RMetS

Royal Meteorological Society

Enhancing the utility of daily GCM rainfall for crop yield prediction

Amor V. M. Ines,* James W. Hansen and Andrew W. Robertson
International Research Institute for Climate and Society, The Earth Institute at Columbia University, Palisades, NY 10964-8000, USA

ABSTRACT: Global climate models (GCMs) are promising for crop yield predictions because of their ability to simulate seasonal climate in advance of the growing season. However, their utility is limited by unrealistic time structure of daily rainfall and biases in rainfall frequency and intensity distributions. Crop growth is very sensitive to daily variations of rainfall; thus any mismatch in daily rainfall statistics could impact crop yield simulations. Here, we present an improved methodology to correct GCM rainfall biases and time structure mismatches for maize yield prediction in Katumani, Kenya. This includes GCM bias correction (BC), to correct over- or under-predictions of rainfall frequency and intensity, and nesting corrected GCM information with a stochastic weather generator, to generate daily rainfall realizations conditioned on a given monthly target. Bias-corrected daily GCM rainfall and generated rainfall realizations were used to evaluate crop response. Results showed that corrections of GCM rainfall frequency and intensity could improve crop yield prediction but yields remain under-predicted. This is strongly attributed to the time structure mismatch in daily GCM rainfall leading to excessively long dry spells. To address this, we tested several ways of improving daily structure of GCM rainfall. First, we tested calibrating thresholds in BC but were found not very effective for improving dry spell lengths. Second, we tested BC-stochastic disaggregation (BC-DisAg) and appeared to simulate more realistic dry spell lengths using bias-corrected GCM rainfall information (e.g., frequency, totals) as monthly targets. Using rainfall frequency alone to condition the weather generator removed biases in dry spell lengths, improved predicted yields, but under-predicted yield variability. Combining rainfall frequency and totals, however, not only produced more realistic yield variability but also corrected under-prediction of yields. We envisaged that the presented method would enhance the utility of daily GCM rainfall in crop yield prediction. Copyright © 2010 Royal Meteorological Society

KEY WORDS   bias correction; GCM; crop models; weather generator; stochastic disaggregation

Received 4 June 2010; Revised 3 August 2010; Accepted 7 August 2010

1. Introduction

Translating seasonal climate forecasts into crop yield response is crucial for food security planning and management as they provide useful information for decision makers to formulate decisions in advance of food insecurity outlooks (Challinor *et al.*, 2005; Hansen *et al.*, 2006). Linking climate forecasts with crop models is often the method used for forecasting crop yields (Hansen and Indeje, 2004; Hansen and Ines, 2005; Semenov and Doblas-Reyes, 2007; Mishra *et al.*, 2008). The procedure, however, is not straightforward because of mismatch between the time structure of climate data required by crop simulation models and the temporal resolution of climate forecasts. Crop models usually require daily weather inputs for simulating crop growth (Tsuji *et al.*, 1994; Jones *et al.*, 2003) while climate forecasts are typically made for seasonal time scale (typically 3 months) (Goddard *et al.*, 2001; Goddard and Mason, 2002). Recent progress in addressing this problem includes linking seasonal climate forecasts with stochastic weather generators

to generate daily rainfall sequences, then used as inputs to crop simulation models (Hansen and Indeje, 2004; Hansen and Ines, 2005; Robertson *et al.*, 2007; Semenov and Doblas-Reyes, 2007; Mishra *et al.*, 2008).

Daily outputs from global or regional climate models have been proposed for bridging the gap between crop simulation models and seasonal climate forecasts (Hansen and Indeje, 2004). Dynamic climate models run in forecast mode could easily provide the daily data required by crop simulation models. However, linking daily GCM rainfall with crop models is not also as straightforward as it seems because of the biases in GCM rainfall relative to station data. Biases in rainfall statistics must be first corrected before linking them with crop models as distortions in rainfall frequency and intensity could adversely impact the simulation of crop growth and yield, resulting in unrealistic outcomes (Mearns *et al.*, 1996; Riha *et al.*, 1996; Mavromatis and Jones, 1999; Hansen and Jones, 2000; Baron *et al.*, 2005). Devising ways for extracting and refining information from daily GCM rainfall for crop simulations are crucial for their successful applications in crop yield predictions.

Ines and Hansen (2006) developed a bias correction (BC) method for daily GCM rainfall for correcting

* Correspondence to: Amor V. M. Ines, International Research Institute for Climate and Society, The Earth Institute at Columbia University, Palisades, NY 10964-8000, USA. E-mail: ines@iri.columbia.edu

biases in rainfall frequency, intensity and totals relative to a station inside a grid cell. They used a two-step approach by correcting rainfall frequency, then the intensity distribution. Rainfall frequency was corrected by truncating the cumulative distribution function (CDF) of daily GCM rainfall from the observation distribution above a calibrated threshold. Mapping the quantiles of the truncated daily GCM rainfall CDF onto the observations was used to correct rainfall intensity. As rainfall total is the product of frequency and intensity, correcting both the biases of intensity and frequency will correct rainfall total bias. Schmidli *et al.* (2006) also proposed a slightly varied method for correcting daily GCM rainfall biases. They used a wet-day threshold to match observed and model rainfall frequency, similar to Ines and Hansen (2006), and a scaling factor, based on the ratio between wet-day intensities of observations and adjusted GCM rainfall, to correct rainfall intensity. A recent comparison between the two methods can be seen in Dobler and Ahrens (2008). GCM BC has been used in other applications as well (e.g. Sharma *et al.*, 2007; Elshamy *et al.*, 2009; Mishra and Singh, 2009; Mishra *et al.*, 2009) aside from agro-meteorological applications (Baigorria *et al.*, 2007, 2008).

Bias-corrected daily GCM rainfall used as input to crop simulation models has made improvements in predicted yields compared to using raw daily GCM rainfall alone (Hansen *et al.*, 2006). Ines and Hansen (2006) observed, however, that a significant bias (negative) in the predicted yields still exists after BCs of daily GCM rainfall. Their analysis suggests that the time structure of daily GCM rainfall does not correspond well with station data and that daily GCM rainfall tends to overestimate dry spell lengths (here, ≤ 1 mm = dry day) exposing the crops to longer periods of dry spells, thus reducing yields. The deterministic GCM BC was not able to correct this bias in dry spell lengths.

The general purpose of this article is to test the notion that by improving the daily structure of GCM rainfall (in bias-corrected or in derived form), the predicted yields by crop simulation models can be improved further. Specifically, we aim (1) to explore strategies for improving the time structure of bias-corrected daily GCM rainfall and (2) to evaluate the performance of the improved daily GCM rainfall for crop yield predictions. We tested the methods in an agricultural experimental station in Katumani, Machakos district of eastern Kenya, using maize as a case study.

2. Materials and methods

2.1. Climate data

We used ECHAMv4.5 (Roeckner *et al.*, 1996) simulated daily rainfall forced by concurrent observed sea-surface temperature (SST) (http://iridl.ldeo.columbia.edu) for BC and crop yield prediction. Note that a GCM run in historical mode delineates the upper bound of climate predictability and that predictive skill in forecast mode (i.e. forced by forecast SSTs) may be lesser than this level (Challinor *et al.*, 2005). For purposes of testing and demonstration, we chose model runs in historical mode. Here, we expanded the study domain considered by Ines and Hansen (2006) from a single to multiple grid cells to test if multi-grid GCM cells can provide better information for crop yield prediction. We selected 3×3 model grid cells (1°23.7′N–4°10.8′S, 33°45′E–39°20.5′E) encompassing the Katumani Dryland Research Center (1°35′S, 37°14′E) (Table I, Figure 1) wherein all 24 ensemble members (GCM simulations from different initial conditions) in each grid cell considered were extracted for analysis for the study period (1970–1995). Rainfall in this region is bi-modal, occurring in February–May and October–December. The latter shorter rainy season is an important maize growing season in the region whose predictability potential has been established in recent studies (e.g. Indeje *et al.*, 2000; Hansen and Indeje, 2004). Daily rainfall observations and other weather data needed for

Table I. Geographic locations, distance from station, weights and uncorrected yield correlations of the selected ECHAMv4.5 grid cells.

Location	Longitude	Latitude	Distance, r (degrees)	Weight, $\omega(r)$ (degrees^{-2})	Uncorrected yield R (−) and MBE (Mg ha^{-1}, in parenthesis)	Corrected yield[†] R (−) and MBE (Mg ha^{-1}, in parenthesis)
Station	37°14′E	1°35′S	0.000	Null	Null	Null
Grid1	33°45′E	4°11.16′S	4.348	0.053	0.60 (−0.32)	0.61 (−0.18)
Grid2	36°33.75′E	4°11.16′S	2.688	0.138	0.68 (−2.18)	0.71 (−0.85)
Grid3	39°22.5′E	4°11.16′S	3.371	0.088	0.43 (−2.84)	0.36 (−0.62)
Grid4	33°45′E	1°23.72′S	3.488	0.082	0.04 (0.76)	0.67 (−0.20)
Grid5	36°33.75′E	1°23.72′S	0.697	2.060	0.61 (−2.35)	0.70 (−0.90)
Grid6	39°22.5′E	1°23.72′S	2.150	0.216	0.44 (−3.13)	0.38 (−0.30)
Grid7	33°45′E	1°23.72′N	4.583	0.048	0.40 (0.52)	0.58 (−0.15)
Grid8	36°33.75′E	1°23.72′N	3.053	0.107	0.65 (−2.10)	0.61 (−0.70)
Grid9	39°22.5′E	1°23.72′N	3.669	0.074	0.59 (−3.17)	0.45 (−0.34)

