Hydrol. Earth Syst. Sci. Discuss., 10, 6847–6896, 2013
www.hydrol-earth-syst-sci-discuss.net/10/6847/2013/
doi:10.5194/hessd-10-6847-2013

Hydrology and Earth System Sciences
Open Access
Discussions

Estimating monthly rainfall in rural river basins under climate change: an improved bias-correcting statistical downscaling approach

D. L. Jayasekera[1] and J. J. Kaluarachchi[2]

[1]Civil and Environmental Engineering Department Utah State University Logan, Utah, 84322-4100, USA
[2]College of Engineering Utah State University Logan, Utah, 84322-4100, USA

Received: 3 May 2013 – Accepted: 12 May 2013 – Published: 3 June 2013

Correspondence to: D. L. Jayasekera (d.jayasekera@aggiemail.usu.edu)

Published by Copernicus Publications on behalf of the European Geosciences Union.

Abstract

This study extended the work of Kim et al. (2008) to generate future rainfall under climate change using a discrete-time/space Markov chain based on historical conditional probabilities. A bias-correction method is proposed by fitting suitable statistical distributions to transform rainfall from the general circulation model (GCM) scale to watershed scale. The demonstration example used the Nam Ngum River Basin (NNRB) in Laos which is a rural river basin with high potential for hydropower generation and significant rain-fed agriculture supporting rural livelihoods. This work generated weekly rainfall for a 100 yr period using historical rainfall data from 1961 to 2000 for ten selected weather stations. The bias-correction method showed the ability to reduce bias of the mean values of GCMs when compared to the observed mean amount at each station. The simulated rainfall series is perturbed using the delta change estimated at each station to project future rainfall for the Special Report on Emission Scenarios (SRES) A2. GCMs consisting of third generation coupled general circulation model (CGCM3.1 T63) and European center Hamburg model (ECHAM5) projected an increasing trend of mean annual rainfall in the NNRB. Seasonal rainfall percent changes showed an increase in the wet and dry seasons with the highest increase in the dry season mean rainfall of about 31 % from 2051 to 2090. While the GCM projections showed good results with appropriate bias corrections, the Providing REgional Climates for Impacts Studies (PRECIS) regional climate model significantly underestimated historical behavior and produced higher mean absolute errors compared to the corresponding GCM predictions.

1 Introduction

A key challenge in water resources planning and management is to estimate the water availability and to adopt management strategies in the presence of climate change. Intergovernmental Panel on Climate Change (IPCC) fourth assessment report AR4

HESSD
10, 6847–6896, 2013

(2007a) defined climate change as “a change in the state of the climate that can be identified by changes using statistical tests in the mean and/or the variability of its properties and that persists for an extended period, typically decades or longer. The change refers to any change in climate over time, whether due to natural variability or as a result of human activity”. Southeast Asia is one such a region vulnerable to climate change and its variability, including rise in sea level, shifts of climatic zones, and the occurrence of extreme events such as droughts and floods (UNFCCC, 2007).

General Circulation Models (GCMs) are used to project future climates under different greenhouse gas emission scenarios (IPCC, 2007b). The key limitation of GCM simulations is the coarse scale grid resolution that prevents the direct use of GCMs for impact assessment studies because GCM results cannot represent sub grid-scale features and dynamics at the watershed scale (IPCC, 2007b; Vicuna et al., 2007). GCMs supported by appropriate downscaling techniques, have long been used to simulate changes in regional climate systems over wide spatiotemporal scales and to allow information from coarser-scale atmospheric simulations to be used in watershed-scale hydrologic models (Wilby and Wigley, 1997; Arnell et al., 2003). There have been many studies and different downscaling techniques developed to transfer the coarse scale GCM output to regional scales. The most common downscaling techniques are (a) dynamical downscaling that uses regional climate models (RCMs) to simulate watershed-scale physical processes (Giorgi et al., 2001; Mearns et al., 2004; Fowler et al., 2007); and (b) statistical downscaling using statistical relationships between the regional climatic conditions and pre-identified large-scale atmospheric parameters (Wilby et al., 2004; Mehrotra and Sharma, 2005; Vrac and Naveau, 2007). Over the past years, a wide range of statistical downscaling techniques have been developed and most techniques fall into a category where response variables (mostly rainfall) are related to a discrete or continuous state, which is modeled as a function of the atmospheric and local-scale predictor variables (Wilby and Wigley, 1997; Stehlík and Bárdossy, 2002; Mehrotra and Sharma, 2005; Vrac and Naveau, 2007; Mehrotra and Sharma, 2010). The limitations and assumptions of both techniques contribute to the uncertainty of results (Fowler et al., 2007). Studies by IPCC (2007b) and Fowler et al. (2007) provided a good discussion of various downscaling techniques. In general, raw GCM rainfall amounts tend to underestimate year-to-year variability and poorly represent extreme events, when compared to the historical rainfall records (Ines and Hansen, 2006; Knutti, 2008), implying that the probability of sustained droughts/low flows or high flows are poorly predicted in future climate projections. This limitation will have a significant impact in water resource planning and management. There is a need to address this limitation and correct the bias of raw GCM outputs for appropriate use in hydrologic modeling.

A commonly used approach in recent studies is the use of change factors (CFs) (Abbaspour et al., 2009; van Roosmalen et al., 2009; Sulis et al., 2011), often called the "perturbation method" (Prudhomme et al., 2002) or "delta change" approach which assumes that the climate model represents relative change more accurately than the absolute climate values and the model bias is constant through time (Fowler et al., 2007). Generally, the CFs are applied to perturb the historical observed time series. A study by Kim et al. (2008) investigated the long-term changes of rainfall by extending the historical rainfall series at multiple sites preserving the historical temporal and spatial correlation structures. The conventional approach is to use the mean of raw GCM grid values over space and few studies investigated the long-term changes of rainfall that persists for an extended period, typically decades or longer. Kim et al. (2008) generated future monthly rainfall series and perturbed the series by the percent change of mean monthly rainfall at the grid nodes of GCMs spatially downscaled to weather stations. However, the question still remains whether the percent change at observed scale (i.e. at a given weather station) is similar to the percent change at the interpolated GCM scale given the different spatial scales. To compare the rainfall changes at a local weather station the coarse scale distribution from the GCM scale needs to be transformed to the observed scale of distribution by using its probability of occurrence. Another limitation of the previous study was its inability to reproduce the months with

HESSD
10, 6847–6896, 2013

zero rainfall (or dry states), because monthly time scale is not adequate to include the rainfall non-occurrence (dry-state) condition.

The need and prior applications of bias-correction methods have been discussed in the recent literature. Johnson and Sharma (2012) developed a nested model for bias correction at multiple time scales. Johnson and Sharma (2011) discussed that bias correction can be performed using parametric and nonparametric approaches. Li et al. (2010) proposed an equidistant quintile matching technique of bias correction for monthly precipitation and temperature using IPCC AR4 models. Fowler et al. (2005), Frei et al. (2006), Christensen et al. (2007), and Schmidli et al. (2007) assessed the ability of RCMs to reproduce credible climate change scenarios for extreme events and climate variability at a regional scale. Fowler et al. (2007) suggested that at least for present-day climates, dynamical downscaling methods provide little advantage over the statistical techniques. Kerr (2013) stated that regional models should be tested to evaluate whether the model outputs are capable of regional scale modeling compared to the use of global models. However, few studies focused on South East Asia to assess the rainfall and temperature changes due to climate change. A study by Lacombe et al. (2013) projected the rainfall and temperature trends of South East Asia. Västilä et al. (2010) simulated the climate change impacts in the Lower Mekong flood plains using re-scaled PRECIS RCM for baseline scenarios. Eastham et al. (2008) used a statistical analysis to quantify the relative ability of each GCM in simulating climate over the Mekong Basin using 24 GCMs used in the AR4 report. Some of the GCMs in the AR4 report showed considerable capability at sub-continental scales even when assessed using daily frequency distributions. This builds confidence in using the GCMs for regional assessment (Perkins et al., 2007) and in some cases for assessing extreme events.

The long-term variability of seasonal and sub-seasonal (e.g. monthly) streamflow is important especially for river basins where primary livelihood is based on rain-fed agriculture. Moreover, a prior understanding of spatial and temporal variability of streamflow is crucial for the sustainability of rural economies especially in a region where

HESSD
10, 6847–6896, 2013

D. L. Jayasekera and J. J. Kaluarachchi

hydropower generation is important. Since streamflow is directly influenced by the rainfall distribution, the estimation of temporal and spatial variability of rainfall is an important consideration. Here we propose to study the influence of climate change on rainfall in rural river basins with limited data using the Nam Ngum River Basin (NNRB) in Laos. The region is well suited for this study because it is undergoing rapid development due to high hydropower generation capacity and population increase while rain-fed agriculture is still a priority. However, developing a reliable rainfall analysis in a rural river basin can be a challenge due to short and missing rainfall records, and limited hydrologic information. First, all of the above discussions identify the need to quantify the rainfall distribution in river basins that are vulnerable to extreme weather conditions. Second, a long term rainfall analysis for any temporal resolution should be able to preserve the temporal and spatial statistics and correlations of historical data so that the projected rainfall distribution is reliable.

The goal of this study is to improve the existing methodology to better project rainfall under climate change. The important considerations are bias correction, limited data availability, and applicability of RCMs. This is an extension of the work proposed by Kim et al. (2008). In the proposed work, the previous stochastic framework is extended for single- and multi-sites that simulates historical rainfall amounts (wet states) at individual locations using a discrete-time/space Markov chain based on historical conditional probabilities. Thereafter, the raw GCM rainfall amount is corrected using statistical bias correction of mean at each station. The proposed methodology is demonstrated for the NNRB to predict the long-term rainfall distribution for two time periods, 2011–2050 and 2051–2090.

2 Description of Nam Ngum River Basin, Laos

2.1 Physical description

The NNRB which originates from the Tran Ninh Plateau, 1000 to 1500 m a.m.s.l. (above mean sea level), is located in Northern Laos (Fig. 1). The drainage area of NNRB at the main outlet close to the confluence with the Mekong River is 16 777 km^2 or 7.3 % of the national area. The elevation of NNRB varies from 6 to 2684 m above m.s.l. The estimated mean slope of the NNRB basin is about 25.5 %. It is the second largest river basin in terms of mean annual flow and population compared to the Sekong and Sebanghieng Basins and the fifth largest basin in terms of land area in the country. The estimated population of the basin is 502 150 in 2005 and this number is approximately 9 % of the population of Laos (WREA, 2008). The major land use types of NNRB are natural forest at 47 %, shrub land at 34 %, agriculture at 8 %, grassland at 7 %, water surface at 3.98 %, and urban area at 0.02 % of the total area (WREA, 2008, 2009).

2.2 Climate and hydrology

The climate of NNRB is subtropical to tropical with a distinct wet season from May to October and mostly dry during the rest of the year. Most of the rainfall in the NNRB is due to the arrival of warm moist air during the south-west monsoon period. The hottest months are March to April during which the mean daily maximum temperature varies between from 28 to 34 °C. The mean minimum daily temperature varies between 14 and 24 °C between December and January at high elevations (ADB, 2008). The mean annual rainfall of NNRB is 2000 mm, varying between 1400 to more than 3500 mm (WREA, 2009). The mean annual Penman–Monteith potential evapotranspiration varies between 1060 and 1360 mm (ADB, 2008).

The mean annual flow to the Mekong River is about 22 billion m^3 (BCM) which is about 14.4 % of the annual flow of the Mekong River. The annual water use of NNRB

HESSD
10, 6847–6896, 2013

D. L. Jayasekera and J. J. Kaluarachchi

is about 0.9 BCM of which 99 % is used by agriculture, 0.52 % is by urban water use, and 0.08 % is for industrial purposes.

3 Rainfall data

3.1 Data sources

There are 40 weather stations available in and around the NNRB (Fig. 1). Daily rainfall data are available for all 40 stations for different periods and the longest daily rainfall data are available at Vientiane from 1951 to 2000. Except for few weather stations, most other stations have missing rainfall records from 1961 to 2000. Luang Prabang, Nong Khai, Xiengkhouang, and Vientiane have daily rainfall records available for 40 yr from 1961 to 2000 and other stations have daily rainfall records varying from 7 to 38 yr for the same period. The period from 1961 to 2000 is comparable to the 20th century experiment (20C3M) period or the baseline period.

The density of weather stations of the study area is low especially in the eastern and north eastern parts of the study area and amounts to about one station per 2100 km^2. Since most weather stations have varying periods of missing rainfall data, it can be challenging to select the weather stations which are representative of rainfall characteristics of the basin.

