from convection following (Derbyshire et al., 2011) as implemented in more recent Hadley Centre models.

The pre-industrial setup follows that of Valdes et al. (2017) with prescribed observed present-day vegetation coverage (Loveland et al., 2000) and pre-industrial levels of greenhouse gases (CO_2, CH_4 and N_2O). We use 1981-2010 climatological sea-surface temperatures (SSTs) and sea-ice from HadISST (Rayner et al., 2003, updated to 2010). For the mid-Holocene, the orbital parameters are modified for the conditions of 6ka before present (BP) (Berger & Loutre, 1991). Sea surface temperatures (SST) and sea-ice are modified by adding the 6 ka - pre-industrial difference, as simulated with the coupled model HadCM3B-M2.1aD, to the pre-industrial HadISST climatology. The simulations setup is summarised in table S1.

Today the Sahara has a surface albedo of 0.35 (Loeb et al., 2012). A reduction in albedo would strengthen the monsoon (Charney, 1975; Street-Perrott et al., 1990; Texier et al., 2000; Boos & Storelvmo, 2016). The mid-Holocene 'greening' involved the northwards expansion of grasses and shrubs (Jolly et al., 1998; Hély et al., 2014). Satellite observations show that these biomes have an albedo of 0.17 - 0.3 when precipitation is in the range reconstructed for the 'greening' (i.e. 1.1 mm day^{-1} Bartlein et al., 2011), see Supporting Information S1 and figure S1. The Sahel which is at the periphery of the present-day West African monsoon is in the upper part of this range (0.2-0.3). This may present the best analogue for the mid-Holocene 'greening'. This higher end is also consistent with mid-Holocene simulations with dynamic vegetation and soils (e.g. Claussen & Gayler, 1997; Vamborg et al., 2011). Many model studies have prescribed a value at the very lower end of 0.15 (Levis et al., 2004; Pausata et al., 2016; Skinner & Poulsen, 2016; Chandan & Peltier, 2020) which is typically seen in regions of higher rainfall of 2.5-3.0 mm day^{-1}.

We do not use the HadCM3B-M2.1aD dynamic vegetation scheme as it is overly sensitive to arid conditions (Hopcroft et al., 2017). Instead bare soil in the Sahara region (from 10-35°N, 30°W-50°E) is replaced with grasses and shrubs with a total fractional coverage of 50%. This produces a surface albedo of 0.27 relative to 0.31 in the pre-industrial simulation. This a relatively conservative change for the period since it is at the upper end of the range discussed above.

2.2 Perturbed parameter ensemble

We introduce a new variable E into HadAM3, which controls the vertical dependence of entrainment and mixing detrainment - the mixing of environmental air into the convecting air, and of convecting air into its environment. By default, the entrainment rate decays with altitude in proportion to pressure. This was intended as an ad-hoc representation of larger clouds, which proportionally mix less with their surrounding, reaching higher (Gregory & Rowntree, 1990). With the new parameter E we relax this assumption. Increasing E increases the upper troposphere entrainment values and reduces those near to the land surface. A value of zero returns the default proportional dependence on pressure (see Supporting Information S2).

We configured a 150-member perturbed parameter ensemble using the REV configuration of HadAM3. Eleven GCM parameters within the convection, boundary layer or large-scale cloud schemes, including three new parameters. These are E which controls the vertical profile of entrainment and detrainment, r_{det} which sets the sensitivity of forced detrainment to the buoyancy gradient (Derbyshire et al., 2011), and α_{det} which sets the sensitivity of detrainment to relative humidity (Derbyshire et al., 2011). The 11 model parameters are assigned different values globally leading to 150 paired ensemble members of pre-industrial and mid-Holocene simulations. The parameter definitions and ranges used are given in table S2 and illustrated schematically in figure S2. The evaluations are selected using a Latin hypercube method (McKay et al., 1979), which distributes the parameter samples optimally across the 11-dimensional state space.

2.3 Statistical modelling and parameter tuning

We used a Gaussian process emulator (e.g. Kennedy & O'Hagan, 2001) to construct a statistical representation of the perturbed parameter ensemble of GCM simulations. Emulators have been extensively used in analysing complex numerical models like GCMs (Rougier et al., 2009; Sexton et al., 2012; McNeall et al., 2016; Edwards et al., 2019). The emulator represents some output as a linear function of the input parameter values combined with a Gaussian process (Roustant et al., 2012). In this way the emulator interpolates in multi-dimensional parameter space to predict the GCM response at any combination of input parameter values. Further details are given in Supporting Information text S3.

We use a Bayesian method (see Supporting Information text S3.2) to update the model parameters based on the mid-Holocene palaeoclimate reconstructions. In a Bayesian method we compute a posterior probability distribution function (PDF) on the model parameters based on the prior PDF and the likelihood (e.g. Rougier, 2007). The prior is taken as the current model version and the likelihood quantifies the performance of the GCM for selected climate outputs such as simulated precipitation. Thus we condition the model parameters with the present-day observed climate variables and optionally the mid-Holocene rainfall increase.

The posterior PDF must be approximated using a Markov chain Monte Carlo method (Gilks et al., 1995) and since the MCMC algorithm requires many thousands of iterations, we use the emulator in place of the full GCM. We perform this process twice. Firstly including four observational targets for the present day (table S3) and secondly adding to this the mid-Holocene absolute precipitation over North Africa inferred from pollen data (Bartlein et al., 2011). Thus we derive a new parameter set suitable for both present and mid-Holocene conditions, which is different to Su & Neelin (2005) who used different parameter sets for the two time periods.

3 Results

3.1 Sampling convection, clouds and boundary layer within a global model

The resultant precipitation anomalies for the 112 simulations that completed 50 model years are averaged over North Africa (20°W-30°E by 15-30°N) in figure 1. This region of North Africa includes many of the fossil pollen sites. All model simulations overestimate present-day precipitation in Africa and in North Africa in particular. This is a systematic bias in HadAM3 (Valdes et al., 2017). Part of which is due to an underestimation of the soil albedo in the Sahara region in the model. The simulated difference mid-Holocene minus pre-industrial in precipitation in this region ranges from 0.7 mm day^{-1} to 2.6 mm day^{-1} for JJAS, when most precipitation falls. Many ensemble members with the pre-industrial precipitation similar to the original (around 0.5 mm day^{-1}), have much higher increases of about 2.0 mm day^{-1} for the mid-Holocene. The weak correlation between the two axes in figure 1 shows that different factors influence precipitation in the pre-industrial compared with the precipitation difference between the two time periods.

