Paper:
Ramirez-Villegas, J and Challinor, AJ (2012) *Assessing relevant climate data for agricultural applications.* Agricultural and Forest Meteorology, 161. 26 – 45.

http://dx.doi.org/10.1016/j.agrformet.2012.03.015

Assessing relevant climate data for agricultural applications

Authors

Julian Ramirez-Villegas

International Center for Tropical Agriculture (CIAT), Cali, Colombia

CGIAR Research Program on Climate Change, Agriculture and Food Security (CCAFS)

Institute for Climate and Atmospheric Science (ICAS), School of Earth and Environment, University of Leeds, Leeds, UK

Andy Challinor

Institute for Climate and Atmospheric Science (ICAS) School of Earth and Environment, University of Leeds, Leeds, UK

CGIAR Research Program on Climate Change, Agriculture and Food Security (CCAFS)

Target journal

Agricultural and Forest Meteorology

Abstract

Climate change is expected to substantially reduce agricultural yields, as reported in the by the Intergovernmental Panel on Climate Change (IPCC). In Sub-Saharan Africa and (to a lesser extent) in South Asia, limited data availability and institutional networking constrain agricultural research and development. Here we performed a review of relevant aspects in relation to coupling agriculture-climate predictions, and a three-step analysis of the importance of climate data for agricultural impact assessment. First, using meta-data from the scientific literature we examined trends in the use of climate and weather data in agricultural research, and we found that despite agricultural researchers' preference for field-scale weather data (50.4% of cases in the assembled literature), large-scale datasets coupled with weather generators can be useful in the agricultural context. Using well-known interpolation techniques, we then assessed the sensitivities of the weather station network to the lack of data and found high sensitivities to data loss only over mountainous areas in Nepal and Ethiopia (random removal of data impacted precipitation estimates by ±1,300 mm/year and temperature estimates by ±3°C). Finally, we numerically compared IPCC Fourth Assessment Report climate models' representation of mean climates and interannual variability with different observational datasets. Climate models were found inadequate for field-scale agricultural studies in West Africa and South Asia, as their ability to represent mean climates and climate variability was limited: more than 50% of the country-model combinations showed <50% adjustment for annual mean rainfall (mean climates), and there were large rainfall biases in GCM outputs (1,000 to 2,500 mm/year), although this varied on a GCM basis (climate variability). Temperature biases were also large for certain areas (5-10°C in the Himalayas and Sahel). All this is expected to improve with IPCC's Fifth Assessment Report; hence, appropriate usage of even these new climate models is still required. This improved usage entails bias reduction (weighting of climate models or bias-correcting the climate change signals), the implementation of methods to match the spatial scales, and the quantification of uncertainties to the maximum extent possible.

Keywords: Sub-Saharan Africa; South Asia; climate modelling; climate model; skill; uncertainty; CMIP3; CMIP5.

1. Introduction

Agriculture is expected to play an important role in the context of climate change, not only because it is considered amongst the most vulnerable sectors, but also because it is part of the solution (i.e. potential to mitigate greenhouse gases [GHGs] emissions) (FAO, 2009; IPCC, 2007). Agriculture will likely be severely affected over the next hundred years due to unprecedented rates of changes in the climate system (IPCC, 2007; Jarvis et al., 2010; Lobell et al., 2008; Thornton et al., 2011). Some of these impacts have already been observed (Battisti and Naylor, 2009; Schlenker and Lobell, 2010). To help cope with such impacts, a framework to assess the effects of climate change on agriculture and food security and to aid with adaptation was established in 2008, as described by Jarvis et al. (2011): The Consultative Group of International Agricultural Research (CGIAR) Research Program on Climate Change, Agriculture and Food Security (CCAFS).

For adaptation to be successful, agricultural and climate data are crucial, and these are scarce in their basic forms (data from research and weather stations, respectively) or not very well managed and/or maintained in certain parts of the world. Most importantly, climate databases and their derived products are sometimes inaccurate, or else lack the documentation necessary to facilitate their use within the agricultural research community. In some instances, this may be indicative of the gap between the agricultural and climate research communities (Pielke et al., 2007; Thornton et al., 2011). Even when the two do collaborate, agricultural researchers face critical constraints when accessing basic sources of climate data (i.e. weather stations) due to a number of factors, from access to data, to weather maintenance and data quality checks, to the weather itself (DeGaetano, 2006).

In the last 10 years, various datasets have been developed by different institutions, usually based on either a combination of weather station data, satellite data, and numerical weather prediction models in addition to interpolation methods, or on the sole application of climate models. The usage of these datasets for agricultural modelling purposes is rather limited for one or more of the following reasons: (1) their time step is long (monthly in the best case); (2) their temporal coverage is limited to an average of several years (Hijmans et al., 2005; New et al., 2002); (3) their spatial resolution is too coarse (Adler et al., 2003; Schneider et al., 2010); (4) their geographic coverage is not wide enough (Di Luzio et al., 2008); and (5) only certain variables (i.e. temperatures, rainfall) are reported whereas other agriculturally relevant measures (e.g. potential and/or reference evapotranspiration, relative humidity, solar radiation) are rarely reported (Di Luzio et al., 2008; Hijmans et al., 2005). Moreover, assessments of these data (particularly climate models) have been done only under a climate-science perspective (Gleckler et al., 2008; Pierce et al., 2009), for a limited number of variables (Jun et al., 2008; Reifen and Toumi, 2009), or for a reduced realm (Walsh et al., 2008).

In this paper, we sought to improve the general knowledge on the available climate data for agricultural research using a three-step thorough analysis on fundamental aspects related to agricultural modelling. First, we perform a meta-analysis on the usage of various data sources for agricultural applications; second, we assess the quality and distribution of weather station records by exploring both the ability of these data to fill geographic information gaps by means of interpolation, and the sensitivities of the different regions to data loss; and finally, we assess the accuracy of climate model outputs against different observational datasets using various metrics reported in previous literature (Gleckler et al., 2008; Pierce et al., 2009). We finally analyse the main implications of our findings on agricultural impact assessment.

2. Review of knowledge and data

2.1. Understanding of processes and crop modelling

Mechanisms to fix carbon in plants (i.e. photosynthesis) are affected by a number of factors (El-Sharkawy, 2005; Prasad et al., 2002), although responses strongly depend on the type of mechanism used by the plant to produce biomass (i.e. C_4, C_3, CAM) and on any other stresses to which the plant could be subjected simultaneously. In crop production, apart from appropriate plant growth it is the amount of biomass accumulated in fruits and seeds and the nutrients in them that matters most (Thuzar et al., 2010). Yields are a direct consequence of photosynthesis and biomass accumulation, and these are directly or indirectly affected by environmental conditions [see (Challinor et al., 2009b) for a review]. Well-watered crops grown under optimal temperature and solar radiation ranges develop to their full production potential (van Ittersum et al., 2003), but growth potential reduces if the crop is stressed during the growing season (Hew et al., 1969; Huntingford et al., 2005).

Therefore, modelling crop growth depends on (1) correct formulation of the simulation model, (2) our ability to understand the effects of environmental factors on growth, and (3) correct measurement of the relevant environmental factors for correct mapping of their interactions (Boote et al., 1996; El-Sharkawy, 2005). Hence, crop modelling largely benefits from accurate measurements of temperatures, rainfall, and solar radiation, as the main factors acting on photosynthesis (Challinor and Wheeler, 2008; Hoogenboom et al., 1994), but even these basic data are often unavailable, messy, or of limited quality. The more available data there exists, the better calibration and evaluation of crop models can be (Adam et al., 2011; Niu et al., 2009; Xiong et al., 2008).

Additionally, most crop models simulate growth of individual plants and then scale out the modelling results to the plot-scale, based on management decisions such as plant and row distances, and plot size (Aggarwal et al., 2006; Boote et al., 1996; Hoogenboom et al., 1994). On the other hand, available weather data (when not measured in the field) is only available at coarse spatial scales. Matching these two spatial scales is not an easy task [see (Challinor et al., 2009a; Jagtap and Jones, 2002; Trnka et al., 2004) for a review]. The challenge is thus to increase the knowledge of the interactions between atmospheric and crop-growth processes (Boote et al., 1996) whilst avoiding model over-parameterisation (Challinor et al., 2009b), improving the accuracy of inputs (Adam et al., 2011), and matching both spatial scales (Challinor et al., 2009a). All this requires closing the gap between crop and climate scientists.

2.2. Weather data

Measurements of weather for a given site are often unavailable because (1) there is no weather station; (2) weather stations are not well maintained so data are either only available for a short period or contain gaps, (3) collected data are not properly stored; (4) data do not pass basic quality checks; and/or (5) access to data is restricted by holding institutions (Figure 1). This all further constrains agricultural impact assessment, highlighting the importance of making data public.

<Insert Figure 1 here>

Apart from the constraints related to access and weather station locations, probably the most important issue regarding weather data is quality (Begert et al., 2008; DeGaetano, 2006) (Figure 1), which also greatly affects the performance of impact models. Therefore, the climate and agricultural community has partly focused on developing methods for either temporal or spatial data gap filling, and on using such methods for developing global or regional datasets with public access (Hijmans et al., 2005; Jones and Thornton, 1999; Soltani et al., 2004).

However, uncertainties in global datasets derived from interpolation methods have been only barely (if at all) estimated (Buytaert et al., 2009; Challinor and Wheeler, 2008; Soria-Auza et al., 2010). Researchers using global datasets and any weather station source need to be aware of these problems and ought to take this into account by testing the sensitivities of their approaches to accuracy issues (i.e. inhomogeneities, discontinuities) and (if possible) providing results within the range of uncertainty in input data (i.e. such as the outputs of cross validated interpolation methods) (Challinor et al., 2005).

2.3. Climate model data

General Circulation Models (GCMs) are currently the best way to model the complex processes that occur at the earth system's level (Huntingford et al., 2005; IPCC, 2007). However, as CGMs are highly complex, they are computationally expensive, so they have only been used for predictions at coarse spatial scales. These predictions therefore involve a number of uncertainties relevant to agriculture [see (Challinor et al., 2009b; Jarvis et al., 2010; Quiggin, 2008) for reviews on the topic].

In short, uncertainty in climate modelling arises from the impossibility of modelling the climate system with complete determinism (Walker et al., 2003). This uncertainty can arise from: context (boundaries of the system modelled), model, inputs, and parameters (Walker et al., 2003). Model uncertainty can be structural or technical: structural uncertainty in models is associated with our lack of understanding of the system, whereas technical uncertainty relates to our inability to implement mathematical formulations in computational systems. Other uncertainties in climate modelling arise from variable driving forces (greenhouse gas emissions and concentrations), initial conditions and parameterised physics (Challinor et al., 2009b; Walker et al., 2003). Rationalisation and quantification of all these uncertainties under the context of agriculture is possible (see Challinor et al., 2009b for a review).

Crop modellers are thus challenged to understand the broad concepts of climate modelling uncertainties and detect the sensitivities of crop models to them, whilst also having a basic understanding of earth processes in order to identify major flaws in climate models and decide the best ways to couple them with crop models.

3. Materials and methods

Throughout this paper, we built upon existing knowledge of agricultural and climate modelling (Sect. 2) and:

1. Performed a meta-analysis on the usage of climate and weather data for agricultural modelling purposes and summarised the desirable characteristics sought when modelling crop production.
2. Analysed the robustness of the existing weather station network by assessing both the ability of these data to correctly fill information gaps via interpolation methods, and the network's sensitivities to information loss.
3. Assessed the accuracy of climate model outputs from the Fourth Assessment Report of the IPCC (IPCC, 2007) against different observational datasets, using metrics and methods reported in the climate-science literature that are also familiar to agricultural researchers.

All calculations were done by means of the software packages R-2.13.1 (available at http://www.r-project.org) and GRASS-GIS 6.4.0 (available at http://grass.fbk.edu) in a 64-bit Red Hat Enterprise Linux 5 box.

3.1. Study area

We focused on the geographic area of Africa and South Asia, where several studies have identified that significant vulnerabilities exist (Aggarwal, 2008; Aggarwal et al., 2004; Barrios et al., 2008; Byjesh et al., 2010; Challinor et al., 2007a; Chipanshi et al., 2003; Jones and Thornton, 2003; Lane and Jarvis, 2007; Liu et al., 2008; Lobell et al., 2008; Thornton et al., 2009; Thornton et al., 2011; Washington et al., 2006). In particular, we concentrate our efforts on West Africa (Senegal, Mali, Burkina Faso, Ghana and Niger), East Africa (Ethiopia, Tanzania, Uganda and Kenya) and the Indo-Gangetic Plains countries (India, Nepal, and Bangladesh), hereafter referred to as WAF, EAF and IGP, respectively (Figure 2).

<Insert Figure 2>

3.2. Analysing the usage of climate data in agricultural studies

3.2.1. Meta-data from agricultural studies

We gathered data from a number of publications on any topic that made use of climate data for any sort of agricultural modelling. We conducted searches using various search engines and downloaded only peer-reviewed publications. Review papers and the Fourth Assessment report of the IPCC were particularly useful in identifying additional published studies. We analysed all publications that in any way involved the usage of climate data for agricultural modelling purposes. As the selection of the impact assessment model is the first decision that any researcher needs to make, we focus on the driving factors of this decision. We recorded different variables from the studies as follows:

(1) Problem and/or topic in question: classified in categories such as impact assessment, seasonal yield forecasting, sole crop modelling, and climate attribution, among others. Each study was classified into only one category by taking into account only the main issue addressed by the paper;
(2) Scale of the approach: includes site, sub-national, country, regional (group of countries), and global;
(3) Use of weather generators: for both present and future, we recorded whether the study did or did not use a weather generator;
(4) Climate dataset (current): GCM when a GCM (regardless of which one) was used, RCM when an RCM (regardless of which one) was used, weather station, satellite (no further discrimination), and important datasets (i.e. CRU, WorldClim, GPCP, among others);
(5) Climate dataset (future): the nature of used future projections was recorded here including the downscaling method, if applicable. Classifications were: GCM "as is" when studies used raw GCM outputs as inputs, pattern scaled GCMs (Mitchell et al., 2004), RCMs, systematic changes to current climate data, statistical downscaling (Wilby et al., 2009), and weather generator downscaled GCM (Jones et al., 2009).

For further details on the above categories the reader is referred to our supplementary material (part 1). We revised a total of 205 peer-reviewed publications (See supplementary material part 2), printed between the years 1983 and 2011. Most of the studies were published immediately before or after the IPCC 4AR was released in 2007. When a certain study made use of two different sources of present-day climate data, it was considered twice (totalling 247 cases).

3.2.2. Analysing the usage of climate data in agricultural studies

We analysed the recent trends in the use of climate data for agriculture: the obvious constraints in the studies, the type of approaches used and the climate data inputs used to drive the chosen agricultural models. By doing this, we ensured that we covered all the main factors driving an agricultural researcher's decision to select a particular approach for a given problem.

3.3. Analysis of weather station data

3.3.1. Worldwide weather station network data

Long term climatological means of monthly precipitation and mean, maximum and minimum temperatures were assembled, as described by Hijmans et al. (2005). However, it is important to note that at the global level the sources of these data are large in number and differ in coverage, availability and quality (Table 1), and thorough quality checks were done only in a sub-set of the sources by original distributing institutions.

<Insert Table 1 here>

Additional sources such as R-Hydronet (http://www.r-hydronet.sr.unh.edu/english/) and Oldeman (1988) database for Madagascar were also included. We discarded any weather station with less than 10 years of data. The final dataset (after quality control and duplicates removal, see Hijmans et al. 2005 for more details) comprised 13,141 locations with monthly precipitation data, 3,744 locations with monthly mean temperature, and 2,684 locations with diurnal temperature range within our study region. This dataset is hereafter referred to as WCL-WS.

3.3.2. Analysing robustness of existing weather station networks

Many methods exist that allow the user to determine (interpolate) the value of a parameter (e.g., monthly rainfall) in a given condition (i.e. in a given site, at a given time, or both), where it had never been measured before. Some of these methods are already popular with researchers using climate data (Hijmans et al., 2005; Hutchinson, 1995; Jones and Thornton, 1999; New et al., 2002) either on a regional or on a global basis. For climate-variable interpolations, the robustness of weather records is critical for an accurate result.

We assessed the robustness of the weather station network by testing both the ability of weather records to yield accurate interpolation results, and the sensitivities of the network to information loss. Towards those ends, we used the WCL-WS dataset to fit a thin plate spline interpolation algorithm (Hutchinson, 1995) for our study region. We investigated the effect of weather station availability by using 100 cross validated folds for four variables (monthly maximum, minimum and mean temperatures and total precipitation) using similar methods as in Hijmans et al. (2005) and New et al. (2002) for each fold. We used longitude, latitude and elevation as independent variables. We used 85% randomly selected data points for fitting the splines and the remaining 15% for evaluating the result for each variable and month. For the evaluation, we calculated the R^2 and the Root Mean Square Error (RMSE) and produced boxplots of the 100-fold-by-12-month interpolations for each of the four variables. As the number of stations considerably exceeded the amount of available memory for processing, we divided the whole region of study in 5 tiles, each with an equivalent number of locations. We then projected the fitted splines onto 30-arc-second gridded datasets of latitude, longitude and altitude (Jarvis et al., 2008), thus producing a total of 4,800 interpolated surfaces (12 months times 4 variables times 100 folds). Finally, we analysed the spatial variability of standard deviations and the performance of the interpolation technique as proxies for sufficient distribution and geographic density of weather stations.

3.4. Assessment of IPCC Fourth Assessment Report (4AR) model data

3.4.1. Long-term observed mean climatology from weather stations

Three different long term climatology datasets were assembled: (1) the Global Historical Climatology Network (GHCN, as in Sect. 3.3.1) version 2 (Peterson and Vose, 1997), available at http://www.ncdc.noaa.gov/pub/data/ghcn/v2. We used GHCN as an independent source because it is a global resource that contributed significantly to WCL-WS and also because it is available at more temporally disaggregated levels (i.e. monthly), thus allowing uniformity with analyses on Sect. 3.4.3 and 3.4.6. This database includes monthly historical totals (1900-2010) of precipitation (20,590 stations), and means of maximum, minimum (4,966) and mean (7,280) temperatures. GHCN data have been subject to quality checks and to a process of "homogenisation" or "adjustment" (Peterson and Easterling, 1994); however, the available data within our analysis domain consisted primarily of “unadjusted” stations. For each location (6,393 stations for rainfall, 1,278 for mean temperature and 549 for minimum and maximum temperature) within our study area, we averaged historical monthly time series for the period 1961-1990 for maximum, minimum and mean temperatures and total rainfall, resulting in a time-averaged dataset of 6,393 locations for rainfall, 1,278 for mean temperature and 549 for minimum and maximum temperature. This dataset will be hereafter referred to as GHCN-CL.

(2) WCL-WS (Sect. 3.3.1); and (3) the Global Surface Summary of the Day (GSOD) was accessed at http://www.ncdc.noaa.gov/cgi-bin/res40.pl. This database contains daily data from ~9,000 weather stations worldwide for 18 variables, including, mean, maximum, minimum and dew point temperature, sea level and location pressure, visibility, wind speed and gust, precipitation, snow depth, and specifications on the occurrence of rain, snow, fog, tornado, thunder, or hail (NOAA, 2011; ftp://ftp.ncdc.noaa.gov/pub/data/gsod/readme.txt). We selected weather stations within our study area (1,999); aggregated daily rainfall, mean, maximum and minimum temperatures to a monthly time scale; and then averaged over the period 1961-1990. This dataset will be hereafter referred to as GSOD-CL.

3.4.2. Long-term observed mean climatology from interpolated surfaces

We gathered high-resolution climatology from two different sources: (1) the high resolution climate surfaces in WorldClim (Hijmans et al., 2005), available at http://www.worldclim.org. WorldClim is a 30 arc-seconds (~1km at the equator) global dataset produced from the interpolation of long-term climatology as measured in weather stations. Global gridded data were downloaded at the 30 arc-second resolution, then masked to our analysis domain, and aggregated to 10 arc-minute using bilinear interpolation in order to reduce computational and storage time; and (2) the University of East Anglia Climatic Research Unit (CRU) dataset (New et al., 2002), available through http://www.cru.uea.ac.uk/cru/data/hrg/ (CRU-CL-2.0). This dataset was developed using the same interpolation method as WorldClim, with the main difference that WorldClim includes many more weather stations, sometimes at the expense of input data quality. CRU-CL-2.0 resolution is 10 arc-minute (~20km at the equator). Data were downloaded at the global level and masked to our analysis domain. WorldClim and CRU-CL-2.0 are hereafter referred to as WCL-IS and CRU-IS (interpolated surfaces), respectively. We used these sources because (1) they are flag products that most researchers use for impact studies; (2) they are much higher resolution than GCMs (and other products such as the Global Precipitation Climatology Project [GPCP] and the Global Precipitation Climatology Centre [GPCC]) and hence permit the capture of small-scale weather patterns (important to agriculture) as well as a direct comparison of their within-GCM-gridcell mean with the actual GCM value; (3) are based only on ground observations of weather and do not incorporate side-products such as reanalysis (Uppala et al., 2005) or satellite data (Huffman et al., 2007), both of whose accuracy is not as good.

3.4.3. Long-term observed time series

Two sources of weather time series were used: (1) long term (1961-1990) series of monthly weather conditions were gathered from GHCN version 2 (Peterson and Vose, 1997). Again, we used mainly unadjusted stations. Mean monthly temperature and total monthly historical rainfall data were used without any further processing; and (2) long-term (1961-1990) series of daily weather as in GSOD (NCDC, 2011). For GSOD, daily precipitation and monthly temperature were aggregated to the monthly level only if all days were reported with data (for rainfall) and if at least 50% of the days had data (for temperatures). This resulted in 1,999 stations within our analysis domain, although not all stations had data for all months and all years. These two data sources are hereafter referred to as GHCN-TS and GSOD-TS, respectively. Lack of data prevented us from including maximum and minimum temperatures in the GHCN-TS and the GSOD-TS datasets. In contrast to GHCN-CL and GSOD-CL, GHCN-TS and GSOD-TS include every month and every year, thus allowing the analysis of inter-annual variability.

3.4.4. Global climate model output

The latest IPCC report (Fourth Assessment Report, 4AR) comprises the sole state-of-the-art public and official source of climate data for use in impact studies (IPCC, 2007; Jarvis et al., 2010). We therefore decided to use IPCC 4AR results.

We downloaded present day (1961-1990) simulations of global climate at original GCM resolution (~100 km) from the CMIP3 (Coupled Model Intercomparison Project phase 3) web data portal at https://esg.llnl.gov:8443/index.jsp (PCMDI, 2007). We downloaded monthly time series of mean, maximum, minimum temperature and precipitation flux in NetCDF format for 24 coupled GCMs (Table 2). Separately for each GCM, we calculated diurnal temperature range for each month and year as the difference between maximum and minimum temperatures and calculated total monthly rainfall as the product between the precipitation rate, the water density at sea level pressure and the number of seconds in a month. We used the each climate model monthly time series (GCM-TS hereafter) and also calculated average 1961-1990 climatology by averaging, for each variable (mean temperature, diurnal temperature range and total rainfall), every month for the whole 1961-1990 period (GCM-CL hereafter). The final datasets (i.e. GCM-TS and GCM-CL, respectively) consisted of three variables (mean temperature, diurnal temperature range and total monthly rainfall) for 24 different GCMs.

<Insert Table 2 here>

3.4.5. Ability to represent long-term climatology

The extent to which GCM predictions are accurate has not been fully explored for some parts of the world, particularly in the context of agriculture (Gleckler et al., 2008; Pierce et al., 2009; Walsh et al., 2008). As previously stated (Sect. 2.1), we compared the most readily available variables from both ground observations and climate models: rainfall, mean temperature and diurnal temperature range. Data for other variables are not available for our study regions in observational datasets. As per our stated objective (Sect. 3), we performed two sets of comparisons:

- First, we compared the GCM-CL dataset with the interpolated climatology in CRU-IS, WCL-IS (Sect. 3.4.2). We performed comparisons on a country basis in order to yield country-specific results. For each GCM gridcell, the mean, maximum and minimum values of all lower scale (CRU-IS, WCL-IS) cells was first calculated and then compared to the GCM value through the determination coefficient (R^2) and corresponding p-value, the slope of a origin-forced (so that a 1:1 relationship was sought) regression curve (S) and the root mean square error (RMSE).

- Second, using the same procedure, we compared the GCM-CL dataset with observed climatology in WCL-WS (Sect. 3.3.1), GHCN-CL and GSOD-CL (Sect. 3.4.1).

We analysed total rainfall, mean temperatures and diurnal temperature ranges over three periods: December-January-February (DJF), June-July-August (JJA) and the whole year (ANN). These months represent the most critical seasons for agriculture in our study regions, and are also the most often assessed in the existing literature (Gleckler et al., 2008; Pierce et al., 2009). Due to space constraints, we present only the results of comparisons between GCM gridcell values and mean values within gridcells, unless otherwise stated. We do, however, discuss other relevant results in more general terms.

3.4.6. Ability to represent long-term monthly climate time series

CMIP3-related GCMs are known to misrepresent certain inter-annual and/or within-decade variations that are important for agricultural systems (Govindan et al., 2002). However, specific aspects of these errors have not been explored in all CMIP3 models in the context of agriculture. Therefore, in order to test the consistency of GCM predictions across time, we compared the GCM-TS (Sect. 3.4.4) dataset against the GHCN-TS and GSOD-TS (Sect. 3.4.3). The comparison was done for three periods (JJA, DJF and ANN, Sect. 3.4.4) by calculating the R^2 and corresponding p-value, the slope of the regression curve as forced to the origin and the RMSE between the two time series (GCM-TS vs. GHCN-TS and GCM-TS vs. GSOD-TS). As a GCM cell contains one or more weather stations, we averaged the monthly time series as needed before comparing the two pairs of series. Finally, we compared the performance of all GCMs across the geographic space of our study area.

4. Results

4.1. Usage of climate data in agricultural studies

4.1.1. Topics of study

The most addressed topic (41.4% of the studies) in our literature review was climate change impact assessment (Figure 3), followed by crop growth simulation (18.5%). Water resources-impact studies round out the top three topics studied (8.1%), followed by climate attribution (6.9%), crop yield forecasting (6.1%), and model assessment (5.7%). Surprisingly, formal studies addressing adaptation were rather scarce (3.6%). Pests and diseases, soils, abiotic stresses and climate risks appeared to be a lot less addressed than impact assessment and crop growth simulation studies, which together accounted for more than 50% of the total publications.

<Insert Figure 3 here>

4.1.2. Scale of studies and type of models

Most of the studies performed their models at a scale less than the size of a country; site-specific or sub-national level together comprised 55% of the studies. Very few (7%) of the studies were performed at the global level, likely because of the type of models used: field-scale mechanistic crop growth models were the most utilised overall (69.2%); followed by statistical and/or empirical approaches (S/E, 21.4%), which most of the crop growth modellers criticise for not being accurate enough (Lobell and Burke, 2010; Lobell et al., 2008); and finally by hydrological models (10%). The frequent use of field-based crop growth models suggests that the time step requirement for input data is rather high (El-Sharkawy, 2005), also confirmed by the usage of weather generators (8.5 and 11.2% for present and future climates, respectively).