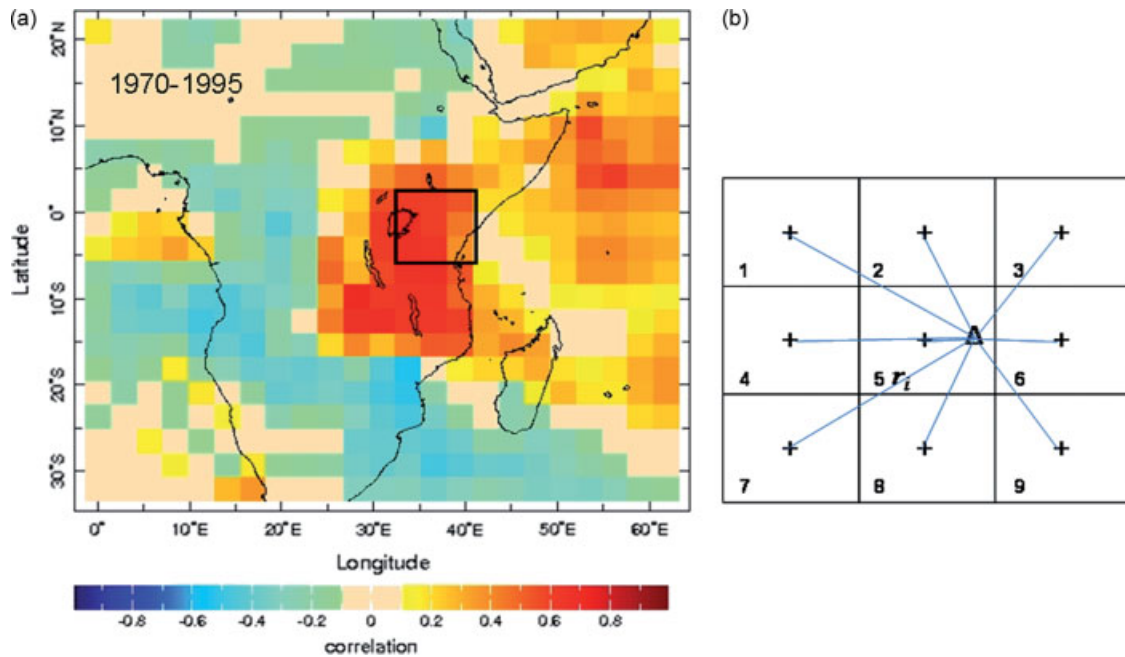[†] Using the new bias correction approach, BC2 (>0 mm threshold; Section 2.2.1).

Figure 1. (a) Correlation map of raw October–December ECHAMv4.5 rainfall and maize yield simulated by observed weather at Katumani Dryland Research Center; (b) schematic of grid cells and spatial kernel ($\omega_i = 1/r_i^2$) used determined the contribution of a grid cell to a station location ($\triangle$); r_i denotes for distance of grid cell at midpoint (+) from station location (Table I).

crop simulations, such as daily maximum (T_{max}) and minimum temperature (T_{min}), and solar radiation (SRAD), were also collected from the research station.

2.2. Improving time structure of GCM rainfall

2.2.1. BC with calibrated thresholds

The first approach used to improve the time structure of daily GCM rainfall is by calibrating rainfall thresholds. The GCM BC (Ines and Hansen, 2006) corrects rainfall frequency by fitting a GCM threshold value $\tilde{x}_{I,GCM}$ to truncate the empirical distribution of uncorrected daily GCM rainfall. As a result, mean rainfall frequency of calendar month I above the threshold matches the observed. The $\tilde{x}_{I,GCM}$ value is calculated by inverting the empirical cumulative rainfall distribution of GCM ($F_{I,GCM}$) as (Equation (1)):

$$\tilde{x}_{I,GCM} = F_{I,GCM}^{-1}[F_{I,obs}(\tilde{x}_{I,obs})] \tag{1}$$

where $\tilde{x}_{I,obs}$ is an observed threshold value for delineating rainy days from observed rainfall (Figure 2; Hansen *et al.*, 2006). Varying $\tilde{x}_{I,obs}$ magnitude during BC runs could alter the time structure of bias-corrected daily GCM rainfall. This is expected because $\tilde{x}_{I,GCM}$ also changes whenever $\tilde{x}_{I,obs}$ is varied.

An update to the GCM BC (BC1) is outlined below; we call the updated version BC2 onwards. The general notion of daily GCM rainfall is that it always over-predicts rainfall frequency. But this is not always the case; hence we implemented a new BC formulation that can handle cases when GCM rainfall frequency is less than or more than the observed. When GCM rainfall frequency is less than the observed, we append a number of wet-day events
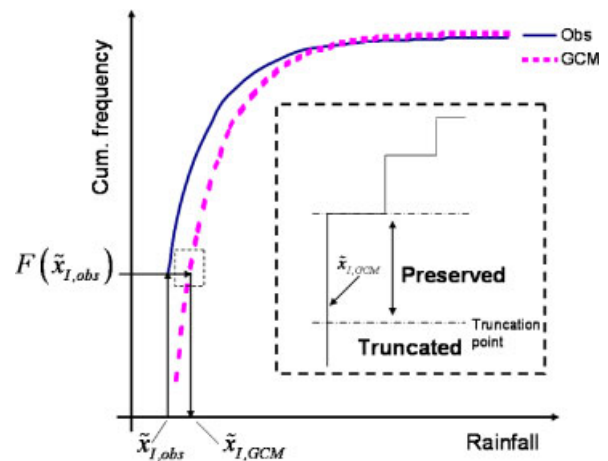

Figure 2. The new bias correction method (BC2) correcting the left side of the empirical distribution if $\tilde{x}_{I,GCM}$ values are too many ('nugget effect') to preserve equal number of wet days with observations (see inset). This figure is available in colour online at wileyonlinelibrary.com/journal/joc

with minimum rainfall amounts (i.e. $\tilde{x}_{I,obs} + 0.1$ mm) in the GCM rainfall CDF to match the observed rainfall frequency. The number of appended wet-days depends on the discrepancy between the number of wet-days in the GCM rainfall CDF above the calibrated threshold and observations. In turn, these wet days are added and distributed evenly (stochastically) in the GCM rainfall daily time series. Tests show, however, that this case is not always the norm, and if it happens, only small number of wet-days per calendar month is added in the time series. The update also handles a case if the left-side of the empirical distribution of GCM rainfall contains more values equivalent to $\tilde{x}_{I,GCM}$ above the truncation

point (Figure 2) by preserving those values instead of eliminating them ('nugget effect') ensuring that rainfall frequency is preserved. This case often occurs during BC, which was not handled in the earlier version.

We used only the two-parameter gamma distribution $G(x; \alpha, \beta)$ to simulate rainfall intensity distributions of both observed and GCM rainfall, both for BC1 and BC2. This so-called gamma–gamma transformation corrects rainfall intensity by mapping gamma-transformed daily GCM rainfall onto the gamma distribution of the observed (truncated) rainfall. GCM rainfall x on day i is corrected as (Equation (2)):

$$x_i' = G_{\text{I,obs}}^{-1}[G_{\text{I,GCM}}(x_i; x_i > \tilde{x}_{\text{I,GCM}})], \text{ otherwise, } 0.$$

(2)

Shape and scale parameters of GCM and observed rainfall gamma distributions are calculated using maximum likelihood method.

2.2.2. BC-stochastic disaggregation (DisAg)

Daily GCM rainfall tends to be highly auto-correlated, hence it tends to cluster and persist but with low intensities (e.g. Ines and Hansen, 2006, 2009). Calibrating thresholds on GCM BC may alter the daily variations of rainfall within rain clusters in the time series but may not be able to correct the mismatch in dry spells especially between rain clusters. For this reason, a stochastic way of redistributing the bias-corrected daily GCM rainfall information within a time series was sought. The purpose is to preserve the corrected GCM wet days in a calendar month, while the time structure of rainfall occurrence is stochastically adjusted based on that corrected probability of wet days. Here, we nested the GCM BC with a stochastic disaggregation method described in Hansen and Ines (2005) to generate sequences of daily rainfall conditioned on the bias-corrected GCM rainfall information (e.g. rainfall frequency) to produce a more realistic dry spell length distribution. In the temporal disaggregation, we focused in using corrected rainfall frequency as conditioning data for generating daily rainfall sequences as this relates directly to daily statistics of GCM rainfall. However, we also used the combination of frequency and rainfall totals for conditioning the stochastic disaggregation process to augment the deficiency of using rainfall frequency information alone.

The rainfall occurrence model of the stochastic weather generator used in the stochastic disaggregation is a two-state, hybrid second-order Markov chain able to simulate rainfall occurrence with a first-order chain if the previous day was wet, or a second-order chain if the previous day was dry (Hansen and Ines, 2005; Stern and Coe, 1984; Wilks, 1999). Literature suggests that higher order Markov chain models better simulate dry spell distributions than first-order models (Wilks and Wilby, 1999). If the Markov model simulates occurrence of rainfall in a given day, the rainfall amount is sampled from a mixture of two exponential distributions (Woolhiser and Roldán, 1982). Details of the temperature and SRAD sub-models are described in Hansen and Mavromatis (2001).

