Manuscript Number: HYDROL17235R2

Title: Comparing Statistically Downscaled Simulations of Indian Monsoon at different spatial Resolutions

Article Type: Research Paper

Keywords: Statistical Downscaling; Climate Change; Precipitation; Bayesian Averaging

Corresponding Author: Dr. Subimal Ghosh, Ph.D.

Corresponding Author's Institution: IIT Bombay

First Author: K Shashikanth, M Tech

Order of Authors: K Shashikanth, M Tech; C G Madhusoodhanan; Subimal Ghosh, Ph.D.; T I Eldho, PhD; K Rajendran, PhD; Raghu Murtugudde, PhD

Abstract: Impacts of climate change are typically assessed with fairly coarse resolution General Circulation Models (GCMs), which are unable to resolve local scale features that are critical to precipitation variability. GCM simulations must be downscaled to finer resolutions, through statistical or dynamic modelling for further use in hydrologic analysis. In this study, we use a linear regression based statistical downscaling method for obtaining monthly Indian Summer Monsoon Rainfall (ISMR) projections at multiple spatial resolutions, viz., 0.050, 0.250 and 0.500, and compare them. We use 19 GCMs of Coupled Model Intercomparison Project Phase 5 (CMIP5) suite and combine them with multi model averaging and Bayesian model averaging. We find spatially non-uniform changes in projections at all resolutions for both combinations of projections. Our results show that the changes in the mean for future time periods (2020s, 2050s, and 2080s) at different resolutions, viz., 0.050, 0.250 and 0.50, obtained with both Multi-Model Average (MMA) and Bayesian Multi-Model Average (BMA) are comparable. We also find that the model uncertainty decreases with projection times into the future for all resolutions. We compute signal to noise ratio (SNR), which represents the climate change signal in simulations with respect to the noise arising from multi-model uncertainty. This appears to be almost similar at different resolutions. The present study highlight that, a mere increase in resolution by a way of computationally more expensive statistical downscaling does not necessarily contribute towards improving the signal strength. Denser data networks and finer resolution GCMs may be essential for producing usable rainfall and hydrologic information at finer resolutions in the context of statistical downscaling.

Comparing Statistically Downscaled Simulations of Indian Monsoon at different spatial Resolutions

**K.Shashikanth^{1,2}, C.G. Madhusoodhanan¹, Subimal Ghosh^{1,3*}, T.I.Eldho^{1,3}, K.Rajendran⁴,
Raghu Murtugudde^{3,5}**

¹ Department of Civil Engineering, Indian Institute of Technology Bombay, Mumbai,
-400076, India

² Department of Civil Engineering, University College of Engineering, Osmania University,
Hyderabad-500007, India

³ Interdisciplinary Programme in Climate Studies, Indian Institute of Technology, Bombay,
Mumbai- 400076. India.

⁴ CSIR Fourth Paradigm Institute (CSIR-4PI), Bangalore, India 560037

⁵ Department of Atmospheric and Oceanic Science, Earth System Science Interdisciplinary Center, University of Maryland College Park, MD 20742, USA.

*Corresponding Author: Department of Civil Engineering, Indian Institute of Technology
Bombay, Powai, Mumbai - 400 076, India. Email: subimal@civil.iitb.ac.in

Phone: +91 22 2576 7319, Fax: +91 22 2576 7302

Comparing Statistically Downscaled Simulations of Indian Monsoon at different spatial Resolutions

K.Shashikanth, C.G. Madhusoodhanan, Subimal Ghosh, T.I.Eldho, K.Rajendran,

Raghu Murtugudde

ABSTRACT

Impacts of climate change are typically assessed with fairly coarse resolution General Circulation Models (GCMs), which are unable to resolve local scale features that are critical to precipitation variability. GCM simulations must be downscaled to finer resolutions, through statistical or dynamic modelling for further use in hydrologic analysis. In this study, we use a linear regression based statistical downscaling method for obtaining monthly Indian Summer Monsoon Rainfall (ISMR) projections at multiple spatial resolutions, viz., 0.05° , 0.25° and 0.50° , and compare them. We use 19 GCMs of Coupled Model Intercomparison Project Phase 5 (CMIP5) suite and combine them with multi model averaging and Bayesian model averaging. We find spatially non-uniform changes in projections at all resolutions for both combinations of projections. Our results show that the changes in the mean for future time periods (2020s, 2050s, and 2080s) at different resolutions, viz., 0.05° , 0.25° and 0.5° , obtained with both Multi-Model Average (MMA) and Bayesian Multi-Model Average (BMA) are comparable. We also find that the model uncertainty decreases with projection times into the future for all resolutions. We compute signal to noise ratio (SNR), which represents the climate change signal in simulations with respect to the noise arising from multi-model uncertainty. This appears to be almost similar at different resolutions. The present study highlight that, a mere increase in resolution by a way of computationally more expensive statistical downscaling does not necessarily contribute

towards improving the signal strength. Denser data networks and finer resolution GCMs may be essential for producing usable rainfall and hydrologic information at finer resolutions in the context of statistical downscaling.

Key Words: Statistical downscaling, Bayesian Multimodel Average, Uncertainty, Signal to Noise Ratio (SNR)

1. Introduction

India receives more than 80% of its total rainfall during the four summer monsoon months; June to September (Webster et al., 1998; Menon et al., 2013). Agriculture, water supply, industrial production and allied activities depend on the Indian summer monsoon rainfall (ISMR hereafter) and hence reliable regional projection of ISMR, under climate change is necessary for adaptation and policy framing. Impacts of climate change are assessed by General Circulation Models (GCMs), which are at present the most reliable tools to simulate the response of the global climate system to increasing the green house gas concentrations (Prudhomme et al., 2002; IPCC, 2007). Since GCMs typically operate at coarse resolutions, they often fail to capture the finer scale processes that affect regional climate. Hence, the outputs from GCMs cannot be directly used in the regional impact assessment studies, specifically in hydrologic applications (Salvi et al., 2013).

The new generation climate models in the CMIP5 (Coupled Model Intercomparison Project Phase 5) suite have undergone several improvements in terms of physics and resolutions (Taylor et al., 2012). Still, outputs from CMIP5 models cannot be directly used for regional scale projections because of the lack of regional specificity in their simulations due to the relatively coarse resolutions. Such coarse resolution projections must be downscaled to obtain

regional/local scale projections. They can be obtained either by statistical or dynamic downscaling approaches. In dynamical downscaling, a higher resolution regional climate model (e.g., RegCM, Limited Area Model (LAM), PRUDUNCE) is embedded within a GCM, which provides large scale boundary conditions. Such regional models also incorporate sub-grid/local scale features to deliver high resolution outputs (Fowler et al., 2007; Maraun et al., 2010). The main limitation of dynamic downscaling method is the requirement of huge computational resources (Wilby and Wigley, 1997). In addition to the limitations of the models themselves in capturing the regional and local land, atmosphere, and ocean interactions (see for e.g., Liang et al., 2012).

Statistical downscaling belongs to the class of data driven approaches, which not only overcomes the limitation of computational requirement but also can be applied to any region of interest given sufficient data for constructing the statistical model. Statistical methods are thus predominantly used when sufficient lengths of historical data are available for establishing the requisite stable statistical relationships (Wilby et al., 2004). The philosophy of statistical downscaling method is based on the development of a relationship between predictors (coarse resolution climate variables/ synoptic scale circulation patterns) and a predictand (e.g., rainfall). They are classified into weather generators (Wilks, 1999; Khalili et al., 2009) weather typing/ classification (Corte-Real et al., 1995) and transfer functions (Ghosh and Mujumdar, 2009). Statistical downscaling methods build different types of empirical relationships (like parametric/nonparametric and/or linear/nonlinear) between large-scale atmospheric predictor variables and the regional climate variable(s) of interest (predictand) (Wilby et al., 2004; Kannan and Ghosh, 2013). In statistical downscaling, the regional climate is assumed to be conditioned by the large-scale climate predictors/variables and regional/local physiographic and region-

specific processes (Von Storch, 1995, 1999; Kannan and Ghosh, 2013; Dobler and Ahrens, 2011). The statistical downscaling methodology has been extensively used for regional applications all over the globe. Complete and comprehensive reviews of downscaling techniques, their advantages and limitations, recent developments can be found in the Maraun et al. (2010), Fowler et al. (2007), Schoof (2013). The various statistical downscaling approaches are summarized in the Table 1.

