Agricultural Systems 114 (2013) 1–5

Contents lists available at SciVerse ScienceDirect

Agricultural Systems

journal homepage: www.elsevier.com/locate/agsy

Short Communication

Generating downscaled weather data from a suite of climate models for agricultural modelling applications

Peter G. Jones [a,1], Philip K. Thornton [b,*]

[a] Waen Associates, Y Waen, Islawrdref, Dolgellau, Gwynedd LL40 1TS, United Kingdom
[b] CGIAR Research Program on Climate Change, Agriculture and Food Security (CCAFS), ILRI, PO Box 30709, Nairobi 00100, Kenya

A R T I C L E I N F O

Article history:
Received 14 January 2012
Received in revised form 30 July 2012
Accepted 10 August 2012
Available online 16 September 2012

Keywords:
Markov models
Climate change
Stochastic generation
Downscaling
DSSAT

A B S T R A C T

We describe a generalised downscaling and data generation method that takes the outputs of a General Circulation Model and allows the stochastic generation of daily weather data that are to some extent characteristic of future climatologies. Such data can then be used to drive any agricultural model that requires daily (or otherwise aggregated) weather data. The method uses an amalgamation of unintelligent empirical downscaling, climate typing and weather generation. We outline a web-based software tool (http://gismap.ciat.cgiar.org/MarkSimGCM) to do this for a subset of the climate models and scenario runs carried out for the 2007 Fourth Assessment Report of the Intergovernmental Panel on Climate Change. We briefly assess the tool and comment on its use and limitations.

1. Introduction

The availability of weather data continues to be a serious constraint to undertaking many applied research activities in the realm of agriculture, particularly in developing countries. Because weather is a primary determinant of agricultural production, weather data are needed for many different types of analysis in agricultural science. In addition to the data availability problem, the format in which data are available may be a considerable constraint to their widespread use. Nowhere is this more apparent than in agricultural impacts modelling, particularly in relation to utilising the outputs of climate models to evaluate possible impacts of climate change on crop and livestock production systems over the coming decades. The outputs from General Circulation Models (GCMs), climate models that project into the future, are almost never in a form that can be used directly to drive agricultural models. Considerable processing has to be gone through before such data can be meaningfully used, both for assessing possible impacts and for evaluating adaptation options. This processing generally involves downscaling the outputs from coarse-scaled GCMs to higher spatial and temporal resolutions. Various methods of downscaling exist, each with its own advantages and disadvantages, and each appropriate for different situations (Wilby et al., 2009). Reliable downscaling depends on the availability of reliable historical weather and climate data. Unfortunately, particularly in many developing countries, ground-based observation has declined considerably in the last several decades (Funk et al., 2011). Satellite technology is advancing rapidly and some aspects of weather and climate can be measured this way but such data are a complement to ground-based observation and not a substitute.

Here we describe a generalised downscaling and data generation method, which takes the outputs of a GCM describing a particular future climatology and allows the stochastic generation of a core set of daily weather data that are to some extent characteristic of this future climatology. This builds on previous methods, outlined and applied in Thornton et al. (2006), which utilised data from a suite of climate models used for the Third Assessment Report of the Intergovernmental Panel on Climate Change (IPCC, 2001). These methods have been modified to use outputs from the later generation of climate models utilised in the IPCC's Fourth Assessment Report (IPCC, 2007). The approach is fast and generalisable, and uses a mixture of methods, including simple interpolation (what Wilby et al. (2009) call "unintelligent downscaling"), climate typing and weather generation. Below we describe a web-based tool that uses these methods with a user interface in Google Earth and provides the user with daily weather data for current and future climatologies, which can then be used directly to run some widely-used crop models.

* Corresponding author. Tel.: +254 131 667 1960.
E-mail addresses: p.jones@cgiar.org (P.G. Jones), p.thornton@cgiar.org (P.K. Thornton).
[1] Tel.: +44 1341 423561.

0308-521X/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.agsy.2012.08.002

2. Materials and methods

2.1. Processing the GCM data

Outputs from many GCMs are available in the public domain, notably in the World Climate Research Program's (WCRP's) Coupled Model Intercomparison Project phase 3 (CMIP3) multi-model dataset. This dataset contains model output from 22 of the GCMs used for the Fourth Assessment (AR4; see Table 8.1 in Randall et al., 2007) and for a range of scenarios including the three scenarios reported on in the IPCC's Special Report on Emission Scenarios (SRES) used in the AR4. They are: A2, a high-greenhouse-gas-emission scenario; A1B, a medium-emission scenario; and B1, a low-emissions scenario. The SRES scenarios are described in detail in Nakicenovic et al. (2000).

Model output data are not available for all combinations of GCM and scenario for the basic "core" variables that are needed to drive many crop and pasture models (precipitation, maximum daily temperature and minimum air temperature). From CMIP3 and the Climate and Environmental Retrieval and Archive (CERA) database at the German Climate Research Centre (DKRZ), we found complete data for the three scenarios for a total of six GCMs (Table 1). As other data become available in the future, they can subsequently be included in the software.

