Spatial disaggregation of bias-corrected GCM precipitation for improved hydrologic simulation: Ping River Basin, Thailand

D. Sharma, A. das Gupta, M. S. Babel

▶ **To cite this version:**

D. Sharma, A. das Gupta, M. S. Babel. Spatial disaggregation of bias-corrected GCM precipitation for improved hydrologic simulation: Ping River Basin, Thailand. Hydrology and Earth System Sciences Discussions, 2007, 4 (1), pp.35-74. hal-00298807

HAL Id: hal-00298807
https://hal.science/hal-00298807

Submitted on 18 Jun 2008

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Hydrol. Earth Syst. Sci. Discuss., 4, 35–74, 2007
www.hydrol-earth-syst-sci-discuss.net/4/35/2007/
© Author(s) 2007. This work is licensed
under a Creative Commons License.

Hydrology and Earth System Sciences Discussions

Papers published in *Hydrology and Earth System Sciences Discussions* are under open-access review for the journal *Hydrology and Earth System Sciences*

Spatial disaggregation of bias-corrected GCM precipitation for improved hydrologic simulation: Ping River Basin, Thailand

D. Sharma, A. Das Gupta, and M. S. Babel

Water Engineering and Management, Asian Institute of Technology, Thailand

Received: 20 November 2006 – Accepted: 17 December 2006 – Published: 17 January 2007

Correspondence to: A. Das Gupta (adg@ait.ac.th)

Abstract

Global Climate Models (GCMs) precipitation scenarios are often characterized by biases and coarse resolution that limit their direct application for basin level hydrological modeling. Bias-correction and spatial disaggregation methods are employed to improve the quality of ECHAM4/OPYC SRES A2 and B2 precipitation for the Ping River Basin in Thailand. Bias-correction method, based on gamma-gamma transformation, is applied to improve the frequency and amount of raw GCM precipitation at the grid nodes. Spatial disaggregation model parameters (β, σ^2), based on multiplicative random cascade theory, are estimated using Mandelbrot-Kahane-Peyriere (MKP) function at q=1 for each month. Bias-correction method exhibits ability of reducing biases from the frequency and amount when compared with the computed frequency and amount at grid nodes based on spatially interpolated observed rainfall data. Spatial disaggregation model satisfactorily reproduces the observed trend and variation of average rainfall amount except during heavy rainfall events with certain degree of spatial and temporal variations. Finally, the hydrologic model, HEC-HMS, is applied to simulate the observed runoff for upper Ping River Basin based on the modified GCM precipitation scenarios and the raw GCM precipitation. Precipitation scenario developed with bias-correction and disaggregation provides an improved reproduction of basin level runoff observations.

1 Introduction

The Southeast Asia is vulnerable to climate change and its variability, including rise in sea level, shifts of climatic zones and occurrence of extreme events such as droughts and floods. The range variation of various climatic model projected mean temperature, precipitation and runoff in 2100 for this part of the world are found to be 1.0 to 4.5°C, –20 to +20% and –10 to +30% respectively (IPCC, 2001). These broad spectrums of projected range imply that the results of the climatic models cannot be directly used in

HESSD

4, 35–74, 2007

analyzing the water resource situation at basin level using high-resolution hydrologic models. The mismatch between coarse resolution projections of global climate models (GCMs) and fine resolution data requirements of hydrologic models is the major obstacle in assessing the impacts of climate change on water resources at the basin level. However, climate model outputs can be utilized more efficiently in climate change impact studies with due consideration on its coarse resolution and biasness. Many GCMs are available with high temporal resolution data output at daily or at even lesser time intervals that satisfy the basic need of temporal scale in hydrological modeling. However, gap still exists in developing proper techniques to minimize these biases and to downscale the spatial resolution of the GCMs. Identification, application and verification of proper methods are therefore necessary to minimize the effect of these drawbacks in basin scale studies.

Despite being important tools to project the expected future scenarios of climatic parameters, GCMs simultaneously contain biases when compared to observed data due to their parameterization systems and large grid size (~300×300 km^2). These types of errors are considered insignificant when applying for the estimation of climate change impact at regional scale. But, such biased climate model scenarios are inadequate for their use in hydrologic models to analyze impact of climate change at basin level. Bias correction methods largely eliminate these problems with added emphasis on statistical characteristics of historical data. Rescaling is the easiest bias-correction method to rectify the systematic error in the mean rainfall amount. A "quantile-based" bias-correction approach is useful to statistically transform rainfall simulated by GCM to bias-corrected data and to make it applicable for use in impact assessment models (Wood et al., 2002; Hamlet et al., 2003). Ines and Jansen (2006) applied empirical-gamma transformation, gamma-gamma transformation and multiplicative shift techniques to correct the frequency and amount distribution of daily GCM rainfall for a particular station and then applied it for maize yield simulation model. All these methods improved the results of maize yield simulation. These studies indicate the significance of bias-correction of GCMs precipitation when applied to impact assessment studies.

Low spatial resolution of GCMs is another critical issue to be considered before applying it to the hydrological impact assessment model. There are numerous techniques available for downscaling the climate models data to a high spatial resolution with their respective advantages and disadvantages. No downscaling method is available with perfection, therefore the techniques used to interpret GCM simulations should be chosen based on the objectives of the research (Hamlet et al., 2003). The disaggregation model, based on multiplicative random cascade theory, is one of the techniques to spatially downscale the projected GCM rainfall. A continuous form of multiplicative random cascades has its major advantage in developing cascades over a continuous interval of scales instead of only a discrete set (Marsan et al., 1996); however, a discrete form of multiplicative random cascade has ability to separate rainy and non-rainy areas (Schertzer and Lovejoy, 1987; Gupta and Waymire, 1993; Over and Gupta, 1994). The main advantage of this approach to precipitation analysis is that it describes the complex rainfall process over a wide range of scales with few parameters. The rainfall modeling based on discrete multiplicative random cascades has been tested for spatial and temporal disaggregation under different climatic conditions with similar conclusion that it was possible to capture rainfall variability of sub grid scale using this approach (Olsson, 1998; Gunter et al., 2001; Molnar and Burlando, 2005). A few studies have been conducted by incorporating spatial heterogeneity in random cascade process and the overall results were found to be acceptable based on important spatial and temporal characteristics of rainfall (Jothityangkoon et al., 2000; Pathirana and Herath, 2002; Tachikawa et al., 2004).

The focus of the present paper is to ameliorate the GCM precipitation for further use in hydrological model by applying the bias-correction and stochastic disaggregation method. Bias-correction method is applied to GCM precipitation at the grid node and then spatially downscaled using multiplicative random cascade model while preserving the spatial heterogeneity for each month. The Ping River Basin, North of Thailand, is selected for the application of the approach.

2 Study area

The Ping River Basin is one of the eight sub-basins in Chao Phraya Basin. It stretches from latitude 19.75° N to 15.75° N and from longitude 98.10° E to 100.20° E, with a catchment area of 34 453 km^2 (Fig. 1). It covers about 22% of the Chao Phraya River Basin and contributes about 24% (9044×10^6 m^3) of the total average annual runoff. Terraced mountains mainly characterize the topography of Ping River Basin. About 55.5% of total basin area is in the elevation range of 500–1500 m. The weather is mainly influenced by the Southwest and Northeast monsoon. It is also influenced by the depression from the South China Sea during July and September, resulting in abundant rain from May to October. The climate is characterized by average annual precipitation of 1097 mm and average annual temperature of 26.7°C. Nearly 90% of rain occurs during rainy season (May–October). Water demand of Ping River Basin in the year 2002 was 5809×10^6 m^3 and it is expected to be 6316×10^6 m^3 after 20 years (Department of Water Resource, 2003). Climate change and its variability can significantly impact on water resources of the basin. The analyze of extreme precipitation and temperature events in this river basin reveal the presence of significant trends in most of the climate indices, with a general tendency for a decrease in precipitation and a significant increase in temperature indices (Sharma et al., 2006).

Daily precipitation data for the period 1991 to 1999 is obtained from the Royal Irrigation Department (RID, Thailand), Thailand Meteorological Department (TMD) and Department of Meteorology and Hydrology (DMH, Myanmar) for 99 stations, as indicated in Fig. 1.

3 Characteristics of ECHAM4/OPYC3

ECHAM4/OPYC3 is selected due to its easy accessibility, high temporal (6 hourly) and spatial resolution as compared to other climate models (Feenstra et al., 1998). ECHAM4/OPYC3 model has been developed at the Max Planck Institute in Hamburg,

Germany, using the weather forecasting model (ECMWF) and it is a model of fourth generation. ECHAM4 uses 2.8°×2.8° grid cells of a 19-layer atmosphere and an 11-layer ocean (Roeckner et al., 1996). Six new greenhouse gases (GHGs) emission reference scenario groups (A1B, A1FI, A1T, A2, B1, and B2), organized into four scenarios "families" (A1, A2, B1, and B2), were developed by the IPCC and published as the Special Report on Emissions Scenarios (SRES). Scenarios A2 and B2, two contrasting future emission scenarios, are used to account for uncertainty in future GHGs/sulphate emissions data. The A2 scenario assumes an emphasis on local traditions, high population growth, and less concern for rapid economic development. The B2 scenario envisages less rapid, and more diverse technological change with emphasis on community initiative and social innovation to find local, rather than global solutions. This scenario is oriented toward environmental protection and social equity (Nakicenovic and Swart, 2000).

