Generating characteristic daily weather data using downscaled climate model data from the IPCC's Fourth Assessment

Project Report

"Supporting the vulnerable: Increasing the adaptive capacity of agro-pastoralists to climatic change in West and Southern Africa using a trans-disciplinary research approach"

August 2009

Generating characteristic daily weather data using downscaled climate model data from the IPCC's Fourth Assessment

Peter G Jones

Waen Associates, Y Waen, Islaw'r Dref, Dolgellau, Gwynedd LL40 1TS Wales

Philip K Thornton,

International Livestock Research Institute (ILRI), PO Box 30709, Nairobi 00100, Kenya

Jens Heinke

Potsdam Institute for Climate Impact Research (PIK), PO Box 60 12 03, 14412 Potsdam, Germany

Contents

Summary

We describe a generalized 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 impacts model that requires daily (or otherwise aggregated) weather data. We outline software to do this for a subset of the climate models and scenario runs carried out for 2007's Fourth Assessment Report of the Intergovernmental Panel on Climate Change (IPCC). We briefly assess the methods used, comment on the limitations of the data, and make suggestions for further work.

1 Introduction

The outputs from General Circulation Models (GCMs), climate models that project into the future, are never in a form that can be used directly to study the impacts of climate change on natural or human systems. There is quite a process that has to be gone through, in general, before such data can be meaningfully used to assess the possible impacts of specific climate scenarios on crop and pasture production in particular places, for example.

In this document, we describe a generalized downscaling and data generation method, that takes the outputs of a GCM that describe some future climatology, and allows the stochastic generation of daily weather data that are to some extent characteristic of this future climatology, that can then be used to drive any impacts model that requires daily (or otherwise aggregated) weather data. This builds on previous methods, outlined and applied in Thornton et al. (2006) which utilised the data set TYN SC 2.0 (Mitchell et al., 2004) based on data from the climate models that were used for the IPCC's Third Assessment Report (IPCC, 2001). Here, we modify these methods to allow us to use a later generation of climate models that were utilised in the IPCC's Fourth Assessment Report (IPCC, 2007).

Section 2 contains a brief outline of different downscaling methods. In section 3, we describe the source of the GCM data, the standardisation of these to a common spatial resolution, and then the methods by which the data were manipulated to form climate grids, and then the methods used to enable daily data to be generated. Section 4 presents a brief description of

the software and indicates how the information maybe used and applied. The paper concludes with a very brief assessment of the limitations of these data and suggestions for further work.

2 Background: Downscaling

GCMs model the atmosphere in stacks of cells that have little to do with the detailed topography of the ground. They therefore do not model weather, which involves local effects, storms, fronts and orogenic effects, but rather they model the average temperature of the cells in the stacks and precipitation is calculated from the latent heat balance as air is transferred from cell to cell.

Downscaling refers to the process of taking these rather esoteric outputs and relating them to real points in the real world. It does not necessarily mean filling in the rather sparse grid that most GCMs operate on, but actually bringing them down to earth. Usually we want to do both (Jones et al., 2005). There are several ways of downscaling, broadly speaking, including the following.

1 Take the difference between what the GCM predicts now, or at some known time in the past, and what it predicts in the future. Add that to the value that you have at the point you are interested in and assume that this will be the value in the future. This will generally involve interpolating the difference values according to some logical rule. This could involve simple interpolation methods such as thin plate smoothing, kriging, or inverse square distance weighting. This is the worst method of downscaling.

2 Take a sequence of real past daily weather data and determine the relationship with the same sequence of years from the nearest GCM point available. The statistical relationship should include regressions on the means and variances of the sequences. This is good, but it begs two questions. First, do you have a real set of data anywhere near the point you want to downscale to? Second, if not, then do you have enough real historic data sets of daily weather data around the point to do a realistic interpolation? This is usually called statistical downscaling.

3 Another option is stochastic downscaling. This involves developing a stochastic model of the time series of data involved, comparing this model with the GCM results, and then using the stochastic model to downscale the results to real earth points. This method has some real advantages, but it suffers from the same problems as in method 2 above.

4 Hierarchical downscaling has a real benefit for those with large purses and small areas to cover. The general idea is this: a GCM can cover the world at various resolutions, with grid cell sizes ranging from 2 degrees latitude-longitude to 0.5 degrees. However, for relatively small areas (say the country of Honduras), much higher resolutions can be achieved by plugging in the boundary conditions from a world-scale GCM and then running a higher-resolution climate model (in this case, for Honduras) on the same sort of software.

5 Then there is climate typing, which involves looking at typical current climate patterns and frequencies (such as the frequency and severity of Atlantic fronts passing over the British Isles). These can then be related to similar events in a GCM and the subsequent model used to predict future climate patterns in more detail. This method seems to have a lot of merit in that it is looking at actual climate mechanisms and how they change under the influence of a GCM. It suffers even more than statistical and stochastic downscaling from problems caused by the lack of reliable historic data.

