



1 **The UK contribution to CMIP6/PMIP4: mid-Holocene and Last**
2 **Interglacial experiments with HadGEM3, and comparison to the pre-**
3 **industrial era and proxy data**

4
5 **Charles J. R. Williams^{1,5}, Maria-Vittoria Guarino², Emilie Capron³, Irene Malmierca-**
6 **Vallet^{1,2}, Joy S. Singarayer^{4,1}, Louise C. Sime², Daniel J. Lunt¹ Paul J. Valdes¹**

7
8 ¹School of Geographical Sciences, University of Bristol, UK (c.j.r.williams@bristol.ac.uk)

9 ²British Antarctic Survey, Cambridge, UK

10 ³Physics of Ice, Climate and Earth, Niels Bohr Institute, University of Copenhagen, Denmark

11 ⁴Department of Meteorology & School of Archaeology, Geography and Environmental
12 Science, University of Reading, UK

13 ⁵NCAS-Climate / Department of Meteorology, University of Reading, UK

14

15

16

17

18

19

20

21

22

23 **Corresponding author address:**

24 Room 1.2n, School of Geographical Sciences,

25 University Road, Bristol, BS8 1SS

26 United Kingdom

27

28 Email: c.j.r.williams@bristol.ac.uk

29

30 Short title: mid-Holocene and Last Interglacial experiments with HadGEM3

31 Keywords: Palaeoclimate, Quaternary change, mid-Holocene, Last Interglacial

32



33 **ABSTRACT**

34 Palaeoclimate model simulations are an important tool to improve our understanding of the
35 mechanisms of climate change. These simulations also provide tests of the ability of models to
36 simulate climates very different to today. Here we present the results from two simulations using the
37 latest version of the UK's physical climate model, HadGEM3-GC3.1; the mid-Holocene (~6 ka) and
38 Last Interglacial (~127 ka) simulations, both conducted under the auspices of CMIP6/PMIP4. These
39 periods are of particular interest to PMIP4 because they represent the two most recent warm periods
40 in Earth history, where atmospheric concentration of greenhouse gases and continental configuration
41 is similar to the pre-industrial period but where there were significant changes to the Earth's orbital
42 configuration, resulting in a very different seasonal cycle of radiative forcing.

43
44 Results for these simulations are assessed against proxy data, previous versions of the UK model, and
45 models from the previous CMIP5 exercise. When the current version is compared to the previous
46 generation of the UK model, the most recent version suggests limited improvement. In common with
47 these previous model versions, the simulations reproduce global land and ocean temperatures (both
48 surface and at 1.5 m) and a West African monsoon that is consistent with the latitudinal and seasonal
49 distribution of insolation. The Last Interglacial simulation appears to accurately capture Northern
50 Hemisphere temperature changes, but without the addition of Last Interglacial meltwater forcing
51 cannot capture the magnitude of Southern Hemisphere changes. Model-data comparisons indicate
52 that some geographical regions, and some seasons, produce better matches to the palaeodata (relative
53 to pre-industrial) than others. Model-model comparisons, relative to previous generations same
54 model and other models, indicate similarity between generations in terms of both the intensity and
55 northward enhancement of the mid-Holocene West African monsoon, both of which are
56 underestimated. On the 'Saharan greening' which occurred the mid-Holocene African Humid Period,
57 simulation results are likewise consistent with other models. The most recent version of the UK
58 model appears to still be unable to reproduce the amount of rainfall necessary to support grassland
59 across the Sahara.

60



61 1. INTRODUCTION

62 Simulating past climates has been instrumental in improving our understanding of the mechanisms of
 63 climate change (e.g. Gates 1976, Haywood *et al.* 2016, Jungclaus *et al.* 2017, Kageyama *et al.* 2017,
 64 Kageyama *et al.* 2018, Kohfeld *et al.* 2013, Lunt *et al.* 2008, Otto-Bliesner *et al.* 2017, Ramstein *et al.*
 65 1997), as well as in identifying and assessing discrepancies in palaeoclimate reconstructions (e.g.
 66 Rind & Peteet 1985). Palaeoclimate scenarios can also provide tests of the ability of models to
 67 simulate climates that are very different to today, often termed ‘out-of-sample’ tests. This notion
 68 underpins the idea that robust simulations of past climates improve our confidence in future climate
 69 change projections (Braconnot *et al.* 2011, Harrison *et al.* 2014, Taylor *et al.* 2011). Palaeoclimate
 70 scenarios have also been used to provide additional tuning targets for models (e.g. Gregoire *et al.*
 71 2011), in combination with historical or pre-industrial conditions.

72

73 The international Climate Model Intercomparison Project (CMIP) and the Palaeoclimate Model
 74 Intercomparison Project (PMIP) have spearheaded the coordination of the international palaeoclimate
 75 modelling community to run key scenarios with multiple models, perform data syntheses, and
 76 undertake model-data comparisons since their initiation twenty-five years ago (Joussaume & Taylor
 77 1995). Now in its fourth incarnation, PMIP4 (part of the sixth phase of CMIP, CMIP6), it includes a
 78 larger set of models than previously, and more palaeoclimate scenarios and experiments covering the
 79 Quaternary (documented in Jungclaus *et al.* 2017, Kageyama *et al.* 2017, Kageyama *et al.* 2018 and
 80 Otto-Bliesner *et al.* 2017) and Pliocene (documented in Haywood *et al.* 2016).

81

82 PMIP4 specifies experiment set-ups for two warm interglacial simulations: the mid-Holocene (MH) at
 83 ~6 ka and the Last Interglacial (LIG) covering ~129-116 ka. These are the two most recent warm
 84 periods in Earth history, and are of particular interest to PMIP4; indeed, the MH experiment is one of
 85 the two entry cards into PMIP (Otto-Bliesner *et al.* 2017). This is because whilst the atmospheric
 86 concentration of greenhouse gases, the extent of land ice, and the continental configuration is similar
 87 in these PMIP4 set-ups compared to the pre-industrial (PI) period, significant changes to the seasonal
 88 cycle of radiative forcing, relative to today, do occur during these periods due to long-term variations
 89 in the Earth’s orbital configuration. The MH and LIG both have higher boreal summer insolation and
 90 lower boreal winter insolation compared to the PI, as shown by Figure 1, leading to an enhanced
 91 seasonal cycle in insolation as well as a change in its latitudinal distribution. The change is more
 92 significant in the LIG than the MH, due to the larger eccentricity of the Earth’s orbit at that time.

93

94 Palaeodata syntheses indicate globally warmer surface conditions of potentially ~0.7°C than PI in the
 95 MH (Marcott *et al.* 2013) and up to ~1.3°C in the LIG (Fischer *et al.* 2018). Recent palaeodata
 96 compilations (Capron *et al.* 2014, Hoffman *et al.* 2017) reveal that the maximum temperatures were



reached asynchronously in the LIG between the Northern and Southern Hemispheres. Furthermore, model simulations suggest that this may have been caused by meltwater induced shutdown of the Atlantic Meridional Overturning Circulation (AMOC) in the early part of the LIG, due to the melting of the Northern Hemisphere ice sheets during the preceding deglaciation (e.g. Stone *et al.* 2016). During both warm periods there is abundant palaeodata evidence indicating enhancement of Northern Hemisphere summer monsoons (e.g. Wang *et al.* 2008) and in the case of the Sahara, replacement of desert by shrubs and steppe vegetation (e.g. Drake *et al.* 2011, Hoelzmann *et al.* 1998) and inland water bodies (e.g. Drake *et al.* 2011, Lezine *et al.* 2011).

The driving mechanism producing the climate and environmental changes indicated by the palaeodata for the LIG and MH is different to current and future anthropogenic warming, as the former results from orbital forcing changes whilst the latter results from increases in greenhouse gases. However, these past warm intervals are a unique opportunity to understand the magnitudes of forcings and feedbacks in the climate system that produce warm interglacial conditions, which can help us understand and constrain future climate projections (e.g. Holloway *et al.* 2016, Rachmayani *et al.* 2017, Schmidt *et al.* 2014). Running the same model scenarios with ever newer models enables the testing of whether model developments are producing improvements in palaeo model-data comparisons, assuming appropriate boundary conditions are used. Previous iterations of PMIP, with older versions of the PMIP4 models, have uncovered persistent shortcomings (Harrison *et al.* 2015) that have not been eliminated despite developments in resolution, model physics, and addition of further Earth system components. One key example of this is the continued underestimation of the increase in rainfall over the Sahara in the MH PMIP simulations (e.g. Braconnot *et al.* 2012).

In this study we run and assess the latest version of the UK's physical climate model, HadGEM3-GC3.1. In Global Coupled (GC) version 3 (and therefore the following GC3.1), there have been many updates and improvements, relative to its predecessors, which are discussed extensively in Williams *et al.* (2017) and a number of companion scientific model development papers (see Section 2.1). As a brief introduction, however, GC3 includes a new aerosol scheme, multilayer snow scheme, multilayer sea ice and several other parametrization changes, including a set relating to cloud and radiation, as well as a revision to the numerics of convection (Williams *et al.* 2017). In addition, the ocean component of GC3 has other changes including a new ocean and sea ice model, a new cloud scheme, and further revisions to all parametrization schemes (Williams *et al.* 2017). See Section 2.1 for further details.

Following the CMIP6/PMIP4 protocol, here the PMIP4 MH and LIG simulations have been conducted and assessed, comparing the results with available proxy data, previous versions of the UK's same physical climate model, and other models from CMIP5. The focus of this paper is on the



134 fidelity of the temperature anomalies globally and the degree of precipitation enhancement in the
 135 Sahara, the latter of which has proved problematic for several generations of models. The results
 136 discussed here are split into two sections: after an assessment of the level of equilibrium gained
 137 during the spin-up phase, the main focus is on the model-data and model-model comparisons using
 138 the production runs. Following this introduction, Section 2 describes the model, the experimental
 139 design and the proxy data used for the model-data comparisons. Section 3 then presents the results,
 140 divided into two subsections: i) equilibrium during the spin-up phase; and ii) model-data and model-
 141 model comparisons from the production runs. Finally, section 4 summarises and concludes.

142

143 2. MODEL, EXPERIMENT DESIGN AND DATA

144 2.1. Model

145 The MH and LIG simulations conducted here (referred to as *midHolocene* and *lig127k*, respectively,
 146 and collectively as the ‘warm climate’ simulations), and indeed the PI simulation (*piControl*,
 147 conducted elsewhere as part of the UK’s CMIP6 runs and used here for comparative purposes) were
 148 all run using the same fully-coupled GCM: the Global Coupled 3 configuration of the UK’s physical
 149 climate model, HadGEM3-GC3.1. Full details on HadGEM3-GC3.1, and a comparison to previous
 150 configurations, are given in Williams *et al.* (2017) and Kuhlbrodt *et al.* (2018). Here, the model was
 151 run using the Unified Model (UM), version 10.7, and including the following components: i) Global
 152 Atmosphere (GA) version 7.1, with an N96 atmospheric spatial resolution (approximately 1.875°
 153 longitude by 1.25° latitude) and 85 vertical levels; ii) the NEMO ocean component, version 3.6,
 154 including Global Ocean (GO) version 6.0 (ORCA1), with an isotropic Mercator grid which, despite
 155 varying in both meridional and zonal directions, has an approximate spatial resolution of 1° by 1° and
 156 75 vertical levels; iii) the Global Sea Ice (GIS) component, version 8.0 (GSI8.0); iv) the Global Land
 157 (GL) configuration, version 7.0, of the Joint UK Land Environment Simulator (JULES); and v) the
 158 OASIS3 MCT coupler. The official title for this configuration of HadGEM3-GC3.1 is HadGEM3-
 159 GC31-LL N96ORCA1 UM10.7 NEMO3.6 (for brevity, hereafter HadGEM3).

160

161 All of the above individual components are summarised by Williams *et al.* (2017) and detailed
 162 individually by a suite of companion papers (see Walters *et al.* 2017 for GA7 and GL7, Storkey *et al.*
 163 2017 for GO6 and Ridley *et al.* 2017 for GIS8). However, a brief description of the major changes
 164 relative to its predecessor are given here. Beginning with GA7 and GL7, a once-in-a-decade
 165 replacement of the model’s dynamical core, implementing ENDGame, was undertaken for the
 166 previous version (GA6) and therefore remains the same in GA7 (Walters *et al.* 2017). In addition, a
 167 number of bottom-up and top-down developments were included in GA7. For the former, these
 168 include improvements to the radiation scheme to allow better treatment of gases absorption,
 169 improvements to how warm rain and ice clouds are treated, and an improvement to the numerics of
 170 the convection scheme (Walters *et al.* 2017). For the latter, these include further improvements to the



171 microphysics as well as an incremental development of ENDGame (Walters *et al.* 2017). Together
 172 these led to reductions in four model errors that were deemed critical in the previous configuration: i)
 173 South Asian monsoon rainfall biases over India; ii) biases in both temperature and humidity in the
 174 tropical tropopause; iii) shortcomings in the numerical conservation; and iv) biases in surface
 175 radiation fluxes over the Southern Ocean (Walters *et al.* 2017). In addition to these developments,
 176 two new parameterisation schemes were introduced in GA7: firstly the UK Chemistry and Aerosol
 177 (UKCA) GLOMAP-mode aerosol scheme, to improve the representation of tropospheric aerosols, and
 178 secondly a multi-layer snow scheme in JULES, to allow the first time inclusion of stochastic physics
 179 in UM climate simulations (Walters *et al.* 2017).

