MarkSim_standalone.V2 for DSSAT users

P. G. Jones November 2013, Waen Associates, Gwynedd.

Introduction

Marksim_standalone.exe combined the basic routines from the original MarkSim version, produced at CIAT (International Center for Tropical Agriculture) and distributed on CD (Jones et.al., 2002), and the new MarkSimGCM web application (Jones et al. 2011a, b) that runs off Google Earth. The standalone version is designed for computer users that need to process a large amount of data. It eliminates their picking at a keyboard and abusing their eyesight by searching an on-screen map for the required data point.

The original MarkSim_Standalone used six GCMs from the fourth approximation of the IPCC, the new one use 17 models from the CMIP5 range that were considered in the last IPCC report.

It is designed to be used with DSSAT http://dssat.net/ and so it uses CLI files (see description below), existing, or constructed by the user, to create new CLI files and WTG files for the simulated weather data under a range of GCMs and Representative Concentration Pathways (RCP).

The EXE runs from the DOS prompt (or a system program call). It will read a directory, work out a list of the CLI files in that directory, and create subdirectories for each CLI file. In each of these it will place up to 99 replicates of yearly weather data and a new CLI file describing the climate under the particular situation of GCM/scenario/year for which the user called.

MarkSim

MarkSim was developed in the 1980s and 1990s to simulate weather from known sources of monthly climate data from around the world (Jones & Thornton 1993, 1997, 1999, 2000). It divides the world into 720 clusters of climate that are distinct from one another and fits a third order Markov model to the precipitation data. The temperature data simulation is derived from SIMMETEO (Geng et al., 1988). The radiation data are based on the model of Donatelli and Campbell (1997).

For each of the about 9,200 stations with usable daily data, the third order Markov chain model of rainfall was developed. The results were grouped according to climate cluster, and regression equations for each of the Markov parameters were calculated using the monthly average rainfall and temperature figures for each station within the cluster. The model can now be fitted to any monthly climate data record by determining to which cluster it belongs, and using the relevant regression equations.

Markov models are excellent simulations of the rainfall process, but they do have limitations. It was found quite early in the analysis (Jones & Thornton, 1993) that a simple first or second order process, while adequate for temperate regions, would not fit well enough for the tropics. Hence we went with the third order model. Another deficiency we found was that Markov processes ordinarily underestimate the variance of rainfall. This is solved in MarkSim by resampling the Markov probability coefficients for each year of simulation. This is because the coefficients themselves are only estimates and have an error term. Once we reincorporate this error term by resampling, the rainfall variance agrees with the observed.

The latest publication we have for MarkSimGCM and for this application is that for the 4th approximation models. Jones & Thornton (2013)

GCM Regression

Table 1 shows the models used in the application. We obtained (Brown, O. 2013) yearly data interpolated by bilinear interpolation from the original GCM pixel sizes. 'Historic' data were for the period ran from 1961 to 2005. Prediction data ran from 2006 to 2099. We calculated means for rainfall, tmax and tmin from the 'historic' data and subtracted them from every year's image of the GCM future data.

The process we used is the delta process in it simplest form, all variables were adjusted arithmetically as a deviation from the 'historic' mean, and bias corrected to WorldClim.v1.3

The GCM data provided annual deviations for the years 2006 to 2099. We fitted 4th order polynomials to every pixel. Inspection of samples of the results showed that this was a rational assumption. The residuals from the regressions matched very closely the mean square residuals of the 'historic' data. We have yet to test this against real historic data, but we feel that the result would be similar.