It is clear that 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. Satellite technology is advancing rapidly, and many things can be measured this way, but such data are a complement to ground-based observation and not a substitute.

There are no real alternatives to the methods outlined above, although there are some methods that combine different approaches that have some merit. Here we describe a suboptimal but fast and generalized downscaling and data generation method, and its use in a previously-described system that incorporates aspects of at least three of the aforementioned downscaling methods.

3. Processing the GCM Data

A considerable amount of GCM data is 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 (see Table 8.1 in Randall et al., 2007), and for a range of scenarios including the three SRES scenarios used in AR4: A2, a high-greenhouse-gas-emission scenario; A1B, a medium-emission scenario; and B1, a low-emissions scenario. The SRES scenarios are described in Table 1. Model output data are not available for all combinations of GCM and scenario, at least not the variables that we were wanting -- the basic "core" variables for many crop and pasture models are precipitation, maximum daily temperature, and minimum air temperature. This severely restricted our choice of GCMs. From the CMIP3 dataset, we could find what we needed for only three GCMs (see Table 2): CNRM-CM3, CSIRO-Mk3.0, and MIROC 3.2 (medium resolution). We were able to obtain maximum and minimum temperature data for the ECHam5 model from another source (the CERA database at DKRZ) for the three SRES scenarios. This gave us data for a total of four GCMs and three emission scenarios. If other data become available in the future (such as from the UK's HadCM3 GCM, for example), these data can subsequently be included in the software.

There are considerable differences between SRES emission scenarios, in terms of projected changes in temperatures and rainfall for different regions. Figure 1, taken from the IPCC's Fourth Assessment Report (AR4), shows global multi-model means of surface warming (relative to 1980–1999) for the scenarios A2, A1B and B1 (Meehl et al., 2007). There is not that much difference between the three scenarios in global warming impacts to 2050, although thereafter the differences become considerable. There are also substantial regional variations in temperature shifts. 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 to 2.5 °C (compare with Figure 1).

In addition to differences between the emission scenarios used to drive the climate models, there can be substantial differences between the GCMs themselves. This is already clear in Figure 1, where the solid lines show the multi-model means, and the shading around each line shows the ±1 standard deviation range of individual model annual means. Some other

Table 1. *The Emissions Scenarios of the Special Report on Emissions Scenarios (SRES) (Nakicenovic et al., 2000)*

_____

A1. The A1 storyline and scenario family describe a future world of very rapid economic growth, global population that peaks in mid-century and declines thereafter, and the rapid introduction of new and more efficient technologies. Major underlying themes are convergence among regions, capacity building and increased cultural and social interactions, with a substantial reduction in regional differences in per capita income. The A1 scenario family develops into three groups that describe alternative directions of technological change in the energy system. The three A1 groups are distinguished by their technological emphasis: fossil intensive (A1FI), non-fossil energy sources (A1T), or a balance across all sources (A1B) (where balanced is defined as not relying too heavily on one particular energy source, on the assumption that similar improvement rates apply to all energy supply and end use technologies).

A2. The A2 storyline and scenario family describe a very heterogeneous world. The underlying theme is self-reliance and preservation of local identities. Fertility patterns across regions converge very slowly, which results in continuously increasing population. Economic development is primarily regionally oriented and per capita economic growth and technological change are more fragmented and slower than in other storylines.

B1. The B1 storyline and scenario family describe a convergent world with the same global population, that peaks in mid-century and declines thereafter, as in the A1 storyline, but with rapid change in economic structures toward a service and information economy, with reductions in material intensity and the introduction of clean and resource-efficient technologies. The emphasis is on global solutions to economic, social and environmental sustainability, including improved equity, but without additional climate initiatives.

B2. The B2 storyline and scenario family describe a world in which the emphasis is on local solutions to economic, social and environmental sustainability. It is a world with continuously increasing global population, at a rate lower than A2, intermediate levels of economic development, and less rapid and more diverse technological change than in the B1 and A1 storylines. While the scenario is also oriented towards environmental protection and social equity, it focuses on local and regional levels.

_____

examples are shown in Appendix 1, which presents June temperature and precipitation anomalies, relative to the twentieth century control (1961-1990) 30 year normal, for the 2046-2065 time slice, for the SRES A2 scenario and four GCMs, from AR4 simulations. Note that the scale for the temperature plots is the same for all GCMs, but there are slight differences between the scales for the rainfall anomalies. These (and many other similar) plots can be generated by any user and are publicly accessible on the website www.ipcc-data.org.

