SUPPLEMENTARY MATERIAL

TEXT

Model description

Here, a description of the major changes in the latest version of UK's physical climate model, HadGEM3-GC3.1, relative to its predecessor, is given, following on directly from Section 2.1.2 in the main manuscript. Beginning with Global Atmosphere (GA7) and Global Land components (GL7), a once-in-a-decade replacement of the model's dynamical core, implementing ENDGame, was undertaken for the previous version (GA6) and therefore remains the same in GA7 (Walters *et al*. 2017)[1]. A number of other developments, since the previous version of the model, have also been included. Improvements were made to the radiation scheme to allow better treatment of gases absorption, as well as improvements to how warm rain and ice clouds are treated, an improvement to the numerics of the convection scheme, and improvements to the microphysics as well as an incremental development of ENDGame (Walters *et al*. 2017). Together these led to reductions in four model errors that were deemed critical in the previous configuration: i) South Asian monsoon rainfall biases over India; ii) biases in both temperature and humidity in the tropical tropopause; iii) shortcomings in the numerical conservation; and iv) biases in surface radiation fluxes over the Southern Ocean (Walters *et al*. 2017). In addition to these developments, two new parameterisation schemes were introduced in GA7: firstly the UK Chemistry and Aerosol (UKCA) GLOMAP-mode aerosol scheme, to improve the representation of tropospheric aerosols, and secondly a multi-layer snow scheme in the Joint UK Land Environment Simulator (JULES), to allow the first time inclusion of stochastic physics in UM climate simulations (Walters *et al*. 2017).

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

[1] Please see Reference section in main manuscript for citations

aforementioned reduced warm bias, the latter of which was deemed primarily due to the tuning of the different mixing (e.g. vertical and isopycnal) parameters (Storkey *et al*. 2018).