The parameter dependence of the mid-Holocene precipitation anomaly is shown in figure S3. The most obvious relationship is with E, for which higher values result in larger changes. E controls the vertical profile of entrainment and detrainment, which is the rate of mixing of convective clouds with the surrounding air masses. In many models entrainment decays with altitude. High values of E increase the upper level entrainment and reduces it near the land surface. This produces a wetter mid-Holocene in North Africa.

The Gaussian process emulator is used to calculate influence of varying each parameter value individually on the simulated precipitation change over North Africa for the mid-Holocene. The result of this is shown in figure 2 and the emulator skill is evaluated in figure S4. We find that E is the dominant parameter. We examined the parameter dependence of the West African monsoon in the pre-industrial ensemble members (over 5-15°N) (figure S5). This shows that the parameters which exert the strongest control on precipitation in this monsoon region (q_{ini}, ct and α_{det}) are not the same as for the mid-Holocene anomaly relative to the pre-industrial simulation (E, F, T_{ini} and q_{ini}).

3.2 Mechanisms of enhanced mid-Holocene precipitation

Given the profound effect of changing E on the mid-Holocene North African precipitation, we ran a simulation with only one change from REV: increasing E from its default value of 0 to 0.25. Figure 3a shows the percentage difference in the mid-Holocene minus pre-industrial (6kaGS-0ka) precipitation anomalies for the pair of simulations with E=0.25 compared to the pair with E=0. With E=0.25 the precipitation anomaly is generally larger across North Africa and is nearly twice as large in the North West (figure 3a). The latitude of the precipitation maximum moves northwards by around 2.5-5° compared to the E=0 simulation (not shown). It produces a more diffuse precipitation band during JJAS which pushes the periphery of the monsoon further into the dry interior.

A key diagnostic of the convection scheme is the updraught mass-flux. The simulated mean convective updraught over North Africa decreases fractionally much more with height than the mean over the wet regions of the tropics as a whole. This is presumably because of dilution by the extremely dry environment in North Africa, which makes it harder for moist convective plumes to persist. In all model versions the increase in rainfall over North Africa at 6 ka BP is accompanied by the updraught mass flux weakening lower down and strengthening aloft (figure 3b), becoming more like tropical moist convection elsewhere in the tropics. The mid-Holocene boundary conditions lead to less dilution of convective plumes low down, so that those that reach their lifting condensation level (LCL: indicated by vertical lines in figure 3b) are more vigorous and end up penetrating higher overall.

The direct effect of increasing the parameter E, i.e. decreasing the entrainment rate near the surface and increasing entrainment rate higher up, is to amplify these mass flux changes for the mid-Holocene relative to the pre-industrial (figure 3b), so that the REV(E=0.25) case has a lower mass-flux near the surface than the REV model version, and a stronger updraught mass flux above the lifting condensation level (LCL: indicated by vertical lines in figure 3b). Interactions between convection and its environment mean the net effect can be very different in other regions (figure 3a), but over North Africa these changes reinforce each other to produce larger amplitude massflux changes, and a stronger rainfall increase at 6 ka BP (figure 3d).

In tandem with this, the circulation (zonal wind) and humidity anomalies associated with heavier downpours in the Sahara are different when E is given a higher value (figure 3c and 3d). For E=0, wetter days north of 15°N are associated with a strengthened tropical easterly jet (TEJ) and a slightly weakened Africa easterly jet (AEJ), as observed (Nicholson, 2009). African Easterly Waves (AEWs) move along the jet and contribute precipitation to the North (Claussen et al., 2017) and this is well represented in HadAM3 (Taylor et al., 2002). For E=0.25, the same precipitation increase is achieved with a 20% smaller increase in the TEJ. This suggests that convection is more effective in this model configuration and this partly explains the increased precipitation response for E=0.25. Some of these downpours are also associated with tropical plumes (Knippertz, 2003), especially in the months of August-October (Skinner & Poulsen, 2016; Dallmeyer et al., 2020). When E is increased to 0.25, plumes contribute more precipitation in the region from 20-35°N despite relatively similar mean climatologies of large-scale circulation and humidity.

3.3 Learning from the model-data mismatch

The mid-Holocene pollen quantitative precipitation reconstruction over North Africa gives an annual-mean precipitation increase of 1.1 ± 0.1 mm day^{-1} (Bartlein et al., 2011) relative to the present day. We use the annual mean reconstruction and model outputs in a probabilistic formulation to optimise the GCM so that it is consistent with the pollen-based precipitation reconstructions and hence the widespread environmental evidence for an invigoration of the hydrological cycle.

The posterior PDFs on the 11 parameters are shown in figure S6. Two cases are considered where the second only differs with the inclusion of the mid-Holocene precipitation target. For the mid-Holocene the algorithm favours high E because as discussed above, it has an extremely strong impact on the response to the mid-Holocene insolation. The emulator predicted mid-Holocene precipitation increase is very different between the two cases, showing sensitivity of the system to parameter combinations and also that the optimisation against present-day observations does not guarantee an improvement for the mid-Holocene.

3.4 New model version

One optimised parameter set derived from the PDFs on the model parameters (figure S6 and tables S4 and S5) was used in new pre-industrial and mid-Holocene GCM simulations and is denoted REVopt. In this we only changed parameters from their original GCM values where there is stronger preference posterior PDF. Whilst the choice of parameters which underline REVopt is based on the posterior PDF sampling, it would be more consistent with the Bayesian paradigm to think in terms of the probability distribution on the parameters, rather than to focus on any single parameter set. However, given computational limitations and to simplify the presentation of the results we mostly focus on the REVopt parameter set.

This simulation is compared with the ORIG and REV models in figure 4. The difference between the two controls (ORIG and REV) is due to changes to the detrainment parametrization following Derbyshire et al. (2011). Figure 4(c) shows that in REVopt the precipitation anomaly over North Africa is approaching double that in the ORIG simulations for JJAS, and the annual mean anomaly is also 67% larger. There is a large region of precipitation increase across North Africa and Arabia and Northern India. This shows that the statistically-based parameter changes also work when re-introduced into the GCM. Crucially, the present-day performance of REVopt is very similar to ORIG and REV for both temperature and precipitation (Supporting Information text S4 and figures S7 and S8). This means that the improvement for the mid-Holocene has not significantly altered the present-day simulation.

4 Discussion

The strong precipitation increase during the early- to mid-Holocene in North Africa presents a unique challenge to climate simulations of tropical precipitation. In this perturbed parameter study, we find that parameters controlling the pre-industrial climatology of precipitation are different from those that determine the anomaly under a climate change scenario. The Bayesian approach demonstrates that a modified vertical profile of convective entrainment can significantly improve the simulation of the mid-Holocene North Africa despite having little impact the simulation of the present day.