Table 2. *AOGCMs used in the work (details from Randall et al., 2007).*

Model Name, Vintage	Institution	Reference	Resolution	Code
CNRM-CM3, 2004	Météo-France/Centre National de Recherches Météorologiques, France	Déqué et al. (1994)	1.9 x 1.9 degrees	CNR
CSIRO-Mk3.0, 2001	Commonwealth Scientific and Industrial Research Organisation (CSIRO) Atmospheric Research, Australia	Gordon et al (2002)	1.9 x 1.9 degrees	CSI
ECHam5, 2005	Max Planck Institute for Meteorology, Germany	Roeckner et al (2003)	1.9 x 1.9 degrees	ECH
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 Developers (2004)	2.8 x 2.8 degrees	MIR

Figure 1. *Multi-model means of surface warming (relative to 1980–1999) for the scenarios A2, A1B and B1, shown as continuations of the 20th-century simulation. Lines show the multi-model means, shading denotes the ±1 standard deviation range of individual model annual means (from Meehl et al., 2007)*
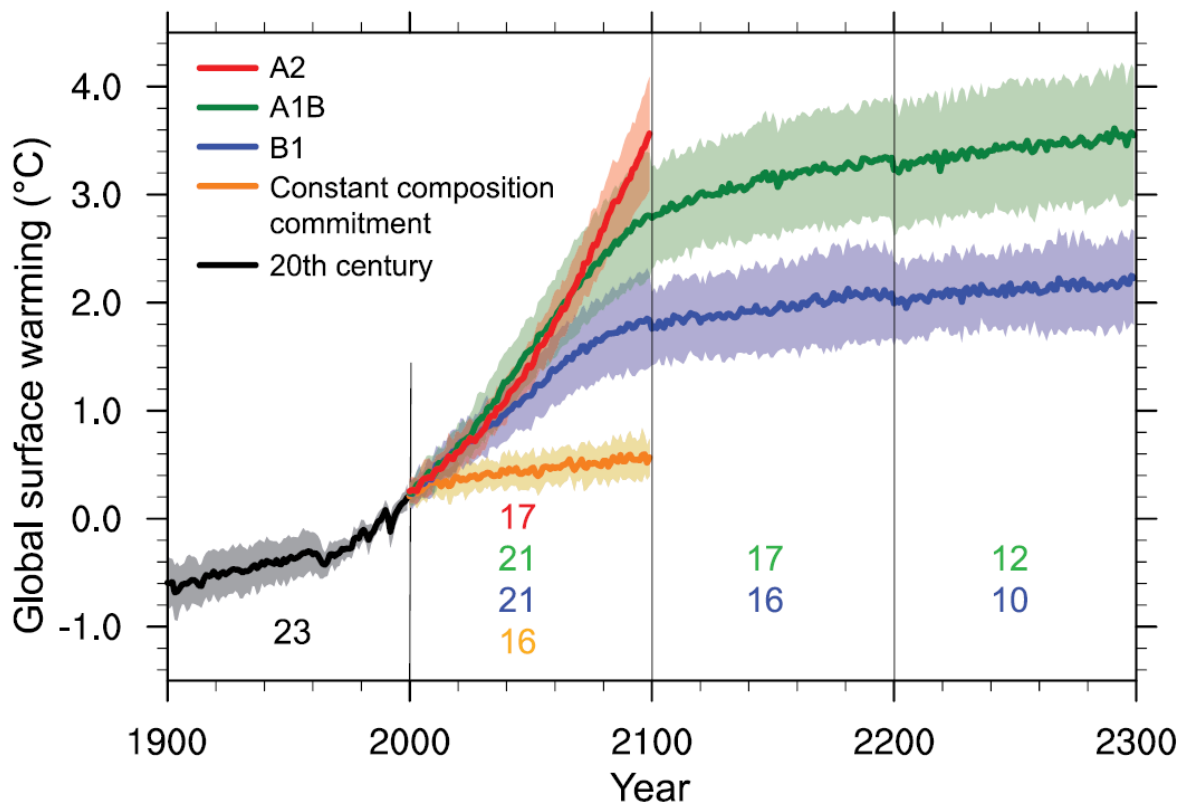
The general scheme of the process is outlined in Figure 2. We obtained data for GCM deviations for five time slices: 1991-2010 (denoted "2000"), 2021-2040 (denoted 2030), 2041-2060 (denoted 2050), 2061-2080 (denoted 2070), and 2081-2100 (denoted 2090) for average monthly precipitation, and maximum (tmax) and minimum (tmin) temperatures. Processing of these data at PIK resulted in calculated mean monthly climatologies for each time slice and for each variable from the original transient daily GCM time series. The mean monthly fields were then interpolated from the original resolution of each GCM to 0.5 degrees latitude-longitude using conservative remapping (which preserves the global averages). We then 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.

On inspection it was clear that the responses of the chosen models 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 four models for scenario A1B were clustered for precipitation, tmax and tmin 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 to 90 percent of the pixels in for any given model.