SRES emission scenarios are considerably different in terms of projected changes in temperatures and rainfall for different regions. Table 2 shows the projected mean impacts on global temperature of these different scenarios from the IPCC multi-model ensemble for different time-slices. Although differences between the three scenarios in global warming impacts to 2050 are limited, thereafter these become considerable. Temperature shifts also vary substantially by region. Many GCMs project mean average temperature increases to 2050 for the East Africa region, for example, that are larger than the global mean: for scenario A2, of between about 1.5–2.5 °C. In addition to differences between the emission scenarios used to drive the climate models, the GCMs themselves can vary greatly. It is straightforward to plot rainfall and temperature patterns from different GCMs using the data and tool on the website www.ipcc-data.org, for example. GCMs differ in consistency for regional climate projections, particularly related to precipitation (IPCC, 2007).

The general scheme of the analysis here is as follows. First, we obtained data from the GCMs for five time slices: 1991–2010 (denoted "2000"), 2021–2040 (denoted "2030"), 2041–2060 (denoted "2050"), 2061–2080 (denoted "2070") and 2081–2100 (denoted "2090") for average monthly precipitation and daily maximum (T_{max}) and minimum (T_{min}) air temperatures. Processing of these data resulted in calculated mean monthly climatologies for each time slice and for each variable from the original daily time series produced by each GCM. The mean monthly fields had been interpolated, by the original agencies from whom we obtained the data, from the original resolution of each GCM to 0.5° latitude–longitude using conservative remapping, which preserves the global averages. Second, we calculated monthly climate anomalies (absolute changes) for monthly rainfall, mean daily maximum temperature and mean daily minimum temperature, for each time slice relative to the baseline climatology (1961–1990). The point of origin was designated 1975, being the mid point of the 30-year climate normals.

Third, we fitted a functional relationship to the climate projections for the variables of interest through time, so that we could interpolate the projections to any year. We inspected the responses of the chosen models and found that they were considerably more complicated than those of the third approximation models used in the previous exercise (Thornton et al., 2006). There it was found by stepwise regression that a cubic term was superfluous to describe the projections over time. In the current case, we made a preliminary investigation of the functional forms of the projections using cluster analysis. All pixels from each of the climate models for scenario A1B were clustered for precipitation, T_{max} and T_{min} using the values of the five periods as clustering variates. We used a leader clustering algorithm (Hartigan, 1975) to cope with the volume of data. The threshold was set to produce from 40 to 100 clusters, which were ranked by the number of pixels, and the cluster means were used to inspect the functional form. The first five clusters normally covered 80–90% of the pixels for any given model.

We fitted polynomials through the cluster means by date (constrained through the origin) and this showed that in many cases a quadratic fit over time would have sufficed but in numerous cases only a fourth-order polynomial would suffice. We therefore decided to fit fourth-order polynomials throughout. We made these fits for all models at all scenarios and made another set for the average of the six models. We constructed world maps of the residual surfaces for every time period for each variate and for each model and scenario. Visual inspection of every map showed that deviations from the fitted curves were within expectations for all the models. Finally, we condensed the polynomial coefficients into a data file structure for ready retrieval on a pixel-by-pixel basis (at a resolution of 30 arc-min) for use in subsequent operations: downscaling the anomalies to a higher resolution, and then generating daily weather data that are characteristic, to some extent, of the future climatologies produced, using a stochastic daily weather generator.

2.2. Generating daily data: MarkSim®

MarkSim® is a third-order markov rainfall generator (Jones and Thornton, 1993, 1997, 1999, 2000; Jones et al., 2002), which has been developed over 20 years. It was not designed as a GCM downscaler, but it does now work as such, employing both stochastic downscaling and climate typing.

The basic algorithm of MarkSim is a daily rainfall simulator that uses a third-order markov process to predict the occurrence of a rain day. A third-order model was shown to be necessary for

Table 1
Atmosphere-Ocean General Circulation Models (AOGMCs) used in the work (details from Randall et al., 2007).

Model name (Date)	Institution	Reference	Resolution	Code
BCCR_BCM2.0 (2005)	Bjerknes Centre for Climate Research (University of Bergen, Norway)	Furevik et al. (2003)	1.9 × 1.9°	BCC
CNRM-CM3 (2004)	Météo-France/Centre National de Recherches Météorologiques, France	Déqué et al. (1994)	1.9 × 1.9°	CNR
CSIRO-Mk3.5 (2005)	Commonwealth Scientific and Industrial Research Organisation (CSIRO) Atmospheric Research, Australia	Gordon et al. (2002)	1.9 × 1.9°	CSI
ECHam5 (2005)	Max Planck Institute for Meteorology, Germany	Roeckner et al. (2003)	1.9 × 1.9°	ECH
INM-CM3_0 (2004)	Institute for Numerical Mathematics, Moscow, Russia	Diansky and Volodin (2002)	4.0 × 5.0°	INM
MIROC3.2 (medres) (2004)	Center for Climate System Research (University of Tokyo), National Institute for Environmental Studies, and Frontier Research Center for Global Change (JAMSTEC), Japan	K-1 Model Developers (2004)	2.8 × 2.8°	MIR
Ensemble average	Average climatology of the above 6 AOGCMs	–	–	AVR