The results of regional projections are always prone to uncertainties stemming from multiple sources. The main sources of uncertainty are model uncertainty, uncertainty from anthropogenic/scenario forcing factors and natural variability (Hawkins and Sutton, 2009; Maraun et al., 2010; Woldsmakel et al., 2012; Yip et al., 2011). The use of multiple downscaling models may also result in downscaling uncertainty (Dibike and Coulibaly, 2005; Ghosh and Katkar, 2012; King et al., 2012). Model uncertainty is the main source of uncertainty and mainly attributed to incomplete understanding of geophysical processes and limitations in solving numerical schemes, parameterizations, and resolutions (Stainforth et al., 2007; Räisänen, 2007). Anthropogenic forcings and scenarios describe how plausible trajectories of different aspects of future climate includes emissions, socio-economic assumptions, energy use, population, air pollutants, land use and consequent radiative forcing and temperature anomalies (Wayne, 2013). The internal uncertainty is due to chaotic climate system and natural forced variability is in fact considered 'unforced' variability since the only source of energy is the solar radiation (Maraun et al., 2010). It is also clear now that higher latitudes are dominated by natural variability and the anthropogenic climate signal may become distinctly identifiable only in the tropics over the next several decades (Deser et al., 2012). Even in the tropics, the relatively small decorrelation scales of precipitation compared to temperatures offer severe challenges to the state-of-the-art GCMs in

delivering reliable projections of rainfall variability. It is thus important to understand the regional and local uncertainties in these projections by combining statistical approaches to extract regional renditions of future projections from these models. This is the main motivation of our study.

The uncertainty stemming from various formulations can be assessed by ensembles of simulations based on different GCMs, or ensembles of particular GCMs with different initial conditions or with perturbed physics. The important roles played by these sources of uncertainty depend on the time scales. In the decades of the immediate future, the projections are argued to be dominated by model uncertainties and internal variabilities whereas in the following decades, scenario uncertainties are significantly larger (Hawkins and Sutton, 2011; Yip et al., 2011; Deser et al. 2012). However, the internal variability increases at smaller spatial scales and shorter time scales (Hawkins and Sutton, 2009). Prudhomme and Davies (2009) reported that uncertainties due to downscaling methods and emission scenarios are of similar magnitude, and generally smaller than GCM uncertainty for the 2080 time horizon. The uncertainties associated with the use of multiple models have been studied by many researchers (Ghosh and Mujumdar, 2009; Watanabe et al., 2012; Woldemeskel et al., 2012). Uncertainty across different GCMs has been measured by a metric/ skill known as Coefficient of variation (CV) (Watanabe et al., 2012; Shashikanth et al., 2013) where higher CV denotes high uncertainty. The CV is defined as the ratio of standard deviation to the mean and it is a dimensionless quantity. Shashikanth et al. (2013) measured the uncertainty resulting from the usage of multiple GCMs from both CMIP3 and CMIP5 suites for ISMR simulations for the future time horizon (2081-2100) and reported that statistically downscaled simulations of CMIP5 show lower uncertainty compared to those of CMIP3.

The projections of climate variables are now presented in a probabilistic framework (Raftery et al., 2005). This has led to the development of new methods like Bayesian Multi-model Averaging (BMA) (Tebaldi et al., 2004, 2005) which is a better alternative to Multi-Model Averaging (MMA), where equal weights are given to all GCMs for both past and future projections. In the BMA, differential weights are assigned to the GCMs, based on performances during the historic period and convergence (agreement with majority of ensemble members in the projection period). Miao et al (2013, 2014) applied the BMA technique for temperature simulations in the different land regions of globe and in the Northern Eurasia and showed that BMA technique is superior to MMA.

Giorgi and Mearns (2002) have developed REA (Reliability Ensemble Average) method for estimating the probability of regional climate change exceeding a given threshold based on ensembles of different AOGCM model simulations. The method takes into account two reliability criteria as mentioned above; model performance for the present climate and convergence of the simulated future changes by all models (convergence criterion). Tebaldi et al. (2004, 2005) developed a Bayesian Multi-Model Averaging (BMA) approach to determine Probability Density Function (PDF) of temperature change from the outputs multi-model ensembles.

Duan and Philips (2010) applied the BMA technique for the estimation of weights assigned to individual historical simulations and future climate projections. In their study, the authors used 17 GCMs from the CMIP3 suite. They also computed the climate change signal from multimodel projections and the Signal to Noise Ratio (SNR), where noise refers to the uncertainty resulting from the disagreement between model simulations. Accurate regional projections are argued to

be a key research challenge in climate change science (Schiermeier, 2010), and in the recent decade, efforts have been made for finer resolution projections. For dynamic downscaling, the key challenge lies in huge computational requirements, while for statistical downscaling, the requirement is the availability of finest resolution data. With the improvements in computational power, as well as with the availability of newer fine resolution gridded products, the resolution for regional modelling is getting enhanced; however, the science question, which remains unaddressed, is whether there is a corresponding improvement in the skills of simulations with the improvements in resolution. It is also important to understand the implications of using higher resolutions on the model uncertainty and specifically, how the SNR changes with the change in resolution. The improvements in ensemble simulations with BMA when compared with that of Multi-Model Averaging (MMA) may also change with the change in resolutions and such evaluations have not been reported in literature. Here, in this present study, we attempt to address these questions with the case study application to the Indian Summer Monsoon Rainfall (ISMR). We apply statistical downscaling, at multiple resolutions, viz., 0.05° , 0.25° and 0.50° for entire India. We use 19 GCMs from the CMIP5 suite, which is a fairly large number of models to project rainfall. The performances of the models are evaluated at multiple scales, as well as with a common scale, where the downscaling is followed by re-gridding/ interpolation. PDF based skill scores are used for evaluation (Perkins et al., 2007) of models. Both MMA and BMA are computed and compared at different resolutions. We also evaluate the models for their SNR (Duan and Philips, 2010) at different resolution to understand the impacts of resolution on SNR. The next section presents the details of the data used for this study.

2.0 Data

A statistical downscaling model (Fig. 1) first develops an empirical relationship between synoptic scale circulation patterns (predictors) and the local variable of interest (predictand). The relationship is then further applied to the GCM simulations of the future to project the regional climate scenarios. Here, we use the National Centre for Environmental Prediction/ National Centre for Atmospheric Research (NCEP/ NCAR) reanalysis data as predictors and the gridded rainfall as the predictand. For calibration of the model, we use the gridded rainfall data, provided by the Asian Precipitation Highly Resolved Observational Data Integration towards Evaluation of Water Resources (APHRODITE) at multiple resolutions.

The selection of predictors is a crucial step in developing a statistical downscaling model. The output of the statistical downscaling method is sensitive to the choice of predictor variables. The criteria used for selections of predictors are that they should be reliably simulated by GCMs, available in the GCM archive and physically associated with the variable of interest (Wilby et al., 2004). Considering these criteria, the predictors used for ISMR downscaling (also used in Salvi et al. (2013); Kannan and Ghosh (2013) and Shashikanth et al. (2013)) are air temperature, wind velocities (U and V wind), specific humidity at both 500 hPa pressure level and the surface and the Mean Sea Level Pressure (MSLP). These predictors are based on the study by Shashikanth et al. (2013) for ISMR.

The NCEP/ NCAR reanalysis data (Kalnay et al.,1996), used for the predictors, are extracted for latitudes 5°N- 40°N and longitude 60° E-120 ° E, encompassing the entire India. This is a global gridded product and continually updated through data assimilation to judiciously blend physical observations and model simulations dating back to 1948. The Reanalysis data is a surrogate for observed data for any predictor variable. The resolution of NCEP/NCAR reanalysis product is 2.5° lat × 2.5° long. For downscaling purposes, India is first subdivided into seven

meteorological subdivisions as per India Meteorological Department (IMD) categorization (Parthasarathy et al., 1996) and the spatial extent of the predictors for different zones are different (Fig. 2). They are selected based on the correlation between the spatially averaged rainfall and gridded predictors (Salvi et al., 2013).