FIGURES

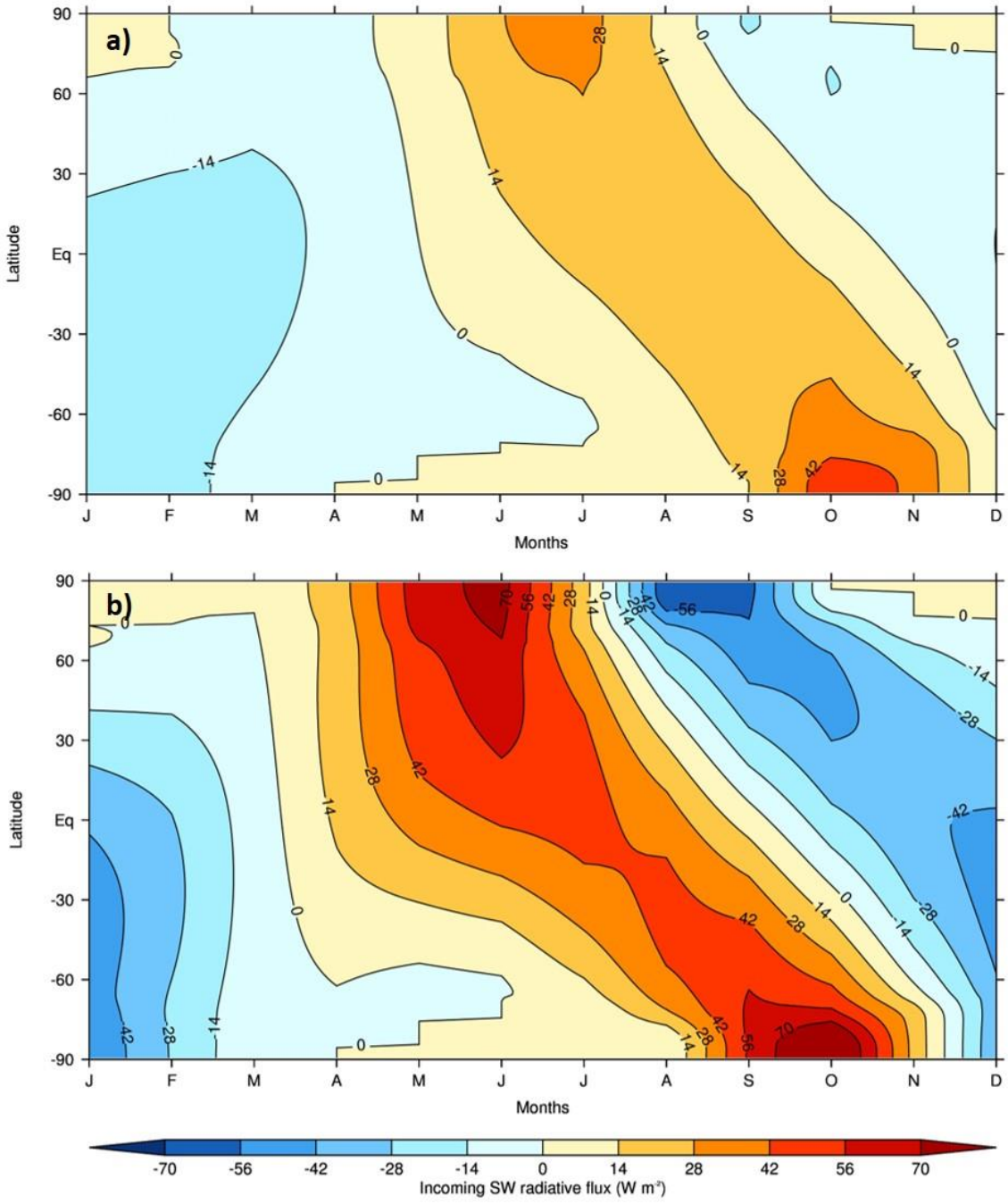

Figure S1 - Latitude-month insolation (incoming SW radiative flux) anomalies, using modern calendar: a) *midHolocene - piControl*; b) *lig127k – piControl*

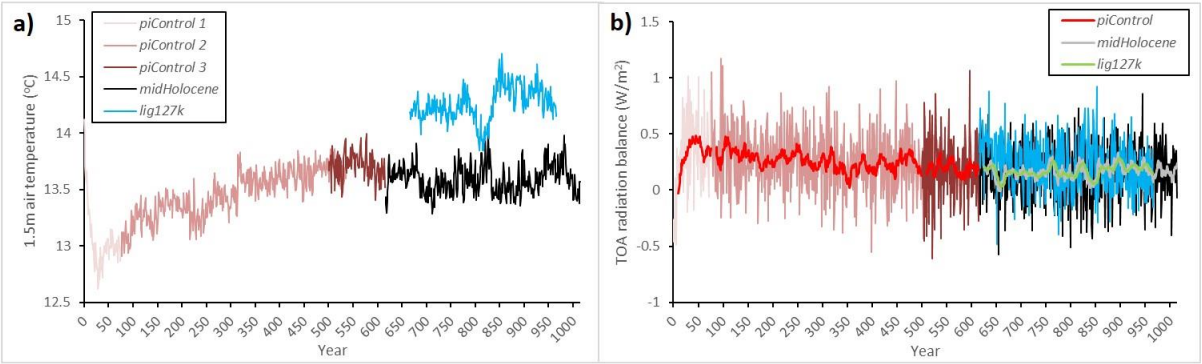

Figure S2 - Annual global mean atmospheric fields from HadGEM3 *piControl*, *midHolocene* and *lig127k* spin-up phases: a) 1.5 m air temperature; b) TOA radiation balance. Thin lines in b) show annual TOA radiation balance, thick lines show 11-year running mean. Note that the *piControl* spin-up phase was run in three separate parts, to accommodate for minor changes/updates in the model as the simulation progressed. Note also that the first ~50 years of the *lig127k* simulation have been deliberately removed from this figure, because a number of model crashes caused the model to be initially unstable and give highly varied global mean temperatures.