3.2 Selection of representative stations

The purpose here is to identify the weather stations from the 40 available that can represent the rainfall pattern of the basin. A non-parametric bootstrap method and the Thiessen polygon spatial interpolation method were used to evaluate the uncertainty in the selection of representative weather stations. To develop a stochastic methodology to generate reliable rainfall data, there should be long historical observed rainfall of at least 30 to 40 yr. Of the 40 weather stations, Luang Prabang, Nong Khai, Xiengkhouang, and Vientiane have 40 yr of rainfall data from 1961 to 2000 whereas

HESSD
10, 6847–6896, 2013

Ban Nasone, Thangone, Sengkhalok, Ban Hinheup, Ban Thouei, and Phonhong have rainfall data for 38, 36, 35, 34, 32, and 30 yr, respectively. Remaining 30 stations have rainfall records for 7 to 38 yr for the same period (Fig. 1). Subsets of weather stations were selected randomly based on the availability of historical data. A subset of 10 weather stations from random sampling was selected for the bootstrapping uncertainty analysis, and to estimate areal mean annual rainfall and arithmetic mean annual rainfall (Fig. 2). The selected 10 representative stations are Luang Prabang, Nong Khai, Xiengkhouang, Vientiane, Ban Nasone, Thangone, Sengkhalok, Ban Hinheup, Ban Thouei, and Phonhong. A non-parametric bootstrap random resampling technique (with a dimension of 1000) was used to evaluate if the selected 10 weather stations can be used to represent historical data both spatially and quantitatively. Figure 2 shows the mean annual rainfall estimated by resampling 10 to 40 stations among the 40 stations. The estimated mean areal annual rainfall using the 10 stations and the arithmetic mean are within the 95 % confidence limit.

3.3 Missing data

A study by Teegavarapu and Chandramouli (2005) provided a detailed discussion of different techniques for the estimation of missing rainfall records. They recommended that coefficient of correlation weighting method is one of the methods conceptually superior to other approaches due to its capability to ensure the existence of spatial autocorrelation in estimating the missing data. This study used coefficient of correlation to estimate the missing rainfall data. A previous study by Kim et al. (2008) showed that a combination of local linear regression and coefficient of correlation methods is good for estimating missing rainfall data. As shown in Table 1, the 10 representative stations selected earlier are significantly correlated (p values ≈ 0) among each other indicating that coefficient of correlation weighting method is suitable for the estimation of missing data for the historical period. The filling is performed at weekly time steps at the 10 representative stations such that complete records are produced from 1961 to 2000. The

HESSD
10, 6847–6896, 2013

total mean annual rainfall changed approximately 2 % from the observed values and the correlation coefficients among stations (not shown here) remained almost similar.

4 Methodology

This work simulates historical weekly rainfall non-occurrence (dry state, 0 mm) and occurrence (wet state, > 0 mm) at individual locations using a discrete-time/space Markov chain based on conditional probabilities. A weekly time step is selected because it can better simulate both rainfall occurrence and non-occurrence compared to monthly time step. The spatial correlations in the simulated amounts are generated using spatially correlated yet serially independent random numbers. The methodology includes the following steps: (a) first representative rainfall stations are selected based on the availability of daily rainfall data to represent the baseline period from 1961 to 2000, (b) a single station (or key station) is selected among the 10 representative stations while for temporal generation preserving the temporal correlation structure, (c) the remaining representative stations are used for spatial generation preserving the spatial correlation, (d) bias correction is performed for the baseline period (20C3M) and future A2 emission scenario using the historical observed and generated rainfall amounts, and (e) perturbation conducted at each station by the CF method to project the future precipitation amounts.

4.1 Single site temporal generation

A correlation analysis was performed to identify a key station that has the highest correlation of annual rainfall with the annual unregulated streamflow. Unregulated streamflow stations are located at Muang Kasi, Vangvieng, Ban Naluang, and the proposed dam site location (Fig. 1). The weather station located at Luang Prabang (Fig. 1) has the highest correlation ($r = 0.99$, p value ≈ 0) with streamflow measured at the proposed dam site location. Therefore, Luang Prabang was selected as the key representative station. Since Luang Prabang is located outside the basin, the rainfall amounts do not physically contribute to the flow at the proposed dam site but the annual rainfall pattern is highly correlated with the annual unregulated streamflow at the proposed dam site (see Fig. 1). Long-term weekly (temporal) rainfall is generated using a discrete-time Markov chain based on conditional probabilities at Luang Prabang and the long-term weekly (spatial) rainfall of remaining nine stations.

4.2 Markov process

Markov process is a special type of stochastic process defined as a family of random variables $\{X(t), t \in T\}$ where t represents time and T is the index set or parameter space that is a subset of $(0, +\infty)$. The values assumed by the random variables $X(t)$ are called states. A special case of this for a discrete-time/discrete-valued (DTDV) random process is called a Markov chain. Specifically, it has the property that the probability of the random process $X[n]$ at time $n = n_0$ only depends upon the outcome or realization of the random process at the previous time $n = n_0 - 1$. This work deals with weekly rainfall amounts at the key rainfall station Luang Prabang.

The conditional probability (P_{ij}) of state i of the current week (w), given the state of the previous week $(w-1)j$, can be written as

$$P_{ij} = \Pr\left[X'(w) \in i | X'(w-1) \in j\right]$$
$$P_{ij} = \frac{\Pr\left[(X'(w) \in i) \cap (X'(w-1) \in j)\right]}{\Pr[X'(w-1) \in j]} \tag{1}$$

where i and j represent current and previous states from 1 to N and N is the number of states corresponding to the standardized weekly rainfall X'. For example, state 1 is defined as 0, state 2 as $0 \leq X' < 1$, state 3 as $1 \leq X' < 2$ and so on. N depends on the range of weekly rainfall data. In this study, N was computed by dividing the range of weekly rainfall by its standard deviation across the historical 40 yr period of 1961 to 2000.

HESSD
10, 6847–6896, 2013

If the chain is previously in state S_j, then it moves to the current state S_i with the probability denoted by P_{ij}, and these probabilities are called conditional or transition probabilities. The conditional or transition probability matrix $\mathbf{P}_\mathrm{w}$ for the current week can be constructed where the elements of $\mathbf{P}_\mathrm{w}$ satisfy the following two properties; $0 \le P_{ij}(w) \le 1$, and $\sum \mathbf{P}_\mathrm{w} = 1$. By using the historical weekly rainfall series for the key station, $\mathbf{P}_\mathrm{w}$ can be computed.

A set of conditioned random numbers is required from a continuous uniform distribution to successively generate a time series of weekly standardized rainfall. Consider the states from 1 to N where each state has a specific probability density. Two sets of discrete uniform random number series from 1 to N are generated and conditioned (i.e. increase or decrease) for each state using a given marginal or conditional probability density. A set of conditioned discrete uniform random numbers $C(1,N)$, can be generated as

$$C(1,N) = \left\{[1]^{d_1}, [2]^{d_2}, \ldots [i]^{d_i}, \ldots [N]^{d} N\right\} \tag{2}$$

where $[i]^{d_i}$ is the set of integer i which represents the state which has a dimension of d_i and d is the dimension of the conditioned discrete uniform random number matrix to be generated which is 1000 in this study. For each week, a series of discrete uniform random numbers were generated. It is considered that monthly values are represented over 4 weeks and the time series were generated for 100 equivalent annual periods consisting of 48 weeks each.

To generate states for week 1, the previous state j (i.e. j-th column of $\mathbf{P}_1$) is decided first from $C(1,N)$ conditioned by the marginal probability of week 48, $\Pr[X(48) \in j]$ where 48 is week 48 which is the previous week. Likewise, current state i can be decided from $C(1,N)$ conditioned by $\mathbf{P}_1$ for a given j. In the same manner, the current state (week 2) i is used to decide the previous state j of week 1 so that the j-th column of $\mathbf{P}_2$ is used to decide the state of week 2, and so on. This process is continued and performed for a time length of 100 yr.

HESSD
10, 6847–6896, 2013

After the generation of continuous uniform random numbers (e.g. 0 to 0.99 for state 2) conditioned by the previous state conditional probabilities, these random numbers need to be restored to its real weekly rainfall amounts, X, in mm by multiplying by the corresponding weekly standard deviations. Here, we assume the historical long-term weekly standard deviations will remain unchanged in future climatic conditions. By considering the conditional probabilities of historical states transitions and randomly generating the amount of rainfall within the range of a particular state of a given month, the discrete-time Markov chain stochastic process can simultaneously address temporal characteristics of historical data between successive weeks and randomness of weekly rainfall. Additional information is available from Kim et al. (2008).

4.3 Multi site spatial generation

For multi-site weekly rainfall generation, the temporal generation used in the single site scenario is extended between the key station and the representative stations except spatially to preserve the spatial correlation structure. For the historical period, as shown in Table 1, the key station Luang Prabang is highly correlated with the other representative nine rainfall stations with correlation coefficients of 0.75 with Station 2 (Vientiane) and 0.95 with Station 7 (Phonhong), and p values close to zero. These statistics indicate that the mean weekly rainfall of the nine stations is closely correlated with the key station, therefore this relationship of conditional probability similar to the temporal condition probability (P_{ij}^{k}) can be written as,

$$P_{ij}^{k} = \Pr\left[X'(w,k) \in i | X'(w,k') \in j\right]$$
$$P_{ij}^{k} = \frac{\Pr\left[(X'(w,k) \in i) \cap (X'(w,k') \in j)\right]}{\Pr[X'(w,k') \in j]} \tag{3}$$

where k is now the target station number, k' is the key station, and other notations are same as given in Eq. (1). As the series of state j for the key station was already generated in the earlier single site temporal generation, state i in the target station

HESSD
10, 6847–6896, 2013

can be iteratively generated using $C(1,N)$ conditioned by $\mathbf{P}_{\mathrm{w}}^{k}$ at the given j of the key station, where $\mathbf{P}_{\mathrm{w}}^{k}$ is the matrix of $\mathbf{P}_{ij}^{k}$ at the current week w. The overall process of generating rainfall for multi-sites is similar to the single site rainfall generation except using the target station number instead of the week.

4.4 GCM and emission scenario

This study used A2 emission scenario which is the most common scenario for mid and high ranges of emissions used in recent climate change impact studies (Abbaspour et al., 2009; van Roosmalen et al., 2009; Anandhi et al., 2011; Sulis et al., 2011), and for South East Asia by Lacombe et al. (2013). The A2 scenario emphasizes on local traditions, high population growth, and less concerns from rapid economic development. Also from an assessment view point, A2 scenario provides probably the worst case scenario for a country such as Laos that is rapidly undergoing development. Eastham et al. (2008) conducted a statistical analysis to quantify the relative ability of each model to simulate climate over the Mekong River Basin using 24 different GCMs. Based on the pattern correlation and root mean square error of temporal and spatial pattern representation of monthly and seasonal rainfall over the Mekong Basin, the authors selected 11 GCMs. In this study, CGCM3.1 T63 (http://www.ec.gc.ca/ccmac-cccma/default.asp?lang=En&n=1299529F-1, accessed March 2012) and ECHAM5 (http://www.mpimet.mpg.de/en/science/models/echam/echam5.html, accessed March 2012) were selected based on the ability to represent the temporal and spatial patterns of rainfall over the Mekong Basin. In addition, a RCM known as PRECIS (Providing REgional Climates for Impacts Studies) developed by the Hadley Center for Climate Change in UK is available for comparison with projections made by other GCMs (http://www.metoffice.gov.uk/precis, accessed March 2012). RCM simulations for the NNRB were conducted by the South East Asia Regional Center (START) in Thailand. (http://www.start.or.th/, accessed March 2012).

HESSD
10, 6847–6896, 2013

Monthly rainfall fluxes for the baseline scenario (20C3M) period and for the future period for A2 scenario (2011–2090) were downloaded from the IPCC Data Distribution Center (DDC) (http://www.mad.zmaw.de/IPCC_DDC/html/SRES_AR4/index.html, accessed December 2010) for CGCM3.1 T63 and ECHAM5. The spatial resolution and the number of GCM grids covering the study area is shown in Table 2. Monthly rainfall amounts from PRECIS were available from the South East Asia-**S**ys**T**em for **A**nalysis, **R**esearch and **T**raining (SEA-START) Center in Thailand for the control period from 1960 to 2004 and for the A2 scenario from 2010 to 2050.

Several methods have been proposed by IPCC (2007a) to apply the GCM outcomes to a small study area. The simplest application is the direct use of the raw GCM grid information to the nearest station in the study area. The main weakness of this method is that rainfall stations located close proximity but falling in different GCM grids, while having similar climatic conditions and characteristics, tend to assign different climatic conditions (Kim et al., 2008). As shown in Table 2, six to nine GCM grids are needed to cover the NNRB with 10 stations whereas 108 RCM grids are needed to cover the same NNRB. The monthly rainfall amounts at each GCM and RCM grid nodes were spatially downscaled to the 10 stations for the baseline scenario and A2 scenario periods. The inverse distance weighted method was used for spatial interpolation.

Regional climate change signals can be significantly different from those projected by GCMs, particularly in regions with complex orography. Normally, RCMs dynamically downscale the climate change signals projected by GCMs. A RCM is driven by sea surface temperatures and atmospheric lateral boundary values from the forcing GCM (Déqué et al., 2005). RCMs are known to better capture the effects of orographic forcing and provide improved simulation of higher moment climate statistics; hence providing more plausible climate change scenarios for extreme events and climate variability at the regional scale. Despite these improvements, there is a need (Leung et al., 2003) for more research examining the statistical structure of climate signals at different spatial scales to establish whether RCMs can accurately predict regional-scale climate.