Like any model, HadCM3 has biases and simplifications. For example, HadCM3 suffers from having too little, too optically bright cloud cover (Massey et al., 2015), a common problem in CMIP5 models (Nam et al., 2012). It also has too much precipitation over Africa and too little over South America, but does not suffer from a double ITCZ bias or weak ENSO variability, which are common problems in many GCMs. Overall its performance compared to observations is typical of GCMs used in recent intercomparisons (Valdes et al., 2017).

Structural limitations mean that many biases could be corrected by varying model parameters, and there are many more than the 11 we varied. Also, the existence of compensating errors means that tuning that improves one bias can actually exacerbate another. Despite this, in the REVopt case we significantly improved the mid-Holocene precipitation in comparison with reconstructions, without affecting the simulation of the present day state. It is possible that with a more comprehensive list of parameters, e.g. of the order of 20-50, some of these other biases may be reduced.

Tropical precipitation in GCMs has recently been improved through the use of adaptive convective entrainment, whereby local entrainment rates reduce following convective activity. On a local scale this could produce a similar effect as in our study, reducing entrainment in the lower troposphere, following precursor convective plumes (Mapes & Neale, 2011; Willett & Whitall, 2017). Future work should consider how such dynamic entrainment parametrizations (e.g. Mapes & Neale, 2011; Hohenegger & Bretherton, 2011) could similarly improve modelling of the mid-Holocene and whether this is consistent with our statistically-derived model changes.

Convection-permitting atmospheric model simulations, with high resolution and no convection parametrization, have highlighted further significant improvements when convection parametrizations are deactivated (Marsham et al., 2013; Birch et al., 2014; Kendon et al., 2019; Finney et al., 2019; Berthou et al., 2019; Pante & Knippertz, 2019). This includes a more realistic diurnal cycle of precipitation (Marsham et al., 2013), improved simulation of wet spells (Berthou et al., 2019; Kendon et al., 2019) and of cloud cover and humidity (Pante & Knippertz, 2019). Making this transition is not without drawbacks and substantial model errors can emerge (Pante & Knippertz, 2019) that are difficult to eliminate because of the high computational cost of test simulations. Future work could compare convection-permitting simulations and ensembles of GCM simulations like those studied here for both present day and palaeoclimate conditions to identify further ways to improve GCM parametrization schemes.

5 Conclusions

Most climate models are developed and calibrated against modern climate. We have shown that there are multiple parameter sets which allow for a good simulation of the present day conditions, but that only a subset of these are also able to satisfy past criteria. This example provides a new quantitative demonstration of how a palaeoclimate state may provide information of relevance to uncertainty in simulating future precipitation change. Palaeoclimate may therefore be a significantly undervalued source of additional information for informing the parameter values and parametrization choices in GCMs as it has rarely been used in this way, see examples by Hopcroft & Valdes (2015) and DiNezio et al. (2016). Well documented climate states in the past thus may have the potential to be used in model development and this approach can include other important palaeoclimate changes.

Acknowledgments

POH was funded by the EPSRC-funded Past Earth Network Feasibility Study (Grant number EP/M008363/1) and by a University of Birmingham Fellowship. WI was supported by the Met Office Hadley Centre Climate Programme funded by BEIS and Defra. This work was carried out using the computational facilities of the Birmingham Environment for Academic Research - `http://www.bear.bham.ac.uk` and the Advanced Computing Research Centre, University of Bristol - `http://www.bris.ac.uk/acrc`. We thank Gethin Williams for installing the climate model code on BEAR.

The Met Office released the HadCM3 source code via the Ported Unified Model release (`https://www.metoffice.gov.uk/research/approach/collaboration/unified-model/partnership`, and um_collaboration@metoffice.gov.uk). Code modifications required to produce the ORIG version of HadAM3/ HadCM3 are available from `https://doi.org/10.5194/gmd-10-3715-2017`. The HadAM3/HadCM3 code changes for the REV and REVopt model versions and the parameter namelist files required to configure these simulations are available from figshare: doi: 10.6084/m9.figshare.12311360.

The ensemble simulation output analysed in the study are available from `www.paleo.bristol.ac.uk/ummodel/scripts/papers/`. The netcdf climatologies of the pre-industrial and mid-Holocene ORIG, REV and REVopt simulations are available from figshare doi: 10.6084/m9.figshare.12505259.

The statistical emulator is available in the R package DiceKriging. The Bartlein et al. (2011) mid-Holocene pollen-based climate reconstruction is available from `pmip3.lsce.ipsl.fr`. The Tierney et al. (2017) precipitation reconstruction is available from `http://www.ncdc.noaa.gov/paleo/study/21091`. CRU precipitation is available from `https://crudata.uea.ac.uk/cru/data/hrg/`. CERES land surface albedo is available from `https://ceres.larc.nasa.gov/data/`. GPCP precipitation is available from `dx.doi.org/10.7289/V56971M6`.

References

Albani, S., & Mahowald, N. (2019). Paleodust Insights into Dust Impacts on Climate. *J Clim*, *32*, 7897-7913. doi: 10.1175/JCLI-D-18-0742.1

Allen, M., & Ingram, W. (2002). Constraints on future changes in climate and the hydrologic cycle. *Nature*, *419*, 224-232.

Bartlein, P. J., Harrison, S., Brewer, S., Connor, S., Davis, B., Gajewski, K., ... Wu, H. (2011). Pollen-based continental climate reconstructions at 6 and 21 ka: a global synthesis. *Climate Dynamics*, *37*(3-4), 775-802. doi: 10.1007/s00382-010-0904-1

Berger, A., & Loutre, M. (1991). Insolation values for the climate of the last 10 million years. *Quaternary Science Reviews*, *10*(4), 297-317.

Berthou, S., Rowell, D., Kendon, E., Roberts, M., Stratton, R., , et al. (2019). Improved climatological precipitation characteristics over West Africa at convection-permitting scales. *Clim Dyn*, *53*, 1991-2011. doi: 10.1007/s00382-019-04759-4

Biasutti, M., Voigt, A., Boos, W., et al. (2018). Global energetics and local physics as drivers of past, present and future monsoons. *Nature Geoscience*, *11*, 392-400. doi: 10.1038/s41561-018-0137-1

Birch, C., Parker, D., J, M., et al. (2014). A seamless assessment of the role of convection in the water cycle of the West African Monsoon. *J Geophys Res*, *119*, 2890-2912. doi: 10.1002/2013JD020887

Boos, W., & Storelvmo, T. (2016). Near-linear response of mean monsoon strength to a broad range of radiative forcings. *Proc Natl Acad Sci*, *113*, 1510-1515.