Table 1. Atmosphere-Ocean General Circulation Models used in the work

	Model	**Institution**	**Resolution, Lat x Long °**	**Reference**
1	BCC-CSM 1.1	Beijing Climate Center, China Meteorological Administration	2.8125 x 2.8125	Wu T (2012). A Mass-Flux Cumulus Parameterization Scheme for Largescale Models: Description and Test with Observations. Clim. Dynam. 38, 725–744
2	BCC-CSM 1.1(m)	Beijing Climate Center, China Meteorological Administration	2.8125 x 2.8125	Wu T (2012). A Mass-Flux Cumulus Parameterization Scheme for Largescale Models: Description and Test with Observations. Clim. Dynam. 38, 725–744
3	CSIRO-Mk3.6.0	Commonwealth Scientific and Industrial Research Organisation and the Queensland Climate Change Centre of Excellence	1.875 x 1.875	Collier MA et al. (2011) The CSIROMk3.6.0 Atmosphere-Ocean GCM: participation in CMIP5 and data publication. MODSIM 2011, Perth, 12–16 December 2011
4	FIO-ESM	The First Institute of Oceanography, SOA, China	2.812 x 2.812	Song Z, Qiao F, Song Y (20120. Response of the equatorial basin-wide SST to wave mixing in a climate model: An amendment to tropical bias, J. Geophys. Res., 117, C00J26
5	GFDL-CM3	Geophysical Fluid Dynamics Laboratory	2.0 x 2.5	Donner LJ et al. (2011). The dynamical core, physical parameterizations, and basic simulation characteristics of the atmospheric component AM3 of the GFDL Global Coupled Model CM3. Journal of Climate, 24(13).
6	GFDL-ESM2G	Geophysical Fluid Dynamics Laboratory	2.0 x 2.5	Dunne JP et al. (2012). GFDL's ESM2 Global Coupled Climate–Carbon Earth System Models. Part I: Physical Formulation and

				Baseline Simulation Characteristics. J. Climate, 25, 6646–6665.
7	GFDL-ESM2M	Geophysical Fluid Dynamics Laboratory	2.0 x 2.5	Dunne JP et al. (2012). GFDL's ESM2 Global Coupled Climate–Carbon Earth System Models. Part I: Physical Formulation and Baseline Simulation Characteristics. J. Climate, 25, 6646–6665.
8	GISS-E2-H	NASA Goddard Institute for Space Studies	2.0 x 2.5	Schmidt GA et al. (2006). Present day atmospheric simulations using GISS ModelE: Comparison to in-situ, satellite and reanalysis data. J. Climate 19, 153-192.
9	GISS-E2-R	NASA Goddard Institute for Space Studies	2.0 x 2.5	Schmidt GA et al. (2006). Present day atmospheric simulations using GISS ModelE: Comparison to in-situ, satellite and reanalysis data. J. Climate 19, 153-192.
10	HadGEM2-ES	Met Office Hadley Centre	1.2414 x 1.875	Collins WJ et al. (2011). Development and evaluation of an Earth-System model-HadGEM2. GMD 4(4):1051–1075.
11	IPSL-CM5A-LR	Institut Pierre-Simon Laplace	1.875 x 3.75	Dufresne JL et al. (2013). Climate change projections using the IPSL-CM5 Earth System Model: from CMIP3 to CMIP5. Climate Dynamics, 1-43.
12	IPSL-CM5A-MR	Institut Pierre-Simon Laplace	1.2587 x 2.5	Dufresne JL et al. (2013). Climate change projections using the IPSL-CM5 Earth System Model: from CMIP3 to CMIP5. Climate Dynamics, 1-43.
13	MIROC-ESM	Atmosphere and Ocean Research Institute (The University of Tokyo), National Institute for Environmental Studies, and Japan Agency for Marine-Earth Science and Technology	2.8125 x 2.8125	Watanabe S et al. (2011). MIROC-ESM2010: model description and basic results of CMIP5-20c3m experiments. Geoscientific Model Development 4 (4), 845–872.
14	MIROC-ESM-CHEM	Atmosphere and Ocean Research Institute (The University of Tokyo), National Institute for Environmental Studies, and Japan Agency for Marine-Earth Science and Technology	2.8125 x 2.8125	Watanabe S et al. (2011). MIROC-ESM2010: model description and basic results of CMIP5-20c3m experiments. Geoscientific Model Development 4 (4), 845–872.
15	MIROC5	Japan Agency for Marine-Earth Science and Technology, Atmosphere and Ocean Research Institute (The University of Tokyo), and National Institute for Environmental Studies	1.4063 x 1.4063	Watanabe M et al. (2010). Improved Climate Simulation by MIROC5: Mean States, Variability, and Climate Sensitivity. J. Climate, 23, 6312–6335.
16	MRI-CGCM3	Meteorological Research Institute	1.125 x 1.125	Yukimoto S (2012). A new global climate model of Meteorological Research Institute: MRI-CGCM3 – Model description and basic performance. J. Meteorol. Soc. Jpn., 90a, 23–64
17	NorESM1-M	Norwegian Climate Centre	1.875 x 2.5	Kirkevag A, Iversen T, Seland O, Debernard JB, Storelvmo T, Kristjansson JE (2008) Aerosol-cloud-climate interactions in the climate model CAM-Oslo.Tellus A