HESSD
10, 6847–6896, 2013

D. L. Jayasekera and J. J. Kaluarachchi

4.5 Bias correction

Kim et al. (2008) compared the generated monthly rainfall series with the coarse scale monthly GCM rainfall amounts and assigned weights for each GCM based on observed accuracy. Comparing amounts from two rainfall series at two different spatial scales is not intuitively correct. Because, the resulting rainfall amounts at a weather station due to regional atmospheric conditions may be different from the atmospheric conditions occurring at the GCM scale. Further, Kim et al. (2008) projected future monthly rainfall amounts by perturbing the generated monthly rainfall series at each location by interpolated percent change of rainfall using the values at the GCM node. The actual change of rainfall at a weather station at regional scale may be different compared to the spatially interpolated change using the GCM nodal percent change values. Therefore, we proposed a bias correction approach to transform GCM signals to the regional scale and to find the delta change at regional scale (at each weather location).

As shown in Fig. 3, a comparison of raw mean monthly rainfall for the baseline scenario (20C3M) with historical observed at the 10 weather stations suggests that the observed and raw GCM mean rainfall amounts are biased and underestimating the historical climatic conditions. A comparison of monthly rainfall is performed here due to the unavailability of daily ECHAM5 rainfall fluxes for the baseline period and A2 emission scenario. A given downscaling method should be able to capture the variability of rainfall at a location. Moreover, the performance of downscaling methods varies across seasons, locations, GCMs, and regional features such as orography, proximity to sea, land use, and vegetation. Therefore, we assume that at a given location, the local climatic effects are reflected by its rainfall distribution. Since the variability of rainfall at a location depends on the amount, statistical properties of GCM values such as mean should be corrected to the statistical properties of observed values at weather stations.

The purpose of bias correction is to reduce the bias between the GCM rainfall amount of the baseline scenario and the historical observed rainfall amount at a given station. Transformation of rainfall distributions from coarse GCM scale to regional

HESSD
10, 6847–6896, 2013

(observed) scale is conducted using the best-fitted parametric probability distribution. First, monthly baseline scenario (20C3M) GCM distribution and the observed historical (1961–2000) distributions are fitted to appropriate parametric probability distributions. For this purpose Exponential, Gamma, Weibull, and Log-normal distributions were considered and the best-fitting distributions are selected using histogram fits, quantile–quantile (Q–Q) plots, and correlation coefficients. Second, using the fitted parametric probability distribution of the GCM baseline rainfall as well as for the historical observed the corresponding cumulative probability function is computed. Third, the cumulative probabilities of the GCM distributions are then used with the fitted parametric probability distribution parameters of historical observed data to estimate the corrected GCM rainfall amounts. The same procedure is repeated for generated precipitation amounts and for future GCM distributions across two time periods, 2011 to 2050 and 2051 to 2090.

Since ECHAM5 data are available as monthly rainfall fluxes, monthly time scale is used for the distribution fitting. Parametric probability distributions are fitted for monthly rainfall amounts greater than zero.

For example, if monthly rainfall amounts follow the Gamma distribution, the probability density function, $f(x,\alpha,\beta)$ is

$$f(x;\alpha,\beta) = \frac{1}{\beta^{\alpha}\Gamma(\alpha)} x^{\alpha-1} \exp\left(-\frac{x}{\beta}\right) \quad \text{for} \quad x > 0 \tag{4}$$

where x is the monthly rainfall amount, and α and β are shape and scale parameters. It should be noted that the shape and scale parameters are station dependent. The shape and scale parameters can be determined using Maximum Likelihood Estimation. The cumulative distribution of the above probability densities can be written as,

$$F(x;\alpha,\beta) = \int_0^x f(t)\mathrm{d}t \tag{5a}$$

$$F_{\mathrm{GCM}}(x_{\mathrm{GCM}};\alpha,\beta|_{\mathrm{GCM}}) \Rightarrow F_{\mathrm{His}}(x_{\mathrm{His}};\alpha,\beta|_{\mathrm{His}}) \tag{5b}$$

HESSD
10, 6847–6896, 2013

where x_{GCM} is the monthly rainfall of the GCM, and x_{His} is the historical observed/generated monthly rainfall. The corrected GCM rainfall for a given month can be estimated by taking the inverse of Eq. (5b)

$$x'_{\text{GCM}} = F^{-1}_{\text{GCM}}\left\{F_{\text{His}}(x_{\text{His}};\alpha,\beta|_{\text{His}})\right\}. \tag{6}$$

The statistical bias-correction method is applied to the rainfall amounts of CGCM3.1 T63 and ECHAM5 baseline scenarios. The inverse distance weighted method was used to estimate rainfall at the 10 stations. This interpolation method was selected because the GCM nodal rainfall amounts are greater than zero for both GCMs when precipitation flux is converted to monthly rainfall amounts.

4.6 Perturbation by CF method

In applying the CF method, it is assumed that the relative and/or absolute changes in rainfall between past and future climatic conditions have a strong physical basis and that rainfall recurrence patterns remain the same between the past and future periods (Akhtar et al., 2008; Kilsby et al., 2007). Therefore, the scaled and baseline scenarios differ only in terms of their respective means, maxima, and minima.

After correcting the raw GCM rainfall for the mean amount at a station, the CF is calculated using the corrected future GCM scenario ($\text{GCM}^{\text{f}}_{\text{corr}}$) and corrected GCM baseline scenario ($\text{GCM}^{\text{b}}_{\text{corr}}$) at monthly time steps at each weather station. The CF of rainfall at a given station is calculated as

$$\text{CF} = \frac{\text{GCM}^{\text{f}}_{\text{corr}}}{\text{GCM}^{\text{b}}_{\text{corr}}}. \tag{7}$$

The future rainfall (R^{f}_{G}) is estimated as

$$R^{\text{f}}_{\text{G}} = R^{\text{b}}_{\text{G}} \cdot \text{CF} \tag{8}$$

where R^{b}_{G} is the generated baseline rainfall.

HESSD
10, 6847–6896, 2013

5 Results and discussion

5.1 Rainfall generation and spatial correlation

The comparison of statistics between historical and generated weekly rainfall data at Luang Prabang (key station) and Nong Khai (furthest station) is shown in Fig. 4. The results show excellent agreement between the historical and generated mean of weekly rainfall data. Although not shown here, a similar excellent agreement of mean weekly rainfall amounts was observed with remaining eight representative stations as well. It was also found that the standard deviations of generated weekly rainfall data are satisfactorily reproduced and the estimated weekly maximum absolute error is 25 mm among the ten stations. The areal mean annual rainfall estimated using the generated values is about 3 % higher compared to the historical value of 1767.6 mm.

Since the multi-site discrete-space Markov chain included the dry state (as zero rainfall) conditional probabilities, it is important to compare the proportion of dry days for the historical period with the proportion of dry days of the generated rainfall series. Figure 5 shows these results of dry days (as 0 mm) at Luang Prabang and Nong Khai. It is noted that the discrete-time/space Markov chain was able to reproduce the historical rainfall patterns with exact proportions of dry weeks across all stations. At Luang Prabang, the average proportion of dry days during dry months (January–April and November–December) and wet months (May–October) for the historical observed period (1961–2000) are 0.58 and 0.07, respectively whereas for the generated 100 yr period, the values are 0.57 and 0.05, respectively. The average proportion of dry days at Nong Khai during dry months and wet months for the historical observed period are 0.35 and 0.06, respectively whereas for the generated 100 yr period it is 0.34 and 0.05, respectively. These statistics clearly shows that this conditional generation method was able to preserve the temporal and spatial correlation structures in terms of rainfall amounts as well as the proportion of dry days for the key station and the other representative stations.

5.2 Bias correction

In most cases, the best fitted distribution is Gamma and in some cases, Weibull and log-normal distributions were best fitted. In this study, statistical bias-correction was performed for both CGCM3.1 T63 and ECHAM5. For the sake of demonstration, the results of CGCM3.1 T63 results are shown in Fig. 6. It is noted that the mean monthly rainfall from raw GCM values are biased probably due to the difference in spatial scales of simulations whereas the observed rainfall distribution is influenced by region-specific climatic conditions. It can be stated that the standard deviations of historical and corrected GCM are similar and have improved compared to the raw GCM statistics. Figure 7 shows the coefficient of variation (CV) for the same results. It is seen that CV is similar between observed and corrected monthly rainfall amounts compared to the raw GCM amounts even if the means are different. Although not shown here, the index of agreement between the corrected GCM and historical values is close to 1 whereas there is poor agreement between the raw GCM and historical values. Figure 8 shows that the statistical bias-correction of raw GCM has reduced the monthly mean absolute errors at multi-sites for the baseline scenario (20C3M) from 1961 to 2000. These results indicate that the statistical bias-correction procedure is capable of preserving the historical statistics of rainfall.

Figure 9 shows the goodness-of-fit results for wet (June) and dry (January) months using the Kolmogorov–Smirnov (K–S) test. These results at Luang Prabang suggest that the difference between the two samples for observed versus bias corrected and observed versus raw GCM is not significant enough to state that they have different distributions at the 5 % significance level. Even though the distributions are not statistically different in the wet month of June, the maximum difference between the curves (k values) are lowest between observed and corrected ($k = 0.09$) as opposed to observed and raw GCM ($k = 0.23$) (Fig. 9a). But, the K–S test for dry month (January) suggests that the difference between the two samples for observed versus raw GCM are statistically significant to state that they are different distributions and the maximum

HESSD
10, 6847–6896, 2013

difference between the curves is high ($k = 0.54$) compared to observed versus corrected ($k = 0.12$). Although not shown here similar results were observed at other representative stations too. This goodness of fit test results suggest that the bias corrected monthly rainfall amounts match better with the observed rainfall amounts and follows the same distribution for a given weather station.

IPCC Report (2007b) states that the most appropriate method to assess the validity of a particular GCM is by examining the historical climatic conditions. The mean absolute errors were used to evaluate the relative accuracy of each GCM and RCM. As shown in Table 2, CGCM3.1 produced the lowest mean absolute error of 0.47 because it simulated both the total amount and the trend of areal monthly rainfall for the historical period with minimum error. ECHAM5 also showed a mean absolute error of 0.65 because of its relative good performance in simulating the trend. The RCM produced a highest mean absolute error of 87.43 indicating its relatively poor performance in simulating the trend compared to the corrected GCMs. Here, the bias correction procedure is not used on raw RCM data because the RCM used dynamical downscaling technique incorporating regional physical and atmospheric processes. A further discussion related to the RCM data will follow in the next sections.

5.3 Projected future rainfall distributions

The CF of rainfall mean at each station was used to perturb the generated baseline scenario to project future rainfall. The results of the perturbed series are given as the percent changes of mean monthly rainfall from the historical observed period in Table A1. The results show that there is a greater variation of percent changes of mean monthly rainfall in the dry season (November to April) compared to the wet season (May to October) at every station. The maximum increase of 112.2 % occurs at Station 4 (Banhinheup) in January whereas the maximum decrease of 88.3 % occurs at Station 9 (Thangone) in December from 2051 to 2090. The maximum variation of 173 % of mean monthly rainfall occurs between at Station 4 (Banhinheup) and Station 9 (Thangone) in January whereas the minimum variation of 44 % of mean monthly

HESSD
10, 6847–6896, 2013

rainfall occurs at Station 4 (Banhinheup) and Station 6 (Ban Thouei) in June from 2051 to 2090. Therefore, it clear that the highest variation of percent changes occur in the dry season and this change is highest in the latter 40 yr of the century.

It is noted from Table A1 that statistical bias-correction reduced the inter-model difference significantly. The maximum percent change between the two GCM projections is about 15 % at Station 7 (Phonhong) in November from 2011 to 2050 whereas the minimum percent change difference between the two GCM projections is almost zero at Station 10 (Luang Prabang) in August from 2051 to 2090.

Table 4 shows the comparison of areal mean monthly rainfall amounts estimated using the historical observed, GCM bias corrected baseline scenario, and RCM control for the period from 1961 to 2000. The perturbed rainfall series for the 10 stations were spatially averaged for each GCM. As shown in Table 4, each GCM shows different increases but with less inter-model difference for both monthly and annual rainfall. Both models projected an increase in the total annual rainfall. The results show that CGCM3.1 produced an increase of total annual rainfall of 12 and 13 %, and ECHAM5 produced a corresponding increase of 11 and 13 % for the time periods of 2011–2050 and 2051–2090, respectively.