Braconnot, P., Harrison, S., Kageyama, M., Bartlein, P., Masson-Delmotte, V., Abe-Ouchi, A., ... Zhao, Y. (2012). Evaluation of climate models using palaeoclimatic data. *Nature Climate Change*, *2*, 417-424. doi: 10.1038/nclimate1456

Braconnot, P., Otto-Bliesner, B., Harrison, S., Joussaume, S., Peterchmitt, J.-Y., Abe-Ouchi, A., ... Zhao, Y. (2007). Results of PMIP2 coupled simulations of the Mid-Holocene and Last Glacial Maximum Part 1: experiments and large-scale features. *Clim Past*, *3*, 261-277.

Brierley, C., Zhao, A., Harrison, S., et al. (2020). Large-scale features and evaluation of the PMIP4-CMIP6 midHolocene simulations . *Clim Past*, *16*, 1847-1872.

Chadwick, R., Good, P., Martin, G., & Rowell, D. (2016). Large rainfall changes consistently projected over substantial areas of tropical land. *Nature Climate Change*, *6*, 177-182. doi: {10.1038/NCLIMATE2805}

Chandan, D., & Peltier, W. (2020). African humid period precipitation sustained by robust 2 vegetation, soil and lake feedbacks. *Geophys Res Lett*, . doi: 10.1029/2020GL088728

Charney, J. (1975). Dynamics of deserts and drought in the sahel. *Quart J Roy Meteor Soc*, *101*, 193-202.

Chen, W., Ciais, P., Zhu, D., et al. (2020). Feedbacks of soil properties on vegetation during the Green Sahara period. *Quat Sci Rev*, *240*, 106389. doi: 10.1016/j.quascirev.2020.106389

Claussen, M., Dallmeyer, A., & Bader, J. (2017). Theory and Modeling of the African Humid Period and the Green Sahara. *Oxford Res Encyclo Clim Sci*, 1-38. doi: {10.1093/acrefore/9780190228620.013.532}

Claussen, M., & Gayler, V. (1997). The Greening of the Sahara during the Mid-Holocene: Results of an Interactive Atmosphere- Biome Model. *Global Ecology and Biogeography Lett*, *6*(5), 369-377.

Dallmeyer, A., Claussen, M., Lorenz, S., & Shanahan, T. (2020). The end of the African humid period as seen by a transient comprehensive Earth system model simulation of the last 8000 years. *Clim Past*, *16*, 117-140. doi: 10.5194/cp-16-117-2020

de Menocal, P., Ortiz, J., Guilderson, T., Adkins, J., Sarnthein, M., Baker, L., &

Yarusinsky, M. (2000). Abrupt onset and termination of the African Humid Period: rapid climate responses to gradual insolation forcing. *Quat Sci Rev*, *19*, 347-361.

Derbyshire, S., Maidens, A., Milton, S., Stratton, R., & Willett, M. (2011). Adaptive detrainment in a convective parametrization. *Q J Royal Met Soc*, *137*, 1856-1871.

DiNezio, P., Timmermann, A., Tierney, J., et al. (2016). The climate response of the Indo-Pacific warm pool to glacial sea level. *Paleoceanography*, *31*, 866-894. doi: 10.1002/2015PA002890

Edwards, T., Brandon, M., Durand, G., et al. (2019). Revisiting Antarctic ice loss due to marine ice-cliff instability. *Nature*, *566*, 58-64. doi: 10.1038/s41586-019-0901-4

Essery, R., Best, M., Betts, R., Cox, P., & Taylor, C. (2003). Explicit representation of subgrid heterogeneity in a GCM land-surface scheme. *J. Hydrometeorol*, *4*, 530-543.

Finney, D., Marsham, J., Jackson, L., et al. (2019). Implications of Improved Representation of Convection for the East Africa Water Budget Using a Convection-Permitting Model. *J Climate*, *32*, 2109-2129. doi: 10.1175/JCLI-D-18-0387.1

Gilks, W., Richardson, S., & Spiegelhalter, D. (1995). Introducing markov chain monte carlo methods. In Gilks, WR and Richardson, S and Spiegelhalter, DJ (Ed.), *Markov Chain Monte Carlo in Practice*. London: Chapman and Hall.

Gordon, C., Cooper, C., Senior, C. A., Banks, H., Gregory, J. M., Johns, T. C., ... Wood, R. A. (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*(2-3), 147-168.

Gregory, D., & Rowntree, P. (1990). A mass-flux convection scheme with representation of cloud ensemble characteristics and stability dependent closure. *Mon. Weather Rev.*, *118*, 1483-1506.

Harris, I., Jones, P., Osborn, T., & Lister, D. (2014). Updated high-resolution grids of monthly climatic observations - the CRU TS3.10 Dataset. *Int J Climtol*, *34*, 623-642.

Harrison, S., Bartlein, P., Izumi, K., Li, G., Annan, J., Hargreaves, J., ... Kageyama, M. (2015). Evaluation of CMIP5 palaeo-simulations to improve climate projections. *Nature Clim Change*, *5*, 735-743.

Hély, C., Lezine, A.-M., & APD contributors. (2014). Holocene changes in African vegetation: tradeoff between climate and water availability. *Clim Past*, *10*, 681-686. doi: 10.5194/cp-10-681-2014

Hoegh-Guldberg, O., Jacob, D., Taylor, M., Bindi, M., Brown, S., et al. (2013). Impacts of 1.5C Global Warming on Natural and Human Systems. In V. Masson-Delmotte, P. Zhai, H.-O. Portner, D. Roberts, et al. (Eds.), *Global Warming of 1.5C. An IPCC Special Report on the impacts of global warming of 1.5C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty* (p. In press). Cambridge, U.K. and New York, U.S.A.: C.U.P.

Hoelzmann, P., et al. (1998). Mid-Holocene land-surface conditions in northern Africa and the Arabian peninsula: A data set for the analysis of biogeophysical feedbacks in the climate system. *Glob Biogeochem Cyc*, *12*, 35-51.

Hohenegger, C., & Bretherton, C. (2011). Simulating deep convection with a shallow convection scheme. *Atmos Chef Phys*, *11*, 10389-10406.