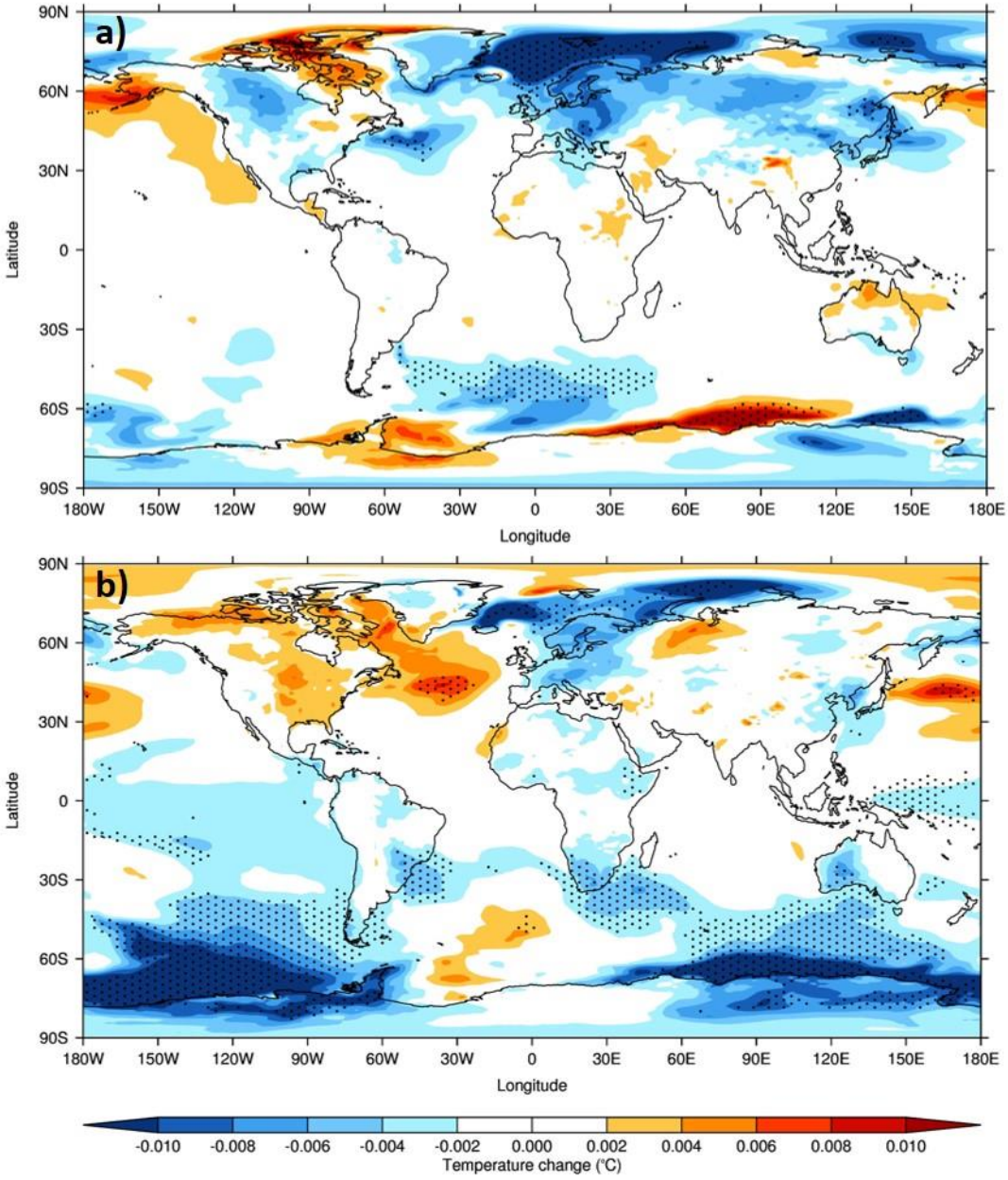

Figure S3 - Centennial trends in 1.5m temperature for HadGEM3 warm climate simulations' spin-up phases, last 100 years only: a) *midHolocene* ; b) *lig127k*. Stippling shows statistical significance (as calculated by a Mann-Kendall test) at the 99% level