Polynomials were fitted through the cluster means by date (constrained through the origin) and showed that in many cases a quadratic fit over time would have sufficed, but there were numerous cases where only a fourth-order polynomial would suffice. We therefore decided to fit fourth-order polynomials throughout. Fourth-order polynomial fits were made for all models at all scenarios and another set was made for the average of the four models. World maps of the residual surfaces were constructed 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 of the models.

Figure 2. *Scheme of the analysis*
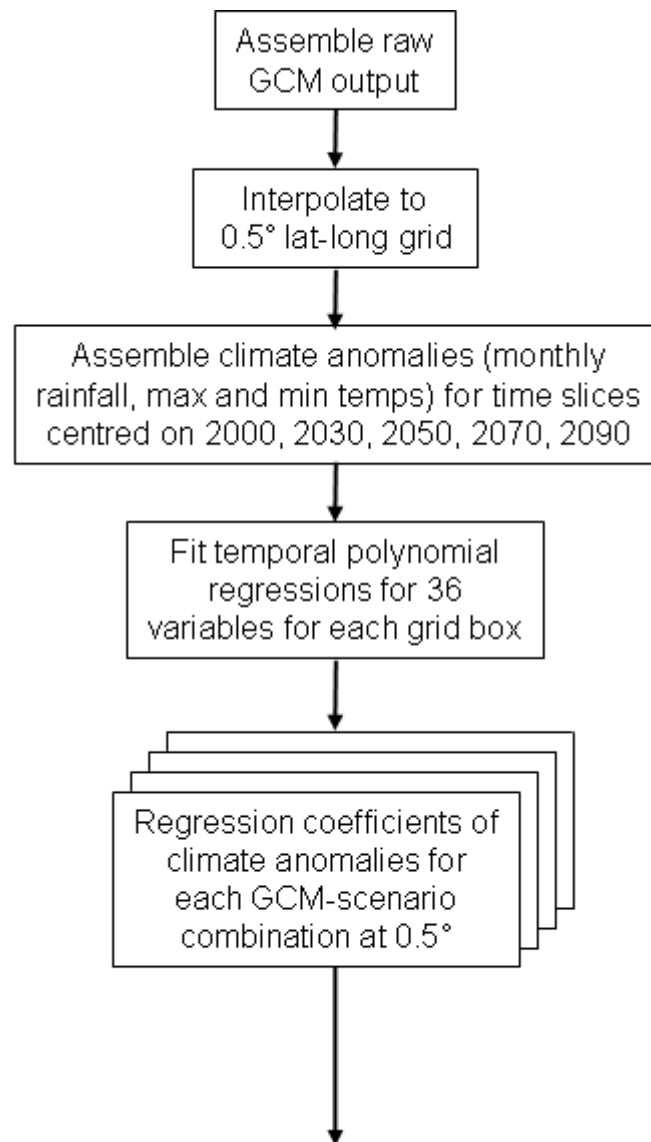
The polynomial coefficients were condensed into a data file structure for ready retrieval on a pixel-by-pixel basis (at a resolution of 30 arc-minutes) for use in subsequent operations. This preliminary process is summarised in Figure 2. The next section describes two processes: downscaling the anomalies to a higher resolution (type 1 above), and generating daily weather data that are characteristic, to some extent, of the future climatologies produced (using aspects of downscaling methods 3 and 5 above), using a stochastic daily weather generator.

4 Generating daily data: MarkSim

MarkSim$^®$ is a third-order Markov rainfall generator (Jones and Thornton, 1993; 1997; 1999; 2000; Jones et al., 2002) that has been developed over 20 years or so. It was not designed as a GCM downscaler, but it does now work as such, employing both stochastic downscaling and weather typing on top of the type 1 (bad) downscaling outlined above in section 2.

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 tropical climates, whereas a lower-order model may suffice for temperate climates (Jones and Thornton, 1993). The crux to the efficiency of MarkSim in simulating the actual variance of rainfall observed both in the tropical and temperate regions is its innovative use of resampling of the markov process parameters. To do this we need the 12 monthly baseline transfer probabilities (i.e., the probability of a wet day following three consecutive dry days), the probability coefficients related to the effect of each of the three previous days, and the correlation matrix of the 12 baseline probabilities.

MarkSim therefore works from a large set of parameters; including those for rainstorm size, this totals 117. To make a globally valid model that does not need recalibration every time that it is used, we have constructed a calibration set of over 10,000 stations worldwide. These were clustered into 702 climate clusters using the 36 values of monthly precipitation and monthly maximum and minimum temperatures. Almost all except a small number of the calibration stations have more than 10 years of (almost) continuous data. Most stations have 15-20 years of data; a few have 100 years or more. The 117 parameters of the MarkSim model are calculated by regression from the cluster most representative of the climate point to be simulated.