The gridded rainfall data, provided by APHRODITE (Yatagai et al., 2012) is used as the predictand, and is obtained at monthly intervals for the spatial resolution 0.5° , 0.25° and 0.05° . They are based on (i) GTS (Global Telecommunication system) data (ii) data compiled by the organizations from the respective countries like, India Meteorological Department (IMD) for India, and (iii) APHRODITE's own data collection system with proper quality control (Yatagai et al., 2012). For historic and future simulations, we selected 19 GCMs, from CMIP5 archives, and obtained the outputs at monthly time scale. The list of GCMs is provided in Table 2. The GCM simulated predictors are obtained from the Program for Climate Model Diagnosis and Intercomparison (PCMDI). For future rainfall projections of ISMR, Representative Concentration Pathways 4.5 (RCP4.5) is considered for the present work. RCP 4.5 represents a stabilization scenario where the total radiative forcing is stabilized at 4.5 W/m^2 at 2100, by employing a range of technologies and strategies to reduce greenhouse gas emissions (Wise et al., 2009).

3.0 Methodology

In the present analysis, we first downscale the monthly rainfall to 0.50° , 0.25° and 0.05° spatial resolutions, with 19 GCMs. The resulting uncertainty is modelled with Bayesian modelling and the SNR is evaluated for each of the resolution. Brief overview of downscaling, uncertainty modelling and estimation of SNR are discussed in the following subsections.

3.1 Statistical Downscaling

The statistical relationship, developed between the reanalysis predictors and gridded rainfall is based on a linear regression, similar to the “Statistical DownScaling Model” (SDSM), developed by Wilby and Dawson (2004). The methodology is presented in Fig 1 and is applied to the seven meteorologically homogeneous zones (Parthasarathy et al., 1996) over India. The gridded multi-dimensional predictors are highly correlated amongst each other and together form high dimensional covariates. This leads to multicollinearity in a linear regression. As the present model is developed for monthly monsoon rainfall and the sample size of calibration set length is small, and the model may over fit the data with high dimensional predictors. Hence, we apply Principal Component Analysis (PCA) to reduce dimensionality by retaining the major fraction of the variability of the predictor field. Here, we use first few principal components (PCs) which together represent 90% of the variability of original predictors. The principal components (PCs) thus obtained are regressed with the rainfall, at individual grids. The set of predictors and hence the principal components, used as regressors, remain unchanged for all the grids in a homogeneous region. The model is calibrated for the period 1970-2005, for all three spatial resolutions. Applying the principal directions obtained from reanalysis for computing the PCs, to the GCM output may introduce new bias. However, this is a standard procedure (Wilby et al., 2004), which is followed for the GCM outputs, to apply the same relationship, which is obtained between reanalysis predictors and observed predictand. Obtaining the individual principal direction from each GCM will not ensure this and hence is not followed. A possible solution would be merging of reanalysis and GCM predictors for obtaining the PCs, and is a potential area of future research.

The calibrated model is then applied to the GCM simulations after bias correction. GCM simulations are first regridded to the reanalysis grids, and then Wilby et al. (2004) method of standardization is applied. The principal directions obtained from the reanalysis data is applied to the bias corrected GCM predictors for obtaining the principal components (PCs) corresponding to GCMs. These PC are then used in the calibrated linear regression model to obtain gridded rainfall simulations for the historic (here, we consider, 1970-2005) and future (2010-2099) periods (RCP 4.5 scenario).

3.2 Evaluating the GCM Simulations

The downscaled GCM simulations are first evaluated based on the deviations from the long term mean and standard deviations for the historic period. However, such evaluations may not compare the observed and projected simulations in terms of all the statistical characteristics and variability. Hence, here we apply the evaluation metric, suggested by Perkins et al., (2007), where the skill of the simulation is measured with respect to the resemblance of its Probability Density Functions (PDFs) to that of observed.

In this evaluation method, first the entire range of observed/ simulations are divided into multiple bins. This methodology calculates the cumulative minimum value between the model and observed data distribution in each of bin; thereby measuring the common area between two PDFs. If a model simulates the observed data perfectly, the skill score is equal to 1. If it simulates the data poorly, the skill score is equal to zero. Mathematically, the Skill score (S-score is given by)

$$\text{S-score} = \sum_{i=1}^n \text{minimum}(Z_m, Z_o) \quad (1)$$

Where n= total number of bins.

271 Z_m = Frequency values in a given bin from model m.

272 Z_o = Frequency values in a given bin from observed data.

273 The S-score is computed for each grid, and then the spatial distribution of this score is obtained
274 in terms of the PDF. The PDFs of S-scores, obtained from multiple resolutions are compared.

275 **3.3 Uncertainty Modelling with Bayesian Approach**

276 Conventional approach for simulating projections with multiple models is to obtain the multi-
277 model average, assuming that the positive and negative errors in the models will cancel each
278 other out (Knutti, 2010), resulting in more reliable projections. Following Duan and Philips
279 (2010), here we briefly discuss the steps for assignment of weights to GCMs, with the Bayesian
280 Model Averaging (BMA). The parameters to be determined are $\theta = \{w_k, \sigma_{PK}^2, \sigma_{FK}^2\}$, where
281 w_k are the weights to each GCM (w_k for k^{th} GCM), σ_{PK}^2 are the variance for each of the
282 past/historical downscaled GCMs and σ_{FK}^2 are the variance for each of downscaled GCM's
283 future projections.

284 **Step 1:** Let us assume the vectors $y = (Y_P, Y_F)$ denote the downscaled data set consisting of past-
285 climate (Y_P) and simulated climate variable for the future (Y_F) from each of the GCMs. The
286 two data sets are assumed to be independent of each other and follow Gaussian distributions.

$$287 \quad g(y|\theta_k) = g(y_P, y_F | \theta_{Pk}, \theta_{Fk}) = g(y_P | \theta_{Pk}) \cdot g(y_F | \theta_{Fk}) \quad (2)$$

288 Where θ_{Pk}, θ_{Fk} denote parameters for past and future downscaled simulated GCM output. The
289 above equation is based on Bayes' theorem.

290 **Step 2:** The log likelihood function for equation 2, becomes

$$l(\theta) = \sum_{(s)} \log(\sum_k w_k \cdot g(Y_{TP} | \theta_{PK}) \cdot g(Y_{TF} | \theta_{FK})) \quad (3)$$

Where Y_{TP} and Y_{TF} are evidentiary targets (observed) and weighted multimodel consensus estimates of climate variable for the future, respectively.

Since $g(Y_P | \theta_P)$ and $g(Y_F | \theta_F)$ are Gaussian functions, the maximum likelihood estimates of the means μ_k of the individual model simulations are equal to the mean of the simulations themselves; i.e., $\mu_{PK} = f_{PK}$ and $\mu_{FK} = f_{FK}$. The f_{PK} and f_{FK} are the means of downscaled simulations of the climate variable for the past and future in each GCM.

Step 3: The parameters $\theta = \{w_k, \sigma_{PK}^2, \sigma_{FK}^2\}$ in equation 3 are estimated using Expectation-Maximization (E-M) Algorithm.

3.4 Expectation-Maximization Algorithm

EM algorithm is an iterative method and consists of two steps; the Expectation (E) step and the Maximization (M) step.

It begins with the initial guess of parameter $\theta^{(0)}$ for the parameter vector $\theta = \{w_k, \sigma_{PK}^2, \sigma_{FK}^2\}$. In the E step, a quantity Z_{kst} estimated given present values of parameters; however Z_{kst} takes two values; either 0 or 1. Where $Z_{kst}=1$ if ensemble member k is best forecast for particular (s, t) . Otherwise $Z_{kst}=0$;

Thus, the E-step yields,

$$Z_{kst}^{(i)} = (g(Y_{TP} | \theta_{PK}^{(i)}) \cdot g(Y_{TF}^{(i)} | \theta_{PK}^{(i)}) / \{\sum_k g(Y_{TP}^{(i)} | \theta_{PK}^{(i)}) \cdot g(Y_{TF}^{(i)} | \theta_{FK}^{(i)})\}) \quad (4)$$

While the M-step leads to,

$$w_k^{(i)} = \sum_s Z_{ks}^{(i)} / S; \quad (5)$$

Where (i) and s are iterations and spatial points, respectively, S is the total number of spatial points. For further details refer to Duan and Philips (2010) and Raftery et al. (2005).