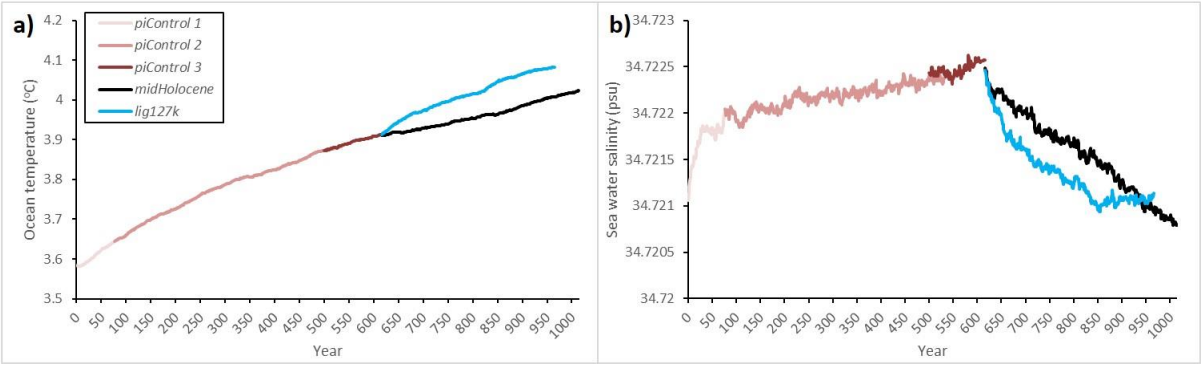

Figure S4 - Annual global mean (full depth) oceanic fields from HadGEM3 *piControl*, *midHolocene* and *lig127k* spin-up phases: a) OceTemp; b) OceSal