4.1.3. Climate data sources

Unlike the model types, which were quite similar, the sources of present climate data varied substantially, with a total of 32 different sources being used for present climate data (Figure 4A). On average, a different present-day-climate dataset was used for every 7 agricultural studies. The most commonly used data source was local (non-public) weather stations (50.4% of the cases), followed by University of East Anglia Climatic Research Unit (CRU) datasets with 13.7% (10.9% for CRU-TS [monthly time series], and 2.8% for CRU-CL [monthly climatology]). Climate model outputs were used in 14.5% of the cases: within this group, 10.5% used GCM data, 4% RCM [Regional Climate Model] data, 3.6% satellite imagery, and 2.8% WorldClim, followed by other less relevant sources. The Global Precipitation Climatology Project (GPCP) (Adler et al., 2003; Huffman et al., 2009), the Global Precipitation Climatology Centre (GPCC) (Schneider et al., 2010) and the Global Historical Climatology Network (GHCN, (Peterson and Vose, 1997)) were rarely reported overall (0.4% each).

<Insert Figure 4 here>

The future climate data used was found to be less variable overall, with only 7 different types of data employed in the 125 cases citing some type of future climate data (Figure 4B). Out of these 125, only one study did not clearly state which type of climate data was used. The vast majority of cases (42. 9%) used GCM data “as is” (AI GCM), meaning that predictions on agricultural yields were based on predicted changes at coarse resolution (~100 km). All other studies involved some type of downscaling, except those that employed the systematic changes approach (SC variables), which can be assumed to be sensitivity analyses rather than impact studies. RCMs (Regional Climate Models) were the most common way of downscaling GCMs, cited in 19% of the studies, followed by statistical downscaling with 17.5% (SD GCM, (Tabor and Williams, 2010)), and pattern scaling with 8.7% (PS GCM, (Mitchell et al., 2004)) (Figure 4B).

Uncertainty, as measured by the number of different future scenarios used (combinations of emissions scenarios and climate models) was explored in only 36.5% of the studies. Additionally, the average number of scenarios per study (rounded to the closest integer) was 3, indicating that climate uncertainties are barely (if at all) studied in agricultural science and highlighting a knowledge gap in agricultural research, an issue previously raised and discussed by other authors (Challinor et al., 2009b; Challinor and Wheeler, 2008), although some studies addressing this aspect are underway (C. Rosenzweig, personal communication).

4.2. Robustness of existing weather station networks

The sensitivities of the network to information loss were found overall to be low. Nevertheless, certain areas, variables and months were found highly sensitive. Agricultural lands (Ramankutty et al., 2008), as visually inspected, are in general less sensitive to data loss than non-agricultural lands. Interpolations' performance varied depending upon the variable, month and parameter used to evaluate them (i.e. R^2, RMSE, and S), but were consistent, statistically significant ($p<0.0001$) and with variability (of R^2, RMSE, and S) between 10–15% in the worst cases. Rainfall presented the lowest R^2 values (Figure 5), particularly in the months of April to August, during which there was a higher variability in the R^2 value and the values reached the absolute minima (0.8). Although it is possible that a high number of weather stations per unit area can improve accuracy, it does not seem to happen in all variables, areas and/or months.

<Insert Figure 5 here>

The DJF period presented significantly lower variability and more predictive power, probably due to overall low climate variability (Cooper et al., 2008). Interestingly, maximum and minimum temperatures showed different interpolation accuracies, even though they were measured in the same places. Maximum RMSE for temperatures was up to 1.7°C, whilst for precipitation it was up to 100 mm/year, as seen in the evaluation data. The effect of geography and the difficulty of fitting unique and complex landscape features cause errors, leading to high standard deviations in some areas (Figure 6). In the highlands of Eastern Africa, particularly in the states of Benshangul-Gumaz, Addis Ababa and Southern Nations in Ethiopia, the central areas of the Eastern and Coast States in Kenya, and the very centre of Tanzania (i.e. regions of Morogoro, Dodoma and Manyara) between-fold variability was found to be high (above 150 mm/year).

<Insert Figure 6 here>

Over IGP, the largest variability was found in the coastal areas of Maharashtra, Karnataka and Kerala in India, where rainfall deviation was up to 600 mm/year, and in Nepal (districts of Gorka, Dhawalagiri, and Lumbini), where rainfall variability can go up to 1,000 mm/year, and temperature uncertainties up to 3°C, probably due to the combined effect of a more complex climate in the Himalayas and low weather station density.

4.3. Accuracy of climate model outputs

4.3.1. Ability to represent mean climate

As expected, the climate models' skill varied on a variable, country and region basis, with certain identifiable patterns (Figure 7, 8). The GCMs represent the observed climatology from weather stations (i.e. WCL-WS, GHCN-CL and GSOD-CL) more poorly than they do interpolated climatology (i.e. WCL-IS, CRU-IS), mainly because GCMs do not account for local-scale variability (Boo et al., 2011). In a broad sense, we found that the more complex the topography, the lower the skill of the GCMs (Gallée et al., 2004; Joubert et al., 1999). We also observed that GCM skill decreased according to the complexity of the variable, with the maximum skill displayed for mean temperatures, followed by temperature range and finally by precipitation. These results agree with those of other studies (Gleckler et al., 2008; Masson and Knutti, 2011; Pierce et al., 2009).

Annual precipitation fit in IGP and WAF was observed to dip as low as 0 in some cases, with a considerable number of cases (23% for WCL-WS, 27% for GHCN-CL and 63% for GSOD-CL) presenting very low adjustment ($R^2 < 0.5$) (Figure 7). In Mali, Niger, India and Bangladesh, model skill in representing precipitation, compared to weather station measurements, was consistently low, an issue also reported in other studies (Douglass et al., 2008; Gleckler et al., 2008; Reichler and Kim, 2008). The Bergen Climate Model (BCCR-BCM2.0) and the INM-CM3.0 model showed very poor performance (R^2<0.5) in more than 25% of the countries when compared with WCL-WS, GHCN-CL and GSOD-CL, while the climate model GISS-ModelE (Hansen et al., 2007) presented the poorest performance.

<Insert Figure 7 here>

When compared with interpolated climatology (i.e. WCL-IS, CRU-IS), annual precipitation R^2 values varied from 0.383 (GISS-ModelE-R in Uganda) to 0.998 (IAP-FGOALS1.0-G in Burkina Faso), whilst for mean temperatures the R^2 varied from 0.195 (GISS-ModelE-R in Nepal) to 0.999 (MIUB-ECHO-G in Burkina Faso), and for temperature range the values were observed between 0.386 (CCCMA-CGCM3.1-T47 in Senegal) to 0.9998 (MPI-ECHAM5 in Burkina Faso) (Figure 7).

<Insert Figure 8 here>

In Ethiopia, mean temperature correlations were lower compared to other countries, despite the relative high density of stations in that area (data not shown). In Senegal, diurnal temperature range was found to be very poorly fitted, particularly for the CCCMA models (Figure 8). This result contrasts with that of other studies, which have marked the CCCMA models as the most skilled (Gleckler et al., 2008; Jun et al., 2008). The ability of GCMs to represent mean climate patterns over a year was neither uniform nor consistent (Table 3), with the lowest performance being observed for precipitation in the DJF period (large number of cases with R^2<0.5, and few cases with R^2>0.8). Performance for temperature range showed almost no cases with R^2<0.5, but fewer cases with R^2>0.8 than for mean temperatures (Table 3).

<Insert Table 3 here>

4.3.2. Ability to represent interannual variability

R square values were above 0.8 in a large number of gridcells (>50%) for all GCMs for both variables (rainfall, mean temperature) (data not shown); however, there were large rainfall biases in GCM outputs (Figure 9, 10), in some cases between 1,000 and 2,500 mm/year, depending on the GCM. These areas were located in Nepal, northern India and EAF. Most of the models' biases were wet-biases (Figure 10) which were found throughout the whole analysis domain, but they were particularly strong over IGP in the models CCCMA-CGCM3.1-T47, CSIRO-Mk3.0 and –Mk3.5, GFDL-CM2.0, all NASA-GISS models, and both UKMO-HadCM3 and –HadGEM1, whereas the opposite signal was observed over the same area for the models MIROC3.2.-HIRES, NCAR-CCSM3.0, INGV-ECHAM4, CNRM-CM3, and GFDL-CM2.1. Over WAF and EAF, almost all GCMs showed a dry-bias, with underestimations of up to 250 mm/year in some cases. Responses varied for seasonal means and totals, with the wet-season (JJA) being more sensitive to wet biases in most GCMs.

Temperature biases were also large for certain areas. In some cases, annual mean temperature biases were greater than 5°C and were observed to go up to 10°C, particularly in the Sahel and in the areas surrounding the Himalayas and the Tibetan Plateau in Nepal (Figure 11). The most evident temperature biases were found in the NASA-GISS models (GISS-AOM, GISS-ModelE-H and GISS-ModelE-R), and in INM-CM3.0, probably due to their coarse resolution. The quality of higher resolution models was in general better, but geographic trends were difficult to identify, as the locations with mean temperature were scant (7,280 locations for the whole study area). The smallest biases were observed in WAF, northern EAF and central India, where temperature biases were below 1.5°C, particularly for the models BCCR-BCM2.0, UKMO-HadCM3, NCAR-PCM1, CCCMA-CGCM3.1-T47 and MIUB-ECHO-G, some of which have been reported to perform well in tropical areas before (Gleckler et al., 2008; Jun et al., 2008). These biases were mostly concentrated in lowlands and were mostly warm-biases, except for UKMO-HadCM3 (Figure 12). Cold-biased models were usually the GISS-NASA models, MIROC3.2-MEDRES, UKMO-HadCM3, IPSL-CM4, MRI-CGCM2.3.2A and IAP-FGOALS1.0-G both for seasons (i.e. JJA, DJF, maps not shown) and for the annual mean (Figure 11, 12).

5. Discussion

5.1. Climate data and agricultural research

Although climate model data ("as is") are often preferred for impact studies, crop modellers and agricultural scientists should be cautious when developing future adaptation strategies based on crop models applied using future predictions of different (and sometimes unknown) nature (Jarvis et al., 2011), given the large uncertainties regarding the agricultural system and plant responses, the underlying uncertainty related to parameterised processes, and the differences in scales, all of which are reported in the impact-assessment literature [e.g. (Challinor and Wheeler, 2008)]. This, however, does not necessarily imply that climate model data cannot or should not be used, but rather means that an adequate treatment of biases needs to be done before climate and crop models can be properly used together (Challinor et al., 2010; Osborne et al., 2007).

Our findings demonstrate that, for regional assessments where large area process-based crop models, statistical, or empirical models are to be used, products such as WorldClim (Jones and Thornton, 2003; Thornton et al., 2009) and CRU (Challinor et al., 2004) coupled with weather generation routines appear to be the best-bet approach (Challinor et al., 2004; Jones and Thornton, 2003), although climate model data can also be used with proper bias treatment (Challinor et al., 2010; Osborne et al., 2007). However, if studies are to be carried out on a site-specific scale (Parry et al., 2005), weather station data is the best means by which to calibrate the modelling approaches. While partnerships are constantly being built and this allows researchers to share data, currently global weather station data such as GSOD and GHCN seem to be good options in cases when no other data is available, particularly when coupled with satellite data or other (country specific) historical weather records (Álvarez-Villa et al., 2010).

Agricultural research requires high quality and high resolution climatological data to yield accurate results, but to date this has been impossible to achieve at detailed scales and with sufficient coverage, partly due to the difficulty in compiling and revising field data and partly due to the limited climatology knowledge of agricultural researchers (with some exceptions). Large-scale datasets can be matched to certain crop models, mostly when these models can be applied at large scales (Challinor et al., 2010) or do not rely on a detailed calibration of varietal-level crop parameters (Lobell et al., 2011; Lobell et al., 2008). However, matching different modelling scales is not a trivial matter (Baron et al., 2005; Challinor et al., 2009a). Two options are available for matching two differing scales:

(1) Decreasing the resolution of the crop model from plot scale to large regions, at the expense of loss of detail in some processes [see (Challinor et al., 2007b; Challinor et al., 2004; Yao et al., 2007)], or
(2) Disaggregating the coarse-resolution climate data, at the expense of introducing noise and possibly propagating uncertainties present in the original climate model data (Tabor and Williams, 2010).

These two choices yield different results that need to be assessed and coupled. Climate data can be aggregated up to any scale to match any intended use (Masson and Knutti, 2011), but agricultural impacts need to be informed at an scale such that information can be used for decision making and adaptation (Jarvis et al., 2011). Hence, governments and international agencies should support common platforms through which data can be shared without restrictions between members of the research community. Best-bet methods can then be applied over such data to produce useable datasets that can be further shared, used and assessed in multidisciplinary and transdisciplinary approaches.

5.2. Robustness of existing weather station network

It is tacitly acknowledged that the use of interpolated surfaces can lead to errors and biases when these data are used for impact assessment (A. Jarvis, pers. comm.). However, we have demonstrated here that the effects on uncertainty are actually rather low in most of the cases, with very few exceptions (highlands of Ethiopia, the Himalayas, and some parts of the Sahara and Southern Africa, Figure 6).

The results of this research suggest that, despite weather station density being important, it may not be the only determining factor for a good ability to fill information gaps (Hijmans et al., 2005). Based on our results, we suggest that, in selecting locations to measure weather, the following factors be taken into account: (1) the nature of the variable (e.g. precipitation might be much more difficult to monitor than temperature), (2) the area where it is measured (topographically complex areas are much more variable), (3) the values of the variable in the areas where it is measured (high values are subjected to larger absolute errors, assuming relative errors are relatively uniform), (4) the relevance of the area for different subjects (i.e. the Sahara might be irrelevant for agriculture but can be of high relevance for other fields such as climate science, ecology or biodiversity and conservation), (5) possible errors in measurements and other underlying factors that can influence the measurability or correctness of estimates of a particular variable, and (6) possible political or social constraints on access to the site. Improving weather station distribution and status, as well as improving the cross-checking, correction and validation of data collected at the different sites, is fundamental for improving climate data for agricultural impact assessment.

5.3. Global climate model accuracy and performance

5.3.1. CMIP3 climate model skill

GCM performance is highly reliant on the type of comparisons performed, on the GCM formulation and on the nature of climate conditions in the analysed areas (Gleckler et al., 2008; Masson and Knutti, 2011). Underlying factors driving GCM performance are indeed difficult to track, given the complexity of the models. IPCC 4AR (CMIP3) models showed varied performance, with a high tendency to being wet-biased and no general trend for temperature. These responses reportedly have their origin in different factors: first, some GCMs have weak forcing on sea surface temperatures (SSTs), whereas climate in Africa and Asia is strongly coupled with the Atlantic and Indian Ocean and with inland water bodies (Gallée et al., 2004; Lebel et al., 2000); second, models do not properly account for the relation between inter-annual variability, ENSO and the monsoonal winds (Gallée et al., 2004; Hulme et al., 2001); third, the resolution of the models prevents acknowledgement of local-scale land use, orographic patterns and small water bodies (Hudson and Jones, 2002); fourth, cloud thickness and latent heat and moisture flux between clouds has not been properly resolved in the models (Gallée et al., 2004); and fifth, convective parameterisations produce an early onset of the seasonal rains and over-prediction of wet days and high-rainfall events (Gallée et al., 2004).

The NASA models GISS-ModelE (-R and -H) consistently presented very low predictive ability, mainly because of the models' coarse spatial resolution in conjunction with the reasons mentioned above (Hansen et al., 2007). These results agree with those of Gleckler et al. (2008), who reported that NCAR-PCM1, GISS-ModelE (-R and –H) and GISS-AOM models are the worst performing in the 24 GCMs of the CMIP3 ensemble. Similar results are reported by other authors that have assessed this or similar model ensembles (Jun et al., 2008; Pierce et al., 2009). Lack of detail in land use and land use changes (Eltahir and Gong, 1996), monsoon winds (Eltahir and Gong, 1996; Gallée et al., 2004), and sea surface temperature anomalies (SSTs) of the Atlantic and the Indian Oceans (Lebel et al., 2000; Sun et al., 1999) also causes the scales at which climate model information is robust to be varied (Masson and Knutti, 2011), and prevents local scale seasonal weather patterns from being modelled consistently (Douglass et al., 2008; Hansen et al., 2007).

5.3.2. Plugging climate model data into agricultural research

GCMs do not provide realistic representations of climate conditions in a particular site, but rather provide estimated conditions for a large area. Our results, in agreement with those from the agricultural community (Baron et al., 2005; Challinor et al., 2003) and the climate community (Jun et al., 2008; Masson and Knutti, 2011), indicate that climate model outputs cannot be input directly into plot-scale (agricultural) models, but support the idea that higher resolution climate modelling largely improves results. Either the CMIP3 (assessed here) or the upcoming CMIP5 (being released at the moment) (Moss et al., 2010) climate model outputs can be adequately used in agricultural modelling if: (1) the scales between the models are matched (see Sect. 5.1), (2) skill of models is assessed and ways to create robust model ensembles are defined, (3) uncertainty and model spread are quantified in a robust way, and (4) decision making in the context of uncertainty is fully understood.

Producing robust (i.e. skilled and certain) ensembles for agriculture is not an easy task, mainly because of the scales at which these have been found to be robust (Masson and Knutti, 2011). Opinions are contrasting: some authors support sub-selecting models based upon performance under present conditions (Matsueda and Palmer, 2011; Pierce et al., 2009), calculating a mean ensemble by weighting models based on skill (Matsueda and Palmer, 2011; Walsh et al., 2008), while others advocate using all available models with no-weighting at all (Reifen and Toumi, 2009). We suggest that until sensitivities of agricultural models to ensemble spread are fully explored (Baigorria et al., 2007), the full CMIP3 (or CMIP5) ensembles should be used.

Strategies for combining plot-scale and large-scale models and for optimising the overall result (including estimation of uncertainties derived from the scale-matching process) need to be further researched. The potential of high-quality and less uncertain climate predictions of current and future climate conditions for agricultural research is expected to have a direct impact on decision-making at different levels and for different purposes: to improve yields on the farm, to direct country level policies and investment, to define research foci, to direct international agencies' investments, and to clarify global greenhouse emissions limits and commitments (Challinor et al., 2009a; Funke and Paetz, 2011; IPCC, 2007).

6. Conclusions

A thorough analysis of different aspects of climate data for agricultural applications was performed. All topics addressed here are of high relevance to agricultural applications, particularly in the global tropics. Several important points were raised: (1) spatial scale is the most important issue for agricultural researchers, as they prefer to use monthly products with higher resolution rather than daily products with very low spatial resolution, or else limit their areas of study to field plots; (2) the sensitivities of Sub-Saharan African and Southeast Asian climate to data loss and poor availability were found to not be limiting factors for the region, with the exceptions of mountainous areas in Nepal and Ethiopia; and (3) climate modelling, although constantly improving and useful, still requires considerable future development.

As such, CMIP3 GCMs can be used with a certain degree of confidence to represent large-area climate conditions for some areas and periods. In areas where predictions lack enough skill for agricultural modelling, models can be bias-corrected using different methods [see (Challinor et al., 2009a; Hawkins et al., 2011; Reifen and Toumi, 2009)]. Whilst model skill is expected to improve with the upcoming IPCC Fifth Assessment Report, climate model ensembles as well as different methods for 'calibrating' (i.e. pre-processing for input into crop models) climate model data both need to be used, as uncertainties go beyond those derived from emissions scenarios (Hawkins et al., 2011). The proper usage of climate projections for agricultural impact assessment is of paramount importance in order to properly inform adaptation.

Finally, it is critical to understand the implications of all this to agriculture. Crops are sensitive to shortages in water and heat stresses during key periods during their development (i.e. flowering, fruit filling). Therefore, lack of skill in representing seasonal and inter-annual variability is expected to produce a significant obstacle to agricultural impact assessment of climate change; several examples in the literature exist that illustrate this (Baigorria et al., 2008; Baigorria et al., 2007). The importance of this factor depends on the strength of the climate signal on yields and the variables that drive this signal. Future impact assessments need to take into account input data and climate model data inaccuracies, sensitivities and uncertainties; make their own assessments of the inaccuracies and uncertainties; and comprehensively quantify and report uncertainties in the impact assessment process.

Acknowledgments

Authors thank Charlotte Lau, from the International Center for Tropical Agriculture (CIAT) for the useful comments in an earlier version of this manuscript, Peter G. Jones and Robert J. Hijmans for the inspiration and the previous assembling of data (WorldClim weather stations) that were used as part of this study, and the two anonymous reviewers for their useful comments and encouragement. Authors also thank the Theme 1 of the CGIAR Research Program on Climate Change, Agriculture and Food Security (CCAFS) for providing the funds for carrying out this study.

References

Adam, M., Van Bussel, L.G.J., Leffelaar, P.A., Van Keulen, H. and Ewert, F., 2011. Effects of modelling detail on simulated potential crop yields under a wide range of climatic conditions. Ecological Modelling, 222(1): 131-143.

Adler, R.F. et al., 2003. The Version-2 Global Precipitation Climatology Project (GPCP) Monthly Precipitation Analysis (1979–Present). Journal of Hydrometeorology, 4(6): 1147-1167.

Aggarwal, P.K., 2008. Global climate change and Indian agriculture: impacts, adaptation and mitigation. Indian Journal of Agricultural Sciences, 78(10): 911-919.

Aggarwal, P.K., Joshi, P.K., Ingram, J.S.I. and Gupta, R.K., 2004. Adapting food systems of the Indo-Gangetic plains to global environmental change: key information needs to improve policy formulation. Environmental Science & Policy, 7(6): 487-498.

Aggarwal, P.K., Kalra, N., Chander, S. and Pathak, H., 2006. InfoCrop: A dynamic simulation model for the assessment of crop yields, losses due to pests, and environmental impact of agro-ecosystems in tropical environments. I. Model description. Agricultural Systems, 89(1): 1-25.

Álvarez-Villa, O.D., Vélez, J.I. and Poveda, G., 2010. Improved long-term mean annual rainfall fields for Colombia. International Journal of Climatology: n/a-n/a.

Baigorria, G.A., Jones, J.W. and O'Brien, J.J., 2008. Potential predictability of crop yield using an ensemble climate forecast by a regional circulation model. Agricultural and Forest Meteorology, 148(8-9): 1353-1361.

Baigorria, G.A., Jones, J.W., Shin, D.W., Mishra, A. and Brien, J.J., 2007. Assessing uncertainties in crop model simulations using daily bias-corrected Regional Circulation Model outputs. Climate Research, 34(3): 211-222.

Baron, C. et al., 2005. From GCM grid cell to agricultural plot: scale issues affecting modelling of climate impact. Philosophical Transactions of the Royal Society B: Biological Sciences, 360(1463): 2095-2108.

Barrios, S., Ouattara, B. and Strobl, E., 2008. The impact of climatic change on agricultural production: Is it different for Africa? Food Policy, 33(4): 287-298.

Battisti, D.S. and Naylor, R.L., 2009. Historical Warnings of Future Food Insecurity with Unprecedented Seasonal Heat. Science, 323(5911): 240-244.

Begert, M., Zenklusen, E., berli, C., Appenzeller, C. and Klok, L., 2008. An automated procedure to detect discontinuities; performance assessment and application to a large European climate data set. Meteorologische Zeitschrift, 17(5): 663-672.

Boo, K.-O., Martin, G., Sellar, A., Senior, C. and Byun, Y.-H., 2011. Evaluating the East Asian monsoon simulation in climate models. J. Geophys. Res., 116(D1): D01109.

Boote, K.J., Jones, J.W. and Pickering, N.B., 1996. Potential uses and limitations of crop models. Agronomy Journal, 88(5): 704-716.

Buytaert, W., Célleri, R. and Timbe, L., 2009. Predicting climate change impacts on water resources in the tropical Andes: Effects of GCM uncertainty. Geophys. Res. Lett., 36(7): L07406.

Byjesh, K., Kumar, S. and Aggarwal, P., 2010. Simulating impacts, potential adaptation and vulnerability of maize to climate change in India. Mitigation and Adaptation Strategies for Global Change, 15(5): 413-431.

Challinor, A., Wheeler, T., Garforth, C., Craufurd, P. and Kassam, A., 2007a. Assessing the vulnerability of food crop systems in Africa to climate change. Climatic Change, 83(3): 381-399.

Challinor, A.J., Ewert, F., Arnold, S., Simelton, E. and Fraser, E., 2009a. Crops and climate change: progress, trends, and challenges in simulating impacts and informing adaptation. Journal of Experimental Botany, 60(10): 2775-2789.

Challinor, A.J. et al., 2009b. Methods and Resources for Climate Impacts Research. Bulletin of the American Meteorological Society, 90(6): 836-848.

Challinor, A.J., Simelton, E.S., Fraser, E.D.G., Hemming, D. and Collins, M., 2010. Increased crop failure due to climate change: assessing adaptation options using models and socio-economic data for wheat in China. Environmental Research Letters, 5(3): 034012.

Challinor, A.J., Slingo, J.M., Wheeler, T.R., Craufurd, P.Q. and Grimes, D.I.F., 2003. Toward a Combined Seasonal Weather and Crop Productivity Forecasting System: Determination of the Working Spatial Scale. Journal of Applied Meteorology, 42(2): 175-192.

Challinor, A.J. and Wheeler, T.R., 2008. Crop yield reduction in the tropics under climate change: Processes and uncertainties. Agricultural and Forest Meteorology, 148(3): 343-356.

Challinor, A.J., Wheeler, T.R., Craufurd, P.Q., Ferro, C.A.T. and Stephenson, D.B., 2007b. Adaptation of crops to climate change through genotypic responses to mean and extreme temperatures. Agriculture, Ecosystems & Environment, 119(1-2): 190-204.

Challinor, A.J., Wheeler, T.R., Craufurd, P.Q., Slingo, J.M. and Grimes, D.I.F., 2004. Design and optimisation of a large-area process-based model for annual crops. Agricultural and Forest Meteorology, 124(1-2): 99-120.

Challinor, A.J., Wheeler, T.R., Slingo, J.M. and Hemming, D., 2005. Quantification of physical and biological uncertainty in the simulation of the yield of a tropical crop using present-day and doubled CO2 climates. Philosophical Transactions of the Royal Society B: Biological Sciences, 360(1463): 2085-2094.

Chipanshi, A.C., Chanda, R. and Totolo, O., 2003. Vulnerability Assessment of the Maize and Sorghum Crops to Climate Change in Botswana. Climatic Change, 61(3): 339-360.

Collins, W.D. et al., 2006. The Community Climate System Model Version 3 (CCSM3). Journal of Climate, 19(11): 2122-2143.

Cooper, P.J.M. et al., 2008. Coping better with current climatic variability in the rain-fed farming systems of sub-Saharan Africa: An essential first step in adapting to future climate change? Agriculture, Ecosystems & Environment, 126(1-2): 24-35.

DeGaetano, A.T., 2006. Attributes of Several Methods for Detecting Discontinuities in Mean Temperature Series. Journal of Climate, 19(5): 838-853.

Delworth, T.L. et al., 2006. GFDL's CM2 Global Coupled Climate Models. Part I: Formulation and Simulation Characteristics. Journal of Climate, 19(5): 643-674.

Déqué, M., Dreveton, C., Braun, A. and Cariolle, D., 1994. The ARPEGE/IFS atmosphere model: a contribution to the French community climate modelling. Climate Dynamics, 10(4): 249-266.

Di Luzio, M., Johnson, G.L., Daly, C., Eischeid, J.K. and Arnold, J.G., 2008. Constructing Retrospective Gridded Daily Precipitation and Temperature Datasets for the Conterminous United States. Journal of Applied Meteorology and Climatology, 47(2): 475-497.

Diansky, N.A. and Zalensky, V.B., 2002. Simulation of present-day climate with a coupled Atmosphere-ocean general circulation model. Izvestiya, Atmospheric and Ocean Physiscs, 38(6): 732-747.

Douglass, D.H., Christy, J.R., Pearson, B.D. and Singer, S.F., 2008. A comparison of tropical temperature trends with model predictions. International Journal of Climatology, 28(13): 1693-1701.

El-Sharkawy, M., 2005. How can calibrated research-based models be improved for use as a tool in identifying genes controlling crop tolerance to environmental stresses in the era of genomics—from an experimentalist's perspective. Photosynthetica, 43(2): 161-176.