180
 181 For the GO and GIS components, a number of improvements to GO6 have been made since the
 182 previous version, the first of which was an upgrade of the NEMO base code (to version 3.6) which
 183 allowed a formulation for momentum advection (from Hollingsworth *et al.* 1983), a Lagrangian
 184 icebergs scheme, and a simulation of circulation below ice shelves (Storkey *et al.* 2018). Other
 185 developments included an improvement to the warm SST bias in the Southern Ocean (as detailed by
 186 Williams *et al.* 2017), as well as tuning to various parameters e.g. the isopycnal diffusion (Storkey *et al.*
 187 2018). For GIS8, along with improvements to the albedo scheme and more realistic semi-implicit
 188 coupling, the biggest development since its predecessor is the inclusion of multilayer
 189 thermodynamics, giving a heat capacity to the sea ice and allowing vertical variation of conduction
 190 (Ridley *et al.* 2018). Testing of these two components produced a better simulation compared to its
 191 predecessor, with more realistic mixed layer depths in the Southern Ocean and the aforementioned
 192 reduced warm bias, the latter of which was deemed primarily due to the tuning of the different mixing
 193 (e.g. vertical and isopycnal) parameters (Storkey *et al.* 2018).

194
 195 When all of these components are coupled together to give GC3, there have been several
 196 improvements relative to its predecessor (GC2), most noticeably to the large warm bias in the
 197 Southern Ocean (which was reduced by 75%), as well as an improved simulation of clouds, sea ice,
 198 the frequency of tropical cyclones in the Northern Hemisphere as well as the AMOC, and the Madden
 199 Julian Oscillation (MJO) (Williams *et al.* 2017). Relative to the previous fully-coupled version of the
 200 model (HadGEM2), which was submitted to the last CMIP5/PMIP3 exercise, many systematic errors
 201 have been improved including a reduction in many regions to the temperature bias, a better simulation
 202 of mid-latitude synoptic variability, and an improved simulation of tropical cyclones and the El Niño
 203 Southern Oscillation (ENSO) (Williams *et al.* 2017).

204
 205 Here, the *midHolocene* and *lig127k* simulations were both run on the UK National Supercomputing
 206 Service, ARCHER, whereas the *piControl* was run on a different platform based within the UK Met
 207 Office's Hadley Centre. While this may mean that anomalies computed against the *piControl* are



potentially influenced by different computing environments, and not purely the result of different climate forcings, the reproducibility of GC3.1 simulations across different platforms has been tested (Guarino *et al.* 2019). It was found that, although a simulation length of 200 years is recommended whenever possible to adequately capture climate variability across different platforms, the main climate variables considered here (e.g. surface temperature) are not expected to be significantly different on a 100- or 50-year timescale (see, for example, Fig. 6 in Guarino *et al.* [2019]) as they are not directly affected by medium-frequency climate processes such as ENSO.

Not including queueing time, both simulations were achieving 3-4 model years per day during the spin-up phase, and 1-2 model years per day during the production run; see below for the differences in output, and therefore speed, between the two phases.

2.2. Experiment design

Full details of the experimental design, and results from the CMIP6 *piControl* simulation, are documented in Menary *et al.* (2018). Both the warm climate simulations followed the experimental design given by Otto-Bliesner *et al.* (2017), and specified at https://pmip4.lsce.ipsl.fr/doku.php/exp_design:index. The primary differences from the *piControl* were to the astronomical parameters and the atmospheric trace greenhouse gas concentrations, summarised in Table 1. For the astronomical parameters, these were prescribed in Otto-Bliesner *et al.* (2017) according to orbital constants from Berger & Loutre (1991). However, in HadGEM3, the individual parameters (e.g. eccentricity, obliquity, etc) use orbital constants based on Berger (1978), according to the specified start date of the simulation. For the atmospheric trace greenhouse gas concentrations, these were based on recent reconstructions from a number of sources (see Table 1 for values, and section 2.2 in Otto-Bliesner *et al.* [2017] for a full list of references/sources).

All other boundary conditions, including solar activity, ice sheets, topography and coastlines, volcanic activity and aerosol emissions, are identical to the CMIP6 *piControl* simulation. Likewise, vegetation was prescribed to present-day values, to again match the CMIP6 *piControl* simulation. As such, the *piControl* and both the warm climate simulations actually include a prescribed fraction of urban land surface. As a result of this, our orbitally- and greenhouse gas-forced simulations should be considered as anomalies to the *piControl*, rather than absolute representations of the MH or LIG climate.

Both the warm climate simulations were started from the end of the *piControl* spin-up phase (which ran for approximately 600 years), after which time the *piControl* was considered to be in atmospheric and oceanic equilibrium (Menary *et al.* 2018). To assess this, four metrics were used, namely net radiative balance at the top of the atmosphere (TOA), surface air temperature (SAT), and full-depth ocean temperature (OceTemp) and salinity (OceSal) Menary *et al.* (2018). See Section 3.1 (and in



particular Table 2) for an analysis of the equilibrium state of both the *piControl* and the warm climate simulations. Starting at the end of the *piControl*, these were then run for their own spin-up phases, 400 and 350 years for the *midHolocene* and *lig127k* respectively. During this phase, ~700 diagnostics were output, containing mostly low temporal frequency (e.g. monthly, seasonal and annual) fields. Once the simulations were considered in an acceptable level of equilibrium (see Section 3.1), a production phase was run for 100 and 200 years for the *midHolocene* and *lig127k* respectively, during which the full CMIP6/PMIP4 diagnostic profile (totalling ~1700 fields) was implemented to output both high and low temporal frequency variables.

253

2.3. Data

Recent data syntheses compiling quantitative surface temperature and rainfall reconstructions were used in order to evaluate the warm climate simulations.

257

For the MH, the global-scale continental surface mean annual temperature (MAT) and rainfall (or mean annual precipitation, MAP) reconstructions from Bartlein *et al.* (2011), with quantitative uncertainties accounting for climate parameter reconstruction methods, were used (see Data Availability for access details). They rely on a combination of existing quantitative reconstructions based on pollen and plant macrofossils and are inferred using a variety of methods (see Bartlein *et al.* 2011 for further details). At each site, the 6 ka anomaly (corresponding to the 5.5-6.5 ka average value), is given relative to the present day, and in the case where modern values could not be directly inferred from the record, modern climatology values (1961-1990) were extracted from the Climate Research Unit historical climatology data set (New *et al.* 2002).

267

For the LIG, two different sets of surface temperature data are available. Firstly, the Capron *et al.* (2017) 127 ka timeslice of SAT and sea surface temperature (SST) anomalies (relative to pre-industrial, 1870-1899), is based on polar ice cores and marine sediment data that are (i) located poleward of 40° latitude and (ii) have been placed on a common temporal framework (see Data Availability for access details). Polar ice core water isotope data are interpreted as annual surface air temperatures, while most marine sediment-based reconstructions are interpreted as summer SST signals. For each site, the 127 ka value was calculated as the average value between 126 and 128 ka using the surface temperature curve resampled every 0.1 ka. Secondly, a global-scale time slice of SST anomalies, relative to pre-industrial (1870-1889), at 127 ka was built, based on the recent compilation from Hoffman *et al.* (2017), which includes both annual and summer SST reconstructions (see Data Availability for access details). The 127 ka values at each site were extracted, following the methodology they proposed for inferring their 129, 125 and 120 ka time slices i.e. the SST value at 127 ka was taken on the provided mean 0.1 ka interpolated SST curve for each core location. Data syntheses from both Capron *et al.* (2014, 2017) and Hoffman *et al.* (2017) are associated with



quantitative uncertainties accounting for relative dating and surface temperature reconstruction methods. Here, the two datasets are treated as independent data benchmarks, as they use different reference chronologies and methodologies to infer temporal surface temperature changes, and therefore they should not be combined. See Capron *et al.* (2017) for a detailed comparison of the two syntheses. A model-data comparison exercise using existing LIG data compilations focusing on continental surface temperature (e.g. Turney and Jones 2010) was not attempted, as they do no benefit yet from a coherent chronological framework, preventing the definition of a robust time slice representing the 127 ka terrestrial climate conditions (Capron *et al.* 2017).

290

291 3. RESULTS

As briefly mentioned above, both the warm climate simulations had a spin-up phase before the main production run was started. The results discussed here are therefore split into two sections: firstly, assessing the level of atmospheric and oceanic equilibrium during (and, in particular, at the end of) the spin-up phase, and secondly assessing the 100-year climatology from the production run.

296

297 3.1. Spin-up

Annual global mean 1.5 m air temperature and TOA radiation from both warm climate simulations, compared to the *piControl*, are shown in Figure 2 and summarised in Table 2. Note that the *piControl* spin-up phase was run in three separate parts, to accommodate for minor changes/updates in the model as the simulation progressed. There is a clear increase in temperature during the beginning of this period, as the *piControl* slowly spins up from its original starting point; this levels off towards the end of the period, however, with a final temperature trend of $0.03^{\circ}\text{C century}^{-1}$ (Table 2 and Fig. 2a). For the warm climate simulations, despite considerable interannual variability (particularly halfway through the *lig127k* simulation) both are showing small long-term trends of $-0.06^{\circ}\text{C century}^{-1}$ and $-0.16^{\circ}\text{C century}^{-1}$ for the last 100 years of the *midHolocene* and *lig127k*, respectively (Table 2 and Fig. 2a). The same is true for TOA, where the *piControl* has a slow downward trend towards zero until equilibrium was reached, whereas the *midHolocene* and *lig127k* are relatively stable (Fig. 2b).

309

For the ocean, annual global mean OcéTemp and OcéSal are shown in Table 2 and Figure 3. There is again a clear increase in OcéTemp during the *piControl* spin-up phase, which again stabilises at $0.035^{\circ}\text{C century}^{-1}$ by the end of the period (Table 2). Whilst OcéTemp stabilises in the *midHolocene* and indeed has a smaller trend than the *piControl* (Table 2), it continues to increase in the *lig127k* until it stabilises within the last ~50 years (Fig. 3a). A similar pattern is shown in OcéSal, with a steady decrease in the *piControl* spin-up phase which continues during the *midHolocene* and, conversely, starts to increase before stabilising during the *lig127k* (Fig. 3b). Concerning the long-term trends, Menary *et al.* (2018) considered values acceptable for equilibrium to be $< \pm 0.035^{\circ}\text{C century}^{-1}$ and $< \pm 0.0001 \text{ psu century}^{-1}$ (for OcéTemp and OcéSal, respectively); as shown in Table 2,



although both warm climate simulations meet the temperature criterion, neither meet the salinity criterion (-0.007 psu and 0.006 psu for the *midHolocene* and *lig127k*, respectively, compared to a criterion of 0.0001 psu). However, running for several thousands of years (and > 5 years of computer time), which would be needed to reach true oceanic equilibrium, was simply unfeasible here given time and resource constraints.

3.2. Production runs results

The warm climate production runs were undertaken following the spin-up phase, with a 100-year climatology of each simulation being compared to that from the *piControl*, as well as available proxy data, using either annual means or summer/winter seasonal means. For the latter, depending on the availability of the proxy data, Northern Hemisphere summer is defined as either June-August (JJA) or July-September (JAS), and Northern Hemisphere winter is defined as either December-February (DJF) or January-March (JFM); and vice versa for Southern Hemisphere summer/winter. Using atmospheric diagnostics, the focus is on three separate measures: i) to describe and understand the differences between the current two warm climate simulations and the *piControl* in terms of temperature, rainfall and atmospheric circulation changes; ii) to compare both current simulations, with existing and newly-available proxy data, and iii) to compare both current simulations with those from previous versions of the UK model (where available), such as HadGEM2-ES or HadCM3, in order to assess any improvements due to model advances. In this aim, previous CMIP3 and 5 versions of the UK model, alongside other CMIP5 models, will be assessed to address the question of whether simulations produce enough rainfall to allow vegetation growth across the Sahara: the mid-Holocene ‘Saharan greening’ problem.

3.2.1. Do the CMIP6 HadGEM3 simulations show temperature, rainfall and circulation differences when compared to the pre-industrial era?