The rainfall frequency of bias-corrected daily GCM rainfall was used to adjust the first- and second-order transition probabilities of the rainfall occurrence model to simulate time series of wet and dry days. The formalism for adjusting these transition probabilities (i.e. first-order probabilities, e.g. wet day following (−>) a dry day (p_{01}), wet->wet (p_{11}), and second-order probabilities, p_{101} and p_{001}) to match a target rainfall frequency (e.g. from bias-corrected GCM rainfall) is shown in Hansen and Ines (2005) and elsewhere. But for the purpose of completeness, we also show it here. The transition probabilities of a first-order, two-state Markov chain rainfall occurrence model are related directly to the unconditional rainfall occurrence probability, π (Equation (3)) and persistence of dry days, ρ_1 (Equation (4)) (Katz and Parlange, 1998; Wilks and Wilby, 1999).

$$\pi = \frac{p_{01}}{1 + p_{01} - p_{11}}$$

(3)

$$\rho_1 = p_{11} - p_{01}$$

(4)

Assuming that persistence of dry days remains constant when a target rainfall occurrence probability changes, π', the first-order adjusted transition probabilities (with apostrophes) are determined by solving Equations (3) and (4) simultaneously, resulting to,

$$p_{01}' = \pi'(1 - \rho_1)$$

(5)

$$p_{11}' = \rho_1 + p_{01}'$$

(6)

Equations (5) and (6) are used to determine a wet day if the previous day was wet. If the previous day was dry, the second-order Markov chain is used to determine if the current day will be wet or dry through the adjusted second-order transition probabilities. The transition probabilities for the hybrid second-order Markov chain rainfall occurrence model can be adjusted for a given rainfall frequency by keeping the first- and second-order persistence of dry days [ρ_1 (Equation (4)) and ρ_2 (Equation (7))] constant.

$$\rho_2 = p_{101} - p_{001}$$

(7)

The adjusted transition probabilities are given as follows (Equations (8)–(10)) (Katz and Parlange, 1998; Hansen and Mavromatis, 2001)

$$p_{11}' = \pi'(1 - \rho_1) + \rho_2$$

(8)

$$p_{001}' = \pi' \frac{(1 - \rho_1)(1 - \pi') - \rho_2(1 - p_{11}')}{1 - \pi'}$$

(9)

$$p_{101}' = p_{001}' + \rho_2$$

(10)

If rainfall frequency and total are used at the same time to condition the stochastic weather generator, (1) transition probabilities (first- and second-order) of the rainfall occurrence model are adjusted based on bias-corrected monthly GCM rainfall frequency, (2) daily

Table II. Summary descriptions of methods used in this study.

Methods	Descriptions
BC1	Based on Ines and Hansen (2006), corrects rainfall frequency (under-prediction) by truncating GCM rainfall distribution at a given threshold, corrects rainfall intensity using gamma–gamma transformation
BC1-DisAg1	BC1 combined with stochastic disaggregation (DisAg) conditioned on the corrected rainfall frequency of each ensemble member, respectively
BC1-DisAg2	BC1 combined with DisAg conditioned on the corrected average rainfall frequency from all ensemble members
BC2	Corrects rainfall frequency (under- and over-prediction) by truncating the GCM rainfall distribution at a given threshold, and also corrects the 'nugget effect' of truncating empirical distribution (Figure 2). Corrects rainfall intensity using gamma–gamma transformation
BC2-DisAg1	BC2 combined with DisAg conditioned on the corrected rainfall frequency of each ensemble member, respectively
BC2-DisAg2	BC2 combined with DisAg conditioned on the corrected average rainfall frequency from all ensemble members
BC2-m	BC2 applied to multiple grid cells
BC2-DisAg2-m	BC2 combined with DisAg conditioned on the corrected average rainfall frequency (and amounts) from all multiple grid cells

rainfall realizations are generated iteratively until generated monthly rainfall total matches 95% of the target value (the bias-corrected monthly GCM rainfall), (3) generated daily values are re-scaled by a ratio of monthly target (R_m) and generated rainfall totals ($\overline{R}_m$) (i.e. $R_m/\overline{R}_m$) such that the monthly rainfall total generated matches the target value. These three steps are repeated in turn, for each calendar month in a year, for all the years considered (Hansen and Ines, 2005).

2.3. Case studies for BC and BC-stochastic disaggregation

Table II shows a summary of the experiments performed in this study. First, we considered a single ECHAMv4.5 grid cell encompassing the agricultural research station (Figure 1) to compare the performances of BC1 and BC2. The 24 ensemble members were individually bias-corrected using several observed threshold values ($>0, 1, 3$ and 5 mm). Monthly rainfall statistics from the bias-corrected daily GCM rainfall were extracted for linking to stochastic disaggregation. In the single grid cell case, only the monthly time series of rainfall frequency (π) was used to condition the stochastic weather generator (as this can directly represent the daily GCM rainfall statistics), (1) for each ensemble member (BC-DisAg1) and (2) aggregated 24 members (BC-DisAg2), where BC is the general term for GCM bias correction (BC) and DisAg for stochastic disaggregation.

The second case involves multiple grid cells from ECHAMv4.5 where all nine grid cells selected were used in the analysis (Figure 1). Using only now, the new BC formulation (BC2), the 24 ensemble members from each GCM grid cell were bias-corrected individually, then analysed to extract monthly rainfall statistics. As will be shown later, individual ensemble member contains lesser information than when combined with other ensemble members in the GCM run. Therefore, we used only the aggregated monthly rainfall frequency, and rainfall totals, for each grid cell to develop a grid-based monthly time series of rainfall frequency and totals for stochastic

disaggregation. This was done using a simple spatial kernel given below (Equations (11)–(13); Figure 1(b)).

$$\overline{\pi}_t = \frac{\sum_{i=1}^{N} \omega_i \pi_{it}}{\sum_{i=1}^{N} \omega_i} \ \forall t \qquad (11)$$

$$\omega_i = \frac{1}{r_i^2} \qquad (12)$$

$$r_i = \sqrt{(x_i - x_{station})^2 + (y_i - y_{station})^2} \qquad (13)$$

where $\overline{\pi}$ is the weighted conditioning data (e.g. frequency, totals), ω is a weighting function, r is the grid cell distance from the weather station, x and y are longitude and latitude, i is an index of grid cell, t an index for time (i.e. months, for 26 years), and N is the number of grid cells. We also used arithmetic averaging (all grid cells have the same weights) for developing the grid-based conditioning data for stochastic disaggregation.

We generated 24 realizations of daily rainfall for all the BC-stochastic disaggregation experiments. The 24 realizations were replicated ten times to minimize the variability of small samples in stochastic modelling. The 24-realization run was chosen for the stochastic disaggregation to give a better basis for comparison with the BC runs applied to 24 ensemble members.

As will be shown in Section 2.4, we linked the corrected information with a crop simulation model. This requires that the corrected daily GCM rainfall must be coupled with other weather variables considering their dependency with rainfall. There are several ways to satisfying this requirement. Baigorria et al. (2007) suggested bias correcting climate model's daily temperature (T) and SRAD in addition to rainfall for inputs to crop yield prediction. This approach attempts to ensure that T and SRAD are consistent with rainfall, but since they are bias-corrected independently, a perfect dependency with

rainfall is not always guaranteed in the end. This can be applied using BC alone. We can also use the weather generator to generate daily realizations of temperature (T_{min} and T_{max}) and SRAD conditioned on the time series of bias-corrected GCM rainfall. In this study, however, we opted for a simpler approach. For all BC and BC-DisAg runs, daily T_{min}, T_{max} and SRAD were generated as monthly mean values conditioned on the occurrence or nonoccurrence of rainfall. In other words, if the current day is wet, we used the wet-day mean values of T_{min}, T_{max} and SRAD of that month as values on that particular day, vice versa. This is analogous to filling missing data with mean + noise, with noise = 0 and the mean is a conditional value. Conditioning the values on a wet or dry day occurrence preserves the co-variation of rainfall with other weather variables. Our earlier tests show that this approximation resulted to yields comparable to those simulated with actual weather.

2.4. Linking daily GCM rainfall with a crop model

The performance of BC and BC-DisAg outputs were quantitatively evaluated by linking them to a crop simulation model. We used CERES-Maize (Ritchie *et al.*, 1998) to simulate and predict maize yields. Data inputs on soil properties, local cultivar '*Katumani composite B*' characteristics and management assumptions were based on a previous study at the same site (Keating *et al.*, 1992). The sandy clay loam soil used in the crop modelling has plant-extractable water-holding capacity of 180 mm m^{-1} of soil. For each simulation year, the soil-water balance was initialized on 17 October with initial soil moisture at 20% water-holding capacity. Sowing was established by the CERES-Maize and was simulated once the soil moisture content of the top 15 cm soil surface exceeded 40% of water-holding capacity. If sowing fails within the prescribed window, sowing will commenced forcedly on 1 November. The plant density was assumed to be 4.4 plants m^{-2} with 50 cm row spacing and 20 kg N ha^{-1} applied as ammonium nitrate (NH$_4$NO$_3$) at sowing. Other details are described in Ines and Hansen (2006).

CERES-Maize was run with observed daily rainfall (>0 mm threshold) for baseline, daily rainfall from ECHAMv4.5 without BC, with BC (all thresholds, >0, 1, 3, 5 mm) under BC1 and BC2, and with variants of BC (all thresholds) - stochastic disaggregation (Table II). Simulated yields for individual years were averaged across the 24 GCM ensemble members and from the 24 realizations (replicated ten times) by the stochastic disaggregation, then compared with the baseline yields. To ensure that yields are simulated with consistent temperature and SRAD, we used the conditional mean values of T_{min}, T_{max} and SRAD in the baseline simulation in lieu of the observed values.

2.5. Analysis of results

Standard goodness-of-fit statistics was used to analyse model performances. We decomposed the mean squared error (Equation (14)) into a random component (not corrected by linear regression) (Equation (15)) and a systematic component (can be corrected by linear regression) (Equation (16)) based on Willmott (1982),

$$MSE = \frac{1}{n} \sum_{i=1}^{n} (\hat{y}_i - y_i)^2 \qquad (14)$$

$$MSE_R = \frac{1}{n} \sum_{i=1}^{n} (\hat{y}_i^* - y_i)^2 \qquad (15)$$

$$MSE_S = MSE - MSE_R \qquad (16)$$

where n is the number of years, i, y and $\hat{y}$ are yields simulated with observed and uncorrected/corrected GCM rainfall, $\hat{y}^*$ is $\hat{y}$ calibrated by linear regression. We also used correlation coefficient (R), mean bias error (MBE), root-mean-squared error (RMSE) and index of agreement d-statistics (Equation (17)) to measure performance, where $\overline{y}$ is average observed yields.

$$d = 1 - \frac{\sum_{i=1}^{n} (\hat{y}_i - y_i)^2}{\sum_{i=1}^{n} (|\hat{y}_i - \overline{y}| + |y_i - \overline{y}|)^2}. \qquad (17)$$

Aside from yields, goodness-of-fit statistics for bias-corrected and generated GCM-based rainfall were calculated to evaluate the performance of the enumerated methods for enhancing the utility of daily GCM rainfall for crop yield prediction. Of particular interest is the evaluation of dry spell lengths among methods, as this was suggested to be the major influencing factor for the underestimation of yields when linking daily GCM rainfall with crop simulation models.

