

Simulation of near-term climate change at target sites in West and East Africa

Working Paper No. 58

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

Arthur M. Greene, Igor Khomyakov



RESEARCH PROGRAM ON
**Climate Change,
Agriculture and
Food Security**



Working Paper

Simulation of near-term climate change at target sites in West and East Africa

Working Paper No. 58

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

Arthur M. Greene
Igor Khomyakov

Correct citation:

Greene A.M., Khomyakov I. 2013. Simulation of near-term climate change at target sites in West and East Africa. CCAFS Working Paper No. 58. CGIAR Research Program on Climate Change, Agriculture and Food Security (CCAFS). Copenhagen, Denmark. Available online at: www.ccafs.cgiar.org

Titles in this Working Paper series aim to disseminate interim climate change, agriculture and food security research and practices and stimulate feedback from the scientific community.

This document is published by the CGIAR Research Program on Climate Change, Agriculture and Food Security (CCAFS), which is a strategic partnership of the CGIAR and the Earth System Science Partnership (ESSP). CCAFS is supported by the CGIAR Fund, the Danish International Development Agency (DANIDA), the Australian Government Overseas Aid Program (AusAid), Irish Aid, Environment Canada, Ministry of Foreign Affairs for the Netherlands, Swiss Agency for Development and Cooperation (SDC), Instituto de Investigação Científica Tropical (IICT), UK Aid, and the European Union (EU). The Program is carried out with technical support from the International Fund for Agricultural Development (IFAD).

Contact:

CCAFS Coordinating Unit - Faculty of Science, Department of Plant and Environmental Sciences, University of Copenhagen, Rolighedsvej 21, DK-1958 Frederiksberg C, Denmark. Tel: +45 35331046; Email: ccaafs@cgiar.org

Creative Commons License



This Working Paper is licensed under a Creative Commons Attribution – NonCommercial–NoDerivs 3.0 Unported License.

Articles appearing in this publication may be freely quoted and reproduced provided the source is acknowledged. No use of this publication may be made for resale or other commercial purposes.

© 2013 CGIAR Research Program on Climate Change, Agriculture and Food Security (CCAFS). CCAFS Working Paper no. 58

DISCLAIMER:

This Working Paper has been prepared as an output for the Theme Integration for Decision Making – Data and Tools under the CCAFS program and has not been peer reviewed. Any opinions stated herein are those of the author(s) and do not necessarily reflect the policies or opinions of CCAFS, donor agencies, or partners.

All images remain the sole property of their source and may not be used for any purpose without written permission of the source.

Abstract

We describe the generation of synthetic sequences of precipitation and maximum and minimum daily temperatures at two locations, in western and eastern Africa respectively. The sequences are generated at the monthly time scale and incorporate both explicitly modelled annual-to-decadal variability, based on the observational record, and long-range (i.e., climate change) trends, as inferred from an ensemble of global climate models. Annual-to-decadal variability is modelled as a first-order vector autoregressive (VAR) process, and the simulations are temporally downscaled to monthly time resolution using a nonparametric resampling scheme. The modelled sequences reproduce well the observed covariances as well as serial autocorrelation in individual variables. The simulations are intended to drive agricultural or other applications models to investigate responses to a range of plausible trends, on which are superimposed decade-scale climate fluctuations whose likelihood of occurrence can be estimated.

Keywords

Stochastic Models ; climate ; Africa South of Sahara

About the authors

Arthur M. Greene, Associate Research Scientist

International Research Institute for Climate and Society (IRI)

Monell building, Lamont Campus, Columbia University

61 Route 9W, Palisades, NY 10964 USA

amg@iri.columbia.edu

Greene is a climate scientist and statistician, specializing in the application of sophisticated statistical models in climate research. He has worked, *inter alia*, with Bayesian models, hidden Markov models and information theory, in the application of climate data for societal benefit.

Igor Khomyakov, Staff Associate

International Research Institute for Climate and Society (IRI)

Monell building, Lamont Campus, Columbia University

61 Route 9W, Palisades, NY 10964 USA

@iri.columbia.edu

Khomyakov is a computer programmer on the IRI staff. For this work he adapted preexisting code written by the first author, translated various ideas into executable form, ran simulations and downscaling routines and assisted with model diagnostics and improvement.

Acknowledgements

Colleagues at the International Research Institute for Climate and Society and at the Lamont-Doherty Earth Observatory of Columbia University provided helpful discussion and suggestions during the production of this report. This work has been supported by donors to CGIAR through the CCAFS secretariat hosted by the Department of Agriculture and Ecology at the Faculty of Life Sciences at University of Copenhagen. This document has been produced with the financial assistance of the European Union, Canadian International Development Agency, World Bank, New Zealand Ministry of Foreign Affairs and Trade and Danida and with the technical support of IFAD. The views expressed herein can in no way be taken to reflect the official opinion of these agencies.

Contents

1	Introduction	8
2	Data acquisition and assessment	10
2.1	Station Kaffrine.....	10
2.2	Station Machakos	14
2.3	Discussion	20
3	Trend: The forced component	21
3.1	Kaffrine	21
3.2	Machakos	24
4	Modelling of the annual-to-decadal component.....	28
4.1	Model structure	28
4.2	Information tests	28
4.3	Reproduction of observed statistics	30
4.4	Considering cross-scale interaction.....	31
5	Projection.....	32
5.1	Future trends	32
5.2	Simulation subsampling.....	37
6	Downscaling.....	39
6.1	Nonparametric resampling scheme	39
6.2	Treatment of out-of-season months	41
7	Example simulation sequences.....	44
7.1	Low trend and fluctuation quantiles.....	44
7.2	Median quantiles	45
7.3	High quantiles	46
8	Quantifying uncertainty.....	48
8.1	Model uncertainty and regression estimates	48
8.2	Scenario uncertainty.....	48
8.3	Sampling variability.....	49
9	Discussion	50
10	Conclusion.....	52
	Appendix Description of simulations provided with the report “Simulation of near-term climate change at target sites in West and East Africa”	53
	Introduction.....	53
	Simulation structure and naming	53
	Out-of-season months	55
	Time span and resolution	55

Location of files	55
Final notes	55
References	56

1 Introduction

We have previously constructed stochastic simulation models, comprising time scales from daily to decadal and including climate change components, for the combined Berg and Breede Water Management Areas in the Western Cape Province of South Africa (*Greene et al.*, 2012) and for south-eastern South America, for a region comprising parts of Uruguay, Argentina and Brazil (*Greene et al.*, 2013). In the present study we fit separate simulation models to what are effectively two point locations on the African continent, one, Kaffrine, Senegal, lying north of the Equator in the West African Sahel, the other, Machakos, Kenya, situated just south of the Equator in East Africa.

The modelling strategy applied to these quite-different regions is based on a common underlying framework, laid out in general terms in *Greene et al.* (2011). This strategy was intended to be adapted as particular regional considerations demand; the present report can serve as an example of such adaptation.

The underlying simulation strategy and resulting models make use of both observational data and simulations from global climate models (GCMs), partitioning “responsibility” for future trends and variability on a range of time and space scales to one or the other (or in some cases, a combination) of these resources. In general, GCMs, such as those participating in the Coupled Model Intercomparison Project (most recently, Phase 5, or CMIP5), which will constitute the climate modelling core of the Fifth Assessment Report of the Intergovernmental Panel on Climate Change (IPCC AR5), tend to be more reliable at large spatial scales and with respect to climatic means, as opposed to variability. These characteristics are implicitly considered when incorporating GCM information into the simulation process.

The issue of spatial scale arises immediately when considering the suitability of individual station records from Kaffrine and Machakos as simulation “targets,” meaning as time series to be used for fitting the statistical models used to generate synthetic future variability. Since individual station records are inherently noisy, longer series are in principal required in order to obtain reasonable estimates of their statistical properties. Because of this limitation, and because of the incomplete nature of temperature data for the two stations considered, we ultimately make use of a gridded dataset for the current analysis, validated for consistency with the initial station records.

The acquisition and assessment of data are described in Section 2. Detrending of the observational records, and some consideration of the problem of separating forced from natural variability, is discussed in Section 3, and modelling of the annual-to-decadal component of variability in Section 4. Generation of future projections is the subject of Section 5, while the temporal downscaling scheme is described in Section 6. Some examples of simulated climate sequences are provided in Section 7, and a consideration of the various contributions to uncertainty in the future projections appears in Section 8. A discussion of some salient issues and a summary of the project appear in Sections 9 and 10, respectively. In general, treatment of the two locations (and for Machakos, the two rainy seasons) is discussed in parallel, often with a single one of these settings serving to exemplify results.

2 Data acquisition and assessment

We utilize data from three sources: weather station records, gridded climate data and, as noted earlier, GCM simulations. We present here the two station records and compare them with gridded data in a small enclosing region, eventually arguing that the analysis is best served by employing the latter, rather than the “raw” underlying station records on which the gridded data is ultimately based. GCM simulations will play a role in defining future precipitation and temperature tendencies, and will be discussed in later sections.

Basic descriptive information for stations Kaffrine and Machakos is provided in Table 1. Note that the Kaffrine data is limited to precipitation. For Machakos both precipitation and temperature are available, but the latter is of more limited extent, covering only the period 1951–1993. Temperature is provided in the form of mean monthly values. As can be seen from the table, Kaffrine lies to the west at about 14° North latitude, while Machakos, in East Africa, is just south of the Equator. As will be discussed, the differing locations experience very different seasonal cycles of precipitation.

Table 1 Basic descriptive information for stations Kaffrine and Machakos

Station	Country	Latitude	Longitude	Variables	Years of record
Kaffrine	Senegal	14.2°N	15.2°W	pr	1951–2010
Machakos	Kenya	1.5°S	37.2°E	pr, T	1894–1985 (pr)

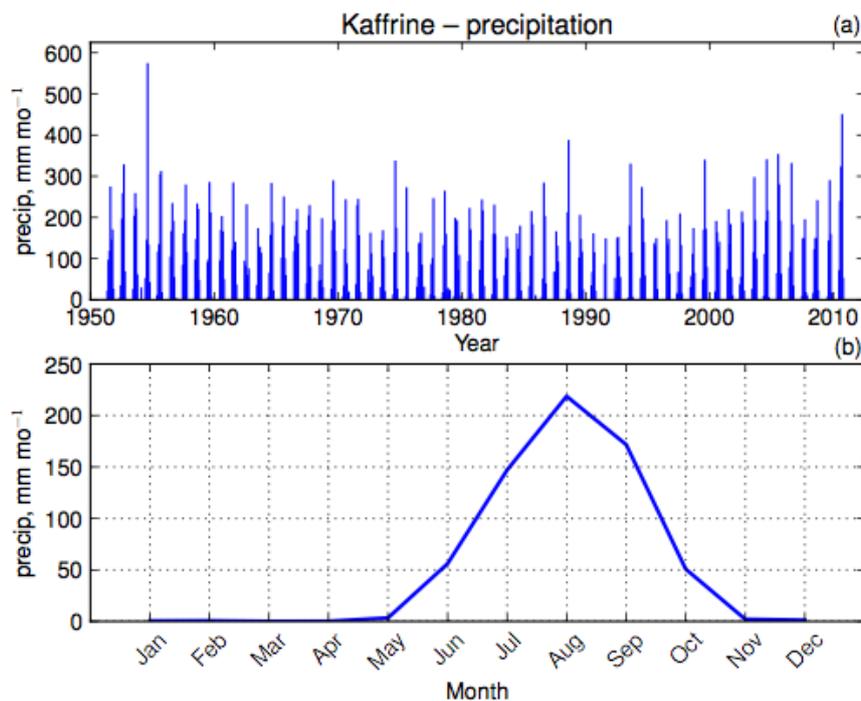
2.1 Station Kaffrine

To obtain the full 60 years of record listed in Table 1 for station Kaffrine, two records were concatenated: The first, extending from 1951–1981 inclusive, was obtained from the NOAA National Climatic Data Center (NCDC) Global Climate Perspectives System (GCPS) (*Baker et al.*, 1995), via the Data Library at the International Research Institute for Climate and Society (IRI), Columbia University. The second, extending from 1981–2010, was obtained from Ousmane Ndiaye at the Senegalese National Weather Service (Agence Nationale de la Météorologie du Sénégal, or ANAMS). These data overlap during 1981 but are very similar during that year; the latter record was utilised for those 12 monthly values. Means and variances of the two records are quite similar, and no adjustment or other calibration was applied in concatenating them. Both records have monthly time resolution. The GCPS does not contain a temperature record for station Kaffrine; neither is such a record available from ANAMS (Ndiaye, pers. comm.)

The Kaffrine record contained a total of 24 missing values (out of a total of 720) grouped in a few clusters, all originating from the GCPS (i.e., 1951–1980) portion of the record. These were filled using “local climatology,” to obtain a complete record. This is an interpolation procedure in which the weighted average of a minimum number of unfilled data points on either side of the missing value (considering only like months) is substituted for it. In the present case the number of points utilised was three. A weighting kernel with values proportional to $1/j$, $j = 1, 2, \dots$, symmetric around the value to be filled, was utilised. It was decided to fill missing values in this manner so as to preserve low-frequency variability, and possible trends, in the record.

A plot of the filled Kaffrine data is shown in Figure 1a, where the single rainy season may be distinguished by the regular, highly periodic sequence of significant monthly rainfall totals separated by dry interludes: Each “spike” in the series corresponds to one year’s rainy season, the intervening months being nearly devoid of precipitation. Figure 1b shows the yearly precipitation climatology for Kaffrine and more clearly illustrates what may be described as the “classically Sahelian” single rainy season, extending approximately from June through October, the core comprising July-August-September, or JAS.

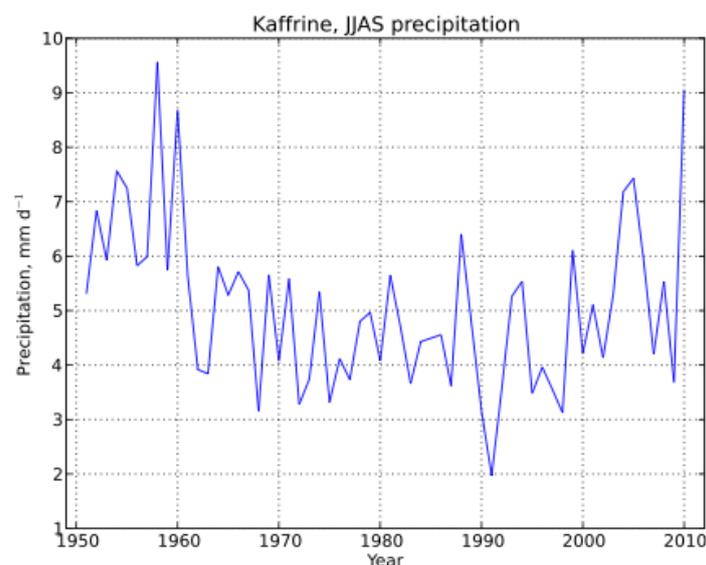
Figure 1 The Kaffrine precipitation record. (a) Time series of monthly values; (b) Seasonal cycle. Some filling has been performed on the monthly values; see text for details.