The ECHAM4/OPYC3 climate model precipitation data for SRES A2 and B2 are obtained from the World Data Center for Climate through CERA (Climate and Environmental Retrieving and Archiving) database system. Two grid boxes of selected model are identified for the study area as shown in Fig. 1. These two grid boxes account for about 96% of total Ping Basin area and out of it 82.4% is within grid 1. Daily rainfall data is prepared by adding the 6 hourly model data.

Figure 2a reveals increasing trend of annual mean temperature for SRES scenarios A2 and B2 for both the grids. The rise in mean annual temperatures for SRES A2 and B2 are found to be about 0.30°C/decade and 0.22°C/decade respectively. Annual temperature across the domain may increase in 2020s by 0.4 to 0.5°C and 1.3°C to 1.5°C in 2050s, according to ECHAM4/OPYC3 with the high emissions scenario (A2). With the low emissions scenario (B2), there may be a rise of 0.3 to 0.4°C by 2020s and 0.9°C to 1.1°C by 2050s. The rate of temperature rise per decade is approximately 1.25 times more in scenario A2 than that in scenario B2. Figure 2b shows the monthly increase in mean temperature per decade for SRES A2 and B2 for both the grids. The variation in monthly increase in temperature is more uniform in grid 2 as compared to

grid 1, under both the scenarios. Figure 3 illustrates the trend of annual standardized anomalies of ECHAM4/OPYC3 precipitation scenarios. There is a tendency toward less abundant rainfall (rainfall close to mean) in future. A comparison of these trends suggests that warmer temperatures and lesser rainfall will affect the present situation of water resources in the study area. Figure 4 shows the mean monthly trend of ECHAM4 model and observed rainfall. The selected model scenarios over predict the mean monthly precipitation at grid 1 (82.4% of Ping Basin area) while it under predict at grid 2 for rainy period. In dry season (November–April), ECHAM4 model over predicts rainfall for the study domain. This shortcoming hinders in direct application of climate model data in impact assessment studies.

4 Methodology

4.1 Bias correction

Variability of rainfall largely depends on its frequency and amount, and it is difficult to estimate average rainfall in a particular region. Gamma-gamma (GG) transformation is used to reduce the gap between the daily GCM simulated and observed rainfalls using GCM bias correction tool – version 0.3a (Ines, 2004). The basic idea of this bias-correction procedure is to reduce the biases from the frequency and amount at grid node before applying downscaling technique. The steps are as follows:

(1) Establish the empirical distributions, $F(x)$, by first classifying long-term daily rainfall data for each month, based on positions of the ordered datasets, using Eq. (1):

$$F(x) = \frac{n}{m}, \tag{1}$$

where n is the position of x in the ordered array, and m is the total number of data in the array. This should be followed by calculation of a threshold value ($x^{\%}_{\mathrm{GCM}}$), derived from the empirical distribution of daily historical rainfall, to truncate the empirical distri-

bution of the raw daily GCM rainfall for that particular month. Basically, F (x_{his}=0.0) is determined and then map to the daily GCM rainfall distribution.

(2) GG transformation method is selected for rainfall amount correction. For GG transformation, the truncated daily GCM rainfall and historical rainfall data are fitted to a two-parameter gamma distribution (Eq. 2) and then the cumulative distribution (Eq. 3) of the truncated daily GCM rainfall is mapped to the cumulative distribution of the truncated historical data (Eq. 4). The shape and scale parameters (α and β) for each gamma distribution are determined using Maximum Likelihood Estimation.

$$f(x;\alpha,\beta) = \frac{1}{\beta^{\alpha}\Gamma(\alpha)} x^{\alpha-1} \exp\left(-\frac{x}{\beta}\right); \quad x \geq x_{\text{Trunc}} \tag{2}$$

$$F(x;\alpha,\beta) = \int_{x_{\text{Trunc}}}^{x} f(t)dt \tag{3}$$

$$F(x_{\text{GCM}};\alpha,\beta|_{\text{GCM}}) \Rightarrow F(x_{\text{His}};\alpha,\beta|_{\text{His}}) \tag{4}$$

The corrected GCM rainfall amount for the particular day can be calculated by taking the inverse of Eq. (4) such that:

$$x'_{\text{GCM}} = F^{-1}\{F(x_{\text{His}};\alpha,\beta|_{\text{His}})\} \tag{5}$$

The bias-correction method is applied to ECHAM4 SRES scenarios A2 and B2 relative to spatial interpolated rainfall at GCM grid node. Inverse Distance Weighting (IDW) method is used to estimate the spatial average rainfall at grid node from observed daily rainfall. Correlation coefficient (R), root mean square error (RMSE), standard deviation (SD) and index of agreement (d) are determined to assess the overall ability of the bias-correction method. The quality of bias-corrected rainfall on monthly scale is evaluated using the mean square error (MSE) skill score with the raw GCM as a reference and formulated as:

$$\text{Skill Score}, SS = 1 - \frac{\text{MSE}_{\text{corrected}}}{\text{MSE}_{\text{raw}}}, \tag{6}$$

HESSD
4, 35–74, 2007

where $MSE_{corrected}$ and MSE_{raw} are mean square errors of the bias-corrected GCM and raw GCM data. The range of MSE skill score varies from negative infinity to 1, and 0 indicates no skill when compared to raw GCM data.

4.2 Spatial disaggregation model

In spatial disaggregation model, rainfall is a combined effect of two processes i.e., a multifractal process that is highly variable in space but statistically uniform at smaller scales, and a process that presents spatial heterogeneity of rainfall. The disaggregation model, based on multiplicative random cascade model, is used to distribute rainfall mass on successive regular subdivisions of a scale in multiplicative manner with a branching number. The branching number (b) is defined as the ratio of the number of segments at cascade stage "$s+1$" to the cascade stage "s". Figure 5a shows the general layout of cascade disaggregation process. At each cascade stage, each segment is divided into b equal parts and each part is multiplied by a weighted value derived from a specified distribution, which is known as the cascade generator (W). The segment after "s" stages of subdivision is denoted by Δ_s^i (i =1, ..., b_s) and rainfall in this particular segment is expressed as $\mu_s\left(\Delta_s^i\right)$ in Eq. (7).

$$\mu_s\left(\Delta_s^i\right) = R_0\lambda_s \prod_{j=1}^{s} W_j(i), \text{for} i = 1, 2, \ldots, b^s; s > 0 \quad (7)$$

where R_0 is rainfall depth at initial stage and λ_s is the dimensionless spatial scale, which equals to b^{-s}. Figure 5b shows the example of disaggregation model output applied to GCM rain field with R_0=3.6 mm at stage s=7.

W is a random positive variable and is considered independent and identically distributed. Properties of W can be estimated from the moment scaling behaviour across scales. The statistical moment is the summation of rainfall volumes and then raise to the power "q" at each segment formed at a particular stage "s". Thiessen polygon method is used to calculate the rainfall volume at each segment at the selected stage.

Sample moment with order q (non-negative) is defined as:

$$M_s(q) = \sum_{i=1}^{b^s} \mu_s^q(\Delta_s^i) \tag{8}$$

In a random cascade, the statistical moments $M_s(q)$ are shown to be a log-log linear function of the scale of resolution λ_s. The slope of the sample moment scaling relationship is:

$$\tau(q) = \lim_{\lambda_s \to 0} \frac{\log M_s(q)}{-\log \lambda_s} \tag{9}$$

Over and Gupta (1996) has proposed an intermittent beta-lognormal model because of its ability to consider dry zone explicitly. It is given as:

$$P(W = b^{\beta - \sigma^2 \frac{\log[b]}{2} + \sigma X}) = b^{-\beta}$$
$$P(W = 0) = 1 - b^{-\beta}, \tag{10}$$

where β and σ^2 are model parameters and X is a standard normal variable. The expected value of W is:

$$E(W) = 1$$
$$E(W) = E[(b^{-\beta} b^{\beta - \sigma^2 \frac{\log(b)}{2} + \sigma X} + (1 - b^{-\beta})0] = E[b^{-\sigma^2 \frac{\log(b)}{2} + \sigma X}] = 1 \tag{11}$$

They also proposed using Mandelbrot–Kahane–Peyriere (MKP) function (Mandelbrot, 1974; Kahane and Peyriere, 1976) to estimate the model parameters.

$$\chi_b(q) = (\beta - 1)(q - 1) + (\sigma^2 \frac{\log(b)}{2})/(q^2 - q) \tag{12}$$

The MKP function, $\chi_b(q)$, is defined as the slope of the statistical moment $M_s(q)$ to the disaggregation stage "s". Considering that the cascade follows the scaling law (Eq. 9),

the MKP function of the cascade is $\tau(q)/d$. The first and second derivatives of $\tau(q)$ with respect to q, indicated by model parameters β and σ^2 respectively, are calculated using the finite difference method.

$$\sigma^2 = \frac{\tau^{(2)}(q)}{d \log(b)}$$

$$\beta = 1 + \frac{\tau^{(1)}(q)}{d} + \frac{\sigma^2 \log(b)(2q - 1)}{2} \quad (13)$$

The model parameters β and σ^2, respectively, represent intermittency and variability of the cascade generator. Values estimated for model parameters are plotted with average rainfall amount (large scale forcing) to find an empirical relation.