3.4.1 BMA consensus Expectation and uncertainty estimation

The BMA expected mean of posterior PDFs for the past and future climates are estimated as a weighted sum of individual model simulations and are calculated as below:

$$E(Y_P | f_{p1}, f_{p2}, \dots, f_{pk}) = \sum_k w_k \cdot E(Y_P | f_{pk}) = \sum_k w_k \cdot f_{pk} \quad (6)$$

$$E(Y_F | f_{F1}, f_{F2}, \dots, f_{FK}) = \sum_k w_k \cdot E(Y_F | f_{FK}) = \sum_k w_k \cdot f_{FK} \quad (7)$$

The total variances are the measures of spreads (uncertainties) of the posterior PDFs and are given by the following expressions:

$$\sigma^2(Y_P | f_{p1}, f_{p2}, \dots, f_{pk}) = \sum_k w_k \cdot (\sum_k f_{PK} - \sum_k w_k \cdot f_{PK})^2 + \sum_k w_k \cdot \sigma^2(Y_{TP} | f_{PK}) \quad (8)$$

$$\sigma^2(Y_F | f_{F1}, f_{F2}, \dots, f_{FK}) = \sum_k w_k \cdot (\sum_k f_{FK} - \sum_k w_k \cdot f_{FK})^2 + \sum_k w_k \cdot \sigma^2(Y_{TF} | f_{FK}) \quad (9)$$

3.5 Estimation of Signal to Noise Ratio (SNR)

The estimation of signal to noise ratio is the measure of the robustness of a projection of a climate variable (rainfall) (Hawkins and Sutton 2011). This is estimated with the help of downscaled climate variables for the past and the future in individual models ($\{f_{p1}, f_{p2}, \dots, f_{PK}$

}, { $f_{F1}, f_{F2}, \dots, f_{FK}$ }) from GCMs along with the BMA parameters $\theta = \{W_k, \sigma_{PK}^2, \sigma_{FK}^2\}$. The steps are as follows (Duan and Philips, 2010);

- i. Select the ensemble size of 1000. Set spatial point $i=1$, and ensemble $j=1$;
- ii. A GCM is selected randomly based on uniform random number and a probability is generated randomly. The values are obtained as $Y_{PK}^{(i,j)}$ from the PDF of normal distribution $g(Y_{PK} | f_{PK}^{(i)}, \sigma_{PK}^2)$ and $Y_{FK}^{(i,j)}$ from the PDF of normal distribution $g(Y_{FK} | f_{FK}^{(i)}, \sigma_{FK}^2)$ using the parameters and compute $\Delta y^{(i,j)} = Y_{FK}^{(i,j)} - Y_{PK}^{(i,j)}$.
- iii. Complete step (ii) for all ensembles (1000) and for all spatial points for a given resolution.
- iv. Signal to Noise ratio (SNR) is computed as; The Difference of future and past climate expected changes / standard deviation (Δy). The expected weighted mean changes are obtained from equation 6 and 7. This measures SNR for local climate change. For further details refer to Duan and Philips (2010).

4.0 Results and Discussion

A linear regression based statistical downscaling method is adopted for rainfall projections of ISMR at multiple resolutions, viz., 0.05° , 0.25° and 0.50° , at monthly time scale to understand their projections and inter-compare them for potential differences. The results are presented in the following subsections which includes a comparison of simulations with observations for the historical period, comparisons of PDF, comparison between BMA and MMA, uncertainty across space and time, and the Signal to Noise ratio (SNR)

4.1 Comparison of results between the model and observations for historical period

348 The results of downscaled product are evaluated with respect to observed rainfall data (Fig.3).
 349 Figures 3 and 4 show the grid-wise details of the mean and standard deviation and are shown in
 350 the order of the observed rainfall, the multi-model average from 19 GCMs, the absolute
 351 percentage errors in the mean and the standard deviation. The results show that characteristics of
 352 the mean are quite faithful to the observational estimates. The spatial pattern of Indian monsoon
 353 rainfall is governed by orography due to the hilly regions of Himalayas, Western Ghats, Sapura,
 354 Vindhya and Aravali ranges (Salvi et al., 2013; Rajendran et al 2008). GCMs being coarse
 355 resolution models, cannot capture the spatial variability of ISMR. Here, we find that the
 356 statistical downscaling captures the occurrences of high rainfall in the windward side of the
 357 Western Ghats and low rainfall for the leeward side. The Rajasthan receives scanty rainfall
 358 because the orientation of Aravali ranges, and is also observed in the downscaling simulations.
 359 The North-east part of India receives heavy rainfall because presence of mountainous regions
 360 and is captured by statistical downscaling.

361 For majority of the grids over the Indian landmass, the absolute percentage errors in the mean
 362 (Figs. 3 (C), (F), and (I)) and the standard deviation (Figs. 4 (C), (F) and (I)) are found to be
 363 around 10%. Similarly the absolute percentage error for standard deviation is also less than 10%
 364 for most of the grids. However, parts of the western and southern regions receive very scanty
 365 rainfall and traditionally, the statistical downscaling model can only poorly capture this
 366 variability of rainfall (Salvi et al 2013). In those regions, the absolute percentage error for
 367 standard deviation is over 30%. The monthly means are expressed as mm/day. Along with the
 368 mean, the spatial pattern and temporal variability of rainfall is also well captured by the
 369 statistical model with a high degree of accuracy. The spatial dependence of rainfall between
 370 rainfall stations has also been well modelled by the model. In the present analysis, for developing

the statistical downscaling, we have used the PCs in each of the seven zones separately and established the regression relation. We have shown in Fig. 5, the cumulative percentage variance explained by the NCEP/NCAR PCs in different zones for the zonal wind (u) at 500 hPa at 90% variability.

The linear regression in statistical downscaling assumes that the residuals follow normal distribution. We check the normality assumptions for the residuals at multiple resolutions (Fig. 6). We find that the errors at majority of the nodes follow normal distribution. We find that for 0.05° , 0.25° and 0.5° resolutions, around 15%, 27% and 35% nodes violate the normality assumptions, respectively. The assumption is violated in the low rainfall regions which are mostly located in the western part and Jammu Kashmir regions of India. This is considered as the limitation of this study. The use of generalized linear regression or robust regression approaches may yield better results. The 0.5 degree downscaling has a slightly higher percentage of around 35% violations.

4.2 Skill Score based PDF comparison

The present statistical downscaling method captures the mean conditions very well for multiple resolutions over the historical period. Here, we also use the Skill score (S-score) for evaluation, which is based on the similarity of the PDFs computed with both observations and simulations for the historic period (Fig. 7). Fig 7a shows that the skill score obtained from the transfer function based downscaling at different resolutions are almost similar; however, such a conclusion is not valid for the interpolation method (Fig. 7 b). We also perform downscaling at different resolutions, viz., 0.50° , 0.25° and 0.05° and then re-grid the first two into the latter one. We

observe that model performance is best for 0.05^0 resolution and reduces with the coarseness of downscaling resolution (Fig. 7(b)).

We apply a transfer function based statistical downscaling approach, which relies on the statistical relationship between the large scale circulation pattern (predictor) and local rainfall (predictand). The other approach of downscaling is a change factor based method, which involves interpolation and bias correction. Here, we regrid 0.5^0 and 0.25^0 datasets to 0.05^0 to check if such an interpolation method is really valid for the spatially varying monsoon rainfall. The results show, that interpolation method fails to attain the skill which is possible to achieve with the transfer function based method. This also points to a very important conclusion on the possible failure of methods such as the Bias Correction Spatial Disaggregation (BCSD), which are widely used and based on interpolation and bias correction. The reason is probably that the high spatial variability of the Indian monsoon rainfall. This is quite expected since the reduction of model performance is primarily due to interpolation without additional information from synoptic scale circulation. This also highlights a possible limitation of widely used change factor based statistical downscaling methods, such as the BCSD (Wood et al., 2004), which is just a combination of interpolation and bias correction.

4.3 Comparison of results based on BMA and MMA

To know the plausible changes in future rainfall, we consider downscaled rainfall projections for RCP4.5 scenarios at 0.50^0 , 0.25^0 and 0.05^0 resolutions. We consider the projections obtained with both MMA and BMA. The weights obtained by using BMA are presented in Fig. 8 at different resolutions for the 2080s. We obtain consistently higher weights for GISS, CMCC, MIROC and MPI at all the resolutions. Similarly, consistently lower weights are assigned to BCC, BNU and CCCMA.