Because the Kaffrine data lack a temperature component we also examined a gridded dataset, the University of East Anglia Climatic Research Unit (CRU) TS3.2 product. This dataset, gridded 0.5° in both latitude and longitude, includes monthly means of daily maximum and minimum temperatures as well as monthly precipitation totals. A reasonably close correspondence between localized CRU precipitation data and the Kaffrine station record would enable the joint modelling of precipitation and temperature variables, resulting in a potentially more useful product than sequences of precipitation alone.

Figure 2 shows the seasonalised (JJAS) Kaffrine precipitation record, here expressed in mm d^{-1} . A four-month season is used in this case in order to include the time of monsoon onset, important for agricultural outcomes. (Correlation between JJAS and JAS seasonalised precipitation indices is 0.97, with JAS exhibiting a slightly higher mean of 6.1 mm d^{-1} , as opposed to 5.1 mm d^{-1} for JJAS). A notable feature of this plot, less evident in Figure 1, is the apparent inhomogeneity in interannual precipitation variance, with relatively high values both preceding and following a period of relatively subdued variance extending from about 1962 to 1986. The difference in variance is statistically significant at better than 0.01, even when the post-1986 portion of the record is detrended.

Figure 2 Station Kaffrine, seasonalised (JJAS) precipitation based on the filled record shown in Figure 1

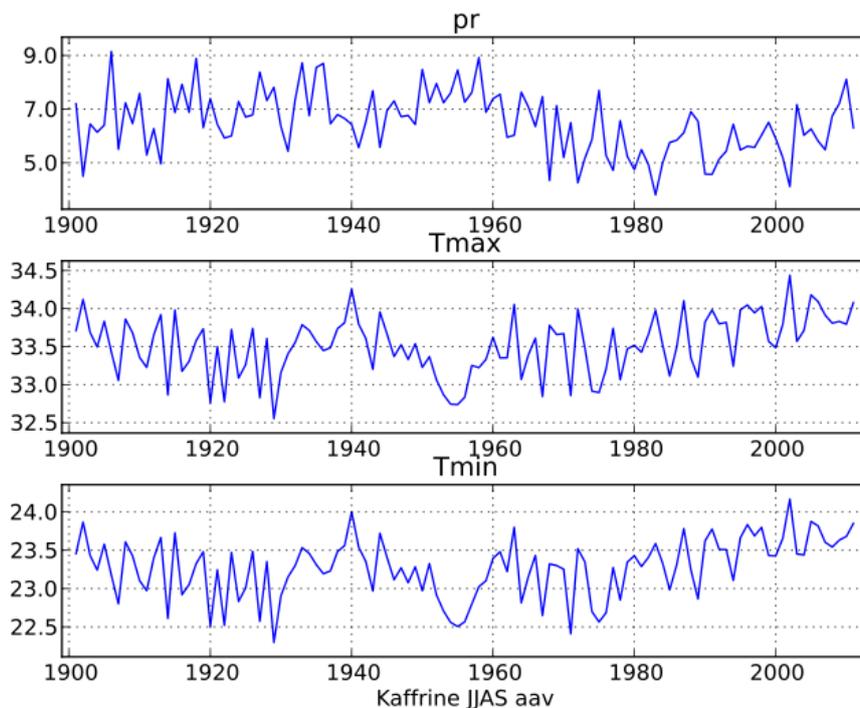


This feature is problematic, since it is not possible to know, in the absence of station metadata, whether the shift represents true climatic variation or some inhomogeneity in recording, or possibly station siting. Even if it could be shown to represent true climatic variation it is not clear how to represent such variance shifts statistically, since there are so few realisations. Is the variance likely to fluctuate cyclically on 25-year periods? Does the

1986 “change point” represent a shift to a new climatic state? Or is this “shift” simply a climatic anomaly, unlikely to persist?

Figure 3 shows seasonalised (JJAS) CRU data area-averaged over a small ($3^{\circ} \times 3^{\circ}$) region extending from 12.75° to 15.25° N and 13.75° to 16.25° W, which encloses the Kaffrine station. The three panels show JJAS mean precipitation as well as JJAS means of the maximum daily temperature (Tmax) and minimum daily temperature (Tmin), respectively. The precipitation record in this case appears more homogeneous, suggesting that the Kaffrine station record may be atypical for the immediate area. Approximately 50 stations contribute to the area-averaged record, although the number is not uniform in time, being lower in the early part of the record and rising to more than 100 stations between about 1960 and 1990. The data include both maximum and minimum daily temperatures (monthly means) and extend over the full century, thus making for a better simulation “target” than the station data alone. For this reason, it was decided to carry out the simulations based on the series shown in Figure 3.

Figure 3 Seasonal (JJAS) mean values of precipitation, Tmax and Tmin, averaged over a 3×3 -gridbox area that includes station Kaffrine.

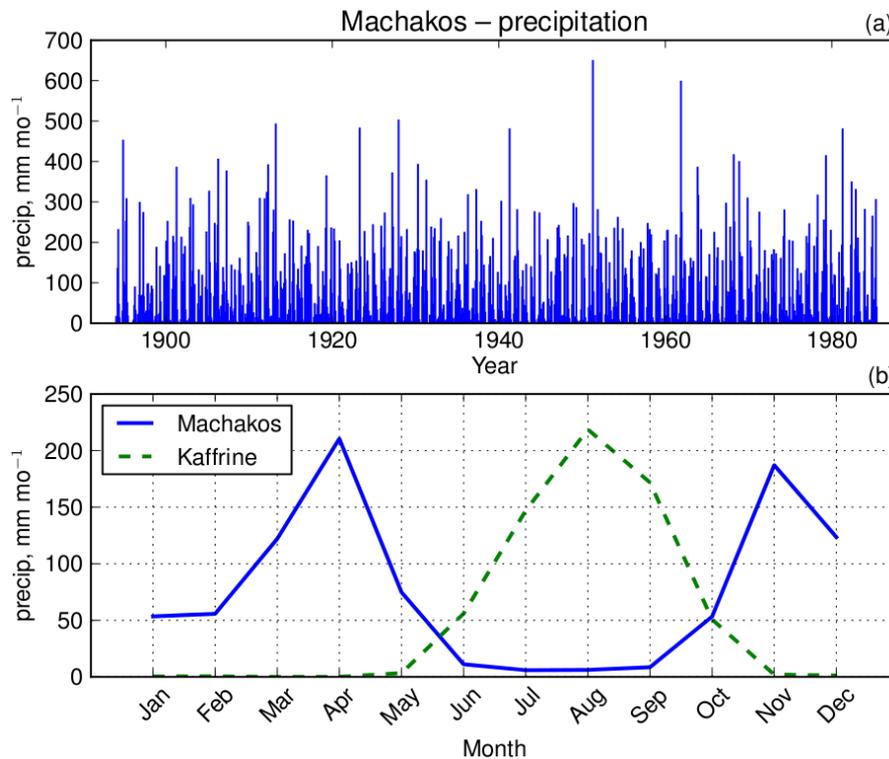


2.2 Station Machakos

The data obtained for station Machakos derive, as with the earlier segment of the Kaffrine precipitation data, from the NOAA NCDC GPCP. Both precipitation and temperature are available in this case. Time resolution, as before, is monthly, with the precipitation data extending from 1894–1985, and the temperature data from 1951–1993. A precipitation plot, corresponding to Figure 1 for station Kaffrine and including both the monthly totals and seasonal cycle, appears as Figure 4. As in the case of the Kaffrine data a small number of missing entries (< 20) was filled, using the same local climatology method and parameters.

Mean annual precipitation for Machakos, at about 78 mm mo^{-1} , is somewhat greater than the corresponding value at Kaffrine, where the mean monthly total is about 56 mm. The seasonal cycle for Machakos (heavy blue line in Figure 4b) exhibits the pattern of long and short rains characteristic of equatorial eastern Africa, reflecting the northward advance, then southward retreat of convection, following the seasonal cycle. The Sahel, as represented by station Kaffrine, lies near the northward extremity of this seasonal migration and experiences only a single rainy season. This linkage between the seasonal migration of convection and the two rainy-season patterns is clearly illustrated in Figure 4b, where the Kaffrine rainy season is seen to be almost perfectly “sandwiched” between the long and short rains at Machakos. That these two well-synchronised stations lie on opposite sides of the African continent simply reflects the global scale of the seasonal cycle, driven by the latitudinal march of solar heating over the course of the year.

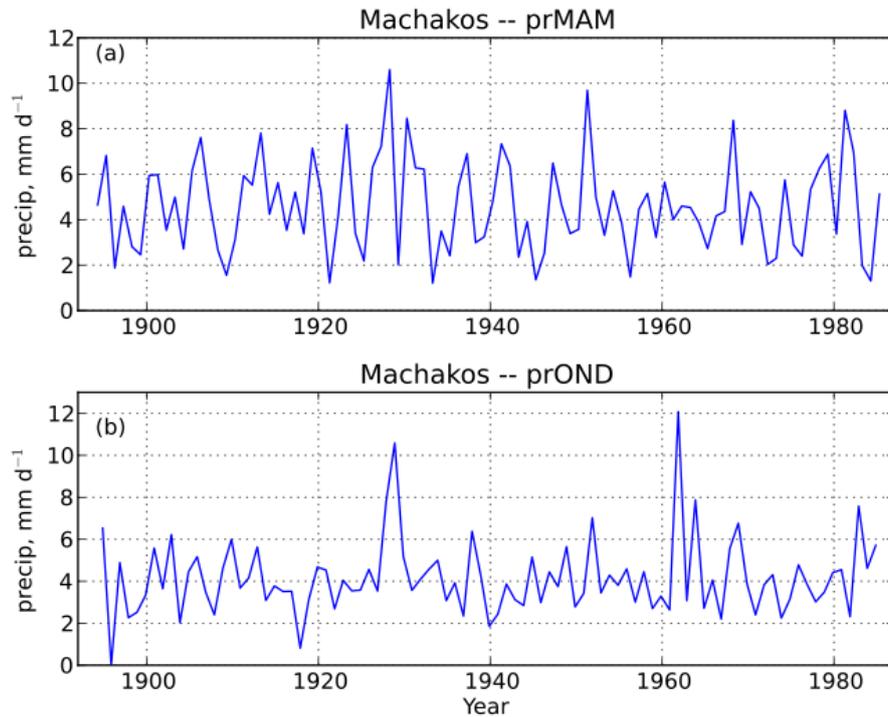
Figure 4 The monthly precipitation record and yearly climatology for station Machakos, corresponding to the two panels of Figure 1. For comparison, the seasonal cycle at Kaffrine is also shown, as a dashed green line, in (b). As in the case of Kaffrine, a small number of data here have been filled.



Beyond this low-order picture, East African rains are known to have a complex dependence on sea-surface-temperature (SST) variations in all three ocean basins; this dependence, furthermore, differs between the two rainy seasons (see, e.g., *Mason and Goddard, 2001; Lyon and DeWitt, 2012*). A detailed analysis of these dependencies constitutes a subject in itself, and lies beyond the scope of this report.

Seasonally-averaged precipitation for the long and short rains, corresponding to seasons March-May (MAM) and October-December (OND), respectively, are shown in Figures 5a and 5b, respectively. Both of these records appear relatively homogeneous, compared with the Kaffrine precipitation record; however the distributional form of the short rains appears somewhat skewed, with quasi-regular spikes punctuating a background variability that is more muted than that of the long rains. Aside from a modest decline from a spike around 1960 in the short rains, neither of these records shows the abrupt decline associated with midcentury Sahel drying. Station Machakos is well outside of the Sahel belt.

Figure 5 Seasonal values for (a) long and (b) short rains for station Machakos.



The NOAA NCDC GCPS also provides a monthly temperature record for station Machakos, but it is essentially just a fragment, covering the period 1951–1993. A plot of annual mean values for this record appears as Figure 6, and shows a period of decline through the late 1960s followed by an increase, with perhaps a levelling off beginning in the 1980s. This decadal-scale fluctuation evidently occupies a large portion of the entire GCPS record. Somewhat analogously to the variance discontinuity observed at station Kaffrine, this presents a conundrum for the simulation exercise because it is not possible to statistically characterise (and thus construct a model based on) a single fluctuation.

Figure 6 Annual mean temperature, station Machakos.