The modeling of rain fields using only multifractal does not consider the existing spatial heterogeneity. To improve the output of multifractal model in term of spatial heterogeneity, a long-term average of rainfall was included in the modeling process (Pathirana and Herath, 2002). The theoretical concept of the applied model is that the rainfall on the segment is a multiplication of two factors namely, parameter of multifractal modeling (M) and component of rainfall (G) that is invariant over a long aggregation. Spatial heterogeneity is incorporated in model by developing the field matrix (G) of normalized long-term monthly average rainfalls using IDW technique. The mathematical notation of the model is:

$$R_{i,j} = M_{i,j} G_{i,j} \quad M_{i,j} = \{0 \text{ for } G_{i,j} = 0,\ R_{i,j}/G_{i,j} \text{ otherwise}\} \quad (14)$$

The bias-corrected ECHAM4/OPYC SRES scenarios A2 and B2 are used in the multiplicative random cascade model for stage, s=7. The branching number (b) and embedding dimension (d) for the study are considered as 4 and 2 respectively. The values of β and σ^2 are evaluated by computing the derivatives of $\tau(q)$ at q=1, separately for each month. Outputs from the spatial disaggregation model are analyzed with respect

to rainfall amount distribution at different points, spatial distribution of rainfall, area distribution of residual and spatial autocorrelation of rain amounts to evaluate the model performance. Spatial autocorrelation is generally used to measure the self-similarity of rainfall field. The Moran's "*I*" and Geary's "*c*" indices are considered to measure the spatial autocorrelation between cells. The value of Moran's "*I*" range between –1 and 1 depending on the degree and direction of correlation. Geary's "*c*" ranges from 0 (maximum positive autocorrelation) to a positive value for high negative autocorrelation. Its expectation in the absence of autocorrelation is 1 (Sokal and Oden, 1978). Both statistics are defined as:

$$\text{Moran's } "I" = \frac{(N/S_0)\sum_i \sum_j w_{ij}(x_i - \mu)(x_j - \mu)}{\sum_i (xi - \mu)^2},$$

$$\text{Geary's } "c" = \frac{(N-1)/2S_0[\sum_i \sum_j w_{ij}(x_i - x_j)^2]}{\sum_i (x_i - \mu)^2}, \quad (15)$$

where μ is the mean of the variable x, w_{ij} are the elements of the spatial weights matrix, and S_0 (sum of the elements of the weights matrix) = $\sum_i \sum_j w_{ij}$.

5 Results and discussion

5.1 GCM rainfall bias correction

The gamma-gamma transformation method is applied to ECHAM4/OPYC3 SRES A2 and B2 precipitation scenarios to reduce the biases from the frequency and amount when compared with the computed frequency and amount at reference grid nodes based on the spatially interpolated field level rainfall data. Figure 6 shows the comparative trend in mean monthly rainfall amount, frequency and intensity for observed data, raw GCM scenarios and bias corrected GCM scenarios at each grid. The raw GCM scenarios over predict the mean monthly precipitation at grid 1 and under predict

at grid 2 for rainy season (May–October). In dry season (November–April), raw GCM scenarios over predict rainfall for both the grids. It is observed that the GCM simulates continuous rain events (rain>0.1 mm) with similar trend throughout the year. This trend of simulated rainfall lacks the demarcation in seasons and results in a large difference in the mean frequency of wet days obtained from model simulation and observed data. The mean rainfall intensity is also found to be low when compared to its corresponding observed value in the study domain except for the period May to August at grid 1. It is because of high difference in mean rainfall amount compared to mean frequency.

Statistical parameters indicating the correspondence of raw GCM and bias-corrected GCM (GG-GCM) scenarios with deduced data from field observation for the monthly mean rainfall amount are provided in Table 1. The standard deviation of GG-GCM precipitation data, when compared to that for the raw GCM data, is closer to the standard deviation deduced for observed values. Also there is increase in the correlation coefficient for all the scenarios in case of GG-GCM compared to raw GCM. For example, there is increase in correlation coefficient value from 0.32 to 0.64 in A2 scenario of grid 1. A correlation coefficient of the order of 0.60 to 0.66 is achieved for grid 1, while the correlation coefficient for grid 2 is in the range of 0.68 to 0.73. The RMSE for grid 1 varies between 2.06 and 2.15, while for grid 2 it varies between 6.0 and 6.45. The effect of bias-correction method is also checked using MSE skill score, as shown in Fig. 7. Positive value of the skill score reveals that the corrected data are better than the raw data, but the improvements vary with month. The improvement is relatively high during dry months compared to rainy months. The high skill score in dry months is the effect of reduction in wet days and in rainfall amount that happened to be the characteristics of raw GCM data as explained earlier. The GG-transformation is, therefore, effective in reducing the biases from the raw GCM precipitation when compared with the observed data. All these biases are taken as sources of uncertainties in climate change studies.

5.2 Spatial disaggregation model

The spatial disaggregation method is applied on bias corrected ECHAM4/OPYC3 SRES A2 and B2 precipitation scenarios to improve the resolution and to incorporate the spatial heterogeneity of rainfall for further use in hydrologic model application

5.2.1 Model parameter

Spatial disaggregation model performance depends mainly on parameters values (β, σ^2), which capture seasonal precipitation characteristics. β determines the growth of intermittency in the field ($0<\beta<1$), and σ^2 determines variance of the cascade generator ($\sigma^2>0$). High β indicates high intermittency while high σ^2 indicates peaks in the field and an increased level of multiscaling. These two parameters are estimated separately for each month based on nine years of observed daily rainfall data. Figure 8 shows the trend in precipitation scaling parameters between rainy and dry seasons. In rainy season, intermittency (β) is low and the variance (σ^2) is high as compared to dry season. The range of σ^2 is from 0.02 to 0.104 in both the grids. Positive values of σ^2 indicate that the structure of precipitation field is multiscaling in all the months, however the extent of multiscaling is less in dry season. Figure 9 shows the existing trend between the model parameters and the regional daily rainfall amount. The equation of best-fit line between parameters and average rainfall amount is given with each figure. Different forms of equation can be used to present this relationship as observed in other studies (Over and Gupta, 1994; Shrestha et al., 2004). Individual values are binned into different classes with N (rainfall events)=10 to reduce the scattered pattern of the parameters values. β values decrease with increase in rainfall while σ^2 values may remain almost not sensitive to rainfall amount (Over and Gupta, 1996; Jothityangkoon et al., 2000; Assela and Herath, 2002). σ^2 values show a seasonal trend but the trend is opposite to that of β values. As can be seen in Fig. 9, high rainfall amounts are less intermittent on an average and there is less effect of rainfall variability. These two model parameters are presented as a function of average spatial rainfall amount and both

show a strong dependency on rainfall amount. Statistics of the two model parameters for rainy months with *SD*, mean and coefficient of variation (*CV*) are summarized in Table 2. The results justify the use of parameters derived for the entire study area, along with their relationship with average rainfall amount, to characterize the random cascade process. In this study, simulation is done based on different σ^2 values for each rainy month to capture the variability of rain.

5.2.2 Model validation

The developed disaggregation model is validated by checking its effectiveness in reproducing the observed rainfall characteristics based on GCM scenarios. The comparative study is done taking into consideration the following statistical approaches: exceedance probability curves for rainfall amount distribution, spatial autocorrelation to measure the self-similarity of rain field, area distribution of pixel residual. Also, daily average rainfall map for the spatial heterogeneity and comparison of observed, raw GCM simulated rain field with bias-corrected downscaled map are considered to verify the spatial disaggregation model output.

Four meteorological stations are selected from different directions (as shown in Fig. 1) in both grids to develop exceedance probability curves between the observed and downscaled rainfall. The downscaled rainfall is the value for a pixel of size 2x2 km in which the selected station is located. Figure 10 shows the exceedance probability curve for May and September months at selected stations. The model satisfactorily reproduces the observed trend and variation of daily rainfall amount in temporal scale, however, the downscaled rainfall amounts during heavy rainfall events show certain degree of variation.

Geary and Moran indices calculated for the adjacent cell are given in Table 3. Moran's "*I*" values are positive and close to 1 in almost all cases. Similarly, the Geary's "*c*" values are positive and close to 0. The spatial correlation (*R*) values vary from 0.82 to 0.99 for all the rainy months. This is a good indicator of high correlation among the observed and simulated rain field as mentioned in previous section. Moran's "*I*"

isotropic autocorrelogram is also calculated for variable separation distances. Moran's "*I*" is weighted by separation distance (*h*) between sample points rather than by simple adjacency. As shown in Figs. 11a–b, Moran's "*I*" autocorrelograms show high spatial autocorrelation for rainy season in both the grids. In all the months, the Moran's "*I*" begin with high positive values and decrease with corresponding increase in lag distance. This observation implies that the pixel data in both the grids are regionalized, smooth and clustered.