Here we focus on mean differences between the HadGEM3 warm climate simulations and the corresponding *piControl*. Seasonal mean summer and winter 1.5 m air temperature anomalies (relative to the *piControl*) from both warm climate simulations are shown in Figure 4. During JJA, the *midHolocene* is showing a widespread increase in temperatures of up to 2°C across the entire Northern Hemisphere north of 30°N, more in some places e.g. Greenland (Fig. 4a), consistent with the increased latitudinal and seasonal distribution of insolation caused by known differences in the Earth’s axial tilt (Berger & Loutre 1991, Otto-Bliesner *et al.* 2017). The only places showing a reduction in temperature are West and central Africa (around 10°N) and northern India; this, as discussed below, is likely related to increased rainfall in response to a stronger summer monsoon, but could also be due to the resulting increase in cloud cover (reflecting more insolation) or a combination of the two. During DJF, only the Northern Hemisphere high latitudes (north of 60°N) continue this



355 warming trend, with the rest of continental Africa and Asia showing a reduction in temperature (Fig.
 356 4b). These patterns are virtually the same during the *lig127k* (Fig. 4c and d), just much more
 357 pronounced (with temperature increases during JJA of 5°C or more); again, this is consistent with the
 358 differences in the Earth's axial tilt, which were more extreme (and therefore Northern Hemisphere
 359 summer experienced larger insolation changes) in the LIG relative to the MH (Berger & Loutre 1991,
 360 Otto-Bliesner *et al.* 2017).

361
 362 Mean JJA rainfall and 850mb wind anomalies (relative to the *piControl*) from both warm climate
 363 simulations are shown in Figure 5, which zooms into Africa. In response to the increased Northern
 364 Hemisphere summer insolation, the West African monsoon is enhanced in both simulations, with
 365 positive (negative) rainfall anomalies across sub-Saharan Africa (eastern equatorial Atlantic)
 366 suggesting a northward displacement of the ITCZ. This is consistent with previous work, with a
 367 northward movement of the rainbelt being associated with increased advection of moisture into the
 368 continent (Huang *et al.* 2001, Singarayer *et al.* 2017, Wang *et al.* 2014). This increased advection of
 369 moisture is shown by the low-level westerlies in Figure 5, drawing in more moisture from the tropical
 370 Atlantic, which are consistent with previous work documenting the intensified monsoon circulation
 371 associated with a greater land-sea temperature contrast (Huang *et al.* 2001, Singarayer *et al.* 2017,
 372 Wang *et al.* 2006). This pattern is enhanced in the *lig127k* relative to the *midHolocene*, again due to
 373 the stronger insolation forcing in the LIG relative to the MH, and the northward displacement of the
 374 ITCZ is more pronounced in the *lig127k* simulation (Fig. 5c). Interestingly, however, regarding very
 375 small anomalies (i.e. < 1 mm day⁻¹), the *midHolocene* is showing wetter conditions further north,
 376 throughout the Sahara and up to the Mediterranean, whereas the *lig127k* simulation has small dry
 377 anomalies in this region (Fig. 5a and b for the *midHolocene* and *lig127k*, respectively).

378
 379 The change to the intensity and the spatial pattern (e.g. latitudinal positioning and extent) of the West
 380 African monsoon is further shown in Figure 6, which shows JJA rainfall anomalies by latitude over
 381 West Africa from both warm climate simulations. Apart from the clear drying relative to the
 382 *piControl* between the Equator and 5°N (which comes almost entirely from the equatorial Atlantic
 383 region), both warm climate simulations are showing a large increase in rainfall (of around 2 and 6 mm
 384 day⁻¹ for the *midHolocene* and *lig127k*, respectively) during the core monsoon region i.e. between
 385 approximately 10-15°N. In terms of the latitudinal extent, an examination of the mean rainfall by
 386 latitude suggests that both warm climate simulations are producing a wider monsoon region (i.e. both
 387 North and South of the Equator), with rainfall only reducing to near zero at 20°N in these simulations
 388 compared to approximately 16°N in the *piControl* (not shown). This is again consistent with previous
 389 work, where various theories are compared as to the reasons behind the latitudinal changes in the
 390 rainbelt's position, one which is a symmetric expansion during boreal summer (Singarayer &
 391 Burrough 2015, Singarayer *et al.* 2017).



392

393 **3.2.2. Model-Data comparison: Do the CMIP6 HadGEM3 simulations reproduce the**
 394 **‘reconstructed’ climate based on available proxy data?**

395 Here we focus on comparison with recent proxy data, focusing on surface temperature and rainfall
 396 (drawing direct comparisons, as well as using the root mean square error (RMSE), between proxy and
 397 simulated data, summarised in Table 4a), to see how well the current warm climate simulations are
 398 reproducing the ‘observed’ approximate magnitudes and patterns of change. It is worth noting that
 399 both simulated and proxy anomalies contain a high level of uncertainty, and in many locations the
 400 uncertainty is often larger than the anomalies themselves (not shown). The following results should
 401 therefore be considered with this caveat in mind.

402

403 Before the spatial patterns are compared, it is useful to assess global means (focusing on 1.5 m air
 404 temperature, calculated both annually and during Northern and Southern Hemisphere summer, JJA
 405 and DJF respectively) for model-model comparisons. Table 3 shows these global means, where it is
 406 clear that when annual means are considered, the *midHolocene* simulation is actually cooler than the
 407 *piControl*; this discrepancy with the palaeodata, which in general suggests a warmer MH relative to
 408 PI, also exists in previous models, and is termed the ‘Holocene temperature conundrum’ by Lui *et al.*
 409 (2014). The *lig127k* simulation is, however, warmer than the *piControl* simulation. Given the
 410 seasonal distribution of insolation in these two simulations, it is expected that the largest difference to
 411 the PI occurs during boreal summer, and indeed it does; during JJA, there is a warmer LIG and a
 412 slightly warmer MH (1.69°C and 0.07°C, respectively). Conversely, the opposite is true during DJF.

413

414 Concerning the spatial patterns during the MH, Figure 7 shows simulated surface MAT and MAP
 415 anomalies from the *midHolocene* simulation versus MH proxy anomalies from Bartlein *et al.* (2011),
 416 both of which have over 600 proxy locations in total (Table 4), although mostly confined to the
 417 Northern Hemisphere. For MAT, globally the simulation looks reasonable (RMSE = 2.45°C), and
 418 appears to be able to reproduce the sign of temperature change for many locations, with both
 419 simulated and proxy anomalies suggesting increases in temperature North of 30°N (Fig. 7a and b).
 420 This is not true everywhere, such as across the Mediterranean where the simulation suggests a small
 421 warming but the proxy data indicates cooling (Fig. 7a and b). However, regarding the magnitude of
 422 change, the *midHolocene* simulation is underestimating the temperature increase across most of the
 423 Northern Hemisphere, with for example increases of up to 1°C across Europe from the simulation
 424 compared to 3–4°C increases from the proxy data (Fig. 7a and b). In the simulation, temperature
 425 anomalies only reach these magnitudes in the Northern Hemisphere polar region (i.e. north of 70°N),
 426 not elsewhere. A similar conclusion can be drawn from MAP (RMSE = 280 mm yr⁻¹), where again
 427 the *midHolocene* simulation is correctly reproducing the sign of change across most of the Northern
 428 Hemisphere, but in some places not the magnitude. Over the eastern US, for example, rainfall



429 decreases of up to 200 mm yr⁻¹ are being shown by the simulation whereas the proxy data suggests a
 430 much stronger drying of up to 400 mm yr⁻¹ (Fig. 7c and d). Elsewhere, such as over Europe and
 431 Northern Hemisphere Africa, the simulation more accurately reproduces the magnitude of rainfall
 432 increases; both simulated and proxy anomalies show increases of 200–400 mm yr⁻¹ (Fig. 7c and d).
 433

434 Concerning the spatial patterns during the LIG, Figure 8 shows simulated mean SST anomalies
 435 (calculated both annually and during JAS/JFM) from the *lig127k* simulation and LIG proxy anomalies
 436 from two sources, Capron *et al.* (2017) and Hoffman *et al.* (2017). When annual anomalies are
 437 considered, despite the lack of reconstructions in the Capron *et al.* (2017) data (Table 4), there is
 438 relatively good agreement (RMSE = 2.44°C and 2.94°C for the Capron *et al.* (2017) and Hoffman *et al.*
 439 *et al.* (2017) data, respectively, and which is within the average uncertainty range), between simulated
 440 and observed SST anomalies in the Northern Hemisphere (and in particular in the North Atlantic),
 441 with both suggesting increased temperatures during the LIG of up to 3°C (Fig. 8a). There are
 442 discrepancies, such as in the Norwegian Sea, where the Hoffman *et al.* (2017) reconstructions suggest
 443 a cooler LIG than preindustrial, whereas the *lig127k* simulation shows a consistent warming; this is,
 444 however, consistent with previous work, and earlier climate models have also failed to capture this
 445 cooling (Capron *et al.* 2014, Stone *et al.* 2016). Note that, over Greenland and Antarctica, the Capron
 446 *et al.* (2017) proxy data show SAT, not SST, and are therefore not compared in this figure;
 447 comparison with simulated SAT, however, suggests that the model is capturing the sign, if not the
 448 magnitude, of annual change over these regions (not shown). During Northern Hemisphere summer,
 449 JAS (during which period Capron *et al.* [2017] has the most proxy locations [Table 4]), the simulated
 450 anomalies are in agreement with many, but not all, of the proxy locations (RMSE = 3.11°C and
 451 2.06°C for the Capron *et al.* (2017) and Hoffman *et al.* (2017) data, respectively); examples of where
 452 they differ, not just in magnitude but also sign, again include the Norwegian and Labrador Seas (Fig.
 453 8b). In Southern Hemisphere summer, JFM, the model suggests a general (but weak) cooling in the
 454 South Atlantic relative to preindustrial and a general (but weak) warming in the Southern Ocean (Fig.
 455 8c). However, certain proxy locations (such as off the coast of southern Africa) suggest a much
 456 warmer LIG than preindustrial (RMSE = 1.94°C and 4.24°C for the Capron *et al.* (2017) and Hoffman
 457 *et al.* (2017) data, respectively), which in stark contrast to the cooling in the same region from the
 458 *lig127k* simulation (Fig. 8c). In the Southern Ocean, the majority of simulated anomalies reproduce
 459 the observed sign of change, but not the magnitude; the *lig127k* simulation suggests temperature
 460 increases of up to 1°C, whereas both proxy datasets suggest SST increases of 2–3°C depending on
 461 location (Fig. 8c).
 462

463 It would therefore be reasonable to say that, for both warm climate simulations, whilst the model is
 464 capturing the sign and magnitude of change (for either temperature or rainfall) in some locations, this
 465 is highly geographically dependent and there are locations where the simulation fails to capture even



the sign of change. The model also appears to be seasonally dependent, with the *lig127k* simulation (but not the *midHolocene* simulation) correctly reproducing both the sign and magnitude of change during Northern Hemisphere summer in some locations, but not during Southern Hemisphere summer or annually.

3.2.3. Model-Model comparison: Do the CMIP6 HadGEM3 simulations show an improvement compared to older CMIP versions of the UK model?

Here we focus on model-model intercomparisons, comparing the HadGEM3 warm climate simulations with firstly those from previous versions of the UK model and secondly with those from other models included in CMIP5. It should be noted that although LIG experiments have been conducted previously with both model-model and model-data comparisons being made (Lunt *et al.* 2013), all of these experiments were carried out using early versions of the models and were thus not included in CMIP5. Moreover, as part of their assessment Lunt *et al.* (2013) considered a set of four simulations, at 130, 128, 125 and 115 ka, none of which are directly comparable to the current HadGEM3 *lig127k* simulation. Instead, a LIG simulation has recently been undertaken using one of the original versions of the UK's physical climate model, HadCM3, and so this is used here to compare with the *lig127k* simulation. As discussed above, this section is divided into two parts: firstly the mean climate state of the warm climate simulations will be compared to the model's predecessors, focusing again on hydroclimate of the West African monsoon (given the known problem of simulated rainfall underestimation in this region, see e.g. Braconnot *et al.* [2007]). Here, both direct comparisons and RMSE values will again be examined, this time calculating the RMSE between the simulated rainfall anomaly from two older versions of the UK model versus the current HadGEM3 *midHolocene* and *lig127k* simulations (summarised in Table 4b). Secondly, previous generation simulations (from all available models included in CMIP5) will be compared to see whether the most recent HadGEM3 *midHolocene* simulation is now providing enough rainfall to allow vegetation growth across the Sahara; something which previous generations of models from CMIP5 did not (Braconnot *et al.* 2007).