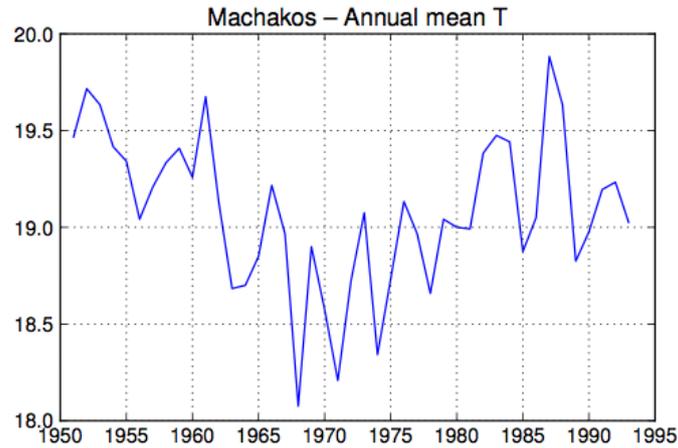
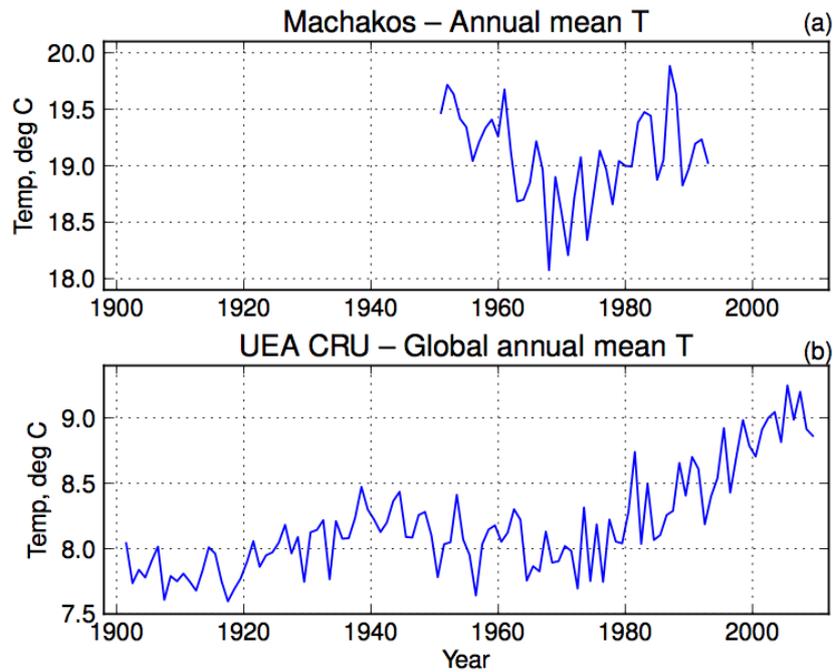


Figure 7a repeats the Machakos temperature record, while Figure 7b displays a record of annual mean, global mean land temperatures for the 20th century from the CRU dataset. The latter record exhibits long and steady decline between about 1940 and 1965, followed by a steady increase. This well-known feature of midcentury climate has been variously attributed to aerosol forcing, particularly over the North Atlantic, natural decadal climate variability or some combination of the two (see, e.g., *Delworth and Knutson, 2000*). The dip in temperature shown for station Machakos can be seen as an expression of this global pattern, both clarifying what the Machakos record may represent and illustrating the potential value of introducing longer and more comprehensive climate records in the analysis, and ultimately synthesis, of climate sequences at individual weather station locations.

Figure 7 Annual mean temperature, station Machakos and global annual mean from the UEA CRU data.



Can we obtain from the CRU data long, homogeneous records of precipitation, Tmax and Tmin for a small area around Machakos, as we did for Kaffrine? Figures 8 and 9 replicate, for the Machakos long and short rain seasons respectively, what is shown in Figure 3 for Kaffrine, in this case for a box extending from 0.26° – 2.75° S, 35.75° – 38.25° E. While the differing seasonal characteristics of station precipitation appear to reasonably well-reproduced in the gridded data the early part of the temperature record is clearly variance-deficient, symptomatic of an overabundance of filled values. However, after about 1950 the series appears reasonably homogeneous and correspond fairly well with the Machakos station record, extending beyond it by some 25 years and thus including potentially important information from the period when anthropogenic forcing is its strongest. (The resultant has a total temporal extent of about 60 years.) Synthesis based on this data would also be more directly comparable to that applied at Kaffrine, and we have therefore adopted this approach for our modelling strategy at Machakos as well.

Figure 8 Machakos seasonal MAM precipitation and maximum and minimum daily temperatures (monthly averages) from the UEA CRU data.

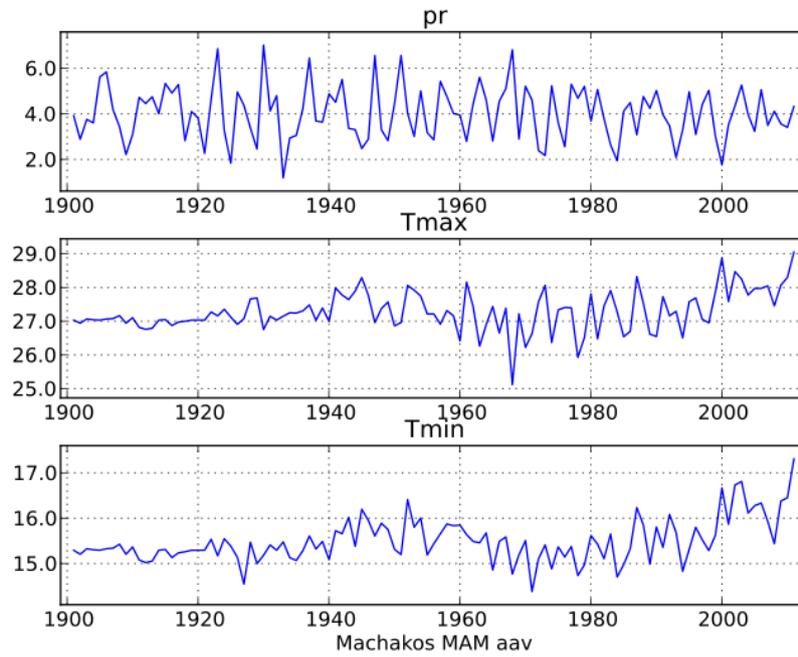
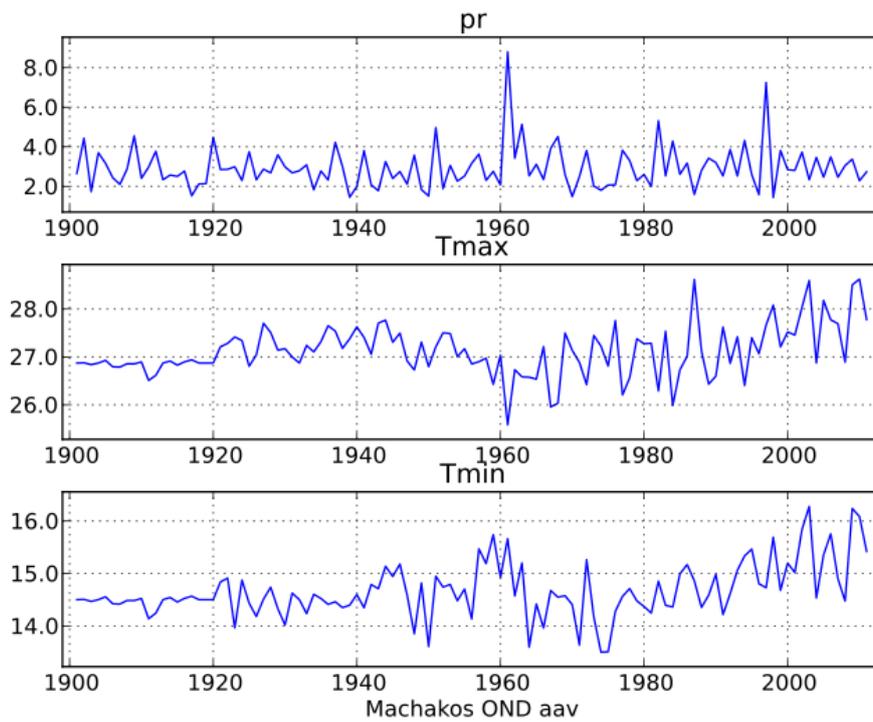


Figure 9 Machakos seasonal OND precipitation and maximum and minimum daily temperatures from the UEA CRU data.



2.3 Discussion

It is clear that the rainfall climatologies at Kaffrine and Machakos differ considerably, as does the character of the long and short rains at the latter station. These diverse behaviours call for the use of distinct models, albeit within our common framework. Treatment of low-frequency variability and trend, and the consideration of anthropogenically forced vs. natural climate variability for station Kaffrine, whose observed precipitation variations are consistent with those of the Sahel generally, will require the adoption of some assumptions, since definitive attribution of Sahelian climate change has not yet been achieved (*Greene et al.*, 2009). We proceed to this step in Section 3.

3 Trend: The forced component

In the treatment of trend, both past and future must be considered: The data must be detrended prior to modelling of variability, requiring an estimate of past trends, while for the future a range of plausible trends must be obtained. In both cases there is an implicit identification of “trend” with what is sometimes referred to as “external forcing,” meaning drivers of climate change originating “outside” the climate system proper. Typically this refers to greenhouse gases and both anthropogenic and volcanic aerosols, as well as variations in solar luminosity. Anthropogenic land use and land cover changes are often included as well. Variations not due to these effects are classified as “natural.” The separation of forced from natural variability is in fact a complex problem (*Solomon et al.*, 2011); however good deal of this complexity is accounted for here by the use of climate models to constrain future trends.

“Trend” and “variability” may in fact not be independent; however their degree of dependence may be tested, at least in the domain of climate models. Typically one is interested in knowing whether precipitation variability will increase as the globe warms; we attempt to answer this question in Section 4.4, by querying the GCMs for regional information.

The other important aspect of trend modelling within the current framework is that the method of forward modelling of temperature differs from that of precipitation. In the former case we use a combination of the observed dependence of Tmax and Tmin and the model dependence of average regional temperature on global mean temperature to project into the future, while for precipitation we take future trends from a distribution over an ensemble of global climate models. Aspects of this difference in treatment persist from the simulations conducted for the Western Cape of South Africa *Greene et al.* (2012).

3.1 Kaffrine

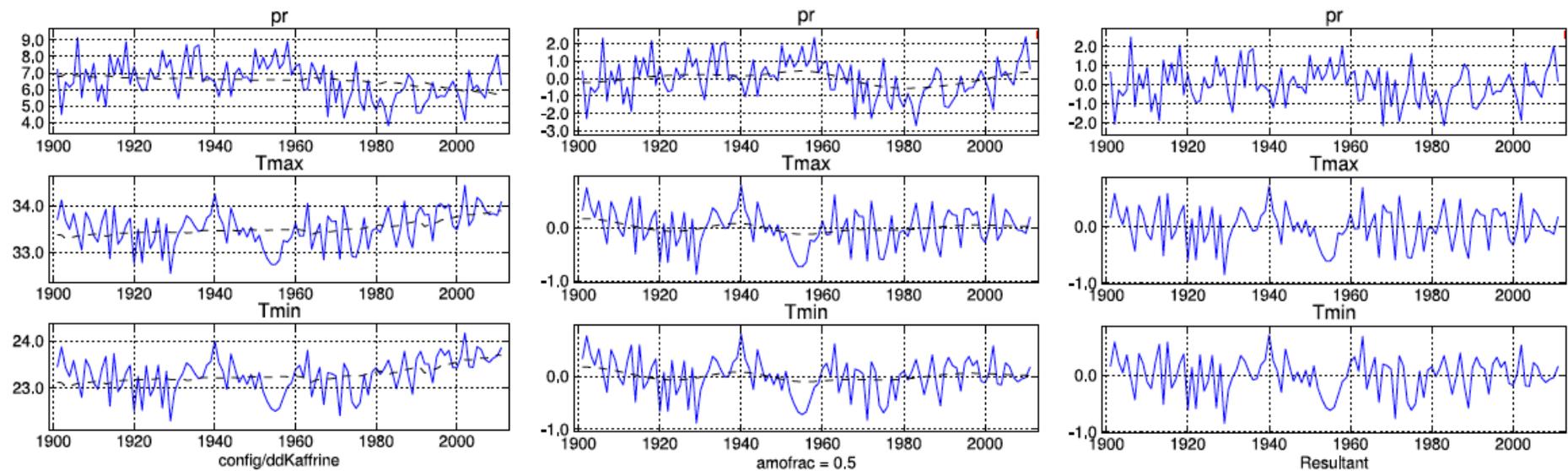
Figure 3 clearly shows the large midcentury decline in precipitation that culminated in the severe Sahelian droughts of the 1970s and 1980s. The question of attribution arises (*Greene et al.*, 2009), and leads us to ask, might such large decadal fluctuations recur in the future? *Biasutti and Giannini* (2006), in showing that there is a consensus for Sahelian drying among 20th-century model simulations but no such consensus in future simulations, suggest that this may not be the case. The difference in response is attributed to aerosol forcing of North Atlantic SSTs, essentially an anthropogenic effect, since the 20th-century simulations are driven in part by common aerosol forcing.

On the other hand, it is also believed that the meridional overturning circulation in the North Atlantic (AMOC) contributes to decadal variability in SST (e.g., *Johnson and Marshall, 2002*). This is primarily a natural phenomenon, and *would* be expected to continue into the future. Thus, both anthropogenic and natural processes may contribute to the precipitation signal we see in Figure 3. The separation of these two influences in a definitive way has not yet been achieved, yet it is desirable that we attempt to project forward only that component of variability that is expected to play a role in future variability.

We resolve this problem by detrending the Kaffrine record in two steps: In the first, a global mean, multimodel mean temperature record, derived from an ensemble of 34 CMIP5 GCMs, is regressed on each member of the Kaffrine series (precipitation, Tmax, Tmin). In principal this extracts that part of the regional forced response owing to global warming (more precisely, the part that is linearly dependent on it). In the second step a lowpass filter is applied to each of the records, extracting variability on time scales of 30 years and longer. Rather than remove *all* of this variability from the records, which would be appropriate if we believed that all low-frequency variability were anthropogenically forced, we remove only a portion, somewhat arbitrarily set here at one half. This procedure leaves an equal portion of low-frequency variation, assigned implicitly to natural causes, in the detrended record that is passed to the time series modelling stage. Along with higher-frequency components, this low-frequency element will be emulated in simulations of future variability.

Figure 10 illustrates this process. Panel (a) shows the raw data (same as Figure 3, above), with trend lines added. These are least squares best-fit lines, where each of the variables is regressed on the same global mean, multimodel mean temperature record. The regression is linear, so the lines are straight when plotted against global mean temperature, but since the temperature does not evolve linearly with time the lines are not straight when plotted against that variable.

Figure 10 Detrending of the trivariate Kaffrine record in two stages: The raw data is shown in (a), along with a trendline that is nonlinear in time but linear in global mean temperature. In (b) is shown the result of subtracting the trendline from the raw data, and lowpass-filtered version scaled so as to include half of the remaining low-frequency component. The final resultant appears in panel (c).