Eltahir, E.A.B. and Gong, C., 1996. Dynamics of Wet and Dry Years in West Africa. Journal of Climate, 9(5): 1030-1042.

FAO, 2001. FAOCLIM 2.0 A World-Wide Agroclimatic Database. In: FAO (Editor), Rome, Italy.

FAO, 2009. The State of Food and Agriculture. FAO, Rome, Italy.

Funke, M. and Paetz, M., 2011. Environmental policy under model uncertainty: a robust optimal control approach. Climatic Change, 107(3): 225-239.

Furevik, T. et al., 2003. Description and evaluation of the bergen climate model: ARPEGE coupled with MICOM. Climate Dynamics, 21(1): 27-51.

Gallée, H. et al., 2004. A high-resolution simulation of a West African rainy season using a regional climate model. J. Geophys. Res., 109(D5): D05108.

Gleckler, P.J., Taylor, K.E. and Doutriaux, C., 2008. Performance metrics for climate models. J. Geophys. Res., 113(D6): D06104.

Gordon, C. et al., 2000. The simulation of SST, sea ice extents and ocean heat transports in a version of the Hadley Centre coupled model without flux adjustments. Climate Dynamics, 16(2): 147-168.

Gordon, H.B. et al., 2002. The CSIRO Mk3 Climate System Model, Aspendale: CSIRO Atmospheric Research.

Govindan, R.B. et al., 2002. Global Climate Models Violate Scaling of the Observed Atmospheric Variability. Physical Review Letters, 89(2): 028501.

Grötzner, A., Sausen, R. and Claussen, M., 1996. The impact of sub-grid scale sea-ice inhomogeneities on the performance of the atmospheric general circulation model ECHAM3. Climate Dynamics, 12(7): 477-496.

Gualdi, S., Scoccimarro, E. and Navarra, A., 2008. Changes in Tropical Cyclone Activity due to Global Warming: Results from a High-Resolution Coupled General Circulation Model. Journal of Climate, 21(20): 5204-5228.

Hansen, J. et al., 2007. Climate simulations for 1880–2003 with GISS modelE. Climate Dynamics, 29(7): 661-696.

Hasumi, H. and Emori, S., 2004. K-1 Coupled GCM (MIROC) Description, CCSR, NIES and FRCGC.

Hawkins, E., Osborne, T.M., Ho, C.K. and Challinor, A.J., 2011. Calibration and bias correction of climate projections for crop modelling: an idealised case study over Europe. Agricultural and Forest Meteorology, in press.

Hew, C.-S., Krotkov, G. and Canvin, D.T., 1969. Effects of Temperature on Photosynthesis and CO2 Evolution in Light and Darkness by Green Leaves. Plant Physiology, 44(5): 671-677.

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. and Jarvis, A., 2005. Very high resolution interpolated climate surfaces for global land areas. International Journal of Climatology, 25(15): 1965-1978.

Hoogenboom, G. et al., 1994. Crop Models. Decision Support System for Agrotechnology Transfer, Version 3.0 (DSSAT V3.0), vol. 2–2. International Benchmark Sites Network for Agrotechnology Transfer (IBSNAT) Project. University of Hawaii, Department of Agronomy and Soil Science, Honolulu, Hawaii.

Hudson, D.A. and Jones, R.G., 2002. Regional Climate Model Simulations of Present-day and Future Climates of Southern Africa, MetOffice, Hadley Centre, UK, Reading, UK.

Huffman, G.J., Adler, R.F., Bolvin, D.T. and Gu, G., 2009. Improving the global precipitation record: GPCP Version 2.1. Geophys. Res. Lett., 36(17): L17808.

Huffman, G.J. et al., 2007. The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-Global, Multiyear, Combined-Sensor Precipitation Estimates at Fine Scales. Journal of Hydrometeorology, 8(1): 38-55.

Hulme, M., Doherty, R., Ngara, T., New, M. and Lister, D., 2001. African climate change: 1900-2100. Climate Research, 17(2): 145-168.

Huntingford, C., Hugo Lambert, F., Gash, J.H.C., Taylor, C.M. and Challinor, A.J., 2005. Aspects of climate change prediction relevant to crop productivity. Philosophical Transactions of the Royal Society B: Biological Sciences, 360(1463): 1999-2009.

Hutchinson, M.F., 1995. Interpolating mean rainfall using thin plate smoothing splines. International Journal of Geographical Information Systems, 9(4): 385 - 403.

IPCC, 2007. IPCC Fourth Assessment Report: Climate Change 2007 (AR4). IPCC, Geneva, Switzerland.

Jagtap, S.S. and Jones, J.W., 2002. Adaptation and evaluation of the CROPGRO-soybean model to predict regional yield and production. Agriculture, Ecosystems & Environment, 93(1-3): 73-85.

Jarvis, A. et al., 2011. An Integrated Adaptation and Mitigation Framework for Developing Agricultural Research: Synergies and Trade-offs. Experimental Agriculture, 47(Special Issue 02): 185-203.

Jarvis, A., Ramirez, J., Anderson, B., Leibing, C. and Aggarwal, P., 2010. Scenarios of Climate Change Within the Context of Agriculture. Climate Change and Crop Production. CAB International.

Jarvis, A., Reuter, H.I., Nelson, A. and Guevara, E., 2008. Hole-filled seamless SRTM data V4. In: I.C.f.T.A. (CIAT) (Editor). International Center for Tropical Agriculture (CIAT).

Johns, T.C. et al., 2006. The New Hadley Centre Climate Model (HadGEM1): Evaluation of Coupled Simulations. Journal of Climate, 19(7): 1327-1353.

Jones, P.G. and Thornton, P.K., 1999. Fitting a third-order Markov rainfall model to interpolated climate surfaces. Agricultural and Forest Meteorology, 97(3): 213-231.

Jones, P.G. and Thornton, P.K., 2003. The potential impacts of climate change on maize production in Africa and Latin America in 2055. Global Environmental Change, 13(1): 51-59.

Jones, P.G., Thornton, P.K. and Heinke, J., 2009. Generating characteristic daily weather data using downscaled climate model data from the IPCC's Fourth Assessment, ILRI, Nairobi, Kenya.

Joubert, A.M., Katzfey, J.J., McGregor, J.L. and Nguyen, K.C., 1999. Simulating midsummer climate over southern Africa using a nested regional climate model. J. Geophys. Res., 104(D16): 19015-19025.

Jun, M., Knutti, R. and Nychka, D.W., 2008. Spatial Analysis to Quantify Numerical Model Bias and Dependence. Journal of the American Statistical Association, 103(483): 934-947.

Jungclaus, J.H. et al., 2006. Ocean Circulation and Tropical Variability in the Coupled Model ECHAM5/MPI-OM. Journal of Climate, 19(16): 3952-3972.

Lane, A. and Jarvis, A., 2007. Changes in Climate will modify the Geography of Crop Suitability: Agricultural Biodiversity can help with Adaptation. Journal of the Semi-Arid Tropics, 4(1): 1-12.

Lebel, T., Delclaux, F., Le Barbé, L. and Polcher, J., 2000. From GCM scales to hydrological scales: rainfall variabiliy in West Africa. Stochastic Environmental Research and Risk Assessment, 14: 275-295.

Liu, J. et al., 2008. A spatially explicit assessment of current and future hotspots of hunger in Sub-Saharan Africa in the context of global change. Global and Planetary Change, 64(3-4): 222-235.

Lobell, D.B., Banziger, M., Magorokosho, C. and Vivek, B., 2011. Nonlinear heat effects on African maize as evidenced by historical yield trials. Nature Clim. Change, 1(1): 42-45.

Lobell, D.B. and Burke, M.B., 2010. On the use of statistical models to predict crop yield responses to climate change. Agricultural and Forest Meteorology, 150(11): 1443-1452.

Lobell, D.B. et al., 2008. Prioritizing Climate Change Adaptation Needs for Food Security in 2030. Science, 319(5863): 607-610.

Marti, O. et al., 2005. The new IPSL climate system model: IPSL-CM4, Institute Pierre Simon Laplace.

Masson, D. and Knutti, R., 2011. Spatial-Scale Dependence of Climate Model Performance in the CMIP3 Ensemble. Journal of Climate, 24(11): 2680-2692.

Matsueda, M. and Palmer, T.N., 2011. Accuracy of climate change predictions using high resolution simulations as surrogates of truth. Geophys. Res. Lett., 38(5): L05803.

Mitchell, T.D., Carter, T.R., Jones, P.D., Hulme, M. and New, M., 2004. A comprehensive set of high-resolution grids of monthly climate for Europe and the globe: the observed record (1901-2000) and 16 scenarios (2001-2100). Tyndall Centre for Climate Change Research, Oxford, UK.

Moss, R.H. et al., 2010. The next generation of scenarios for climate change research and assessment. Nature, 463(7282): 747-756.

NCDC, 2011. National Climatic Data Center, NESDIS, NOAA, U.S. Department of Commerce. Global Surface Summary of the Day - GSOD. In: NCDC, NIESDIS, NOAA and USDC (Editors). National Climatic Data Center (NCDC), Asheville, NC, US.

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

Niu, X., Easterling, W., Hays, C.J., Jacobs, A. and Mearns, L., 2009. Reliability and input-data induced uncertainty of the EPIC model to estimate climate change impact on sorghum yields in the U.S. Great Plains. Agriculture, Ecosystems & Environment, 129(1-3): 268-276.

Oldeman, L.R., 1988. An Agroclimatic Characterization of Madagascar. International Soil Reference and Information Centre, Wageningen, Netherlands.

Osborne, T.M., Lawrence, D.M., Challinor, A.J., Slingo, J.M. and Wheeler, T.R., 2007. Development and assessment of a coupled crop–climate model. Global Change Biology, 13(1): 169-183.
Parry, M., Rosenzweig, C. and Livermore, M., 2005. Climate change, global food supply and risk of hunger. Philosophical Transactions of the Royal Society B: Biological Sciences, 360(1463): 2125-2138.
PCMDI, 2007. IPCC Model Output. Available at: http://www.pcmdi.llnl.gov/ipcc/about_ipcc.php.
Peterson, T.C. and Easterling, D.R., 1994. Creation of homogenous composite climatological reference series. International Journal of Climatology, 14: 671-679.
Peterson, T.C. and Vose, R.S., 1997. An Overview of the Global Historical Climatology Network Temperature Database. Bulletin of the American Meteorological Society, 78(12): 2837-2849.
Pielke, S.R.A. et al., 2007. A new paradigm for assessing the role of agriculture in the climate system and in climate change. Agricultural and Forest Meteorology, 142(2-4): 234-254.
Pierce, D.W., Barnett, T.P., Santer, B.D. and Gleckler, P.J., 2009. Selecting global climate models for regional climate change studies. Proceedings of the National Academy of Sciences, 106(21): 8441-8446.
Prasad, P.V.V., Boote, K.J., Allen, L.H. and Thomas, J.M.G., 2002. Effects of elevated temperature and carbon dioxide on seed-set and yield of kidney bean (Phaseolus vulgaris L.). Global Change Biology, 8(8): 710-721.
Quiggin, J., 2008. Uncertainty and Climate Change Policy. Economic Analysis & Policy, 38(2): 203-210.
Ramankutty, N., Evan, A.T., Monfreda, C. and Foley, J.A., 2008. Farming the planet: 1. Geographic distribution of global agricultural lands in the year 2000. Global Biogeochem. Cycles, 22(1): GB1003.
Reichler, T. and Kim, J., 2008. How Well Do Coupled Models Simulate Today's Climate? Bulletin of the American Meteorological Society, 89(3): 303-311.
Reifen, C. and Toumi, R., 2009. Climate projections: Past performance no guarantee of future skill? Geophys. Res. Lett., 36(13): L13704.
Russell, G., Miller, J.R. and Rind, D., 1995. A Coupled Atmosphere-Ocean Model for Transient Climate Change Studies. Atmosphere-Ocean, 33(4): 683-730.
Salas-Mélia, D. et al., 2005. Description and validation of the CNRM-CM3 global coupled model, CNRM.
Schlenker, W. and Lobell, D.B., 2010. Robust negative impacts of climate change on African agriculture. Environmental Research Letters, 5(1): 014010.
Schmidt, G.A. et al., 2006. Present-Day Atmospheric Simulations Using GISS ModelE: Comparison to In Situ, Satellite, and Reanalysis Data. Journal of Climate, 19(2): 153-192.
Schneider, U., Becker, A., Meyer-Christoffer, A., Ziese, M. and Rudolf, B., 2010. Global Precipitation Analysis Products of the GPCC, Global Preciptation Climatology Centre (GPCC).
Scinocca, J.F., McFarlane, N.A., Lazare, M., Li, J. and Plummer, D., 2008. The CCCma third generation AGCM and its extension into the middle atmosphere. Atmos. Chem. Phys. Discuss., 8(2): 7883-7930.
Soltani, A., Meinke, H. and de Voil, P., 2004. Assessing linear interpolation to generate daily radiation and temperature data for use in crop simulations. European Journal of Agronomy, 21(2): 133-148.
Soria-Auza, R.W. et al., 2010. Impact of the quality of climate models for modelling species occurrences in countries with poor climatic documentation: a case study from Bolivia. Ecological Modelling, 221(8): 1221-1229.
Sun, L., Semazzi, F.H.M., Giorgi, F. and Ogallo, L., 1999. Application of the NCAR Regional Climate Model to eastern Africa 2. Simulation of interannual variability of short rains. J. Geophys. Res., 104(D6): 6549-6562.
Tabor, K. and Williams, J.W., 2010. Globally downscaled climate projections for assessing the conservation impacts of climate change. Ecological Applications, 20(2): 554-565.

Thornton, P.K., Jones, P.G., Alagarswamy, G. and Andresen, J., 2009. Spatial variation of crop yield response to climate change in East Africa. Global Environmental Change, 19(1): 54-65.

Thornton, P.K., Jones, P.G., Ericksen, P.J. and Challinor, A.J., 2011. Agriculture and food systems in sub-Saharan Africa in a 4°C+ world. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 369(1934): 117-136.

Thuzar, M., Puteh, A.B., Abdullah, N.A.P., Lassim, M.B.M. and Jusoff, K., 2010. The Effects of Temperature Stress on the Quality and Yield of Soya Bean [(Glycine max L.) Merrill.]. Journal of Agricultural Science, 2(1): 172-179.

Trnka, M., Dubrovský, M., Semerádová, D. and Žalud, Z., 2004. Projections of uncertainties in climate change scenarios into expected winter wheat yields. Theoretical and Applied Climatology, 77(3): 229-249.

Uppala, S.M. et al., 2005. The ERA-40 re-analysis. Quarterly Journal of the Royal Meteorological Society, 131(612): 2961-3012.

van Ittersum, M.K., Howden, S.M. and Asseng, S., 2003. Sensitivity of productivity and deep drainage of wheat cropping systems in a Mediterranean environment to changes in CO2, temperature and precipitation. Agriculture, Ecosystems & Environment, 97(1-3): 255-273.

Walker, W.E. et al., 2003. Defining Uncertainty: A Conceptual Basis for Uncertainty Management in Model-Based Decision Support. Integrated Assessment, 4(1): 5-17.

Walsh, J.E., Chapman, W.L., Romanovsky, V., Christensen, J.H. and Stendel, M., 2008. Global Climate Model Performance over Alaska and Greenland. Journal of Climate, 21(23): 6156-6174.

Washington, R. et al., 2006. African Climate Change: Taking the Shorter Route. Bulletin of the American Meteorological Society, 87(10): 1355-1366.

Washington, W.M. et al., 2000. Parallel climate model (PCM) control and transient simulations. Climate Dynamics, 16(10): 755-774.

Wilby, R.L. et al., 2009. A review of climate risk information for adaptation and development planning. International Journal of Climatology, 29(9): 1193-1215.

Xiong, W., Holman, I., Conway, D., Lin, E. and Li, Y., 2008. A crop model cross calibration for use in regional climate impacts studies. Ecological Modelling, 213(3-4): 365-380.

Yao, F., Xu, Y., Lin, E., Yokozawa, M. and Zhang, J., 2007. Assessing the impacts of climate change on rice yields in the main rice areas of China. Climatic Change, 80(3): 395-409.

Yongqiang, Y., Xuehong, Z. and Yufu, G., 2004. Global coupled ocean-atmosphere general circulation models in LASG/IAP. Advances in Atmospheric Sciences, 21(3): 444-455.

Yukimoto, S. et al., 2001. The New Meteorological Research Institute Coupled GCM (MRI-CGCM2) Model Climate and Variability. Papers in Meteorology and Geophysics, 51(2): 47-88.

Figure captions

Figure 1 Cascade of constraints to climate data, as normally observed in agricultural impact assessment

Figure 2 Areas of study. Bold-outlined areas indicate the areas on which the study focused (SN: Senegal, ML: Mali, NE: Niger, BF: Burkina Faso, GH: Ghana, UG: Uganda, ET: Ethiopia, KE: Kenya, TZ: Tanzania, NP: Nepal, BD: Bangladesh, IN: India)

Figure 3 Topics treated in the analysed agricultural studies. WG: weather generators.

Figure 4 Frequency of use of the different data sources in agricultural studies. A. Present-day climates. B. Future climates. Datasets acronyms are as follows: CRU-TS: Climatic Research Unit monthly time series product at 0.5 degree, GCM: global climate model output, RCM: regional climate model, CRU-CL: CRU monthly climatology product at 10 arc-minute, MARS: Data from the MARS European project, GSOD: Global summary of the day, ARTES: Africa rainfall and temperature evaluation system, VEMAP: United States comprehensive dataset, ATEAM: Advanced Terrestrial Ecosystem Analysis and Modelling, PRISM: United States dataset, GPCP: Global Precipitation Climatology Project, GPCC: Global Precipitation Climatology Centre, GHCN: Global Historical Climatology Network, AI GCM: GCM data "as is", SD GCM: statistically downscaled GCM, PS GCM: pattern scaled GCM, WG GCM: GCM data through a weather generator, SC Variables: systematic changes in target key variables, Unclear: not specified clearly in study, ARPEGE: the ARPEGE Atmospheric GCM (Déqué et al., 1994).

Figure 5 Performance of the interpolations for all variables and months as measured by the R-square value. A. Rainfall, B. Mean temperature, C. Maximum temperature, D. Minimum temperature

Figure 6 Uncertainties in WorldClim expressed as standard deviations from the mean of the 100 cross-validated folds for (A) total annual rainfall (in mm), and (B) annual mean temperature (in ºC).

Figure 7 Comparison (R-square based) of observed climatology (CL-WS [w], GHCN-CL [g] and GSOD-CL [o]) and each of the GCMs (GCM-CL) for each of the countries in the study area for mean temperature (top), temperature range (middle) and precipitation (bottom), for the annual and two seasonal (DJF, JJA) means or totals. All R^2 values were statistically significant at $p<0.0001$

Figure 8 Comparison (R-square based) of interpolated climatology (i.e. CRU-IS [c], WCL-IS [w]), and each of the GCMs (GCM-CL) for each of the countries in the study area for mean temperature (top), temperature range (middle) and precipitation (bottom) for the annual mean or total and two seasons (DJF, JJA). All R^2 values were statistically significant at $p<0.001$.

Figure 9 Root mean squared error (RMSE), in millimetres, between observed (GHCN-TS) and GCM (GCM-TS) time series, for the 24 GCMs in Table 2, for annual total rainfall between the years 1961-1990.

Figure 10 Mean bias of GCM (GCM-TS) time series compared to observed time series (GHCN-TS), for the 24 GCMs in Table 2, for annual total rainfall between the years 1961-1990. Bias is expressed as the slope of the regression curve between observed and climate-model series. Values below 1 (light grey areas) indicate that GCMs are wet-biased, whereas values above 1 (dark grey areas) indicate that GCMs are dry-biased.

Figure 11 Root mean squared error (RMSE), in Celsius degree, between observed (GHCN-TS) and GCM (GCM-TS) time series, for the 24 GCMs in Table 2, for annual mean temperature between the years 1961-1990

Figure 12 Mean bias of GCM (GCM-TS) time series compared to observed time series (GHCN-TS), for the 24 GCMs in Table 2, for annual mean temperature between the years 1961-1990. Bias is expressed as the slope of the regression curve between observed and climate-model series. Values below 1 (light grey areas) indicate that GCMs are warm-biased, whereas values above 1 (dark grey areas) indicate that GCMs are cold-biased.

Table 1 Number of locations per data source (global)

Source*	Precipitation stations	Mean temperature stations	Min., Max. temperature stations	Period
GHCN v2	20,590	7,280	4,966	1950-2000
WMO CLINO	4,261	3,084	2,504	1961-1990
FAOCLIM 2.0	27,372	20,825	11,543	1960-1990
CIAT	18,895	13,842	5,321	1950-2000

*GHCN v2: Global Historical Climatology Network version 2 (Peterson and Vose, 1997); WMO CLINO: World Meteorological Organization Climatology Normals; FAOCLIM 2.0: Food and Agriculture Organization of the United Nations Agro-Climatic database (FAO, 2001); CIAT: Database assembled by Peter J. Jones at the International Center for Tropical Agriculture (CIAT).

Table 2 Available GCMs, resolutions, and main references

Model	Country	Atmosphere	Ocean	Reference
BCCR-BCM2.0	Norway	T63, L31	1.5x0.5, L35	(Furevik et al., 2003)
CCCMA-CGCM3.1 (T47)	Canada	T47 (3.75x3.75), L31	1.85x1.85, L29	(Scinocca et al., 2008)
CCCMA-CGCM3.1 (T63)	Canada	T63 (2.8x2.8), L31	1.4x0.94, L29	(Scinocca et al., 2008)
CNRM-CM3	France	T63 (2.8x2.8), L45	1.875x(0.5-2), L31	(Salas-Mélia et al., 2005)
CSIRO-Mk3.0	Australia	T63, L18	1.875x0.84, L31	(Gordon et al., 2002)
CSIRO-Mk3.5	Australia	T63, L18	1.875x0.84, L31	(Gordon et al., 2002)
GFDL-CM2.0	USA	2.5x2.0, L24	1.0x(1/3-1), L50	(Delworth et al., 2006)
GFDL-CM2.1	USA	2.5x2.0, L24	1.0x(1/3-1), L50	(Delworth et al., 2006)
GISS-AOM	USA	4x3, L12	4x3, L16	(Russell et al., 1995)
GISS-MODEL-EH	USA	5x4, L20	5x4, L13	(Schmidt et al., 2006)
GISS-MODEL-ER	USA	5x4, L20	5x4, L13	(Schmidt et al., 2006)
IAP-FGOALS1.0-G	China	2.8x2.8, L26	1x1, L16	(Yongqiang et al., 2004)
INGV-ECHAM4	Italy	T42, L19	2x(0.5-2), L31	(Gualdi et al., 2008)
INM-CM3.0	Russia	5x4, L21	2.5x2, L33	(Diansky and Zalensky, 2002)
IPSL-CM4	France	2.5x3.75, L19	2x(1-2), L30	(Marti et al., 2005)
MIROC3.2-HIRES	Japan	T106, L56	0.28x0.19, L47	(Hasumi and Emori, 2004)
MIROC3.2-MEDRES	Japan	T42, L20	1.4x(0.5-1.4), L43	(Hasumi and Emori, 2004)
MIUB-ECHO-G	Germany/Korea	T30, L19	T42, L20	(Grötzner et al., 1996)
MPI-ECHAM5	Germany	T63, L32	1x1, L41	(Jungclaus et al., 2006)
MRI-CGCM2.3.2A	Japan	T42, L30	2.5x(0.5-2.0)	(Yukimoto et al., 2001)
NCAR-CCSM3.0	USA	T85L26, 1.4x1.4	1x(0.27-1), L40	(Collins et al., 2006)
NCAR-PCM1	USA	T42 (2.8x2.8), L18	1x(0.27-1), L40	(Washington et al., 2000)
UKMO-HADCM3	UK	3.75x2.5, L19	1.25x1.25, L20	(Gordon et al., 2000)
UKMO-HADGEM1	UK	1.875x1.25, L38	1.25x1.25, L20	(Johns et al., 2006)

Table 3 Summarised performance of all GCMs with available data for each of the variables and periods in the study countries for different ranges of the R^2 skill evaluation parameter.

Variable	**Period**	**Dataset***	**R^2<0.5 (%)***	**0.5<R^2<0.7 (%)***	**R^2>0.8 (%)***	**R^2>0.9 (%)***
Rainfall	Annual	IS	2.8	6.6	77.8	54.3
		WS	37.5	19.4	30.8	17.0
		ALL	23.6	14.3	49.6	31.9
	DJF	IS	17.7	19.3	49.1	25.9
		WS	38.1	17.2	31.4	15.7
		ALL	29.9	18.1	38.5	19.8
	JJA	IS	12.8	17.2	58.9	40.1
		WS	15.2	19.1	52.1	34.5
		ALL	14.2	18.3	54.8	36.7
Diurnal temperature range	Annual	IS	0.4	2.2	81.8	73.1
		WS	0.4	1.2	54.5	46.1
		ALL	0.4	1.7	68.1	59.6
	DJF	IS	0.4	2.2	80.4	71.2
		WS	0.4	2.4	53.1	47.7
		ALL	0.4	2.3	66.8	59.4
	JJA	IS	0.4	2.0	80.7	67.2
		WS	0.4	1.2	54.5	46.1
		ALL	0.4	1.6	67.6	56.6
Mean temperature	Annual	IS	0.7	1.2	96.4	95.7
		WS	2.4	1.9	93.5	91.0
		ALL	1.7	1.6	94.7	92.8
	DJF	IS	3.5	1.9	93.2	91.5
		WS	2.3	2.3	93.9	91.2
		ALL	2.8	2.2	93.6	91.3
	JJA	IS	0.0	0.0	100.0	98.8
		WS	0.0	0.1	99.8	98.5
		ALL	0.0	0.1	99.9	98.6

* Values are expressed as percent of country-GCM combinations for comparisons of GCM-CL and different observational datasets: interpolated surfaces (IS), namely, WCL-IS and CRU-IS; weather stations (WS), namely, GHCN-CL, WCL-WS, GSOD-CL; and as the average of IS and WS (ALL)

Assessing relevant climate data for agricultural applications

Authors

Julian Ramirez-Villegas
International Center for Tropical Agriculture (CIAT), Cali, Colombia
CGIAR Research Program on Climate Change, Agriculture and Food Security (CCAFS)
Institute for Climate and Atmospheric Science (ICAS), School of Earth and Environment, University of Leeds, Leeds, UK

Andy Challinor
Institute for Climate and Atmospheric Science (ICAS) School of Earth and Environment, University of Leeds, Leeds, UK
CGIAR Research Program on Climate Change, Agriculture and Food Security (CCAFS)

Target journal

Agricultural and Forest Meteorology

Abstract

Climate change is expected to substantially reduce agricultural yields, as reported in the by the Intergovernmental Panel on Climate Change (IPCC). In Sub-Saharan Africa and (to a lesser extent) in South Asia, limited data availability and institutional networking constrain agricultural research and development. Here we performed a review of relevant aspects in relation to coupling agriculture-climate predictions, and a three-step analysis of the importance of climate data for agricultural impact assessment. First, using meta-data from the scientific literature we examined trends in the use of climate and weather data in agricultural research, and we found that despite agricultural researchers' preference for field-scale weather data (50.4% of cases in the assembled literature), large-scale datasets coupled with weather generators can be useful in the agricultural context. Using well-known interpolation techniques, we then assessed the sensitivities of the weather station network to the lack of data and found high sensitivities to data loss only over mountainous areas in Nepal and Ethiopia (random removal of data impacted precipitation estimates by ±1,300 mm/year and temperature estimates by ±3°C). Finally, we numerically compared IPCC Fourth Assessment Report climate models' representation of mean climates and interannual variability with different observational datasets. Climate models were found inadequate for field-scale agricultural studies in West Africa and South Asia, as their ability to represent mean climates and climate variability was limited: more than 50% of the country-model combinations showed <50% adjustment for annual mean rainfall (mean climates), and there were large rainfall biases in GCM outputs (1,000 to 2,500 mm/year), although this varied on a GCM basis (climate variability). Temperature biases were also large for certain areas (5-10°C in the Himalayas and Sahel). All this is expected to improve with IPCC's Fifth Assessment Report; hence, appropriate usage of even these new climate models is still required. This improved usage entails bias reduction (weighting of climate models or bias-correcting the climate change signals), the implementation of methods to match the spatial scales, and the quantification of uncertainties to the maximum extent possible.