415 The projections obtained with both MMA (Fig. 9a-c) and BMA (Fig. 9d-f) show spatially non-
416 uniform projected changes of ISMR for the 21st century. Increases in precipitation are projected
417 for the Himalayan foothills, the west coast, and the northern regions, while decreasing trends are
418 observed in the Gangetic West Bengal and regions of Central India with mixed results in
419 Southern India. The increase/decrease in rainfall for future time slices is consistent with other
420 studies with fine resolution models (Salvi et al., 2013; Shashikanth et al., 2013; Kumar et al.,
421 2006; Rajendran and Kitoh 2008). The orographic rainfall is expected to increase in west coast
422 and the northeast part of India in the future in those studies also. We find spatially non-uniform
423 projected changes of ISMR which is not visible in the original simulations from GCMs
424 (Shashikanth et al, 2013). Rajendran and Kitoh (2008) suggested that there will be changes in
425 the moisture holding capacity of atmosphere in the future; mean changes in the wind and
426 convergence of moisture at 850 hPa, evaporation, and precipitable water (PWAT) are some of
427 the causes for future changes in mean rainfall over India. In general, we expect intensification of
428 land-ocean temperature gradient, a strengthening of the of low pressure trough, and more
429 moisture in atmosphere (as per the Clausius-Clapeyron relation), which may intensify the ISMR
430 in the future (Kripalani et al 2007). However, uneven heating may result in uneven distribution
431 of rainfall resulting non-uniform changes and the past decades have shown a decreasing trend in
432 ISMR despite the rapid warming of the Indian Ocean and the models are unable to capture this
433 trend (Saha et al. 2013). For the present study, we have considered a blend of predictors which
434 increases our confidence in climate projections for the future. The statistical downscaling
435 assumes stationarity in the relation between predictors and predictands. This is the main
436 limitation of the present work.

We find that the changes projected with MMA are slightly more on the positive side, compared to BMA. This suggests the possibility of the existence of models with low weights projecting high positive changes. The differences between BMA and MMA are not sensitive to the resolution of downscaling. For the application of BMA, the changes in rainfall should follow normal distribution. Here we find (Fig. 10) 25%, 28%, 27% of the grids at resolutions 0.05° , 0.25° and 0.5° , respectively, violate this assumption. The BMA should be applied for those grids with a modified maximum likelihood function. This is considered as the potential scope for future research.

The climate projections at global/ regional scale are subject to uncertainty associated with the use of multiple models. Here we use the coefficient of variation (CV) (Shashikanth et al., 2013; Woldsmakel et al., 2012) as a measure of uncertainty for ISMR projections at different resolutions and for different time windows 2020s (2010-2039) , 2050s (2040-2069) and 2080s (2070-2099). The relative changes in uncertainties across different resolutions are measured with respect to 2020s. We find that the model uncertainty is highest for the 2020s and then it gradually decreases with the lead time (Fig. 11). This is true at all resolutions. The possible reason is that with the increase in lead time, the climate change signal becomes more prominent compared to internal variability and model uncertainty due to the imposed radiative forcings (Hawkins and Sutton, 2009), resulting in a decrease in the CV values.

4.5 Estimation of SNR at multiple resolutions

The SNR in climate change projections indicates the strength of the climate signal with respect to the uncertainty associated with the use of multiple models. Here we compute SNR for the three resolutions (0.05° , 0.25° , and 0.50°) for different time periods (2020s, 2050s and 2080s)

(Fig. 12). With the reduction of model uncertainty, the SNR increases from the 2020s to 2080s for all resolutions. This is in agreement with Hawkins and Sutton (2009), where the prominence of climate change signal is argued to emerge at longer lead times. We also find that SNR does not depend on the resolution of downscaling. Statistical downscaling uses the statistical relationship between the fine resolution rainfall and coarse resolution circulation patterns and hence this is mostly affected by the uncertainty in the circulation patterns due to the use of multiple GCMs. Increase in resolution does not add any source of uncertainty into the downscaling method and hence the SNR and uncertainty remain unaffected by the change in resolution. Though fine resolution modelling improves the simulation of local scale climatology (Fig. 7), it does not necessarily add any value in terms of changes (Fig. 9), uncertainty (Fig. 11) and SNR (Fig. 12). The primary purpose for downscaling in climate change studies is to model the projected changes and here we do not find any significant improvements or differences in the projected changes. However, the selection of optimal resolution for downscaling is essential for specifically capturing the orography and region-specific land, ocean, atmosphere interactions which are generally not considered in coarse resolution GCM simulations. The present analysis points to the need for research activities in understanding the optimum resolution required in downscaling (specific to a region), to capture the orographic and other local effects, and at the same time, not to increase the resolution unless beneficial effects are expected on the SNR to justify the high computational efforts.

5.0 Conclusions

In this study, we use statistical downscaling for projections of ISMR at multiple resolutions. The primary objective is to understand the improvements obtained by increasing the spatial resolution from 0.50° to 0.05° . Following are the main conclusions from the present study

- 1 The skill score plots (Fig. 7) show improvements in simulating climatology at the finest resolution (0.05°) with respect to observed data, when we use direct statistical downscaling at that finest resolution in comparison to that of simulations (at 0.05°) interpolated from coarser resolution downscaled data (0.5° and 0.25°). This result also highlights the failure of the interpolation method, which are often coupled with a bias correction (Wood et al., 2004), without any specific regional modelling.
- 2 Enhanced resolution increases the computational requirements for any downscaling and hence it is necessary to evaluate the finest resolution model for payoff in terms of uncertainty and SNR. We observe that enhanced resolution does not affect the uncertainty in statistical downscaling. This is specifically because the key uncertainty in statistical downscaling is due to the uncertainty in synoptic scale circulation pattern simulated by the parent models or the reanalysis data employed in developing the statistical relations (Kannan et al., 2014). As the observed rainfall data at multiple resolutions are obtained from the same agency, rainfall data does not add resolution-dependent uncertainty. However, the contrary is not necessarily true; i.e., increasing resolutions does not necessarily guarantee the usability or the usefulness of the downscaled product at arbitrarily high resolutions. Caution is thus needed in using other metrics for determining optimal resolutions for statistical downscaling. This is reemphasized below in the context of SNR.

3 We also find that the SNR of climate change projections does not improve by increasing resolution. Here SNR denotes the strength of the signal of climate change with respect to the uncertainty. As there is no significant change in uncertainty with the enhanced resolution, the SNR also remains the same. This highlights that a finer resolution in statistical downscaling does not improve the reliability of projected changes; however an optimal resolution for downscaling is necessary to capture the major local and orographic factors and other region-specific convective processes related to the land-atmosphere-ocean interactions.

4 The present study shows that there is no significant difference between the BMA and MMA projections at any resolution; though, the projected changes with MMA are slightly more on the positive side, compared to BMA. This is because of the use of a relatively large number of models. With the increase in the number of models, the individual weights are assigned to lower values. For example, if there are 5 models, with MMA, the individual weights will be 0.2; but with 20 models, the weights are reduced to 0.05. The change in the weights, where they themselves are very low, because of the use of a 'sufficient' number of models, may not result in significant changes in the ultimate multi-model projections.

Further work is needed in determining what this 'sufficient' number is and what the optimal resolution for statistical downscaling would be as the resolution of the regional and global GCMs and the available data continue to increase. The real ultimate decider will of course be the use and usability of the products in decision-making and policy frameworks. The exercise here is but the first step in exploring these requirements for decision-support.

525

526 **ACKNOWLEDGEMENT**

527 We acknowledge the World Climate Research Programme's working Group on coupled
528 Modelling, which is responsible for CMIP, and we thank the modelling groups for
529 producing and making available their model output. For CMIP the U.S. Department of
530 Energy's Program for Climate Model Diagnosis and Intercomparison (PCMDI) provides
531 coordinating support and led development of software infrastructure in partnership with the
532 Global Organization for Earth System Science Portals. We would like to thank APHRODITE,
533 Japan for making available 0.050, 0.250 and 0.500 resolution observed data. We particularly
534 thank Dr. Kamaguchi of APHRODITE, for making available the 0.05⁰ observed data. Raghu
535 Murtugudde gratefully acknowledges the National Monsoon Mission funding for partial support.
536 K Shashikanth acknowledges Qingyun Duan for his help in understanding Bayesian Multimodel
537 Averaging.