(a) Raw + trend

(b) Detrended + smooth

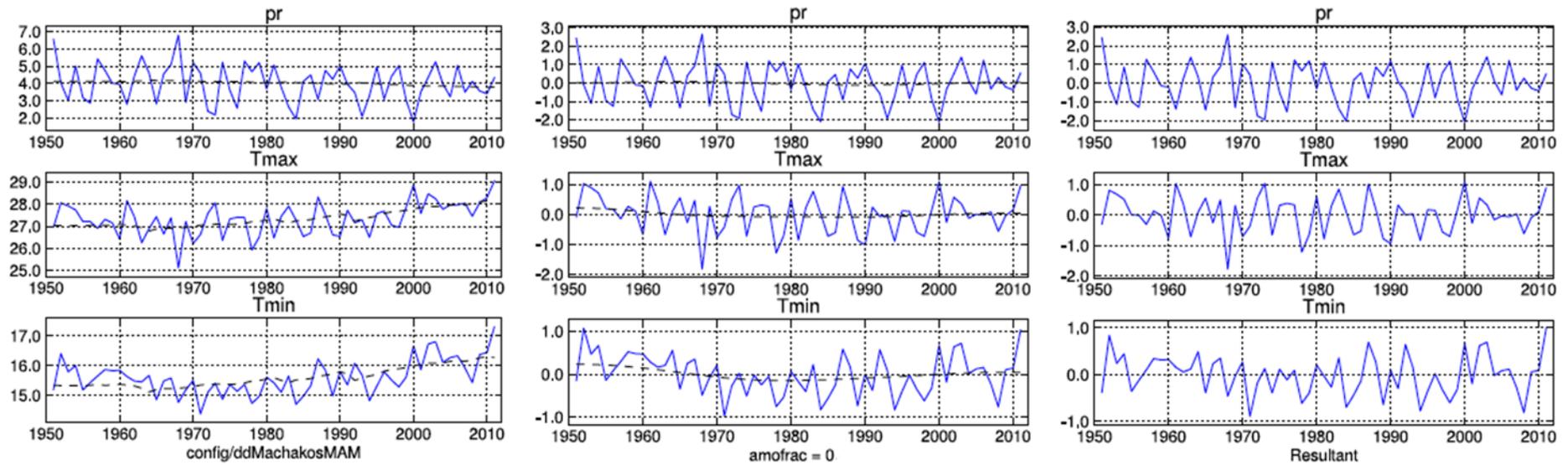
(c) Resultant

Panel (b) shows results of the first detrending step, i.e., the raw data minus the trendlines shown in panel (a). Plotted on top of these traces, but at half amplitude, are 30-year lowpassed versions of the same series. This is the output from a digital filter constructed so as to attenuate signal components with periods shorter than 30 years. The half-amplitude lowpassed signal can be seen to capture some, but not all, of the low-frequency variation in the records, particularly precipitation. Panel (c) shows the final resultant, in which the half-amplitude lowpass has been subtracted from the detrended series. The resultant retains some, but not all, of the original low-frequency component of the original data, emulating variations that contain a degree of natural variability, but this component is attenuated, as compared with the original signal. It is the resultant series of panel (c) that constitute the modelling target for the annual-to-decadal variability component of the simulation.

3.2 Machakos

Figure 11 shows the same detrending procedure applied to the 1951-2011 portion of the Machakos long rains (MAM seasonal average) record. The precipitation record in this case does not exhibit the dramatic mid-century decline observed at Kaffrine. In common with that station, both Tmax and Tmin show upward tendencies. Here we do *not* assume that there is a substantial forced low-frequency component; rather, all of the low-frequency variability (of which there is relatively little, as compared with Kaffrine) is attributed to natural processes, and allowed to remain in the final detrended resultant.

Figure 11 Detrending of the Machakos MAM record.



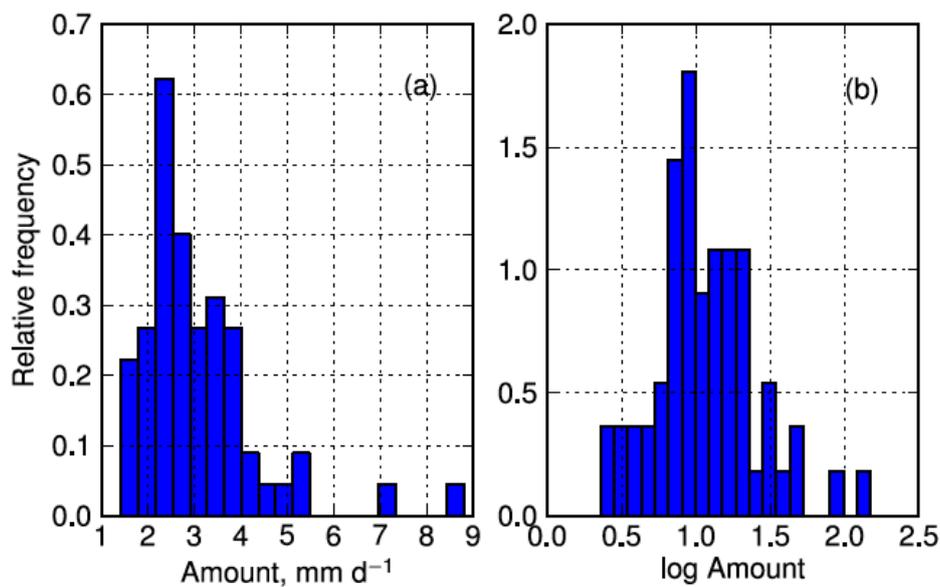
(a) Raw + trend

(b) Detrended + smooth

(c) Resultant

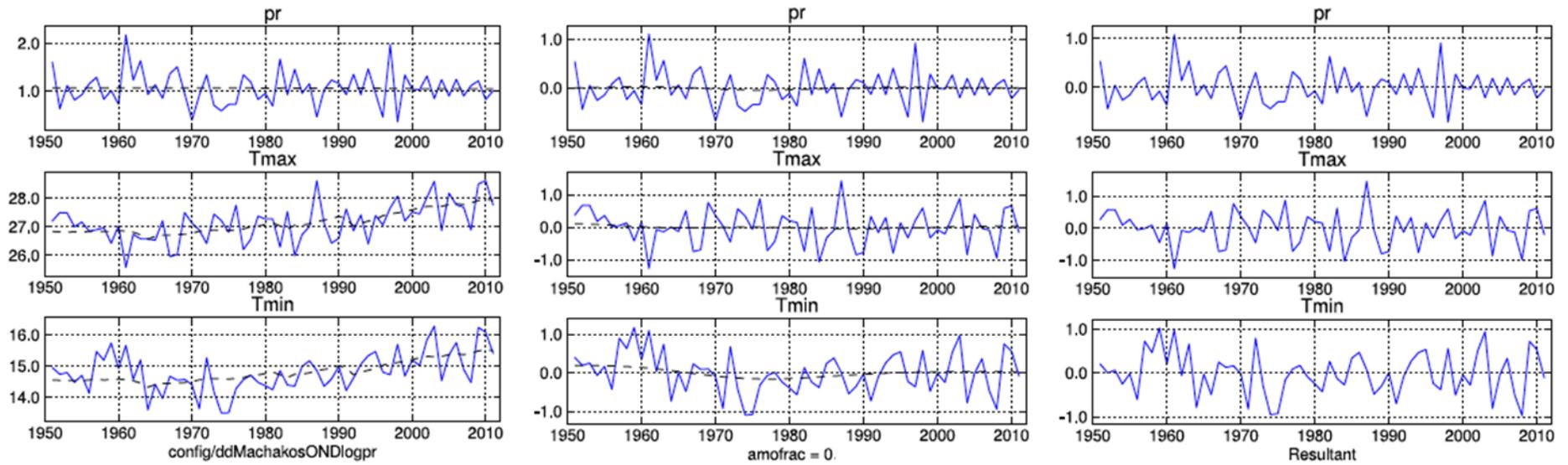
It was noted earlier that the probability distribution function (PDF) of the short rains is somewhat skewed, and to take account of this a slight procedural modification is implemented, in which the precipitation values are log-transformed prior to detrending. The annual-to-decadal model is fit to the resultant, which may more reasonably be assumed to be Gaussian. The transformation is illustrated in Figure 12 where the PDF of the raw OND precipitation is shown in its raw form in Figure 12a and transformed in Figure 12b. The inverse transform is applied prior to the downscaling procedure described in Section 6.

Figure 12 Probability distribution function for the Machakos short rains. Raw data in (a), log-transformed in (b).



Detrending of the Machakos short rain (i.e., OND) data, with log-transformed precipitation, is illustrated in Figure 13, where the residual low-frequency variability is most evident in the T_{min} component. As with the two records illustrated in Figures 10 and 11, it is the resultant trivariate time series that is taken as representative of annual-to-decadal variability. We turn in Section 4 to the modelling of these series.

Figure 13 Detrending of the Machakos OND record.



(a) Raw + trend

(b) Detrended + smooth

(c) Resultant

4 Modelling of the annual-to-decadal component

The aim in this step is to generate long trivariate data sequences having the key statistical properties of the (detrended) resultants from the previous step. As in *Greene et al. (2012)* the data for each location and season is modelled as a vector autoregressive (VAR) process, using the dynamical systems estimation (dse) package (*Gilbert, 1993*) in the R programming language (*R Core Team, 2013*). We note that there are a variety of packages and modes for VAR modelling that might usefully be deployed for the purpose of multivariate climate simulation; space and time limitations preclude a more comprehensive exploration of those alternatives in this report.

4.1 Model structure

The general, VAR(m) (order- m) model can be written

$$\mathbf{Y}_t = \sum_{k=1,m} \mathbf{A}_k \mathbf{Y}_{t-k} + \boldsymbol{\epsilon}_t \quad (1)$$

where \mathbf{Y}_t is the three-component (pr, Tmax, Tmin) climate vector at time (i.e., year) t , \mathbf{Y}_{t-k} the k -year lagged value of this vector and $\boldsymbol{\epsilon}_t$ a random, vector “noise” process that is uncorrelated in time but may be correlated across variables. The m 3×3 matrices \mathbf{A}_k hold the coefficients relating \mathbf{Y}_t to its m predecessors in time, the nine elements of each matrix characterising lag- k relationships between the nine possible pairwise combinations of pr, Tmax and Tmin, thus including serial autocorrelation as well as lag- k cross correlations among the variables.

Because of the explicit time and cross-variable dependence included in the VAR structure, as well as prior results indicating the suitability of such models, we limited our survey to VAR models of various orders, entertaining a maximum possible m of five. Results are discussed in Section 4.2.

4.2 Information tests

Table 2 shows the results of goodness-of-fit tests for VAR models of orders one through five, applied to the Kaffrine JJAS data. The column labelled “-log(L)” is the negative log-likelihood of the data, given the model. As m (leftmost column) increases, so does the number of parameters in the model, since each increment of m requires an additional matrix \mathbf{A}_k .

Adding parameters in general results in increasing the log-likelihood, and this is seen to be the case here. While such an outcome is nominally desirable, however, the log-likelihood does not discriminate between fitting the signal and fitting the noise, i.e., that portion of the data which is essentially random, and does not exhibit systematic lagged relationships. Thus, adding too many lags to the model results in *overfitting*.

The next two columns, “AIC” and “BIC,” give values for two alternate metrics (*Akaike, 1973; Schwarz, 1978*), each of which takes into account not just the model fit, in terms of log-likelihood, but also the number of parameters in the model, penalising the (negative) log-likelihood by adding a term that increases with the number of parameters. The more complex the model, the more likely it is to be fitting not simply the signal of interest but also the noise. The penalty term thus acts as a check on overfitting. What we then seek is a minimum, not of the log-likelihood, but of the AIC and/or BIC.

Table 2 Information tests for a range of models, applied to the detrended JJAS Kaffrine data.

Lags (m)	$-\log(L)$	AIC	BIC
1	42.70	103.40	137.70
2	37.25	110.50	179.05
3	32.51	119.03	221.85
4	29.23	130.46	267.55
5	25.86	141.73	313.10

Table 2 shows such minima, for both AIC and BIC, for a model with just a single lag. As additional lags are included, both AIC and BIC increase, indicating that the best fit is provided by an order-one model. A number of other metrics (not shown, for brevity) are in agreement. Similar statistics are produced when models are fit to both of the Machakos datasets, suggesting that for all three situations a VAR(1) model gives the best fit to data, and that including additional lags does not yield improvements in model fit.

While such a result might seem coincidental it is in fact not entirely surprising. Climate variability is largely a product of interaction between the atmosphere and ocean, and it was hypothesised some time ago (see, e.g., *Hasselmann, 1976*) that the dynamics (and resulting spectra) of the atmosphere-ocean system are similar in principal to that of Brownian motion, in which rapid, random fluctuations of the molecules in a fluid induce slower, random-walk motion in particles suspended in that fluid. The rapid molecular motion is taken as an analog of rapid, quasi-random atmospheric motions, while the comparatively sluggish movement of the suspended particles (in the initial observational context these were pollen grains) is

identified with movements of the ocean, which is characterised by substantially greater thermal and mechanical inertia. The spectrum of the resulting oceanic movements is “red,” meaning that the motions, like those of the pollen grains, are characterised by what is essentially “first-order” memory. Since terrestrial climate fluctuations are largely governed by (sometimes remote) forcing arising from sea-surface temperature (SST) variations, surface climate fluctuations over land also tend to be first-order, or in modelling terms, first-order autoregressive, corresponding to the VAR(1) modelling structure. All of this is essentially a way of saying that “red” (i.e., first-order autoregressive) noise often provides quite a good model for climate variability.

4.3 Reproduction of observed statistics

Equation 1 expresses a relationship between present and lagged values of the three variables; the reader may have noted that it does not explicitly include contemporaneous relationships among them. However, these relationships are well-reproduced in simulated data generated with the model, in this case a trivariate series of length 10000. Table 3 compares observed and simulated correlation matrices, where an excellent correspondence can be observed. These correlations (or more accurately, covariances) are encoded in Σ_e , the noise covariance matrix, which is computed as part of the model fitting routine.

Table 3 Contemporaneous correlations for the Kaffrine JJAS observations and a long simulation sequence generated by the VAR model.

	Observations			Simulation		
	pr	Tmax	Tmin	pr	Tmax	Tmin
Pr	1.000			1.000		
Tmax	-0.457	1.000		-0.460	1.000	
Tmin	-0.413	0.969	1.000	-0.419	0.968	1.000

Table 4 shows the lag-1 autocorrelation coefficients for the three JJAS-seasonalised variables for Kaffrine, and shows that the model replicates the observed statistics fairly well. It should be noted that none of these coefficients are statistically significant, meaning that the three detrended series are individually similar to white noise: each is essentially unpredictable from a knowledge of only its own past.

Table 4 Lag-1 autocorrelation coefficients for pr, Tmax, Tmin, for the Kaffrine JJAS observations and corresponding VAR model simulation.

Source	pr	Tmax	Tmin
Observations	0.039	-0.021	0.022
Simulation	0.061	-0.016	0.030

Examination of the two Machakos seasonal observational records and the corresponding simulations reveals similar agreement between simulated and observed statistics. Most, but not all, of the lag-1 autocorrelation coefficients are not statistically significant at the 0.05 level, suggesting that once anthropogenic trends are accounted for, a good deal of the residual variance in the individual climate time series is essentially uncorrelated from year to year.