The seasonal variation of rainfall is important information where variation of streamflow can occur due to the changes in rainfall. Wet season rainfall contributes 76 % whereas dry season rainfall contributes 24 % to the mean annual rainfall from 1961 to 2000. Both CGCM and ECHAM projected that the wet season rainfall contributes 75 % whereas the dry season rainfall contributes 25 % to the mean annual rainfall from 2011 to 2050 while the wet season contributes 72 % and the dry season contributes 28 % from 2051 to 2090. These statistics indicate that during the latter half of the century there will be an increase in mean dry season rainfall compared to the first half of the century. Figure 10 shows the spread and variations of annual, seasonal, and mean monthly rainfall for the baseline and A2 scenarios. Both future GCM projections show a minimum inter-model difference for annual and seasonal variations and show an increase in mean annual rainfall for both time periods. The wet season rainfall (Fig. 10c)

HESSD
10, 6847–6896, 2013

Estimating monthly rainfall in rural river basins

D. L. Jayasekera and J. J. Kaluarachchi

contributes significantly to the variation of mean annual rainfall (Fig. 10a) from each GCM. It is noticed that rainfall is distributed in a wide range in the dry season for both GCMs compared to the historical rainfall. The median rainfalls for both seasons have increased compared to the historical amounts. Figure 10b shows that the 25th percentile of dry season rainfall has decreased during 2051 to 2090 whereas it has increased in the wet season compared to the historical amounts.

Table 5 provides a quantitative comparison of statistics of projected areal rainfall for wet and dry seasons of the study area. The maximum and minimum percent change of mean annual rainfall is 14.7 and 8.7 % for CGCM3.1 and ECHAM5 scenarios, respectively, from 2011 to 2050. Both GCM projections are in agreement to show that maximum and minimum mean annual rainfalls will increase during the next 80 yr and the maximum and minimum mean annual will significantly increase during the second half of the century.

The variability of seasonal variation of rainfall is useful in long-term planning and management as it can affect agricultural activities, hydropower generation, and ecosystem functions. Table 5 shows that the maximum rainfall in the wet season will increase about 14 % from 2051 to 2090 and the minimum rainfall will decrease about 3 % according to the CGCM3.1 projections. Similarly, the maximum rainfall in the dry season will increase about 38 % from 2051 to 2090 and the minimum rainfall will decrease about 25 % according to the CGCM3.1 projections. Therefore, fluctuation of extremes rainfall events are highest during the dry season compared to the wet season.

The spatial distribution of percent changes of projected mean annual rainfall is shown in Fig. 11. Both GCMs projected an increasing trend of mean annual rainfall in the NNRB. The downscaled GCM mean annual rainfall shows that the northern and north eastern parts of the basin have the highest projected change of 14 to 17 % in the next 80 yr. The central region will have a change of 13 to 14 % in mean annual rainfall. The lowest percent change of mean annual rainfall is projected in the southern and south western parts of the basin. As shown in Table 5, the percent change of mean annual rainfall is about 12 and 13 % from 2011 to 2050 and from 2051 to 2090,

HESSD
10, 6847–6896, 2013

D. L. Jayasekera and J. J. Kaluarachchi

respectively, according to both CGCM and ECHAM projections. These areal averages of mean annual rainfall estimated using the downscaled GCMs and spatial interpolated percentage change results are in good agreement for these future time periods. The increasing trend of mean annual rainfall could help to improve hydropower generation.

5.4 Comparison with the RCM

PRECIS uses a dynamical downscaling approach for a wide range of GCM scenarios for which the lateral boundary conditions (LBCs) have been included. The rainfall output from PRECIS was derived using the ECHAM4 LBCs as initial data for downscaling. As shown on Fig. 12, the RCM model outputs underestimate the annual rainfall amounts at the selected 10 stations across all 40 yr. The observed and total annual rainfalls of PRECIS were compared to minimize the random effects. It is clear from Fig. 12 that the PRECIS results cannot be directly used for climate change impact studies even though the outputs are available at much finer spatial scales. Also, the PRECIS results at monthly time scale produces relatively higher mean absolute errors compared to bias-corrected ECHAM5 results (Fig. 13). Despite this discrepancy from PRECIS, Fig. 13 provides a comparative insight of bias correction of GCMs for the study of climate change. The results clearly show that the mean absolute error is highest during the wet season in most stations. The results of this work shows that results from RCMs may not be directly applicable at the regional-scale and may need bias correction. This comparison also shows that GCMs projections can be used after bias correction that produce minimal mean absolute errors especially during the wet season.

6 Summary and conclusions

The focus of this study is to develop an appropriate methodology to project future rainfall under climate change with limited data for rural river basins while preserving the historical temporal and spatial characteristics. The NNRB located in the Mekong River

Basin was selected to demonstrate the applicability of the methodology where rain-fed agriculture and hydropower generation are priority economic activities. Ten stations from 40 available weather stations were selected to represent the temporal and spatial characteristics of rainfall using a non-parametric bootstrapping technique. Missing rainfall data for the ten selected stations were filled using the coefficient of correlation weighting method to maintain a complete data record of 40 yr which is similar to the temporal domain of the IPCC AR4 baseline scenario from 1961 to 2000.

The proposed methodology simulated weekly rainfall non-occurrence (dry state) and occurrence (wet state) at ten selected locations by preserving the historical temporal and spatial correlation structures using a discrete-time/space Markov chain based on conditional probabilities. At each location, the stochastically generated weekly rainfall series which consists of dry states and wet states were aggregated to monthly temporal scale. GCM rainfall bias at each station was corrected by transforming the coarse scale rainfall distribution to the region specific rainfall distribution. The bias-correction was performed by fitting statistical distributions to GCM and regional scale (observed or generated) monthly rainfall amounts. The main assumptions are (a) the historical temporal and spatial correlation structures remain unchanged, and (b) the location specific regional climatic conditions are representative of its rainfall distributions.

The bias correction approach reduced the error of mean monthly rainfall, relative frequency, and intensity of raw GCM rainfall amounts at ten selected stations hence reduced the inter-model differences and spatial heterogeneity of rainfall CFs of GCMs. The CFs estimated using the corrected GCM scenarios were perturbed to generate 100 yr rainfall rainfall amounts. Both GCMs, ECHAM and CGCM, projected an increase in the mean annual rainfall in the next 80 yr. The highest percent changes of annual rainfall are about 15 % from CGCM for 2011 through 2050 and 12 % from ECHAM for 2051 through 2090, respectively. The results showed a highest rainfall increase in the dry season amounts to 31 % from 2051 to 2090. The spatial distribution of projected mean annual rainfall showed a significant increase in the north eastern part of the study area. The RCM, PRECIS, provides rainfall projections from 2011 to 2050 while rainfall from 2051 to 2090 is not available. The spatial distribution of mean annual rainfall projected from 2011 to 2050 using PRECIS showed the highest annual rainfall amounts in the south eastern part of the basin. A comparison of RCM areal mean annual rainfall estimates for the baseline scenario and for A2 scenario from 2011 to 2050 showed that there will be only 0.7 % increase. Compared to this 0.7 % increase of areal mean rainfall, the bias corrected CGCM and ECHAM rainfall estimates showed 12 and 11 % increase, respectively for the same time period.

It is a challenging task to assess the impacts of climate change in rural river basins where data and hydrologic information are limited. In the presence of these limitations, this study was able to use available data and information and demonstrate the applicability of the proposed methodology that projects reliable future rainfall patterns assuming that the historical correlation structure is preserved. In situations where climate models show noticeable bias in reproducing regional climate for the historical period, their capacity to represent future may be questionable. This study focused on bias-correction of raw GCM outputs even though the RCM outputs are available at much finer spatial scale. The methodology proposed in this study was able to minimize bias in reproducing regional climate for the historical (or baseline scenario) period. The estimated future rainfall amounts produced in this study can be easily used to investigate regional impacts due to climate change.

Acknowledgements. The authors like to acknowledge the financial support for travel provided by the International Water Management Institute and the support and assistance of Hoanh Chu-Thai, Vladimir Smakhtin, and Guillaume Lacombe and the IWMI-Laos staff. Valuable discussions with Ungtae Kim are highly appreciated. The authors also like to thank the staff of Mekong River Commission for providing hydrologic data for this research especially by Vanthong Khamloonvilayvong, Huon Rath, Somchay Sichanthalath, and the SEA-START Center in Thailand for providing PRECIS data. Funding for this work was partially supported by the Inland Northwest Research Alliance of the US Department of Energy.

References

Abbaspour, K. C., Faramarzi, M., Ghasemi, S. S., and Yang, H.: Assessing the impact of climate change on water resources in Iran, Water Resour. Res., 45, W10434, doi:10.1029/2008WR007615, 2009.

ADB – Asian Development Bank: Lao People's Democratic Republic: Preparing the Cumulative Impact Assessment for the Nam Ngum 3 Hydropower Project, Technical Assistance Consultant's Report, Project Number 40514, Financed by the Japan Special Fund) Prepared by Vattenfall Power Consultant AB in Association with Ramboll Natura AB and Earth Systems Lao, February 2008, Sweden, 2008.

Akhtar, M., Ahmad, N., and Booij, M. J.: The impact of climate change on the water resources of Hindukush–Karakorum–Himalaya region under different glacier coverage scenarios, J. Hydrol., 255, 148–163, doi:10.1016/j.jhydrol.2008.03.015, 2008.

Anandhi, A., Frei, A., Pierson, D. C., Schneiderman, E. M., Zion, M. S., Lounsbury, D., and Matonse, A. H.: Examination of change factor methodologies for climate change impact assessment, Water Resour. Res., 47, W03501, doi:10.1029/2010WR009104, 2011.

Arnell, N. W., Hudson, D. A., and Jones, R. G.: Climate change scenarios from a regional climate model: estimating change in runoff in southern Africa, J. Geophys. Res., 108, 4519, doi:10.1029/2002JD002782, 2003.

Christensen, J. H., Carter, T. R., Rummukainen, M., and Amanatidis, G.: Evaluating the performance and utility of regional climate models: the PRUDENCE project, Climatic Change, 81 (Suppl.), 1–6, 2007.

Déqué, M., Jones, R. G., Wild, M., Giorgi, F., Christensen, J. H., Hassell, D. C., Vidale, P. L., Rockel, B., Jacob, D., Kjellstrom, E., de Castro, M., Kucharski, F., and van den Hurk, B.: Global high resolution versus Limited Area Model climate change projections over Europe: quantifying confidence level from PRUDENCE results, Clim. Dynam., 25, 653–670, doi:10.1007/s00382-005-0052-1, 2005.

Eastham, J., Mpelasoka, F., Mainuddin, M., Ticehurst, C., Dyce, P., Hodgson, G., Ali, R., and Kirby, M.: Mekong River Basin Water Resources Assessment: Impacts Of Climate Change, CSIRO: Water for a Healthy Country National Research Flagship, Australia, 2008.

Efron, B. and Tibshirani, R. J.: An Introduction to the Bootstrap, Chapman and Hall, New York, 1993.

Fowler, H. J., Ekström, M., Kilsby, C. G., and Jones, P. D.: New estimates of future changes in extreme rainfall across the UK using regional climate model integrations: 1. Assessment of control climate, J. Hydrol., 300, 212–233, 2005.

Fowler, H. J., Blenkinsop, S., and Tebaldi, C.: Linking climate change modeling to impacts studies: recent advances in downscaling techniques for hydrological modeling, Int. J. Climatol., 27, 1547–1578, 2007.

Frei, C., Schöll, R., Fukutome, S., Schmidli, J., and Vidale, P. L.: Future change of rainfall extremes in Europe: an intercomparison of scenarios from regional climate models, J. Geophys. Res., 111, D06105, doi:10.1029/2005JD005965, 2006.

Giorgi, F., Hewitson, B., Christensen, J., Hulme, M., Von Storch, H., Whetton, P., Jones, R., Mearns, L., and Fu, C.: Regional climate information: evaluation and projections (chapt. 10), in: Climate Change 2001: The Scientific Basis, Contribution of Working Group I to the Third Assessment Report of the IPCC, edited by: Houghton, J. T., Ding, Y., Griggs, D. J., Noguer, M., van der Linden, P. J., and Xiaosu, D., Cambridge Univ. Press, Cambridge, 739–768, 2001.

Grotch, S. L. and MacCracken, M. C.: The use of general circulation models to predict regional climatic change, J. Climate, 4, 286–303, 1991.

Huth, R.: Potential of continental-scale circulation for the determination of local daily surface variables, Theor. Appl. Climatol., 56, 165–186, doi:10.1007/BF00866425, 1997.

Ines, A. V. M. and Hansen, J. W.: Bias correction of daily GCM rainfall for crop simulation studies, Agr. Forest Meteorol., 138, 44–53, 2006.

IPCC – Intergovernmental Panel on Climate Change: IPCC Special Report, Emissions Scenarios: A Special Report of Working Group III of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, UK, 2000.

IPCC – fourth assessment report (AR4): Climate Change 2007: Synthesis Report, Contribution of Working Groups I, II and III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, Core Writing Team, edited by: Pachauri, R. K. and Reisinger, A., IPCC, Geneva, Switzerland, 104 pp., 2007a.

IPCC – Inter Governmental Panel on Climate Change: Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, edited by: Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K. B., Tignor, M., and Miller, H. L., Cambridge Univ. Press, Cambridge, UK and New York, NY, USA, 2007b.