3.2.3.1. Mean climate state from predecessors of HadGEM3

Regarding the magnitude and latitudinal extent of the West African monsoon, Figure 9 shows the JJA rainfall differences averaged over West Africa from the current *midHolocene* and *lig127k* simulation versus two of the model's predecessors. During the MH, the two most recent generations of the model (HadGEM3 and HadGEM2-ES) generally agree on drier conditions over the equatorial Atlantic and then wetter conditions over West Africa, however the oldest generation model (HadCM3) does not reproduce the Atlantic drying. Likewise the two most recent generations share a similar latitudinal distribution of rainfall above $\sim 5^{\circ}\text{N}$, with a wetter MH over land, peaking at $\sim 2\text{--}3\text{ mm day}^{-1}$ at $\sim 11\text{--}12^{\circ}\text{N}$. Interestingly, the previous version of the model (HadGEM2-ES) shows the



strongest and most northwardly displaced rainfall peak, as discussed in previous work (e.g. Huag *et al.* 2001, Otto-Bliesner *et al.* 2017, Singarayer *et al.* 2017, Wang *et al.* 2014); the most recent version, HadGEM3, has lower northward displacement compared to the two older versions of the model. Both recent versions suggest that the monsoon region extends to $\sim 17^\circ\text{N}$, above which the differences between the MH and PI reduce to near zero. In contrast, HadCM3 suggests a generally weaker, but latitudinally more extensive, monsoon region, suggesting a wetter MH (by $\sim 1 \text{ mm day}^{-1}$) as far north as 20°N and beyond. For the LIG, HadGEM3 is showing a much stronger monsoon region relative to the *piControl*, compared to HadCM3. However, in terms of extent, similar results are shown to those for the MH, with HadCM3 showing a generally weaker, but more northwardly displaced, monsoon region. In this older generation model, positive rainfall anomalies of $\sim 2\text{--}3 \text{ mm day}^{-1}$ extend as far north as $17\text{--}18^\circ\text{N}$, whereas in HadGEM3 they fall to $\sim 1 \text{ mm day}^{-1}$ at these latitudes.

In terms of the spatial patterns of the West African monsoon, Figure 10 and Figure 11 show the JJA daily rainfall climatology differences from the same three model generations for the MH and LIG, respectively. During the MH, consistent with Figure 9, the two most recent simulations generally agree ($\text{RMSE} = 0.46 \text{ mm day}^{-1}$) and show similar spatial patterns, with a drier equatorial Atlantic during the MH and then increased rainfall around 10°N (Fig. 10a and b for HadGEM3 and HadGEM2-ES, respectively). Both simulations also suggest that the increases in rainfall extend longitudinally across the entire continent, with the largest changes not only occurring across western and central regions but also further east. In contrast, HadCM3 is less consistent than HadGEM3 ($\text{RMSE} = 0.53 \text{ mm day}^{-1}$) and only suggests a wetter MH over West Africa; moreover, again consistent with Figure 9, HadCM3 suggests that although the West African monsoon region is longitudinally narrower, it is latitudinally wider than the other two simulations (Fig. 10c). HadCM3 also differs from the other simulations over the equatorial Atlantic, showing a region of drying that is not only stronger in magnitude (with the MH being over 5 mm day^{-1} drier than the PI in HadCM3, compared to $\sim 2\text{--}3 \text{ mm day}^{-1}$ in the two most recent simulations), but also larger in terms of latitude and longitude extent (Fig. 10c).

During the LIG, only the most recent and oldest version of the model can be compared, as a LIG simulation using HadGEM2-ES is unavailable. In Figure 11 there is a noticeable difference between generations and the level of agreement is the lowest across all simulation combinations ($\text{RMSE} = 1.57 \text{ mm day}^{-1}$), with the most recent HadGEM3 showing greatly increased rainfall across all of northern Africa, centred on 10°N but extending from $\sim 5^\circ\text{N}$ to almost 20°N and beyond (Fig. 11a), again consistent with Figure 9. In contrast, and similar to the MH results, in HadCM3 the largest rainfall increases are confined to Western Africa only, rather than extending longitudinally across the continent (Fig. 11b). However, in terms of latitudinal extent, HadCM3 is showing weak wet anomalies all the way to the Mediterranean, whereas the monsoon region diminishes further south (at



~30°N) in HadCM3 and dry anomalies are suggested North of this. Another noticeable difference is the region of drying, with the most recent generation model placing this over the equatorial Atlantic (consistent with the MH) but HadCM3 shifting this further east, over most of central Africa (Fig. 11b). The region of equatorial Atlantic drying shown by the more recent versions of the model is actually wetter during this HadCM3 LIG simulation.

It would therefore appear that, for the MH, whilst there is less difference between the most recent two configurations of the model (in terms of a more localised West African monsoon region), there nevertheless has been improvement since the oldest version of the UK's physical climate model. For the LIG, where unfortunately there is no intermediate generation, it would be reasonable to say that again considerable change has occurred since the oldest generation model, with the suggestion that, although HadCM3 is identifying an enhanced monsoon which extends to the Mediterranean (albeit with very weak anomalies), at lower latitudes it is not showing the level of northward displacement as the most recent version, apart from in the far western regions.

3.2.3.2. *Rainfall across the Sahara*

Given that the warm climate simulations, and indeed the *piControl*, did not use interactive, but rather prescribed, vegetation, it is not possible to directly test if the model is reproducing the 'Saharan greening' that proxy data suggest. For example, Jolly *et al.* (1998a, 1998b) analysed MH pollen assemblages across northern Africa and suggested that some areas south of 23°N (characterised by desert today) were grassland and xerophytic woodland/scrubland during the MH (Joussaume *et al.* 1999). To circumvent this caveat, Joussaume *et al.* (1999) developed a method for indirectly assessing Saharan greening, based on the annual mean rainfall anomaly relative to a given model's modern simulation. Using the water-balance module from the BIOME3 equilibrium vegetation model (Haxeltine & Prentice 1996), Joussaume *et al.* (1999) calculated the increase in mean annual rainfall, zonally averaged over 20°W-30°E, required to support grassland at each latitude from 0 to 30°N, compared to the modern rainfall at that latitude. This was then used to create maximum and minimum estimates, within which bounds the model's annual mean rainfall anomaly must lie to suggest enough of an increase to support grassland (Joussaume *et al.* 1999).

Therefore, an adapted version of Figure 3a in Joussaume *et al.* (1999) is shown here in Figure 12, which includes the above mean annual rainfall anomalies from not only the current *midHolocene* simulation, but also all previous MH simulations from CMIP5. Firstly of note is that, despite the equatorial Atlantic drying that all the models show (seen, for example, in Figure 5), the HadGEM3 *midHolocene* simulation is showing a peak in rainfall further south compared to many other CMIP5 models, suggesting less northward displacement of the rainbelt relative to the other models (Fig. 12). Concerning the threshold required to support grassland, it is clear that although the current



midHolocene simulation is showing an increase in mean annual rainfall further north than some of the models, including its predecessor HadGEM2-ES, and is just within the required bounds at lower latitudes (e.g. up to 17°N), north of this the current *midHolocene* simulation is not meeting the required threshold, neither are any of the other CMIP5 models after ~18°N (Fig. 12). It would therefore appear that, although some improvement has been made since CMIP5 and earlier models, the latest version of the UK's physical climate model it is still unable to reproduce the amount of rainfall necessary to give the 'Saharan greening' suggested by proxy data during the MH.

4. SUMMARY AND CONCLUSIONS

This study has conducted and assessed the mid-Holocene and Last Interglacial simulations using the latest version of the UK's physical climate model, HadGEM3-GC3.1, comparing the results with available proxy data, previous versions of the same model, and other models from CMIP's previous iteration, CMIP5. Both the *midHolocene* and *lig127k* simulations followed the experimental design defined in Otto-Bliesner *et al.* (2017) and under the auspices of CMIP6/PMIP4. Both simulations were run for a 350-400 year spin-up phase, during which time atmospheric and oceanic equilibrium was assessed, and once an acceptable level of equilibrium had been reached, the production runs were started.

Concerning the results from the spin-up phase, comparison to the metrics used to assess the CMIP6 *piControl* suggest that both warm climate simulations reached an acceptable state of equilibrium, in the atmosphere at least, to allow the production runs to be undertaken. From these, both simulations are showing global temperatures consistent with the latitudinal and seasonal distribution of insolation, and with previous work (e.g. Otto-Bliesner *et al.* 2017). Globally, whilst both the simulations are mostly capturing the sign and, in some places, magnitude of change relative to the PI, similar to previous model simulations this is geographically and seasonally dependent. It should be noted that the proxy data (against which the simulations are evaluated) also contain a high level of uncertainty in both space and time, and so it is encouraging that the simulations are generally reproducing the large-scale sign of change, if not at an individual location. Likewise, the behaviour of the West African monsoon in both simulations is consistent with current understanding (e.g. Huag *et al.* 2001, Singarayer *et al.* 2017, Wang *et al.* 2014), which suggests a wetter (and possibly latitudinally wider, and/or northwardly displaced) monsoon during the MH and LIG, relative to the PI. Regarding model development in simulating the West African monsoon, although there has been an improvement since the oldest version of the UK's physical climate model (HadCM3), the two most recent version of the model yield similar results in terms of both intensity and position. Lastly, regarding the well-documented 'Saharan greening' during the MH, results here suggest that the most recent version of the UK's physical climate model is consistent with all other previous models to date.



614 In conclusion, the results suggest that the most recent version of the UK's physical climate model is
615 reproducing climate conditions consistent with the known changes to insolation during these two
616 warm periods, and is consistent with previous versions of the same model, and other models. Even
617 though the *lig127k* simulation did not contain any influx of Northern Hemisphere meltwater, shown
618 by previous work to be a critical forcing in LIG warming, it is still nevertheless showing increased
619 temperatures in certain regions. A potential caveat of this conclusion, however, is the matter of spin-
620 up and the fact that neither of the current warm climate simulations were in oceanic equilibrium when
621 the production runs were undertaken. The production runs were undertaken nevertheless because the
622 resources required to run for several thousands of years (needed to reach true oceanic equilibrium)
623 would have been impossible to obtain, but future simulations using this model should endeavour to
624 obtain a better level of oceanic equilibrium. Another limitation of using this particular version of the
625 model is that certain processes, such as vegetation and atmospheric chemistry, were prescribed, rather
626 than allowed to be dynamically evolving. Moreover, for reasons of necessity some of the boundary
627 conditions were left as PI, such as vegetation, surface like, anthropogenic deforestation and aerosols; a
628 better simulation might be achieved if these were prescribed for the MH. Processes and boundary
629 conditions such as these may be of critical importance regarding climate sensitivity during the MH
630 and the LIG, and therefore ongoing work is underway to repeat both of these experiments using the
631 most recent version of the UK's Earth Systems model, UKESM1. Here, although the atmospheric
632 core is HadGEM3, UKESM1 contains many other earth system components (e.g. dynamic
633 vegetation), and therefore in theory should be able to better reproduce these paleoclimate states.

635 DATA AVAILABILITY

636 For the MH reconstructions, the data can be found within the Supplementary Online Material of
637 Bartlein *et al.* (2011), at <https://link.springer.com/article/10.1007/s00382-010-0904-1>. For the LIG
638 reconstructions, the data can be found within the Supplementary Online Material of Capron *et al.*
639 (2017), at <https://www.sciencedirect.com/science/article/pii/S0277379117303487?via%3Dihub>, and
640 the Supplementary Online Material of Hoffman *et al.* (2017), at
641 <https://science.sciencemag.org/content/suppl/2017/01/23/355.6322.276.DC1>. The model simulations
642 will be uploaded in early 2020 to the Earth System Grid Federation (ESGF) WCRP Coupled Model
643 Intercomparison Project (Phase 6), but are not yet available. The simulations are currently available
644 by directly contacting the lead author.

646 COMPETING INTERESTS

647 The authors declare that they have no conflict of interest.

649 AUTHOR CONTRIBUTION



650 CJRW conducted the *midHolocene* simulation, carried out the analysis, produced the figures, wrote
651 the majority of the manuscript, and led the paper. MVG conducted and provided the *lig127k*
652 simulation, and contributed to some of the analysis and writing. EC provided the proxy data, and
653 contributed to some of the writing. IMV provided the HadCM3 LIG simulation. PJV provided the
654 HadCM3 MH simulation. JS contributed to some of the writing. All authors proofread the
655 manuscript and provided comments.

656

657 **ACKNOWLEDGEMENTS**

658 CJRW acknowledges the financial support of the UK Natural Environment Research Council-funded
659 SWEET project (Super-Warm Early Eocene Temperatures), research grant NE/P01903X/1. CJRW
660 also acknowledges the financial support of the Belmont-funded PACMEDY (PALaeo-Constraints on
661 Monsoon Evolution and Dynamics) project, as does JS. MVG and LCS acknowledge the financial
662 support of the NERC research grants NE/P013279/1 and NE/P009271/1. EC acknowledges financial
663 support from the ChronoClimate project, funded by the Carlsberg Foundation.