4.4 Considering cross-scale interaction

There is some expectation that precipitation variability may increase in a warming climate, owing to the rapid and nonlinear increase of saturation water vapour pressure with temperature, according to the well-known Clausius-Clapeyron relation (*Held and Soden, 2006*). We tested this hypothesis using the model ensemble listed in Table 5, comparing, in each model, regionally averaged precipitation variance for 2010–2039 with that of 2040–2079, in both the RCP4.5 and RCP8.5 simulations. All three regional/seasonal combinations were tested.

For both scenarios there is some tendency for the variance to increase with global temperatures, but these changes were not found to be statistically significant (i.e., zero, one or two models out of 34 tested as significant at $\alpha = 0.05$, which would be expected even with no change in variance). Results were similar across the two scenarios. A comparison of 2010–2039 with 2070–2099 yielded a somewhat greater number of models (five) showing significant increases in variance for the Machakos OND season; otherwise the changes were not significantly different from zero. Since our simulations run only through 2050 (and since even this may exceed application needs), we have not attempted to incorporate a precipitation variance response to temperature in our statistical model.

The ability of the VAR(1) structure to reproduce essential observational statistics gives us some confidence in the annual-to-decadal variability simulated by these models. We turn in Section 5 to forward projection, the simulation of future climates at the study locations, in which future trends and simulated variability are combined.

5 Projection

Projection of the seasonalised data involves several steps, as follows:

- Future trends are inferred using an ensemble of CMIP5 models. Regional sensitivity to global temperature change is computed independently in the domain of each model, without reference to the observational data.
- The long simulation sequence described in Section 4 is subsampled in accordance with user needs. For example, if there is particular interest in 10- year mean fluctuations in rainfall (i.e., wet and dry decades), representative samples are drawn in which decadal-mean precipitation falls at or near prespecified quantiles.
- Trends and subsampled variations are combined, yielding annual trivariate series (i.e., one value per year for each variable, representing seasonal averages.)

A final simulation step involves temporal downscaling, where is here performed at a monthly time step. This is described in Section 6.

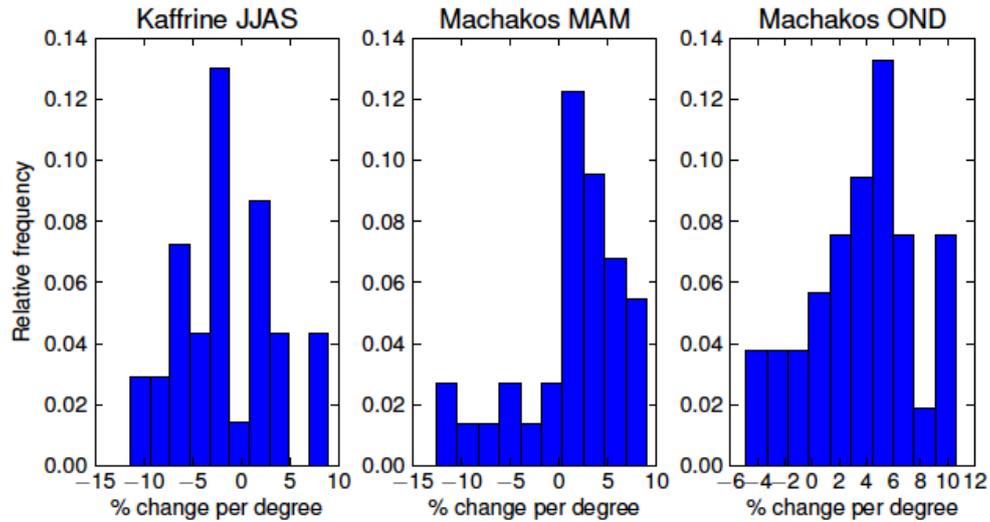
5.1 Future trends

To gauge the sensitivity of climate at our study locations to changes in global mean temperature we analysed simulations from the ensemble of CMIP5 models listed in Table 5. The results for precipitation for each of our experimental settings are shown in Figure 14. For each of the panels, logarithmically-transformed local precipitation averaged over somewhat larger regions than those defined earlier (10° – 20° N, 20° – 10° W for Kaffrine, 5° S– 5° N, 32° – 42° E for Machakos) was regressed on global mean temperature, for the years 2012–2072 in the RCP8.5 IPCC scenario. Larger regions were utilised in this case because climate models are less reliable at small scales and may “misplace” features of interest, whereas in the CRU dataset we are fairly confident that geographical patterns actually correspond to the locations where they are expressed. Results are not sensitive to either the scenario chosen or the precise period used for the computation, but scatter among the models is somewhat reduced in RCP8.5, compared with the lower-emissions scenario tested (RCP4.5), and also when using longer time periods for the computation. Sensitivity was calculated independently within the domain of each of the models; it is the regression coefficients, multiplied by a factor of 100 to give percentage values, whose distributions are shown. The coefficients express, for each model, the percentage change in local precipitation per degree of global warming, and are thus a measure of regional precipitation sensitivity.

Table 5 The ensemble of CMIP5 models utilised in the estimation of precipitation sensitivity distributions. See Section 5.1 for discussion.

Modeling Center	Model
Commonwealth Scientific and Industrial Research Organization (CSIRO) and Bureau of Meteorology (BOM), Australia	ACCESS1-0 ACCESS1-3
Beijing Climate Center, China	bcc-csm1-1
Beijing Climate Center, China	bcc-csm1-1-m
College of Global Change and Earth System Science, China	BNU-ESM
National Center for Atmospheric Research, USA	CCSM4
Community Earth System Model Contributors	CESM1-BGC
Community Earth System Model Contributors	CESM1-CAM5
Centro Euro-Mediterraneo per I Cambiamenti Climatici, Italy	CMCC-CESM
Centro Euro-Mediterraneo per I Cambiamenti Climatici	CMCC-CM
Centro Euro-Mediterraneo per I Cambiamenti Climatici	CMCC-CMS
Centre National de Recherches Météorologiques, France	CNRM-CM5
Commonwealth Scientific and Industrial Research Organisation	CSIRO-Mk3-6-0
Canadian Centre for Climate Modelling and Analysis	CanESM2
Institute of Atmospheric Physics, Chinese Academy of Sciences	FGOALS-g2
The First Institute of Oceanography, SOA, China	FIO-ESM
Geophysical Fluid Dynamics Laboratory, USA	GFDL-CM3
Geophysical Fluid Dynamics Laboratory, USA	GFDL-ESM2G
Geophysical Fluid Dynamics Laboratory, USA	GFDL-ESM2M
NASA Goddard Institute for Space Studies, USA	GISS-E2-H
NASA Goddard Institute for Space Studies	GISS-E2-R
Met Office Hadley Centre, UK	HadGEM2-CC
Met Office Hadley Centre	HadGEM2-ES
Institute for Numerical Mathematics, Russia	INM-CM4
Institut Pierre-Simon Laplace, France	IPSL-CM5A-LR
Institut Pierre-Simon Laplace	IPSL-CM5A-MR
Institut Pierre-Simon Laplace	IPSL-CM5B-LR
Atmosphere and Ocean Research Institute,	MIROC-ESM
National Institute for Environmental Studies and	MIROC-ESM-CHEM
National Institute for Environmental Studies, Japan	MIROC5
Max Planck Institute for Meteorology, Germany	MPI-ESM-LR
Max Planck Institute for Meteorology	MRI-CGCM3
Norwegian Climate Centre	NorESM1-M
Norwegian Climate Centre	NorESM1-ME

Figure 14 Probability distribution functions for local precipitation sensitivity for each of the regional/seasonal targets. The values express the percent change in local precipitation per degree global warming.



The sample means and standard deviations for our three settings are Kaffrine: -1.7 ± 5.0 ; Machakos MAM: 1.4 ± 5.3 ; Machakos OND: 3.7 ± 3.92 . Statistically speaking the multimodel mean sensitivities do not differ significantly from zero, although the models do not divide evenly according to sign of the precipitation change: For Kaffrine only 13 of 34 models get wetter with warming global temperatures, while for Machakos the numbers are 25 of 34 and 28 of 34, for MAM and OND, respectively. Viewed in the context of ensemble spread, these “consensus” judgments are not very strong, although Machakos OND is arguably a bit more definitive than the other two settings. The plots reflect this uncertainty, and also indicate that even with 34 members the sample size is effectively small, since the distributions are not very smooth, or well-formed. Machakos MAM in particular exhibits a somewhat pathological distributional form, as though the models were divided into two groups, one, containing the drying models, following a quasi-uniform distribution and the remainder a distribution that is skewed to the right. Such irregularities may also reflect to some degree the still relatively small sizes of the zones surrounding the two stations, over which the climate variables are averaged.

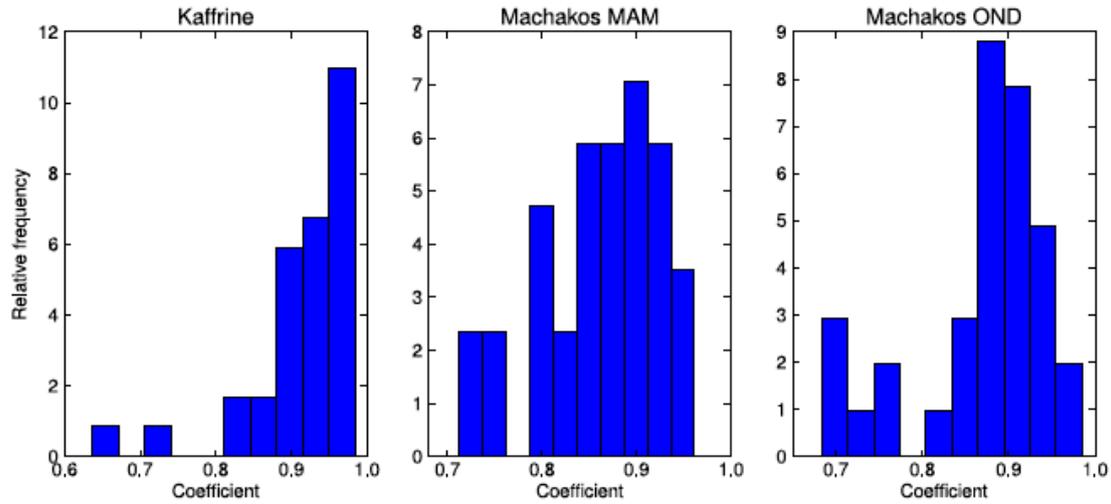
Rather than fit some assumed distributional form, such as a Gaussian, to these distributions as was done in *Greene et al. (2012)*, we allow the data to “speak for themselves,” and compute quantiles of future precipitation sensitivity directly from the raw coefficients. (The Kaffrine JJAS record might have been an exception, being relatively Gaussian in appearance, but for the sake of uniformity we have decided to treat all three locations similarly here.) The quantile function utilised interpolates, if necessary, between the raw point values, providing at

least a modicum of smoothing for simulation purposes. Future trends are then computed based on the evolution of the global mean, multimodel mean temperature record referred to in Section 3.1, which is extended into the 21st century using the RCP8.5 emissions scenario from CMIP5. This represents a high-emissions outcome, with global mean temperature increasing by about 4°C by the end of the century. (The simulations presented here extend only to 2050.)

For temperature we utilise a similar procedure, with the small difference that mean temperature trends are inferred from the same ensemble of GCMs and for the same geographical boxes, but the *ratio* between the sensitivities of Tmax and Tmin are inferred from regressing the observational records of these quantities onto a global mean, multimodel mean temperature record. This slight complexity was necessitated by the absence of Tmax and Tmin (but the availability of their average) in the CMIP5 archive being utilised, but has a possible advantage in that it blends observational and model data, the former being utilised in the estimation of the higher-order statistics represented by departures from changes in the ordinary mean.

Figure 15 shows distributions of the coefficients for local mean seasonal temperatures regressed on the global mean for 2012-2072, similar to what is shown for precipitation in Figure 14. Mean values for the three distributions are 0.91, 0.86 and 0.87, for Kaffrine JJAS, Machakos MAM and Machakos OND, respectively. These mean values, disaggregated into Tmax and Tmin, are used to impose future temperature trends for each of the study regions. As noted, the disaggregation procedure conserves the ratio between Tmax and Tmin coefficients, as derived from regressions of the observational Tmax and Tmin on a global mean, multimodel mean temperature record. The model PDFs of regional temperature sensitivity (i.e., the spread in mean temperature sensitivity estimates) are not exploited in the current modelling scheme.

Figure 15 Probability distribution functions for local mean temperature sensitivity for each of the regional/seasonal targets. The values express the linear dependence of local temperature change on global temperature.



The resulting coefficients, relating seasonal temperature changes to global mean, annual mean changes are given in Table 6. It is of interest that for two of the study areas Tmin increases slightly more rapidly than Tmax. The implied reduction in the diurnal temperature range, or DTR, has been studied as a corollary of global warming (e.g., *Stone and Weaver, 2003*) and has been hypothesised to be an effect of changing cloudiness; we do not attempt to diagnose its cause here.

Table 6 Dependence of Tmax and Tmin on global mean temperature, for each of the three settings considered. See text for details.

Setting	Tmax	Tmin
Kaffrine JJAS	0.86	0.96
Machakos MAM	0.87	0.85
Machakos DJF	0.82	0.92

All of the coefficients in Table 6 are less than unity (although they approach it, particularly in the case of Tmin). This means that, in the model world at least, both regions are warming, in the RCP8.5 scenario, slightly more slowly than is the globe as a whole. In both the 20th century and in future simulations there is typically an acceleration of warming in the high latitudes, the so-called “polar amplification,” owing to the so-called ice albedo feedback (e.g., *Dickinson. et al., 1987*). This means that other regions, on average, must be warming *less rapidly* than the mean. On the other hand it has also been noted that land areas are likely to

warm more rapidly than oceanic regions, because of both the ocean's greater heat capacity and its ability to sequester heat below the surface. These two processes would potentially have opposing effects on the coefficients in Table 6: polar amplification would tend to drive them down, while attenuated oceanic warming would drive them up. Other factors, such as land surface changes, may also contribute. At any rate it should not be surprising that the coefficients differ from unity: All regions of the globe are not guaranteed to warm at the same rate.

5.2 Simulation subsampling

Selection, fitting and validation of the VAR model were discussed in Section 4, where statistical comparisons were made using a simulated (trivariate) data sequence. Sequences were generated for each of the three region/season combinations, each of length 10000 yr. These have a single value per year for each variable, representing seasonal mean quantities. (As mentioned, in the case of Machakos OND the model was fit, and the simulations generated, using the logarithmically transformed precipitation.)