3. Results and discussions

3.1. Crop yield predictions

3.1.1. Without BC

Due to lack of available actual yield data, we opted to pursue the analysis with maize yields simulated by observed weather, with the exception of temperature and SRAD values (Section 2.4). ECHAMv4.5 October–December (raw) rainfall is moderately–highly correlated with the baseline yield (Figure 1(a)). This is interesting to note because it suggests the good predictability of maize yields in the region.

Figure 3 shows the predicted yields using uncorrected daily GCM rainfall from the selected 3 × 3 grid cells. Except for grid cell 4 (Figure 1(b)), a moderate to strong correlation (0.40–0.68, Table I) exists between yields simulated by observed station rainfall and those simulated by uncorrected GCM rainfall from surrounding grid cells, suggesting that the inter-annual variability of rainfall was more or less captured by the climate model for the majority of grid cells selected, as evident in Figure 1(a).

But most of the yields simulated by uncorrected daily GCM rainfall extremely under-predicted those simulated by observed station rainfall (Table I, MBE uncorrected). This can be explained in part by the rainfall characteristics of the selected grid cells. The seasonal rainfall (October–December) was extremely underestimated (Table III, uncorrected).

For comparison purposes, we selected three grid cells representing the three clusters of simulated yield levels (Figure 3) for in-depth analysis. Here, we selected grid cells 3, 4 and 5 from the bottom, top and middle clusters, respectively; grid cell 5 was the chosen study domain in Ines and Hansen (2006) and kept it here for better comparison of methods. Figure 4 shows the rainfall characteristics (total, intensity and frequency) of grid cells 3, 4 and 5 compared to observed station weather. On the average, although rainfall frequency was over-predicted in all three grid cells during the growing season (sowing to anthesis, October–December), all the ensemble members in grid cells 3 and 5 underestimated rainfall intensities and totals. Apparently, rainfall intensities from these two grid cells were too low to saturate the rootzone that even with too many rainfall events during the growing season, did not improve the simulated yields. Ines and Hansen (2006) suggested that some of this yield bias may be attributed to the longer dry spells (≤ 1 mm = dry day) associated with daily GCM rainfall as the low intensity rains, even if they are frequently occurring, do not satisfy crop water requirements as they tend to be only lost by

Figure 3. Uncorrected yields from the 3×3 ECHAMv4.5 grid cells compared to station data. This figure is available in colour online at wileyonlinelibrary.com/journal/joc

canopy interception or evaporated before reaching the soil surface. In contrast, grid cell 4 showed a relatively lesser yield mean bias error (MBE) than grid cells 3 and 5, but poses the worst yield correlation (Table I; Figure 3). Rainfall intensity in grid cell 4 was fairly simulated well by the GCM, but the frequency was extremely over-predicted resulting to wetter growing seasons (Figure 4) and perhaps causing this higher than average but lowly correlated yields (poor inter-annual variability) (Table I; Figure 3). The above discussion exhibits the complexity of interactions between rainfall (and its time structure) and crop growth.

3.1.2. BC-stochastic disaggregation, Case 1: single grid cell

Using grid cell 5 as a case study, we tested and evaluated the BC-stochastic disaggregation methods (Table II) designed for extracting useful information from daily GCM rainfall for crop prediction. This grid cell encompasses the agricultural experimental station (Figure 1) used in this study. In Section 3.1.1, we showed that grid cell 5 always over-predicts rainfall frequency (Figure 4). The inability of BC1 to correct the 'nugget effect' of truncating empirical rainfall distributions (Figure 2) can undermine BC of rainfall frequency. This grid cell therefore is a good test-bed for the performance of the improved BC method (BC2, Table II) developed in this study.

Overall, BC and its combination with stochastic disaggregation for tuning daily GCM rainfall information improved the simulation of crop yields, compared to no corrections made in the GCM rainfall (Figure 5). BC of rainfall frequency and intensity of GCM rainfall could improve the systematic and random errors in the predicted yields (Table IV). The yield correlations are higher and MBE lower after BCs. This trend is shown by the increased d-statistics and reduced MSE of the predicted yields. But a majority of the total error (MSE) after BC is still systematic in nature (MSE_s) due to mean bias (Figure 5), corroborating the earlier findings of Ines and Hansen (2006).

Calibrating thresholds on GCM BC had little impact to the performance of the predicted yields (Table IV;
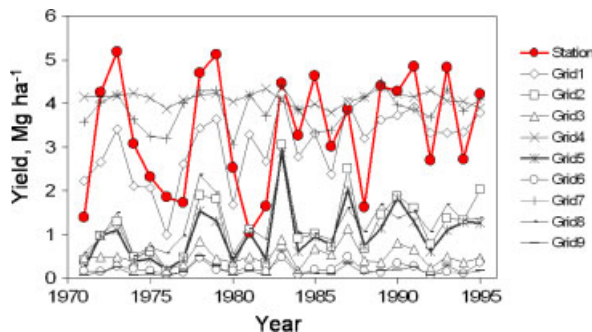
Table III. Seasonal GCM rainfall (October–December) from different grid cells.

|	Station obs.	1	2	3	4	5	6	7	8	9
Uncorrected										
Totals, mm d^{-1}	2.96	2.71	1.34	0.93	5.51	1.17	0.53	2.75	1.22	0.40
Intensity, mm wd^{-1}	7.34	3.49	2.00	1.85	6.17	1.70	1.23	3.41	2.06	1.03
Frequency, wd d^{-1}	0.36	0.70	0.56	0.43	0.87	0.59	0.35	0.75	0.50	0.30
Corrected[†]										
Totals, mm d^{-1}	2.96	2.99	2.97	3.01	2.97	2.97	3.29	2.99	2.94	3.40
Intensity, mm wd^{-1}	7.34	7.19	6.62	7.03	7.02	6.55	7.53	7.42	6.74	7.87
Frequency, wd d^{-1}	0.36	0.36	0.36	0.36	0.36	0.37	0.36	0.36	0.36	0.36

[†] Using BC2 (Table II).
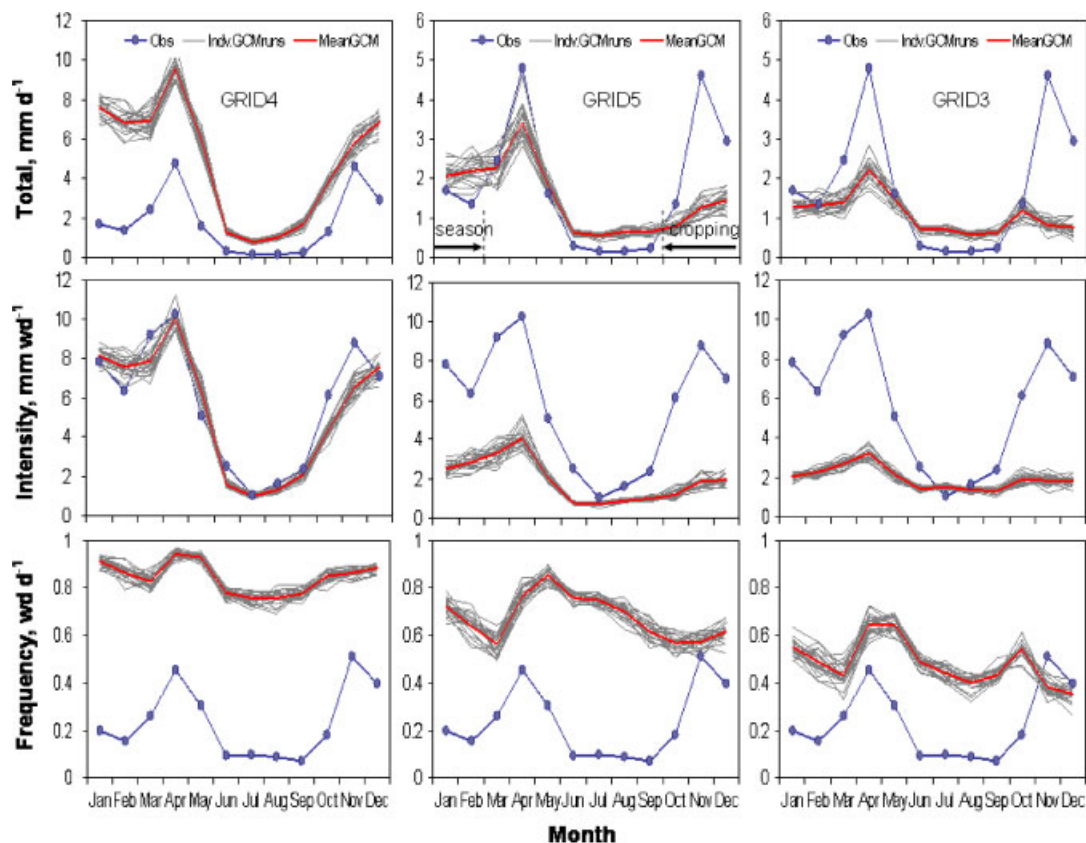
Figure 4. Sample uncorrected rainfall totals (top row), intensity (middle row) and frequency (bottom row) of selected grid cells, 4 (first column), 5 (second column; Ines and Hansen, 2006) and 3 (third column), compared to station data. This figure is available in colour online at wileyonlinelibrary.com/journal/joc
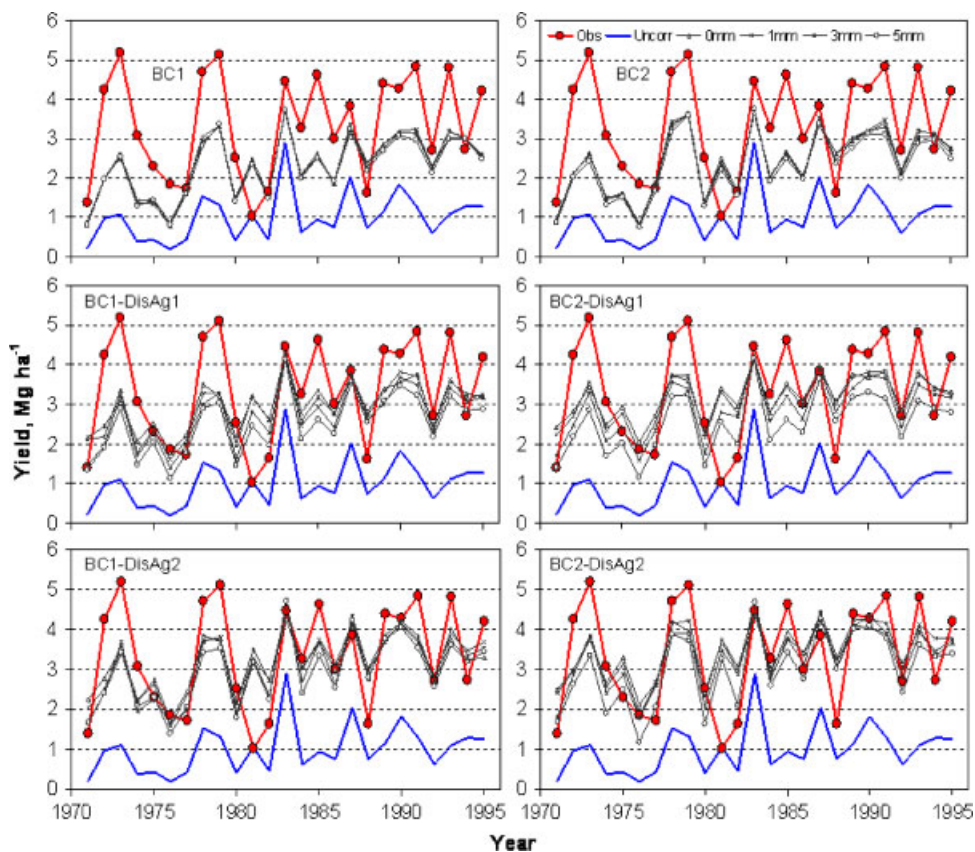