Johnson, F. and Sharma, A.: Accounting for interannual variability: a comparison of options for water resources climate change impact assessments, Water Resour. Res., 47, W04508, doi:10.1029/2010WR009272, 2011.
Johnson, F. and Sharma, A.: A nesting model for bias correction of variability at multiple time scales in general circulation model precipitation simulations, Water Resour. Res., 48, W01504, doi:10.1029/2011WR010464, 2012.
Kerr, R. A.: Forecasting regional climate change flunks its first test, Science, 339, p. 638, 2013.
Kilsby, C. G., Jones, P. D., Burton, A., Ford, A. C., Fowler, H. J., Harpham, C., James, P., Smith, A., and Wilby, R. L.: A daily weather generator for use in climate change studies, Environ. Modell. Softw., 22, 1705–1719, doi:10.1016/j.envsoft.2007.02.005, 2007.
Kim, U., Kaluarachchi, J. J., and Smakhtin, V. U.: Generation of monthly rainfall under climate change for the Upper Blue Nile River Basin, Ethiopia, J. Am. Water Resour. As., 44, 1231–1247, doi:10.1111/j.1752-1688.2008.00220.x, 2008.
Knutti, R.: Should we believe model predictions of future climate change?, Philos. T. Roy. Soc. Lond., 366, 4647–4664, 2008.
Lacombe, G., Smakhtin, V., and Hoanh, C. T.: Wetting tendency in the Central Mekong Basin consistent with climate change-induced atmospheric disturbances already observed in East Asia, Theor. Appl. Climatol. 111, 251–263, doi:10.1007/s00704-012-0654-6, 2013.
Leung, L. R., Qian, Y., Bian, X., and Hunt, A.: Hydroclimate of the western United States based on observations and regional climate simulation of 1981–2000, Part II: Mesoscale ENSO anomalies, J. Climate, 16, 1912–1928, 2003.
Li, H., Sheffield, J., and Wood, E. F.: Bias correction of monthly precipitation and temperature fields from Intergovernmental Panel on Climate Change AR4 models using equidistant quantile matching, J. Geophys. Res., 115, D10101, doi:10.1029/2009JD012882, 2010.
Mearns, L. O., Giorgi, F., Whetton, P., Pabon, D., Hulme, M., and Lal, M.: Guidelines for Use of Climate Scenarios Developed from Regional Climate Model Experiments, Tech. Rep., Data Distribution Center of the IPCC, Geneva„ 2004.
Mehrotra, R. and Sharma, A.: A nonparametric nonhomogeneous hidden Markov model for downscaling of multi-site daily rainfall occurrences, J. Geophys. Res., 110, D16108, doi:10.1029/2004JD005677, 2005.
Mehrotra, R. and Sharma, A.: Development and application of a multisite rainfall stochastic downscaling framework for climate change impact assessment, Water Resour. Res., 46, W07526, doi:10.1029/2009WR008423, 2010.

Perkins, S. E., Pitman, A. J., Holbrook, N. J., and McAneney, J.: Evaluation of the AR4 climate models' simulated daily maximum temperature, minimum temperature, and rainfall over Australia using probability density functions, J. Climate, 20, 4356–4376, doi:10.1175/JCLI4253.1, 2007.

Prudhomme, C., Reynard, N., and Crooks, S.: Downscaling of global climate models for flood frequency analysis: where are we now?, Hydrol. Process., 16, 1137–1150, doi:10.1002/hyp.1054, 2002.

Schmidli, J., Goodess, C. M., Frei, C., Haylock, M. R., Hundecha, Y., Ribalaygua, J., and Schmith, T.: Statistical and dynamical downscaling of rainfall: an evaluation and comparison of scenarios for the European Alps, J. Geophys. Res., 112, D04105, doi:10.1029/2005JD007026, 2007.

Stehlík, J. and Bárdossy, A.: Multivariate stochastic downscaling model for generating daily rainfall series based on atmospheric circulation, J. Hydrol., 256, 120–141, 2002.

Sulis, M., Paniconi, M. C., Rivard, C., Harvey, R., and Chaumont, D.: Assessment of climate change impacts at the catchment scale with a detailed hydrological model of surface-subsurface interactions and comparison with a land surface model, Water Resour. Res., 47, W01513, doi:10.1029/2010WR009167, 2011.

Teegavarapu, R. S. V. and Chandramouli, V.: Improved weighting methods, deterministic and stochastic data-driven models for estimation of missing rainfall records, J. Hydrol., 312, 191–206, 2005.

UNFCCC – United Nations Framework Convention on Climate Chang): Climate Change: Impacts, Vulnerabilities and Adaptation in Developing Countries, 20, Germany, 2007.

van Roosmalen, L., Sonnenborg, T. O., and Jensen, K. H.: Impact of climate and land use change on the hydrology of a large-scale agricultural catchment, Water Resour. Res., 45, W00A15, doi:10.1029/2007WR006760, 2009.

Västilä, K., Kummu, M., Sangmanee, C., and Chinvanno, S.: Modeling climate change impacts on the flood pulse in the Lower Mekong floodplains, J. Water Clim. Change, 1, 67–86, doi:10.2166/wcc.2010.008, 2010.

Vicuna, S., Maurer, E. P., Joyce, B., Dracup, J. A., and Purkey, D.: The sensitivity of California water resources to climate change scenarios, J. Am. Water Resour. As., 43, 482–498, doi:10.1111/j.1752-1688.2007.00038.x, 2007.

Vrac, M. and Naveau, P.: Stochastic downscaling of rainfall: from dry to heavy rainfalls, Water Resour. Res., 43, W07402, doi:10.1029/2006WR005308, 2007.

HESSD

10, 6847–6896, 2013

Estimating monthly rainfall in rural river basins

D. L. Jayasekera and J. J. Kaluarachchi

Wilby, R. L. and Wigley, T. M. L.: Downscaling general circulation model output: a review of methods and limitations, Prog. Phys. Geogr., 21, 530–548, doi:10.1177/030913339702100403, 1997.
Wilby, R. L., Charles, S. P., Zorita, E., Timbal, B., Whetton, P., and Mearns, L. O.: Guidelines for Use of Climate Scenarios Developed from Statistical Downscaling Methods, Tech. Rep., Data Distribution Center of the IPCC, Geneva, 2004.
WREA – Water Resources and Environment Administration: Nam Ngum River Basin Profile, Prepared Under the Nam Ngum River Basin Development Project, Supported by the Asian Development Bank (Loan 1933-LAO) and Agence Francaise de Development, Grant CLA 1048 01H, Laos, 2008.
WREA – Water Resources and Environment Administration: Nam Ngum River Basin Integrated Water Resources Management Plan, Nam Ngum River Basin Development Sector Project, Supported by Agence Française de Développement, Asian Development Bank, Laos, 2009.

HESSD
10, 6847–6896, 2013

Estimating monthly rainfall in rural river basins

D. L. Jayasekera and J. J. Kaluarachchi

Table 1. Correlation coefficient matrix of the 10 representative stations for the historical period (lower triangle) and generated weekly rainfall (upper triangle).

Station	1	2	3	4	5	6	7	8	9	10
1	1.00	0.94	0.80	0.92	0.95	0.94	0.81	0.93	0.82	0.80
2	0.98	1.00	0.78	0.94	0.97	0.94	0.77	0.90	0.79	0.73
3	0.82	0.81	1.00	0.80	0.77	0.85	0.71	0.83	0.65	0.75
4	0.93	0.96	0.82	1.00	0.95	0.93	0.81	0.90	0.78	0.75
5	0.95	0.97	0.81	0.96	1.00	0.93	0.80	0.90	0.80	0.74
6	0.93	0.94	0.88	0.94	0.94	1.00	0.79	0.91	0.78	0.78
7	0.84	0.82	0.73	0.83	0.81	0.81	1.00	0.79	0.97	0.92
8	0.92	0.91	0.90	0.90	0.91	0.94	0.80	1.00	0.78	0.82
9	0.83	0.81	0.69	0.82	0.82	0.80	0.98	0.78	1.00	0.90
10	0.78	0.75	0.77	0.77	0.76	0.79	0.95	0.82	0.93	1.00

HESSD
10, 6847–6896, 2013

Estimating monthly rainfall in rural river basins

D. L. Jayasekera and J. J. Kaluarachchi

Table 2. Description of the selected GCMs and PRECIS.

GCM/RCM[1]	Spatial resolution[2]	Number of grids[3]	Mean absolute error[4]
CGCM3.1_T63	2.79, 2.81	6	0.47
ECHAM5	1.865, 1.875	9	0.65
PRECIS	0.2, 0.2	108	81.43

[1] From the IPCC DDC. CCCMA_CGCM3.1_T63, Canadian Centre for Climate Modeling and Analysis (Third generation), ECHAM5, European Center Hamburg Model (5th generation), PRECIS_RCM, Providing REgional Climates for Impacts Studies, Regional Climate Model.

[2] Mean resolution of GCMs and RCM in latitudinal and longitudinal degrees.

[3] Number of grids covering the NNRB.

[4] Computed using areal monthly rainfall absolute error of each GCM obtained for its baseline scenario (from 1961 to 2000) compared to the historical observed value.

HESSD
10, 6847–6896, 2013

Table 3. Mean absolute relative error statistics from the Markov chain process.

Site	Mean	STD[1]	CV[2]	SK[3]	Range
Ban Hinheup	0.15	0.10	0.10	0.23	0.08
Ban Nasone	0.13	0.11	0.11	0.38	0.10
Ban Thouei	0.17	0.11	0.10	0.24	0.07
Nong Khai	0.16	0.11	0.11	0.19	0.08
Phonhong	0.15	0.10	0.11	0.22	0.09
Sengkhalok	0.15	0.10	0.10	0.17	0.09
Thangone	0.19	0.11	0.13	0.20	0.08
Vientiane	0.13	0.09	0.10	0.22	0.08
Xiengkhouang	0.15	0.10	0.09	0.19	0.10
Luang Prabang	0.14	0.10	0.11	0.17	0.09

[1] STD: standard deviation.
[2] CV: coefficient of variation.
[3] SK: skewness coefficient.

HESSD
10, 6847–6896, 2013

Table 4. Comparison of historical and generated areal mean monthly rainfall (in mm). The values in parentheses are percent changes from the baseline scenario.

Month	Historical	RCM	CGCM3.1			ECHAM5		
	(1961–2000)	(1961–2000)	20C3M Baseline Scenario	A2 2011–2050	A2 2051–2090	20C3M Baseline Scenario	A2 2011–2050	A2 2051–2090
Jan	53.9	4.1	53.9	57.2 (6)	82.0 (52)	53.7	57.8 (8)	82.4 (53)
Feb	53.8	6.4	53.8	64.2 (19)	65.7 (22)	53.6	63.9 (19)	66.0 (23)
Mar	75.3	18.1	75.3	79.0 (5)	88.8 (18)	76.0	78.3 (3)	89.2 (17)
Apr	117.7	55.1	117.5	145.5 (24)	148.9 (27)	119.7	144.2 (20)	144.7 (21)
May	226.2	132.1	225.8	269.8 (20)	233.9 (4)	226.5	268.8 (19)	234.3 (3)
Jun	257.5	322.6	257.2	264.4 (3)	252.2 (−2)	257.3	264.6 (3)	251.4 (−2)
Jul	259.2	314.1	258.9	266.0 (3)	299.5 (16)	258.9	266.7 (3)	299.7 (16)
Aug	283.9	277.6	283.7	326.4 (15)	284.7 (0)	283.8	326.0 (15)	284.8 (0)
Sep	194.0	146.1	192.9	211.0 (9)	226.8 (18)	195.1	207.5 (6)	225.4 (16)
Oct	98.8	36.9	98.9	110.5 (12)	123.4 (25)	98.9	109.9 (11)	123.5 (25)
Nov	57.2	7.5	57.0	85.1 (49)	93.3 (64)	55.9	87.3 (56)	95.9 (71)
Dec	54.4	2.2	54.4	56.6 (4)	60.5 (11)	54.1	57.3 (6)	60.7 (12)
Annual	1732.1	1322.8	1729.1	1935.8 (12)	1959.7 (13)	1733.7	1932.4 (11)	1957.9 (13)

HESSD
10, 6847–6896, 2013

Estimating monthly rainfall in rural river basins

D. L. Jayasekera and J. J. Kaluarachchi

Table 5. Percent changes of projected areal rainfall and historical rainfall.

Season	Statistic	CGCM3.1		ECHAM5	
		2011–2050	2051–2090	2011–2050	2051–2090
Annual	Max	14.7	12.0	9.6	11.9
	Mean	11.8	13.1	11.6	13.0
	Min	19.4	22.3	8.7	15.4
Wet	Max	9.1	13.8	8.8	14.0
	Mean	9.7	7.6	9.4	7.5
	Min	1.1	−2.8	0.4	−3.0
Dry	Max	23.3	38.0	22.2	37.9
	Mean	18.2	30.7	18.5	30.6
	Min	−1.6	−25.5	−1.4	−25.5

HESSD
10, 6847–6896, 2013

Table A1. Computed percent changes of downscaled mean monthly rainfalls for the 10 selected stations from 2011 to 2090.