To generate simulations spanning a predefined range of years at one of the locations, a segment of suitable length is subsampled from the corresponding long sequence. The length of this sequence permits the selection to be highly customised, or "tailored." Say the user is interested in fluctuations in 10-year mean precipitation, perhaps for longer-range planning purposes or for an impact study. A segment can be selected in which the mean precipitation fluctuation for a specified decade falls at a particular percentile, relative to all 10-year means in the long sequence. Since the annual-to-decadal statistics are known, the risk for exceeding such a threshold can be quantified, at least to the extent of uncertainty in the underlying data.

The precipitation and temperature variables are generally anticorrelated (see, e.g., Table 3), meaning that a dry decade will tend to be a warm one. However for any particular 10-year sequence there is no guarantee that the "typical" correlation will be replicated. For this reason, and because we intend that the selected decade be characterised principally by a precipitation deviation, we conduct secondary screenings of Tmax and Tmin, selecting for decades where these variates lie close to their conditional mean values, given the specified precipitation quantile. A third level of screening is also applied, in which the precipitation fluctuation during the decade *prior* to the target decade is constrained to be small. This condition is applied to minimise the possibility that prior conditions will unduly influence results assumed to be based on climate fluctuations during the target decade.

The constraints on decadal mean precipitation, Tmax and Tmin, as well as the previous decade's precipitation, are implemented in terms of standard deviations, with the strictest control applied to the target decade's precipitation. Generally this is constrained to lie within 0.05 to 0.1 standard deviations (σ) of the specified quantile, with the precise limits depending on quantile. Thus, for quantiles near the centre of the distribution the precipitation constraint might be closer to 0.05σ , while near the tails it may be relaxed toward 0.1σ , since by constructions there are fewer available sequences from which to choose, the further one departs from the mean. Constraints on the other variables are somewhat more relaxed, ranging from 0.25σ to 1σ , depending on quantile.

The screening parameters are adjusted so that a few dozen candidate sequences pass all the designated criteria. These are further screened visually in order to make a final selection. It is worthwhile noting that, owing to the stochastic nature of all of the variables, there are many possible within-decade patterns that match the selection criteria enumerated above. (Examples will be shown in Section 7). A dry decade may contain one or more normal, or even wet years, and vice versa, and such years may be clustered or may appear sporadically within the decade. Depending on the application (reservoir management vs. rain-fed agriculture, for example), these within-decade patterns of climatic fluctuations may or may not make a difference to the experimental outcome. In the visual selection of example fluctuations we have tried to choose those that seem more "typical," but it is certainly possible that a particular user might be interested, say, in decades with near-normal precipitation (given long term trends) but with large interannual fluctuations, or in other types of behaviour. The screening criteria utilised can be modified or adjusted as necessary, in order to identify behaviours of particular interest. Additionally, the criterion interval need not be ten years. This length was chosen because it corresponds to the nominal decade, but is otherwise arbitrary.

The screened segments are combined with projected future trends, resulting in annually-resolved sequences of seasonal mean values. In the temporal downscaling step, discussed in Section 6, these data are disaggregated to monthly time resolution.

6 Downscaling

The data produced by the procedures described so far has annual time resolution, which may not be sufficiently fine for some applications. Here, we describe the temporal disaggregation of this data to monthly resolution. In general, the generation of data at a particular time resolution, under the current scheme, is dependent on the availability of observational data at that resolution: Daily data is required in order to generate simulations at the daily time step, and so on. (This would be equally true were some type of weather generator used to generate daily values; such generators require extensive parameter sets describing local variability if their output is to replicate well the observed statistics.) The CRU data utilized here has monthly resolution; this is the resolution at which the simulations are produced.

6.1 Nonparametric resampling scheme

The method of disaggregation utilised here is the k -nearest-neighbour approach, sometimes designated k -NN. In essence it consists of searching among the existing observational data for years whose seasonal characteristics are “close,” in a particular statistical sense, to those of the year being simulated. The data are compared at the seasonal level (at this point in the simulation scheme, this is the only form of data that is available), and the k closest candidate years from among the observed are selected. A random procedure is used to choose from among these candidates, and the monthly values from chosen year are utilised in the simulation. A rescaling step adjusts the individual monthly values so that their averages — one for each variable — precisely match those of the year being simulated. The number of neighbours utilised here, k , is set to five, and the final choice made according to the toss of a weighted die, whose five faces appear with probabilities $1/j$, $j = 1, 2, \dots, 5$. The choice of k represents a balance between finding neighbours that are reasonably “close” to what is simulated and achieving a realistic degree of stochasticity, or quasi-randomness, in the resulting sequences.

The metric utilised to compute the distance between the simulation vector and the observational climate vector is the Mahalanobis distance, which can be written

$$d = \sqrt{(\mathbf{x}_1 - \mathbf{x}_2)^T \mathbf{W} \boldsymbol{\Sigma}^{-1} \mathbf{W}^T (\mathbf{x}_1 - \mathbf{x}_2)}. \quad (2)$$

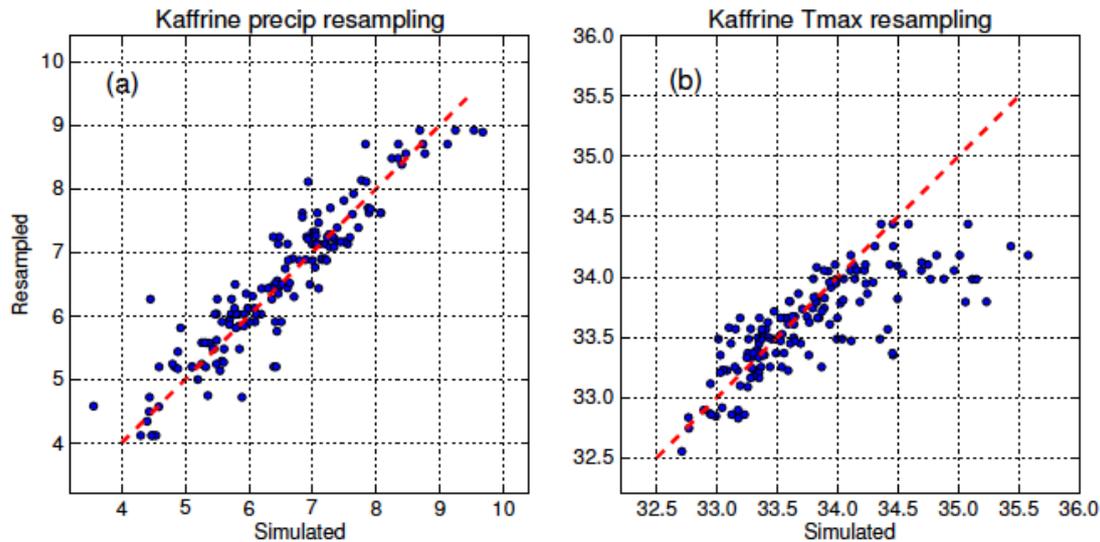
Here, \mathbf{x}_1 and \mathbf{x}_2 represent two (precip, Tmax, Tmin) triples, one the simulation target vector, the other the observational data vector for a particular year. The matrix $\boldsymbol{\Sigma}^{-1}$ represents the inverse covariance of the observations (although that of the simulations should be similar, see

the metrics discussed in Section 4.3), and \mathbf{W} is a diagonal weighting matrix. As currently constituted, the weights in \mathbf{W} are $(\frac{2}{3}, \frac{1}{6}, \frac{1}{6})$, so precipitation is effectively given twice the weight of the combined temperature variables in the distance computation. The inverse covariance has the effect of emphasising “directions” in the difference vector $\mathbf{x}_1 - \mathbf{x}_2$ that are not aligned with the principal axes of variation in the data, while having the opposite effect on differences that do align with these axes. The result is a distance measure that takes into account the internal structure of the data. The “normal” Euclidean distance is a special case of the Mahalanobis distance, in which it is assumed (or possibly known) that the data components are independent.

Once the observational year to be resampled is chosen, additive corrections are computed for its three vector components — the seasonal means of precipitation, Tmax and Tmin — so as to bring them into agreement with the values simulated for the year in question. These corrections are applied equally across individual monthly values within the season being simulated, in order to generate the final disaggregated simulation sequences.

Figure 16a shows a plot of “raw” (i.e., uncorrected) resampled precipitation values for Kaffrine JJAS against the simulated values they were selected to mimic, and provides an idea of how well the resampling works in this setting. Resampling before correction accounts for 87% of the variance in the simulated precipitation sequence, meaning that after-the-fact corrections are small, relatively speaking. Note that a portion of the scatter around the 1:1 line is due to stochastic variability introduced by the resampling procedure; it is not all due to mismatch between the nearest of the neighbours and the target vector.

Figure 16 Resampled values of (a) precipitation and (b) Tmax prior to correction, plotted against the corresponding simulated values. The dashed lines represent a 1:1 correspondence.



It should be mentioned that k-NN has its limitations. These may become significant when nonstationary simulated sequences begin to depart from the climatological domain of the observations being resampled. Here we have chosen a relatively high emissions scenario, and as a result temperatures do tend to climb above their climatological ranges, even during the relatively short course of the future simulations, which are produced for the period 2012–2050. The resampling problem is illustrated by Figure 16b, which is similar to 16a except that here it is Tmax that is resampled. There are no values of Tmax above about 34.5°C in the observational record, so as the region warms beyond this level the resampled values cannot “keep up,” and the scatter of points drifts to the right of the 1:1 line. The additive corrections that are applied as part of the downscaling scheme do bring the resampled values into agreement with the simulated values, but here these corrections are applied to a greater degree of mismatch than is the case with precipitation. A good deal of the variation in simulated Tmax (62%) is still accounted for by the resampling alone, however, and the largest temperature corrections are of order 1°C, which seems reasonable. The fact that our simulations stop in midcentury no doubt mitigates this well-known resampling issue.

6.2 Treatment of out-of-season months

In order to “concentrate” the signals to be simulated, the annual-to-decadal models have been fit to seasonalised, rather than annual-mean data, since low-precipitation months outside the rainy-season window would tend to dilute systematic patterns of wet-season variability with

low-level noise. A corresponding assumption is that follow-on modelling efforts will focus on seasons of agricultural importance. Accordingly, when the temporal disaggregation is performed it is applied only to the months within the simulated season; monthly data outside this window is incorporated in the form of climatology, and is invariant from year to year, with the exception that it is shifted, in the sense of long-term trends, either as a ratio (for precipitation) or as a fixed difference (for temperature) with respect to the wet-season trend. Thus, as temperatures increase during the 21st century both the simulated temperature variability and the static climatological temperature patterns of the out-of-season months experience the same long-term warming, but only the rainy-season months will exhibit year-to-year fluctuations. The same applies to precipitation, except that in this case the out-of-season trend is computed as a fixed ratio with respect to the wet-season trend. This simulation modality applies to all of the experimental settings.

An example of the fully-resolved monthly output is shown in Figure 17. This 8-year precipitation sequence for Kaffrine exhibits time-varying JJAS precipitation, with differing values for each month within that season, while the out-of-season months return to climatological values. As the mean JJAS precipitation shifts, over the course of years, the out-of-season trend follows along, held at a fixed interseasonal ratio with respect to the rainy season trend. The simulations then permit the focus to remain on variability in the season of agricultural interest.

Figure 17 A portion of one of the Kaffrine simulations, showing the full monthly-resolved sequence of precipitation.

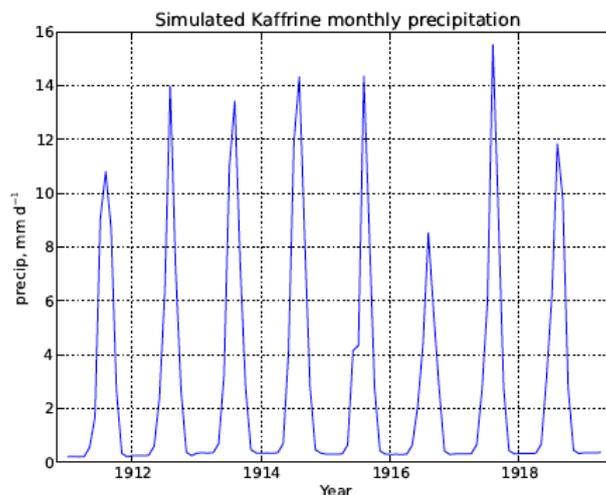
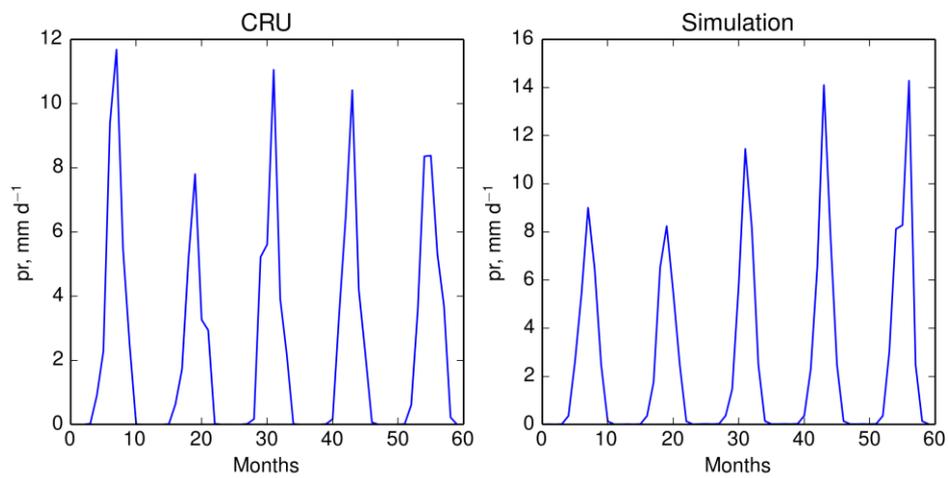


Figure 18 shows two 60-month sequences for Kaffrine, with observations (“CRU”) at left and a simulation at right. (The simulation is not designed to match observed values year for year.) Here the out-of-season months have precipitation totals approximating zero. Negative

monthly precipitation values may occasionally be simulated when the climatological out-of-season pattern is combined with the computed out-of-season trend, in particularly dry periods. Such values are set to zero. The left panel of Figure 18 suggests that this is a reasonable procedure: observed precipitation in the out-of-season months can indeed be null.

Figure 18 Observed and simulated sequences for Kaffrine.