Figure 5. Performance of bias correction – stochastic disaggregation in crop yield prediction, grid cell 5. This figure is available in colour online at wileyonlinelibrary.com/journal/joc

Table IV. Performance of the bias correction-stochastic disaggregation methods on crop yield prediction when correcting only for over-predictions in GCM rainfall frequency (BC1) (grid cell 5).

Thresholds ($\tilde{x}_{I,obs}$)	R (−)	MBE (Mg ha^{-1})	d (−)	MSE (Mg ha^{-1})^2	MSE$_R$ (Mg ha^{-1})^2	MSE$_S$ (Mg ha^{-1})^2
Uncorrected						
–	0.61	−2.35	0.50	6.61	1.06	5.55
BC1						
>0 mm	0.70	−1.04	0.67	1.95	0.86	1.09
>1 mm	0.70	−0.98	0.67	1.84	0.86	0.98
>3 mm	0.72	−1.01	0.68	1.86	0.82	1.04
>5 mm	0.71	−1.07	0.67	2.01	0.84	1.17
BC1-DisAg1						
>0 mm	0.63	−0.41	0.66	1.22	1.01	0.21
>1 mm	0.65	−0.42	0.68	1.18	0.97	0.21
>3 mm	0.66	−0.64	0.69	1.37	0.95	0.42
>5 mm	0.67	−0.84	0.68	1.65	0.93	0.72
BC1-DisAg2						
>0 mm	0.73	−0.20	0.74	0.91	0.79	0.12
>1 mm	0.68	−0.14	0.75	0.95	0.92	0.04
>3 mm	0.72	−0.26	0.80	0.88	0.81	0.07
>5 mm	0.69	−0.39	0.75	1.04	0.89	0.16

Uncorrected and BC1 at threshold >0 mm values are excerpt from Ines and Hansen (2006).

Figure 5). Increasing observed threshold values ($\tilde{x}_{I,obs}$) tend to reduce a little the total error (MSE) in predicted yields (up to a certain threshold). But we noticed that increasing the observed threshold value ($\tilde{x}_{I,obs}$) for a wet-day reduces the rainfall frequency (both observed and GCM) and shifts the rainfall intensity distribution to the right, causing lesser frequent but higher intensity bias-corrected GCM rainfall (compared to observation at >0 mm threshold). The good performance of >3 mm threshold for predicting crop yields could be an artifact of this process. Moderately higher threshold value could lead to favourable rainfall conditions for crop growth due to resulting wetter growing season, caused by the more intense but relatively lesser frequent rainfall, compared to bias-corrected GCM rainfall conditions generated by >0 mm threshold. Extreme threshold values (e.g. >5 mm), however, can produce significantly (negatively) biased number of wet days (compared to observation at >0 mm) that even with very intense rainfall (distribution extremely shifted to the right) could not improve crop yields due to longer dry spells caused by extremely lesser rainfall events.

Applying BC and stochastic disaggregation to individual ensemble members (BC1-DisAg1) reduced more the systematic errors of predicted yields than from using BC1 alone, but yield correlations are lower, hence more random errors (MSE$_R$) in the predicted yields (Table IV). It should be noted that ensemble members contained noisier information individually than when pooled collectively, as shown by the results in BC1-DisAg2. Although a redundant way of exploration, conducting the BC1-DisAg1 confirms the need for multi-ensemble simulations to capture better physical processes. When averaged rainfall frequency from all ensemble members was used to condition the stochastic disaggregation, it did not only preserve higher yield correlations, but also improved systematic errors in the predicted yields (see also d-statistics; Table IV).

More improvements in the predicted yields were observed when correcting the 'nugget effect' (Figure 2) of truncating empirical distributions (Table V; Figure 5). Note that BC2 is able to correct both over- and under-prediction of rainfall frequency in the GCM data (Section 2.2.1) although the latter capability was not used in this case because the number of wet days in grid cell 5 was always over-predicted (Figure 4). But since BC1 cannot correct the 'nugget effect' of truncating empirical distributions, the corrected long-term rainfall frequency tends to be underestimated when the 'nugget effect' is present during BC. BC2 ensures that the number of wet days is consistent between calibrated GCM and observed rainfall when this situation occurs during BC. The impact of this GCM BC update is mostly observed in the systematic errors of predicted yields (Tables IV and V; Figure 5, BC2 *vs* BC1) suggesting the sensitivity of number of wet days in the crop simulations.

Outcomes of BC2 improvements are noticed more in the performance of individual ensemble members when combined with stochastic disaggregation (BC2-DisAg1). When BC2-improved rainfall frequency for each ensemble member was used to condition the stochastic weather generator, yield correlations were higher and systematic errors lower than BC1-DisAg1 (Tables IV and V). Using BC2-DisAg2 (Table II) further reduced mean bias errors compared to BC1-DisAg2, although other goodness-of-fit indicators gave somewhat mixed results (Tables IV and V).

Overall, BC-stochastic disaggregation using averaged GCM information (BC-DisAg2) gave the best results

Table V. Performance of the bias correction-stochastic disaggregation methods on crop yield prediction when correcting for both under- or over-predictions and 'nugget effects' in GCM rainfall frequency (BC2) (grid cell 5).

Thresholds ($\tilde{x}_{I,obs}$)	R (−)	MBE (Mg ha^{-1})	d (−)	MSE (Mg ha^{-1})^2	MSE$_R$ (Mg ha^{-1})^2	MSE$_S$ (Mg ha^{-1})^2
Uncorrected						
−	0.61	−2.35	0.50	6.61	1.06	5.55
BC2						
>0 mm	0.70	−0.90	0.69	1.69	0.87	0.82
>1 mm	0.71	−0.89	0.70	1.64	0.83	0.81
>3 mm	0.72	−0.93	0.70	1.69	0.81	0.87
>5 mm	0.71	−1.04	0.68	1.93	0.85	1.09
BC2-DisAg1						
>0 mm	0.67	−0.14	0.66	1.06	0.94	0.12
>1 mm	0.70	−0.20	0.71	0.98	0.87	0.11
>3 mm	0.70	−0.44	0.71	1.10	0.86	0.24
>5 mm	0.68	−0.79	0.68	1.57	0.92	0.65
BC2-DisAg2						
>0 mm	0.70	0.06	0.71	0.96	0.86	0.09
>1 mm	0.69	0.13	0.76	0.92	0.88	0.04
>3 mm	0.72	−0.03	0.79	0.84	0.82	0.02
>5 mm	0.69	−0.33	0.77	0.99	0.88	0.11

Uncorrected values are excerpt from Ines and Hansen (2006).

Table VI. Performance of the bias correction (>0 mm threshold)-stochastic disaggregation on crop yield prediction with BC2 (multiple grid cells).

Methods	R (−)	MBE (Mg ha^{-1})	d (−)	MSE (Mg ha^{-1})^2	MSE$_R$ (Mg ha^{-1})^2	MSE$_S$ (Mg ha^{-1})^2
Uncorrected-m[†]	0.62	−2.25	0.50	6.19	1.03	5.15
BC2-m[†]	0.69	−0.78	0.68	1.51	0.88	0.63
BC2-DisAg2-m ($\overline{\pi}$)[‡]	0.72	0.10	0.71	0.95	0.82	0.13
BC2-DisAg2-m ($\overline{\pi} + \overline{R_m}$)[‡]	0.70	0.13	0.84	0.98	0.86	0.11
BC2-DisAg2-m ($\overline{\pi}$)[§]	0.70	0.13	0.66	1.05	0.87	0.18
BC2-DisAg2-m ($\overline{\pi} + \overline{R_m}$)[§]	0.67	0.36	0.80	1.07	0.92	0.15

[†] Weighted average yields.
[‡] Weighted average frequency ($\overline{\pi}$), and frequency ($\overline{\pi}$) + totals ($\overline{R_m}$).
[§] Arithmetic average $\overline{\pi}$, and $\overline{\pi} + \overline{R_m}$.

among the methods tested. Not only it produced higher yield correlations but also reduced drastically systematic bias in predicted yields, suggesting that post-process corrections of systematic errors by linear regression (e.g. Baigorria *et al.*, 2008) may no longer be necessary, if the GCM bias-correction-stochastic disaggregation was applied in crop yield prediction.