				60(3):492–512. Seland O, Iversen T, Kirkevag A, Storelvmo T (2008). Aerosol-climate interactions in the CAM-Oslo atmospheric GCM and investigation of associated basic shortcomings. Tellus A 60(3):459–491.

Data

You need to have a space of a bit more than 2Gb to take the driving data. The GCM coefficients are 17 sets of four binary files for each GCM name coded to four characters, plus one which is the ensemble of all 17, denoted 'ALL_'. For other ensemble options the software will create one on the fly as you specify it. The other files are the MarkSim basic parameters which specify the model characteristics based on 702 climate clusters. They are all in the download as DATA.rar and should all be expanded into the same directory that you will reference as data in your subsequent calls to MarkSim_Standalone_V1 as Path1

Running the EXE

MarkSim_standalone.exe can be run from any directory under the DOS prompt or by system call. It requires six arguments with an optional seventh; the call from DOS is:

MarkSim_standalone path1 path2 template RCP year nreps seed

The arguments are:

Path1: The path to the directory in which you have placed the MarkSimGCM data file (see Data above).

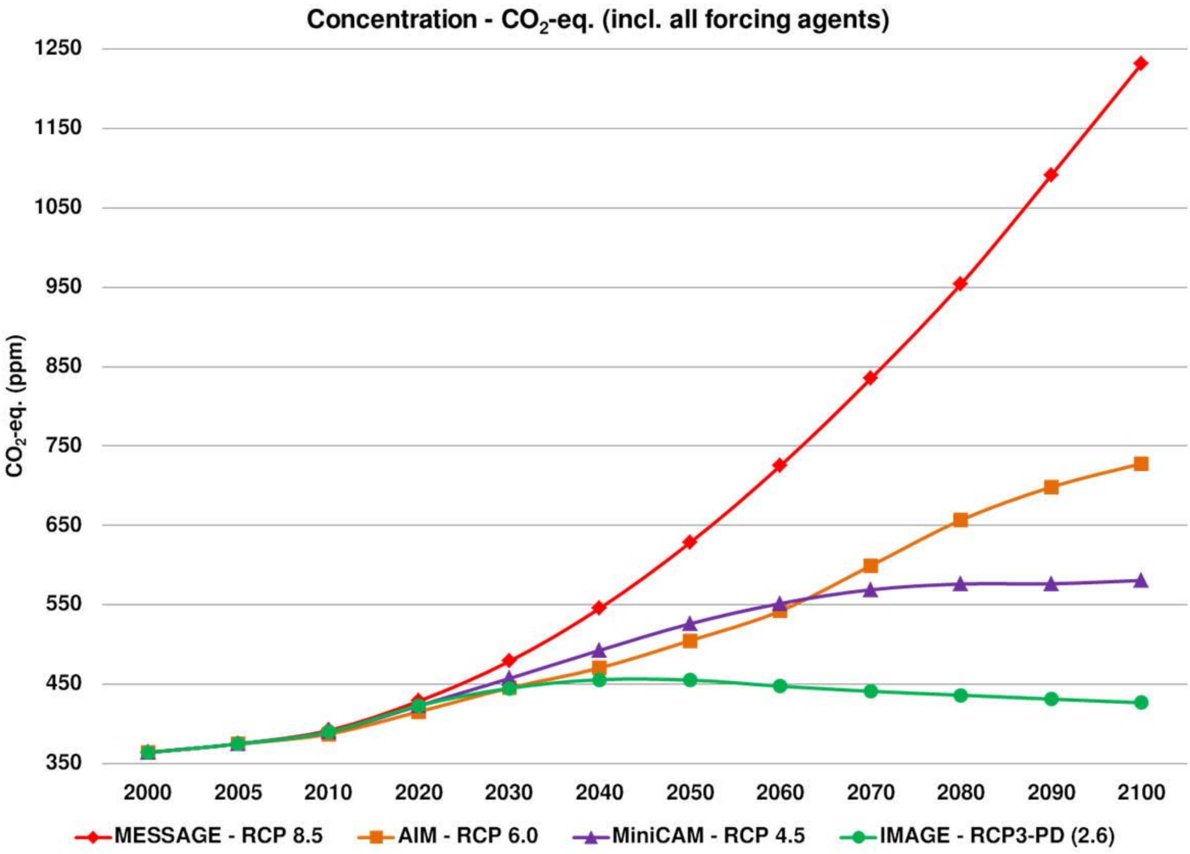
Path2: The path to the directory in which the CLI files reside. Note that this directory will be expanded by a subdirectory for every CLI file and for every use of MarkSim_standalone with different GCM, scenario, and year arguments.