The percent area distribution of spatial residuals (observed-downscaled) from the average monthly observed and simulated rain field is presented in Table 4. The residual range is divided into seven classes. The values corresponding to these classes, as in the table, are the percent area of grid under particular class. Major portion of the two grids is falling in classes 4 and 5. In grid 1, more than 75% of the total area is in these two classes. The inference is that the most of the downscaled values are close to the observed rainfall within the absolute residual of 1 mm/day. The downscaled model over predicts in grid 1 and under predicts in grid 2. There is a strong correlation between magnitude of residual and rainfall amount for both grids. The area with high residual is scattered over all classes in grid 2 compared to grid 1. This is due to high rainfall amount, its variability and less number of stations in grid 2 as compared to grid 1.

The existence of spatial heterogeneity in the output of disaggregation model is verified by comparing the observed and downscaled rain fields. Figure 12 shows the spatial heterogeneity for May and September months. The trend in rainfall distribution pattern is much similar in observed and downscaled rain fields. In grid 2, a large portion of the grid is showing dark colour (heavy rain) without much variation in rainfall pattern. This is the reflection of the lack of data availability (Thailand–Myanmar border) and a larger area of grid 2 is covered by Andaman Sea.

This spatial disaggregation methodology is further applied for downscaling the raw GCM data and the spatial distribution pattern is compared to bias-corrected downscaled GCM. The average downscaled maps of raw GCM and bias-corrected downscaled GCM with their spatial differences for the month of May are shown in Fig. 13a.

The raw GCM downscaled rain field produces more rain in spatial sense compared to bias-corrected rain field. The bias-corrected downscaled map is found closer to the average observed rain field with less residual range variation, as shown in Fig. 13b. The values of residual are found to be more in the region with high rainfall amount.

In next section, these processed GCM precipitation scenarios are used in hydrological model to compare their skill with respect to observed flow.

6 Application: Upper Ping River Basin

A hydrological model is used to simulate the observed flow at a selected station in the basin using processed precipitation scenarios. The United States Army Corps of Engineers (USACE) watershed model HEC-HMS (version 3.0.1) is used as the hydrologic model. The Ping River Basin is modeled with a focus on the upper watershed that provides inflow to the Bhumibol reservoir. The location map of the upper Ping River Basin and Bhumibol reservoir is indicated in Fig. 1. A detail description of HEC-HMS set up can be found in HEC (2005). HEC-GeoHMS v-1.1, as described in HEC (2003), is used to develop the input basin file for the HEC-HMS and it contains the hydrologic and hydraulic parameters of the basin. HYDRO-1K digital elevation model, developed at the U.S. Geological Survey's (USGS), is used to prepare the basin file. HEC-HMS is applied with the deficit and constant loss method keeping the monthly baseflow constant. MODCLARK transformation is used with Standard Hydrologic Grid (SHG) size of $2 \times 2\,km^2$ for incorporating the distributed precipitation data for the model. The distributed precipitation data for Ping Basin are clipped from the bias-corrected downscaled scenarios using *ArcInfo* GIS tool. The program named *asc2dssGrid* is used for integration between data storage system (DSS) and *ArcInfo* through the use of an intermediate ASCII text file.

The HEC-HMS model is calibrated for the year 1999 and is verified for the year 2000 with observed daily inflow at Bhumibol reservoir. The coefficient of determination (R^2), Nash-Sutcliffe efficiency (*EI*) and absolute percentage volume error (*APVE*) are calcu-

lated to evaluate the model performance. The R^2 values for calibration and validation period are found to be 0.71 and 0.76, respectively. The *EI* values for calibration and validation period are found to be 0.65 and 0.75, respectively. The *APVE* for calibration and verification period are found to be 1.9% and 9.3%, respectively.

This calibrated model is then used to simulate the flow for two years (1999-2000) from four modified precipitation scenarios as given in Table 5. Figure 14 shows the average monthly inflow to Bhumibol reservoir for the observed condition and the four scenarios. These simulations illustrate how bias-correction and downscaling methods are able to capture the essential precipitation features required for better simulation of flow in the Ping River. It is observed that the raw GCM precipitation (scenario 1) misses the peak flow period and hydrograph trend. The peak flows for scenario 1 occur in July and August, while it occurs in November for observed flow. Minimum monthly flow occurs in March for all scenarios. According to scenario 1, the simulated flows in three months (June–August) are found to be 4.2 (A2) and 3.5 (B2) times of observed flow. But, scenario 4 reduced it to 1.4 (A2) and 1.3 (B2) times of observed flow. This concludes that the simulated hydrologic flows from the downscaled scenarios are able to capture the peak better than the raw and bias-corrected precipitation scenarios. Scenario 4 (bias-corrected downscaled precipitation) is found to be accurate in simulating the flow peak and trend as compared to other scenarios. This improvement is a result of incorporating the spatial heterogeneity in precipitation, which retains the monthly variability of flow. The large-scale precipitation is able to provide the annual and interseasonal information when it is downscaled at high spatial resolution. The bias-correction and downscaling techniques can profoundly affect the hydrograph, which cannot be simulated with raw GCM precipitation.

7 Conclusions

Precipitation biases and low scale resolution are the two main factors related to application of GCM scenarios to assess the impact of climate change on water resources

at basin level. Gamma-Gamma (GG) transformation and spatial disaggregation model are applied to the ECHAM4/OPYC3 SRES A2 and B2 precipitation scenarios for use in high-resolution impact assessment models. Gamma-Gamma transformation method reduces the biases from raw GCM precipitation whereas the spatial disaggregation model is used in dealing with coarse resolution problem in GCM. In spatial disaggregation model, β values decrease with increase in rainfall while σ^2 values remain almost not sensitive to rainfall amount. The decreasing value of β with increase in average rainfall amount indicates that the fields with more rainfall are less intermittent. The bias-corrected downscaled precipitation scenarios show more realistic hydrologic simulation when compared to observed flow data. The applied approach can be further used for other GCM scenarios to make them more applicable for impact assessment research. Lack of observed rainfall data particularly in Thailand–Myanmar border and larger area covered by Andaman Sea are the two limiting factors for better estimation of disaggregation model parameters. The scattered characteristics of downscaled rain field for heavy rainfall show the need to develop the disaggregation model by considering the different classes of rainfall amount.

Acknowledgements. The authors would like to thank H. Ludhart, Max-Planck-Institut fuer Meteorologie, and the World Data Center for Climate (WDCC) for providing ECHAM4/OPYC climate model datasets.

References

Department of Water Resource: Participatory watershed management for Ping River Basin, Technical Proposal, office of natural resources and environmental policy and planning, Thailand, 2004.

Feenstra, J. F., Burton, I., Smith, J. B., and Tol, R. S. J.: Handbook on methods for climate change impact assessment and adaptation strategies, UNEP and Institute for Environmental Studies, http://www.falw.vu.nl/images_upload/151E6515-C473-459C-85C59441A0F3FB49.pdf, 1998.

Güntner, A., Olsson, J., Calver, A., and Gannon, B.: Cascade-based disaggregation of continuous rainfall time series: the influence of climate, Hydrol. Earth Syst. Sci., 5, 145–164, 2001, http://www.hydrol-earth-syst-sci.net/5/145/2001/.

Gupta, V. K. and Waymire, E. C.: A statistical analysis of mesoscale rainfall as a random cascade, J. Appl. Meteorol., 32, 251–267, 1993.

Gupta, V. K. and Waymire, E. C.: Multiscaling properties of spatial rainfall and river flow distributions, J. Geophys. Res., 95, 1999–2009, 1990.

Hamlet, A. F., Snover, A., and Lettenmaier, D. P.: Climate change scenarios for Pacific Northwest water planning studies: motivation, methodologies, and a user's guide to applications, Technical Document, 1–8, http://www.hydro.washington.edu/Lettenmaier/permanent_archive/hamleaf/bams_paper/technical_documentation.pdf, 2003.

Hydrologic Engineering Center (HEC): Geospatial hydrologic modeling extension HEC-GeoHMS: user's manual (version 1.1), U.S. Army Corps of Engineers, Davis, California, 2003.

Hydrologic Engineering Center (HEC): Hydrologic modeling system HEC-HMS: user's manual (version 3.0.0), U.S. Army Corps of Engineers, Davis, California, 2005.

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

Ines, A. V. M.: GCM bias correction tool (version 0.3a), IRI-Columbia university, New York, USA, 2004.

IPCC (Inter Governmental Panel on Climate Change): Climate Change 2001: Impacts, Adaptation and Vulnerability, Cambridge University Press, http://www.grida.no/climate/ipcc_tar/wg2, 2001.

Jothityangkoon, C., Sivapalan, M., and Viney, N. R.: Tests of a space-time model of daily rainfall in southwestern Australia based on non-homogenous random cascades, Water Resour. Res., 36, 267–284, 2000.

Kahane, J. P. and Peyriere, J.: Sur certaines mertingales de Benoit Mandelbrot, Adv. Math, 22, 131–145, 1976.