Keywords: Sub-Saharan Africa; South Asia; climate modelling; climate model; skill; uncertainty; CMIP3; CMIP5.

1. Introduction

Agriculture is expected to play an important role in the context of climate change, not only because it is considered amongst the most vulnerable sectors, but also because it is part of the solution (i.e. potential to mitigate greenhouse gases [GHGs] emissions) (FAO, 2009; IPCC, 2007). Agriculture will likely be severely affected over the next hundred years due to unprecedented rates of changes in the climate system (IPCC, 2007; Jarvis et al., 2010; Lobell et al., 2008; Thornton et al., 2011). Some of these impacts have already been observed (Battisti and Naylor, 2009; Schlenker and Lobell, 2010). To help cope with such impacts, a framework to assess the effects of climate change on agriculture and food security and to aid with adaptation was established in 2008, as described by Jarvis et al. (2011): The Consultative Group of International Agricultural Research (CGIAR) Research Program on Climate Change, Agriculture and Food Security (CCAFS).

For adaptation to be successful, agricultural and climate data are crucial, and these are scarce in their basic forms (data from research and weather stations, respectively) or not very well managed and/or maintained in certain parts of the world. Most importantly, climate databases and their derived products are sometimes inaccurate, or else lack the documentation necessary to facilitate their use within the agricultural research community. In some instances, this may be indicative of the gap between the agricultural and climate research communities (Pielke et al., 2007; Thornton et al., 2011). Even when the two do collaborate, agricultural researchers face critical constraints when accessing basic sources of climate data (i.e. weather stations) due to a number of factors, from access to data, to weather maintenance and data quality checks, to the weather itself (DeGaetano, 2006).

In the last 10 years, various datasets have been developed by different institutions, usually based on either a combination of weather station data, satellite data, and numerical weather prediction models in addition to interpolation methods, or on the sole application of climate models. The usage of these datasets for agricultural modelling purposes is rather limited for one or more of the following reasons: (1) their time step is long (monthly in the best case); (2) their temporal coverage is limited to an average of several years (Hijmans et al., 2005; New et al., 2002); (3) their spatial resolution is too coarse (Adler et al., 2003; Schneider et al., 2010); (4) their geographic coverage is not wide enough (Di Luzio et al., 2008); and (5) only certain variables (i.e. temperatures, rainfall) are reported whereas other agriculturally relevant measures (e.g. potential and/or reference evapotranspiration, relative humidity, solar radiation) are rarely reported (Di Luzio et al., 2008; Hijmans et al., 2005). Moreover, assessments of these data (particularly climate models) have been done only under a climate-science perspective (Gleckler et al., 2008; Pierce et al., 2009), for a limited number of variables (Jun et al., 2008; Reifen and Toumi, 2009), or for a reduced realm (Walsh et al., 2008).

In this paper, we sought to improve the general knowledge on the available climate data for agricultural research using a three-step thorough analysis on fundamental aspects related to agricultural modelling. First, we perform a meta-analysis on the usage of various data sources for agricultural applications; second, we assess the quality and distribution of weather station records by exploring both the ability of these data to fill geographic information gaps by means of interpolation, and the sensitivities of the different regions to data loss; and finally, we assess the accuracy of climate model outputs against different observational datasets using various metrics reported in previous literature (Gleckler et al., 2008; Pierce et al., 2009). We finally analyse the main implications of our findings on agricultural impact assessment.

2. Review of knowledge and data

2.1. Understanding of processes and crop modelling

Mechanisms to fix carbon in plants (i.e. photosynthesis) are affected by a number of factors (El-Sharkawy, 2005; Prasad et al., 2002), although responses strongly depend on the type of mechanism used by the plant to produce biomass (i.e. C_4, C_3, CAM) and on any other stresses to which the plant could be subjected simultaneously. In crop production, apart from appropriate plant growth it is the amount of biomass accumulated in fruits and seeds and the nutrients in them that matters most (Thuzar et al., 2010). Yields are a direct consequence of photosynthesis and biomass accumulation, and these are directly or indirectly affected by environmental conditions [see (Challinor et al., 2009b) for a review]. Well-watered crops grown under optimal temperature and solar radiation ranges develop to their full production potential (van Ittersum et al., 2003), but growth potential reduces if the crop is stressed during the growing season (Hew et al., 1969; Huntingford et al., 2005).

Therefore, modelling crop growth depends on (1) correct formulation of the simulation model, (2) our ability to understand the effects of environmental factors on growth, and (3) correct measurement of the relevant environmental factors for correct mapping of their interactions (Boote et al., 1996; El-Sharkawy, 2005). Hence, crop modelling largely benefits from accurate measurements of temperatures, rainfall, and solar radiation, as the main factors acting on photosynthesis (Challinor and Wheeler, 2008; Hoogenboom et al., 1994), but even these basic data are often unavailable, messy, or of limited quality. The more available data there exists, the better calibration and evaluation of crop models can be (Adam et al., 2011; Niu et al., 2009; Xiong et al., 2008).

Additionally, most crop models simulate growth of individual plants and then scale out the modelling results to the plot-scale, based on management decisions such as plant and row distances, and plot size (Aggarwal et al., 2006; Boote et al., 1996; Hoogenboom et al., 1994). On the other hand, available weather data (when not measured in the field) is only available at coarse spatial scales. Matching these two spatial scales is not an easy task [see (Challinor et al., 2009a; Jagtap and Jones, 2002; Trnka et al., 2004) for a review]. The challenge is thus to increase the knowledge of the interactions between atmospheric and crop-growth processes (Boote et al., 1996) whilst avoiding model over-parameterisation (Challinor et al., 2009b), improving the accuracy of inputs (Adam et al., 2011), and matching both spatial scales (Challinor et al., 2009a). All this requires closing the gap between crop and climate scientists.

2.2. Weather data

Measurements of weather for a given site are often unavailable because (1) there is no weather station; (2) weather stations are not well maintained so data are either only available for a short period or contain gaps, (3) collected data are not properly stored; (4) data do not pass basic quality checks; and/or (5) access to data is restricted by holding institutions (Figure 1). This all further constrains agricultural impact assessment, highlighting the importance of making data public.

<Insert Figure 1 here>

Apart from the constraints related to access and weather station locations, probably the most important issue regarding weather data is quality (Begert et al., 2008; DeGaetano, 2006) (Figure 1), which also greatly affects the performance of impact models. Therefore, the climate and agricultural community has partly focused on developing methods for either temporal or spatial data gap filling, and on using such methods for developing global or regional datasets with public access (Hijmans et al., 2005; Jones and Thornton, 1999; Soltani et al., 2004).

However, uncertainties in global datasets derived from interpolation methods have been only barely (if at all) estimated (Buytaert et al., 2009; Challinor and Wheeler, 2008; Soria-Auza et al., 2010). Researchers using global datasets and any weather station source need to be aware of these problems and ought to take this into account by testing the sensitivities of their approaches to accuracy issues (i.e. inhomogeneities, discontinuities) and (if possible) providing results within the range of uncertainty in input data (i.e. such as the outputs of cross validated interpolation methods) (Challinor et al., 2005).

2.3.Climate model data

General Circulation Models (GCMs) are currently the best way to model the complex processes that occur at the earth system's level (Huntingford et al., 2005; IPCC, 2007). However, as CGMs are highly complex, they are computationally expensive, so they have only been used for predictions at coarse spatial scales. These predictions therefore involve a number of uncertainties relevant to agriculture [see (Challinor et al., 2009b; Jarvis et al., 2010; Quiggin, 2008) for reviews on the topic].

In short, uncertainty in climate modelling arises from the impossibility of modelling the climate system with complete determinism (Walker et al., 2003). This uncertainty can arise from: context (boundaries of the system modelled), model, inputs, and parameters (Walker et al., 2003). Model uncertainty can be structural or technical: structural uncertainty in models is associated with our lack of understanding of the system, whereas technical uncertainty relates to our inability to implement mathematical formulations in computational systems. Other uncertainties in climate modelling arise from variable driving forces (greenhouse gas emissions and concentrations), initial conditions and parameterised physics (Challinor et al., 2009b; Walker et al., 2003). Rationalisation and quantification of all these uncertainties under the context of agriculture is possible (see Challinor et al., 2009b for a review).

Crop modellers are thus challenged to understand the broad concepts of climate modelling uncertainties and detect the sensitivities of crop models to them, whilst also having a basic understanding of earth processes in order to identify major flaws in climate models and decide the best ways to couple them with crop models.

3. Materials and methods

Throughout this paper, we built upon existing knowledge of agricultural and climate modelling (Sect. 2) and:

1. Performed a meta-analysis on the usage of climate and weather data for agricultural modelling purposes and summarised the desirable characteristics sought when modelling crop production.
2. Analysed the robustness of the existing weather station network by assessing both the ability of these data to correctly fill information gaps via interpolation methods, and the network's sensitivities to information loss.
3. Assessed the accuracy of climate model outputs from the Fourth Assessment Report of the IPCC (IPCC, 2007) against different observational datasets, using metrics and methods reported in the climate-science literature that are also familiar to agricultural researchers.

All calculations were done by means of the software packages R-2.13.1 (available at http://www.r-project.org) and GRASS-GIS 6.4.0 (available at http://grass.fbk.edu) in a 64-bit Red Hat Enterprise Linux 5 box.

3.1. Study area

We focused on the geographic area of Africa and South Asia, where several studies have identified that significant vulnerabilities exist (Aggarwal, 2008; Aggarwal et al., 2004; Barrios et al., 2008; Byjesh et al., 2010; Challinor et al., 2007a; Chipanshi et al., 2003; Jones and Thornton, 2003; Lane and Jarvis, 2007; Liu et al., 2008; Lobell et al., 2008; Thornton et al., 2009; Thornton et al., 2011; Washington et al., 2006). In particular, we concentrate our efforts on West Africa (Senegal, Mali, Burkina Faso, Ghana and Niger), East Africa (Ethiopia, Tanzania, Uganda and Kenya) and the Indo-Gangetic Plains countries (India, Nepal, and Bangladesh), hereafter referred to as WAF, EAF and IGP, respectively (Figure 2).

<Insert Figure 2>

3.2. Analysing the usage of climate data in agricultural studies

3.2.1. Meta-data from agricultural studies

We gathered data from a number of publications on any topic that made use of climate data for any sort of agricultural modelling. We conducted searches using various search engines and downloaded only peer-reviewed publications. Review papers and the Fourth Assessment report of the IPCC were particularly useful in identifying additional published studies. We analysed all publications that in any way involved the usage of climate data for agricultural modelling purposes. As the selection of the impact assessment model is the first decision that any researcher needs to make, we focus on the driving factors of this decision. We recorded different variables from the studies as follows:

(1) Problem and/or topic in question: classified in categories such as impact assessment, seasonal yield forecasting, sole crop modelling, and climate attribution, among others. Each study was classified into only one category by taking into account only the main issue addressed by the paper;
(2) Scale of the approach: includes site, sub-national, country, regional (group of countries), and global;
(3) Use of weather generators: for both present and future, we recorded whether the study did or did not use a weather generator;
(4) Climate dataset (current): GCM when a GCM (regardless of which one) was used, RCM when an RCM (regardless of which one) was used, weather station, satellite (no further discrimination), and important datasets (i.e. CRU, WorldClim, GPCP, among others);
(5) Climate dataset (future): the nature of used future projections was recorded here including the downscaling method, if applicable. Classifications were: GCM "as is" when studies used raw GCM outputs as inputs, pattern scaled GCMs (Mitchell et al., 2004), RCMs, systematic changes to current climate data, statistical downscaling (Wilby et al., 2009), and weather generator downscaled GCM (Jones et al., 2009).

For further details on the above categories the reader is referred to our supplementary material (part 1). We revised a total of 205 peer-reviewed publications (See supplementary material part 2), printed between the years 1983 and 2011. Most of the studies were published immediately before or after the IPCC 4AR was released in 2007. When a certain study made use of two different sources of present-day climate data, it was considered twice (totalling 247 cases).

3.2.2. Analysing the usage of climate data in agricultural studies

We analysed the recent trends in the use of climate data for agriculture: the obvious constraints in the studies, the type of approaches used and the climate data inputs used to drive the chosen agricultural models. By doing this, we ensured that we covered all the main factors driving an agricultural researcher's decision to select a particular approach for a given problem.

3.3. Analysis of weather station data

3.3.1. Worldwide weather station network data

Long term climatological means of monthly precipitation and mean, maximum and minimum temperatures were assembled, as described by Hijmans et al. (2005). However, it is important to note that at the global level the sources of these data are large in number and differ in coverage, availability and quality (Table 1), and thorough quality checks were done only in a sub-set of the sources by original distributing institutions.

<Insert Table 1 here>

Additional sources such as R-Hydronet (http://www.r-hydronet.sr.unh.edu/english/) and Oldeman (1988) database for Madagascar were also included. We discarded any weather station with less than 10 years of data. The final dataset (after quality control and duplicates removal, see Hijmans et al. 2005 for more details) comprised 13,141 locations with monthly precipitation data, 3,744 locations with monthly mean temperature, and 2,684 locations with diurnal temperature range within our study region. This dataset is hereafter referred to as WCL-WS.

3.3.2. Analysing robustness of existing weather station networks

Many methods exist that allow the user to determine (interpolate) the value of a parameter (e.g., monthly rainfall) in a given condition (i.e. in a given site, at a given time, or both), where it had never been measured before. Some of these methods are already popular with researchers using climate data (Hijmans et al., 2005; Hutchinson, 1995; Jones and Thornton, 1999; New et al., 2002) either on a regional or on a global basis. For climate-variable interpolations, the robustness of weather records is critical for an accurate result.

We assessed the robustness of the weather station network by testing both the ability of weather records to yield accurate interpolation results, and the sensitivities of the network to information loss. Towards th~~a~~ose~~t~~ ends, we used the WCL-WS dataset to fit a thin plate spline interpolation algorithm (Hutchinson, 1995) for our study region. We investigated the effect of weather station availability by using 100 cross validated folds for four variables (monthly maximum, minimum and mean temperatures and total precipitation) using similar methods as in Hijmans et al. (2005) and New et al. (2002) for each fold. We used longitude, latitude and elevation as independent variables. We used 85% randomly selected data points for fitting the splines and the remaining 15% for evaluating the result for each variable and month. For the evaluation, we calculated the R^2 and the Root Mean Square Error (RMSE) and produced boxplots of the 100-fold-by-12-month interpolations for each of the four variables. As the number of stations considerably exceeded the amount of available memory for processing, we divided the whole region of study in 5 tiles, each with an equivalent number of locations. We then projected the fitted splines onto 30-arc-second gridded datasets of latitude, longitude and altitude (Jarvis et al., 2008), thus producing a total of 4,800 interpolated surfaces (12 months times 4 variables times 100 folds). Finally, we analysed the spatial variability of standard deviations and the performance of the interpolation technique as proxies for sufficient distribution and geographic density of weather stations.

3.4. Assessment of IPCC Fourth Assessment Report (4AR) model data

3.4.1. Long-term observed mean climatology from weather stations

Three different long term climatology datasets were assembled: (1) the Global Historical Climatology Network (GHCN, as in Sect. 3.3.1) version 2 (Peterson and Vose, 1997), available at http://www.ncdc.noaa.gov/pub/data/ghcn/v2. We used GHCN as an independent source because it is a global resource that contributed significantly to WCL-WS and also because it is available at more temporally disaggregated levels (i.e. monthly), thus allowing uniformity with analyses on Sect. 3.4.3 and 3.4.6. This database includes monthly historical totals (1900-2010) of precipitation (20,590 stations), and means of maximum, minimum (4,966) and mean (7,280) temperatures. GHCN data have been subject to quality checks and to a process of "homogenisation" or "adjustment" (Peterson and Easterling, 1994); however, the available data within our analysis domain consisted primarily of "unadjusted" stations. For each location (6,393 stations for rainfall, 1,278 for mean temperature and 549 for minimum and maximum temperature) within our study area, we averaged historical monthly time series for the period 1961-1990 for maximum, minimum and mean temperatures and total rainfall, resulting in a time-averaged dataset of 6,393 locations for rainfall, 1,278 for mean temperature and 549 for minimum and maximum temperature. This dataset will be hereafter referred to as GHCN-CL.

(2) WCL-WS (Sect. 3.3.1); and (3) the Global Surface Summary of the Day (GSOD) was accessed at http://www.ncdc.noaa.gov/cgi-bin/res40.pl. This database contains daily data from ~9,000 weather stations worldwide for 18 variables, including, mean, maximum, minimum and dew point temperature, sea level and location pressure, visibility, wind speed and gust, precipitation, snow depth, and specifications on the occurrence of rain, snow, fog, tornado, thunder, or hail (NOAA, 2011; ftp://ftp.ncdc.noaa.gov/pub/data/gsod/readme.txt). We selected weather stations within our study area (1,999); aggregated daily rainfall, mean, maximum and minimum temperatures to a monthly time scale; and then averaged over the period 1961-1990. This dataset will be hereafter referred to as GSOD-CL.

3.4.2. Long-term observed mean climatology from interpolated surfaces

We gathered high-resolution climatology from two different sources: (1) the high resolution climate surfaces in WorldClim (Hijmans et al., 2005), available at http://www.worldclim.org. WorldClim is a 30 arc-seconds (~1km at the equator) global dataset produced from the interpolation of long-term climatology as measured in weather stations. Global gridded data were downloaded at the 30 arc-second resolution, then masked to our analysis domain, and aggregated to 10 arc-minute using bilinear interpolation in order to reduce computational and storage time; and (2) the University of East Anglia Climatic Research Unit (CRU) dataset (New et al., 2002), available through http://www.cru.uea.ac.uk/cru/data/hrg/ (CRU-CL-2.0). This dataset was developed using the same interpolation method as WorldClim, with the main difference that WorldClim includes many more weather stations, sometimes at the expense of input data quality. CRU-CL-2.0 resolution is 10 arc-minute (~20km at the equator). Data were downloaded at the global level and masked to our analysis domain. WorldClim and CRU-CL-2.0 are hereafter referred to as WCL-IS and CRU-IS (interpolated surfaces), respectively. We used these sources because (1) they are flag products that most researchers use for impact studies; (2) they are much higher resolution than GCMs (and other products such as the Global Precipitation Climatology Project [GPCP] and the Global Precipitation Climatology Centre [GPCC]) and hence permit the capture of small-scale weather patterns (important to agriculture) as well as a direct comparison of their within-GCM-gridcell mean with the actual GCM value; (3) are based only on ground observations of weather and do not incorporate side-products such as reanalysis (Uppala et al., 2005) or satellite data (Huffman et al., 2007), both of whose accuracy is not as good.

3.4.3. Long-term observed time series

Two sources of weather time series were used: (1) long term (1961-1990) series of monthly weather conditions were gathered from GHCN version 2 (Peterson and Vose, 1997). Again, we used mainly unadjusted stations. Mean monthly temperature and total monthly historical rainfall data were used without any further processing; and (2) long-term (1961-1990) series of daily weather as in GSOD (NCDC, 2011). For GSOD, daily precipitation and monthly temperature were aggregated to the monthly level only if all days were reported with data (for rainfall) and if at least 50% of the days had data (for temperatures). This resulted in 1,999 stations within our analysis domain, although not all stations had data for all months and all years. These two data sources are hereafter referred to as GHCN-TS and GSOD-TS, respectively. Lack of data prevented us from including maximum and minimum temperatures in the GHCN-TS and the GSOD-TS datasets. In contrast to GHCN-CL and GSOD-CL, GHCN-TS and GSOD-TS include every month and every year, thus allowing the analysis of inter-annual variability.

3.4.4. Global climate model output

The latest IPCC report (Fourth Assessment Report, 4AR) comprises the sole state-of-the-art public and official source of climate data for use in impact studies (IPCC, 2007; Jarvis et al., 2010). We therefore decided to use IPCC 4AR results.

We downloaded present day (1961-1990) simulations of global climate at original GCM resolution (~100 km) from the CMIP3 (Coupled Model Intercomparison Project phase 3) web data portal at https://esg.llnl.gov:8443/index.jsp (PCMDI, 2007). We downloaded monthly time series of mean, maximum, minimum temperature and precipitation flux in NetCDF format for 24 coupled GCMs (Table 2). Separately for each GCM, w~~W~~e calculated diurnal temperature range for each month and year as the difference between maximum and minimum temperatures and calculated total monthly rainfall as the product between the precipitation rate, the water density at sea level pressure and the number of seconds in a month. We used the each climate model monthly time series (GCM-TS hereafter) and also calculated average 1961-1990 climatology by averaging, for each variable (mean temperature, diurnal temperature range and total rainfall), every month for the whole 1961-1990 period (GCM-CL hereafter). The final datasets (i.e. GCM-TS and GCM-CL, respectively) consisted of three variables (mean temperature, diurnal temperature range and total monthly rainfall) for 24 different GCMs.

<Insert Table 2 here>

3.4.5. Ability to represent long-term climatology

The extent to which GCM predictions are accurate has not been fully explored for some parts of the world, particularly in the context of agriculture (Gleckler et al., 2008; Pierce et al., 2009; Walsh et al., 2008). As previously stated (Sect. 2.1), we compared the most readily available variables from both ground observations and climate models: rainfall, mean temperature and diurnal temperature range. Data for other variables are not available for our study regions in observational datasets. As per our stated objective (Sect. 3), we performed two sets of comparisons:

- First, we compared the GCM-CL dataset with the interpolated climatology in CRU-IS, WCL-IS (Sect. 3.4.2). We performed comparisons on a country basis in order to yield country-specific results. For each GCM gridcell, the mean, maximum and minimum values of all lower scale (CRU-IS, WCL-IS) cells was first calculated and then compared to the GCM value through the determination coefficient (R^2) and corresponding p-value, the slope of a origin-forced (so that a 1:1 relationship was sought) regression curve (S) and the root mean square error (RMSE).

- Second, using the same procedure, we compared the GCM-CL dataset with observed climatology in WCL-WS (Sect. 3.3.1), GHCN-CL and GSOD-CL (Sect. 3.4.1).

We analysed total rainfall, mean temperatures and diurnal temperature ranges over three periods: December-January-February (DJF), June-July-August (JJA) and the whole year (ANN). These months represent the most critical seasons for agriculture in our study regions, and are also the most often assessed in the existing literature (Gleckler et al., 2008; Pierce et al., 2009). Due to space constraints, we present only the results of comparisons between GCM gridcell values and mean values within gridcells, unless otherwise stated. We do, however, discuss other relevant results in more general terms.

3.4.6. Ability to represent long-term monthly climate time series

CMIP3-related GCMs are known to misrepresent certain inter-annual and/or within-decade variations that are important for agricultural systems (Govindan et al., 2002). However, specific aspects of these errors have not been explored in all CMIP3 models in the context of agriculture. Therefore, in order to test the consistency of GCM predictions across time, we compared the GCM-TS (Sect. 3.4.4) dataset against the GHCN-TS and GSOD-TS (Sect. 3.4.3). The comparison was done for three periods (JJA, DJF and ANN, Sect. 3.4.4) by calculating the R^2 and corresponding p-value, the slope of the regression curve as forced to the origin and the RMSE between the two time series (GCM-TS vs. GHCN-TS and GCM-TS vs. GSOD-TS). As a GCM cell contains one or more weather stations, we averaged the monthly time series as needed before comparing the two pairs of series. Finally, we compared the performance of all GCMs across the geographic space of our study area.

4. Results

4.1. Usage of climate data in agricultural studies

4.1.1. Topics of study

The most addressed topic (41.4% of the studies) in our literature review was climate change impact assessment (Figure 3), followed by crop growth simulation (18.5%). Water resources-impact studies round out the top three topics studied (8.1%), followed by climate attribution (6.9%), crop yield forecasting (6.1%), and model assessment (5.7%). Surprisingly, formal studies addressing adaptation were rather scarce (3.6%). Pests and diseases, soils, abiotic stresses and climate risks appeared to be a lot less ~~important~~ addressed than impact assessment and crop growth simulation studies, which together accounted for more than 50% of the total publications.

<Insert Figure 3 here>

4.1.2. Scale of studies and type of models

Most of the studies performed their models at a scale less than the size of a country; site-specific or sub-national level together comprised 55% of the studies. Very few (7%) of the studies were performed at the global level, likely because of the type of models used: field-scale mechanistic crop growth models were the most utilised overall (69.2%); followed by statistical and/or empirical approaches (S/E, 21.4%), which most of the crop growth modellers criticise for not being accurate enough (Lobell and Burke, 2010; Lobell et al., 2008); and finally by hydrological models (10%). The frequent use of field-based crop growth models suggests that the time step requirement for input data is rather high (El-Sharkawy, 2005), also confirmed by the usage of weather generators (8.5 and 11.2% for present and future climates, respectively).

4.1.3. Climate data sources

Unlike the model types, which were quite similar, the sources of present climate data varied substantially, with a total of 32 different sources being used for present climate data (Figure 4A). On average, a different present-day-climate dataset was used for every 7 agricultural studies. The most commonly used data source was local (non-public) weather stations (50.4% of the cases), followed by University of East Anglia Climatic Research Unit (CRU) datasets with 13.7% (10.9% for CRU-TS [monthly time series], and 2.8% for CRU-CL [monthly climatology]). Climate model outputs were used in 14.5% of the cases: within this group, 10.5% used GCM data, 4% RCM [Regional Climate Model] data, 3.6% satellite imagery, and 2.8% WorldClim, followed by other less relevant sources. The Global Precipitation Climatology Project (GPCP) (Adler et al., 2003; Huffman et al., 2009), the Global Precipitation Climatology Centre (GPCC) (Schneider et al., 2010) and the Global Historical Climatology Network (GHCN, (Peterson and Vose, 1997)) were rarely reported overall (0.4% each).

<Insert Figure 4 here>

The future climate data used was found to be less variable overall, with only 7 different types of data employed in the 125 cases citing some type of future climate data (Figure 4B). Out of these 125, only one study did not clearly state which type of climate data was used. The vast majority of cases (42. 9%) used GCM data “as is” (AI GCM), meaning that predictions on agricultural yields were based on predicted changes at coarse resolution (~100 km). All other studies involved some type of downscaling, except those that employed the systematic changes approach (SC variables), which can be assumed to be sensitivity analyses rather than impact studies. RCMs (Regional Climate Models) were the most common way of downscaling GCMs, cited in 19% of the studies, followed by statistical downscaling with 17.5% (SD GCM, (Tabor and Williams, 2010)), and pattern scaling with 8.7% (PS GCM, (Mitchell et al., 2004)) (Figure 4B).

Uncertainty, as measured by the number of different future scenarios used (combinations of emissions scenarios and climate models) was explored in only 36.5% of the studies. Additionally, the average number of scenarios per study (rounded to the closest integer) was 3, indicating that climate uncertainties are barely (if at all) studied in agricultural science and highlighting a knowledge gap in agricultural research, an issue previously raised and discussed by other authors (Challinor et al., 2009b; Challinor and Wheeler, 2008), although some studies addressing this aspect are underway (C. Rosenzweig, personal communication).