Hopcroft, P., & Valdes, P. (2015). Last Glacial Maximum constraints on the Earth System model HadGEM2-ES. *Climate Dynamics*, *45*(5), 1657-1672. doi: 10.1007/s00382-014-2421-0

Hopcroft, P., & Valdes, P. (2019). On the role of dust-climate feedbacks during the mid-Holocene. *Geophysical Research Letters*, *46*, 1612-1621. doi: 10.1029/

2018GL080483
Hopcroft, P., Valdes, P., Harper, A., & Beerling, D. (2017). Multi vegetation model evaluation of the Green Sahara climate regime. *Geophys Res Lett*, *44*, 6804-6813. doi: 10.1002/2017GL073740
Jolly, D., Prentice, I., Bonnefille, R., et al. (1998). Biome reconstruction from pollen and plant macrofossil data for Africa and the Arabian peninsula at 0 and 6000 years. *J Biogeography*, *25*, 1007-1027.
Joussaume, S., et al. (1999). Monsoon changes for 6000 years ago: Results of 18 simulations from the Paleoclimate Modeling Intercomparison Project (PMIP). *Geophys Res Lett*, *26*(7), 859-862.
Kendon, E., Stratton, R., Tucker, S., et al. (2019). Enhanced future changes in wet and dry extremes over Africa at convection-permitting scale. *Nature Comms*, *10*, 1794.
Kennedy, M., & O'Hagan, A. (2001). Bayesian calibration of computer models. *J R Statist Soc B*, *63(3)*, 425-464.
Kent, C., Chadwick, R., & Rowell, D. (2015). Understanding uncertainties in future projections of seasonal tropical precipitation. *J Clim*, *28*, 4390-4413.
Knippertz, P. (2003). Tropical-extratropical interactions causing precipitation in northwest Africa: Statistical analysis and seasonal variations. *Mon Weather Rev*, *131*, 3069-3076.
Kohfeld, K., & Harrison, S. (2000). How well can we simulate past climates? Evaluating the models using palaeoenvironmental datasets. *Quaternary Science Reviews*, *19*, 321-346.
Kutzbach, J., & Street-Perrott, F. (1985). Milankovitch forcing of fluctuations in the level of tropical lakes from 18 to 0 kyr BP. *Nature*, *317*, 130–134.
Levis, S., Bonan, G., & Bonfils, C. (2004). Soil feedback drives the mid-Holocene North African monsoon northward in fully coupled CCSM2 simulations with a dynamic vegetation model. *Clim Dyn*, *23*, 791-802.
Loeb, N., Lyman, J., Johnson, G., Allan, R., Doeling, D., Wong, T., ... Stephenes, G. (2012). Observed changes in top-of-the-atmosphere radiation and upper-ocean heating consistent within uncertainty. *Nature Geosciences*, *5*, 110-113. doi: 10.1038/ngeo1375
Loveland, T., Reed, B., Brown, J., Ohlen, D., Zhu, Z., Yang, L., & Merchant, J. (2000). Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data. *Int. J. Remote Sens*, *21*(6-7), 1303-1330.
Lu, Z., Miller, P., Zhang, Q., et al. (2018). Dynamic Vegetation Simulations of the Mid-Holocene Green Sahara. *Geophys Res Lett*, *45*, 1-10. doi: 10.1029/2018GL079195
Manning, K., & Timpson, A. (2014). The demographic response to Holocene climate change in the Sahara. *Quat Sci Rev*, *101*, 28–35. doi: 10.1016/j.quascirev.2014.07.003
Mapes, B., & Neale, R. (2011). Parameterizing Convective Organization to Escape the Entrainment Dilemma. *J Adv Model Earth Syst*, *3*, M06004.
Maraun, D., & Widmann, M. (2018). *Statistical downscaling and bias correction for climate research* (1st ed.). Cambridge, UK: Cambridge University Press. doi: 10.1017/9781107588783
Marsham, J., Dixon, N., Garcia-Carreras, L., et al. (2013). The role of moist convection in the West African monsoon system: Insights from continental-scale convection-permitting simulations. *Geophys Res Lett*, *40*, 1843-1849. doi: 10.1002/grl.50347
Massey, N., Jones, R., Otto, F., et al. (2015). weather@home development and validation of a very large ensemble modelling system for probabilistic event attribution . *Q J R Meteorol Soc*, *14*, 1528-1545. doi: 10.1002/qj.2455
McKay, M., Beckman, R., & Conover, W. (1979). A Comparison of Three Methods

for Selecting Values of Input Variables in the Analysis of Output from a Computer Code. *Technometrics*, *32*(2), 239-245.

McNeall, D., Williams, J., Booth, B., et al. (2016). The impact of structural error on parameter constraint in a climate model. *Earth Syst Dynam*, *7*, 917-935.

Nam, C., Bony, S., Dufresne, J.-L., & Chepfer, H. (2012). The 'too few, too bright' tropical low-cloud problem in cmip5 models. *Geophys Res Lett*(L21801). doi: 10.1029/2012GL053421

Nicholson, S. (2009). A revised picture of the structure of the monsoon and land ITCZ over West Africa. *Clim Dyn*, *32*, 1155-1171.

Pachur, H.-J., & Holzmann, P. (1991). Paleoclimatic Implications of Late Quaternary Lacustrine Sediments in Western Nubia, Sudan. *Quat Res*, *36*, 257-276.

Pante, G., & Knippertz, P. (2019). Resolving Sahelian thunderstorms improves mid-latitude weather forecasts. *Nature Comm*, *10*, 3487. doi: 10.1038/s41467-019-11081-4

Pausata, F., Messori, G., & Zhang, Q. (2016). Impacts of dust reduction on the northward expansion of the African monsoon during the Green Sahara period. *Earth Planet Sci Lett*, *434*, 298-307.

Perez-Sanz, A., Li, G., Gonzalez-Samperiz, P., & Harrison, S. (2014). Evaluation of modern and mid-Holocene seasonal precipitation of the Mediterranean and northern Africa in the CMIP5 simulations. *Clim Past*, *10*, 551-568.

Peyron, O., Jolly, P., D nad Braconnot, Bonnefille, R., Guiot, J., Wirrmann, D., & Chalie, F. (2006). Quantitative reconstructions of annual rainfall in Africa 6000 years ago: Model-data comparison. *J Geophys Res*, *111*, D24110.

Pope, V. D., Gallani, M. L., Rowntree, P. R., & Stratton, R. A. (2000). The impact of new physical parametrizations in the Hadley Centre climate model: HadAM3. *Climate Dynamics*, *16*(2-3), 123-146.

Rayner, N., Parker, D., Horton, E., Folland, C., Alexander, L., Rowell, D., … 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*, D14,4407.

Ritchie, J., & Haynes, C. (1987). Holocene vegetation zonation in the eastern Sahara. *Nature*, *330*, 645-647.