MarkSim estimates daily maximum and minimum air temperatures and daily solar radiation values from monthly means of these variables, using the methods originating with Richardson (1981). Monthly solar radiation values are estimated from temperatures, longitude and latitude using the model of Donatelli and Campbell (1997).

MarkSim guarantees that in the long run the values used as a starting point for a simulation series will be returned as the average of the simulated series. This is to be expected in a valid

weather simulator. If this were the only thing it could do, it would not be judged a good downscaler. When GCM differentials are added to the starting values, not only may the regression values for the coefficients change, but also they may completely change the climate cluster that is associated with that point. This means that the simulated climate has been shifted to a different type. Thus we have a form of "climate typing", although not one the original coiners of that phrase would recognise as such.

This does raise a question that we are currently addressing: when does a GCM differential addition take us out of our current cluster space? As yet we do not know. We can calculate just how far any given climate on earth is outside the MarkSim current cluster space, and we have found that about 20% are more that two standard deviations from a calibrated cluster, based on WorldClim, a 1-km interpolated climate grid for the globe (Hijmans et al., 2005). There are two points to make here. First, we can improve the current calibration considerably. We already have a wealth of new data to incorporate in the next MarkSim calibration, and this can be done given appropriate time and resources. Second, we need to look carefully at the climates that are going to occur with global warming. This is quite a problem. We have quite good estimates of future climates from GCMs, but we have no good estimates of future weather. All downscaling relies on the fact that we have got something here and now to scale it to. When a GCM differential puts a point out of the range of MarkSim's simulation clusters then we can only extrapolate from the nearest climate we have. We can hope that not that many pixels on the earth fall into that situation in the near future, but for more distant future climates, the situation is highly uncertain.

5 Using Marksim

A central part of MarkSim is the concept of a climate record. This is independent of the scale of the data, but is constant in its form and acceptability to the rest of the MarkSim software. A climate record includes the information shown in Table 3. It includes the temporal phase angle, that is, the degree by which the climate record is "rotated" in date. This rotation is done to eliminate timing differences in climate events, such as the seasons in the northern and southern hemispheres, so that analysis can be done on standardised climate data. The climate record is rotated to a standard date, using the 12-point Fast Fourier transform, on the basis of the first phase angle calculated using both rainfall and temperature. A discussion of the

methods used can be found in Jones (1987) and Jones et al. (2002). In MarkSim, almost all operations are done in rotated date space.

Table 3. Standard climate record format

Variables	FORTRAN format
latitude (decimal), longitude (decimal), elevation (m), temporal phase angle	2f10.5,i5,f7.4
monthly rainfall (mm)	12f5.0
mean daily temperature for each month (°C)	12f.5.1
diurnal temperature range for each month (°C)	12f5.1

The estimated GCM differential values are added to the rotated record. This is (bad) downscaling of type 1 above; inverse square distance weighting is used over the valid elements of the nearest nine GCM cells. This can be done with a climate database such as WorldClim (Hijmans et al., 2005), although prerotated MarkSim datasets are available. WorldClim may be taken to be representative of current climatic conditions (most of the data cover the period 1960-1990). It uses historical weather data from a number of databases. WorldClim uses thin plate smoothing with a fixed lapse rate employing the program ANUSPLIN. The algorithm is described in Hutchinson (1997).

The GCM4_MODULE

A series of FORTRAN data structures were developed along with the relevant operational programs in a FORTRAN object oriented module. This module can be used from a FORTRAN 90 program by simply specifying

USE GCM4_MODULE

A current climate record (defined as CLIMATE_RECORD in the module) can then be used to generate data for any year (more properly, any time slice, with the year the centre of the time slice), for any of the four GCMs, and for any of the three SRES scenarios included.

A simple function statement,

$$M = \text{GCM4 } (Q, \textit{Year}, \textit{Model}, \textit{Scenario})$$

returns a modified climate record *M* (see Table 3), where *Q* is the current climate record (where *M* may equal *Q* if desired), *Year* is the centre of the time slice wanted (between 2000 and 2099), *Model* is one of the GCM codes shown in column 5 of Table 2, and *Scenario* is one of A2, A1 (i.e., A1B), and B1. The variable *Model* can also take the value "av", which will provide the mean values for the four GCMs included.

An example of this process is shown in Figure 3, for the grid cell at latitude 7.0 °N, longitude 37.0 °E in south-central Ethiopia. The climate diagram, generated by the MarkSim software itself, shows a strongly uni-modal rainfall distribution with a peak in July, on average. The climate record is shown below the map, and this climate falls within a specific cluster in MarkSim, number 134. The second record in Figure 3 shows the rotated climate record for this grid cell (note that the peak rainfall month in rotated phase space is the second month, rather than the seventh month as in the unrotated record). The third record in Figure 3 shows the generated record for 2050 using the ECHam5 GCM for the A2 scenario. The rainfall amount and distribution is not projected to change much, but the average temperature throughout the year is increased by about 2°C, it seems. The future climate for 2050 for this grid cell now belongs to another MarkSim cluster, number 56.