3.1.3. BC-stochastic disaggregation, Case 2: multiple grid cells

Because of scale and process aggregation is GCM schemes, sometimes climate signals from a GCM are geographically shifted and that grid cells other than the grid cell containing the study location may better predict a local phenomenon (Wilks, 1995). Here, we used all nine surrounding grid cells to predict maize yields at the station using now only BC2 (Figure 1). Interestingly, grid cell 4, which initially had very low yield correlations improved yield simulations drastically ($R = 0.60$) after

BC (Table I; Figure 3), suggesting further that correcting daily GCM rainfall for rainfall frequency and intensity biases could improve systematic errors as well as random errors in predicted yields provided there is a skill of the GCM. Table III shows the improvements in seasonal (October–December) rainfall for each grid cell after BCs.

Averaging schemes (distance-based and equal weighting) of yields from the bias-corrected grid cells had little impact on the total error (MSE) of predicted yields (Tables V and VI; Figures 6 and 5; see BC2-m and BC2). Some surrounding grid cells have lower skills than grid cell 5 (Table I) and that blending them with this cell, which is nearest to the station and happens to be predicting yield better (Table I), did not impact the final outcome. Goodness-of-fit statistics showed that results from the multiple grid cell analysis are more or less similar with the single
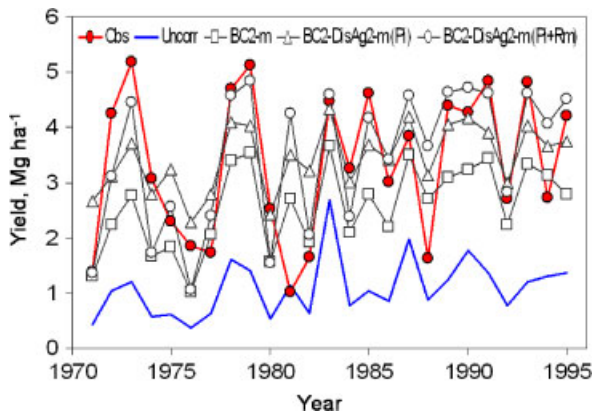
Figure 6. Performance of the bias-correction stochastic disaggregation using multiple grid cells. This figure is available in colour online at wileyonlinelibrary.com/journal/joc

Table VII. Average, standard deviation (std. dev.) and coefficient of variation (cv) in predicted yields using multiple grid cell information (>0 mm threshold).

Methods	Average (Mg ha^{-1})	Std. dev. (Mg ha^{-1})	c.v. (−)
Observed	3.34	1.33	0.40
Uncorrected	1.09	0.53	0.49
BC2-m	2.57	0.76	0.30
BC2-DisAg2-m ($\overline{\pi}$)[†]	3.44	0.59	0.17
BC2-DisAg2-m ($\overline{\pi} + \overline{R_m}$)[†]	3.47	1.25	0.36

[†] Weighted average frequency ($\overline{\pi}$), and frequency ($\overline{\pi}$) + totals ($\overline{R_m}$).

grid cell case (Tables V and VI). But this could be an artifact of the inverse-distance weighting method used in summarizing multiple grid cell yield simulations.

Using all corrected rainfall frequency (π alone) from surrounding grid cells to condition the weather generator improved all goodness-of-fit statistics in yield predictions (Table VI; Figure 6). Conditioning rainfall frequency while constraining the generation of monthly rainfall totals at the same time ($\pi + R_m$) also improved the degree of similarity (*d*-statistics) in predicted yields (Table VI). Using $\pi + R_m$ for conditioning, the weather

generator not only improved the mean of predicted yield distribution but also its variance (Table VII) suggesting that for a more realistic yield distribution, it is necessary to make sure that frequency of rainfall and its time structure are corrected, as well as the amount of rainfall generated. Giving the same weights to each grid cell's average π and R_m for stochastic disaggregation did not make any improvements in crop yield prediction as compared to inverse-distance-weighting results (Table VI; Section 2.3).

3.2. Properties of bias-corrected and stochastically-disaggregated daily GCM rainfall

Figure 7 shows the effects of the BC methods on rainfall total, intensity and frequency (grid cell 5). The
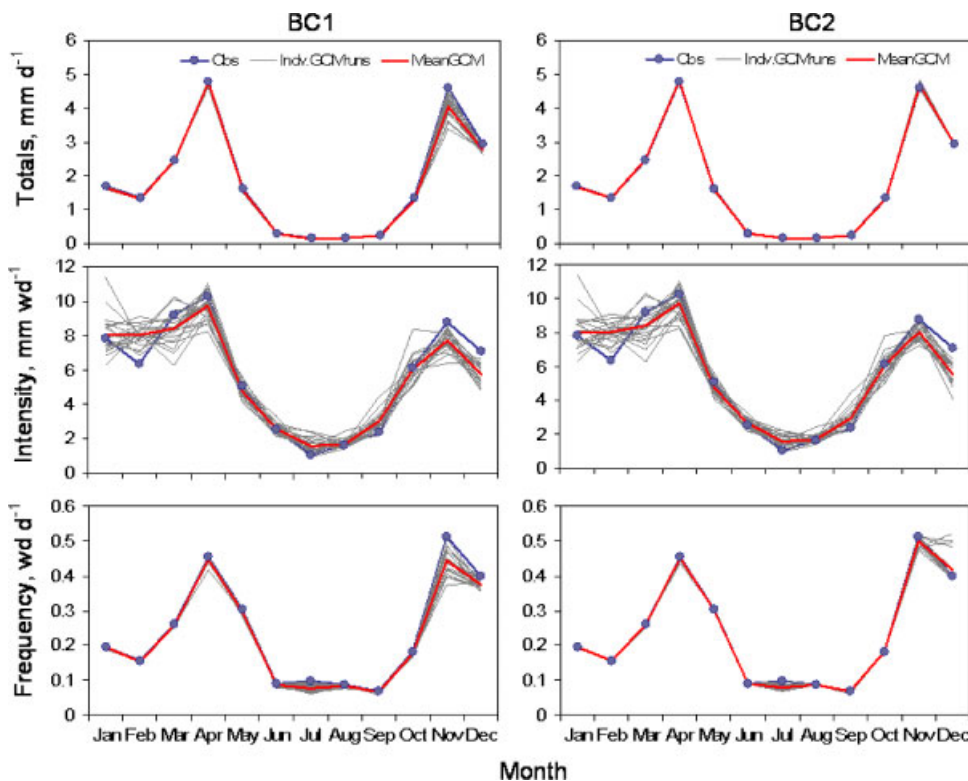

Figure 7. Corrected rainfall totals (top row), intensity (middle row) and frequency (bottom row) of grid cell 5 (>0 mm threshold) using BC1 (first column; Ines and Hansen, 2006) and BC2 (second column), compared to station data. This figure is available in colour online at wileyonlinelibrary.com/journal/joc
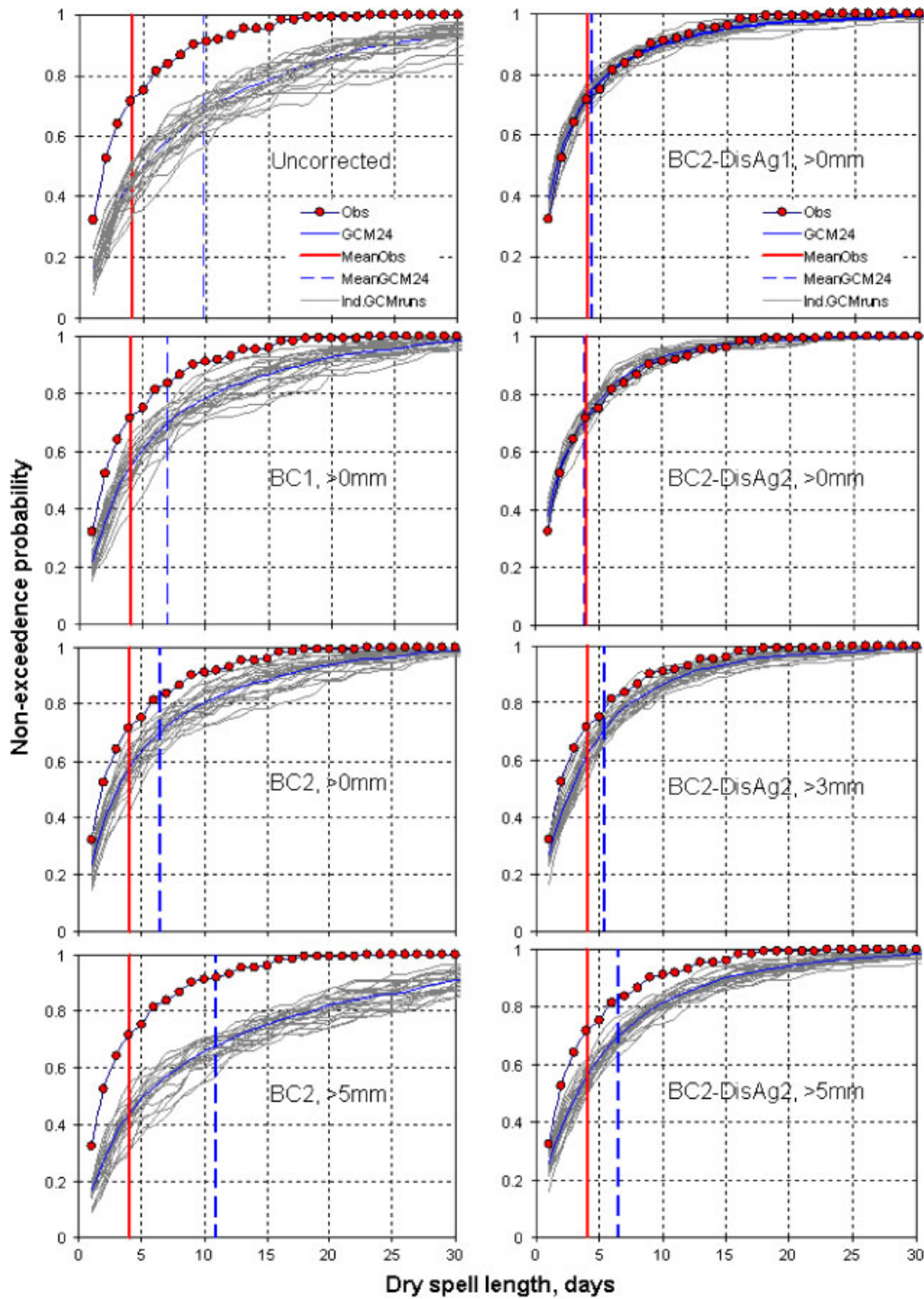
Figure 8. Dry spell length distributions (15 November–31 December) of Uncorrected GCM rainfall (grid cell 5), corrected by BC1, BC2, BC2-DisAg1 and BC2-DisAg2 at various threshold values compared to station data. Uncorrected and BC1 results based on Ines and Hansen (2006). (Note: ≤1 mm is dry day). This figure is available in colour online at wileyonlinelibrary.com/journal/joc

advantage of BC2 over BC1 (Table II) is mostly apparent at rainfall frequency and total corrections (Figure 7, BC2). As BC1 was not able to correct the 'nugget effect' of truncating empirical distributions (Figure 2), the corrected GCM rainfall frequency underestimated reality (e.g. October–December) (Figure 7, BC1). BC2 corrected this by implementing the 'nugget effect' correction although some room for improvements is apparent (Figure 7, BC2).