Rougier, J. (2007). Probabilistic inference for future climate using an ensemble of climate model evaluations. *Climatic Change*, *81*, 247-264.

Rougier, J., Sexton, D., Murphy, J., & Stainforth, D. (2009). Analyzing the climate sensitivity of the hadsm3 climate model using ensembles from different but related experiments. *J Clim*, *22*, 1327-1353.

Roustant, O., Ginsbourger, D., & Deville, Y. (2012). DiceKriging, DiceOptim: Two R Packages for the Analysis of Computer Experiments by Kriging-Based Metamodeling and Optimization. *J Stat Software*, *51*(1), 1-55.

Rowell, D., Senior, C., Vellinga, M., & Graham, R. (2016). Can climate projection uncertainty be constrained over Africa using metrics of contemporary performance? *Clim Change*, *134*(621-633). doi: 10.1007/s10584-015-1554-4

Schmidt, G., Annan, J., Bartlein, P., Cook, B., Guilyardi, E., Hargreaves, J., … Yiou, P. (2014). Using palaeo-climate comparisons to constrain future projections in CMIP5. *Climate of the Past*, *10*, 221-250. doi: 10.5194/cp-10-221-2014

Sexton, D., Murphy, J., Collins, M., & Webb, M. (2012). Multivariate probabilistic projections using imperfect climate models part I: outline of methodology. *Clim Dyn*, *38*, 2513-2542.

Skinner, C., & Poulsen, C. (2016). The role of fall season tropical plumes in enhancing Saharan rainfall during the African Humid Period. *Geophys Res Lett*, *43*, 349-358. doi: 10.1002/2015GL066318

Skonieczny, C., Paillou, A., P nad Bory, Bayon, G., Biscara, L., Crosta, X., Eynaud, F., … Grousset, F. (2015). African humid periods triggered the reactivation of a large river system in Western Sahara. *Nature Communications*, *6*, 8751-8756.

Stensrud, D. (2007). *Parameterization schemes: Keys to understanding numerical weather prediction models* (1st ed.). Cambridge, UK: Cambridge University Press.

Street-Perrott, F., Mitchell, J., Marchand, D., & Brunner, J. (1990). Milankovitch and albedo forcing of the tropical monsoons: a comparison of geological evidence and numerical simulations for 9000 yBP. *Earth and Environmental Science Transactions of The Royal Society of Edinburgh*, *81*(4), 407-427.

Su, H., & Neelin, J. (2005). Dynamical mechanisms for African monsoon changes during the mid-Holocene. *J Geophys Res*, *110*, D19105. doi: 10.1029/2005JD005806

Taylor, C., Lambin, E., Stephenne, N., et al. (2002). The Influence of Land Use Change on Climate in the Sahel. *J Clim*, *15*, 3615-3629.

Texier, D., De Noblet, N., & Braconnot, P. (2000). Sensitivity of the African and Asian Monsoons to Mid-Holocene Insolation and Data-Inferred Surface Changes. *J Climate*, *13*, 164-181.

Thompson, A., Skinner, C., Poulsen, C., & Zhu, J. (2019). Modulation of Mid-Holocene African Rainfall by Dust-Aerosol Direct and Indirect Effects. *Geophys Res Lett*, *46*, 3917-3926. doi: 10.1029/2018GL081225

Tierney, J., Pausata, F., & deMenocal, P. (2017). Rainfall regimes of the Green Sahara. *Science Advances*, *3*, e1601503.

Valdes, P. (2011). Built for stability? *Nature Geosci*, *4*, 414-416.

Valdes, P., Armstrong, E., Badger, M., Bradshaw, C., Bragg, F., Crucifix, M., … Williams, J. (2017). The BRIDGE HadCM3 family of climate models: HadCM3@Bristol v1.0. *Geosci Model Dev*, *10*, 3715-3743. doi: 10.5194/gmd-10-3715-2017

Vamborg, F., Brovkin, V., & Claussen, M. (2011). The effect of a dynamic background albedo scheme on Sahel/Sahara precipitation during the mid-Holocene. *Clim Past*, *7*, 117-131.

Willett, M., & Whitall, M. (2017). A Simple Prognostic based Convective Entrainment Rate for the Unified Model: Description and Tests. *Met Office Forecasting Research Technical Report*, *617*, 1-54.

Williams, R., McGee, D., Kinsley, C., Ridley, D., Hu, S., Fedorov, A., … de Menocal, P. (2016). Glacial to Holocene changes in trans-Atlantic Saharan dust transport and dust-climate feedbacks. *Science Advances*, *2*, e1600445.

Wu, H., Guiot, J., Brewer, S., & Guo, Z. (2007). Climatic changes in Eurasia and Africa at the last glacial maximum and mid-Holocene: reconstruction from pollen data using inverse vegetation modelling. *Clim Dyn*, *29*, 211–229. doi: 10.1007/s00382-007-0231-3

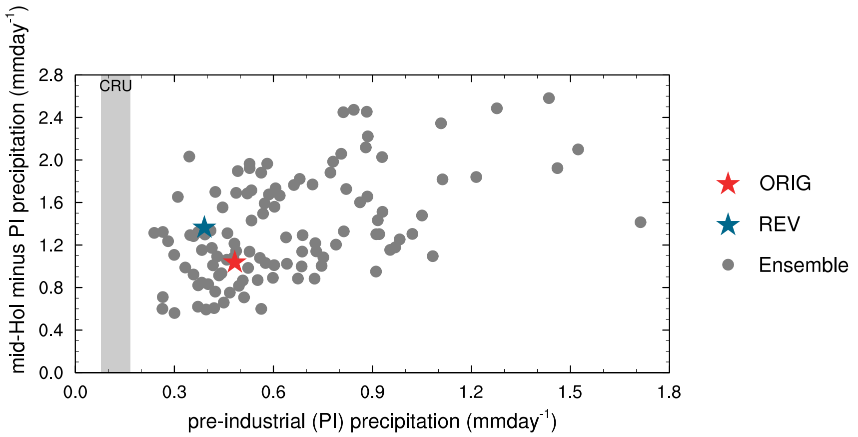

Figure 1. ORIG (red), REV (blue) and ensemble (grey) simulated JJAS precipitation in North Africa (20°W-30°E, 15-30°N) against the simulated mid-Holocene minus pre-industrial precipitation (mm day^{-1}) change. The observed precipitation in this region from CRU (Harris et al., 2014) based on years 1961-1990, is indicated by the shaded grey bar. The impact of using present-day versus pre-industrial precipitation observations is likely to be small and less important than differences due to model biases and due to significant spatial gaps in the early instrumental observations.