Mandelbrot, B. B.: Intermittent turbulence in self-similar cascades: divergence of high moments and dimension of the carrier, J. Fluid Mech., 62, 331–358, 1974.

Marsan, D., Schertzer, D., and Lovejoy, S.: Causal space-time multifractal modeling of rain, J. Geophys. Res., 31, 333–346, 1996.

Molnar, P. and Burlando, P.: Preservation of rainfall properties in stochastic disaggregation by

a simple random cascade model, Atmos. Res., 77, 137–151, 2005.
Nakicenovic, N. and Swart, R.: Emissions Scenarios. 2000, Special Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, 2000.
Olsson, J.: Evaluation of a scaling cascade model for temporal rainfall disaggregation, Hydrol. Earth Syst. Sci., 2, 19–30, 1998, http://www.hydrol-earth-syst-sci.net/2/19/1998/.
Over, T. M. and Gupta, V. K.: Statistical analysis of mesoscale rainfall: dependence of a random cascade generator on large scale forcing, J. Appl. Meteorol., 33, 1526–1542, 1994.
Over, T. M. and Gupta, V. K.: A space-time theory of mesoscale rainfall using random cascades, J. Geophys. Res., 101, 319–331, 1996.
Pathirana, A. and Herath, S.: Multifractal modeling and simulation of rain fields exhibiting spatial heterogeneity, Hydrol. Earth Syst. Sci., 6, 695–708, 2002, http://www.hydrol-earth-syst-sci.net/6/695/2002/.
Roeckner, E., Arpe, K., Bengtsson, L., Christoph, M., Claussen, M., Dümenil, L., Esch, M., Giorgetta, M., Schlese, U., and Schulzweida, U.: The atmospheric general circulation model ECHAM4: model description and simulation of present-day climate, Max-Planck-Institut für Meteorologie, Hamburg, Report No. 218, 1996.
Schertzer, D. and Lovejoy, S.: Physical modeling and analysis of rain and clouds by anisotropic scaling multiplicative processes, J. Geophys. Res., 92, 9693–9714, 1987.
Sharma, D., Gupta, A. D., and Babel, M. S.: Trends in extreme rainfall and temperature indices for two river basins of Thailand, Proc., 3rd APHW International Conference, Bangkok, 2006.
Shrestha, R. K., Tachikawa, Y., and Takara, K.: Downscaling spatial rainfall field from global scale to local scale using improved multiplicative random cascade method, Ann. Disas. Prev. Res. Inst., Kyoto Univ., 47, 2004.
Sokal, R. R. and Oden, N. L.: Spatial autocorrelation in biology. 1. Methodology. 2. Some biological implications and four applications of evolutionary and ecological interest, Biol. J. Linn. Soc., 10, 199–228, 1978.
Tachikawa, Y., Hiwasa, M., and Takara, K.: Spatial rainfall field simulation with random cascade introducing orographic effects on rainfall, Proc., 2nd APHW International Conference, Singapore, 2004.
Wood, A. W., Maurer, E. P., Kumar, A., and Lettenmaier, D. P.: Long range experimental hydrologic forecasting for the Eastern U.S., J. Geophys. Res., 107, 4429, doi:10.1029/2001JD000659, 2002.

Table 1. Performance of bias-correction method on monthly rainfall amount.

	SRES-A2				SRES-B2			
	SD	R	RMSE	d	SD	R	RMSE	d
Grid 1								
Observed	2.48				2.48			
Raw-GCM	3.28	0.32	3.64	0.74	3.47	0.27	3.87	0.47
GG-GCM	2.52	0.66	2.06	0.81	2.30	0.60	2.15	0.76
Grid 2								
Observed	8.48				8.48			
Raw-GCM	4.08	0.34	7.02	0.83	4.42	0.32	7.06	0.61
GG-GCM	8.24	0.73	6.00	0.84	8.20	0.68	6.45	0.81

Table 2. Statistics of estimated parameters for random cascade model.

	Grid 1						Grid 2					
	σ^2			β			σ^2			β		
Month	SD	Mean	CV	SD	Mean	CV	SD	Mean	CV	SD	Mean	CV
May	0.045	0.062	0.002	0.154	0.229	0.024	0.081	0.077	0.007	0.173	0.263	0.030
June	0.044	0.081	0.002	0.130	0.212	0.017	0.063	0.095	0.004	0.107	0.185	0.012
July	0.042	0.077	0.002	0.125	0.201	0.016	0.061	0.104	0.004	0.104	0.176	0.011
August	0.045	0.092	0.002	0.125	0.154	0.016	0.052	0.098	0.003	0.104	0.165	0.011
September	0.043	0.081	0.002	0.136	0.176	0.019	0.068	0.097	0.005	0.120	0.176	0.014
October	0.046	0.063	0.002	0.157	0.228	0.025	0.072	0.076	0.005	0.178	0.266	0.032
Annual	0.084	0.060	0.007	0.158	0.276	0.030	0.075	0.078	0.006	0.171	0.251	0.029
Rainy	0.045	0.076	0.002	0.140	0.201	0.019	0.064	0.094	0.004	0.129	0.192	0.017

Table 3. Spatial statistics of observed and simulated data relative to adjacent cells.

		Grid-1			Grid-2		
Month	Data	Moran's "I"	Geary "c"	R	Moran's "I"	Geary "c"	R
May	Simulated	0.936	0.054		0.974	0.024	
	Observed	0.989	0.001	0.923	0.997	0.001	0.973
June	Simulated	0.976	0.016		0.995	0.006	
	Observed	0.990	0.001	0.946	1.001	0.001	0.993
July	Simulated	0.975	0.018		0.986	0.012	
	Observed	0.991	0.001	0.969	0.998	0.001	0.981
August	Simulated	0.976	0.015		0.984	0.011	
	Observed	0.993	0.002	0.983	0.988	0.003	0.984
September	Simulated	0.967	0.025		0.978	0.020	
	Observed	0.990	0.001	0.972	0.997	0.001	0.972
October	Simulated	0.873	0.123		0.911	0.084	
	Observed	0.991	0.002	0.816	0.992	0.001	0.874

Table 4. Percent area distribution of pixel residual for rainy months (May–October).

		Grid-1						Grid-2					
Class	Residual* [mm]	May	June	July	Aug	Sept	Oct	May	June	July	Aug	Sept	Oct
1	>3	0.5	0.8	4.9	0.1	0.1	0.0	0.2	9.8	12.0	15.2	4.2	0.0
2	2 to 3	1.4	1.6	3.2	0.8	0.7	0.2	1.7	3.9	5.0	4.5	5.9	0.2
3	1 to 2	8.7	2.5	4.5	3.3	5.3	3.1	9.5	5.8	7.4	5.3	0.8	6.3
4	0 to 1	39.0	41.9	17.8	37.3	43.4	46.5	40.5	12.6	19.4	8.2	22.3	44.7
5	–1 to 0	38.9	49.1	58.2	57.9	44.3	47.8	35.8	35.0	29.3	20.7	34.6	41.0
6	–2 to –1	10.7	3.2	10.1	0.7	5.3	2.4	9.9	18.0	11.1	20.4	16.5	7.4
7	< – 3	0.8	0.9	1.2	0.0	0.9	0.0	2.4	14.9	15.8	25.8	5.7	0.5

*Residual = (observed – downscaled)

Table 5. Description of processed precipitation scenarios.

Scenario	Scenario_code	Description
Scenario 1	A2R, B2R	*Raw* ECHAM4/OPYC3 daily precipitation with SRES A2/B2
Scenario 2	A2GG, B2GG	*Bias-corrected* ECHAM4/OPYC3 daily precipitation with SRES A2/B2
Scenario 3	A2R_d, B2R_d	*Raw downscaled* ECHAM4/OPYC3 daily precipitation with SRES A2/B2
Scenario 4	A2GG_d, B2GG_d	*Bias-corrected downscaled* ECHAM4/OPYC3 daily precipitation with SRES A2/B2

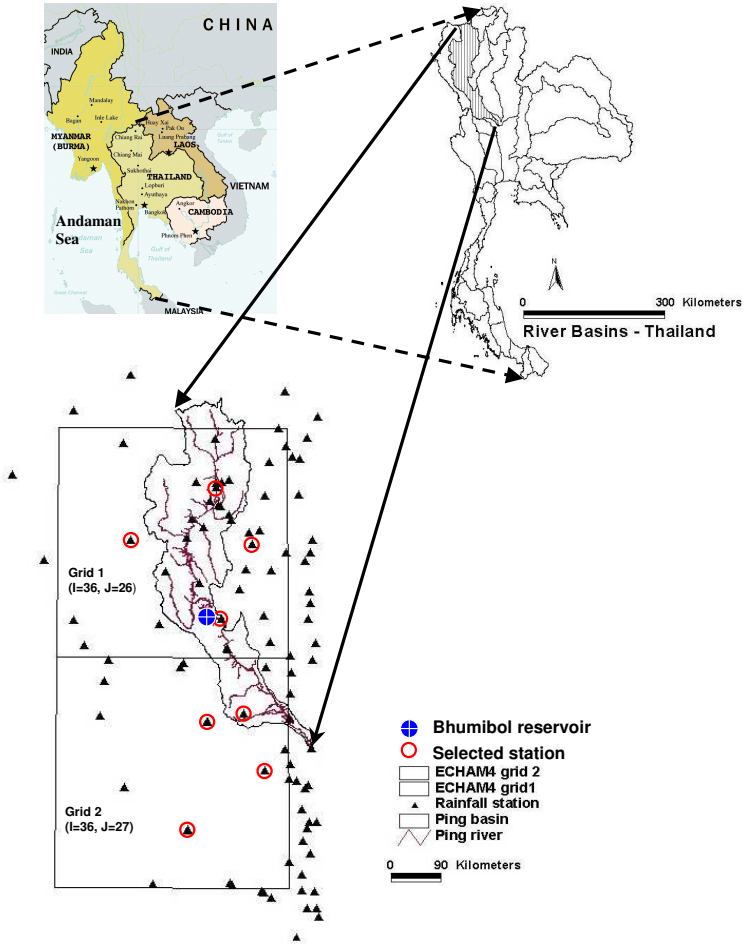

Fig. 1. Location map of Ping River Basin.