664 REFERENCES

- 665 Bartlein, P. J., Harrison, S. P., Brewer, S., et al. (2011). ‘Pollen-based continental climate
 666 reconstructions at 6 and 21 ka: a global synthesis’. *Clim. Dyn.* 37: 775–802. DOI:10.1007/s00382-
 667 010-0904-1
- 668
- 669 Berger, A. L. (1978). ‘Long-term variations of daily insolation and Quaternary climatic changes’. *J.*
 670 *Atmos. Sci.* 35: 2362-2367. [https://doi.org/10.1175/1520-0469\(1978\)035<2362:LTVODI>2.0.CO;2](https://doi.org/10.1175/1520-0469(1978)035<2362:LTVODI>2.0.CO;2)
 671
- 672 Berger, A. L. & Loutre, M. F. (1991). Insolation values for the climate of the last 10,000,000 years.
 673 *Quaternary Sci. Rev.* 10: 297–317. [https://doi.org/10.1016/0277-3791\(91\)90033-Q](https://doi.org/10.1016/0277-3791(91)90033-Q)
 674
- 675 Braconnot, P., Harrison, S. P., Kageyama, M., et al. (2012). ‘Evaluation of climate models using
 676 palaeoclimatic data’. *Nature Climate Change.* 2 (6): 417. DOI: 10.1038/NCLIMATE1456
 677
- 678 Braconnot, P., Harrison, S. P., Otto-Bliesner, B., et al. (2011). ‘The palaeoclimate modelling
 679 intercomparison project contribution to CMIP5’. *CLIVAR Exch. Newsl.* 56: 15–19
 680
- 681 Braconnot, P., Otto-Bliesner, B., Harrison, S., Joussaume, S. et al. (2007). ‘Results of PMIP2 coupled
 682 simulations of the Mid-Holocene and Last Glacial Maximum – Part 1: experiments and large-scale
 683 features’. *Clim. Past.* 3: 261-277. <https://doi.org/10.5194/cp-3-261-2007>
 684
- 685 Capron, E., Govin, A., Stone, E. J., et al. (2014). ‘Temporal and spatial structure of multi-millennial
 686 temperature changes at high latitudes during the Last Interglacial’. *Quat. Sci. Rev.* 103: 116-133.
 687 <https://doi.org/10.1016/j.quascirev.2014.08.018>
 688
- 689 Capron, E., Govin, A., Feng R. et al. (2017). ‘Critical evaluation of climate syntheses to benchmark
 690 CMIP6/PMIP4 127 ka Last Interglacial simulations in the high-latitude regions’. *Quat. Sci. Rev.*
 691 168: 137-150. DOI: 10.1016/j.quascirev.2017.04.019
 692
- 693 Drake, N. A., Blench, R. M., Armitage, S. J., et al. (2011). ‘Ancient watercourses and biogeography
 694 of the Sahara explain the peopling of the desert’. *Proceedings of the National Academy of Sciences.*
 695 108 (2): 458-462. DOI: 10.1073/pnas.1012231108
 696
- 697 Fischer, H., Meissner, K. J., Mix, A. C. et al. (2018). ‘Palaeoclimate constraints on the impact of 2°C
 698 anthropogenic warming and beyond’. *Nature Geoscience.* 11: 474-485.
 699 <https://doi.org/10.1038/s41561-018-0146-0>
 700



- 701 Gates, W. L. (1976). ‘The numerical simulation of ice-age climate with a global general circulation
 702 model’. *J. Atmos. Sci.* 33: 1844–1873. DOI: 10.1175/1520-
 703 0469(1976)033<1844:TNSOIA>2.0.CO;2
 704
- 705 Gregoire, L. J., Valdes, P. J., Payne, A. J. & Kahana, R. (2011). ‘Optimal tuning of a GCM using
 706 modern and glacial constraints’. *Clim Dyn.* 37: 705-719. DOI:10.1007/s00382-010-0934-8
 707
- 708 Guarino, M.-V., Sime, L., Schroeder, D., et al. (2019). ‘Machine dependence and reproducibility for
 709 coupled climate simulations: The HadGEM3-GC3.1 CMIP Preindustrial simulation’. *GMD*. In press.
 710 <https://doi.org/10.5194/gmd-2019-83>
 711
- 712 Harrison, S. P., Bartlein, P. J., Brewer, S., et al. (2014). ‘Climate model benchmarking with glacial
 713 and mid-Holocene climates’. *Clim. Dyn.* 43: 671–688. <https://doi.org/10.1007/s00382-013-1922-6>
 714
- 715 Harrison, S. P., Bartlein, P. J., Izumi, K., et al. (2015). ‘Evaluation of CMIP5 palaeo-simulations to
 716 improve climate projections’. *Nature Climate Change*. 5: 735. DOI: 10.1038/nclimate2649
 717
- 718 Haxeltine, A. & Prentice, I. C. (1996). ‘BIOME3: an equilibrium terrestrial biosphere model based on
 719 ecophysiological constraints, resource availability, and competition among plant functional types’.
 720 *Global Biogeochemical Cycles*. 10 (4): 693-709. DOI: 10.1029/96GB02344
 721
- 722 Haywood, A. M., Dowsett, H. J., Dolan, A. M. et al. (2016). ‘The Pliocene Model Intercomparison
 723 Project (PlioMIP) Phase 2: scientific objectives and experimental design’. *Clim. Past*. 12: 663-675.
 724 <https://doi.org/10.5194/cp-12-663-2016>
 725
- 726 Haug, G., Hughen, K. A., Sigman, D. M., et al. (2001). ‘Southward migration of the intertropical
 727 convergence zone through the Holocene’. *Science*. 293: 1304-1308. DOI: 10.1126/science.1059725
 728
- 729 Hoelzmann, P., Jolly, D., Harrison, S. P., et al. (1998). ‘Mid-Holocene land-surface conditions in
 730 northern Africa and the Arabian Peninsula: A data set for the analysis of biogeophysical feedbacks in
 731 the climate system’. *Global Biogeochemical Cycles*. 12 (1): 35-51.
 732 <https://doi.org/10.1029/97GB02733>
 733
- 734 Hoffman, J. S., Clark, P. U., Parnell, A. C., et al. (2017). ‘Regional and global sea-surface
 735 temperatures during the last interglaciation’. *Science*. 355: 276-279. DOI: 10.1126/science.aai8464
 736



- 737 Hollingsworth, A., Kållberg, P., Renner, V. & Burridge, D. M. (1983). ‘An internal symmetric
 738 computational instability’. QJRM. 109: 417–428. <https://doi.org/10.1002/qj.49710946012>
 739
- 740 Holloway, M. D., Sime, L. C., Singarayer, J. S., et al. (2016). ‘Antarctic last interglacial isotope peak
 741 in response to sea ice retreat not ice-sheet collapse’. Nature Comms. 7: 12293.
 742 <https://doi.org/10.1038/ncomms12293>
 743
- 744 Jungclaus, J. H., Bard, E., Baroni, M. et al. (2017). ‘The PMIP4 contribution to CMIP6 – Part 3: The
 745 last millennium, scientific objective, and experimental design for the PMIP4 past1000 simulations’.
 746 GMD. 10: 4005-4033. <https://doi.org/10.5194/gmd-10-4005-2017>
 747
- 748 Jolly, D., Harrison, S. P., Damnati, B. & Bonnefille, R. (1998a). ‘Simulated climate and biomes of
 749 Africa during the Late Quaternary: Comparison with pollen and lake status data’. Quat. Sci. Rev. 17
 750 (6-7): 629-657. [https://doi.org/10.1016/S0277-3791\(98\)00015-8](https://doi.org/10.1016/S0277-3791(98)00015-8)
 751
- 752 Jolly, D., Prentice, I. C., Bonnefille, R. et al. (1998b). ‘Biome reconstruction from pollen and plant
 753 macrofossil data for Africa and the Arabian peninsula at 0 and 6000 years’. J. Biogeography. 25 (6):
 754 1007-1027
 755
- 756 Joussaume, S. & Taylor K. E. (1995). ‘Status of the Paleoclimate Modeling Intercomparison Project’.
 757 In: Proceedings of the First International AMIP Scientific Conference, WCRP-92 425. 430 Monterey,
 758 USA
 759
- 760 Joussaume, S., Taylor, K. E., Braconnot, P. et al. (1999). ‘Monsoon changes for 6000 years ago:
 761 Results of 18 simulations from the Paleoclimate Modeling Intercomparison Project (PMIP)’. GRL.
 762 26 (7): 859-862. <https://doi.org/10.1029/1999GL900126>
 763
- 764 Kageyama, M., Albani, S., Braconnot, P. et al. (2017). ‘The PMIP4 contribution to CMIP6 – Part 4:
 765 Scientific objectives and experimental design of the PMIP4-CMIP6 Last Glacial Maximum
 766 experiments and PMIP4 sensitivity experiments’. GMD. 10: 4035-4055.
 767 <https://doi.org/10.5194/gmd-10-4035-2017>
 768
- 769 Kageyama, M., Braconnot, P., Harrison, S. P. et al. (2018). ‘The PMIP4 contribution to CMIP6 –
 770 Part 1: Overview and over-arching analysis plan’. GMD. 11: 1033-1057.
 771 <https://doi.org/10.5194/gmd-11-1033-2018>
 772



- 773 Kuhlbrodt, T., Jones, C. G., Sellar, A. et al. (2018). ‘The low resolution version of HadGEM3 GC3.1:
 774 Development and evaluation for global climate’. *J. Adv. Model. Earth Sy.* 10: 2865-2888.
 775 <https://doi.org/10.1029/2018MS001370>
 776
- 777 Kohfeld, K. E., Graham, R. M., de Boer, A. M. et al. (2013). ‘Southern Hemisphere westerly wind
 778 changes during the Last Glacial Maximum: paleo-data synthesis’. *Quat. Sci. Rev.* 68: 76-95.
 779 <https://doi.org/10.1016/j.quascirev.2013.01.017>
 780
- 781 Lézine, A. M., Hély, C., Grenier, C. et al. (2011). ‘Sahara and Sahel vulnerability to climate changes,
 782 lessons from Holocene hydrological data’. *Quat. Sci. Rev.* 30 (21-22): 3001-3012.
 783 DOI:10.1016/j.quascirev.2011.07.006
 784
- 785 Liu, Z., Zhu, J., Rosenthal, Y. et al. (2014). ‘The Holocene temperature conundrum’. *PNAS.* 111
 786 (34): 3501-3505. DOI: 10.1073/pnas.1407229111
 787
- 788 Lunt, D. J., Abe-Ouchi, A., Bakker, P. et al. (2013). ‘A multi-model assessment of last interglacial
 789 temperatures’. *Clim. Past.* 9: 699–717. <https://doi.org/10.5194/cp-9-699-2013>
 790
- 791 Lunt, D. J., Foster, G. L., Haywood, A. M. & Stone, E. J. (2008). ‘Late Pliocene Greenland
 792 glaciation controlled by a decline in atmospheric CO₂ levels’. *Nature.* 454 (7208): 1102. DOI:
 793 10.1038/nature07223.
 794
- 795 Marcott, S. A., Shakun, J. D., Clark, P. U. & Mix, A. C. (2013). ‘A reconstruction of regional and
 796 global temperature for the past 11,300 years’. *Science.* 399 (6124): 1198-1201. DOI:
 797 10.1126/science.1228026.
 798
- 799 Menary, M. B., Kuhlbrodt, T., Ridley, J. et al. (2018). ‘Pre-industrial control simulations with
 800 HadGEM3-GC3.1 for CMIP6’. *JAMES.* 10: 3049–3075. <https://doi.org/10.1029/2018MS001495>
 801
- 802 New, M., Lister, D., Hulme, M. & Makin, I. (2002). ‘A high-resolution data set of surface climate
 803 over global land areas’. *Clim Res.* 21: 2217–2238. DOI:10.3354/cr021001
 804
- 805 Otto-Bliesner, B. L., Braconnot, P., Harrison, S. P. et al. (2017). ‘The PMIP4 contribution to CMIP6
 806 – Part 2: Two interglacials, scientific objective and experimental design for Holocene and Last
 807 Interglacial simulations’. *GMD.* 10: 3979-4003. <https://doi.org/10.5194/gmd-10-3979-2017>
 808



- 809 Rachmayani, R., Prange, M., Lunt, D. J., et al. (2017). ‘Sensitivity of the Greenland Ice Sheet to
 810 interglacial climate forcing: MIS 5e versus MIS11’. *Paleoceanography*. 32 (11): 1089-1101.
 811 <https://doi.org/10.1002/2017PA003149>
 812
- 813 Ramstein, G., Fluteau, F., Besse, J. & Joussaume, S. (1997). ‘Effect of orogeny, plate motion and
 814 land–sea distribution on Eurasian climate change over the past 30 million years’. *Nature*. 386 (6627):
 815 788. <https://doi.org/10.1038/386788a0>
 816
- 817 Ridley, J., Blockley, E., Keen, A. B. et al. (2017). ‘The sea ice model component of HadGEM3-
 818 GC3.1’. *GMD*. 11: 713-723. <https://doi.org/10.5194/gmd-11-713-2018>
 819
- 820 Rind, D. & Peteet, D. (1985). ‘Terrestrial conditions at the last glacial maximum and CLIMAP sea-
 821 surface temperature estimates: Are they consistent?’ *Quat. Res.* 2: 1-22. DOI:10.1016/0033-
 822 5894(85)90080-8
 823
- 824 Schmidt, G. A., Annan, J. D., Bartlein, P. J. et al. (2014). ‘Using paleo-climate comparisons to
 825 constrain future projections in CMIP5’. *Clim. Past*. 10: 221-250. [https://doi.org/10.5194/cp-10-221-](https://doi.org/10.5194/cp-10-221-2014)
 826 2014
 827
- 828 Singarayer, J. S. & Burrough, S. L. (2015). ‘Interhemispheric dynamics of the African rainbelt during
 829 the late Quaternary’. *Quaternary Science Reviews*. 124: 48-67.
 830 DOI: 10.1016/j.quascirev.2015.06.021
 831
- 832 Singarayer, J. S., Valdes, P. J. & Roberts, W. H. G. (2017). ‘Ocean dominated expansion and
 833 contraction of the late Quaternary tropical rainbelt’. *Nature Scientific Reports*. 7: 9382.
 834 DOI:10.1038/s41598-017-09816-8
 835
- 836 Stone, E. J., Capron, E., Lunt, D. J., et al. (2016). ‘Impact of meltwater on high-latitude early Last
 837 Interglacial climate’. *Clim. Past*. 12: 1919–1932. <https://doi.org/10.5194/cp-12-1919-2016>
 838
- 839 Storkey, D., Megann, A., Mathiot, P. et al. (2017). ‘UK Global Ocean GO6 and GO7: A traceable
 840 hierarchy of model resolutions’. *GMD*. 11: 3187-3213. <https://doi.org/10.5194/gmd-11-3187-2018>
 841
- 842 Taylor, K. E., Stouffer, R. J. & Meehl, G. A. (2011). ‘An overview of CMIP5 and the experiment
 843 design’. *Bull. Am. Meteorol. Soc.* 93: 485-498. <https://doi.org/10.1175/BAMS-D-11-00094.1>
 844