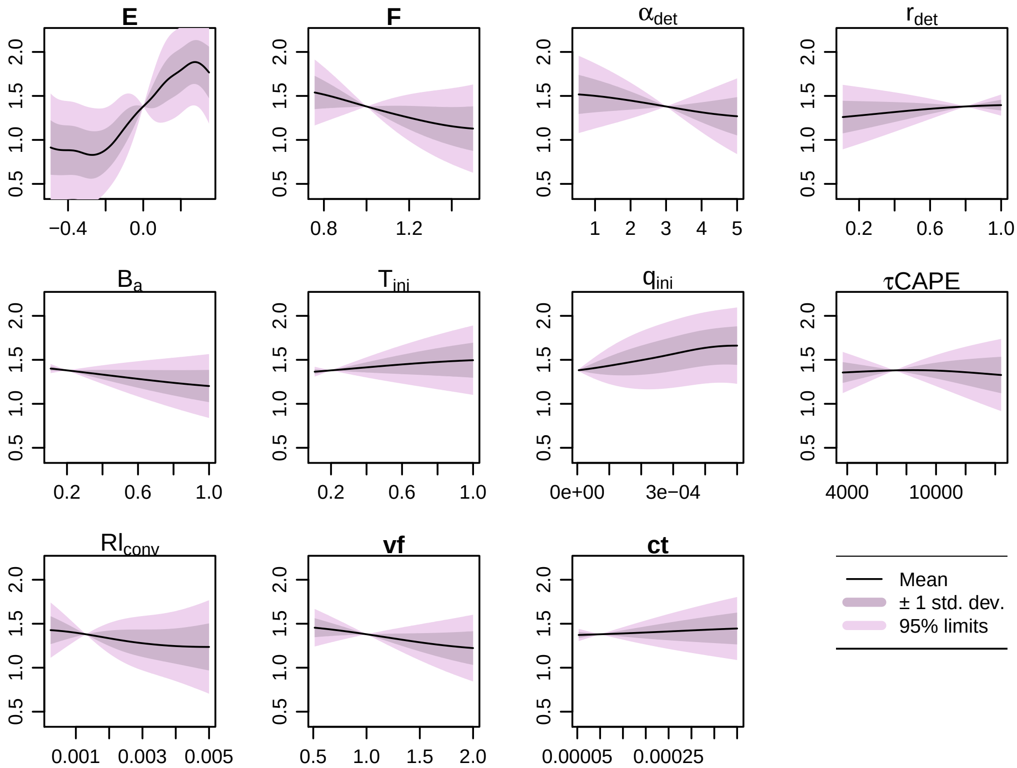

Figure 2. Emulator prediction of JJAS mean precipitation change (mid-Holocene minus pre-industrial) over 15-30°N by 20°W-30°E. Dependence of the JJAS precipitation difference on each individual model parameter. In each panel that parameter is varied across the range, whilst the remaining 10 parameters are held at their default values. The uncertainty ranges (± 1 standard deviation and the 95% intervals) are as reported by the emulator and collapse to zero at the point in parameter space at which the climate model has been run before. These single parameter sampling evaluations are from the emulator based on all members of the ensemble simulations performed with the GCM. The error bars are a function of distance from points in state-space that have already been evaluated with the GCM.

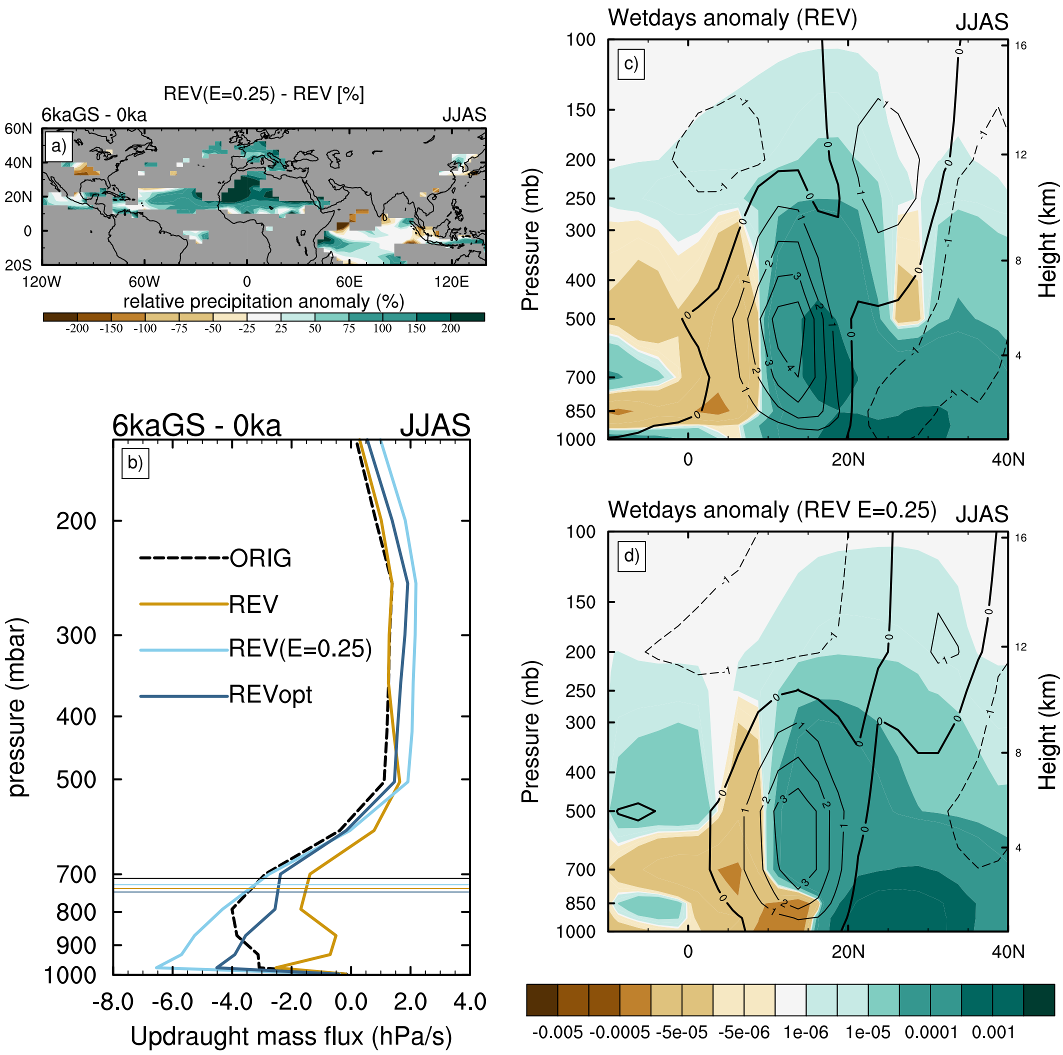

Figure 3. a) relative difference in the mid-Holocene JJAS precipitation anomaly (6kaGS-0ka) for REV model version E=0.25 minus E=0. (b) Mid-Holocene minus pre-industrial (6kaGS-0ka) JJAS mean vertical profiles of updraught mass flux (hPa s^{-1}). Horizontal lines show the lifting condensation level for an undilute parcel of surface air for the mid-Holocene averaged over the same domain. (c) and (d) The REV JJAS zonally averaged daily-mean anomalies of humidity (shading: $kgkg^{-1}$) and zonal wind (contour lines: ms^{-1}) for wetdays (>4.0 mm day^{-1}) for E=0 and E=0.25 respectively.