4.2. Robustness of existing weather station networks

The sensitivities of the network to information loss were found overall to be low. Nevertheless, certain areas, variables and months were found highly sensitive. Agricultural lands (Ramankutty et al., 2008), as visually inspected, are in general less sensitive to data loss than non-agricultural lands. Interpolations' performance varied depending upon the variable, month and parameter used to evaluate them (i.e. R^2, RMSE, and S), but were consistent, statistically significant ($p<0.0001$) and with variability (of R^2, RMSE, and S) between 10–15% in the worst cases. Rainfall presented the lowest R^2 values (Figure 5), particularly in the months of April to August, during which there was a higher variability in the R^2 value and the values reached the absolute minima (0.8). Although it is possible that a high number of weather stations per unit area can improve accuracy, it does not seem to happen in all variables, areas and/or months.

<Insert Figure 5 here>

The DJF period presented significantly lower variability and more predictive power, probably due to overall low climate variability (Cooper et al., 2008). Interestingly, maximum and minimum temperatures showed different interpolation accuracies, even though they were measured in the same places. Maximum RMSE for temperatures was up to 1.7°C, whilst for precipitation it was up to 100 mm/year, as seen in the evaluation data. The effect of geography and the difficulty of fitting unique and complex landscape features cause errors, leading to high standard deviations in some areas (Figure 6). In the highlands of Eastern Africa, particularly in the states of Benshangul-Gumaz, Addis Ababa and Southern Nations in Ethiopia, the central areas of the Eastern and Coast States in Kenya, and the very centre of Tanzania (i.e. regions of Morogoro, Dodoma and Manyara) between-fold variability was found to be high (above 150 mm/year).

<Insert Figure 6 here>

Over IGP, the largest variability was found in the coastal areas of Maharashtra, Karnataka and Kerala in India, where rainfall deviation was up to 600 mm/year, and in Nepal (districts of Gorka, Dhawalagiri, and Lumbini), where rainfall variability can go up to 1,000 mm/year, and temperature uncertainties up to 3°C, probably due to the combined effect of a more complex climate in the Himalayas and low weather station density.

4.3. Accuracy of climate model outputs

4.3.1. Ability to represent mean climate

As expected, the climate models' skill varied on a variable, country and region basis, with certain identifiable patterns (Figure 7, 8). The GCMs represent the observed climatology from weather stations (i.e. WCL-WS, GHCN-CL and GSOD-CL) more poorly than they do interpolated climatology (i.e. WCL-IS, CRU-IS), mainly because GCMs do not account for local-scale variability (Boo et al., 2011). In a broad sense, we found that the more complex the topography, the lower the skill of the GCMs (Gallée et al., 2004; Joubert et al., 1999). We also observed that GCM skill decreased according to the complexity of the variable, with the maximum skill displayed for mean temperatures, followed by temperature range and finally by precipitation. These results agree with those of other studies (Gleckler et al., 2008; Masson and Knutti, 2011; Pierce et al., 2009).

Annual precipitation fit in IGP and WAF was observed to dip as low as 0 in some cases, with a considerable number of cases (23% for WCL-WS, 27% for GHCN-CL and 63% for GSOD-CL) presenting very low adjustment ($R^2 < 0.5$) (Figure 7). In Mali, Niger, India and Bangladesh, model skill in representing precipitation, compared to weather station measurements, was consistently low, an issue also reported in other studies (Douglass et al., 2008; Gleckler et al., 2008; Reichler and Kim, 2008). The Bergen Climate Model (BCCR-BCM2.0) and the INM-CM3.0 model showed very poor performance ($R^2<0.5$) in more than 25% of the countries when compared with WCL-WS, GHCN-CL and GSOD-CL, while the climate model GISS-ModelE (Hansen et al., 2007) presented the poorest performance.

<Insert Figure 7 here>

When compared with interpolated climatology (i.e. WCL-IS, CRU-IS), annual precipitation R^2 values varied from 0.383 (GISS-ModelE-R in Uganda) to 0.998 (IAP-FGOALS1.0-G in Burkina Faso), whilst for mean temperatures the R^2 varied from 0.195 (GISS-ModelE-R in Nepal) to 0.999 (MIUB-ECHO-G in Burkina Faso), and for temperature range the values were observed between 0.386 (CCCMA-CGCM3.1-T47 in Senegal) to 0.9998 (MPI-ECHAM5 in Burkina Faso) (Figure 7).

<Insert Figure 8 here>

In Ethiopia, mean temperature correlations were lower compared to other countries, despite the relative high density of stations in that area (data not shown). In Senegal, diurnal temperature range was found to be very poorly fitted, particularly for the CCCMA models (Figure 8). This result contrasts with that of other studies, which have marked the CCCMA models as the most skilled (Gleckler et al., 2008; Jun et al., 2008). The ability of GCMs to represent mean climate patterns over a year was neither uniform nor consistent (Table 3), with the lowest performance being observed for precipitation in the DJF period (large number of cases with $R^2<0.5$, and few cases with $R^2>0.8$). Performance for temperature range showed almost no cases with $R^2<0.5$, but fewer cases with $R^2>0.8$ than for mean temperatures (Table 3).

<Insert Table 3 here>

4.3.2. Ability to represent interannual variability

R square values were above 0.8 in a large number of gridcells (>50%) for all GCMs for both variables (rainfall, mean temperature) (data not shown); however, there were large rainfall biases in GCM outputs (Figure 9, 10), in some cases between 1,000 and 2,500 mm/year, depending on the GCM. These areas were located in Nepal, northern India and EAF. Most of the models' biases were wet-biases (Figure 10) which were found throughout the whole analysis domain, but they were particularly strong over IGP in the models CCCMA-CGCM3.1-T47, CSIRO-Mk3.0 and –Mk3.5, GFDL-CM2.0, all NASA-GISS models, and both UKMO-HadCM3 and –HadGEM1, whereas the opposite signal was observed over the same area for the models MIROC3.2.-HIRES, NCAR-CCSM3.0, INGV-ECHAM4, CNRM-CM3, and GFDL-CM2.1. Over WAF and EAF, almost all GCMs showed a dry-bias, with underestimations of up to 250 mm/year in some cases. Responses varied for seasonal means and totals, with the wet-season (JJA) being more sensitive to wet biases in most GCMs.

Temperature biases were also large for certain areas. In some cases, annual mean temperature biases were greater than 5°C and were observed to go up to 10°C, particularly in the Sahel and in the areas surrounding the Himalayas and the Tibetan Plateau in Nepal (Figure 11). The most evident temperature biases were found in the NASA-GISS models (GISS-AOM, GISS-ModelE-H and GISS-ModelE-R), and in INM-CM3.0, probably due to their coarse resolution. The quality of higher resolution models was in general better, but geographic trends were difficult to identify, as the locations with mean temperature were scant (7,280 locations for the whole study area). The smallest biases were observed in WAF, northern EAF and central India, where temperature biases were below 1.5°C, particularly for the models BCCR-BCM2.0, UKMO-HadCM3, NCAR-PCM1, CCCMA-CGCM3.1-T47 and MIUB-ECHO-G, some of which have been reported to perform well in tropical areas before (Gleckler et al., 2008; Jun et al., 2008). These biases were mostly concentrated in lowlands and were mostly warm-biases, except for UKMO-HadCM3 (Figure 12). Cold-biased models were usually the GISS-NASA models, MIROC3.2-MEDRES, UKMO-HadCM3, IPSL-CM4, MRI-CGCM2.3.2A and IAP-FGOALS1.0-G both for seasons (i.e. JJA, DJF, maps not shown) and for the annual mean (Figure 11, 12).

5. Discussion

5.1.Climate data and agricultural research

Although climate model data ("as is") are often preferred for impact studies, crop modellers and agricultural scientists should be cautious when developing future adaptation strategies based on crop models applied ~~over~~ using future predictions of different (and sometimes unknown) nature (Jarvis et al., 2011), given the large uncertainties regarding the agricultural system and plant responses, the underlying uncertainty related to parameterised processes, and the differences in scales, all of which are reported in the impact-assessment literature [e.g. (Challinor and Wheeler, 2008)]. This, however, does not necessarily imply that climate model data cannot or should not be used, but rather means that an adequate treatment of biases needs to be done before climate and crop models can be properly used together (Challinor et al., 2010; Osborne et al., 2007).

Our findings demonstrate that, for regional assessments where large area process-based crop models, statistical, or empirical models are to be used, products such as WorldClim (Jones and Thornton, 2003; Thornton et al., 2009) and CRU (Challinor et al., 2004) coupled with weather generation routines appear to be the best-bet approach (Challinor et al., 2004; Jones and Thornton, 2003), although climate model data can also be used with proper bias treatment (Challinor et al., 2010; Osborne et al., 2007). However, if studies are to be carried out on a site-specific scale (Parry et al., 2005), weather station data is the best means by which to calibrate the modelling approaches. While partnerships are constantly being built and this allows researchers to share data, ~~G~~currently global weather station data such as GSOD and GHCN seem to be good options in ~~such~~ cases when no other data is available, particularly when coupled with satellite data or other (country specific) historical weather records (Álvarez-Villa et al., 2010).

Agricultural research requires high quality and high resolution climatological data to yield accurate results, but to date this has been impossible to achieve at detailed scales and with sufficient coverage, partly due to the difficulty in compiling and revising field data and partly due to the limited climatology knowledge of agricultural researchers (with some exceptions). Large-scale datasets can be matched to certain crop models, mostly when these models can be applied at large scales (Challinor et al., 2010) or do not rely on a detailed calibration of varietal-level crop parameters (Lobell et al., 2011; Lobell et al., 2008). However, matching different modelling scales is not a trivial matter (Baron et al., 2005; Challinor et al., 2009a). Two options are available for matching two differing scales:

(1) Decreasing the resolution of the crop model from plot scale to large regions, at the expense of loss of detail in some processes [see (Challinor et al., 2007b; Challinor et al., 2004; Yao et al., 2007)], or
(2) Disaggregating the coarse-resolution climate data, at the expense of introducing noise and possibly propagating uncertainties present in the original climate model data (Tabor and Williams, 2010).

These two choices yield different results that need to be assessed and coupled. Climate data can be aggregated up to any scale to match any intended use (Masson and Knutti, 2011), but agricultural impacts need to be informed at an scale such that information can be used for decision making and adaptation (Jarvis et al., 2011). Hence, governments and international agencies should support common platforms through which data can be shared without restrictions between members of the research community. Best-bet methods can then be applied over such data to produce useable datasets that can be further shared, used and assessed in multidisciplinary and transdisciplinary approaches.

5.2. Robustness of existing weather station network

It is tacitly acknowledged that the use of interpolated surfaces can lead to errors and biases when these data are used for impact assessment (A. Jarvis, pers. comm.). However, we have demonstrated here that the effects on uncertainty are actually rather low in most of the cases, with very few exceptions (highlands of Ethiopia, the Himalayas, and some parts of the Sahara and Southern Africa, Figure 6).

The results of this research suggest that, despite weather station density being important, it may not be the only determining factor for a good ability to fill information gaps (Hijmans et al., 2005). Based on our results, we suggest that, in selecting locations to measure weather, the following factors be taken into account: (1) the nature of the variable (e.g. precipitation might be much more difficult to monitor than temperature), (2) the area where it is measured (topographically complex areas are much more variable), (3) the values of the variable in the areas where it is measured (high values are subjected to larger absolute errors, assuming relative errors are relatively uniform), (4) the relevance of the area for different subjects (i.e. the Sahara might be irrelevant for agriculture but can be of high relevance for other fields such as climate science, ecology or biodiversity and conservation), (5) possible errors in measurements and other underlying factors that can influence the measurability or correctness of estimates of a particular variable, and (6) possible political or social constraints on access to the site. Improving weather station distribution and status, as well as improving the cross-checking, correction and validation of data collected at the different sites, is fundamental for improving climate data for agricultural impact assessment.

5.3. Global climate model accuracy and performance

5.3.1. CMIP3 climate model skill

GCM performance is highly reliant on the type of comparisons performed, on the GCM formulation and on the nature of climate conditions in the analysed areas (Gleckler et al., 2008; Masson and Knutti, 2011). Underlying factors driving GCM performance are indeed difficult to track, given the complexity of the models. IPCC 4AR (CMIP3) models showed varied performance, with a high tendency to being wet-biased and no general trend for temperature. These responses reportedly have their origin in different factors: first, some GCMs have weak forcing on sea surface temperatures (SSTs), whereas climate in Africa and Asia is strongly coupled with the Atlantic and Indian Ocean and with inland water bodies (Gallée et al., 2004; Lebel et al., 2000); second, models do not properly account for the relation between inter-annual variability, ENSO and the monsoonal winds (Gallée et al., 2004; Hulme et al., 2001); third, the resolution of the models prevents acknowledgement of local-scale land use, orographic patterns and small water bodies (Hudson and Jones, 2002); fourth, cloud thickness and latent heat and moisture flux between clouds has not been properly resolved in the models (Gallée et al., 2004); and fifth, convective parameterisations produce an early onset of the seasonal rains and over-prediction of wet days and high-rainfall events (Gallée et al., 2004).

The NASA models GISS-ModelE (-R and -H) consistently presented very low predictive ability, mainly because of the models' coarse spatial resolution in conjunction with the reasons mentioned above (Hansen et al., 2007). These results agree with those of Gleckler et al. (2008), who reported that NCAR-PCM1, GISS-ModelE (-R and –H) and GISS-AOM models are the worst performing in the 24 GCMs of the CMIP3 ensemble. Similar results are reported by other authors that have assessed this or similar model ensembles (Jun et al., 2008; Pierce et al., 2009). Lack of detail in land use and land use changes (Eltahir and Gong, 1996), monsoon winds (Eltahir and Gong, 1996; Gallée et al., 2004), and sea surface temperature anomalies (SSTs) of the Atlantic and the Indian Oceans (Lebel et al., 2000; Sun et al., 1999) also causes the scales at which climate model information is robust to be varied (Masson and Knutti, 2011), and prevents local scale seasonal weather patterns from being modelled consistently (Douglass et al., 2008; Hansen et al., 2007).

5.3.2. Plugging climate model data into agricultural research

GCMs do not provide realistic representations of climate conditions in a particular site, but rather provide estimated conditions for a large area. Our results, in agreement with those from the agricultural community (Baron et al., 2005; Challinor et al., 2003) and the climate community (Jun et al., 2008; Masson and Knutti, 2011), indicate that climate model outputs cannot be input directly into plot-scale (agricultural) models, but support the idea that higher resolution climate modelling largely improves results. Either the CMIP3 (assessed here) or the upcoming CMIP5 (being released at the moment) (Moss et al., 2010) climate model outputs can be adequately used in agricultural modelling if: (1) the scales between the models are matched (see Sect. 5.1), (2) skill of models is assessed and ways to create robust model ensembles are defined, (3) uncertainty and model spread are quantified in a robust way, and (4) decision making in the context of uncertainty is fully understood.

Producing robust (i.e. skilled and certain) ensembles for agriculture is not an easy task, mainly because of the scales at which these have been found to be robust (Masson and Knutti, 2011). Opinions are contrasting: some authors support sub-selecting models based upon performance under present conditions (Matsueda and Palmer, 2011; Pierce et al., 2009), calculating a mean ensemble by weighting models based on skill (Matsueda and Palmer, 2011; Walsh et al., 2008), while others advocate using ~~as many as~~all available models with no-weighting at all (Reifen and Toumi, 2009). We suggest that until sensitivities of agricultural models to ensemble spread are fully explored (Baigorria et al., 2007), the full CMIP3 (or CMIP5) ensembles should be used.

Strategies for combining plot-scale and large-scale models and for optimising the overall result (including estimation of uncertainties derived from the scale-matching process) need to be further researched. The potential of high-quality and less uncertain climate predictions of current and future climate conditions for agricultural research is expected to have a direct impact on decision-making at different levels and for different purposes: to improve yields on the farm, to direct country level policies and investment, to define research foci, to direct international agencies' investments, and to clarify global greenhouse emissions limits and commitments (Challinor et al., 2009a; Funke and Paetz, 2011; IPCC, 2007).

6. Conclusions

A thorough analysis of different aspects of climate data for agricultural applications was performed. All topics addressed here are of high relevance to agricultural applications, particularly in the global tropics. Several important points were raised: (1) spatial scale is the most important issue for agricultural researchers, as they prefer~~red~~ to use monthly products with higher resolution rather than daily products with very low spatial resolution, or else limit~~ed~~ their areas of study to field plots; (2) the sensitivities of Sub-Saharan African and

Southeast Asian climate to data loss and poor availability were found to not be limiting factors for the region, with the exceptions of mountainous areas in Nepal and Ethiopia; and (3) climate modelling, although constantly improving and useful, still requires considerable future development.

As such, CMIP3 GCMs can be used with a certain degree of confidence to represent large-area climate conditions for some areas and periods. In areas where predictions lack enough skill for agricultural modelling, models can be bias-corrected using different methods [see (Challinor et al., 2009a; Hawkins et al., 2011; Reifen and Toumi, 2009)]. Whilst model skill is expected to improve with the upcoming IPCC Fifth Assessment Report, climate model ensembles as well as different methods for 'calibrating' (i.e. pre-processing for input into crop models) climate model data both need to be used, as uncertainties go beyond those derived from emissions scenarios (Hawkins et al., 2011). The proper usage of climate projections for agricultural impact assessment is of paramount importance in order to properly inform adaptation.

Finally, it is critical to understand the implications of all this to agriculture. Crops are sensitive to shortages in water and heat stresses during key periods during their development (i.e. flowering, fruit filling). Therefore, lack of skill in representing seasonal and inter-annual variability is expected to produce a significant obstacle to agricultural impact assessment of climate change; several examples in the literature exist that illustrate this (Baigorria et al., 2008; Baigorria et al., 2007). The importance of this factor depends on the strength of the climate signal on yields and the variables that drive ~~such~~ this signal. Future impact assessments need to take into account input data and climate model data inaccuracies, sensitivities and uncertainties; make their own assessments of the inaccuracies and uncertainties; and comprehensively quantify and report uncertainties in the impact assessment process.

Acknowledgments

Authors thank Charlotte Lau, from the International Center for Tropical Agriculture (CIAT) for the useful comments in an earlier version of this manuscript, ~~and~~ Peter G. Jones and Robert J. Hijmans for the inspiration and the previous assembling of data (WorldClim weather stations) that were used as part of this study, and the two anonymous reviewers for their useful comments and encouragement. Authors also thank the Theme 1 of the CGIAR Research Program on Climate Change, Agriculture and Food Security (CCAFS) for providing the funds for carrying out this study.

References

Adam, M., Van Bussel, L.G.J., Leffelaar, P.A., Van Keulen, H. and Ewert, F., 2011. Effects of modelling detail on simulated potential crop yields under a wide range of climatic conditions. Ecological Modelling, 222(1): 131-143.

Adler, R.F. et al., 2003. The Version-2 Global Precipitation Climatology Project (GPCP) Monthly Precipitation Analysis (1979–Present). Journal of Hydrometeorology, 4(6): 1147-1167.

Aggarwal, P.K., 2008. Global climate change and Indian agriculture: impacts, adaptation and mitigation. Indian Journal of Agricultural Sciences, 78(10): 911-919.

Aggarwal, P.K., Joshi, P.K., Ingram, J.S.I. and Gupta, R.K., 2004. Adapting food systems of the Indo-Gangetic plains to global environmental change: key information needs to improve policy formulation. Environmental Science & Policy, 7(6): 487-498.

Aggarwal, P.K., Kalra, N., Chander, S. and Pathak, H., 2006. InfoCrop: A dynamic simulation model for the assessment of crop yields, losses due to pests, and environmental impact of agro-ecosystems in tropical environments. I. Model description. Agricultural Systems, 89(1): 1-25.

Álvarez-Villa, O.D., Vélez, J.I. and Poveda, G., 2010. Improved long-term mean annual rainfall fields for Colombia. International Journal of Climatology: n/a-n/a.

Baigorria, G.A., Jones, J.W. and O'Brien, J.J., 2008. Potential predictability of crop yield using an ensemble climate forecast by a regional circulation model. Agricultural and Forest Meteorology, 148(8-9): 1353-1361.

Baigorria, G.A., Jones, J.W., Shin, D.W., Mishra, A. and Brien, J.J., 2007. Assessing uncertainties in crop model simulations using daily bias-corrected Regional Circulation Model outputs. Climate Research, 34(3): 211-222.

Baron, C. et al., 2005. From GCM grid cell to agricultural plot: scale issues affecting modelling of climate impact. Philosophical Transactions of the Royal Society B: Biological Sciences, 360(1463): 2095-2108.

Barrios, S., Ouattara, B. and Strobl, E., 2008. The impact of climatic change on agricultural production: Is it different for Africa? Food Policy, 33(4): 287-298.

Battisti, D.S. and Naylor, R.L., 2009. Historical Warnings of Future Food Insecurity with Unprecedented Seasonal Heat. Science, 323(5911): 240-244.

Begert, M., Zenklusen, E., berli, C., Appenzeller, C. and Klok, L., 2008. An automated procedure to detect discontinuities; performance assessment and application to a large European climate data set. Meteorologische Zeitschrift, 17(5): 663-672.

Boo, K.-O., Martin, G., Sellar, A., Senior, C. and Byun, Y.-H., 2011. Evaluating the East Asian monsoon simulation in climate models. J. Geophys. Res., 116(D1): D01109.

Boote, K.J., Jones, J.W. and Pickering, N.B., 1996. Potential uses and limitations of crop models. Agronomy Journal, 88(5): 704-716.

Buytaert, W., Célleri, R. and Timbe, L., 2009. Predicting climate change impacts on water resources in the tropical Andes: Effects of GCM uncertainty. Geophys. Res. Lett., 36(7): L07406.

Byjesh, K., Kumar, S. and Aggarwal, P., 2010. Simulating impacts, potential adaptation and vulnerability of maize to climate change in India. Mitigation and Adaptation Strategies for Global Change, 15(5): 413-431.

Challinor, A., Wheeler, T., Garforth, C., Craufurd, P. and Kassam, A., 2007a. Assessing the vulnerability of food crop systems in Africa to climate change. Climatic Change, 83(3): 381-399.

Challinor, A.J., Ewert, F., Arnold, S., Simelton, E. and Fraser, E., 2009a. Crops and climate change: progress, trends, and challenges in simulating impacts and informing adaptation. Journal of Experimental Botany, 60(10): 2775-2789.

Challinor, A.J. et al., 2009b. Methods and Resources for Climate Impacts Research. Bulletin of the American Meteorological Society, 90(6): 836-848.

Challinor, A.J., Simelton, E.S., Fraser, E.D.G., Hemming, D. and Collins, M., 2010. Increased crop failure due to climate change: assessing adaptation options using models and socio-economic data for wheat in China. Environmental Research Letters, 5(3): 034012.

Challinor, A.J., Slingo, J.M., Wheeler, T.R., Craufurd, P.Q. and Grimes, D.I.F., 2003. Toward a Combined Seasonal Weather and Crop Productivity Forecasting System: Determination of the Working Spatial Scale. Journal of Applied Meteorology, 42(2): 175-192.

Challinor, A.J. and Wheeler, T.R., 2008. Crop yield reduction in the tropics under climate change: Processes and uncertainties. Agricultural and Forest Meteorology, 148(3): 343-356.

Challinor, A.J., Wheeler, T.R., Craufurd, P.Q., Ferro, C.A.T. and Stephenson, D.B., 2007b. Adaptation of crops to climate change through genotypic responses to mean and extreme temperatures. Agriculture, Ecosystems & Environment, 119(1-2): 190-204.

Challinor, A.J., Wheeler, T.R., Craufurd, P.Q., Slingo, J.M. and Grimes, D.I.F., 2004. Design and optimisation of a large-area process-based model for annual crops. Agricultural and Forest Meteorology, 124(1-2): 99-120.

Challinor, A.J., Wheeler, T.R., Slingo, J.M. and Hemming, D., 2005. Quantification of physical and biological uncertainty in the simulation of the yield of a tropical crop using present-day and doubled CO2 climates. Philosophical Transactions of the Royal Society B: Biological Sciences, 360(1463): 2085-2094.

Chipanshi, A.C., Chanda, R. and Totolo, O., 2003. Vulnerability Assessment of the Maize and Sorghum Crops to Climate Change in Botswana. Climatic Change, 61(3): 339-360.

Collins, W.D. et al., 2006. The Community Climate System Model Version 3 (CCSM3). Journal of Climate, 19(11): 2122-2143.

Cooper, P.J.M. et al., 2008. Coping better with current climatic variability in the rain-fed farming systems of sub-Saharan Africa: An essential first step in adapting to future climate change? Agriculture, Ecosystems & Environment, 126(1-2): 24-35.

DeGaetano, A.T., 2006. Attributes of Several Methods for Detecting Discontinuities in Mean Temperature Series. Journal of Climate, 19(5): 838-853.

Delworth, T.L. et al., 2006. GFDL's CM2 Global Coupled Climate Models. Part I: Formulation and Simulation Characteristics. Journal of Climate, 19(5): 643-674.

Déqué, M., Dreveton, C., Braun, A. and Cariolle, D., 1994. The ARPEGE/IFS atmosphere model: a contribution to the French community climate modelling. Climate Dynamics, 10(4): 249-266.

Di Luzio, M., Johnson, G.L., Daly, C., Eischeid, J.K. and Arnold, J.G., 2008. Constructing Retrospective Gridded Daily Precipitation and Temperature Datasets for the Conterminous United States. Journal of Applied Meteorology and Climatology, 47(2): 475-497.

Diansky, N.A. and Zalensky, V.B., 2002. Simulation of present-day climate with a coupled Atmosphere-ocean general circulation model. Izvestiya, Atmospheric and Ocean Physiscs, 38(6): 732-747.

Douglass, D.H., Christy, J.R., Pearson, B.D. and Singer, S.F., 2008. A comparison of tropical temperature trends with model predictions. International Journal of Climatology, 28(13): 1693-1701.

El-Sharkawy, M., 2005. How can calibrated research-based models be improved for use as a tool in identifying genes controlling crop tolerance to environmental stresses in the era of genomics—from an experimentalist's perspective. Photosynthetica, 43(2): 161-176.

Eltahir, E.A.B. and Gong, C., 1996. Dynamics of Wet and Dry Years in West Africa. Journal of Climate, 9(5): 1030-1042.

FAO, 2001. FAOCLIM 2.0 A World-Wide Agroclimatic Database. In: FAO (Editor), Rome, Italy.

FAO, 2009. The State of Food and Agriculture. FAO, Rome, Italy.

Funke, M. and Paetz, M., 2011. Environmental policy under model uncertainty: a robust optimal control approach. Climatic Change, 107(3): 225-239.

Furevik, T. et al., 2003. Description and evaluation of the bergen climate model: ARPEGE coupled with MICOM. Climate Dynamics, 21(1): 27-51.

Gallée, H. et al., 2004. A high-resolution simulation of a West African rainy season using a regional climate model. J. Geophys. Res., 109(D5): D05108.

Gleckler, P.J., Taylor, K.E. and Doutriaux, C., 2008. Performance metrics for climate models. J. Geophys. Res., 113(D6): D06104.

Gordon, C. et al., 2000. The simulation of SST, sea ice extents and ocean heat transports in a version of the Hadley Centre coupled model without flux adjustments. Climate Dynamics, 16(2): 147-168.