Raw ECHAMv4.5 daily rainfall over-predicts dry spell lengths (Figure 8; Uncorrected), which was attributed as one of the main causes of the large negative mean bias in predicted yields after BC (Ines and Hansen, 2006). Dry spell length bias, however, is attributed to the unrealistic time structure of daily GCM rainfall, which deterministic GCM BC methods may not be able to correct. Corrections of rainfall frequency, however, led to small improvements in dry spell length bias (Figure 8; BC1, >0 mm). The

updated BC algorithm (BC2) slightly improved the dry spell length distribution compared to BC1 (Figure 8; BC2, >0 mm). Increasing the threshold did not improve dry spell length (Figure 8; BC2, >5 mm).

And this all boils down to the research question of: Does improving time structure of bias-corrected GCM rainfall effectively improve yield simulations? Table VIII suggests that calibrating thresholds could not correct dry spell length bias. But there is evidence that BC itself, and the improvements made with the GCM BC method (BC2) could correct some of the systematic bias in dry spell lengths, however, this was not enough. The improvements in simulated yields using bias-corrected daily GCM rainfall are a compounded effect of corrected rainfall frequency and intensity (and hence total), including the slight nudge to the bias of dry spell lengths especially during flowering stage. Figure 9 shows the improvements in seasonal rainfall (October–December) after BC.

But distributing better the bias-corrected number of wet days within a calendar month using a hybrid-order (i.e. first-, when yesterday was wet, and second-order, when yesterday was dry) Markov chain rainfall occurrence model could lead to a more realistic daily rainfall distribution, including simulations of dry spells (Figure 8). Notice that even with the case of BC2-DisAg1 (>0 mm; one GCM member, 24 realizations) the generated daily

Table VIII. Summary of dry spell length (days) statistics of bias-corrected daily GCM rainfall, grid cell 5 (15 November–31 December).

| Thresholds ($\tilde{x}_{\mathrm{I,obs}}$)	Mean	Std. Dev.	c.v.	Std. dev. of mean across ensemble
Uncorrected				
–	9.82	10.24	1.04	1.24
Observed				
–	4.05	4.18	1.03	–
BC1				
>0 mm	7.05	7.53	1.07	1.05
>1 mm	7.20	7.64	1.06	0.85
>3 mm	9.10	9.65	1.06	0.91
>5 mm	11.01	11.33	1.03	0.91
BC2				
>0 mm	6.50	6.99	1.07	0.86
>1 mm	6.97	7.36	1.06	0.84
>3 mm	8.81	9.48	1.08	1.04
>5 mm	10.93	11.46	1.05	0.86
BC1-DisAg2				
>0 mm	3.99	4.40	1.10	0.35
>1 mm	4.53	4.78	1.05	0.29
>3 mm	5.78	6.18	1.07	0.40
>5 mm	6.64	7.32	1.10	0.51
BC2-DisAg2				
>0 mm	3.85	4.25	1.10	0.22
>1 mm	4.33	4.54	1.05	0.37
>3 mm	5.37	5.88	1.09	0.43
>5 mm	6.47	7.12	1.10	0.49

≤1 mm is equivalent to dry day.
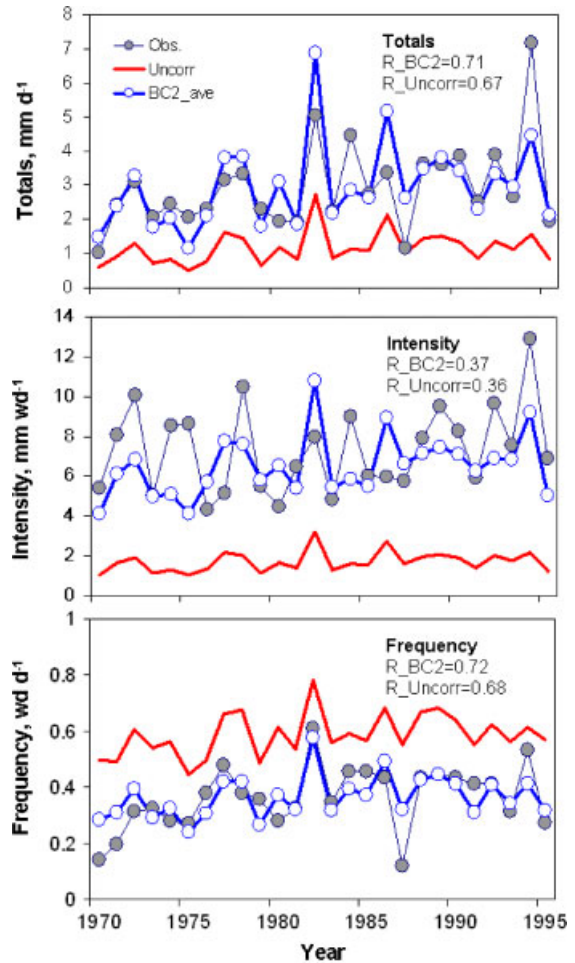
Figure 9. Seasonal rainfall characteristics (October-December) after bias correction, grid cell 5 (BC2, >0 mm threshold). This figure is available in colour online at wileyonlinelibrary.com/journal/joc

rainfall already matched the dry spell length distribution of the station rainfall. There was only a slight over-prediction of the tail of the dry spell distribution (i.e. more longer dry spells from individual realization). When using the averaged ensemble members' corrected rainfall frequency (BC2-DisAg2, >0 mm) to condition the weather generator, this led to a better fit of the simulated dry spell length distribution. Similar results were obtained from BC1-DisAg runs, thus it is not shown here. The only difference is that, using individual member's information alone (BC1-DisAg1) over-predicted the longer dry spells more than BC1-DisAg2. Dry spell length distributions from the multi-grid cell analysis were mostly similar to BC2-DisAg2 results using a single grid cell. As mentioned earlier, this could be an artifact of the inverse distance weighting method used to summarize multi-grid cell information.

Based on these results, we suggest that the improvements made in the predicted yields, from BC to BC-DisAg (Tables II, IV and V), could be attributed to the better time structure of the generated daily rainfall obtained from the bias-corrected daily GCM information. This more realistic time structure corrected the bias
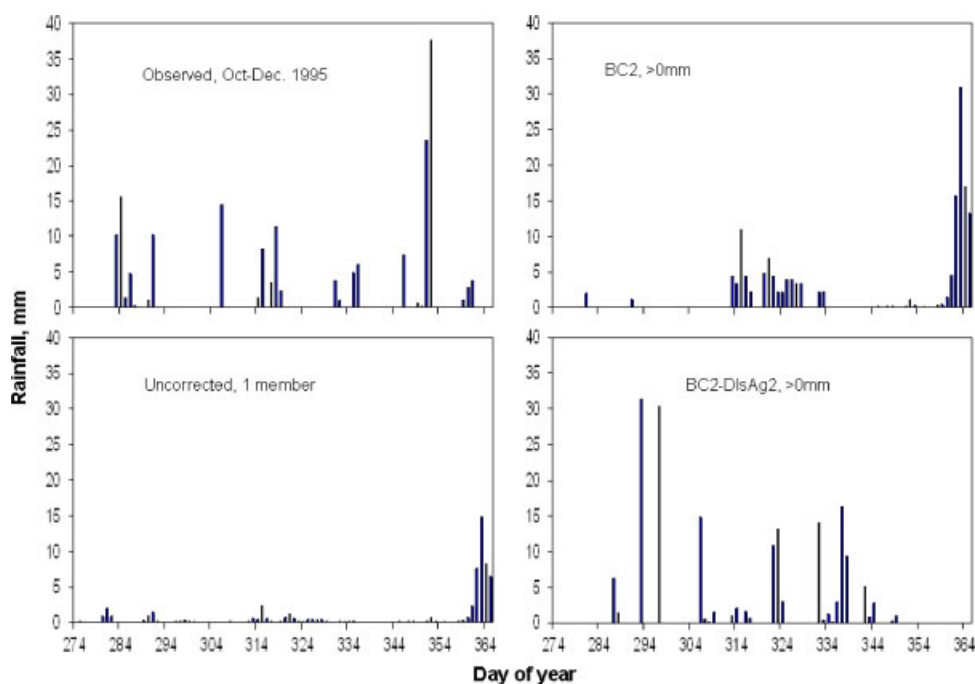
Figure 10. Sample daily rainfall data, (October-December 1995) before and after bias correction, stochastic disaggregation; grid cell 5. This figure is available in colour online at wileyonlinelibrary.com/journal/joc

in dry spell lengths, thus reducing the mean bias and eliminated most of the systematic errors in the predicted yields (Tables IV and V). Figure 10 shows a sample of bias corrected daily GCM rainfall (October–December, 1995) and how BC-DisAg distributed the number of rainfall events using the bias-corrected GCM information (>0 mm threshold). Table VIII also shows the dry spell length statistics with BC-DisAg using various threshold levels. Until >1 mm threshold, the corrected mean dry spell length approximates the observed. Beyond that, we observed a significant deviation but this is more pronounced with BC alone (Figure 8, BC-DisAg, higher thresholds).