Figure S2

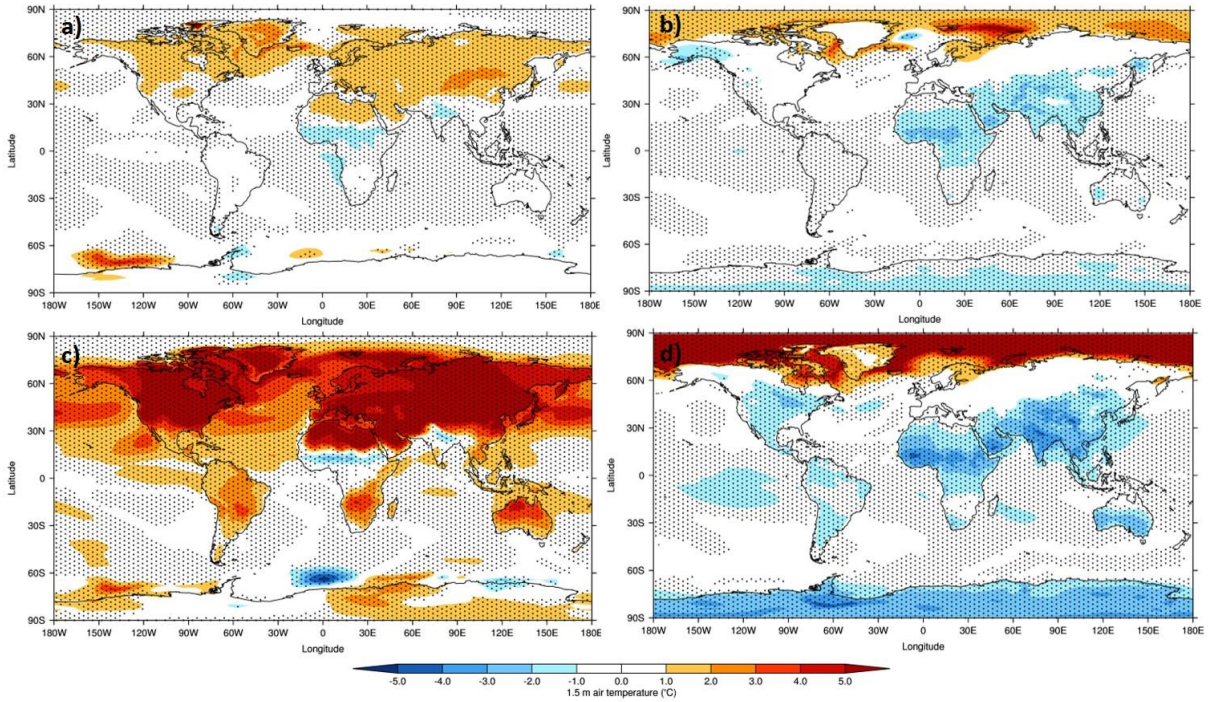

Figure S5 - Modern calendar 1.5 m air temperature climatology differences, HadGEM3 *midHolocene* and *lig127k* production runs versus HadGEM3 *piControl* production run: a) *midHolocene – piControl*, JJA; b) *midHolocene – piControl*, DJF; c) *lig127k – piControl*, JJA; d) *lig127k – piControl*, DJF. Stippling shows statistical significance (as calculated by a Student's T-test) at the 99% level

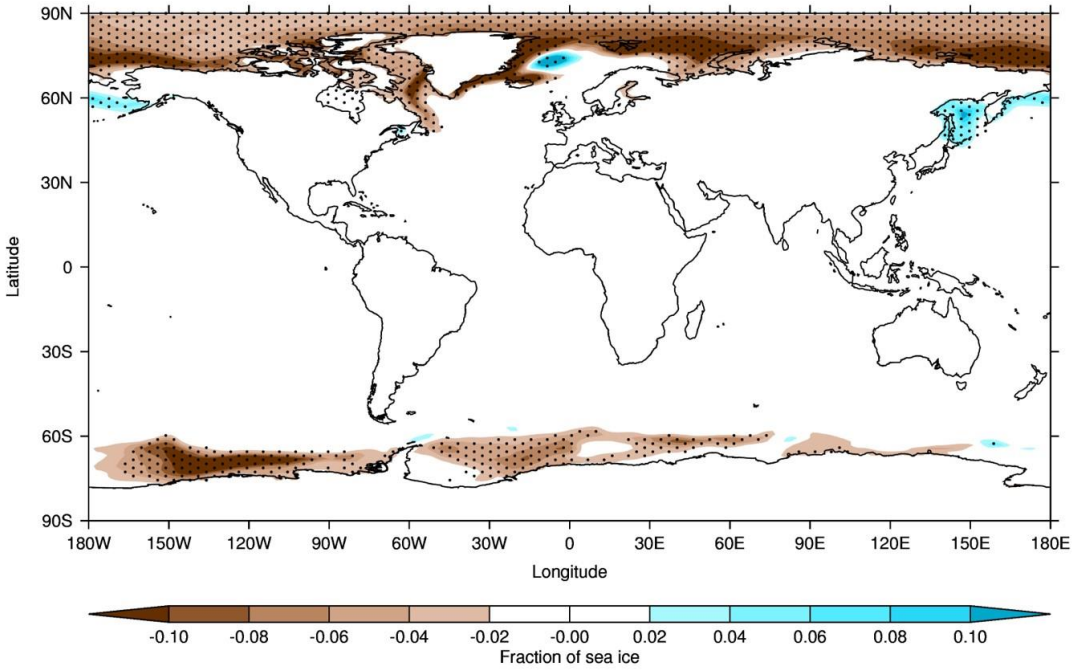

Figure S6 - Annual mean sea-ice climatology differences, HadGEM3 *midHolocene* production run versus HadGEM3 *piControl* production run. Stippling shows statistical significance (as calculated by a Student's T-test) at the 99% level