Gordon, H.B. et al., 2002. The CSIRO Mk3 Climate System Model, Aspendale: CSIRO Atmospheric Research.
Govindan, R.B. et al., 2002. Global Climate Models Violate Scaling of the Observed Atmospheric Variability. Physical Review Letters, 89(2): 028501.
Grötzner, A., Sausen, R. and Claussen, M., 1996. The impact of sub-grid scale sea-ice inhomogeneities on the performance of the atmospheric general circulation model ECHAM3. Climate Dynamics, 12(7): 477-496.
Gualdi, S., Scoccimarro, E. and Navarra, A., 2008. Changes in Tropical Cyclone Activity due to Global Warming: Results from a High-Resolution Coupled General Circulation Model. Journal of Climate, 21(20): 5204-5228.
Hansen, J. et al., 2007. Climate simulations for 1880–2003 with GISS modelE. Climate Dynamics, 29(7): 661-696.
Hasumi, H. and Emori, S., 2004. K-1 Coupled GCM (MIROC) Description, CCSR, NIES and FRCGC.
Hawkins, E., Osborne, T.M., Ho, C.K. and Challinor, A.J., 2011. Calibration and bias correction of climate projections for crop modelling: an idealised case study over Europe. Agricultural and Forest Meteorology, in press.
Hew, C.-S., Krotkov, G. and Canvin, D.T., 1969. Effects of Temperature on Photosynthesis and CO2 Evolution in Light and Darkness by Green Leaves. Plant Physiology, 44(5): 671-677.
Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. and Jarvis, A., 2005. Very high resolution interpolated climate surfaces for global land areas. International Journal of Climatology, 25(15): 1965-1978.
Hoogenboom, G. et al., 1994. Crop Models. Decision Support System for Agrotechnology Transfer, Version 3.0 (DSSAT V3.0), vol. 2–2. International Benchmark Sites Network for Agrotechnology Transfer (IBSNAT) Project. University of Hawaii, Department of Agronomy and Soil Science, Honolulu, Hawaii.
Hudson, D.A. and Jones, R.G., 2002. Regional Climate Model Simulations of Present-day and Future Climates of Southern Africa, MetOffice, Hadley Centre, UK, Reading, UK.
Huffman, G.J., Adler, R.F., Bolvin, D.T. and Gu, G., 2009. Improving the global precipitation record: GPCP Version 2.1. Geophys. Res. Lett., 36(17): L17808.
Huffman, G.J. et al., 2007. The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-Global, Multiyear, Combined-Sensor Precipitation Estimates at Fine Scales. Journal of Hydrometeorology, 8(1): 38-55.
Hulme, M., Doherty, R., Ngara, T., New, M. and Lister, D., 2001. African climate change: 1900-2100. Climate Research, 17(2): 145-168.
Huntingford, C., Hugo Lambert, F., Gash, J.H.C., Taylor, C.M. and Challinor, A.J., 2005. Aspects of climate change prediction relevant to crop productivity. Philosophical Transactions of the Royal Society B: Biological Sciences, 360(1463): 1999-2009.
Hutchinson, M.F., 1995. Interpolating mean rainfall using thin plate smoothing splines. International Journal of Geographical Information Systems, 9(4): 385 - 403.
IPCC, 2007. IPCC Fourth Assessment Report: Climate Change 2007 (AR4). IPCC, Geneva, Switzerland.
Jagtap, S.S. and Jones, J.W., 2002. Adaptation and evaluation of the CROPGRO-soybean model to predict regional yield and production. Agriculture, Ecosystems & Environment, 93(1-3): 73-85.
Jarvis, A. et al., 2011. An Integrated Adaptation and Mitigation Framework for Developing Agricultural Research: Synergies and Trade-offs. Experimental Agriculture, 47(Special Issue 02): 185-203.
Jarvis, A., Ramirez, J., Anderson, B., Leibing, C. and Aggarwal, P., 2010. Scenarios of Climate Change Within the Context of Agriculture. Climate Change and Crop Production. CAB International.
Jarvis, A., Reuter, H.I., Nelson, A. and Guevara, E., 2008. Hole-filled seamless SRTM data V4. In: I.C.f.T.A. (CIAT) (Editor). International Center for Tropical Agriculture (CIAT).

Johns, T.C. et al., 2006. The New Hadley Centre Climate Model (HadGEM1): Evaluation of Coupled Simulations. Journal of Climate, 19(7): 1327-1353.

Jones, P.G. and Thornton, P.K., 1999. Fitting a third-order Markov rainfall model to interpolated climate surfaces. Agricultural and Forest Meteorology, 97(3): 213-231.

Jones, P.G. and Thornton, P.K., 2003. The potential impacts of climate change on maize production in Africa and Latin America in 2055. Global Environmental Change, 13(1): 51-59.

Jones, P.G., Thornton, P.K. and Heinke, J., 2009. Generating characteristic daily weather data using downscaled climate model data from the IPCC's Fourth Assessment, ILRI, Nairobi, Kenya.

Joubert, A.M., Katzfey, J.J., McGregor, J.L. and Nguyen, K.C., 1999. Simulating midsummer climate over southern Africa using a nested regional climate model. J. Geophys. Res., 104(D16): 19015-19025.

Jun, M., Knutti, R. and Nychka, D.W., 2008. Spatial Analysis to Quantify Numerical Model Bias and Dependence. Journal of the American Statistical Association, 103(483): 934-947.

Jungclaus, J.H. et al., 2006. Ocean Circulation and Tropical Variability in the Coupled Model ECHAM5/MPI-OM. Journal of Climate, 19(16): 3952-3972.

Lane, A. and Jarvis, A., 2007. Changes in Climate will modify the Geography of Crop Suitability: Agricultural Biodiversity can help with Adaptation. Journal of the Semi-Arid Tropics, 4(1): 1-12.

Lebel, T., Delclaux, F., Le Barbé, L. and Polcher, J., 2000. From GCM scales to hydrological scales: rainfall variabiliy in West Africa. Stochastic Environmental Research and Risk Assessment, 14: 275-295.

Liu, J. et al., 2008. A spatially explicit assessment of current and future hotspots of hunger in Sub-Saharan Africa in the context of global change. Global and Planetary Change, 64(3-4): 222-235.

Lobell, D.B., Banziger, M., Magorokosho, C. and Vivek, B., 2011. Nonlinear heat effects on African maize as evidenced by historical yield trials. Nature Clim. Change, 1(1): 42-45.

Lobell, D.B. and Burke, M.B., 2010. On the use of statistical models to predict crop yield responses to climate change. Agricultural and Forest Meteorology, 150(11): 1443-1452.

Lobell, D.B. et al., 2008. Prioritizing Climate Change Adaptation Needs for Food Security in 2030. Science, 319(5863): 607-610.

Marti, O. et al., 2005. The new IPSL climate system model: IPSL-CM4, Institute Pierre Simon Laplace.

Masson, D. and Knutti, R., 2011. Spatial-Scale Dependence of Climate Model Performance in the CMIP3 Ensemble. Journal of Climate, 24(11): 2680-2692.

Matsueda, M. and Palmer, T.N., 2011. Accuracy of climate change predictions using high resolution simulations as surrogates of truth. Geophys. Res. Lett., 38(5): L05803.

Mitchell, T.D., Carter, T.R., Jones, P.D., Hulme, M. and New, M., 2004. A comprehensive set of high-resolution grids of monthly climate for Europe and the globe: the observed record (1901-2000) and 16 scenarios (2001-2100). Tyndall Centre for Climate Change Research, Oxford, UK.

Moss, R.H. et al., 2010. The next generation of scenarios for climate change research and assessment. Nature, 463(7282): 747-756.

NCDC, 2011. National Climatic Data Center, NESDIS, NOAA, U.S. Department of Commerce. Global Surface Summary of the Day - GSOD. In: NCDC, NIESDIS, NOAA and USDC (Editors). National Climatic Data Center (NCDC), Asheville, NC, US.

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

Niu, X., Easterling, W., Hays, C.J., Jacobs, A. and Mearns, L., 2009. Reliability and input-data induced uncertainty of the EPIC model to estimate climate change impact on sorghum yields in the U.S. Great Plains. Agriculture, Ecosystems & Environment, 129(1-3): 268-276.

Oldeman, L.R., 1988. An Agroclimatic Characterization of Madagascar. International Soil Reference and Information Centre, Wageningen, Netherlands.

Osborne, T.M., Lawrence, D.M., Challinor, A.J., Slingo, J.M. and Wheeler, T.R., 2007. Development and assessment of a coupled crop–climate model. Global Change Biology, 13(1): 169-183.

Parry, M., Rosenzweig, C. and Livermore, M., 2005. Climate change, global food supply and risk of hunger. Philosophical Transactions of the Royal Society B: Biological Sciences, 360(1463): 2125-2138.

PCMDI, 2007. IPCC Model Output. Available at: http://www.pcmdi.llnl.gov/ipcc/about_ipcc.php.

Peterson, T.C. and Easterling, D.R., 1994. Creation of homogenous composite climatological reference series. International Journal of Climatology, 14: 671-679.

Peterson, T.C. and Vose, R.S., 1997. An Overview of the Global Historical Climatology Network Temperature Database. Bulletin of the American Meteorological Society, 78(12): 2837-2849.

Pielke, S.R.A. et al., 2007. A new paradigm for assessing the role of agriculture in the climate system and in climate change. Agricultural and Forest Meteorology, 142(2-4): 234-254.

Pierce, D.W., Barnett, T.P., Santer, B.D. and Gleckler, P.J., 2009. Selecting global climate models for regional climate change studies. Proceedings of the National Academy of Sciences, 106(21): 8441-8446.

Prasad, P.V.V., Boote, K.J., Allen, L.H. and Thomas, J.M.G., 2002. Effects of elevated temperature and carbon dioxide on seed-set and yield of kidney bean (Phaseolus vulgaris L.). Global Change Biology, 8(8): 710-721.

Quiggin, J., 2008. Uncertainty and Climate Change Policy. Economic Analysis & Policy, 38(2): 203-210.

Ramankutty, N., Evan, A.T., Monfreda, C. and Foley, J.A., 2008. Farming the planet: 1. Geographic distribution of global agricultural lands in the year 2000. Global Biogeochem. Cycles, 22(1): GB1003.

Reichler, T. and Kim, J., 2008. How Well Do Coupled Models Simulate Today's Climate? Bulletin of the American Meteorological Society, 89(3): 303-311.

Reifen, C. and Toumi, R., 2009. Climate projections: Past performance no guarantee of future skill? Geophys. Res. Lett., 36(13): L13704.

Russell, G., Miller, J.R. and Rind, D., 1995. A Coupled Atmosphere-Ocean Model for Transient Climate Change Studies. Atmosphere-Ocean, 33(4): 683-730.

Salas-Mélia, D. et al., 2005. Description and validation of the CNRM-CM3 global coupled model, CNRM.

Schlenker, W. and Lobell, D.B., 2010. Robust negative impacts of climate change on African agriculture. Environmental Research Letters, 5(1): 014010.

Schmidt, G.A. et al., 2006. Present-Day Atmospheric Simulations Using GISS ModelE: Comparison to In Situ, Satellite, and Reanalysis Data. Journal of Climate, 19(2): 153-192.

Schneider, U., Becker, A., Meyer-Christoffer, A., Ziese, M. and Rudolf, B., 2010. Global Precipitation Analysis Products of the GPCC, Global Preciptation Climatology Centre (GPCC).

Scinocca, J.F., McFarlane, N.A., Lazare, M., Li, J. and Plummer, D., 2008. The CCCma third generation AGCM and its extension into the middle atmosphere. Atmos. Chem. Phys. Discuss., 8(2): 7883-7930.

Soltani, A., Meinke, H. and de Voil, P., 2004. Assessing linear interpolation to generate daily radiation and temperature data for use in crop simulations. European Journal of Agronomy, 21(2): 133-148.

Soria-Auza, R.W. et al., 2010. Impact of the quality of climate models for modelling species occurrences in countries with poor climatic documentation: a case study from Bolivia. Ecological Modelling, 221(8): 1221-1229.

Sun, L., Semazzi, F.H.M., Giorgi, F. and Ogallo, L., 1999. Application of the NCAR Regional Climate Model to eastern Africa 2. Simulation of interannual variability of short rains. J. Geophys. Res., 104(D6): 6549-6562.

Tabor, K. and Williams, J.W., 2010. Globally downscaled climate projections for assessing the conservation impacts of climate change. Ecological Applications, 20(2): 554-565.

Thornton, P.K., Jones, P.G., Alagarswamy, G. and Andresen, J., 2009. Spatial variation of crop yield response to climate change in East Africa. Global Environmental Change, 19(1): 54-65.

Thornton, P.K., Jones, P.G., Ericksen, P.J. and Challinor, A.J., 2011. Agriculture and food systems in sub-Saharan Africa in a 4°C+ world. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 369(1934): 117-136.

Thuzar, M., Puteh, A.B., Abdullah, N.A.P., Lassim, M.B.M. and Jusoff, K., 2010. The Effects of Temperature Stress on the Quality and Yield of Soya Bean [(Glycine max L.) Merrill.]. Journal of Agricultural Science, 2(1): 172-179.

Trnka, M., Dubrovský, M., Semerádová, D. and Žalud, Z., 2004. Projections of uncertainties in climate change scenarios into expected winter wheat yields. Theoretical and Applied Climatology, 77(3): 229-249.

Uppala, S.M. et al., 2005. The ERA-40 re-analysis. Quarterly Journal of the Royal Meteorological Society, 131(612): 2961-3012.

van Ittersum, M.K., Howden, S.M. and Asseng, S., 2003. Sensitivity of productivity and deep drainage of wheat cropping systems in a Mediterranean environment to changes in CO2, temperature and precipitation. Agriculture, Ecosystems & Environment, 97(1-3): 255-273.

Walker, W.E. et al., 2003. Defining Uncertainty: A Conceptual Basis for Uncertainty Management in Model-Based Decision Support. Integrated Assessment, 4(1): 5-17.

Walsh, J.E., Chapman, W.L., Romanovsky, V., Christensen, J.H. and Stendel, M., 2008. Global Climate Model Performance over Alaska and Greenland. Journal of Climate, 21(23): 6156-6174.

Washington, R. et al., 2006. African Climate Change: Taking the Shorter Route. Bulletin of the American Meteorological Society, 87(10): 1355-1366.

Washington, W.M. et al., 2000. Parallel climate model (PCM) control and transient simulations. Climate Dynamics, 16(10): 755-774.

Wilby, R.L. et al., 2009. A review of climate risk information for adaptation and development planning. International Journal of Climatology, 29(9): 1193-1215.

Xiong, W., Holman, I., Conway, D., Lin, E. and Li, Y., 2008. A crop model cross calibration for use in regional climate impacts studies. Ecological Modelling, 213(3-4): 365-380.

Yao, F., Xu, Y., Lin, E., Yokozawa, M. and Zhang, J., 2007. Assessing the impacts of climate change on rice yields in the main rice areas of China. Climatic Change, 80(3): 395-409.

Yongqiang, Y., Xuehong, Z. and Yufu, G., 2004. Global coupled ocean-atmosphere general circulation models in LASG/IAP. Advances in Atmospheric Sciences, 21(3): 444-455.

Yukimoto, S. et al., 2001. The New Meteorological Research Institute Coupled GCM (MRI-CGCM2) Model Climate and Variability. Papers in Meteorology and Geophysics, 51(2): 47-88.

Figure captions

Figure 1 Cascade of constraints to climate data, as normally observed in agricultural impact assessment

Figure 2 Areas of study. Bold-outlined areas indicate the areas on which the study focused (SN: Senegal, ML: Mali, NE: Niger, BF: Burkina Faso, GH: Ghana, UG: Uganda, ET: Ethiopia, KE: Kenya, TZ: Tanzania, NP: Nepal, BD: Bangladesh, IN: India)

Figure 3 Topics treated in the analysed agricultural studies. WG: weather generators.

Figure 4 Frequency of use of the different data sources in agricultural studies. A. Present-day climates. B. Future climates. Datasets acronyms are as follows: CRU-TS: Climatic Research Unit monthly time series product at 0.5 degree, GCM: global climate model output, RCM: regional climate model, CRU-CL: CRU monthly climatology product at 10 arc-minute, MARS: Data from the MARS European project, GSOD: Global summary of the day, ARTES: Africa rainfall and temperature evaluation system, VEMAP: United States comprehensive dataset, ATEAM: Advanced Terrestrial Ecosystem Analysis and Modelling, PRISM: United States dataset, GPCP: Global Precipitation Climatology Project, GPCC: Global Precipitation Climatology Centre, GHCN: Global Historical Climatology Network, AI GCM: GCM data "as is", SD GCM: statistically downscaled GCM, PS GCM: pattern scaled GCM, WG GCM: GCM data through a weather generator, SC Variables: systematic changes in target key variables, Unclear: not specified clearly in study, ARPEGE: the ARPEGE Atmospheric GCM (Déqué et al., 1994).

Figure 5 Performance of the interpolations for all variables and months as measured by the R-square value. A. Rainfall, B. Mean temperature, C. Maximum temperature, D. Minimum temperature

Figure 6 Uncertainties in WorldClim expressed as standard deviations from the mean of the 100 cross-validated folds for (A) total annual rainfall (in mm), and (B) annual mean temperature (in ºC).

Figure 7 Comparison (R-square based) of observed climatology (CL-WS [w], GHCN-CL [g] and GSOD-CL [o]) and each of the GCMs (GCM-CL) for each of the countries in the study area for mean temperature (top), temperature range (middle) and precipitation (bottom), for the annual and two seasonal (DJF, JJA) means or totals. All R^2 values were statistically significant at $p<0.0001$

Figure 8 Comparison (R-square based) of interpolated climatology (i.e. CRU-IS [c], WCL-IS [w]), and each of the GCMs (GCM-CL) for each of the countries in the study area for mean temperature (top), temperature range (middle) and precipitation (bottom) for the annual mean or total and two seasons (DJF, JJA). All R^2 values were statistically significant at $p<0.001$.

Figure 9 Root mean squared error (RMSE), in millimetres, between observed (GHCN-TS) and GCM (GCM-TS) time series, for the 24 GCMs in Table 2, for annual total rainfall between the years 1961-1990.

Figure 10 Mean bias of GCM (GCM-TS) time series compared to observed time series (GHCN-TS), for the 24 GCMs in Table 2, for annual total rainfall between the years 1961-1990. Bias is expressed as the slope of the regression curve between observed and climate-model series. Values below 1 (light grey areas) indicate that GCMs are wet-biased, whereas values above 1 (dark grey areas) indicate that GCMs are dry-biased.

Figure 11 Root mean squared error (RMSE), in Celsius degree, between observed (GHCN-TS) and GCM (GCM-TS) time series, for the 24 GCMs in Table 2, for annual mean temperature between the years 1961-1990

Figure 12 Mean bias of GCM (GCM-TS) time series compared to observed time series (GHCN-TS), for the 24 GCMs in Table 2, for annual mean temperature between the years 1961-1990. Bias is expressed as the slope of the regression curve between observed and climate-model series. Values below 1 (light grey areas) indicate that GCMs are warm-biased, whereas values above 1 (dark grey areas) indicate that GCMs are cold-biased.

Table 1 Number of locations per data source (global)

Source*	Precipitation stations	Mean temperature stations	Min., Max. temperature stations	Period
GHCN v2	20,590	7,280	4,966	1950-2000
WMO CLINO	4,261	3,084	2,504	1961-1990
FAOCLIM 2.0	27,372	20,825	11,543	1960-1990
CIAT	18,895	13,842	5,321	1950-2000

*GHCN v2: Global Historical Climatology Network version 2 (Peterson and Vose, 1997); WMO CLINO: World Meteorological Organization Climatology Normals; FAOCLIM 2.0: Food and Agriculture Organization of the United Nations Agro-Climatic database (FAO, 2001); CIAT: Database assembled by Peter J. Jones at the International Center for Tropical Agriculture (CIAT).

Table 2 Available GCMs, resolutions, and main references

Model	Country	Atmosphere	Ocean	Reference
BCCR-BCM2.0	Norway	T63, L31	1.5x0.5, L35	(Furevik et al., 2003)
CCCMA-CGCM3.1 (T47)	Canada	T47 (3.75x3.75), L31	1.85x1.85, L29	(Scinocca et al., 2008)
CCCMA-CGCM3.1 (T63)	Canada	T63 (2.8x2.8), L31	1.4x0.94, L29	(Scinocca et al., 2008)
CNRM-CM3	France	T63 (2.8x2.8), L45	1.875x(0.5-2), L31	(Salas-Mélia et al., 2005)
CSIRO-Mk3.0	Australia	T63, L18	1.875x0.84, L31	(Gordon et al., 2002)
CSIRO-Mk3.5	Australia	T63, L18	1.875x0.84, L31	(Gordon et al., 2002)
GFDL-CM2.0	USA	2.5x2.0, L24	1.0x(1/3-1), L50	(Delworth et al., 2006)
GFDL-CM2.1	USA	2.5x2.0, L24	1.0x(1/3-1), L50	(Delworth et al., 2006)
GISS-AOM	USA	4x3, L12	4x3, L16	(Russell et al., 1995)
GISS-MODEL-EH	USA	5x4, L20	5x4, L13	(Schmidt et al., 2006)
GISS-MODEL-ER	USA	5x4, L20	5x4, L13	(Schmidt et al., 2006)
IAP-FGOALS1.0-G	China	2.8x2.8, L26	1x1, L16	(Yongqiang et al., 2004)
INGV-ECHAM4	Italy	T42, L19	2x(0.5-2), L31	(Gualdi et al., 2008)
INM-CM3.0	Russia	5x4, L21	2.5x2, L33	(Diansky and Zalensky, 2002)
IPSL-CM4	France	2.5x3.75, L19	2x(1-2), L30	(Marti et al., 2005)
MIROC3.2-HIRES	Japan	T106, L56	0.28x0.19, L47	(Hasumi and Emori, 2004)
MIROC3.2-MEDRES	Japan	T42, L20	1.4x(0.5-1.4), L43	(Hasumi and Emori, 2004)
MIUB-ECHO-G	Germany/Korea	T30, L19	T42, L20	(Grötzner et al., 1996)
MPI-ECHAM5	Germany	T63, L32	1x1, L41	(Jungclaus et al., 2006)
MRI-CGCM2.3.2A	Japan	T42, L30	2.5x(0.5-2.0)	(Yukimoto et al., 2001)
NCAR-CCSM3.0	USA	T85L26, 1.4x1.4	1x(0.27-1), L40	(Collins et al., 2006)
NCAR-PCM1	USA	T42 (2.8x2.8), L18	1x(0.27-1), L40	(Washington et al., 2000)
UKMO-HADCM3	UK	3.75x2.5, L19	1.25x1.25, L20	(Gordon et al., 2000)
UKMO-HADGEM1	UK	1.875x1.25, L38	1.25x1.25, L20	(Johns et al., 2006)

Table 3 Summarised performance of all GCMs with available data for each of the variables and periods in the study countries for different ranges of the R^2 skill evaluation parameter.

Variable	**Period**	**Dataset***	**R^2<0.5 (%)***	**0.5<R^2<0.7 (%)***	**R^2>0.8 (%)***	**R^2>0.9 (%)***
Rainfall	Annual	IS	2.8	6.6	77.8	54.3
		WS	37.5	19.4	30.8	17.0
		ALL	23.6	14.3	49.6	31.9
	DJF	IS	17.7	19.3	49.1	25.9
		WS	38.1	17.2	31.4	15.7
		ALL	29.9	18.1	38.5	19.8
	JJA	IS	12.8	17.2	58.9	40.1
		WS	15.2	19.1	52.1	34.5
		ALL	14.2	18.3	54.8	36.7
Diurnal temperature range	Annual	IS	0.4	2.2	81.8	73.1
		WS	0.4	1.2	54.5	46.1
		ALL	0.4	1.7	68.1	59.6
	DJF	IS	0.4	2.2	80.4	71.2
		WS	0.4	2.4	53.1	47.7
		ALL	0.4	2.3	66.8	59.4
	JJA	IS	0.4	2.0	80.7	67.2
		WS	0.4	1.2	54.5	46.1
		ALL	0.4	1.6	67.6	56.6
Mean temperature	Annual	IS	0.7	1.2	96.4	95.7
		WS	2.4	1.9	93.5	91.0
		ALL	1.7	1.6	94.7	92.8
	DJF	IS	3.5	1.9	93.2	91.5
		WS	2.3	2.3	93.9	91.2
		ALL	2.8	2.2	93.6	91.3
	JJA	IS	0.0	0.0	100.0	98.8
		WS	0.0	0.1	99.8	98.5
		ALL	0.0	0.1	99.9	98.6

* Values are expressed as percent of country-GCM combinations for comparisons of GCM-CL and different observational datasets: interpolated surfaces (IS), namely, WCL-IS and CRU-IS; weather stations (WS), namely, GHCN-CL, WCL-WS, GSOD-CL; and as the average of IS and WS (ALL)

Figure 1

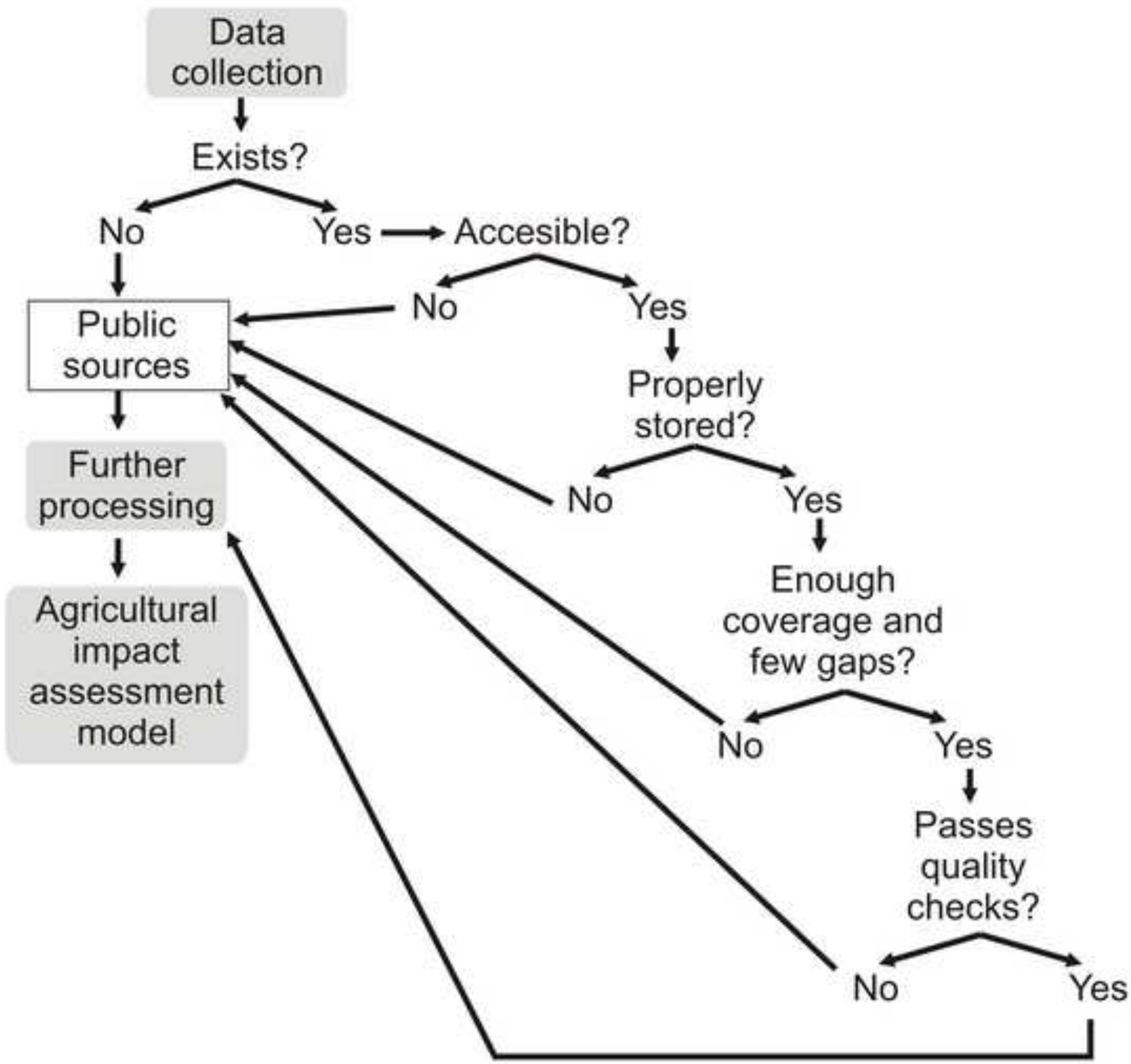
Data collection
Exists?
No
Yes
Accesible?
Public sources
No
Yes
Properly stored?
Further processing
No
Yes
Enough coverage and few gaps?
Agricultural impact assessment model
No
Yes
Passes quality checks?
No
Yes