- 845 Turney, C. S. M. & Jones, R. T. (2010). ‘Does the Agulhas Current amplify global temperatures
 846 during super-interglacials?’ J. Quat. Sci. 25 (6): 839-843. <https://doi.org/10.1002/jqs.1423>
 847
- 848 Walters, D. N., A., Baran, I., Boutle, M. E. et al. (2017). ‘The Met Office Unified Model Global
 849 Atmosphere 7.0/7.1 and JULES Global Land 7.0 configurations’. GMD. 12: 1909-1923.
 850 <https://doi.org/10.5194/gmd-12-1909-2019>
 851
- 852 Wang, X., et al. (2006). ‘Interhemispheric anti-phasing of rainfall during the last glacial period’.
 853 Quat. Sci. Rev. 25: 3391-3403. DOI: 10.1016/j.quascirev.2006.02.009
 854
- 855 Wang, Y., Cheng, H., Edwards, R.L., et al. (2008). ‘Millennial-and orbital-scale changes in the East
 856 Asian monsoon over the past 224,000 years’. Nature. 451 (7182): 1090. DOI: 10.1038/nature06692
 857
- 858 Wang, P. X., et al. (2014). ‘The global monsoon across timescales: coherent variability of regional
 859 monsoons’. Clim. Past. 10: 2007-2052. <https://doi.org/10.5194/cp-10-2007-2014>
 860
- 861 Williams, K. D., Copsey, D., Blockley E. W., et al. (2017). ‘The Met Office Global Coupled Model
 862 3.0 and 3.1 (GC3.0 and GC3.1) Configurations’. JAMES. 10 (2): 357-380.
 863 <https://doi.org/10.1002/2017MS001115>
 864



865 LIST OF TABLES

866 Table 1 - Astronomical parameters and atmospheric trace gas concentrations used in HadGEM3
 867 *midHolocene* and *lig127k* simulations

868

869 Table 2 - Trends (per century) in global mean measures of climate equilibrium for the last hundred
 870 years of the simulations, adapted from and including *piControl* results from Menary *et al.* (2018).

871 Note - For temperature, Menary *et al.* (2018) provide SAT. For OceTemp and OceSal, these were
 872 calculated using the full-depth ocean for the *piControl*, whereas in the other two simulations these
 873 fields were calculated down to a depth of 1045m

874

875 Table 3 - Global 1.5 m air temperature means and anomalies from HadGEM3 *piControl*,
 876 *midHolocene* and *lig127k* production runs (100-year climatology)

877

878 Table 4 - RMSE values for *midHolocene* and *lig127k* production runs (100-year climatology) versus:
 879 a) proxy data from Bartlein *et al.* (2011) for the MH and Capron *et al.* (2017) / Hoffman *et al.* (2017)
 880 for the LIG; b) MH and LIG simulations from previous versions of UK model. Regarding the proxy
 881 data comparisons in a), for JAS the simulated SST anomalies are compared to Northern Hemisphere
 882 summer reconstructions and for JFM the simulated SST anomalies are compared to Southern
 883 Hemisphere summer reconstructions. Note that, as shown in Figure 8, proxy locations show SST over
 884 ocean and SAT over Greenland/Antarctica; to calculate RMSE values, however, only SST data were
 885 used

886

887 LIST OF FIGURES

888 Figure 1 - Latitude-month insolation (incoming SW radiative flux) anomalies: a) *midHolocene* -
 889 *piControl*; b) *lig127k* - *piControl*

890

891 Figure 2 - Annual global mean atmospheric fields from HadGEM3 *piControl*, *midHolocene* and
 892 *lig127k* spin-up phases: a) 1.5 m air temperature; b) TOA. Thin lines in b) show annual TOA, thick
 893 lines show 11-year running mean

894

895 Figure 3 - Annual global mean oceanic fields from HadGEM3 *piControl*, *midHolocene* and *lig127k*
 896 spin-up phases: a) OceTemp down to 1045m; b) OceSal down to 1045m

897

898 Figure 4 - 1.5 m air temperature climatology differences, HadGEM3 *midHolocene* and *lig127k*
 899 production runs versus HadGEM3 *piControl* production run, 100-year climatology from each: a)
 900 *midHolocene* - *piControl*, JJA; b) *midHolocene* - *piControl*, DJF; c) *lig127k* - *piControl*, JJA; d)
 901 *lig127k* - *piControl*, DJF



902

903 Figure 5 – JJA rainfall & 850mb wind climatology differences, HadGEM3 *midHolocene* and *lig127k*
 904 production runs versus HadGEM3 *piControl* production run, 100-year climatology from each: a)
 905 *midHolocene* – *piControl*, JJA; b) *lig127k* – *piControl*; c) *lig127k* – *midHolocene*

906

907 Figure 6 – JJA rainfall differences by latitude, averaged over West Africa (20°W-30°E, including both
 908 land and ocean points), HadGEM3 *midHolocene* and *lig127k* production runs versus HadGEM3
 909 *piControl* production run, 100-year climatology from each year

910

911 Figure 7 – Simulated versus proxy MAT and MAP anomalies. Left-hand side panels show simulated
 912 gridded anomalies from HadGEM3 (*midHolocene* production run – *piControl* production run, 100-
 913 year climatology from each), right-hand side panels show proxy data from Bartlein *et al.* (2011) (MH
 914 – preindustrial). Proxy data locations are projected onto model grid: a) Simulated MAT; b) Proxy
 915 MAT; c) Simulated MAP; d) Proxy MAP

916

917 Figure 8 – Simulated versus proxy SST anomalies. Background gridded data show simulated
 918 anomalies (*lig127k* production run – *piControl* production run) from HadGEM3 (100-year
 919 climatology), circles show proxy data (LIG – preindustrial) from Capron *et al.* (2017) and triangles
 920 show proxy data (LIG – preindustrial) from Hoffman *et al.* (2017). Proxy data locations are projected
 921 onto model grid: a) Annual data; b) Northern Hemisphere summer (JAS); c) Southern Hemisphere
 922 summer (JFM). Note that proxy locations show SST over ocean and SAT over Greenland/Antarctica

923

924 Figure 9 – JJA daily rainfall climatology differences (MH and LIG-PI) by latitude, averaged over
 925 West Africa (20°W-30°E, including both land and ocean points), for the various generations of the
 926 UK's physical climate model, 100-year climatology from each (50-year climatology for HadCM3
 927 LIG). Solid lines show MH simulations, dotted lines show LIG simulations. Note that due to the low
 928 spatial resolution in HadCM3, values in between latitude points have been interpolated

929

930 Figure 10 – JJA daily rainfall climatology differences (MH-PI) for the various generations of the
 931 UK's physical climate model, 100-year climatology from each: a) HadGEM3; b) HadGEM2-ES; c)
 932 HadCM3

933

934 Figure 11 – JJA daily rainfall climatology differences (LIG-PI) for the various generations of the
 935 UK's physical climate model, 100-year climatology from HadGEM3, 50-year climatology from
 936 HadCM3: a) HadGEM3; b) HadCM3

937



938 Figure 12 – Annual mean rainfall over West Africa, zonally averaged from 20°W-30°E, HadGEM3
939 and CMIP5 *midHolocene* production run minus corresponding *piControl* production runs, 100-year
940 climatology. Solid line shows HadGEM3, dotted lines show CMIP5 simulations. Grey dashes show
941 maximum and minimum bounds of the increase in rainfall required to support grassland at each
942 latitude, within which simulations must lie if producing enough rainfall to support grassland
943



944 TABLES

	<i>piControl</i>	<i>midHolocene</i>	<i>lig127k</i>
Astronomical parameters			
Eccentricity	0.016764	0.018682	0.039378
Obliquity	23.459	24.105°	24.04°
Perihelion-180°	100.33	0.87°	275.41°
Date of vernal equinox	March 21 at noon	March 21 at noon	March 21 at noon
Trace gases			
CO₂	284.3 ppm	264.4 ppm	275 ppm
CH₄	808.2 ppb	597 ppb	685 ppb
N₂O	273 ppb	262 ppb	255 ppb
Other GHG gases	CMIP DECK <i>piControl</i>	CMIP DECK <i>piControl</i>	CMIP DECK <i>piControl</i>

945

946 Table 1 - Astronomical parameters and atmospheric trace gas concentrations used in HadGEM3
 947 simulations

948

949

950

Variable	<i>piControl</i>	<i>midHolocene</i>	<i>lig127k</i>
TOA (W m²)	-0.002	-0.05	-0.06
1.5 m air temp (°C)	0.03	-0.06	-0.16
OceTemp (°C)	0.035	0.0002	0.02
OceSal (psu)	0.0001	-0.007	0.006

951

952 Table 2 - Trends (per century) in global mean measures of climate equilibrium for the last hundred
 953 years of the simulations, adapted from and including *piControl* results from Menary *et al.* (2018).

954 Note - For temperature, Menary *et al.* (2018) provide SAT. For OceTemp and OceSal, these were
 955 calculated using the full-depth ocean for the *piControl*, whereas in the other two simulations these
 956 fields were calculated down to a depth of 1045m

957

958

Time period	Means (°C)			Anomalies (°C)	
	<i>piControl</i>	<i>midHolocene</i>	<i>lig127k</i>	<i>midHolocene – piControl</i>	<i>lig127k – piControl</i>
Annual	13.8	13.67	14.29	-0.12	0.49
JJA	15.68	15.75	17.37	0.07	1.69
DJF	11.86	11.55	11.39	-0.31	-0.47

959

960 Table 3 - Global 1.5 m air temperature means and anomalies from HadGEM3 *piControl*,
 961 *midHolocene* and *lig127k* production runs (100-year climatology)

962

963

964

965

966



967
968
969
970
971
972
973

Metric	a) Simulations versus proxy data						
	MH	LIG					
MAT (°C)	2.45	Capron <i>et al.</i> (2017)			Hoffman <i>et al.</i> (2017)		
No. of proxy locations	638						
MAP (mm year ⁻¹)	280						
No. of proxy locations	651						
SST (°C)	Yearly	JAS	JFM	Yearly	JAS	JFM	
	2.44	3.11	1.94	2.94	2.06	4.24	
No. of proxy locations	7	24	15	86	12	6	
JJA rainfall (mm day ⁻¹)	b) Simulations versus simulations						
	MH			LIG			
	HadGEM2-ES v HadGEM3	HadCM3 v HadGEM3		HadCM3 v HadGEM3			
	0.46	0.53		1.57			