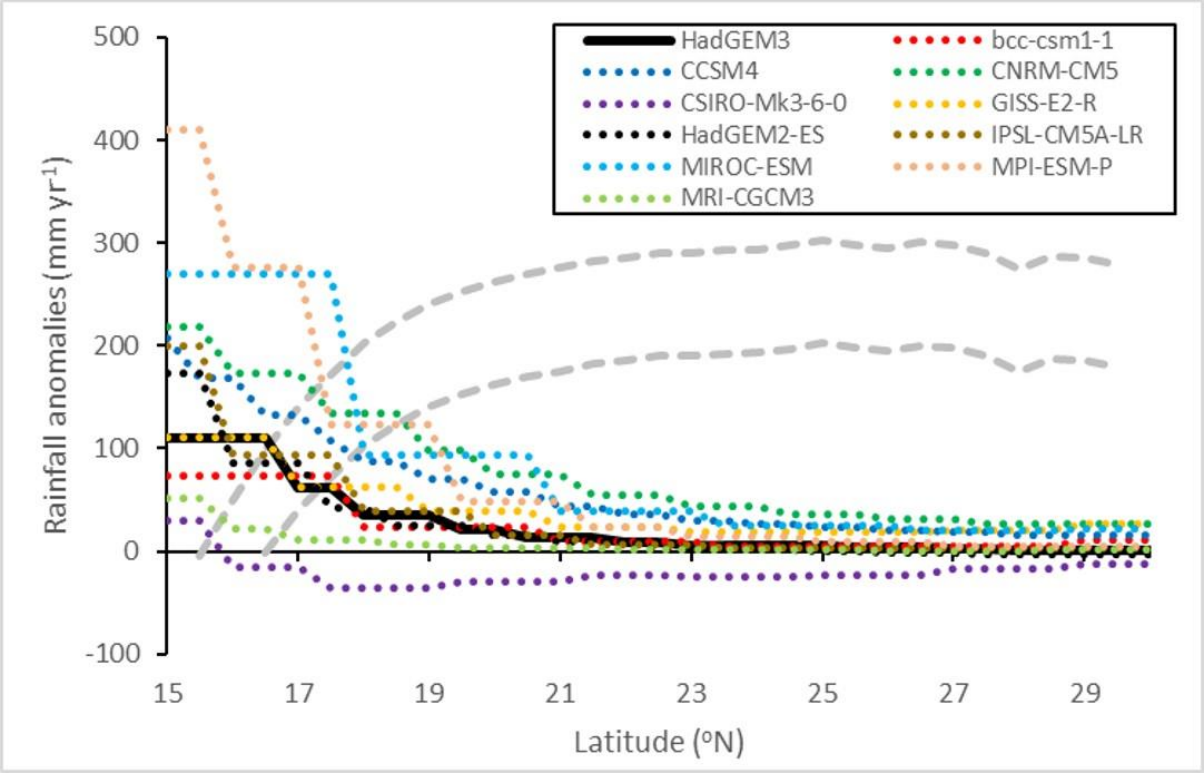

Figure S7 - Annual mean rainfall over West Africa (averaged over 20°W-30°E, consistent with Joussaume *et al*. [1999]), HadGEM3 *midHolocene* simulation minus corresponding *piControl*, and likewise for previous models from CMIP5. Solid line shows HadGEM3, dotted lines show CMIP5 simulations. Grey dashes show maximum and minimum bounds of the increase in rainfall required to support grassland at each latitude, within which simulations must lie if producing enough rainfall to support grassland (adapted from Figure 3a in Joussaume *et al*. [1999])