template: There is a big difference in this parameter to that which we had previously. Now you have 17 GCMs to choose from, not singly, you can choose any combination, including ALL or NONE. So we have a character array as a parameter. 0 for out , 1 for in. It's 17 characters long and can only take the values 1 or 0.

Thus:
'00001000000000000' means GFDL-CM3
'00000000000000000' means Present day climate – no GCM
'11111111111111111' means Ensemble of all 17 GCMs
'10001000100001000' means Ensemble of BCC-CSM 1.1 , GFDL-CM3, GISS-E2-R and MIROC5

I'm sorry if this seems complicated. There are actually 131072 different combinations of the 17 CGMs. It's actually fairly simple to combine the regression coefficients, so I thought it might be of some use to those of you who want different combinations of GCMs.

RCP: **Representative Concentration Pathways (RCPs)** are the four greenhouse gas concentration trajectories adopted by the IPCC for its fifth assesment



Copied, courtesy of Wikipedia
http://en.wikipedia.org/wiki/Representative_Concentration_Pathways

Coding for this parameter drops the decimal point and the space, so it takes the values rcp26, rcp45, rcp60 or rcp85.

Year: Is the year you wish to simulate. This can take the values 2013 to 2099

If you wish to simulate ‘present’ climate, as represented by WordClim.v1.3 (i.e. average of a varying number of years over the last 50) use the ‘000000000000000000’ option on the template parameter. In this case the year parameter still needs to be there as it is checked but then ignored.

Nreps: Is the number of years to be simulated. NOTE these are replicates of the year that you specify and are NOT a series of years from that date!

Seed: This was originally planned as an optional parameter, but it does not appear to work as such. Therefore allways use a random number seed as this parameter. It must be an odd integer, preferably 4 or more characters long.

CLI files and naming conventions

DSSAT convention is for four character CLI names, thus the program cannot add the codes for the model, scenario, or simulation date. Therefore it creates a new subdirectory with the name of the 'CLI file//GCM code//scenario code//year' and writes the modified CLI file therein under its own name along with the required number of years of simulated data. Since users may be using a CLI name greater than four characters outside of DSSAT, a file name length of 20 characters is permitted.

The simulated data files are named ccccnn01.WTG where cccc are the first four characters of the CLI name and nn the year number 01-99 to conform to DDSAT naming. *This means that all files in a run with multiple CLI files carry the same names and are differentiated only by the directory in which they reside*. Note that in MarkSim output these do not represent a series of years, but are replicates of the target year.