There are many ways in which this basic information can be used (with care). For example, daily data can be used to calculate lengths of growing period for current conditions and future climate scenarios. Recently we used such information to identify areas of sub-Saharan Africa in which cereal cropping may become increasingly risky in the future, where the increased probabilities of failed seasons may mean that people might need to shift from cropping and increase their dependence on livestock (Jones and Thornton, 2009). GCM4_MODULE can also be used to generate daily weather for future climates that can be used by impact modelling software to drive various models. An example is the Decision Support System for Agrotechnology Transfer (DSSAT; ICASA, 2007), with which a wide range of crop models can be run (such as CERES-Maize, for example). We recently used these methods to assess possible changes in yields of maize and beans in East Africa (Thornton et al., 2009).

Figure 3. *A climate record for the grid cell at 7.0 °N 37.0 °E in south-central Ethiopia, its rotated record (MarkSim cluster 134), and the generated record for 2050 using the ECHam5 model and the SRES A2 scenario (MarkSim cluster 56).*
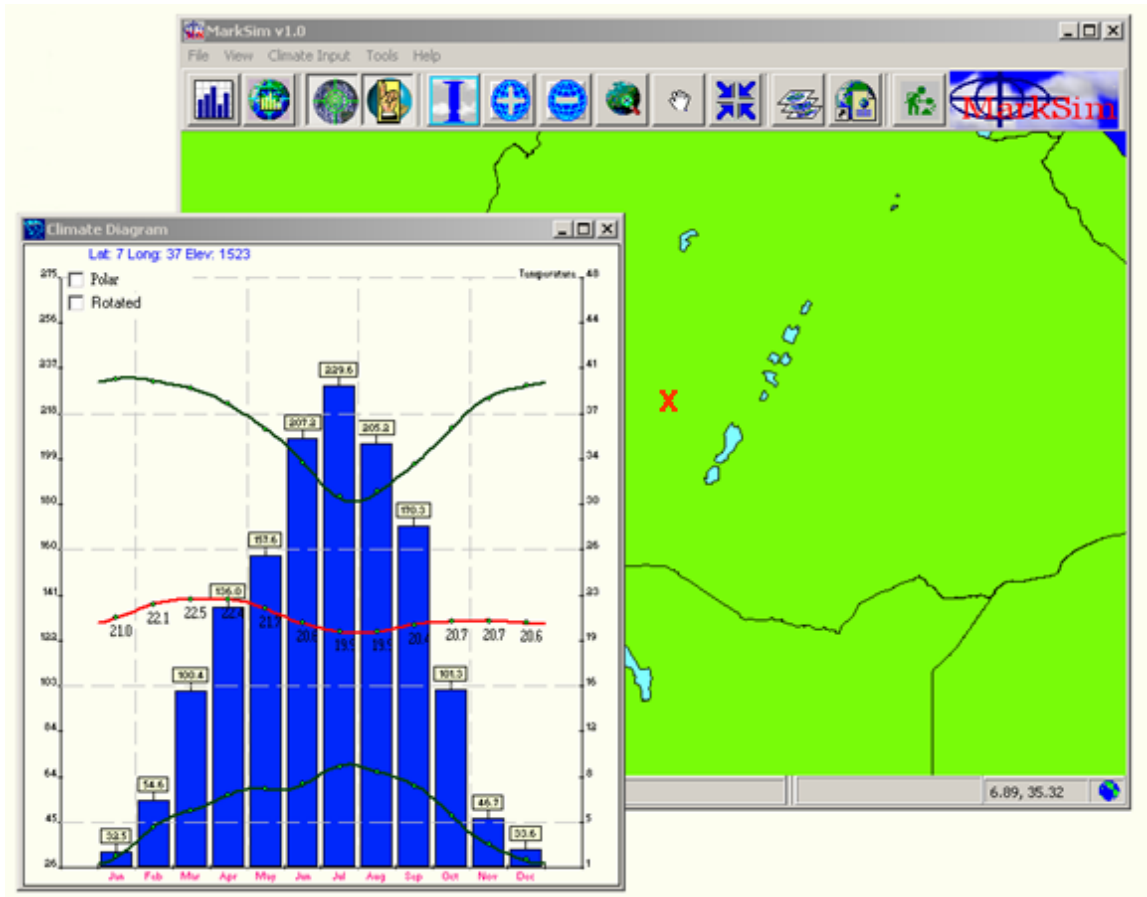

```
Current climate, "calendar" format

Lat   7.000000    Long   37.00000    Elev       1862 Phase  0.000
Rain    37.  56. 116. 179. 179. 189. 211. 193. 173. 126.  69.  33.
Temp   19.7 20.5 20.7 20.1 19.5 18.7 17.8 17.8 18.5 18.9 19.4 19.2
Diurn  16.9 16.3 15.4 13.9 13.0 11.5 10.1 10.3 11.9 14.3 15.5 16.7
```

```
Current climate, rotated format

Lat   7.000000    Long   37.00000    Elev       1862 Phase  2.772
Rain   198. 209. 187. 164. 108.  55.  30.  42.  67. 140. 184. 178.
Temp   18.4 17.7 18.0 18.6 19.1 19.4 19.2 20.0 20.6 20.6 19.9 19.3
Diurn  10.9 10.0 10.6 12.6 14.8 15.8 17.0 16.7 16.1 14.9 13.6 12.7
```

```
"Future" climate, rotated format

Lat   7.000000    Long   37.00000    Elev       1862 Phase  2.869
Rain   201. 229. 207. 158.  98.  51.  36.  47.  74. 151. 177. 158.
Temp   21.0 19.7 19.9 20.4 21.1 21.4 21.3 22.0 22.7 22.5 21.8 21.8
Diurn  11.3  9.8 10.5 12.8 14.4 15.5 16.5 16.0 16.0 14.8 13.3 13.1
```