4. Summary and conclusions

We present a methodology for eliciting useful information from daily GCM rainfall. The BC of daily GCM rainfall can improve crop yield prediction compared to no correction, although the predicted yields are still negatively biased. This improvement in yield simulations was associated with better rainfall frequency, intensity and total of the corrected GCM rainfall. However, deterministic BC cannot correct the time structure of daily GCM rainfall, although it was shown that rainfall frequency correction could slightly improve dry spell length distributions.

Calibrating thresholds on GCM BC was not successful to improve the time structure of GCM rainfall. In fact, as we increased the observed threshold value delineating a wet day, we also decreased the number of wet days, contributing more to longer dry spells in the resulting daily GCM rainfall. Threshold calibration also tends to shift the rainfall intensity distribution to the right.

Coupling GCM BC with stochastic disaggregation (BC-DisAg) was more successful for simulating better dry spell length distributions from the corrected daily GCM rainfall information. Using monthly series of bias-corrected GCM rainfall frequency to adjust transition probabilities of the hybrid-order Markov chain rainfall occurrence model corrected the dry spell length bias adequately. This correction of time structure removed most of the mean bias error in predicted yields, but not the spread of the yield distribution. Combining bias-corrected GCM rainfall frequency and totals for stochastic disaggregation improved both the mean of predicted yield distribution and its variance.

Furthermore, the improvements made to the GCM BC method accounting for over- and under-prediction of rainfall frequency, and the 'nugget effect' correction for truncating empirical distributions, resulted to better GCM rainfall frequency correction and in the correction of dry spell lengths when coupled with stochastic disaggregation. We did not get better information from using multiple grid cells because the grid cell nearest to the agricultural experimental station dominated the solutions. Using more robust aggregation methods for summarizing information better for the multi-grid cell case (e.g. optimized weighting of grid cell information (not based on distance alone), Model Output Statistics (MOS) correction, etc.) may circumvent the caveat of distance-weighted method.

the extraction of ECHAMv4.5 daily GCM rainfall from the IRI data library. We thank the editor and two anonymous reviewers for helping us improve the quality of the article.

References

Baigorria GA, Jones JW, Shin DW, Mishra A, O'Brien JJ. 2007. Assessing uncertainties in crop model simulations using daily bias-corrected Regional Circulation Model outputs. *Climate Research* **34**: 211–222.

Baigorria GA, Jones JW, O'Brien JJ. 2008. Potential predictability of crop yield using an ensemble climate forecast by a regional circulation model. *Agricultural and Forest Meteorology* **148**: 1353–1361.

Baron C, Sultan B, Balme M, Sarr B, Traore S, Lebel T, Janicot S, Dingkuhn M. 2005. From GCM grid cell to agricultural plot: scale issues affecting modelling of climate impact. *Philosophical Transactions of the Royal Society B* **360**: 2095–2108.

Challinor AJ, Slingo JM, Wheeler TR, Doblas-Reyes FJ. 2005. Probabilistic simulations of crop yield over western India using DEMETER seasonal hindcasts ensembles. *Tellus A* **57**: 498–512.

Dobler A, Ahrens B. 2008. Precipitation by a regional climate model and bias correction in Europe and South Asia. *Meteorologische Zeitschrift* **17**: 499–509.

Elshamy ME, Seierstad IA, Sorteberg A. 2009. Impacts of climate change on Blue Nile flows using bias-corrected GCM scenarios. *Hydrology and Earth System Sciences* **13**: 551–565.

Goddard L, Mason SJ. 2002. Sensitivity of seasonal climate forecasts to persisted SST anomalies. *Climate Dynamics* **19**: 619–632.

Goddard L, Mason SJ, Zebiak SE, Ropelewski CF, Basher R, Cane MA. 2001. Current approaches to seasonal to interannual climate predictions. *International Journal of Climatology* **21**: 1111–1152.

Hansen JW, Challinor A, Ines AVM, Wheeler T, Moron V. 2006. Translating climate forecasts into agricultural terms: advances and challenges. *Climate Research* **33**: 27–41.

Hansen JW, Indeje M. 2004. Linking dynamic seasonal climate forecasts with crop simulation for maize yield prediction in semi-arid Kenya. *Agricultural and Forest Meteorology* **125**: 143–157.

Hansen JW, Ines AVM. 2005. Stochastic disaggregation of monthly rainfall data for crop simulation studies. *Agricultural and Forest Meteorology* **131**: 233–246.

Hansen JW, Jones JW. 2000. Scaling-up crop models for climate variability applications. *Agricultural Systems* **65**: 43–72.

Hansen JW, Mavromatis T. 2001. Correcting low-frequency variability bias in stochastic weather generators. *Agricultural and Forest Meteorology* **109**: 297–310.

Indeje M, Semazzi FHM, Ogallo LJ. 2000. ENSO signals in East African rainfall and their prediction potentials. *International Journal of Climatology* **20**: 19–46.

Ines AVM, Hansen JW. 2006. Bias correction of daily GCM rainfall for crop simulation studies. *Agricultural and Forest Meteorology* **138**: 44–53.

Ines AVM, Hansen JW. 2009. Extracting useful information from daily GCM rainfall for cropping system modeling. AgSAP Conference 2009. Egmond Aan Zee: The Netherlands.

Jones JW, Hoogenboom G, Porter CH, Boote KJ, Batchelor WD, Hunt LA, Wilkens PW, Singh U, Gijsman AJ, Ritchie JT. 2003. The DSSAT cropping system model. *European Journal of Agronomy* **18**: 235–265.

Katz RW, Parlange MB. 1998. Overdispersion phenomenon in stochastic modeling of precipitation. *Journal of Climate* **11**: 591–601.

Keating BA, Wafula BM, Watiki JM. 1992. Exploring strategies for increased productivity – the case for maize in semi-arid Eastern Kenya. In *A Search for Strategies for Sustainable Dryland Cropping in Semi-arid Eastern Kenya*, ACIAR Proceedings, No. 41. Probert ME (ed). Australian Centre for International Agricultural Research: Canberra, 90–101.

Mavromatis T, Jones PD. 1999. Evaluation of HADCM2 and direct use of daily GCM data in impact assessment studies. *Climatic Change* **41**: 583–614.

Mearns LO, Rosenzweig C, Goldberg R. 1996. The effects of changes in daily and interannual climatic variability on CERES-Wheat: a sensitivity study. *Climatic Change* **32**: 257–292.

Mishra A, Hansen JW, Dingkuhn M, Baron C, Traore SB, Ndiaye O, Ward MN. 2008. Sorghum yield prediction from seasonal rainfall forecasts in Burkina Faso. *Agricultural and Forest Meteorology* **148**: 1798–1814.

Mishra AK, Singh VP. 2009. Analysis of drought severity-area-frequency curves using a general circulation model and scenario uncertainty. *Journal of Geophysical Research D: Atmospheres* **114**(6): D06120.

Mishra AK, Özger M, Singh VP. 2009. Trend and persistence of precipitation under climate change scenarios for Kansabati basin, India. *Hydrological Processes* **23**: 2345–2357.

Riha SJ, Wilks DS, Simeons P. 1996. Impacts of temperature and precipitation variability on crop model predictions. *Climatic Change* **32**: 293–311.

Ritchie JT, Singh U, Godwin DC, Bowen WT. 1998. Cereal growth, development and yield. In *Understanding Options for Agricultural Production*, Tsuji GY, Hoogenboom G, Thornton PK (eds). Kluwer Academic Publishers: Dordrecht, 79–98.

Robertson AW, Ines AVM, Hansen JW. 2007. Downscaling of seasonal precipitation for crop simulation. *Journal of Applied Meteorology and Climatology* **46**: 677–693.

Roeckner E, Arpe K, Bengtsson L, Claussen CM, Dümenil L, Esch M, Giorgetta M, Schiese U, Schulzweida U. 1996. The atmospheric general circulation model ECHAM-4: model description and simulation of present-day climate. Report No. 218, Max Planck Institute for Meteorology. Hamburg.

Schmidli J, Frei C, Vidale PL. 2006. Downscaling from GCM precipitation: a benchmark for dynamical and statistical downscaling methods. *International Journal of Climatology* **26**: 679–689.

Semenov MA, Doblas-Reyes FJ. 2007. Utility of dynamical seasonal forecasts in predicting crop yield. *Climate Research* **34**: 71–81.

Sharma D, Gupta AD, Babel MS. 2007. Spatial disaggregation of bias-corrected GCM precipitation for improved hydrologic simulation: Ping River Basin, Thailand. *Hydrology and Earth System Sciences* **11**: 1373–1390.

Stern RD, Coe R. 1984. A model fitting analysis of daily rainfall data. *Journal of Royal Statistical Society A* **147**: 1–34.

Tsuji GT, Uehara G, Salas S. (eds). 1994. *DSSATv3.0, Vol. 3*. University of Hawaii: Honolulu, Hawaii, p. 286.

Wilks DS. 1995. *Statistical Methods in the Atmospheric Sciences*, Academic Press: San Diego.

Wilks DS. 1999. Interannual variability and extreme-value characteristics of several stochastic daily precipitation models. *Agricultural and Forest Meteorology* **93**: 153–169.

Wilks DS, Wilby RL. 1999. The weather generation game: a review of stochastic weather models. *Progress in Physical Geography* **23**: 329–357.

Willmott CJ. 1982. Some comments on the evaluation of model performance. *Bulletin of the American Meteorological Society* **63**: 1309–1313.

Woolhiser DA, Roldán J. 1982. Stochastic daily precipitation models. 2. A comparison of distributions of amounts. *Water Resources Research* **18**: 1461–1468.