MarkSim_standalone will search a directory for all the CLI files it contains. Other files in the directory are ignored. A maximum of 32,000 CLI files is allowed in each directory. If more than these are needed, then we suggest that you place them in multiple directories and call the EXE for each directory.

On a reasonable PC, the EXE takes about 4 seconds to load the GCM regression data and exceeds the simulation run time by more than 10^6. This happens once for each directory call, so it is advantageous to place as many CLI files together as possible to minimize loading time. *Placing a single CLI file in each directory and calling the EXE with a BAT that produces one simulation run per directory is therefore a slow option.*

Error reporting

The EXE is designed to run in the DOS prompt and will hard stop and write error messages. There has been a major upheaval in the source code and the possible error messages have multiplied, but the normal user should not be put off by the following list. It is the list of all possible error messages that could occur from the MarkSim system. Most of these are only of interest to the systems programmer, i.e. p.jones@cgiar.org

Those that are in the realm of the user are denoted in red, the others are green. If you see a green one please let me know because something is definitely wrong. If one appears that is not in the list please call me immediately because that should be definitely impossible. . Quite a lot of these are duplicates, but they come from different parts of the application. It helps me to know where they're coming from if you get into problems.

1. Error. Cannot construct clx with an unrotated climate record
2. Error. Radiation out of range in camdon, cannot be negative
3. Error. Non-convergence in root
4. Error. Av unusually high
5. Error. You have not called markset_direct

6. Error. You have not loaded the cluster data with loadclus

20. Error. MarkSim Data file not available, Make sure path is correct' , return_code

The return code (7-20) denotes the faulty data, but since all base data should be in one directory it either means that the directory address is wrong or that you haven't loaded the full data

21. Error. Error reading CLI file', return_code

In this case the return code is informative.

21	means	Error opening CLI file
22	means	Error reading CLI file EOF on first record
24	means	Unable to read @ INSI
25	means	Unable to read @START
33	means	Unable to read @GSST
34	means	Unable to read @MONTH

Which might help you if you've written the CLI file yourself

23. Error. Finding nearest met record

26. Error. Maths error in dcpmat , return_code

A serious math error that should not happen

29. Error. Probability out of range in probit

Another one, but possibly more likely

30. Error reading WTG file in read_year_met_data

31. Error in mean_climate. Climates not from same place must all be rotated

32. Error in mean_climate. Climates from same place must, all not be rotated, or all rotated the same'

40. Error, No data loaded at rmse_grid

41. Error, No data loaded at gcm5

42. Error, Climate structure rotated in gcm5

43. Error, Opening bin file in read_in_data

Make sure your directory for the MarkSim data is correct

44. Error, Reading bin file in read_in_data

Make sure your directory for the MarkSim data is correct

45. Error, Opening bin in read_in_emsemble

Make sure your directory for the MarkSim data is correct

46. Error, In template in read_in_ensemble

The template must contain exactly 17 characters, only 0's (zero) or 1's (one)

47. Error, Reading bin in read_in_ensemble

Make sure your directory for the MarkSim data is correct

101. Error. Unable to write new CLI file

Either an unformed directory, or bad filename. Should not happen

103. Error. Opening WTG file

Should not happen as the program opens the directory

104. Error. Writing WTG file

Can only happen with a hardware failure

105. Error. GCM path incorrect or longer than 80 characters

106. Error. Incorrect path to CLI Directory, may be longer than 80 characters

107. Error. In GCM template, May be the wrong length

108. Error. RCP must be a 5 character code rcp26, or RCP85 etc

109. Error. The Year of simulation parameter could not be read.

110. Error. You can only simulate between 2013 and 2099

Unless you have specified a 00000000000000000 template, in which case your simulation is WorldClim v1.3

111. Error. Reading number of replications field

This error means the designated argument could not be read

112. Error. Number of replicates out of range. 1 to 99 are allowed

113. Error. Seed is not in satisfactory format

114. Error. Too many digits for seed

115. Error. Seed out of range. Try positive odd integer

117 Error. Incorrect entry in template at GCM , 'template'

Only characters 0 or 1 are allowed

301. Error. Path specified is not a valid path, 'path'

302. Error. Unable to list directory, error code, return code

The program uses a system call to list the directory. If this fails then this is the return code