Figure 2

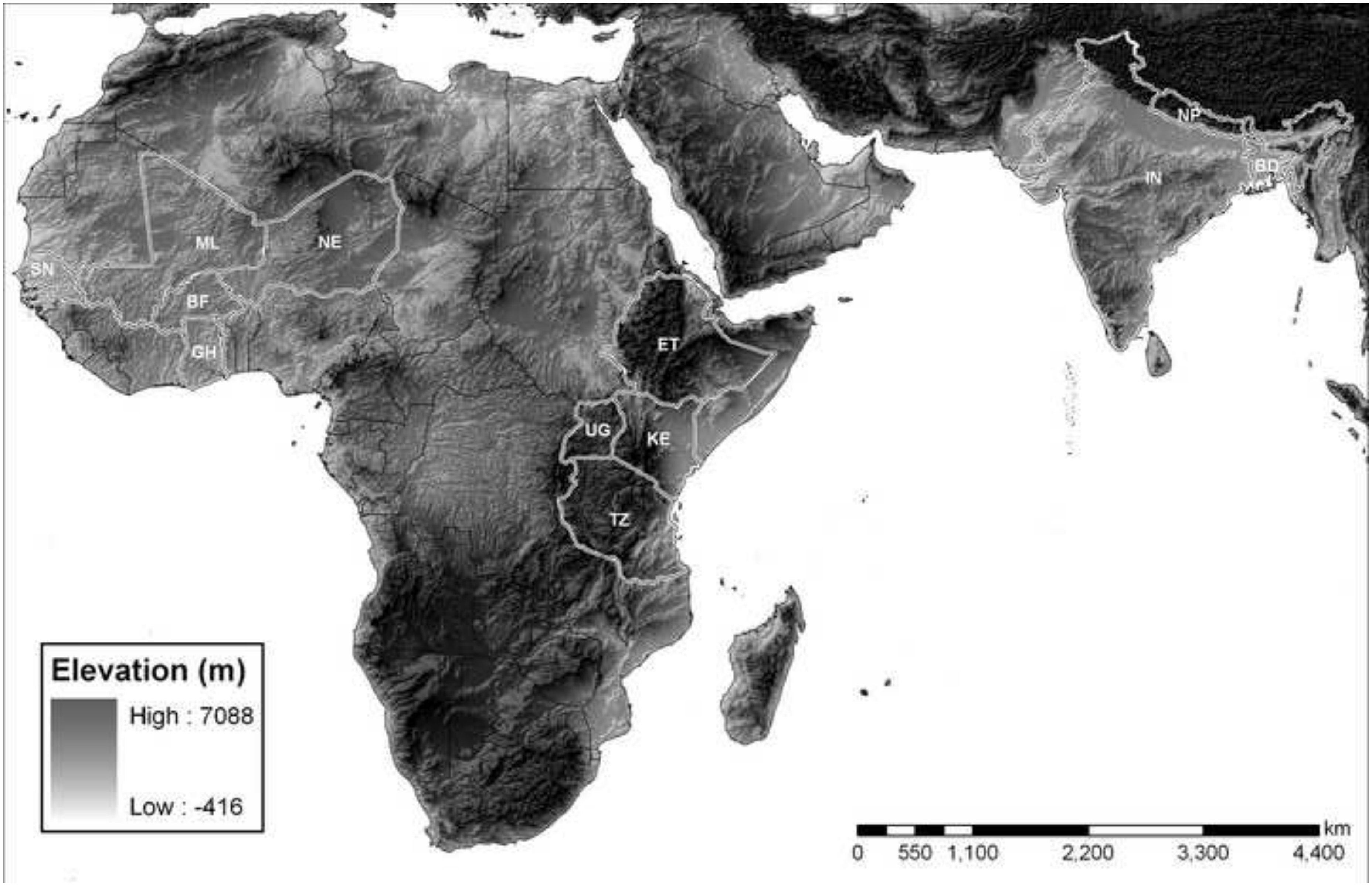
ML
NE
SN
BF
GH
ET
UG
KE
TZ
NP
IN
BD
Elevation (m)
High : 7088
Low : -416
0
550
1,100
2,200
3,300
4,400
km

Figure 3

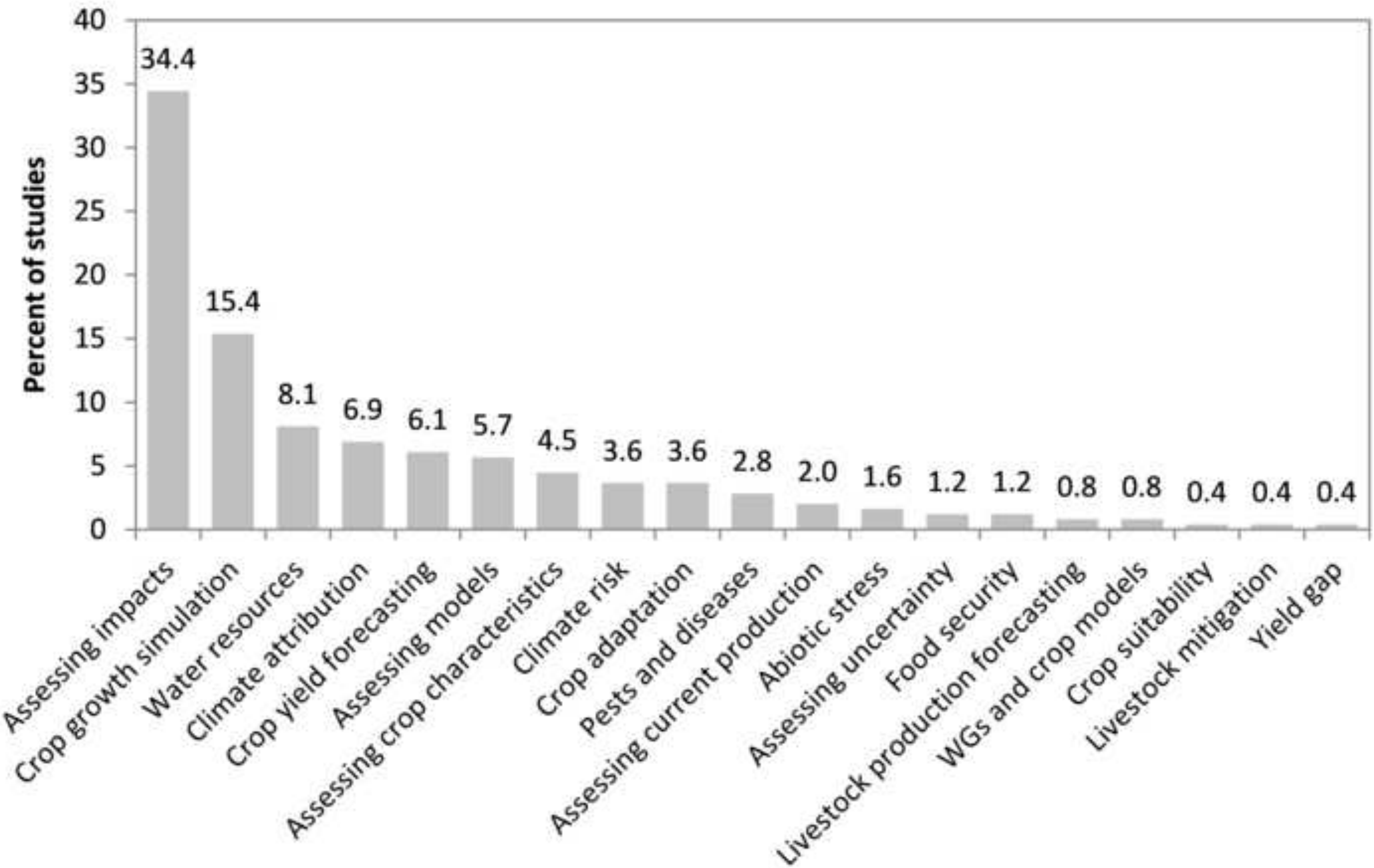
Percent of studies
40
35
30
25
20
15
10
5
0
34.4
15.4
8.1
6.9
6.1
5.7
4.5
3.6
3.6
2.8
2.0
1.6
1.2
1.2
0.8
0.8
0.4
0.4
0.4
Assessing impacts
Crop growth simulation
Water resources
Climate attribution
Crop yield forecasting
Assessing models
Assessing crop characteristics
Climate risk
Crop adaptation
Pests and diseases
Assessing current production
Abiotic stress
Assessing uncertainty
Food security
Livestock production forecasting
WGs and crop models
Crop suitability
Livestock mitigation
Yield gap