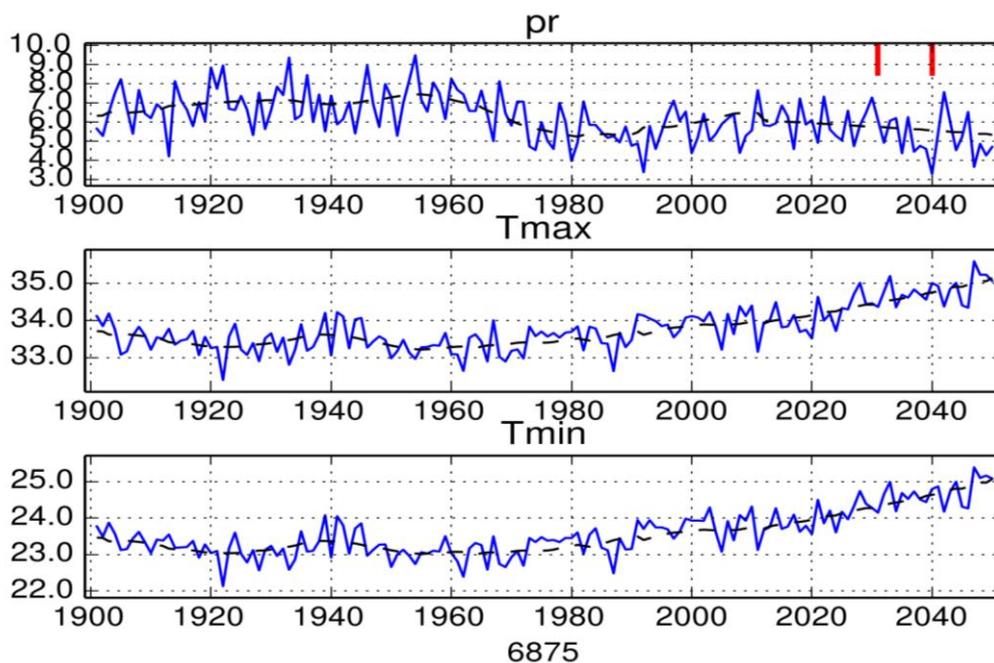
7 Example simulation sequences

We show here three example simulations to illustrate both the nature of the model output and the diversity of possible choices of simulation parameters. The three cases represent low, median and high quantiles for both future precipitation trends *and* a decadal excursion, the latter being located in the 2031-2040 window. There is no requirement that these quantiles be matched; decadal fluctuations below the 5th and above the 95th quantiles are equally likely, for example, irrespective of the trend quantile that is selected. The exact simulation specifications are at the disposal of the user, and will ultimately be governed by the applications context.

7.1 Low trend and fluctuation quantiles

Figure 19 shows a trivariate JJAS sequence (i.e., one value per year, representing the JJAS means for the three variables) for the Kaffrine area. The future precipitation trend lies at the fifth percentile of the multimodel distribution (cf. Figure 14, left panel), corresponding to a sensitivity of -9.2% per degree global temperature increase. The resultant drying can be seen in the falling off of the precipitation curve, particularly after 2020. At the same time, a decadal precipitation fluctuation lying at the fifth percentile has been located in the 2031-2040 window. This period is indicated by short vertical red lines at the top of the upper panel.

Figure 19 A simulation for Kaffrine, having low quantiles (0.05) for both the trend and a decadal precipitation fluctuation occurring during 2031-2040.



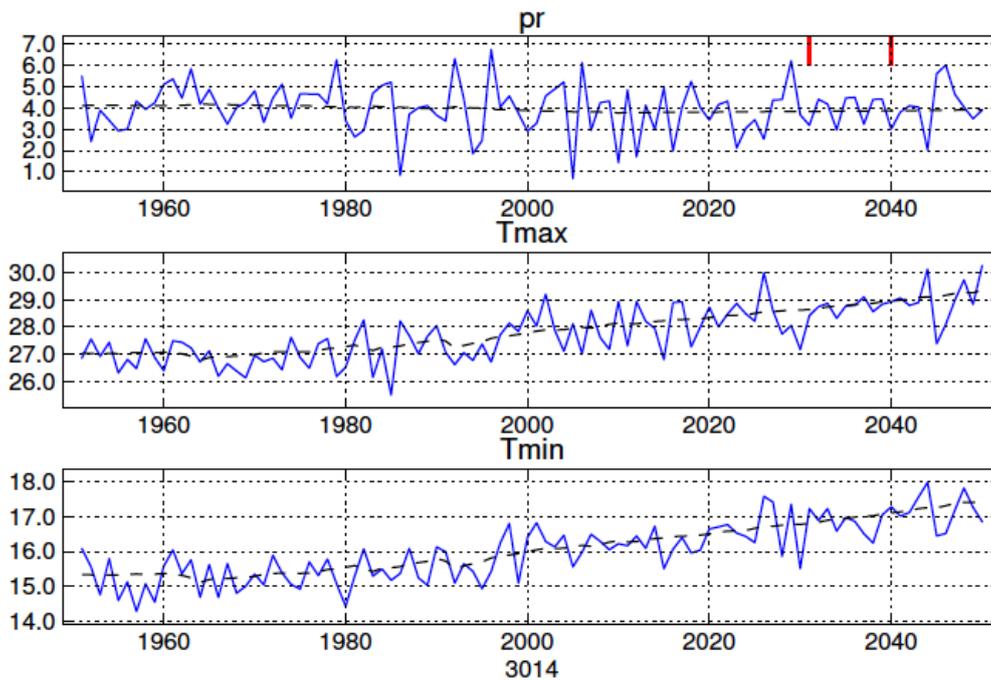
Despite a significant downward trend, the absolute shift in precipitation remains small when compared with interannual fluctuations. This situation is typical of many regions, and may have ramifications for water management, especially if some sort of storage capacity exists that can “smooth” these fluctuations. Nevertheless, the interannual variations are centred on a declining trend that will inexorably decrease water availability. Ultimately, it is only through follow-on application-based modelling that the agricultural, hydrologic and ultimately, economic impacts of such features may be evaluated in a quantitative manner.

Another feature of interest is the downward precipitation spike that occurs at 2040, which attains the lowest value in the entire record. This is an example in which trend and fluctuation combine, to produce extreme seasonal values lying outside the range of what has previously been experienced.

7.2 Median quantiles

Figure 20 shows a simulation for the Machakos area for the MAM (long rains) season. Here, both the trend and fluctuation quantiles have been specified at 0.5, the median level. The resulting precipitation trend, at 2.4% per degree global temperature increase, rises gently from its beginning in 2012 (recall that the observational record extends through 2011), while the 2013-2040 decade is marked by a series of quasi-regular precipitation cycles, with short plateaus interrupted by dry years. Again, this is just one example of the variety of within-decade patterns that may occur, regardless of the specified fluctuation quantile. If differences among such patterns are expected to have significant impacts on agricultural outcomes, the appropriate strategy would be to run a crop model, or models, over an ensemble of such sequences, in order to obtain estimates of the *distribution* of possible outcomes at a given quantile.

Figure 20 A simulation for the Machakos long rains season (MAM), having median quantiles (0.5) for both the trend and a decadal precipitation fluctuation.

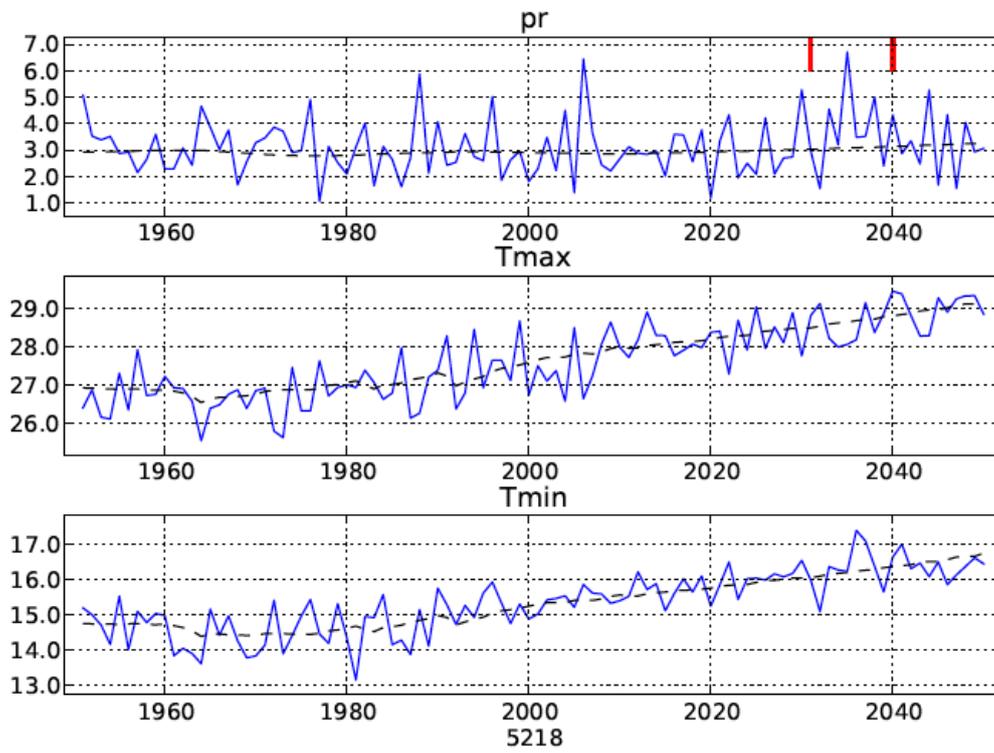


7.3 High quantiles

Figure 21 illustrates a simulation for the short rains (OND) season at Machakos. In this case the 95th percentile has been specified, again for both the precipitation trend (9.7% per degree global temperature increase) and decadal fluctuation at 2031 – 2040. Of interest here is the character of the simulated precipitation variable, which differs from those in the two previous plots. Here it has a more punctuated character, with occasional upward spikes interrupting an otherwise normally-distributed-looking sequence of values. This is a consequence of modelling the log-transformed, rather than the raw precipitation, and mimics the character of the observed OND precipitation sequence (cf. Figure 5b). Even the 2031-2040 fluctuation is coloured to some extent by this behaviour.

One can also observe, in all of these simulations, the tendency of the temperature variables to covary in the opposite sense from precipitation. For example, in Figure 21 many of the upward precipitation spikes are mirrored with downward excursions in the temperature variables. This behaviour reflects the covariances represented in the VAR model, which are estimated in turn from the observational data.

Figure 21 A simulation for the Machakos short rains, having high quantiles (0.95) for both the trend and a decadal precipitation fluctuation occurring during 2031-2040.



8 Quantifying uncertainty

8.1 Model uncertainty and regression estimates

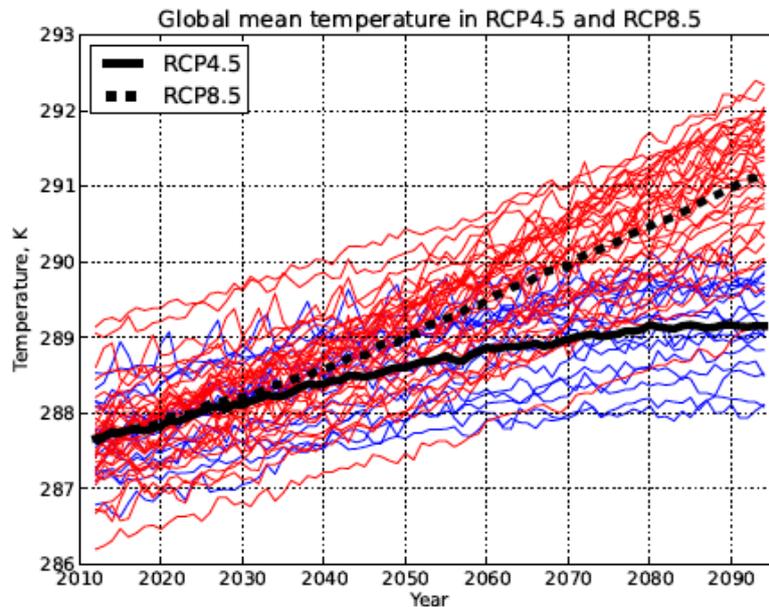
As illustrated in the above examples, when simulations are generated we specify a quantile for future trends, based on the distribution of those trends as represented in the CMIP5 ensemble listed in Table 5 (recall Figure 14). This distribution represents model uncertainty, i.e., differences owing to variations in the design and structure of climate models, as expressed in terms of regional precipitation sensitivity.

The values plotted in Figure 14 represent estimated regression coefficients, meaning that there is also a potential contribution to trend uncertainty from the errors associated with these estimates. Examination of these errors suggests that they are relatively (but not trivially) small compared with intermodel uncertainty, meaning that the spread of the true coefficient distribution is slightly wider than is represented in our model. Accordingly, trends determined by the outer quantiles of this distribution may be slightly on the conservative side. Error in regression estimates appears smaller still when scenario uncertainty is taken into account.

8.2 Scenario uncertainty

We have assumed a particular emissions scenario for these simulations, denoted as Representative Concentration Pathway 8.5 (RCP8.5). The numerical value refers to projected radiative forcing of the climate system owing to changes in trace gases, notably the so-called “greenhouse gases,” in the Earth’s future atmosphere. Of the four scenarios (*Vuuren et al.*, 2011) to be considered in the upcoming IPCC assessment, RCP8.5 is the most extreme in terms of emissions, and represents a future with “high energy demand and [greenhouse gas] emissions in [the] absence of climate change policies” (*Riahi et al.*, 2011). The various scenarios represent different hypothetical pathways for the social and economic development of human society on the planet, including sources and intensity of energy production, the degree of adoption of sustainable development practices, and so on. Figure 22 illustrates the evolution of global mean temperature under the RCP8.5 and RCP4.5 scenarios, the latter having considerably lower emissions. The two have quite different outcomes by the end of the 21st century, with RCP8.5 warming the planet post–2012 by about 3.5°C and the corresponding warming in RCP4.5 being about 1.4°C. However, the differences tend to be less pronounced prior to 2050 than later in the century, mitigating to some degree this contribution to uncertainty in our “near-term” simulations.

Figure 22 Evolution of global mean temperature, 2012–2095 in the RCP4.5 (blue) and RCP8.5 (red) scenarios. Thin traces represent individual climate models; heavy lines are the multimodel means.



8.3 Sampling variability

The examples presented in Section 7 included decadal fluctuations placed in the 2031 – 2040 window. These fluctuations are specified, as are the trends, as quantiles, and are computed on the distribution of decadal averages of the simulated variables, which, as we have seen, reproduce well the statistical properties of the observations. Of course the observational record is itself a sample, from a very long record, most of which is unobserved. It is thus likely that the statistical properties of the observed sample do not represent perfectly the true properties of the observed climate. This is a fundamental limitation associated with the short length of the instrumental record.