303. Error, Unable to open directory listing

304. Error, Unable to delete list file, error code, return code

DOS return code from delete command

305. Error. Too many CLI files in directory, directory

Limit is 32000 files

306. Error, CLI name contains no characters

401. Error. Output path invalid, output path

DSSAT CLI file description for MarkSim_standalone

The CLI file required for MarkSim_standalone is a reduced version of the standard DSSAT CLI file. It does not contain either the `WGEN` parameters, the `RANGE CHECK VALUES`, or the `FLAGGED DATA COUNT`. It therefore stops after the listing of `MONTHLY AVERAGES`.

MarkSim ignores many fields, but some are transferred to the new CLI file constructed with the future climate.

```
*CLIMATE : PALMIRA-CIAT
@ INSI      LAT     LONG  ELEV   TAV   AMP  SRAY  TMXY  TMNY  RAIY
  PALM    3.500  -76.300   965  24.1  11.6 199.4  29.4  18.8  1147
```

In this line only the `INSI`, `LAT`, `LONG` and `ELEV` fields are actually used. All the others are recalculated from the new climate. If necessary they can take the value `-99.0` or `-99` and they will be replaced in the new file.

```
@START  DURN  ANGA  ANGB REFHT WNDHT SOURCE
     0     0  0.25  0.50  0.00  0.00 Calculated_from_daily_data
@ GSST  GSDU
    1   365
```

None of the values in these two lines are used, however they are transferred to the new file. The values shown here are the standard default values and should always be set to these if actual values are not available. The exception is the source field. This will be replaced with the relevant GCM model, scenario and simulation year information.

```
*MONTHLY AVERAGES
@MONTH  SAMN  XAMN  NAMN  RTOT  RNUM  SHMN
     1  18.0  29.7  18.7  81.0   9.1 -99.0
     2  18.0  29.8  18.8  75.0   7.8 -99.0
                  . . . .

    12  17.4  29.1  18.6 102.0  10.6 -99.0
```

SAMN and RNUM are recalculated for the new file and can take the missing value -99.0 if desired. SHMN is not used and will be transferred to the new file. It is, however, read and needs to be present, even if only missing value.

Selected weather data codes relevant to CLI as above

```
 AMP     Temperature amplitude, monthly averages, C
AMTH     Angstrom 'a' coefficient, monthly, unitless
ANGA     Angstrom 'a' coefficient, yearly, unitless
ANGB     Angstrom 'b' coefficient, yearly, unitless
BMTH     Angstrom 'b' coefficient, monthly, unitless
DURN     Duration of summarization period for climate files, Yr
ELEV     Elevation, m
GSDU     Growing season duration, Day
GSST     Growing season start day, Doy
INSI     Institute and site code
LAT      Latitude, degrees (decimals)
LONG     Longitude, degrees (decimals)
MTH      Month, #
NAMN     Temperature minimum,all days,monthly average, C
RAIY     Rainfall, yearly total, mm
REFHT    Reference height for weather measurements, m
RNUM     Rainy days, # month-1
RTOT     Rainfall total, mm month-1
SAMN     Solar radiation,all days,monthly average, MJ m-2 d-1
SHMN     Daily sunshine duration, monthly average, percent
SOURCE   Source of daily weather data, text
SRAY     Solar radiation,yearly average, MJ m-2 day-1
START    Start of summary period for climate (CLI) files, Year
TAV      Temperature average for whole year, C
TMNY     Temperature minimum, yearly average, C
TMXY     Temperature maximum, yearly average,
WIND     Daily wind speed (km d-1)
WNDHT    Reference height for windspeed measurements, m
XAMN     Temperature maximum,all days,monthly average, C
```

References

Brown, O. (2013) Personal communication

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

Diansky, N.A. and Volodin, Ye. M. 2002. The reproduction of modern climate by model of atmosphere-ocean general circulation model (in Russian). Izv. Atmos. Oceanic Phys. 38: 824-840.