Station	Period	GCM	Month											
			Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1	2011–2050	CGCM	1.1	33.7	21.2	−7.0	−17.6	18.6	12.9	14.3	28.9	17.6	43.9	−14.6
		ECHAM	1.4	30.0	21.3	−7.4	−17.7	18.5	13.7	13.9	26.1	16.9	43.5	−14.3
	2051–2090	CGCM	10.8	30.6	4.6	27.8	−14.7	−2.4	9.2	16.7	−1.7	38.6	104.8	0.2
		ECHAM	11.5	29.0	3.8	24.0	−14.5	−1.6	9.0	16.7	−3.4	38.1	105.5	−0.1
2	2011–2050	CGCM	43.1	23.5	49.6	47.2	13.6	−10.8	−3.1	−1.3	36.2	12.2	−7.6	6.1
		ECHAM	33.9	24.2	48.2	45.4	13.6	−10.8	−3.3	−1.3	35.2	11.4	−4.7	6.4
	2051–2090	CGCM	38.0	5.0	39.0	3.7	−21.1	6.5	−24.2	−27.0	25.4	3.1	1.6	−7.2
		ECHAM	39.2	5.4	37.3	1.0	−21.0	5.2	−24.0	−27.0	24.7	3.0	3.5	−6.8
3	2011–2050	CGCM	−0.2	46.0	14.2	3.2	41.2	12.6	7.0	14.8	21.2	−0.5	66.1	37.2
		ECHAM	3.7	47.4	13.2	2.3	40.5	12.3	7.7	14.6	16.3	−0.9	75.5	39.2
	2051–2090	CGCM	3.6	24.0	24.9	44.8	1.8	11.9	24.7	−19.6	37.8	−17.3	80.8	26.9
		ECHAM	5.0	25.0	29.0	40.6	1.8	11.4	24.8	−19.4	36.6	−17.5	93.7	26.5
4	2011–2050	CGCM	12.5	5.5	−11.1	64.5	6.8	−3.6	13.6	22.8	4.1	12.3	55.2	−20.3
		ECHAM	12.1	3.7	−11.3	63.7	6.4	−3.3	14.3	22.7	2.5	11.7	54.6	−19.3
	2051–2090	CGCM	111.8	38.7	1.2	32.1	3.1	−20.5	13.9	3.6	12.3	48.0	82.4	19.8
		ECHAM	112.2	39.6	−0.3	28.1	3.2	−20.4	14.0	3.6	11.0	48.4	83.2	20.5
5	2011–2050	CGCM	33.3	17.1	−0.9	13.7	22.1	23.4	−0.2	9.5	10.3	49.4	15.7	7.5
		ECHAM	33.5	14.2	−0.3	12.5	22.4	23.3	0.4	9.0	7.6	48.6	15.5	7.2
	2051–2090	CGCM	3.1	−1.2	16.3	−11.6	26.2	−3.7	5.6	30.1	−13.6	1.3	30.3	14.0
		ECHAM	3.4	−3.7	15.3	−14.5	26.5	−3.3	5.7	30.1	−14.1	1.5	30.2	13.9
6	2011–2050	CGCM	12.0	25.8	24.9	27.4	28.2	−3.3	2.9	−1.3	42.4	20.1	20.5	42.3
		ECHAM	12.6	24.0	25.4	24.2	27.6	−3.3	2.7	−1.4	39.9	19.1	23.2	43.9
	2051–2090	CGCM	40.8	2.1	25.3	17.8	0.1	24.3	12.6	4.8	14.1	55.1	−12.0	5.1
		ECHAM	41.4	2.1	23.4	14.6	1.0	24.4	13.0	5.0	13.3	55.0	−10.3	5.7
7	2011–2050	CGCM	−42.0	−5.2	3.6	−13.3	15.0	10.2	−10.0	15.2	11.1	28.5	17.3	6.6
		ECHAM	−42.3	−3.1	−2.7	−14.1	14.6	10.3	−11.1	15.1	10.4	27.2	32.6	8.0
	2051–2090	CGCM	−32.8	−33.1	30.6	−6.4	−4.1	8.0	22.7	−0.3	12.1	8.7	2.2	−25.0
		ECHAM	−32.8	−31.5	28.5	−8.8	−4.0	6.2	22.6	−0.4	11.6	8.5	7.7	−24.2
8	2011–2050	CGCM	45.5	19.7	4.8	18.1	26.1	−5.8	−13.3	18.8	22.0	1.0	36.9	23.5
		ECHAM	52.9	18.2	5.5	16.9	25.1	−6.1	−13.4	18.4	20.6	0.7	39.9	23.7
	2051–2090	CGCM	−36.2	−9.6	−0.1	28.5	23.2	6.0	−9.7	25.8	19.2	81.8	39.0	−2.4
		ECHAM	−37.0	−9.3	0.0	25.4	22.7	5.5	−9.7	26.5	19.1	82.0	40.0	−1.4
9	2011–2050	CGCM	−32.7	18.8	−4.7	8.7	27.2	2.1	−22.3	−3.3	−22.0	3.1	3.2	−34.4
		ECHAM	−32.1	20.1	−4.1	7.9	27.3	2.2	−21.7	−3.4	−22.3	2.3	11.2	−33.9
	2051–2090	CGCM	−61.2	20.9	61.1	−19.9	10.6	3.9	6.7	10.1	19.5	25.7	62.9	−88.3
		ECHAM	−60.7	19.4	58.8	−22.1	10.9	3.8	6.9	9.9	20.5	25.7	72.2	−88.1
10	2011–2050	CGCM	15.8	28.9	17.6	10.0	2.8	2.4	17.8	37.0	19.5	10.8	43.4	12.6
		ECHAM	25.7	29.6	15.1	9.0	2.5	2.3	18.2	36.9	18.5	10.8	46.5	14.1
	2051–2090	CGCM	−24.2	−0.2	−36.7	36.7	30.1	20.8	13.3	27.0	24.3	20.9	5.2	−52.6
		ECHAM	−23.3	2.6	−37.3	33.5	30.5	20.8	12.9	27.0	24.3	21.2	−4.9	−51.0

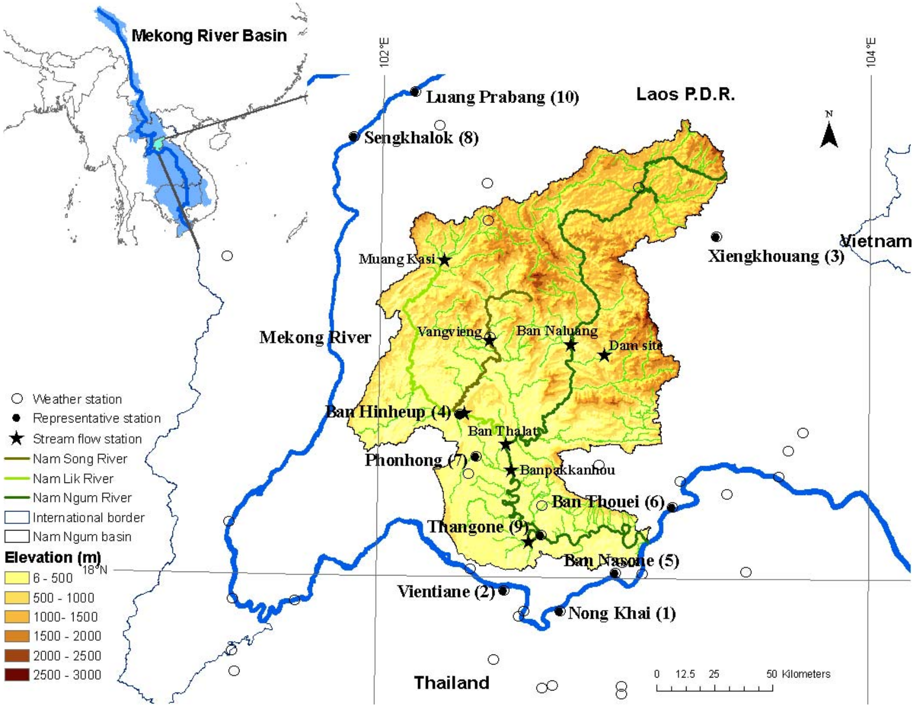

Fig. 1. Layout of the Nam Ngum River Basin in Laos. The number following the station name indicate the ten representative stations used in the analysis while all 40 weather stations are shown in blank circles.

HESSD
10, 6847–6896, 2013

Estimating monthly rainfall in rural river basins

D. L. Jayasekera and J. J. Kaluarachchi

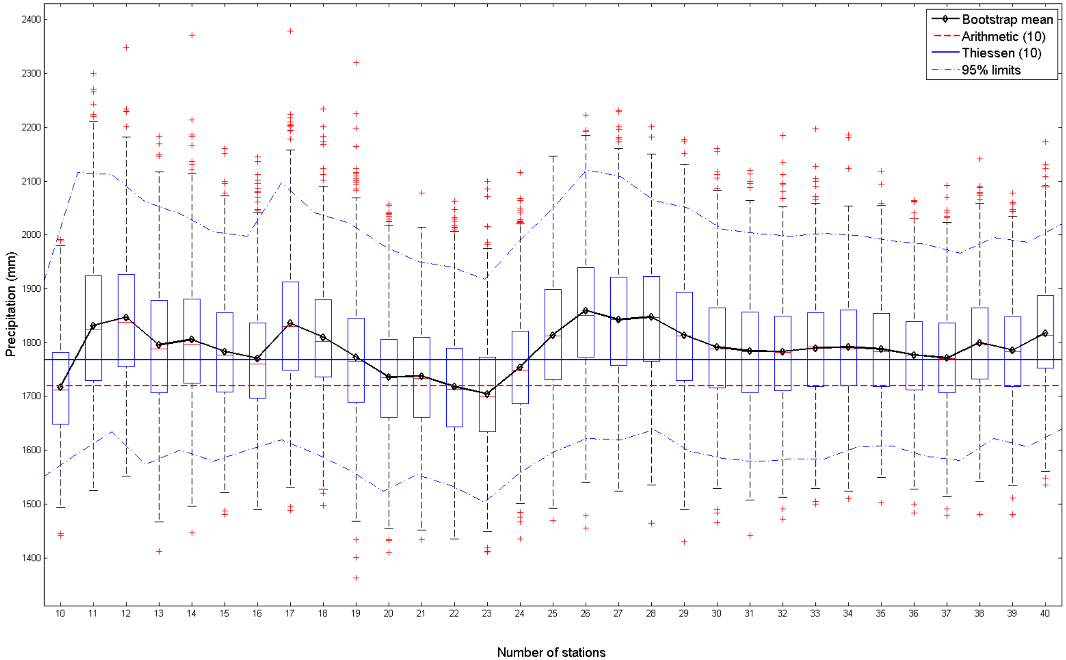

Fig. 2. Box plots showing the estimated mean annual rainfall using different number of stations by the bootstrap method.

HESSD
10, 6847–6896, 2013

Estimating monthly rainfall in rural river basins

D. L. Jayasekera and J. J. Kaluarachchi

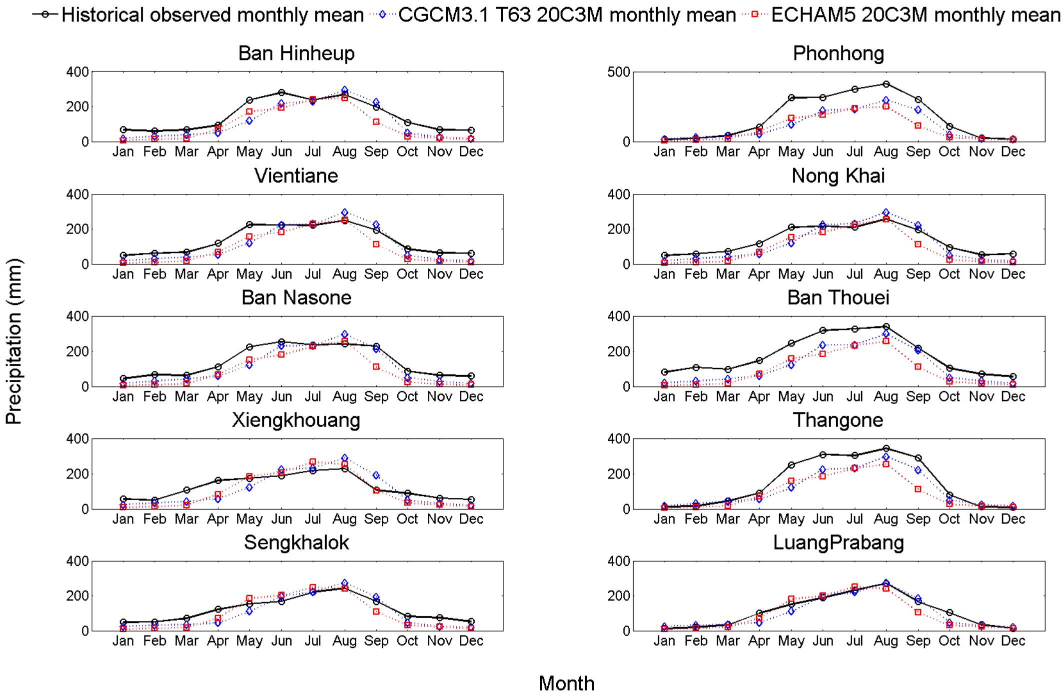

Fig. 3. Comparison of raw GCM mean monthly rainfall for the baseline scenario (20C3M) with historical observed for the 40 yr period from 1961 to 2000.