HESSD
4, 35–74, 2007

GCM precipitation for improved hydrologic simulation

D. Sharma et al.

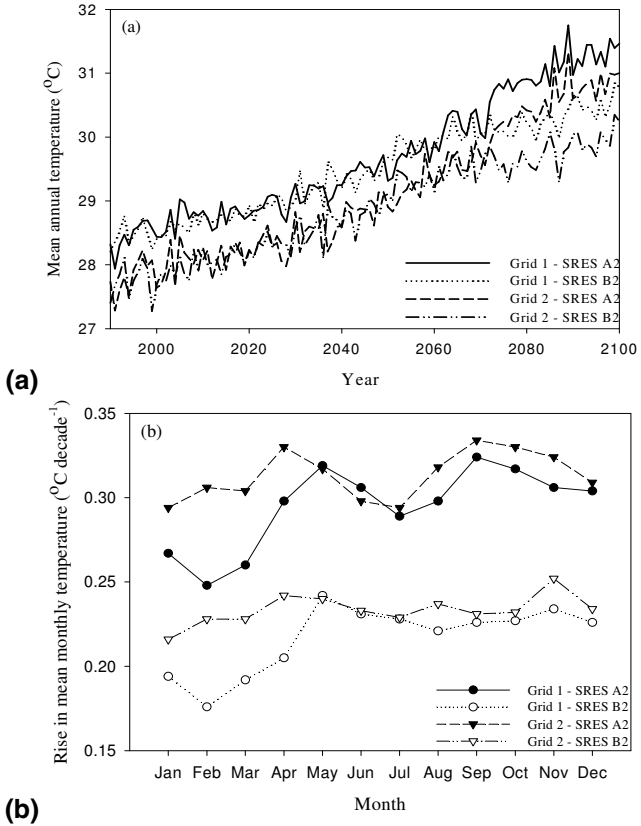

Fig. 2. Trend in ECHAM4 SRES A2 and B2 mean temperature from 1990 to 2100 **(a)** mean annual temperature variation in two grids **(b)** Decadal rise in monthly temperature in two grids.

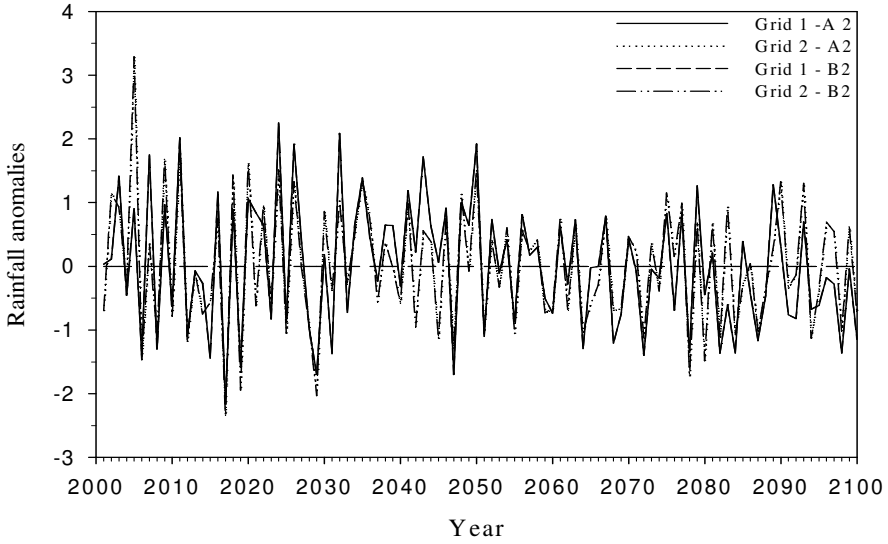

Fig. 3. Annual variability of the standardized rainfall anomalies for 100 years (2000–2100), [(ECHAM4 simulated – mean)/standard deviation].

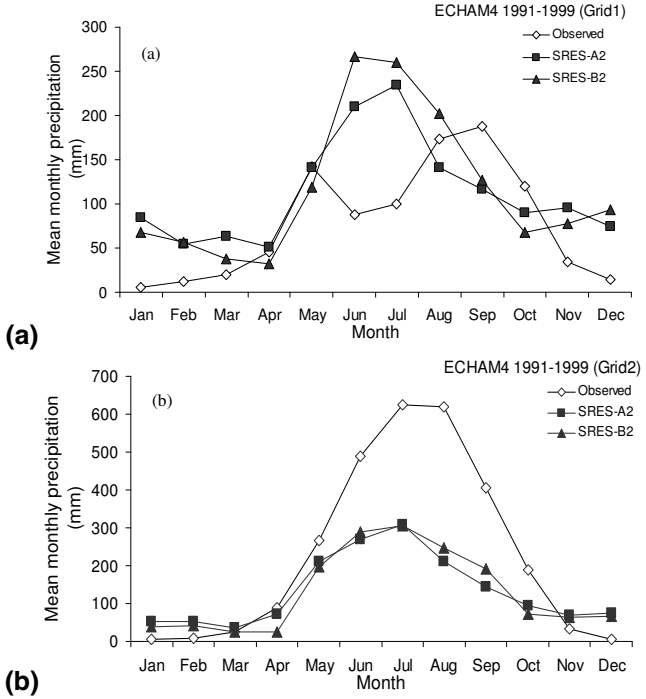

Fig. 4. Mean monthly rainfall trend of ECHAM4 SRES A2 and B2 scenarios compared with observed rainfall data (1990–1999) for **(a)** grid 1, and **(b)** grid 2.

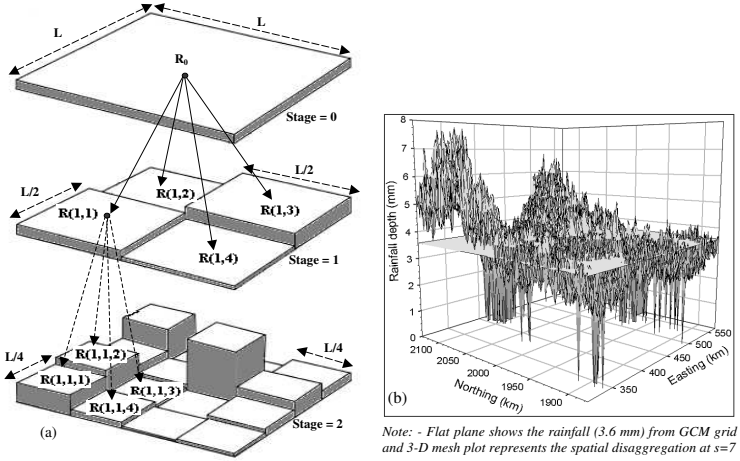

Fig. 5. Spatial Disaggregation model **(a)** representation of disaggregation process at various stages, and **(b)** spatial disaggregation of GCM rain field (horizontal plane) with R_0= 3.6 mm at stage s=7.