Donatelli, M. and Campbell. G.S. 1997. A simple model to estimate global solar radiation. PANDA Project, Subproject 1, Series 1, Paper 26. Research Institute for Industrial Crops (ISCI), Bologna, Italy. 3 pp.

Furevik, T., Bentsen, M., Drange, H., Kindem, I.K.T., Kvamstø, N.G., Sorteberg, A. 2003. Description and evaluation of the Bergen climate model: ARPEGE coupled with MICOM. Clim. Dyn. 21: 27-51.

Geng, S., Auburn, J., Brandstetter, E., Li, B. 1988. A program to simulate meteorological variables. Documentation for SIMMETEO. (Agronomy Report No. 204). University of California, Davis Crop Extension, Davis, California.

Gordon, H.B., Rotstayn, L.D., McGregor, J.L., Dix, M.R., Kowalczyk, E.A., O'Farrell, S.P., Waterman, L.J., Hirst, A.C., Wilson, S.G., Collier, M.A., Watterson, I.G., Elliot, T.I. 2002. The CSIRO Mk3 Climate System Model. Commonwealth Scientific and Industrial Research Organisation (CSIRO) Atmospheric Research Technical Paper No. 60. CSIRO, Aspendale, Victoria, Australia. 130 pp. http://www.cmar.csiro.au/e-print/open/gordon_2002a.pdf (accessed December 2009).

Jones, P.G. and Thornton, P.K. 1993. A rainfall generator for agricultural applications in the tropics. Agric. For. Meteorol. 63: 1-19.

Jones, P.G. and Thornton, P.K. 1997. Spatial and temporal variability of rainfall related to a third order Markov model. Agric. For. Meteorol. 86: 127-138.

Jones, P.G. and Thornton, P K. 1999. Fitting a third order Markov rainfall model to interpolated climatic surfaces. Agric. For. Meteorol. 97: 213-231.

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

Jones, P.G and Thornton, P.K. 2013 . Generating downscaled weather data from a suite of climate models for agricultural modeling applications. Agric. Systems 114:1-5

Jones P.G., Thornton, P.K., Diaz, W., Wilkins, P.W. 2002. MarkSim. A computer tool that generates simulated weather data for crop modeling and risk assessment. Centro Internacional de Agricultura Tropical (CIAT), Cali, CO. Theory section available at: http://gisweb.ciat.cgiar.org/marksim/download/Theory.pdf

Jones, P.G., Thornton P.K., Heinke, J. 2011a. Generating characteristic daily weather data using downscaled climate model data from the Intergovernmental Panel on Climate Change (IPCC) Fourth Assessment. Available at: https://hc.box.com/shared/f2gk053td8

Jones, P.G., Thornton, P.K., Giron, E. 2011b. Web application. MarkSimGCM, A weather simulator. Available at: http://gismap.ciat.cgiar.org/MarkSimGCM/

K-1 Model Developers. 2004. K-1 Coupled Model (MIROC) Description. K-1 Technical Report No. 1, Hasumi, H., Emori, S. (Eds.) Center for Climate System Research(CCSR), National Institute for Environmental Studies (NIES), Frontier Research Center for Global Change (FRCGC), University of Tokyo, Tokyo, Japan. http://www.ccsr.u-tokyo.ac.jp/kyosei/hasumi/MIROC/tech-repo.pdf (accessed December 2009).

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: Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K.B., Tignor, M., Miller, H.L. (Eds.) 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, UK and New York, NY, USA.

Roeckner, E., Bäuml, G., Bonaventura, L., Brokopf, R., Esch, M., Giorgetta, S., Hagemann, S., Kirchner, I., Kornbluch, L., Mauzini, E., Rhodin, A., Schlese, U., Schulzweida, U., Tompkins, A. 2003. The Atmospheric General Circulation Model ECHAM5. Part I: Model Description. MPI Report 349, Max Planck Institute for Meteorology, Hamburg, Germany. 127 pp.