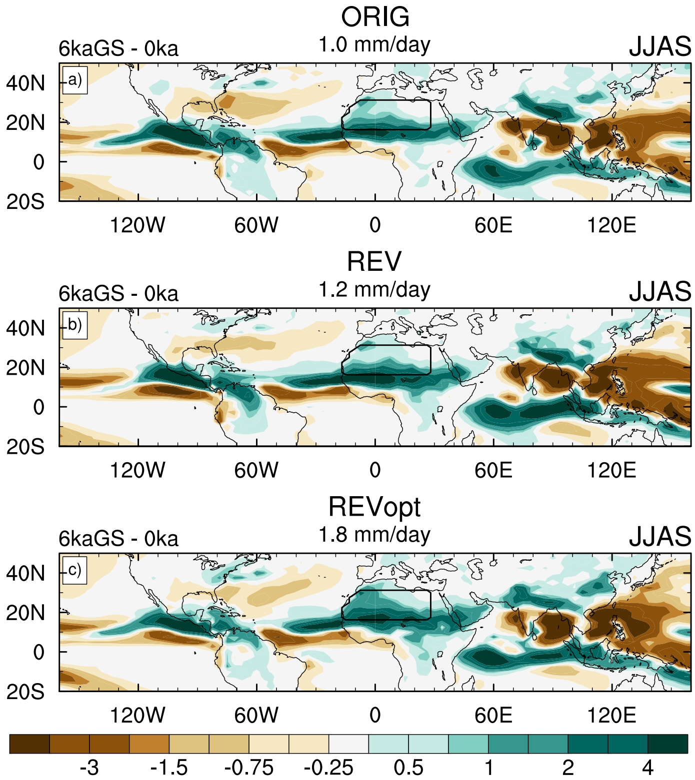

Figure 4. Simulated Northern Hemisphere summer (JJAS) precipitation anomalies (mm day^{-1}): (a) ORIG (Valdes et al., 2017); (b) REV: the modified version used as a starting configuration in this study; and (c) REVopt, the optimised version based on the probabilistic approach. The mean simulated difference in JJAS precipitation (mm day^{-1}) for North Africa in the area of the box is given above each panel.

References From the Supporting Information

References

Adler, R. et al. (2003). The Version-2 Global Precipitation Climatology Project (GPCP) monthly precipitation analysis (1979present). *J. Hydrometeorol. 4*(6), 1147–1167.

Bonan, G. (2015). *Ecological Climatology* (3rd ed.). Cambridge, UK: Cambridge University Press.

Fritsch, J. and C. Chappell (1980). Numerical prediction of convectively driven mesoscale pressure systems. Part I: convective parameterization. *J Atmos Sci 37*, 1722–1733.

Harris, I., P. Jones, T. Osborn, and D. Lister (2014). Updated high-resolution grids of monthly climatic observations - the CRU TS3.10 Dataset. *Int J Climtol 34*, 623–642.

Harrison, S., P. Bartlein, S. Brewer, I. Prentice, M. Boyd, I. Hessler, K. Holmgren, K. Izumi, and K. Willis (2014). Climate model benchmarking with glacial and mid-Holocene climates. *Clim Dyn 43*, 671–688.

Hegglin, M., D. Plummer, T. Shepherd, J. Scinocc, J. Anderson, L. Froidevaus, B. Funke, D. Hurst, A. Rozanov, J. Urban, T. von Clarmann, K. Walker, H. Wang, S. Tegtmeier, and K. Weigel (2014). Vertical structure of stratospheric water vapour trends derived from merged satellite data. *Nature Geoscience 7*, 768–766.

Jones, P. and C. Harpham (2013). Estimation of the absolute surface air temperature of the Earth. *J Geophys Res 118*, 3213–3217.

Joshi, M., M. Webb, A. Maycock, and M. Collins (2010). Stratospheric water vapour and high climate sensitivity in a version of the HadSM3 climate model. *Atmos Chem Phys 10*, 7161–7167.

Lowe, J., D. Bernie, P. Bett, et al. (2018). UK Climate Projections UKCP18 Science Overview Report, v2. *Met Office Hadley Centre*, 1–73.

Murphy, J. M., D. M. H. Sexton, D. N. Barnett, G. S. Jones, M. J. Webb, and M. Collins (2004). Quantification of modelling uncertainties in a large ensemble of climate change simulations. *Nature 430*(7001), 768–772.

New, M., D. Lister, M. Hulme, and I. Makin (2002). A high-resolution data set of surface climate over global land areas. *Climate Research 21*, 1–25.

Reichler, T. and J. Kim (2008). How well do coupled models simulated today's climate? *Bull Amer Meteor Soc 89*, 303–311.

Sanderson, B., K. Shell, and W. Ingram (2010). Climate feedbacks determined using radiative kernels in a multi-thousand member ensemble of AOGCMs. *Clim Dyn 35*, 1219–1236.

Schaller, N., K. Alison, R. Lamb, et al. (2016). Human influence on climate in the 2014 southern England winter floods and their impacts. *Nature Clim Change 6*, 627–634.

Sime, L., P. Hopcroft, and R. Rhodes (2019). Impact of abrupt sea ice loss on Greenland water isotopes during the last glacial period. *Proceedings of the National Academy of Sciences 116*(10), 4099–4104.

Sun, Q., C. Miao, Q. Duan, H. Ashouri, S. Sorooshian, and K.-L. Hsu (2018). A Review of Global Precipitation Data Sets: Data Sources, Estimation, and Intercomparisons. *Rev Geophys 56*, 79–107.

Taylor, K. (2001). Summarizing multiple aspects of model performance in a single diagram. *J Geophys Res 106*(D7), 7183–7192.