538 **REFERENCES**

539

540 Anandhi, A., V. V. Srinivas, R. S. Nanjundiah, and D. N. Kumar (2008), Downscaling
541 precipitation to river basin for IPCC SRES scenarios using support vector machines, Int. J.
542 Climatol., 28(3), 401–420.10.1002/joc.1529.

543 Bardaossy, A., and E.J. Plate (1992) Space-time model for daily rainfall using atmospheric
544 circulation patterns Water Resources Research 28 (5) 1247-1259 DOI: 10.1029/91WR02589

545 Convoy D., and P.D.Jones (1998) The use of weather types and air flow indices for GCM
 546 downscaling , Journal of Hydrology 212-213,348-361.

547 Corte Real, J., X. Zhang and X. Wang (1995) Downscaling GCM information to regional scales:
 548 A non-parametric multivariate regression approach Clim.Dyn., 11, 413-424

549 Coulibaly, P.,Y. B. Dibike, and F. Anctil 2005) Downscaling Precipitation and Temperature with
 550 Temporal Neural Networks, J. Hydrometeor, 6, 483–496.

551 Deser,C., Knutti, R., Solomon, S., Phillips, A.S., 2012. Communication of the role of natural
 552 variability in future North American climate, Nature Clim. change, vol. 2, pp. 775-779, 2012.
 553 <http://dx.doi.org/10.1038/nclimate1562>.

554 Dibike, Y.B., Coulibaly, P., 2005. Hydrologic impact of climate change in the Saguenay
 555 watershed: comparison of downscaling methods and hydrologic models. J. Hydrology 307: 144–
 556 163.

557 Dobler A, Ahrens B. 2011. Four climate change scenarios for the Indian summer monsoon by the
 558 regional climate model COSMO-CLM. J. Geophys. Res. 116: D24104, DOI:
 559 10.1029/2011JD016329

560 Duan, Qingyun, Thomas J. Philips, 2010. Bayesian estimation of local signal and noise in
 561 multimodel simulations of climate change J.Geophys. Res. 115, D18123, DOI
 562 10.1029/2009/JD013654

563 Fowler, H. J., Blenkinsop, S., Tebaldi, C., 2007. Linking climate change modelling to impacts
 564 studies: recent advances in downscaling techniques for hydrological modelling, Int. J. Climatol.
 565 **27**: 1547–1578 2007

566 Ghosh, S. (2010), SVM-PGSL coupled approach for statistical downscaling to predict rainfall
567 from GCM output, J. Geophys. Res., 115, D22102, doi:10.1029/2009JD013548

568 Ghosh, S., Mujumdar, P. P., 2009. Climate Change Impact Assessment- Uncertainty Modeling
569 with Imprecise Probability, J. Geophys. Res.Atmos., 114, D18113, doi:10.1029/2008JD011648.

570 Ghosh, S., and Mujumdar, P. P. (2006), Future rainfall scenario over Orissa with GCM
571 projections by statistical downscaling, Curr. Sci. India, 90(3), 396–404.

572 Ghosh, S., Katkar, S., 2012. Modeling Uncertainty Resulting from Multiple Downscaling
573 Methods in Assessing Hydrological Impacts of Climate Change, Water Resour Manage 2012
574 26:3559–3579, DOI 10.1007/s11269-012-0090-5

575 Giorgi, F., Mearns L.O., 2002. Calculation of average, uncertainty range, and reliability of
576 regional climate changes from AOGCM simulations via the reliability ensemble averaging REA
577 method. J Clim 15:1141–1158

578 Groppelli, B., D. Bocchiola and R.Rosso., 2011. Spatial downscaling of precipitation from
579 GCMs for climate change projections using random cascades: a case study in Italy, Water
580 Resour. Res., doi. 10.1029/2010WR009437.

581 Gutmann et al., 2014. An intercomparison of statistical downscaling methods used for water
582 resource assessments in the United States accepted article in Water Resources Research, doi
583 10.1002/2014/WR015559.

584 Hawkins, Ed., Rowan Sutton, 2009. The Potential to Narrow Uncertainty in Regional Climate
585 Predictions. Bull. Amer. Meteor. Soc., 90, 1095–1107. doi:
586 <http://dx.doi.org/10.1175/2009BAMS2607.1>

587 Hawkins, Ed., Rowan Sutton, 2011. The Potential to Narrow Uncertainty in Regional Climate
588 Predictions, *Clim Dyn*, 37:407–418DOI 10.1007/s00382-010-0810-6.

589 Hewitson, B. C. and R. G. Crane, 1992, Large-scale atmospheric controls on local precipitation
590 in tropical Mexico, *Geophysical Research Letter* DOI: 10.1029/92GL01423

591 IPCC. 2007. *Climate Change 2007: The Physical Science Basis Contribution of Working Group*
592 *I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, Solomon
593 S, Qin D, Manning M, Chen Z, Marquis M, Averyt K.B, Tignor, M, Miller HL eds. Cambridge
594 University Press: Cambridge, UK/New York,

595 Kalnay, N.Y. et al., 1996 The NCEP/NCAR 40- years reanalysis project. *Bull. Amer. Meteor.*
596 *Soc.*, **77**: 437471

597 Kannan, S, Ghosh. S., 2013. A nonparametric kernel regression model for downscaling
598 multisite daily precipitation in the Mahanadi basin. *Water Resour. Res.* 49, DOI
599 10.1002/wrcr.20118

600 Kannan, S, Ghosh, S, Mishra, V. Salvi K., 2014, Uncertainty Resulting from Multiple Data
601 Usage in Statistical Downscaling, *Geophys. Res. Lett.*, doi: 10.1002/2014GL060089

602 Khalili, M., Brissette, F. Leconte, R., 2009. Stochastic multi-site generation of daily weather
603 data, *Stoch Environ Res Risk Assess* 23, 837-849, doi:10.1007/s00477-008-0275-x

604 King, L., Irwin S, Sarwar. R, McLeod A.I., Simonovic, S.P., 2012. The effects of climate
605 change on extreme precipitation events in the Upper Thames River Basin: A comparison of
606 downscaling approaches. *Canadian Water Resources Journal* **37**: 253–274.

607 Knutti, R., 2010. The end of model democracy, *Climatic Change* 102:395–404DOI
608 10.1007/s10584-010-9800-2

609 Kripalani , H.R., J.H.Oh, A.kulkarni, S.S.Sabade, and H.S. Chaudari , 2007, South Asian
 610 summer monsoon precipitation variability: Coupled climate model simulations and projections
 611 under IPCC AR4, Theor.Appl.Climatol doi. 10.1007/s0074-006-282-0

612 Liang et al., 2012. Regional climate weather research and forecasting model, Bull. Amer.
 613 Meteor. Soc., <http://dx.doi.org/10.1175/BAMS-D-11-00180.1>

614 Maraun, D., et al., 2010. Precipitation downscaling under climate change: Recent developments
 615 to bridge the gap between dynamical models and the end user, Rev. Geophys., 48, RG3003,
 616 doi:10.1029/2009RG000314.

617 Meherotra, R and A.Sharma, 2005, A non parametric non-homogeneous hidden Markov model
 618 for downscaling of multisite daily rainfall occurrences, J.Geophys.Res. 110,D16108, doi.
 619 10.1029/2004/JD005677.

620 Meherotra, R and A.Sharma, 2006, A non parametric stochastic downscaling frame work for
 621 daily rainfall at multiple locations, J. Geophys.Res. 111, D15101, DOI: 10.1029/2005JD006637.

622 Mehrotra, R., and A. Sharma, 2010, Development and Application of a Multisite Rainfall
 623 Stochastic Downscaling Framework for Climate Change Impact Assessment, Water Resour.
 624 Res.46, W07526, doi:10.1029/2009WR008423.

625 Menon, A., Levermann, A., Schewe, J., 2013. Enhanced future variability during India's rainy
 626 season, Geophys. Res. Lett., 40,3242–3247, doi:10.1002/grl.50583.