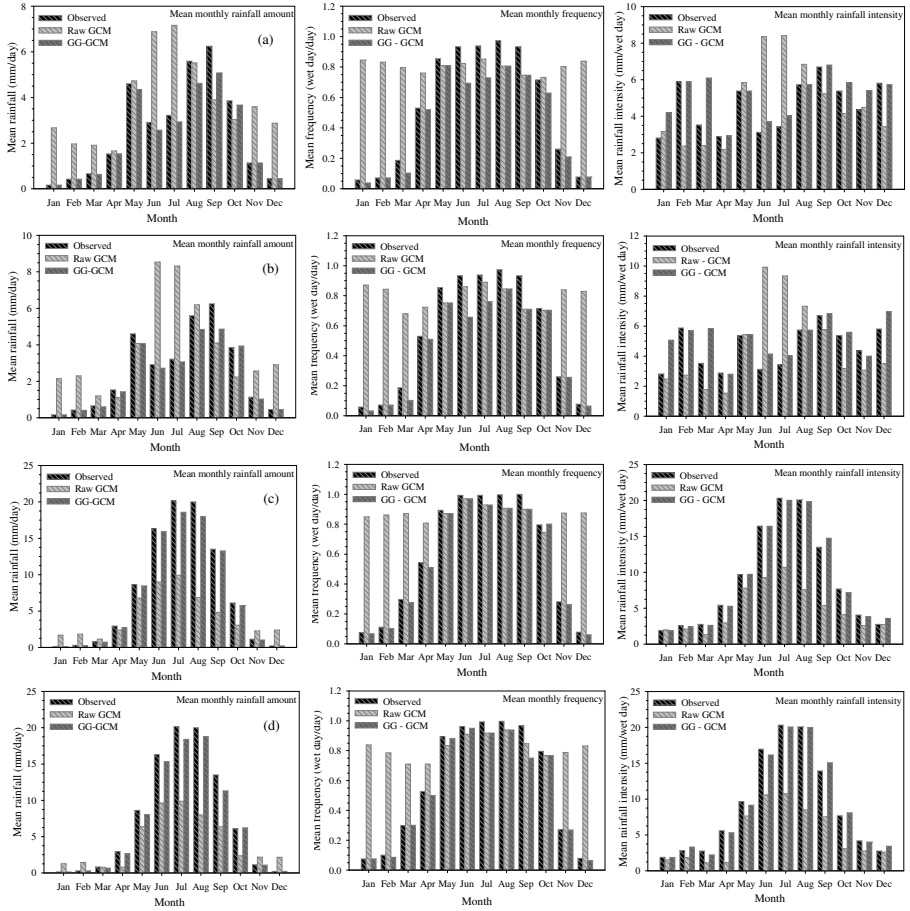

Fig. 6. Comparison of observed, raw GCM and bias-corrected GCM scenarios **(a)** grid 1-A2, **(b)** grid 1-B2, **(c)** grid 2-A2, and **(d)** grid 2-B2.

HESSD
4, 35–74, 2007

EGU

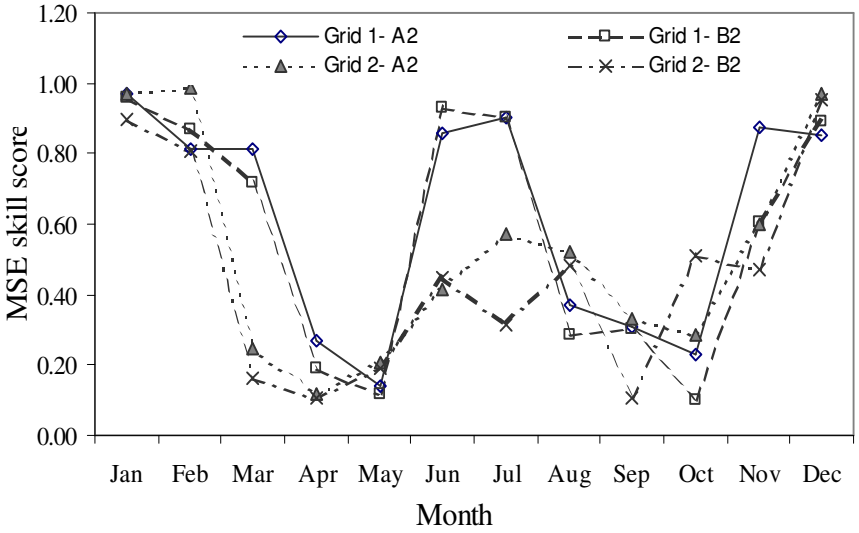

Fig. 7. Monthly variation in MSE skill scores for ECHAM4/OPYC rainfall with bias correction method.

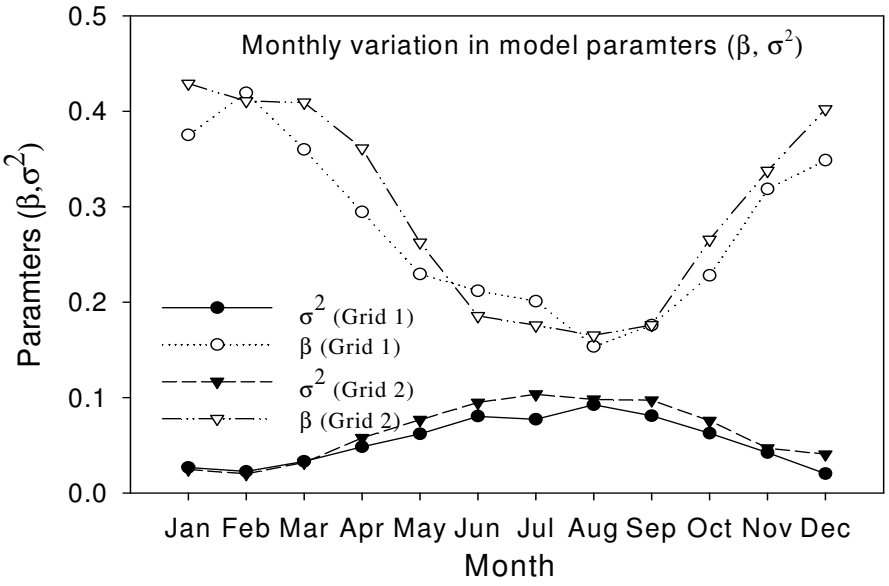

Fig. 8. Monthly variation in average spatial disaggregation model parameters (β, σ^2).

HESSD
4, 35–74, 2007

EGU

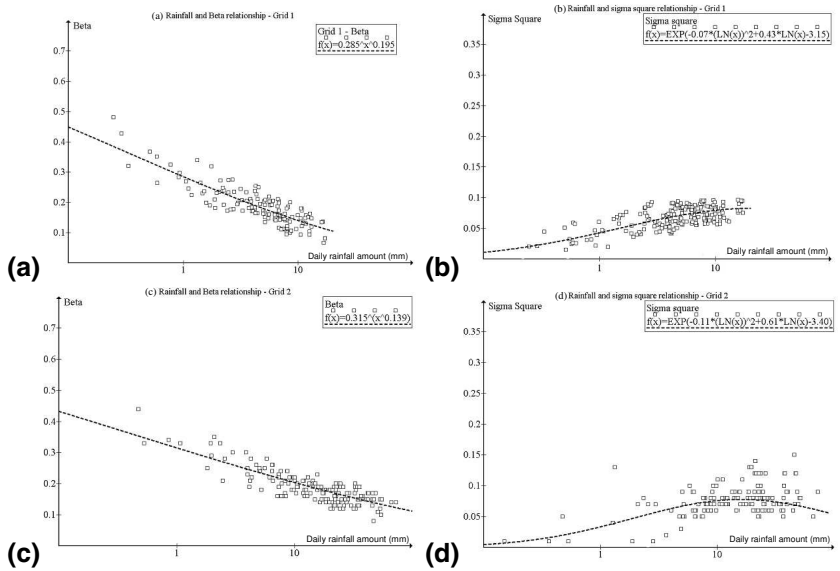

Fig. 9. Graph presents the relation of model parameters (β, σ^2) with average rainfall **(a)** grid 1-β, **(b)** grid 1-σ^2, **(c)** grid 2-β, and **(d)** grid 2-σ^2.

HESSD
4, 35–74, 2007

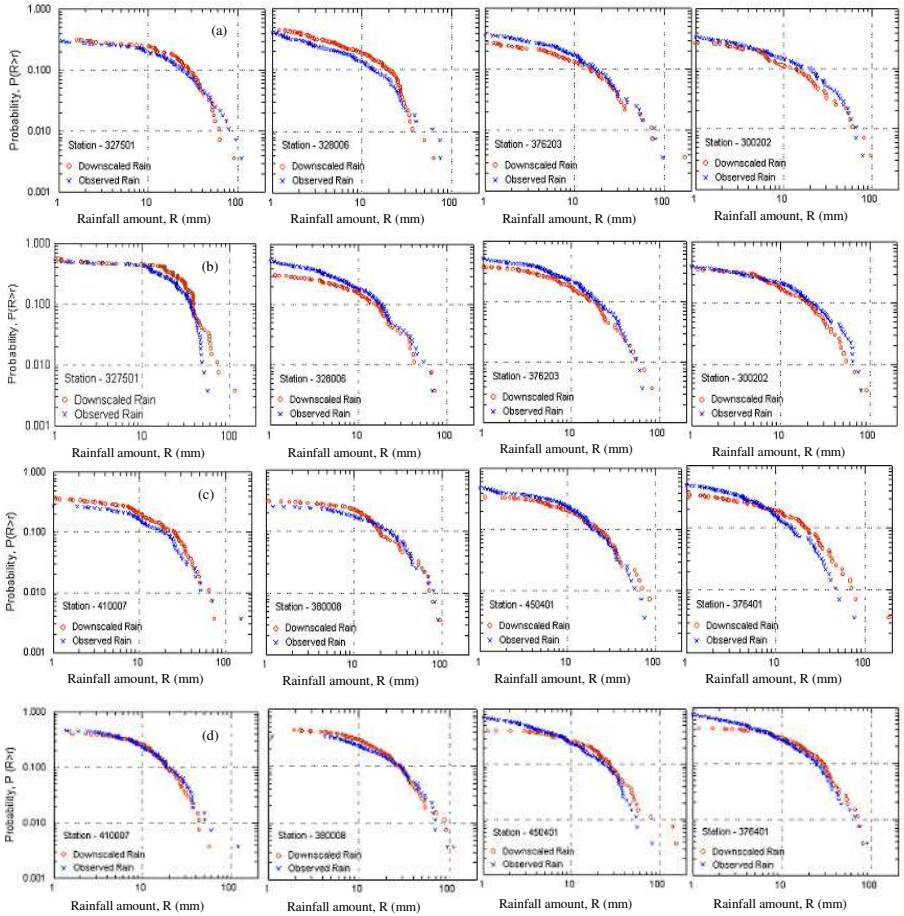

Fig. 10. Exceedance probability curve of rainfall amount at four rainfall stations from observed and downscaled data **(a)** grid 1-May, **(b)** grid 1-September, **(c)** grid 2-May, and **(d)** grid 2-September.