Aside from the variance, there are also the intervariable correlations and serial autocorrelations of the simulated data. Again, the simulations represent a subsample of the long sequence (which itself is based on a subsample of the real climate). Statistics will therefore vary from sample to sample, just as they might if we were able to resample the real climate (i.e., by “re-running” the 20th century). The comparison statistics presented in Tables 3 and 4 show that the simulations, on average, reproduce the observed variability. Given the implicit sampling variability of the instrumental record, this is quite a reasonable result.

9 Discussion

The modelling discussed in this report, although it does involve a predictive component, must be distinguished from prediction per se. Philosophically, one could say that the predictive part of our approach is less about “making predictions” than it is about passing along predictive information, along with its associated uncertainty envelope. In the present scheme this information is gleaned from the CMIP5 ensemble, and is used to inform the trend components of the simulation.

The annual-to-decadal simulation component is generated by a statistical model, whose parameters are estimated from the detrended observational record. The resulting variations are not to be construed as predictions, nor are they structured so as to reproduce specific observed variations or events. (We are not speaking of the trend component here.) One should not attempt, therefore, to identify particular observed climate events in the simulated sequences.

For the Machakos area we have treated the long- and short-rains seasons independently, because there is some evidence that they behave in this way. It is known, for example that the short rains respond to ENSO variations (*Mason and Goddard, 2001*) while the long rains are less influenced, if at all, by ENSO (*Lyon and DeWitt, 2012*). The recent downward trend of the long rains is not reflected in the short rains, and correlation between rainfall totals in two seasons is not significant. Modelling of full years, rather than separate seasonal treatment implemented here, would require that trends be unified across the two seasons, which does not seem to be the case in nature. Alternatively, one could produce a model having different precipitation trends for the two rainy seasons, but this would require more complex specifications and assumptions than it has been possible to implement in the context of the present project. Accordingly we have elected to treat the two seasons separately. This means that in simulations of the long rains (MAM), the remainder of the year, including the short rains period (OND) is modelled as climatology, and vice versa.

The modelled decadal fluctuations occurring during 2031 – 2040 (demarcated by red lines in Figures 19, 20 and 21) represent arbitrary benchmarks: One could as easily model five-year mean excursions at the 33rd and 66th percentiles and locate specified examples in the 2021-2025 window, for example. The relevant averaging lengths and time placements, as is the case for specified trend quantiles, are at the user’s disposal, and will ultimately be decided by follow-on modelling needs or preferences.

As we have noted, a constraint that 10-year mean precipitation lie at a specified quantile does not constrain the form of *within-decade* variations. There are in principle an infinite number

of 10-year sequences that correspond to a particular decadal mean specification. The extent to which differences among these sequences are material to agricultural outcomes may depend on infrastructure, such as irrigation systems or reservoirs, cultivars or other factors. As we have noted, such dependences can be explored by driving the appropriate agricultural models with an ensemble of simulations having the same decadal properties but differing within-decade patterns.

10 Conclusion

We have described the generation of synthetic data sequences at two locations, Kaffrine in the West African Sahel and Machakos in equatorial East Africa. At the latter location two seasonal models are employed, for the long and short rains respectively. The method incorporates information about long-range trends, including their uncertainty, from an ensemble of 34 GCMs contributing to the CMIP5 project. This project forms the basis of much of the upcoming IPCC assessment report.

Data at the annual-to-decadal scale is generated by an order-one vector autoregressive (VAR) model, fitted to detrended, seasonalised observational data. It was demonstrated that the first-order model is optimal, in the sense of balancing goodness-of-fit with model complexity, for the data to be simulated. Very long ($N = 10000$) sequences are generated, using the fitted models for each of the three regional/seasonal settings. These sequences are then subsampled according to specified fluctuation criteria, here for decadal precipitation lying at the 5th, 50th and 95th percentiles. The selected sequences, identified through a multilevel screening process, were then positioned in the 2031 – 2040 window. It is expected that these criteria will be adjusted according to user needs.

Pursuant to the current exercise a set of 27 simulation examples has been generated, nine for each of the regional/seasonal settings. These simulations, with full monthly time resolution, combine precipitation trends and decadal fluctuations lying at the above-mentioned percentile values (i.e., three percentile values for the 2031 – 2040 decadal fluctuation are simulated for each of the selected trend percentiles). These simulations are being provided as part of this report. Their location, organization and naming are described in the Appendix to this document.

Appendix Description of simulations provided with the report “Simulation of near-term climate change at target sites in West and East Africa”

Introduction

The report, “Simulation of near-term climate change at target sites in West and East Africa,” produced under a CCAFS small grant, described methods for generating stochastic simulations extending to 2050 for two stations, Kaffrine in Senegal and Machakos in Kenya. To complement the report document, simulations using the described methodologies were generated. This document describes these simulations and provides a Uniform Resource Locator (URL) to their location.

Simulation structure and naming

A total of 27 simulations are provided for the two locations, 9 for the Kaffrine location and 18 for Machakos. The reason for this unequal division is described in the report document, but briefly, it reflects the fact that there are two rainy seasons at the latter station. Because the linkages between these seasons are unclear they have been simulated separately, leading effectively to three “location-season” simulation targets, rather than simply two target sites defined by location alone.

For both locations the simulation method focuses on a predefined rainy season (in the case of Machakos, two rainy seasons). For the Kaffrine location this is June-September (JJAS); for Machakos the two seasons are March-May (MAM, the period of the so-called “long rains”) and October-December (OND, for the “short rains.”). Each of the 27 simulations corresponds to an individual data file, labelled by location and these seasons, followed by three additional identifiers. An example file name is

`sim_Machakos_MAM_05-05_02024.dat`

The location here is obviously Machakos, and the simulated season March-April-May. As described in the report, this means that the simulated stochastic variability will appear in the season of the long rains, with months outside MAM represented by climatological values superimposed on a gradually shifting baseline.

Future precipitation trends are modelled in terms of the local/regional response to global warming. This response is assumed known only as far as the distribution exhibited by an ensemble of CMIP5 Global Climate Models (GCMs). When a simulation is generated the user selects a quantile from this distribution; in the file name this quantile is represented by the first two-digit number following the season, in this case the value “05,” or fifth percentile. Thus, future precipitation in this simulation follows a trajectory defined by GCMs that simulate drier conditions, relatively speaking, for the Machakos long rains.

In addition to a specified trend quantile, the user may also impose a decadal precipitation anomaly; in the present group of simulations these anomalies are located in the 2031 – 2040 decade. The anomaly is also specified in terms of a quantile, in this case referring to decadal precipitation fluctuations whose statistical properties are inferred from observations. This quantile appears as the second two-digit identifier following the season abbreviation; for the simulation denoted above this is also “05.” This means that the simulation indexed above not only follows a relatively dry trajectory with respect to future climate changes, but that the 2031 – 2040 decade is also a relatively dry decade.

What about the case where the trend is toward drier conditions, but the user wishes to consider the effect of a wetter-than-normal decade? This scenario is represented in the file named

`sim_Machakos_MAM_05-95_05615.dat,`

where the second quantile identifier is “95,” meaning that the 2031-2040 decadal fluctuation in this simulation lies at the 95th percentile. The trend in this case is also toward drier conditions; in fact it is the same trend that underlies the simulation discussed above, lying at the fifth percentile.

To cover a range of possibilities, simulations have been generated for each location-season combination with trends lying at the 5th, 50th and 95th percentiles, and for each of these, decadal fluctuations during 2031 – 2040 also lying at the 5th, 50th and 95th percentiles, yielding the previously-mentioned total of 27 simulations. With this range of possible outcomes in hand, the user can explore a wide range of plausible (if not equally likely) future scenarios.

The final identifier is a 5-digit code that refers to the exact location in the long (10,000-year) simulation sequence from which the decadal fluctuation was taken. This is of less

consequence to the end user but permits the authors to reference that location should questions arise in the future.

Out-of-season months

Early versions of the simulation code filled months not included in the simulated rainy seasons with climatological values. As explained in the report, this procedure permits the annual-to-decadal statistical model to be fitted to an “undiluted” climate signal, extracting more of the meaningful variability that resides in the total (i.e., full annual) values. This procedure poses a potential problem under climate change, however, since it may result in increasingly large “jumps” between in- and out-of-season parts of the year as precipitation (or likely more significantly) temperature rises. To ameliorate this problem, climatologies of the out-of-season months are now superimposed on a slowly evolving trend. In the simulations described herein this trend is modelled in two ways: For precipitation it is computed as a fixed ratio to the in-season trend; for temperature the computation uses a fixed difference with the in-season trend. These choices were based analyses of both methods (“fixed ratio” and “fixed difference”), and comparisons of root mean square error (RMSE) differences between the out-of-season trends they produced and the actual observed out-of-season trend.

Time span and resolution

For reasons given in the report, the simulation models are based on observational years 1901-2011 for Kaffrine and 1951-2011 for Machakos. In each case the simulations begin with the first observational year, and extend through 2050. Thus, the Kaffrine simulations are of length 150 yr, those for Machakos 100 yr. The data have monthly time resolution; data for all 12 months are provided in each file.

Location of files

The simulation files can be accessed at <http://iri.columbia.edu/~amg/CCAFS/2013/>. If problems are encountered with either the files or the URL the author may be contacted at the address given near the beginning of this report.

Final notes

The deposit of these files nominally completes the set of deliverables under the CCAFS small grant that funded this project. We note that this project has focused on the methodology of simulation and the generation of the herein-described simulation data, but does not extend to the utilization of these simulations in conjunction with the household surveys conducted by CCAFS. We hope and expect that the provided files will prove useful in that regard.

References

- Akaike, H. (1973), Information theory and an extension of the maximum likelihood principle, in *Second International Symposium on Information Theory*, edited by B. N. Petrov and F. Csaki, pp. 267–281, Akademiai Kiado, Budapest.
- Baker, C. B., J. K. Eischeid, T. R. Karl, and H. F. Diaz (1995), The quality control of long-term climatological data using objective data analysis, in *Preprints of AMS Ninth Conference on Applied Climatology*, Am. Meteorol. Soc., Dallas, TX.
- Biasutti, M., and A. Giannini (2006), Robust Sahel drying in response to late 20th century forcings, *Geophys. Res. Lett.*, *33*, doi:10.1029/2006GL026067.
- Delworth, T. L., and T. R. Knutson (2000), Simulation of early 20th century global warming, *Science*, pp. 2246–2250, doi:10.1126/science.287.5461.2246.
- Dickinson, R. E., G. A. Meehl, and W. M. Washington (1987), Ice-albedo feedback in a CO₂-doubling simulation, *Clim. Change*, *10*(3), 241–248, doi:10.1007/BF00143904.
- Gilbert, P. D. (1993), State space and ARMA models: An overview of the equivalence, *Working Paper 93*, Bank of Canada.
- Greene, A. M., A. Giannini, and S. E. Zebiak (2009), Drought return times in the Sahel: A question of attribution, *Geophys. Res. Lett.*, *36*, L12701, doi: 10.1029/2009GL03886.
- Greene, A. M., L. Goddard, and J. W. Hansen (2011), A framework for the simulation of regional decadal variability for agricultural and other applications, *Tech. rep.*, International Research Institute for Climate and Society, Palisades, NY.
- Greene, A. M., M. Hellmuth, and T. Lumsden (2012), Stochastic decadal climate simulations for the Berg and Breede Water Management Areas, Western Cape province, South Africa, *Water Resour. Res.*, *48*, W06504, doi: 10.1029/2011WR011152.
- Greene, A. M., L. Goddard, and P. L. Gonzalez (2013), Stochastic simulations for agricultural planning in southeastern South America, *Int. J. Climatol.*, in prep.
- Hasselmann, K. (1976), Stochastic climate models part 1. Theory, *Tellus*, *28*(6), 473 – 485, doi:10.1111/j.2153-3490.1976.tb00696.x.
- Held, I. M., and B. J. Soden (2006), Robust responses of the hydrological cycle to global warming, *J. Climate*, *19*, 5686–5699.

- Johnson, H. L., and D. P. Marshall (2002), A theory for the surface Atlantic response to thermohaline variability, *J. Phys. Oceanogr.*, *32*, 1121–1132, doi:10.1175/15200485(2002)032<1121:ATFTSA>2.0.CO; 2.
- Lyon, B., and D. G. DeWitt (2012), A recent and abrupt decline in the East African long rains, *Geophys. Res. Lett.*, *39*, L02702, doi:10.1029/2011GL050337.
- Mason, S. J., and L. Goddard (2001), Probabilistic precipitation anomalies associated with ENSO, *Bull. Am. Met. Soc.*, *82*(4), 619–638.
- R Core Team (2013), *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria.
- Riahi, K., S. Rao, V. Krey, C. Cho, V. Chirkov, G. Fischer, G. Kindermann, N. Nakicenovic, and P. Rafaj (2011), RCP 8.5—a scenario of comparatively high greenhouse gas emissions, *Clim. Change*, *109*(1–2), 33–57, doi:10.1007/s10584-011-0149-y.
- Schwarz, G. (1978), Estimating the dimension of a model, *Ann. Stat.*, *6*(6), 461–464.
- Solomon, A., et al. (2011), Distinguishing the roles of natural and anthropogenically forced decadal climate variability: Implications for prediction, *Bull. Am. Met. Soc.*, *92*, 141–156, doi:10.1175/2010BAMS2962.1.
- Stone, D., and A. Weaver (2003), Factors contributing to diurnal temperature range trends in twentieth and twenty-first century simulations of the CCCma coupled model, *Clim. Dyn.*, *20*(5), 435–445, doi:10.1007/s00382-002-0288-y.
- Vuuren, D. P., et al. (2011), The representative concentration pathways: an overview, *Clim. Change*, *109*(1–2), 5–31, doi:10.1007/s10584-011-0148-z.



RESEARCH PROGRAM ON
**Climate Change,
Agriculture and
Food Security**



The CGIAR Research Program on Climate Change, Agriculture and Food Security (CCAFS) is a strategic initiative of CGIAR and the Earth System Science Partnership (ESSP), led by the International Center for Tropical Agriculture (CIAT). CCAFS is the world's most comprehensive global research program to examine and address the critical interactions between climate change, agriculture and food security.

For more information, visit www.ccafs.cgiar.org

Titles in this Working Paper series aim to disseminate interim climate change, agriculture and food security research and practices and stimulate feedback from the scientific community.