627 Miao Chiyuan , Qingyun Duan, Qiaohong Sun, Yong Huang, Dongxian Kong, Tiantian Yang,
 628 Aizhong Ye, Zhenhua Di, and Wei Gong (2014) Assessment of CMIP5 climate models and
 629 projected temperature changes over Northern Eurasia, Environ. Res. Lett., doi 10.1088/1748-
 630 9326/9/ 055007.

631 Miao Chiyuan , Qingyun Duan, Qiaohong Sun, Jianduo Li (2013) Evaluation and application of
632 Bayesian multi-model estimation in temperature simulations, Progress in Physical geography 1-
633 18, doi 10.1177/030913313494961.

634 Parthasarathy, B., Munot, A. A., Kothawale, D. R., 1996. Monthly and Seasonal Rainfall Series
635 for All-India Homogeneous Regions and Meteorological Subdivisions, 1871 – 1994 pp., Indian
636 Inst. of Trop. Meteorol., *Pune*, India.

637 Perkins S.E., Pitman, A.J., Holbrook, N.J. Mc. Aneney, J., 2007. Evaluation of the AR4 Climate
638 models simulated Daily maximum temperature, minimum temperature and precipitation over
639 Australia using probability density functions , J.Clim.,Vol.20,4356 DOI10.1175/JCLI4253.1

640 Prudhomme, C., Davies.H., 2009. Assessing uncertainties in climate change impact analyses
641 on the river flow regimes in the UK. Part 2: Future climate, Clim. Change, 93, 197–222.

642 Prudhomme, C., Reynard, N., Crooks., S. 2002. Downscaling of global climate models for
643 flood frequency analysis: where are we now? Hydrol. Process., 16, 1137-1150.

644 Raftery, A. E., Gneiting, T., Balabdaoui, F., Polakowski, M., 2005. Using Bayesian model
645 averaging to calibrate forecast ensembles, Mon. Weather Rev., 133, 1155–1174,
646 doi:10.1175/MWR2906.1.

647 Räisänen, J., 2007. Objective comparison of patterns of CO₂ induced climate change in coupled
648 GCM experiments. Clim. Dyn., 13, 197-211

649 Raje, D., and P.P. Mujumdar 2009, A conditional Random field based downscaling method for
650 assessment of climate change impact on multisite daily precipitation in the Mahanadi basin,
651 Water Resour.Res., 45, W10404, doi. 1029/2008WR007487.

652 Rajendran K, Kitoh A. 2008, Indian summer monsoon in future climate projection by a super
653 high resolution global model. *Current Science* **95**(11): 1560–1569.

654 Rupa Kumar KR, Sahai AK, Kumar KK, Patwardhan SK, Mishra PK, Revadekar JV, Kamala K.
655 2006, High-resolution climate change scenarios for India for the 21st century. *Current Science*
656 **90**(3).

657 Saha, S K, S Pokhrel, H.S. Chaudhari, A Dhakate, S Shewlale, CT Sabeerali, K Salunke, A
658 Hazra (2013) Improved Simulation of Indian summer monsoon in latest NCEP climate forecast
659 system free run, *Int. J. Climatol.* Doi 10.1002/joc.3791.

660 Salvi, K., Kannan, S. Ghosh S., 2013. High-resolution multisite daily rainfall projections in
661 India with statistical downscaling for climate change impacts assessment, *J. Geophys.*
662 *Res.Atmos.*, 118, doi:10.1002/jgrd.50280.

663 Schiermeier, Q. 2010. The real Holes in the Climate science *Nature*, Vol .463 284-287

664 Schoof J.T., 2013, Statistical downscaling in climatology, *Geography Compass* 7/4:249-265,
665 doi10.1111/gec3.12036

666 Shashikanth, K., Kaustubh, S., Ghosh, S., Rajendran, K., 2013. Do CMIP5 simulations of Indian
667 summer monsoon rainfall differ from those of CMIP3? *Atmospheric Science Letters*,
668 10.1002/asl2.466, 10.1002.

669 Stainforth, D.A., Allen, M.R, Tredger, E.R., Smith, L.A, 2007. Confidence, uncertainty and
670 decision-support relevance in climate predictions *Phil. Trans. R. Soc.* 2007 365, doi:
671 10.1098/rsta.2007.2074.

672 Taylor, Karl E., et al., 2012. An Overview of CMIP5 and the Experiment Design, *Bull Amer.*
673 *Meteor. Soc.*, 93, 485–498. doi: <http://dx.doi.org/10.1175/BAMS-D-11-00094.1>.

674 Tebaldi, C., Mearns, L. O., Nychkam D. Smith, R. W., 2005. Quantifying uncertainty in
675 projections of regional climate change: A Bayesian approach to the analysis of multimodel
676 ensembles, *J. Clim.*, 18, 1524–1540.

677 Tebaldi, C., Mearns, L. O., Nychkam D., Smith, R. W., 2004. Regional probabilities of
678 precipitation change: A Bayesian analysis of multi-model simulations, *Geophys. Res. Lett.*, 31,
679 L24213, doi: L10.1029/2004GL021276

680 Von Storch, H., 1995. Inconsistencies at the interface of climate impact studies and global
681 climate research, *Meteor. Z.*, 4 NF, 72-80.

682 Von Storch, H., 1999. On the use of “inflation” in downscaling, *J. Clim.*, 12, 3505-3506

683 Watanabe, S., Kanae, S., Seto, S., Yeh, P. J.-F, Hirabayashi, Y. Oki, T., 2012. Intercomparison
684 of bias-correction methods for monthly temperature and precipitation simulated by multiple
685 climate models, *J. Geophys. Res.*, 117, D23114, doi:10.1029/2012JD018192.

686 Wayne, G.P. 2013. Beginners guide to Representative concentration pathways Skeptical Science
687 version 1.

688 Webster, P.J., Magana, V.O., Palmer, T.N., Shukla, J., Tomas, R.A., Yanai, M. Yasunari, T.,
689 1998. Monsoons: Processes, predictability and the prospects for prediction. *J. Geophys. Res.*,
690 103, 14,451-14, 510.

691 Wilby R.L., Dawson C.W. 2004. Using SDSM Version 3.1—a decision support tool for the
692 assessment of regional climate change impacts, User manual.

693 Wilby, R. et al., 2004. Guidelines for use of climate scenarios developed from statistical
694 downscaling methods. <http://www.narccap.ucar.edu/doc/tgica-guidance-2004.pdf> accessed on
695 719 10/08/2013.

696 Wilby, R. L., Wigley, T. M. L., 1997. Downscaling general circulation model output: A review
697 of methods and limitations, *Prog. Phys. Geogr.*, 21, 530–548

698 Wilks, D.S. (1999) Multisite downscaling of daily precipitation with a stochastic weather
699 generator, *Clim.Res.*, 11, 125-136

700 Wilks, D.S. 1999. Multi-site downscaling of daily precipitation with a stochastic weather
701 generator, *Clim. Res.*, 11, 125–136.

702 Wise, M.A, Calvin, K.V., Thomson, A.M., Clarke, L.E., Bond-Lamberty, B., Sands, R.D.,
703 Smith, S.J., Janetos, A.C., Edmonds., J.A., 2009 Implications of Limiting CO₂ Concentrations
704 for Land Use and Energy, *Science.*, 324:1183-1186. May 29, 2009.

705 Woldemeskel, F. M., Sharma, A., Sivakumar, B., Mehrotra, R., 2012. An error estimation
706 method for precipitation and temperature projections for future climates, *J. Geophys. Res.*, 117,
707 D22104, doi: 10.1029/2012JD018062.

708 Wood, A. W., Leung, I. R., Sridhar, V., Lettenmaier, D. P., 2004. Hydrologic implications of
709 dynamical and statistical approaches to downscaling climate model outputs climatic change 62:
710 189–216, 2004.

711 Yatagai, A. et al. 2012 APHRODITE: Constructing a Long-Term Daily Gridded Precipitation
712 Dataset for Asia Based on a Dense Network of Rain Gauges, *Bull. Amer. Meteor. Soc.*, 939:
713 1401-1415, 727 10.1175/BAMS-D-11-00122.1.