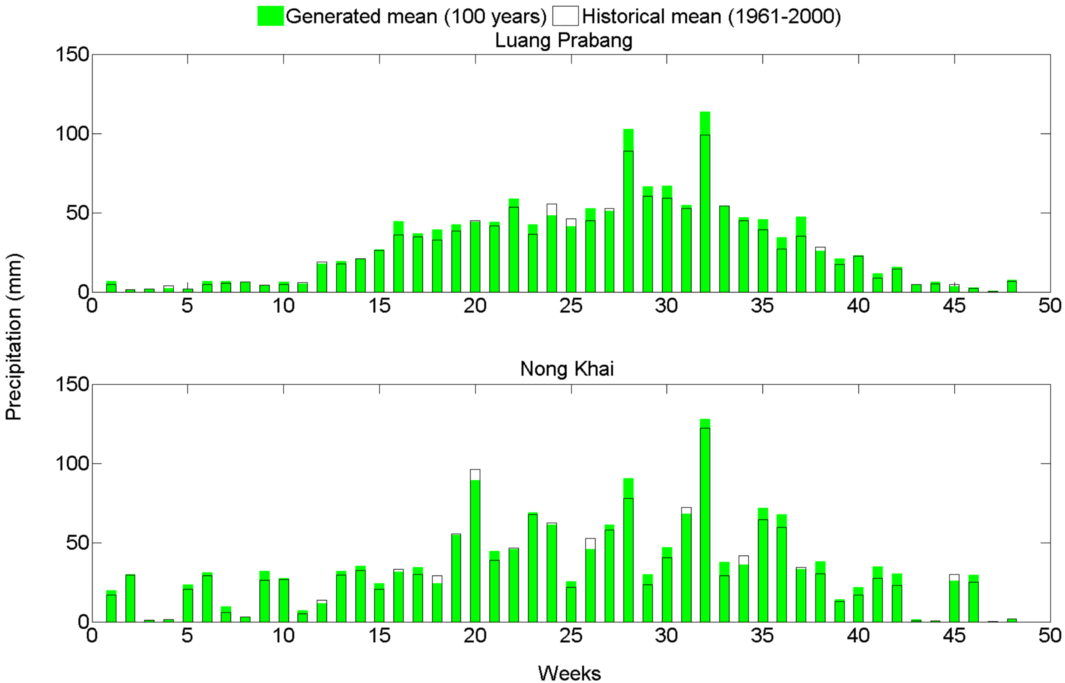

Fig. 4. Comparison of mean historical and generated rainfall series. Blank and shaded bar graphs represent the mean of historical observed and generated values, respectively.

HESSD
10, 6847–6896, 2013

D. L. Jayasekera and J. J. Kaluarachchi

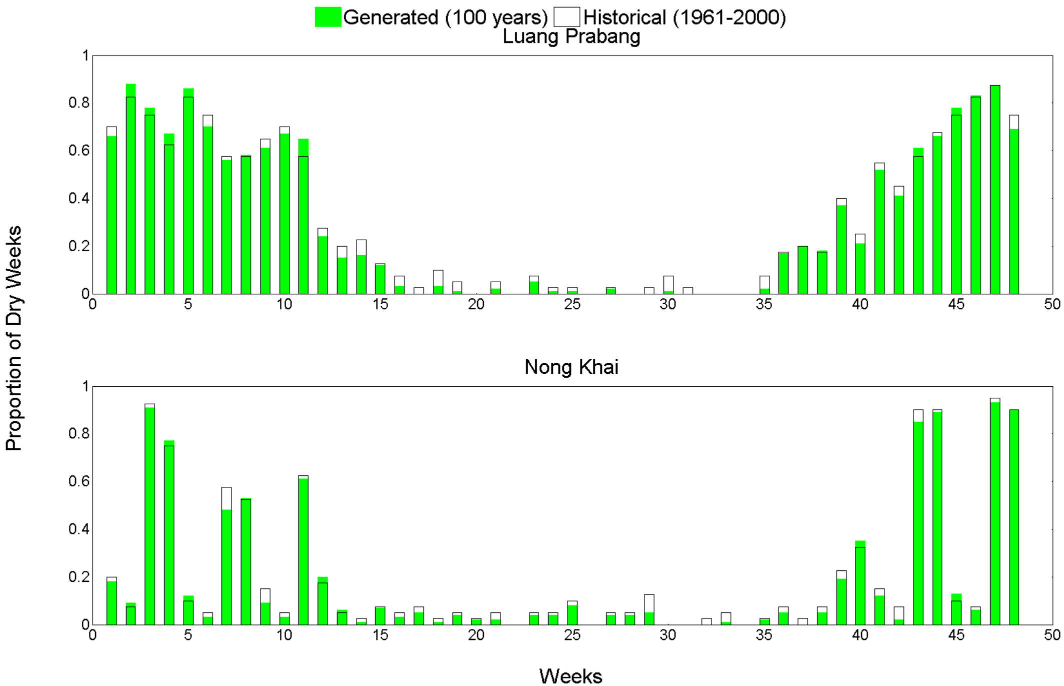

Fig. 5. Comparison of historical and generated proportion of dry days.

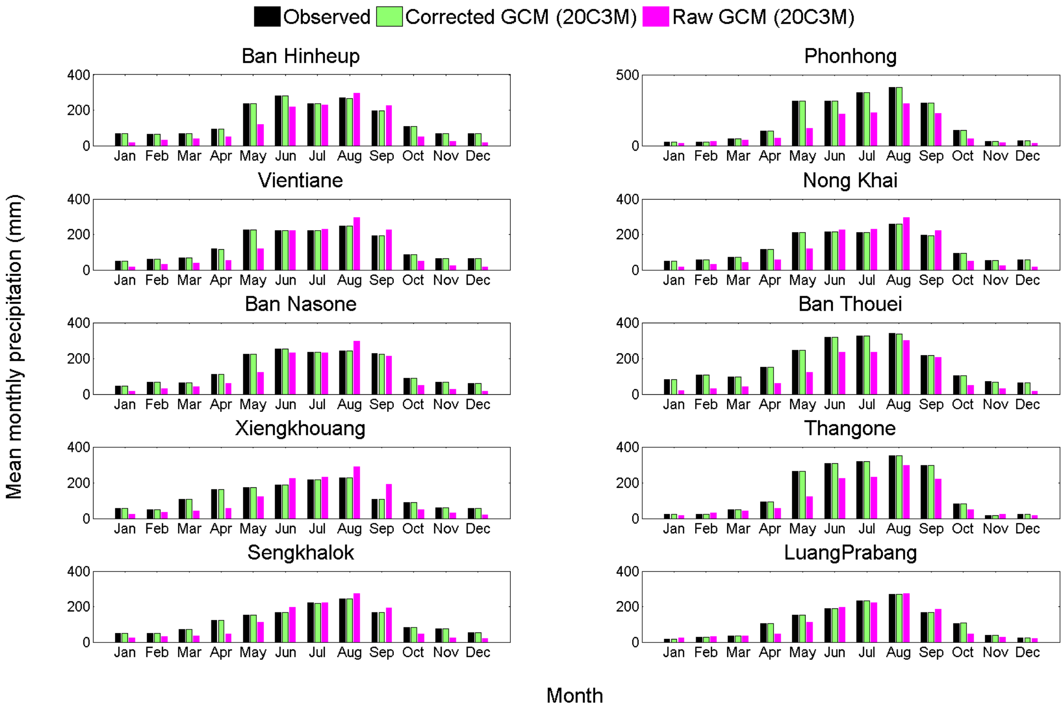

Fig. 6. Comparison of observed, corrected, and raw GCM mean monthly rainfall for the baseline scenario (20C3M) from 1961 to 2000.

HESSD
10, 6847–6896, 2013

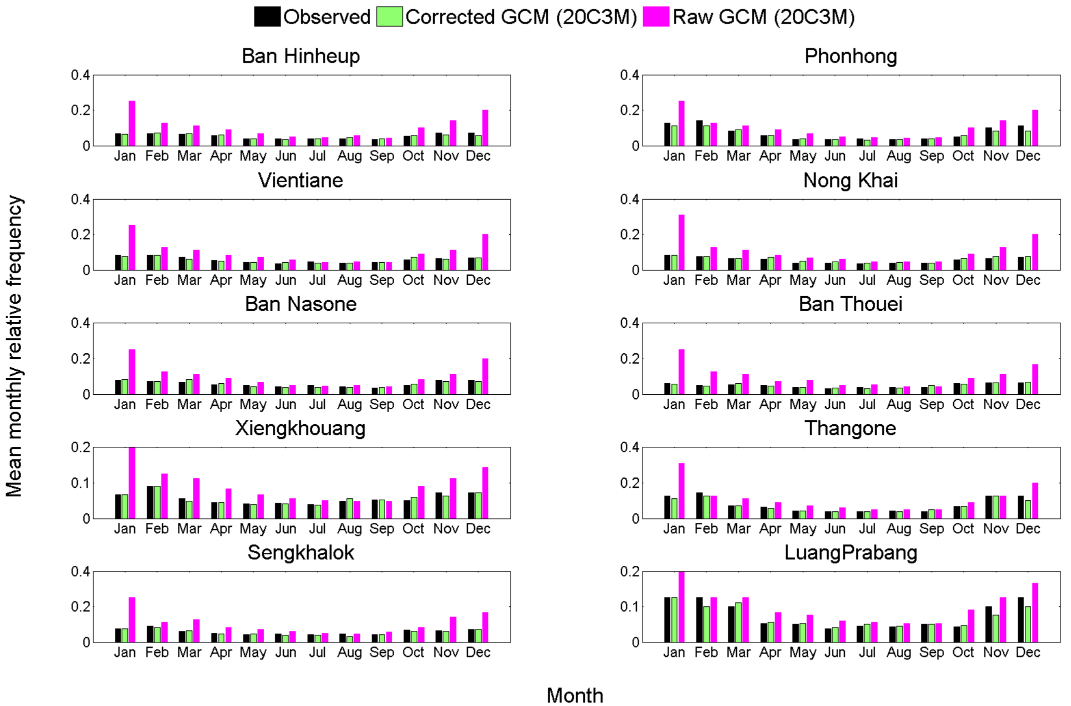

Fig. 7. Comparison of coefficient of variation for observed, corrected, and raw GCM results of monthly rainfall for the baseline scenario (20C3M) from 1961 to 2000.

HESSD
10, 6847–6896, 2013

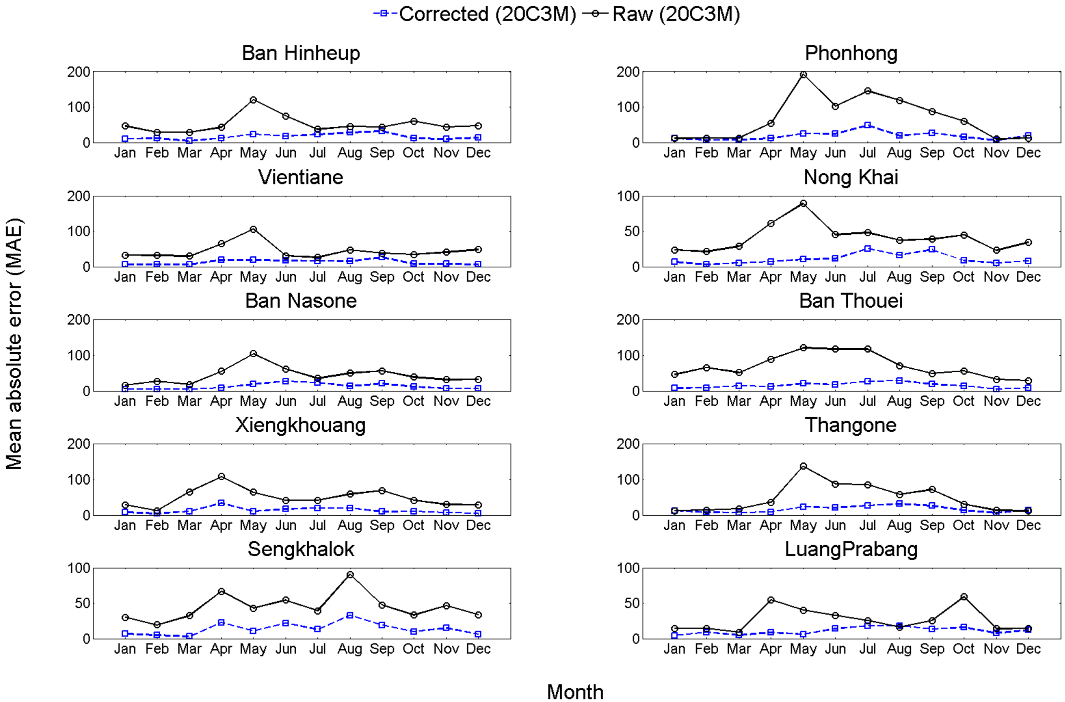

Fig. 8. Comparison of monthly mean absolute error for the baseline scenario (20C3M) from 1961 to 2000.