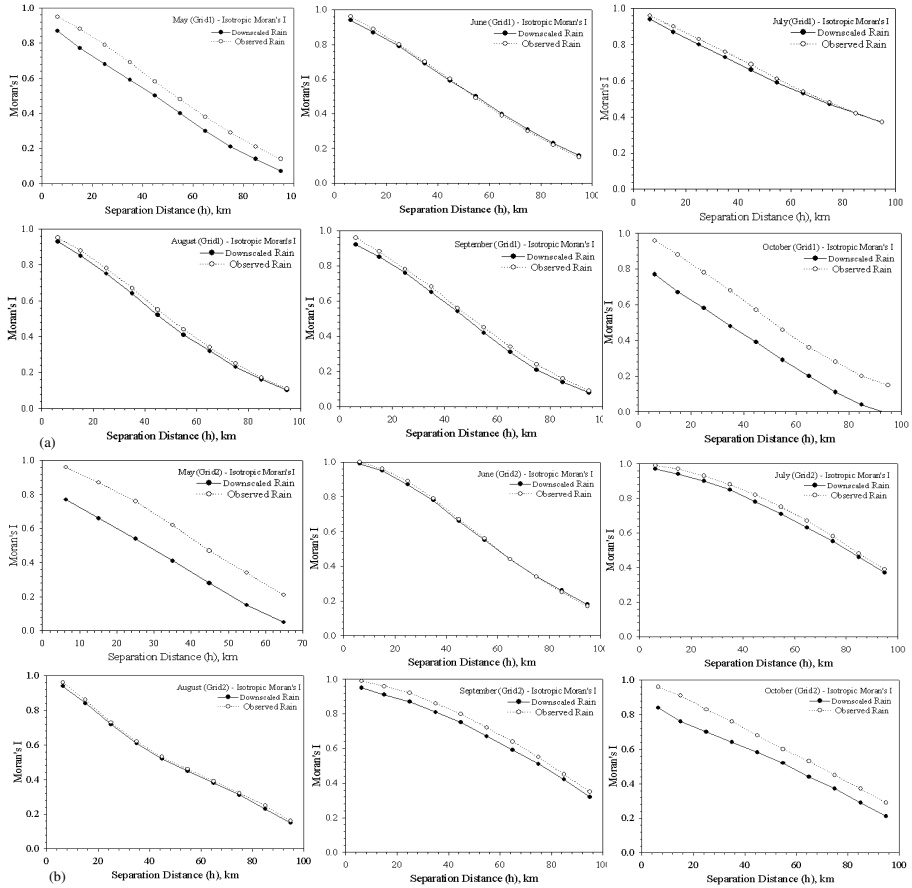

Fig. 11. Moran's '/' isotropic autocorrelogram between ECHAM4 SRES A2 downscaled precipitation rain field and observed rain field for **(a)** grid 1-wet period, and **(b)** grid 2-wet period.

HESSD
4, 35–74, 2007

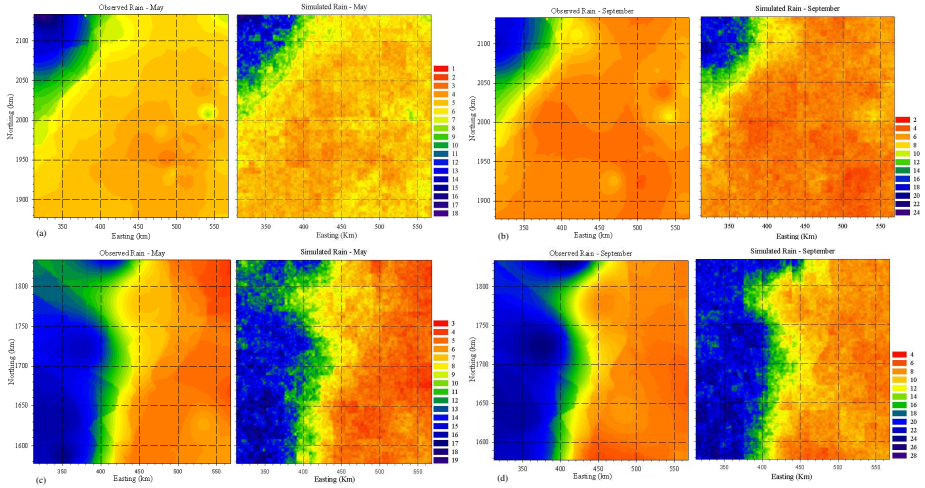

Fig. 12. Spatial heterogeneity between observed and simulated rain field **(a)** grid 1- May month, **(b)** grid 1- September month, **(c)** grid 2- May month, **(d)** grid 2- September month.

HESSD
4, 35–74, 2007

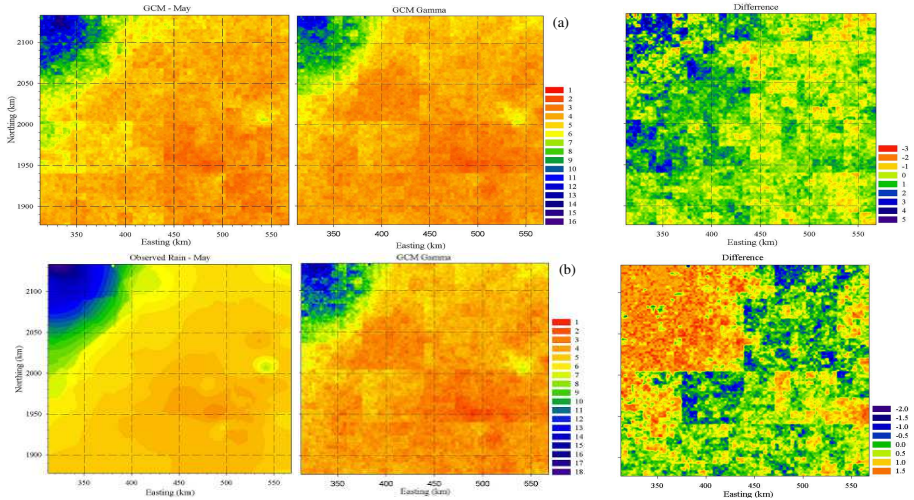

Fig. 13. Spatial distribution pattern shows **(a)** average downscaled maps of raw GCM and bias-corrected downscaled GCM **(b)** average observed rain field and bias-corrected downscaled map.

HESSD
4, 35–74, 2007

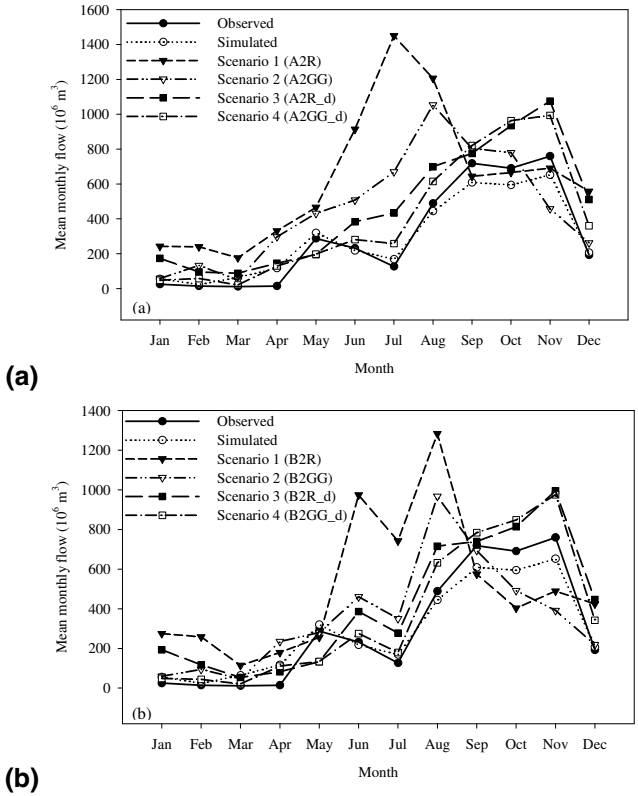

Fig. 14. Average monthly inflow to Bhumibol reservoir for the observed condition and the four modified precipitation scenarios **(a)** SRES A2, and **(b)** SRES B2.