Using the information in different ways can have substantial effects on the time taken to run the programmes, especially over large spatial areas at higher resolutions. For example, characterising Africa at 5-minute resolution in terms of average temperatures to 2050 (say), if that is all that is wanted, can be done in a few hours. Generating 100 years of daily data for each agricultural pixel in Africa at this resolution may take many hours (even days) of computer time. Running the software on a computer cluster could obviously cut the time required enormously, and we are starting to be able to do this.

6 Concluding comments

All downscaling activity is affected by considerable uncertainties of different types. First, even from the GCMs themselves, it is clear that present and future predictability of climate variability and climate change is not the same everywhere, and that gaps in knowledge of basic climatology are revealed by a lack of agreement between climate models in some regions (Wilby, 2007). While there is now higher confidence in projected patterns of warming and sea-level rise, there is less confidence in projections of the numbers of tropical storms and of regional patterns of rainfall over large areas of Africa, south Asia and Latin America. This highlights the importance of using different scenarios and different models to assess likely climate changes and their impacts. Second, our understanding is limited of what the local-level impacts of climate change are likely to be, which means that it is very difficult to evaluate the adequacy of different downscaling techniques. Third, there is a significant gap between the information that we currently have at seasonal time scales and the information we have at longer time scales -- information about what is likely over the next three to 20 years is largely missing (Washington et al., 2006). This is problematic, as the medium-term time scale is vital for political negotiation, for assessing vulnerability, and for agricultural planning, for example.

As noted above, the utility of MarkSim as a downscaling tool could be considerably strengthened by the addition of large numbers of additional calibration stations. This might lead to more information being extractable form downscaled GCM data on the nature of the variability of weather that is associated with different climate clusters. Without this, the lack of information on future weather variability associated with future climatologies is likely to remain a stumbling block to impact assessment studies. In the meantime, it may have to be incorporated on the basis of sensitivity analyses that involve manual changes to the

parameters in weather generators that control variability. How this can be done in any sensible fashion is not that clear, however.

References

Déqué M, Dreveton C, Braun A, Cariolle D, 1994. The ARPEGE/IFS atmosphere model: A contribution to the French community climate modeling. Climate Dynamics 10, 249–266.

Donatelli M, Campbell G S, 1997. A simple model to estimate global solar radiation. PANDA Project, Subproject 1, Series 1, paper 26, ISCI, Bologna, Italy. 3 pp.

Gordon H B et al., 2002. The CSIRO Mk3 Climate System Model. CSIRO Atmospheric Research Technical Paper No. 60, Commonwealth Scientific and Industrial Research Organisation Atmospheric Research, Aspendale, Victoria, Australia, 130 pp, http://www.cmar.csiro.au/e-print/open/gordon_2002a.pdf.

Hartigan J A, 1975. Clustering Algorithms. John Wiley & Sons, New York. 351 pp.

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

Hutchinson M F, 1997. ANUSPLIN Version 3.2 Users Guide. The Australian National University, Centre for Resource and Environmental Studies, Canberra, Australia, 39 pp.

ICASA, 2007. The International Consortium for Agricultural Systems Applications website. Online at http://www.icasa.net/index.html

IPCC (Intergovernmental Panel on Climate Change), 2001. Climate Change 2001: the Scientific Basis. Cambridge University Press, Cambridge, UK.