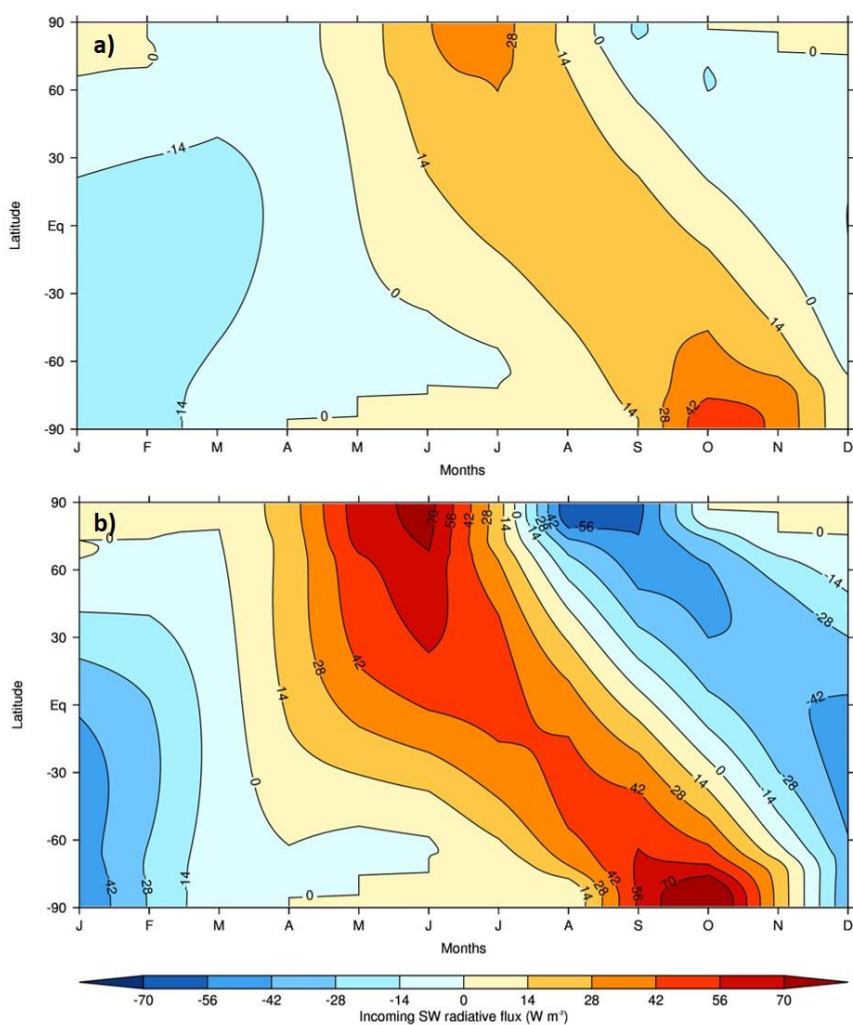
974

975 Table 4 - RMSE values for *midHolocene* and *lig127k* production runs (100-year climatology) versus:
 976 a) proxy data from Bartlein *et al.* (2011) for the MH and Capron *et al.* (2017) / Hoffman *et al.* (2017)
 977 for the LIG; b) MH and LIG simulations from previous versions of UK model. Regarding the proxy
 978 data comparisons in a), for JAS the simulated SST anomalies are compared to Northern Hemisphere
 979 summer reconstructions and for JFM the simulated SST anomalies are compared to Southern
 980 Hemisphere summer reconstructions. Note that, as shown in Figure 8, proxy locations show SST over
 981 ocean and SAT over Greenland/Antarctica; to calculate RMSE values, however, only SST data were
 982 used

983
984



985 **FIGURES**



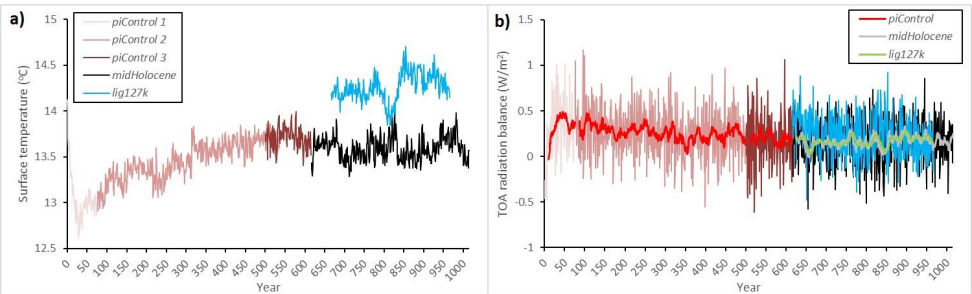
986

987 Figure 1 - Latitude-month insolation (incoming SW radiative flux) anomalies: a) *midHolocene* -
 988 *piControl*; b) *lig127k* - *piControl*

989

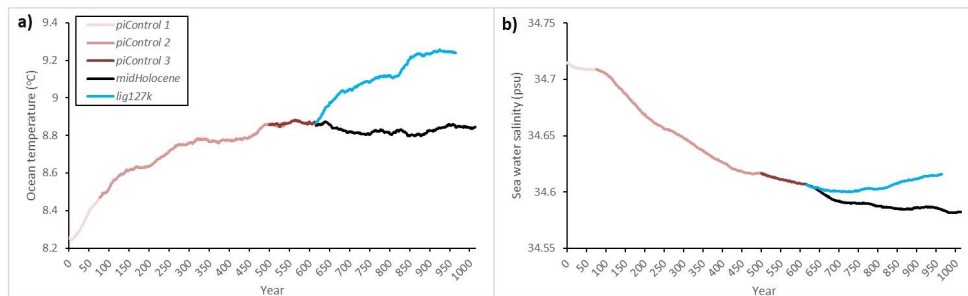


990



991

992 Figure 2 - Annual global mean atmospheric fields from HadGEM3 *piControl*, *midHolocene* and
993 *lig127k* spin-up phases: a) 1.5 m air temperature; b) TOA. Thin lines in b) show annual TOA, thick
994 lines show 11-year running mean
995



996

997 Figure 3 - Annual global mean oceanic fields from HadGEM3 *piControl*, *midHolocene* and *lig127k*

998 spin-up phases: a) *OceTemp* down to 1045m; b) *OceSal* down to 1045m

999

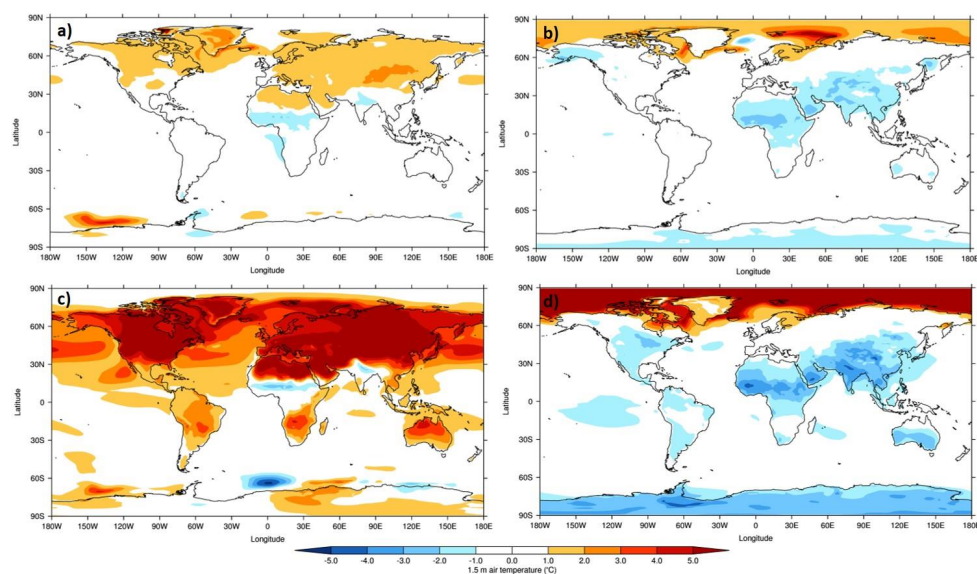
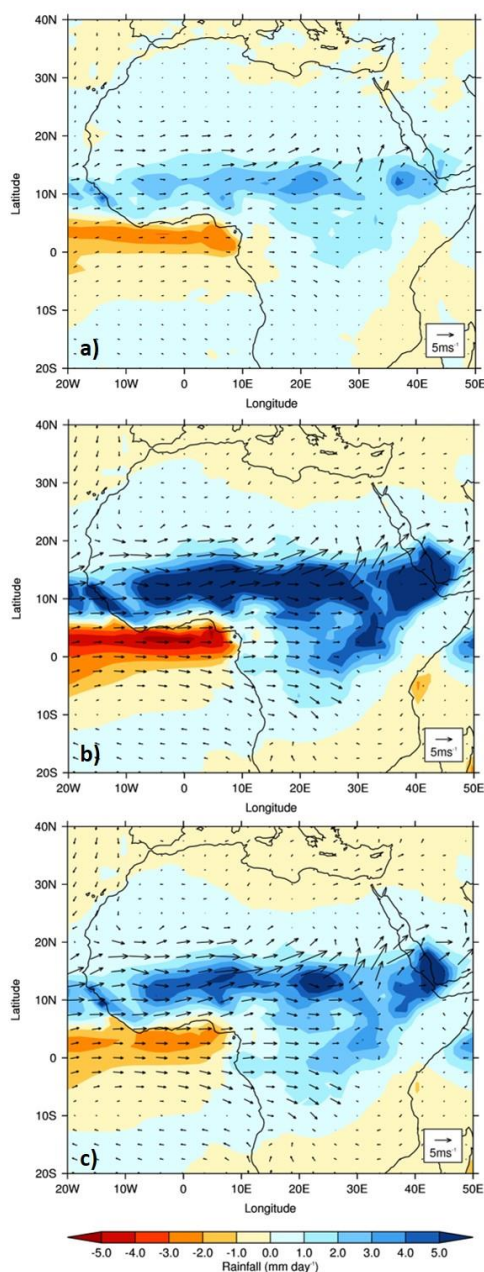
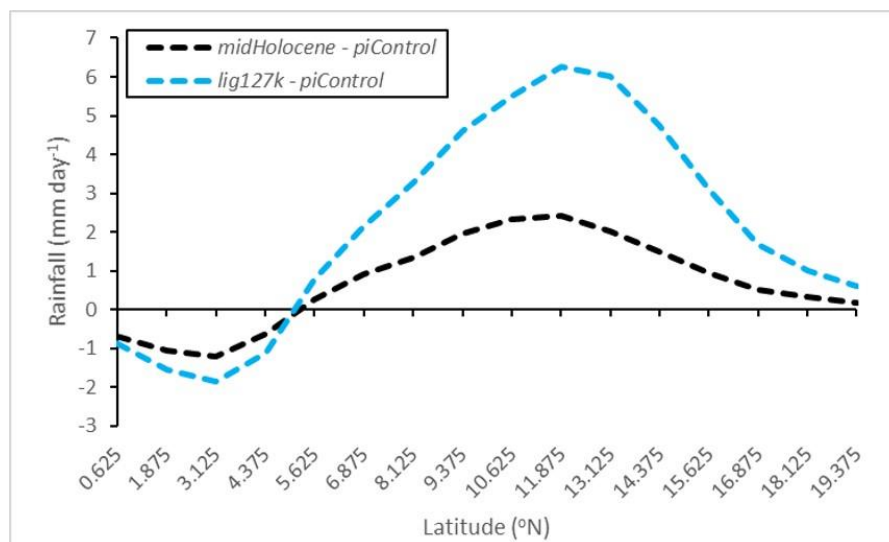


Figure 4 – 1.5 m air temperature climatology differences, HadGEM3 *midHolocene* and *lig127k* production runs versus HadGEM3 *piControl* production run, 100-year climatology from each: a) *midHolocene* – *piControl*, JJA; b) *midHolocene* – *piControl*, DJF; c) *lig127k* – *piControl*, JJA; d) *lig127k* – *piControl*, DJF



1006

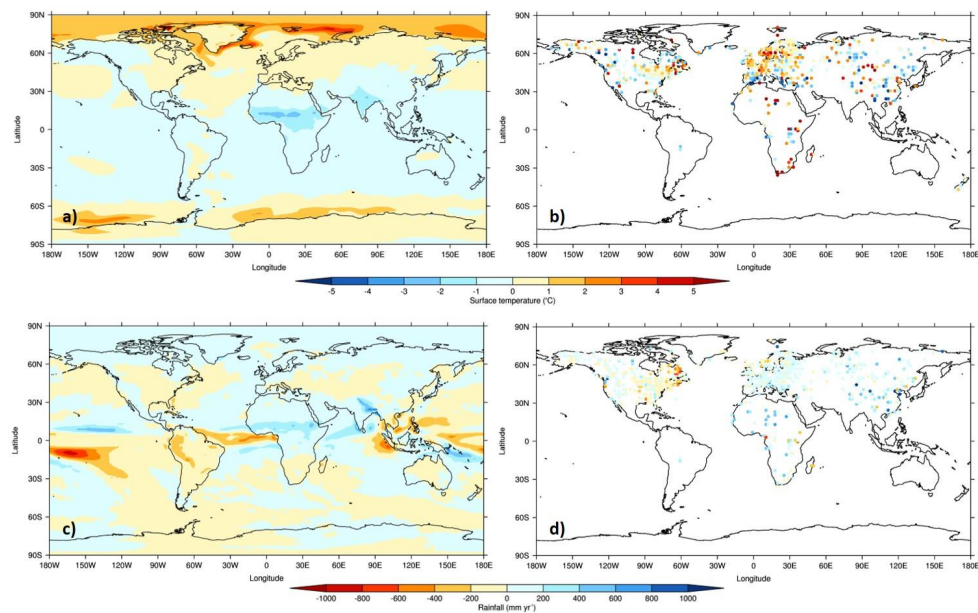
1007 Figure 5 – JJA rainfall & 850mb wind climatology differences, HadGEM3 *midHolocene* and *lig127k*
 1008 production runs versus HadGEM3 *piControl* production run, 100-year climatology from each: a)
 1009 *midHolocene* – *piControl*, JJA; b) *lig127k* – *piControl*; c) *lig127k* – *midHolocene*



1010
 1011 Figure 6 – JJA rainfall differences by latitude, averaged over West Africa (20°W-30°E, including both
 1012 land and ocean points), HadGEM3 *midHolocene* and *lig127k* production runs versus HadGEM3
 1013 *piControl* production run, 100-year climatology from each year
 1014