Figure 4a
Click here to download high resolution image

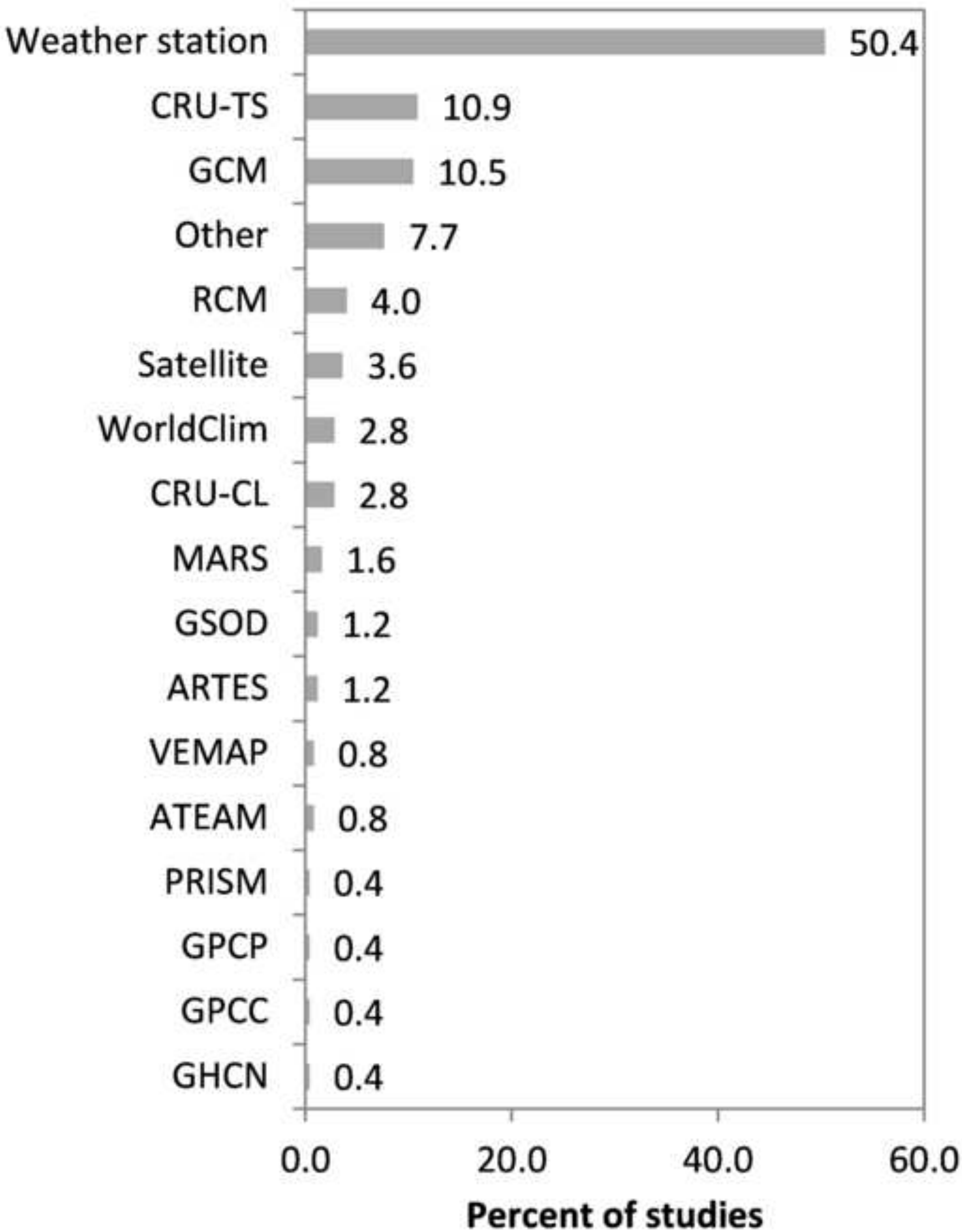

Figure 4b
Click here to download high resolution image

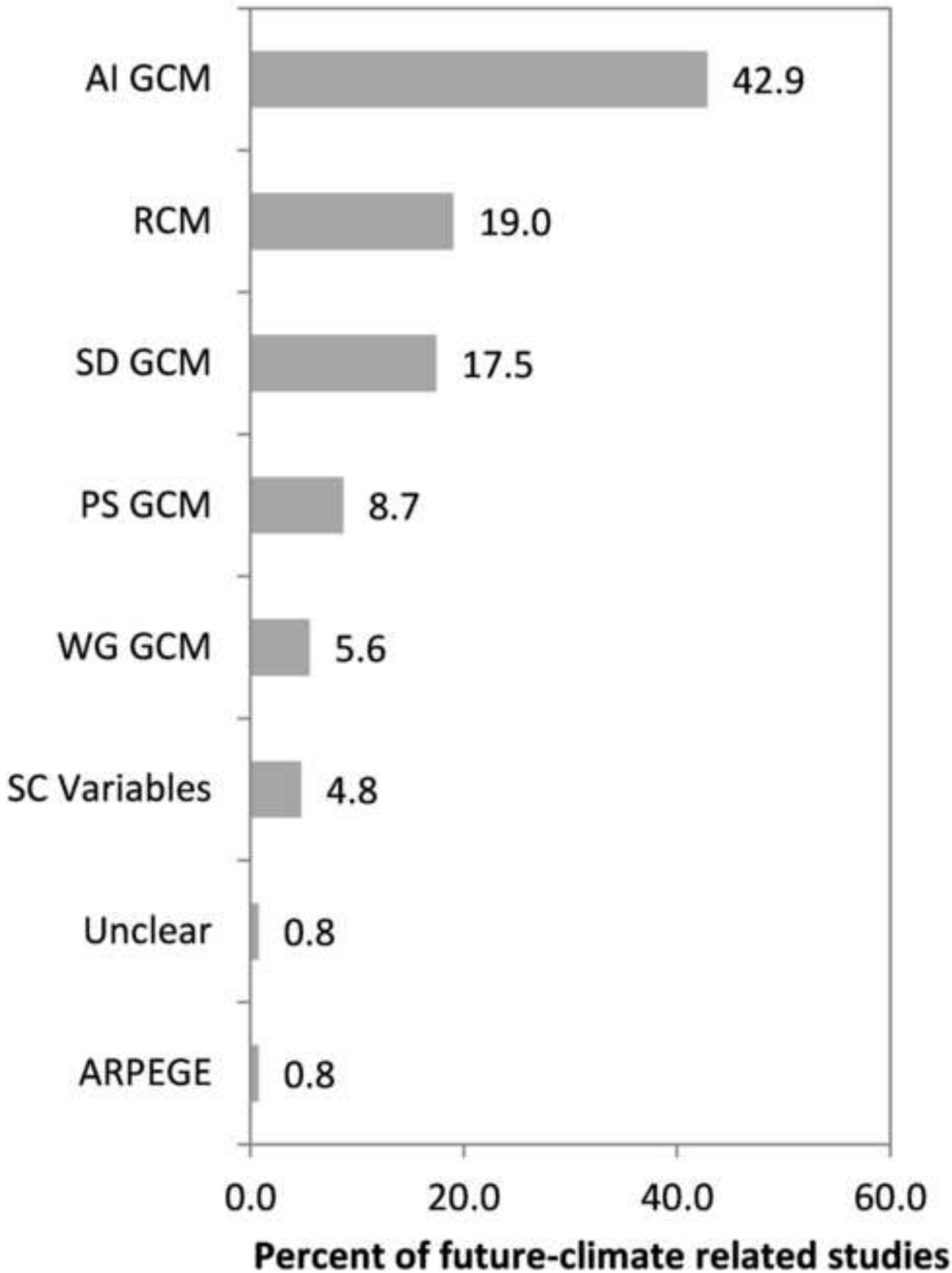

Figure 5a
Click here to download high resolution image

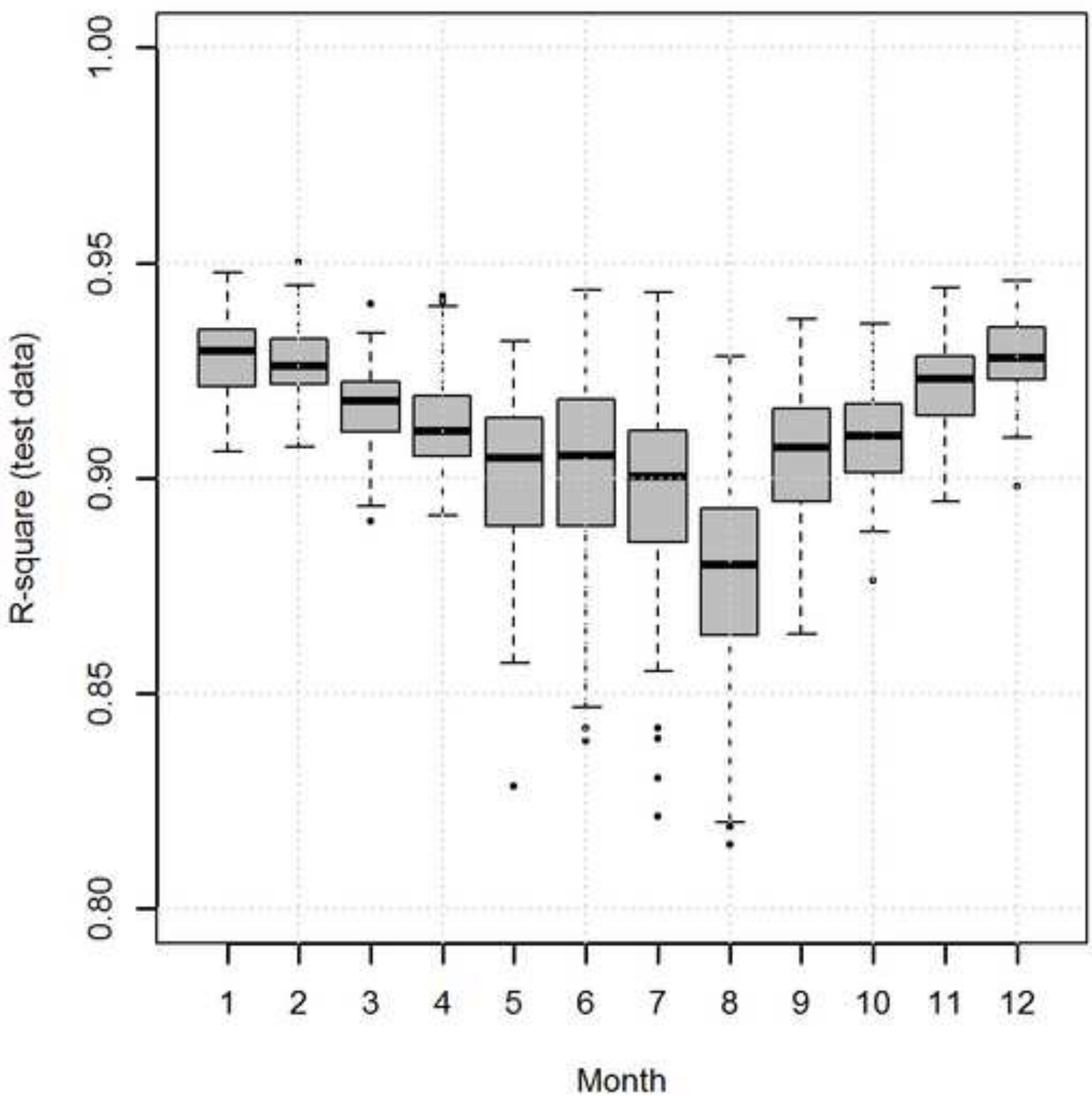

Figure 5b
Click here to download high resolution image

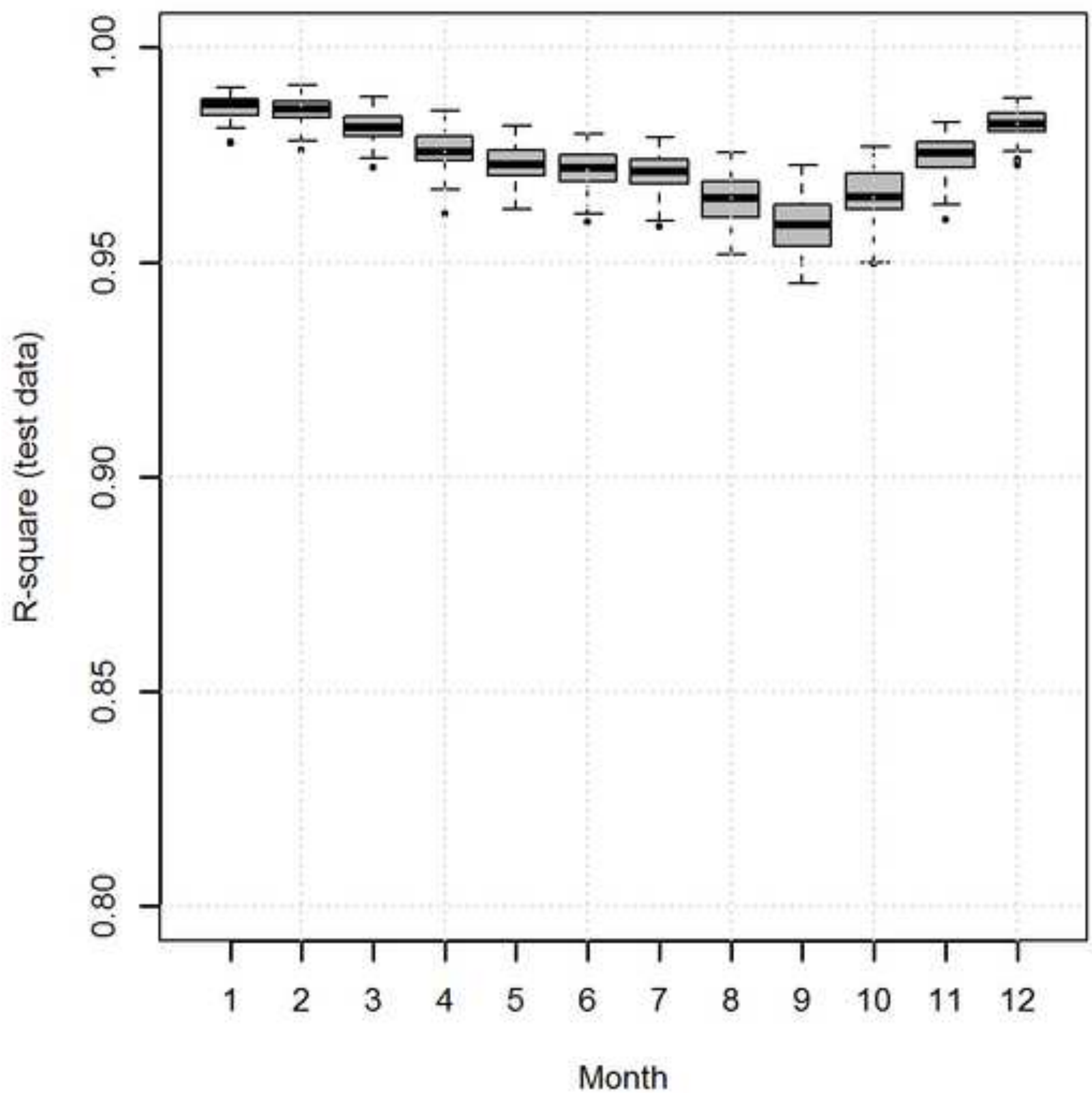

Figure 5c
Click here to download high resolution image

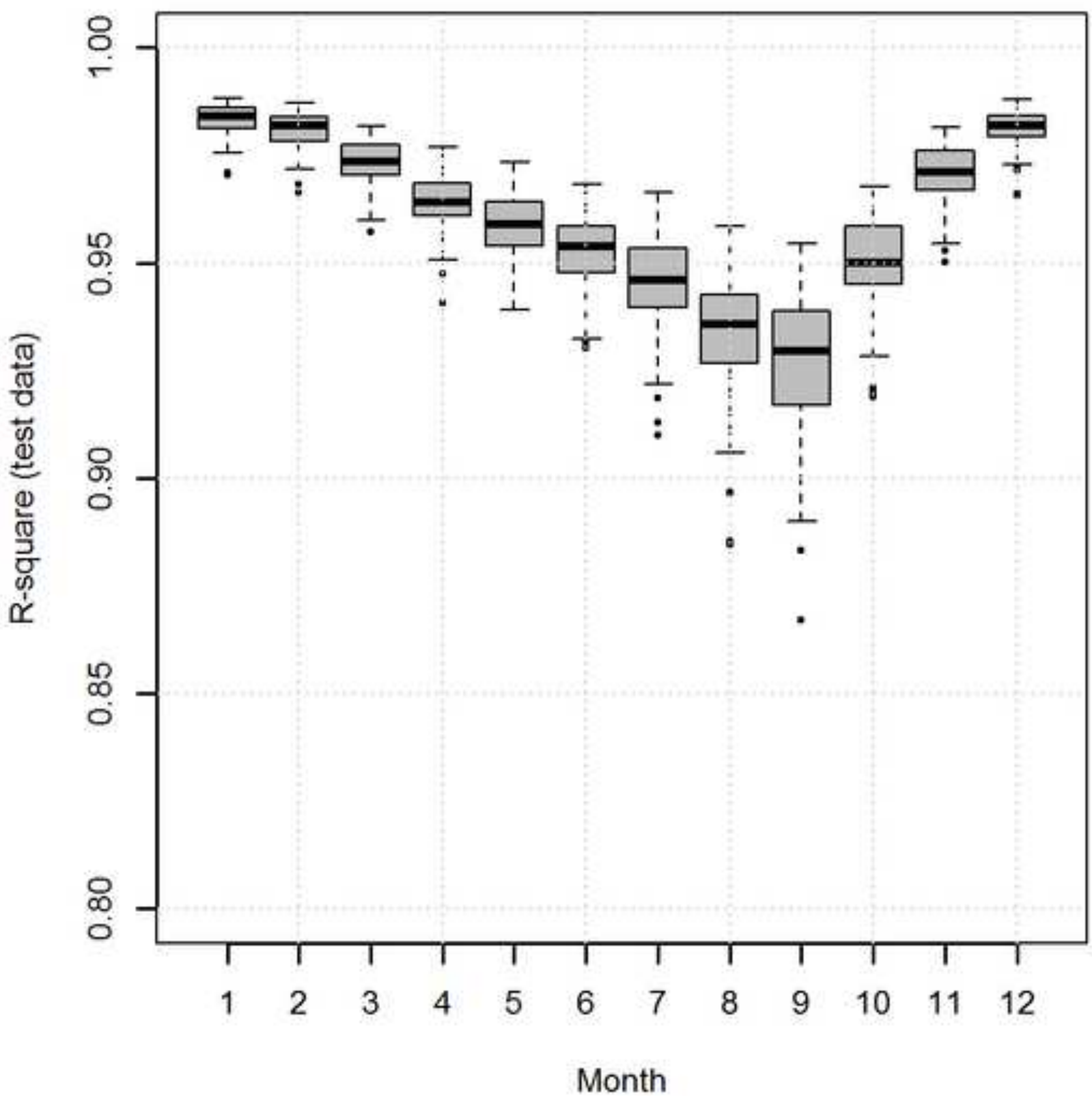

Figure 5d
Click here to download high resolution image

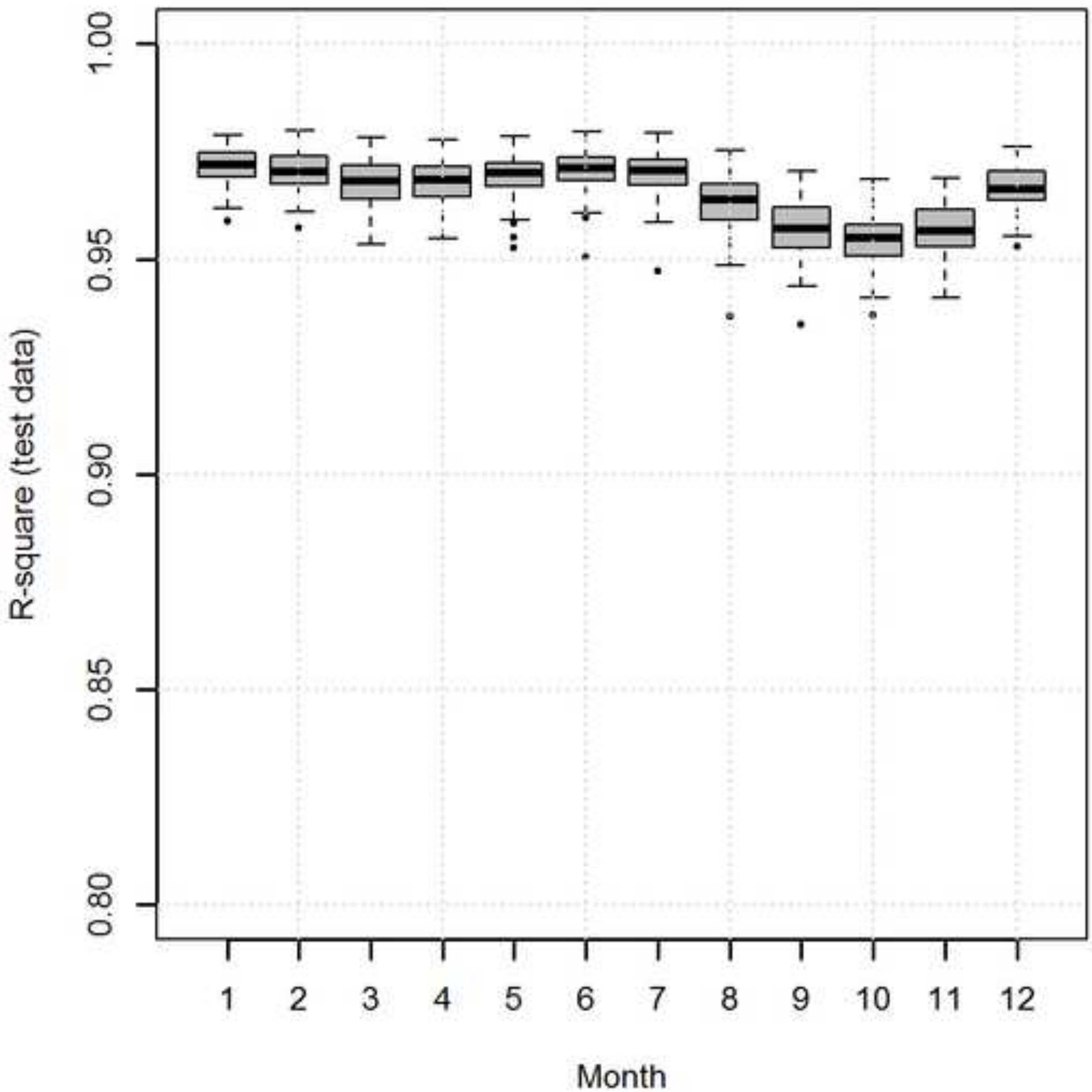

Figure 6a
Click here to download high resolution image

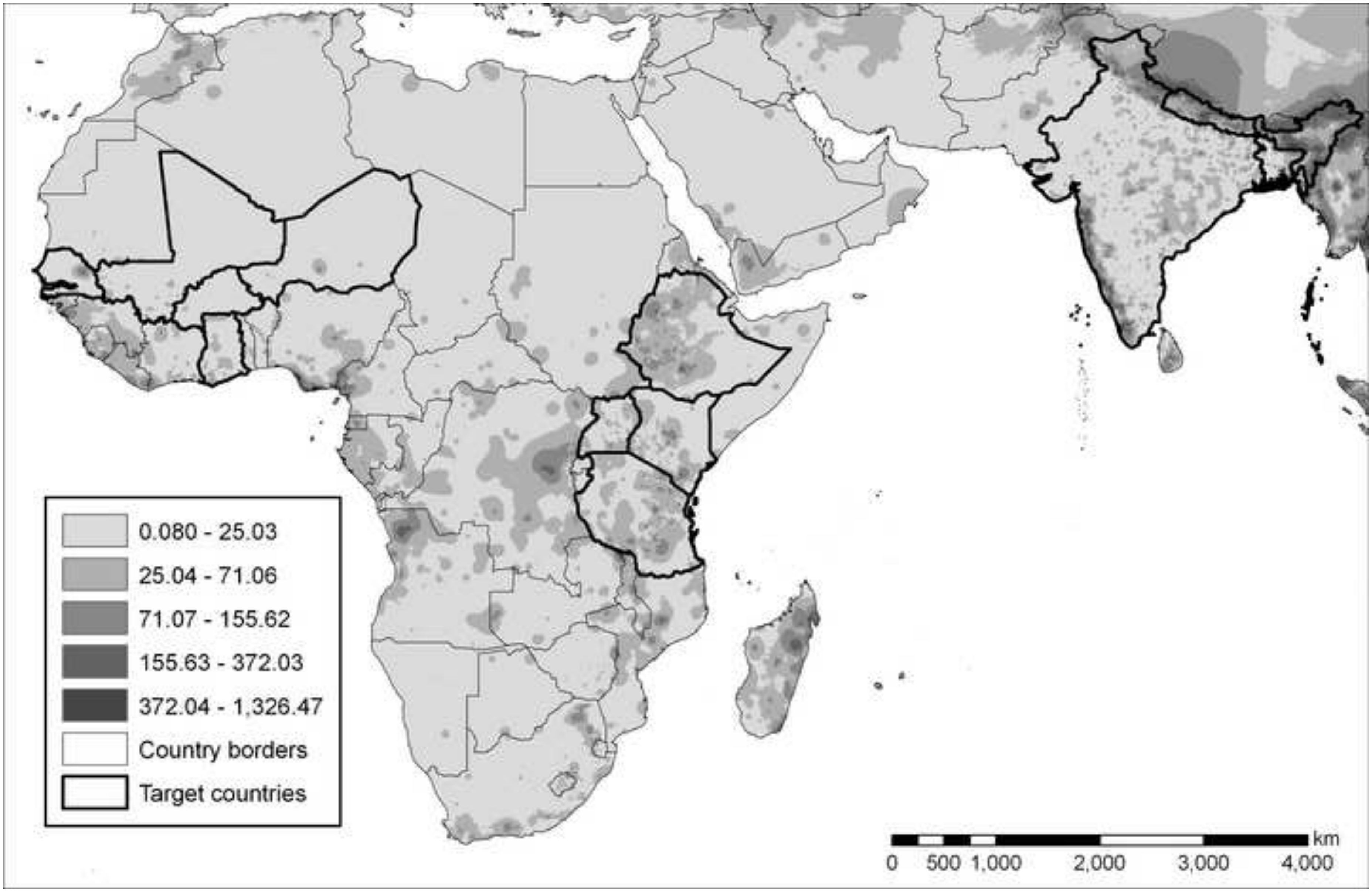

Figure 6b
Click here to download high resolution image

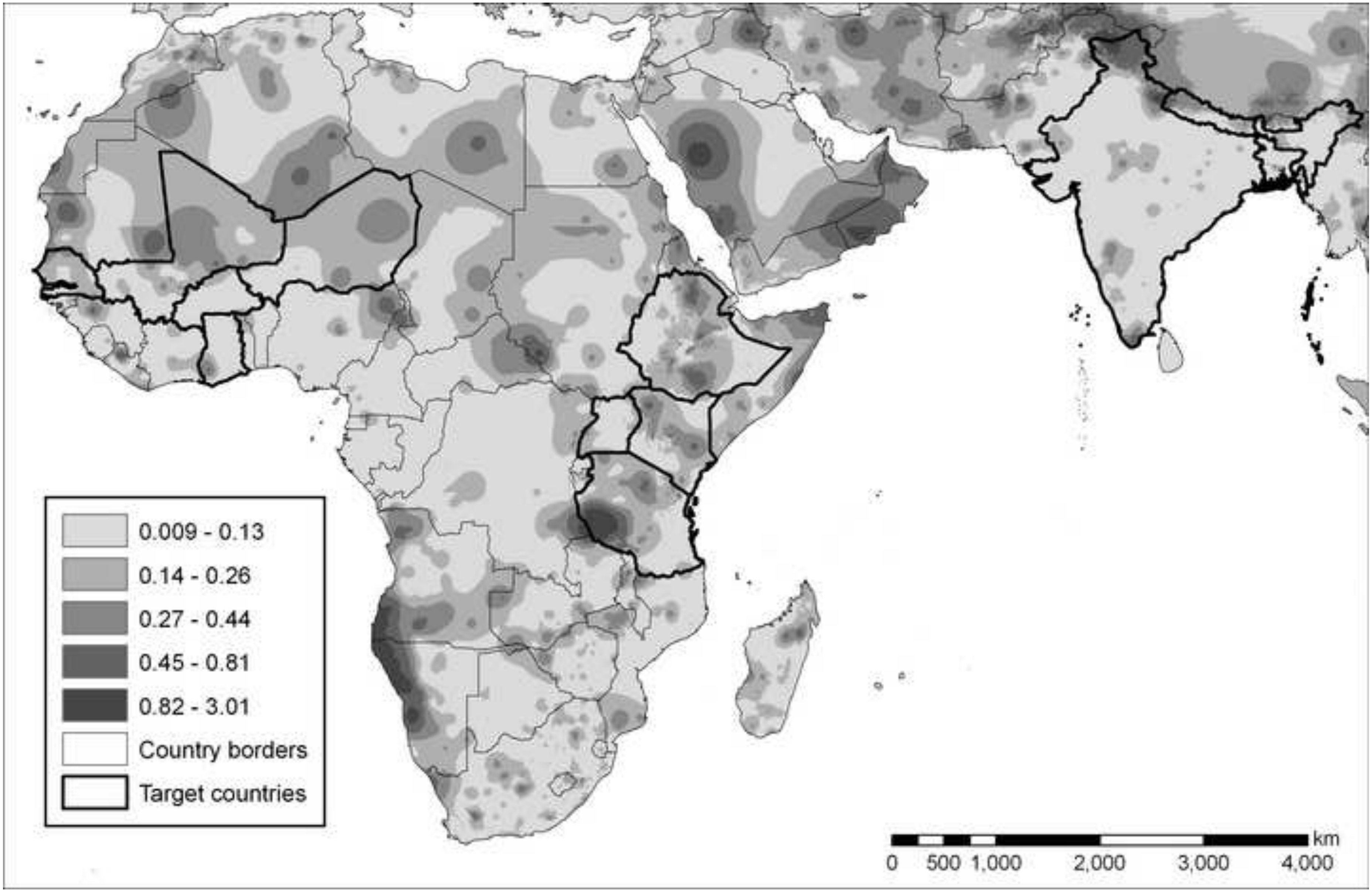

Figure 7
Click here to download high resolution image

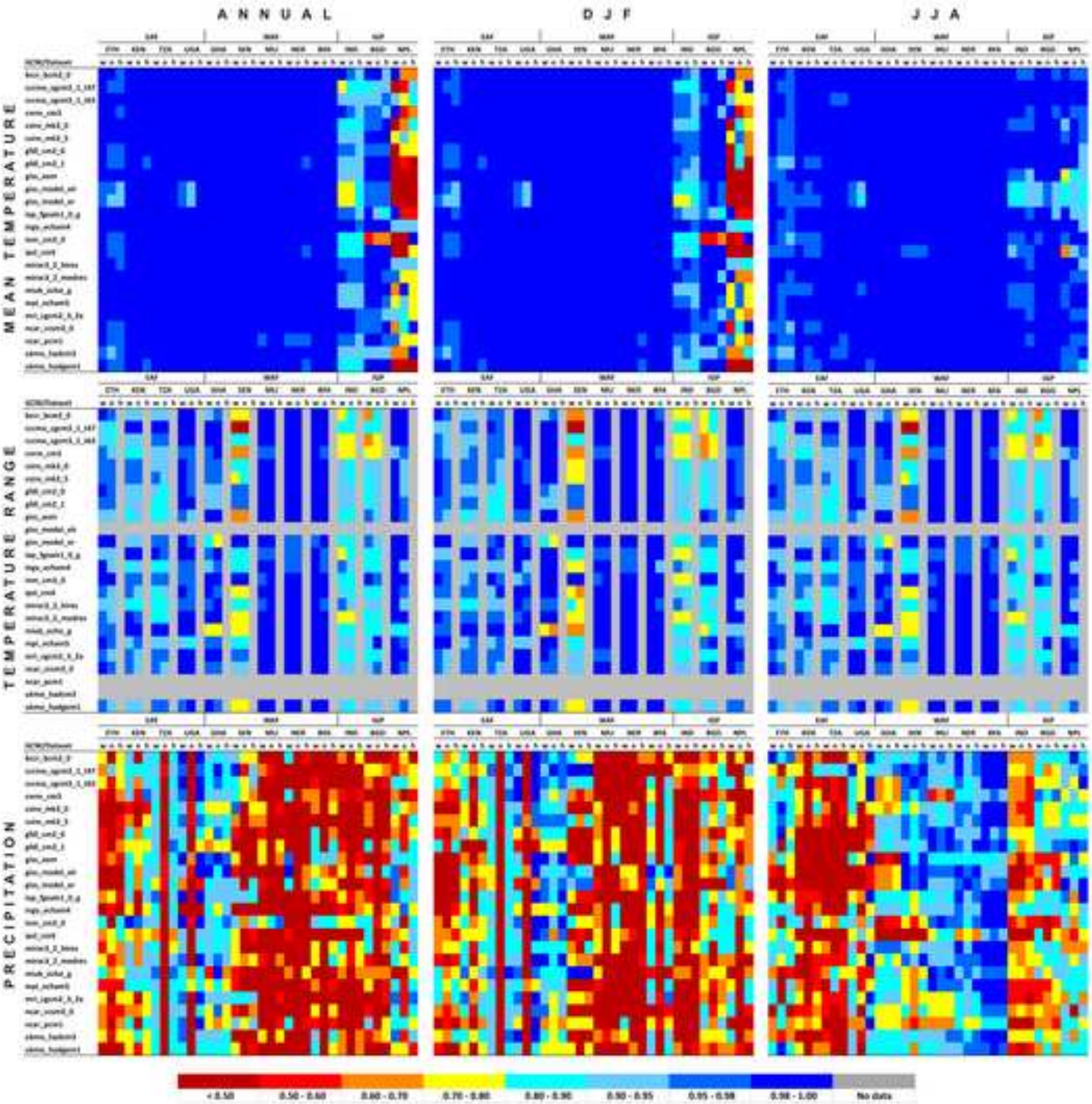

Figure 8
Click here to download high resolution image

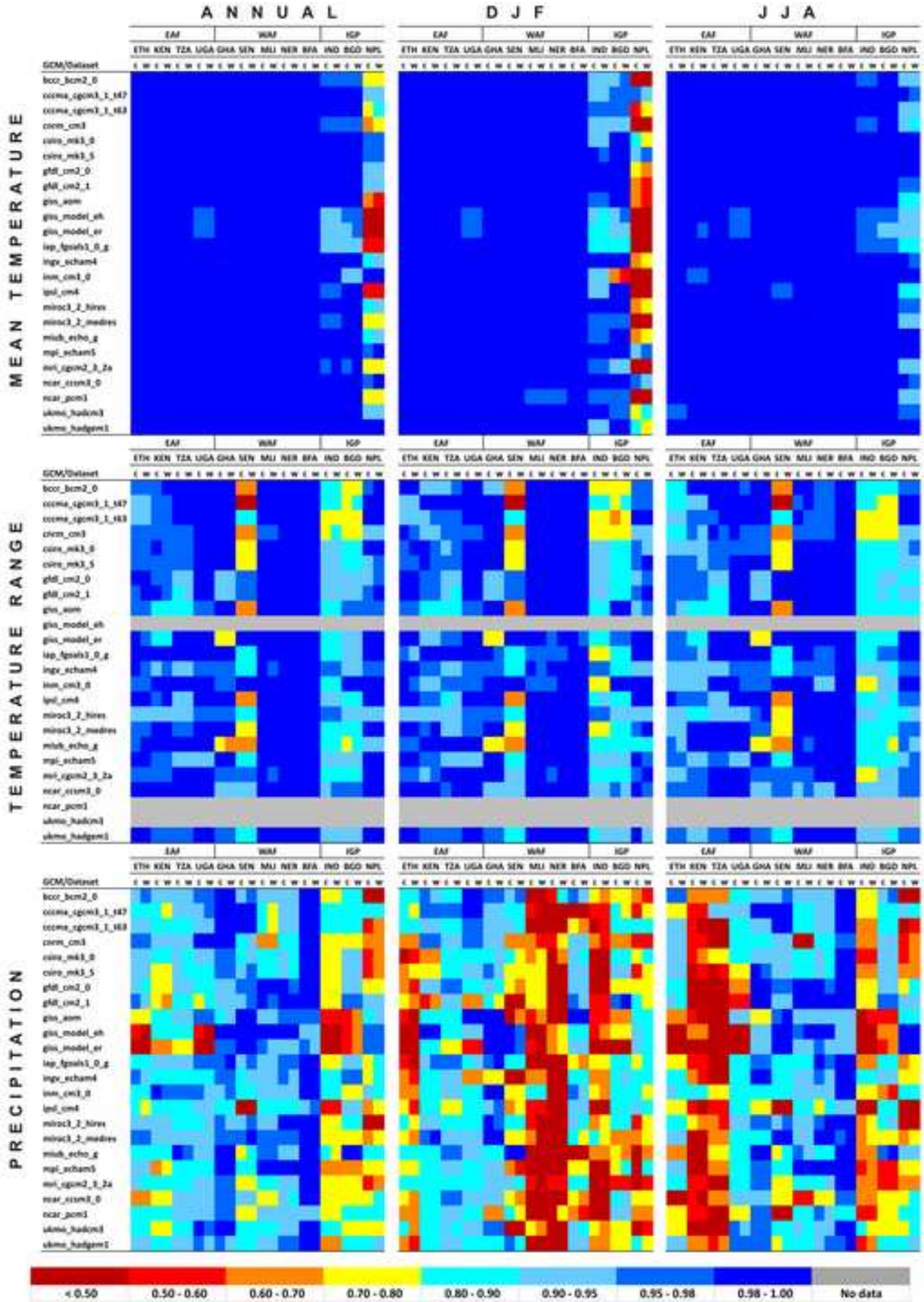

Figure 9
Click here to download high resolution image

Figure 10
Click here to download high resolution image

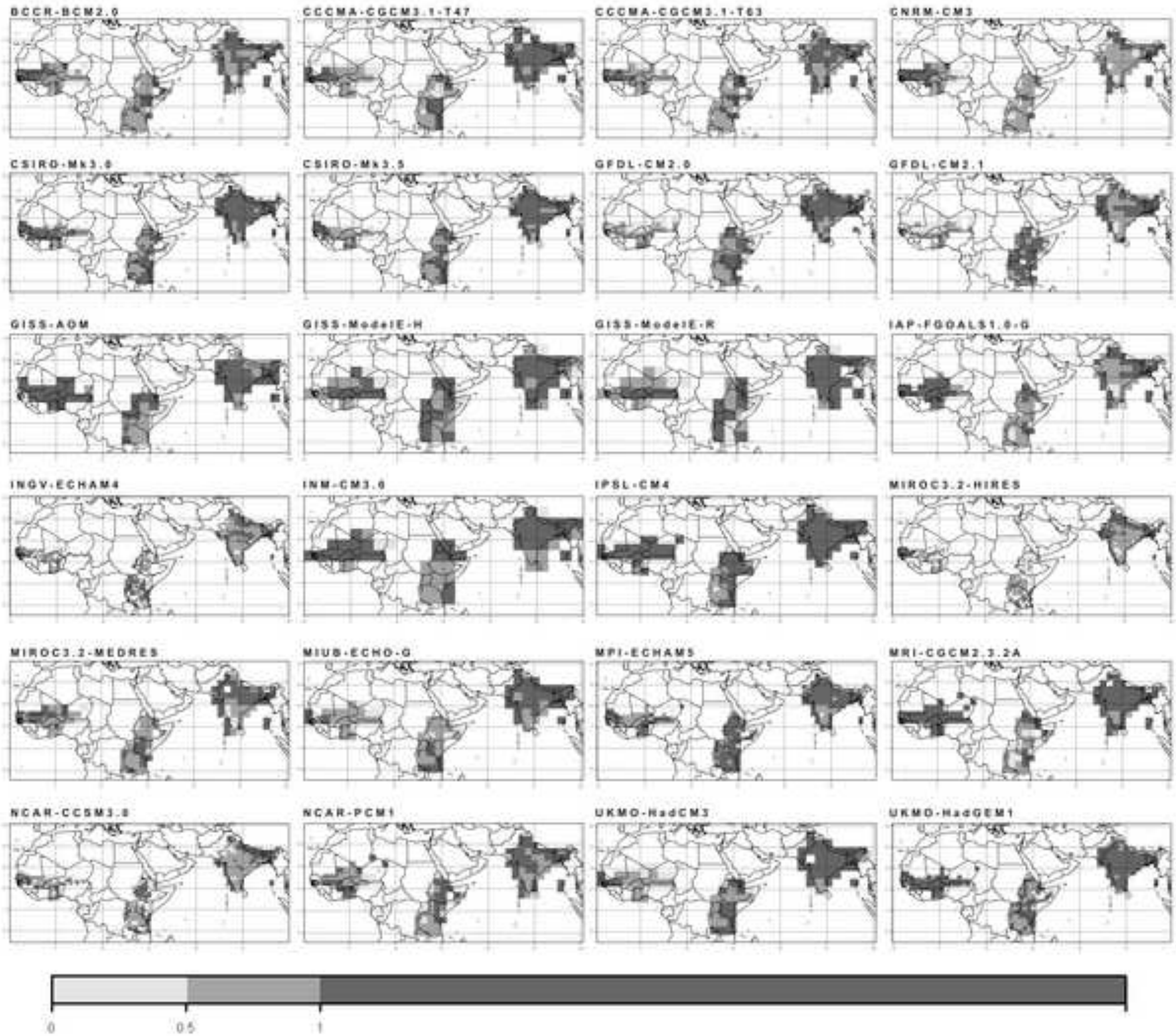

Figure 11
Click here to download high resolution image

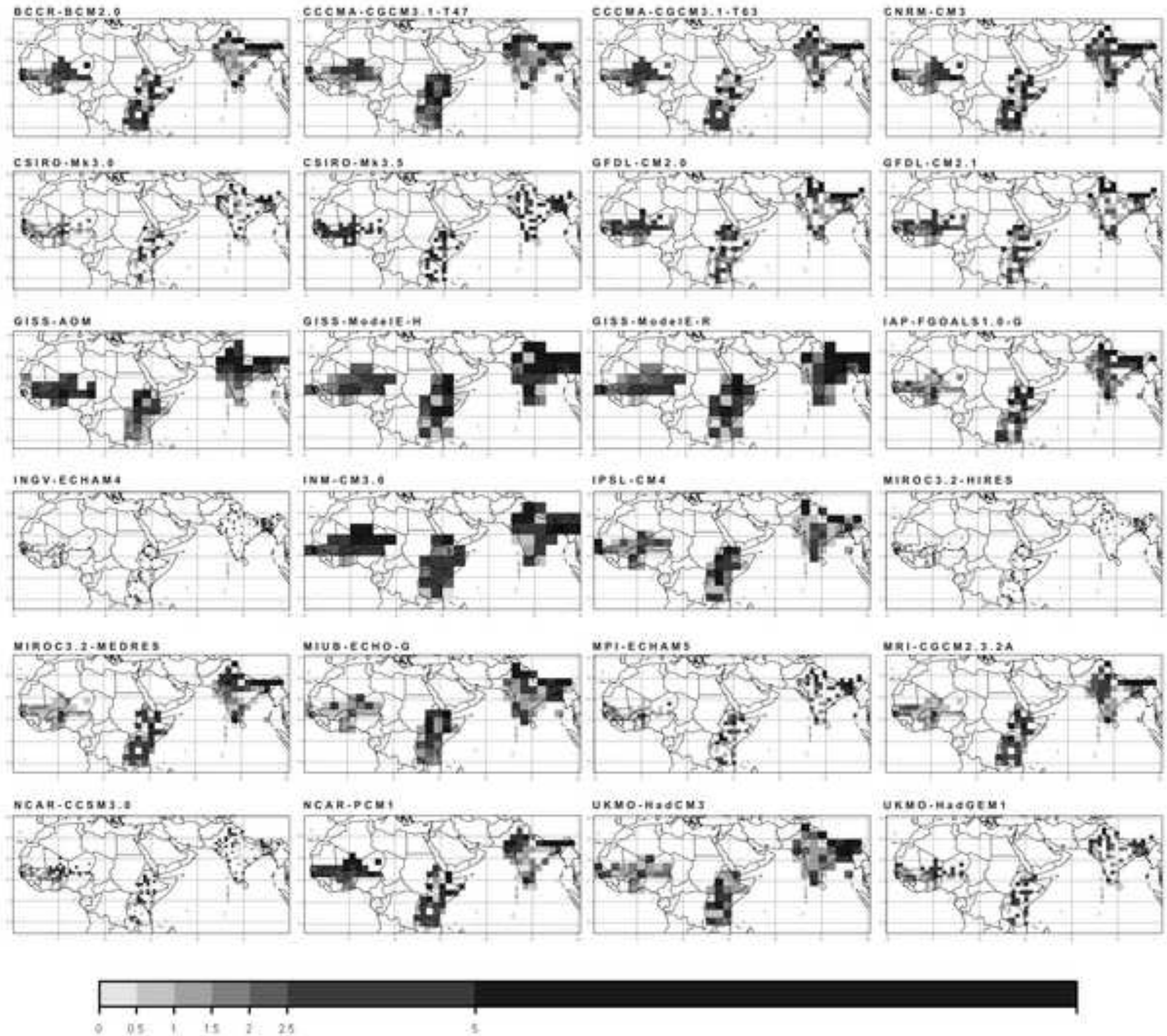

Figure 12

Click here to download high resolution image