HESSD
10, 6847–6896, 2013

Estimating monthly rainfall in rural river basins

D. L. Jayasekera and J. J. Kaluarachchi

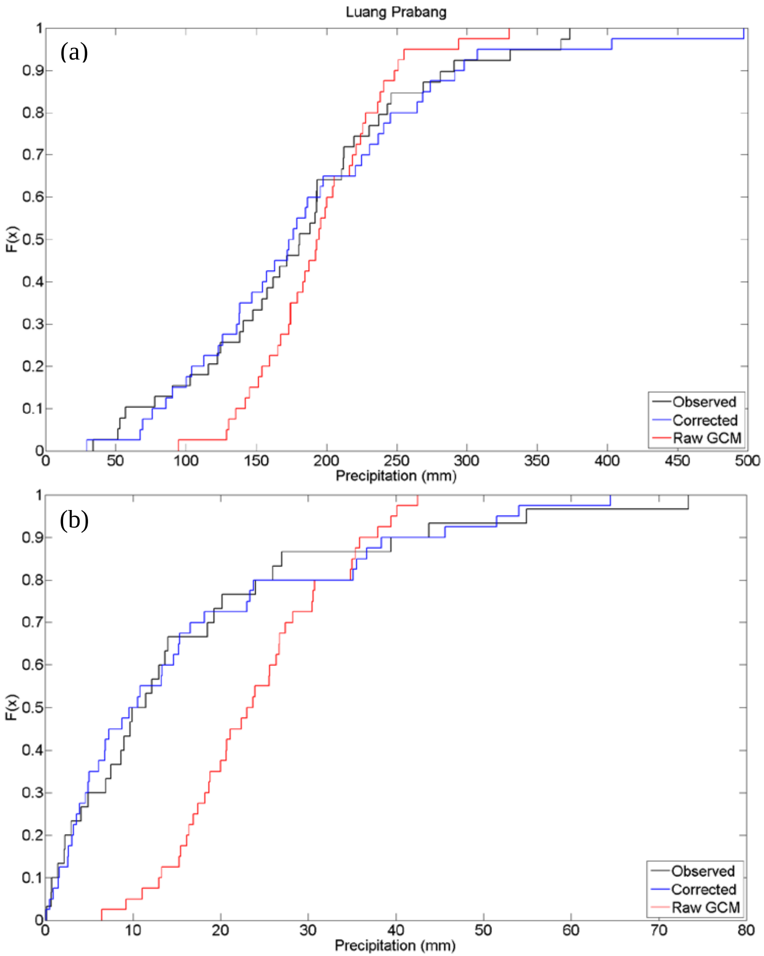

Fig. 9. Comparison of results of goodness-of-fit (K–S) test for Luang Prabang (key station): **(a)** wet month (June), **(b)** dry month (January).

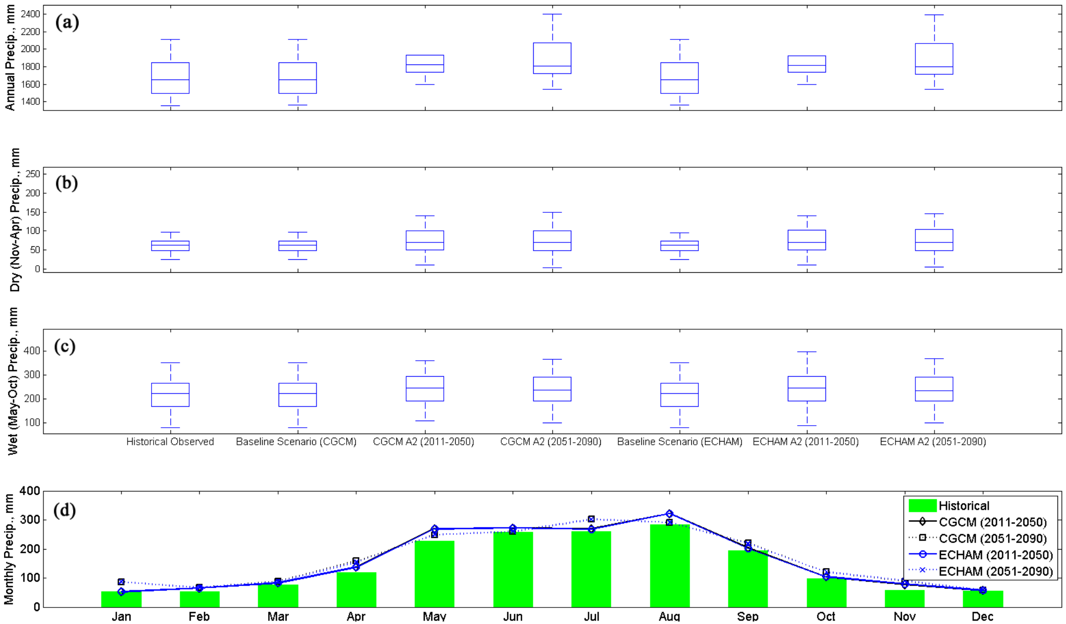

Fig. 10. Box plot comparison of temporal characteristics of historical and projected rainfall from 2011 to 2090: **(a)** mean annual, **(b)** dry season, **(c)** wet season, and **(d)** mean monthly rainfall.

HESSD
10, 6847–6896, 2013

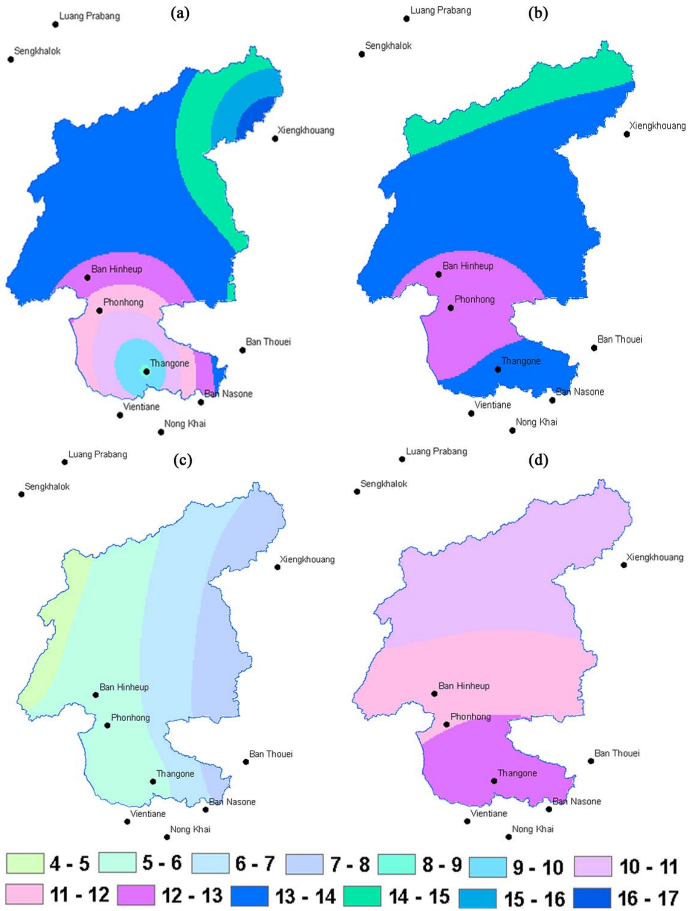

Fig. 11. Spatial distributions of percent changes in mean annual rainfall from 2011 to 2090: **(a)** CGCM (2011–2050), **(b)** CGCM (2051–2090), **(c)** ECHAM (2011–2050), and **(d)** ECHAM (2051–2090).

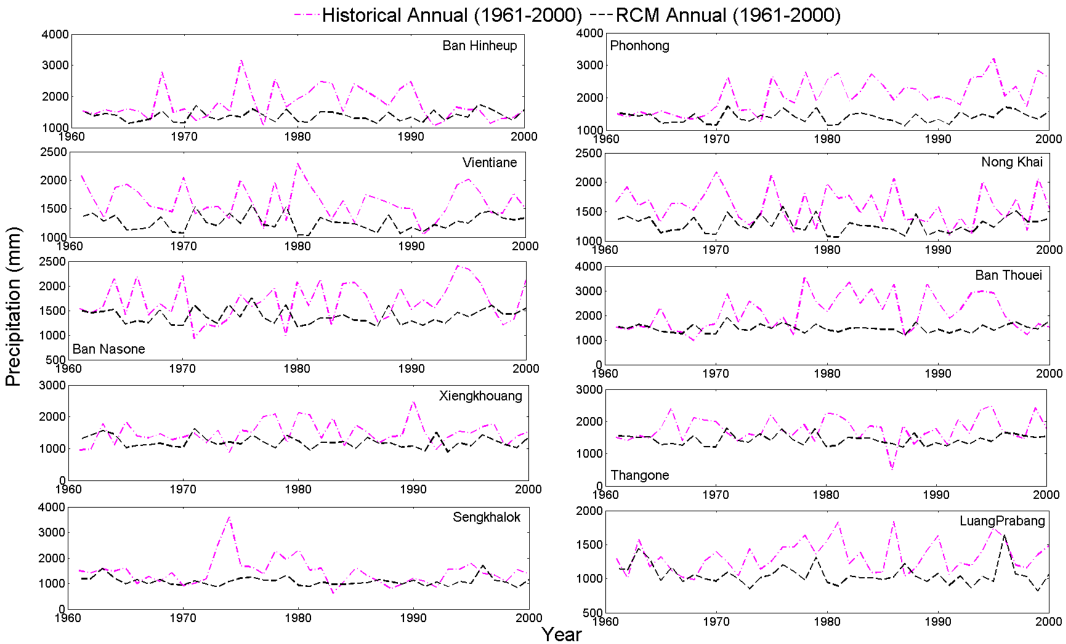

Fig. 12. Comparison of time series of annual rainfall between the RCM outputs and historical data from 1961 to 2000.

HESSD
10, 6847–6896, 2013

Estimating monthly rainfall in rural river basins

D. L. Jayasekera and J. J. Kaluarachchi

Title Page
Abstract | Introduction
Conclusions | References
Tables | Figures
Back | Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion

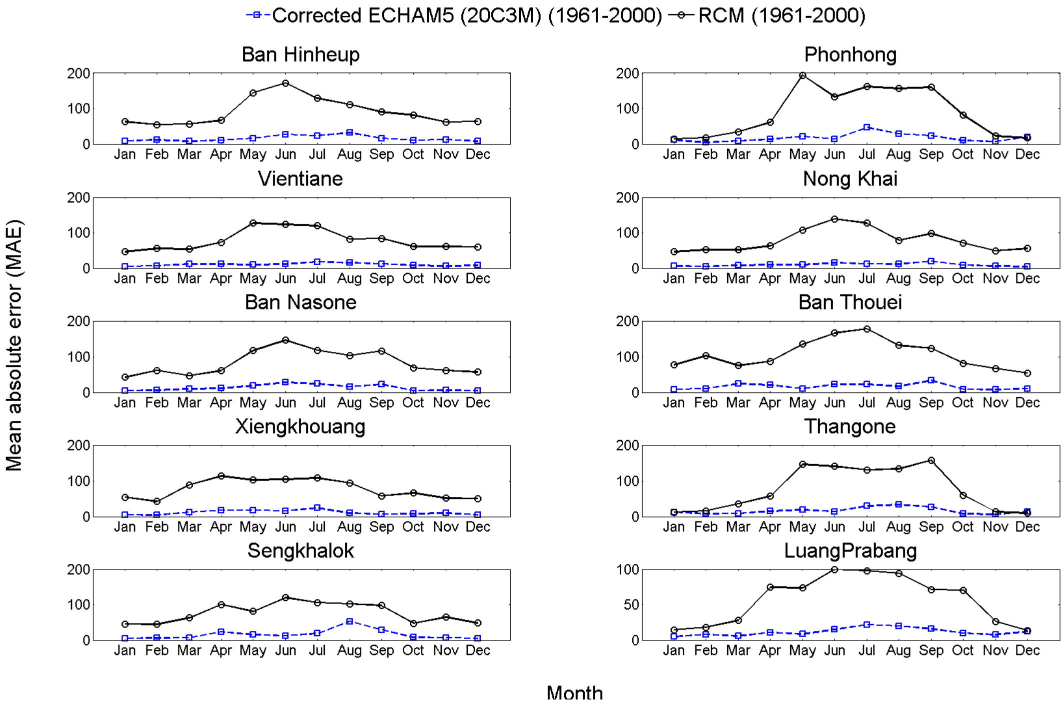

Fig. 13. Comparison of MAE of the RCM and corrected ECHAM outputs at ten selected stations from 1961 to 2000.

HESSD
10, 6847–6896, 2013