714 Yip, Stan, Christopher A. T., Ferro, David B., Stephenson, Ed Hawkins, 2011. A Simple,
715 Coherent Framework for Partitioning Uncertainty in Climate Predictions. J. Clim. 24, 4634–
716 4643. doi: <http://dx.doi.org/10.1175/2011JCLI4085.1>.

717

718

Table 1: Statistical Downscaling approaches to different case studies

S.No.	Literature Citation	Case study	Features of the Model
1	Bardossy and Plate, 1992	River Ruhr (Germany)	A multidimensional stochastic model is developed for the space-time distribution of daily precipitation.
2	Hewitson and Crane, 1992	Tropical Mexico region	Empirical relationships between sea level and 500 mb circulation fields, and the local precipitation in Chiapas, Mexico, are derived using an Artificial Neural Networks (ANN)
3	Corte-Real et al. 1995	Portugal	The multivariate adaptive regression splines (MARS) model was used to capture the observed relationships between sea level pressure (SLP) anomalies over the Euro-Atlantic sector and the winter time (December-February) monthly rainfall.
4	Conway and Jones, 1998	British Isles	Analysis using Lamb weather types, counts of weather fronts and air flow indices show strong relationships with daily rainfall characteristics.
5	Wilks, 1999	30 stations located in the state of Oregon in US	A stochastic weather generator model is developed and used for downscaling.
6	Mehrotra and Sharma 2005, 2006	Sydney Australia	K-nearest neighbor based nonparametric, non-homogeneous hidden Markov model is developed to generate rainfall sequences at multiple locations using synoptic scale circulation patterns.
7	Ghosh and Mujumdar, 2006	Orissa, India	A linear regression developed between GCM simulated predictors Vs rainfall data. Principal component analysis (PCA) is used to reduce the dimensionality and fuzzy clustering algorithm is used to classify the components.
8	Anandhi et al., 2008	Malaprabha reservoir catchment in India.	The monthly precipitation is downscaled using support vector machine.

9	Raje and Mujumdar, 2009	Mahanadi catchment Orissa, India	Developed conditional random field based algorithm to downscale rainfall. Direct classification of dry/wet days as well as precipitation amount is achieved within a single modeling framework.
10	Mehrotra and Sharma 2010	Network of 45 rain gauges near Sydney, Australia.	Used stochastic downscaling framework for simulation of multisite daily rainfall occurrences. Rainfall occurrences are modeled using a modified Markov mode and rainfall amounts are estimated using nonparametric kernel density simulator conditioned on previous time step rainfall using large scale predictors.
11	Ghosh, 2010	Assam and Meghalaya, India	SVM based relationship between predictors and predictand. Probabilistic Global search algorithm, a probabilistic search technique is used to compute the optimum number of parameters for SVM; uncertainty addressed using modified Reliability Ensemble Average (REA) technique.
12	Groppelli et al., 2011	Oglio River basin in Italy	Stochastic space random cascade (SSRC) approach is developed to downscale rainfall from GCM data under climate change scenario.
13	Kannan and Ghosh, 2013	Mahanadi River Basin, Orissa, India	Developed the Kernel Regression based downscaling model conditioned on weather state based on Classification And Regression Tree (CART) method.
14	Salvi et al., 2013	India	Used CART and Kernel Regression technique to project rainfall using CCCma GCM.
15	Gutmann et al., 2014	USA	The downscaling algorithms considered are daily and monthly-disaggregated –to –daily Bias and Bias Corrected Constructed Analog (BCCA), Asynchronous Regression (AR), methods.

Table 2: List of GCMs used in monthly Rainfall Downscaling

S.No	Name	Institution	Resolution	
			Lat(degrees)	Long(degrees)
1	BCC- CSM1	Beijing Climate Center	2.7906	2.8125
2	BNU-ESM	College of Global Change and Earth System Science, Beijing Normal University	2.79	2.81
3	CCCma-CAN-ESM2	Canadian Centre for Climate Modelling and Analysis	2.8	2.8
4	CMCC-CMS	Centro Euro-Mediterraneo per I Cambiamenti Climatici	1.87	1.88
5	CNRM-CM5	Centre National de Recherches Meteorologiques	1.4	1.4
6	CSIR-ACCESS 1-0	CSIRO (Commonwealth Scientific and Industrial Research Organisation, Australia), and BOM (Bureau of Meteorology, Australia)	1.25	1.875
7	CSIRO-Mk3-6-0	Commonwealth Scientific and Industrial Research Organization in collaboration with Queensland Climate Change Centre of Excellence	1.87	1.88
8	FIO-ESM	The First Institute of Oceanography, SOA, China	2.79	2.81
9	GISS-E2-R-CC	NASA Goddard Institute for Space Studies	2.00	2.5
10	Had-GEM2-E2S	Met Office Hadley Centre	1.875	1.25
11	INM-CM4	Inst. for Numerical Mathematics	1.5	2
12	IPSL-CM5A-LR	Institute Pierre-Simon Laplace	1.89	3.75
13	IPSL-CM5A-MR	Institute Pierre-Simon Laplace	1.89	3.75
14	MIROC5	Atmosphere and Ocean Research Institute (-University of Tokyo)and others	1.4	1.41
15	MIROC-ESM	Atmosphere and Ocean Research Institute (University of Tokyo)and others	2.8	2.8
16	MPI-ESM-MR	Max-Planck-Inst. For Meteorology	1.8652	1.875
17	NASA-GISS-E2-R	NASA Goddard Institute for Space Studies	2	2.5
18	NorESM1-ME	Norwegian Climate Centre	1.8947	2.5
19	NorESM1-M	Norwegian Climate Centre	1.5	1.9

Fig 1: Statistical downscaling (SD) algorithm used for the study.

Fig 2 Meteorological homogeneous regions of India identified by IMD (A) [*Parthasarathy et al.* 1996] and the areal extent of predictors for the central zone (B).

Fig.3. The cumulative percentage of variance explained by NCEP/NCAR principal components (PCs) in the different zones for the zonal wind (u) at 500 hpa (90% variability shown along with principal component number).

Fig.4. The absolute percentage errors in the mean of multi-model averaged downscaled simulations at 0.05° , 0.25° and 0.50° spatial resolutions, respectively.

Fig.5. The absolute percentage errors in the standard deviation of multi model averaged downscaled simulations at 0.05° , 0.25° and 0.50° spatial resolutions, respectively.

Fig.6. Grids, where the residuals follow a normal distribution at 0.05° , 0.25° and 0.50° spatial resolutions, respectively. Fractions of grids, where residuals do not follow a normal distribution, are also shown.

Fig 7 Skill scores for downscaled projections at three resolutions, viz., 0.05° , 0.25° and 0.50° . The spatial distribution is represented as a PDF. Panel A represents the skill score at different resolutions. Panel B shows the skill score at 0.05° , where the downscaled simulations of 0.25° and 0.50° are further interpolated.

Fig. 8. Weights obtained for different GCMs at different resolutions with Bayesian modelling for 2080s.

Fig 9: Projected changes in mean rainfall for a future period (2070-2099) with respect to the base line period (1970-2005), with multi-model average (MMA) at 0.05^0 (A), 0.25^0 (B) and 0.50^0 (C) and with BMA (D, E and F, respectively). The percentage difference between MMA and BMA are presented in panels G, H and I.

Fig.10: Grids, where the projected changes in rainfall follow normal distribution at 0.05^0 , 0.25^0 and 0.50^0 spatial resolutions, respectively. Fractions of grids, where changes do not follow normal distribution, are also shown.

Fig. 11: Multi-model uncertainty during 2020s and its changes with time. The uncertainty is measured as Coefficient of Variation.

Fig. 12; The Signal to Noise Ratio (SNR) for projections at multiple resolutions for different time periods. The pie chart shows the fraction of grid points having SNR more/ less than 0.5.

Figure

Figure

(A)

(B)

Seven Meteorologically homogeneous zones of Indian Region as per India Meteorological Department (IMD), (Parthasarathy et al., 1996) with spatial extent of predictors

Figure

Figure

Figure

Figure

Figure

Figure

0.05°

0.25°

0.50°

Uncertainty

2020s

Changes in uncertainties measured with respect to 2020s

2050s

2080s

0.05 °

0.25 °

0.50 °

S
N
R
2
0
2
0 S
N
R
2
0
5
0 S
N
R
2
0
8
0

SNR > 0.5

SNR < 0.5