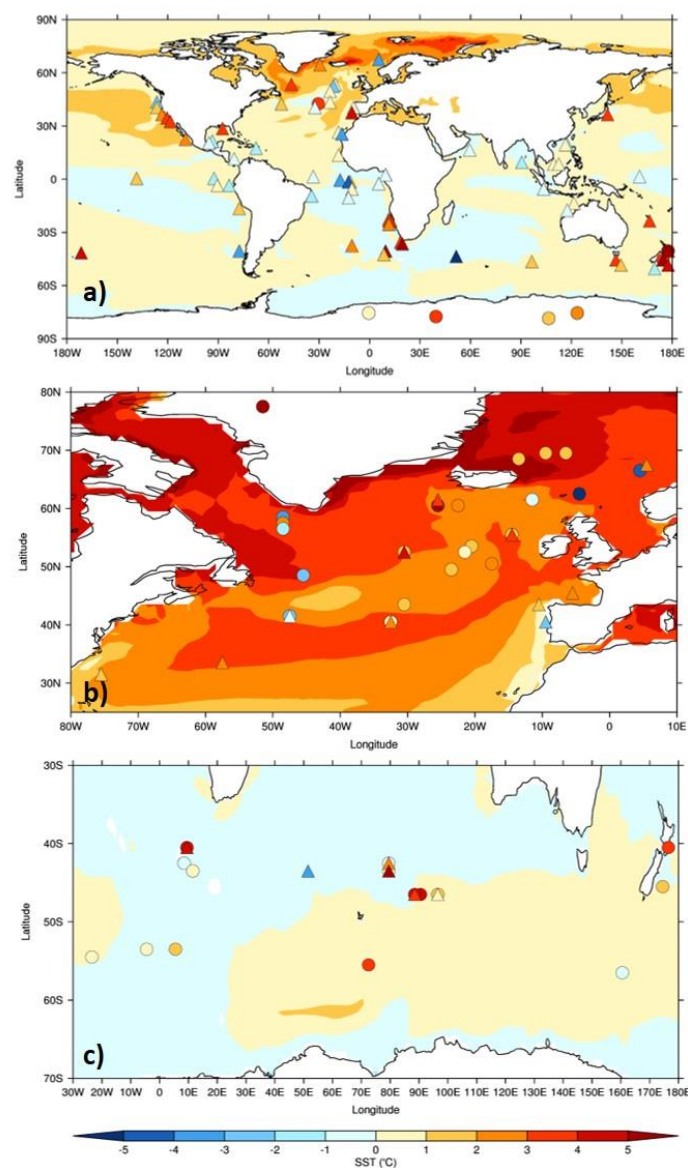


1015

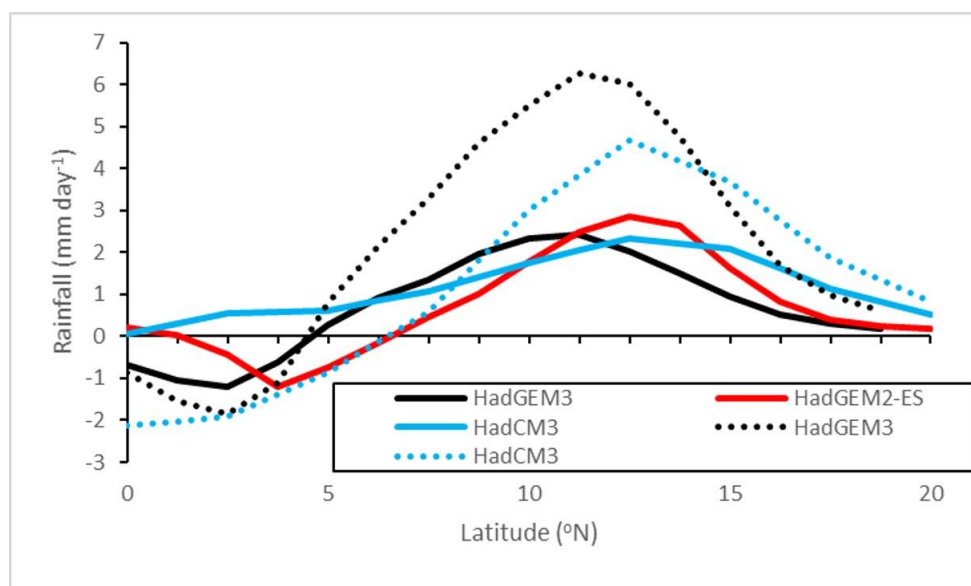


1016

1017 Figure 7 – Simulated versus proxy MAT and MAP anomalies. Left-hand side panels show simulated
1018 gridded anomalies from HadGEM3 (*midHolocene* production run – *piControl* production run, 100-
1019 year climatology from each), right-hand side panels show proxy data from Bartlein *et al.* (2011) (MH
1020 – preindustrial). Proxy data locations are projected onto model grid: a) Simulated MAT; b) Proxy
1021 MAT; c) Simulated MAP; d) Proxy MAP
1022



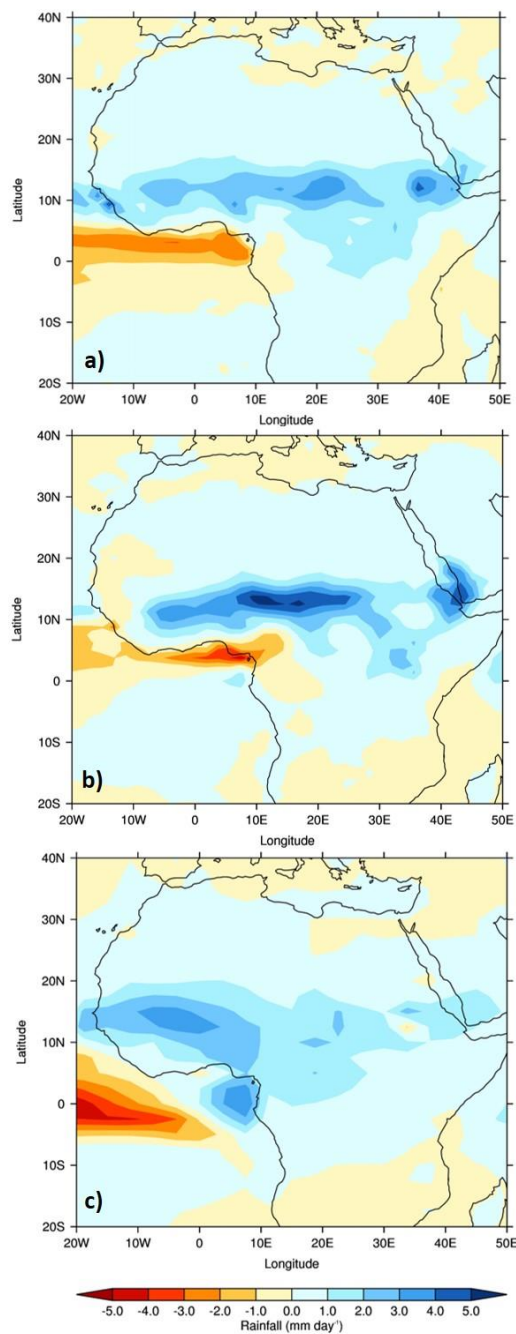
1023
 1024 Figure 8 – Simulated versus proxy SST anomalies. Background gridded data show simulated
 1025 anomalies (lig127k production run – piControl production run) from HadGEM3 (100-year
 1026 climatology), circles show proxy data (LIG – preindustrial) from Capron et al. (2017) and triangles
 1027 show proxy data (LIG – preindustrial) from Hoffman et al. (2017). Proxy data locations are projected
 1028 onto model grid: a) Annual data; b) Northern Hemisphere summer (JAS); c) Southern Hemisphere
 1029 summer (JFM). Note that proxy locations show SST over ocean and SAT over Greenland/Antarctica



1030

1031 Figure 9 – JJA daily rainfall climatology differences (MH and LIG-PI) by latitude, averaged over
 1032 West Africa (20°W-30°E, including both land and ocean points), for the various generations of the
 1033 UK's physical climate model, 100-year climatology from each (50-year climatology for HadCM3
 1034 LIG). Solid lines show MH simulations, dotted lines show LIG simulations. Note that due to the low
 1035 spatial resolution in HadCM3, values in between latitude points have been interpolated

1036

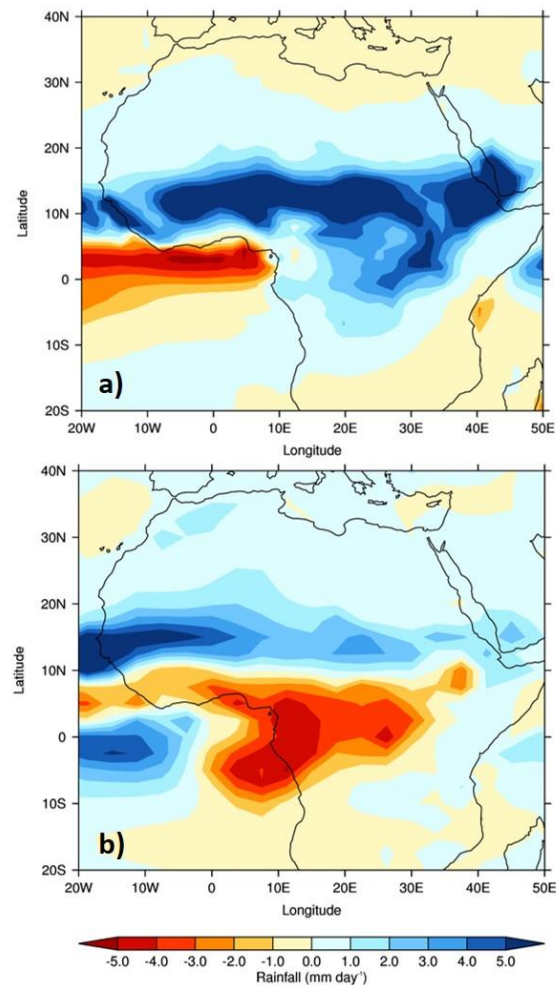


1037

1038 Figure 10 – JJA daily rainfall climatology differences (MH-PI) for the various generations of the
1039 UK's physical climate model, 100-year climatology from each: a) HadGEM3; b) HadGEM2-ES; c)
1040 HadCM3



1041



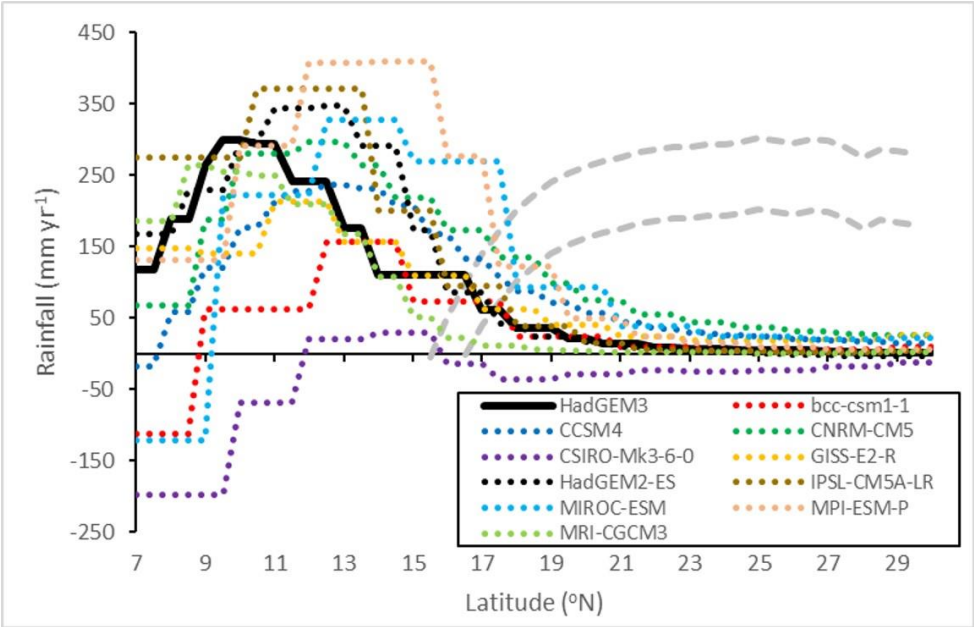
1042

1043 Figure 11 – JJA daily rainfall climatology differences (LIG-PI) for the various generations of the
1044 UK's physical climate model, 100-year climatology from HadGEM3, 50-year climatology from
1045 HadCM3: a) HadGEM3; b) HadCM3

1046



1047



1048

1049 Figure 12 – Annual mean rainfall over West Africa, zonally averaged from 20°W-30°E, HadGEM3
1050 and CMIP5 *midHolocene* production run minus corresponding *piControl* production runs, 100-year
1051 climatology. Solid line shows HadGEM3, dotted lines show CMIP5 simulations. Grey dashes show
1052 maximum and minimum bounds of the increase in rainfall required to support grassland at each
1053 latitude, within which simulations must lie if producing enough rainfall to support grassland

1054

1055