IPCC (Intergovernmental Panel on Climate Change), 2007. Climate Change 2007: Impacts, Adaptation and Vulnerability. Summary for policy makers. Online at http://www.ipcc.cg/SPM13apr07.pdf

Jones P G, 1987. Current availability and deficiencies data relevant to agro-ecological studies in the geographical area covered in IARCS. In: A H Bunting (ed.), Agricultural environments: Characterisation, classification and mapping. CAB International, UK. p. 69-82.

Jones P G, Thornton P K, 1993. A rainfall generator for agricultural applications in the tropics. Agricultural and Forest Meteorology 63, 1 19.

Jones P G, Thornton P K, 1997. Spatial and temporal variability of rainfall related to a third-order Markov model. Agricultural and Forest Meteorology 86, 127-138.

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

Jones P G, Thornton P K, 2000. MarkSim: Software to generate daily weather data for Latin America and Africa. Agronomy Journal 93, 445-453.

Jones P G, Thornton P K, 2009. Croppers to livestock keepers: Livelihood transitions to 2050 in Africa due to climate change. Environmental Science and Policy 12, 427-437.

Jones P G, Thornton P K, Diaz W, Wilkens P W, 2002. MarkSim: A computer tool that generates simulated weather data for crop modeling and risk assessment. CD-ROM Series, CIAT, Cali, Colombia, 88 pp.

Jones P G, Romero J, Amador J, Campos M, Hayhoe K, Marin M, Fischlin A, 2005. Generating high-resolution climatic change scenarios for impact studies and adaptation with a focus on developing countries. Pp 38-56 in: C Robledo, M Kanninen, L Pedroni (eds), Tropical forests and adaptation to climate change: in search of synergies. CIFOR, Bogor, Indonesia.

K-1 Model Developers (2004). K-1 Coupled Model (MIROC) Description. K-1 Technical Report 1, H Hasumi, S Emori (eds). Center for Climate System Research, University of Tokyo, Tokyo, Japan, 34 pp., http://www.ccsr.u-tokyo.ac.jp/kyosei/hasumi/MIROC/tech-repo.pdf.

Meehl G A, Stocker T F, Collins W D, Friedlingstein P, Gaye A T, Gregory J M, Kitoh A, Knutti R, Murphy J M, Noda A, Raper S C B, Watterson I G, Weaver A J, Zhao Z C, 2007. Global Climate Projections. In: Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Mitchell T D, Carter T R, Jones P D, Hulme M, 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 Working Paper 55.

Nakicenovic N, Alcamo J, Davis G, de Vries B, Fenhann J, Gaffin S, Gregory K, Grübler A, Jung T Y, Kram T, Lebre la Rovere E, Michaelis L, Mori S, Morita T, Pepper W, Pitcher H, Price L, Riahi K, Roehrl A, Rogner H H, Sankovski A, Schlesinger M, Shukla P, Smith S, Swart R, van Rooijen S, Victor N, Dadi Z, 2000. Emissions Scenarios: A Special

Report of the Intergovernmental Panel on Climate Change (IPCC). Cambridge University Press, Cambridge, 509 pp. Online at http://www.grida.no/climate/ipcc/ emission/index.htm

Randall D A, Wood R A, Bony S, Colman R, Fichefet T, Fyfe J, Kattsov V, Pitman A, Shukla J, Srinivasan J, Stouffer R J, Sumi A, Taylor K E, 2007. Climate Models and Their Evaluation. In: Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Richardson C W, 1981. Stochastic simulation of daily precipitation, temperature and solar radiation. Water Resources Research 17 (1), 182-190.

Roeckner E et al., 2003. The Atmospheric General Circulation Model ECHAM5. Part I: Model Description. MPI Report 349, Max Planck Institute for Meteorology, Hamburg, Germany, 127 pp.

Thornton P K, Jones P G, Owiyo T, Kruska R L, Herrero M, Kristjanson P, Notenbaert A, Bekele N, Omolo A, with contributions from Orindi V, Ochieng A, Otiende B, Bhadwal S, Anantram K, Nair S, Kumar V and Kelkar U, 2006. Mapping climate vulnerability and poverty in Africa. Report to the Department for International Development, ILRI, Nairobi, Kenya, May 2006, 200 pp. Online at http://www.dfid.gov.uk/research/mapping-climate.pdf

Thornton P K, Jones P G, Alagarswamy A, Andresen J, 2009. Spatial variation of crop yield responses to climate change in East Africa. Global Environmental Change 19, 54-65.

Washington R, Harrison M, Conway D, Black E, Challinor A, Grimes D, Jones R, Morse A, Kay G, Todd M, 2006. African climate change: taking the shorter route. Bulletin of the American Meteorological Society, 87, 1355-1366.

Wilby R, 2007. Decadal forecasting techniques for adaptation and development planning. Report to DFID, August 2007.