Projection of future rainfall changes over northeast Bangladesh: A Bayesian Model Averaging Approach

M. Basher, P.-S. Chu*, M. Stiller-Reeve#, and S. Islam

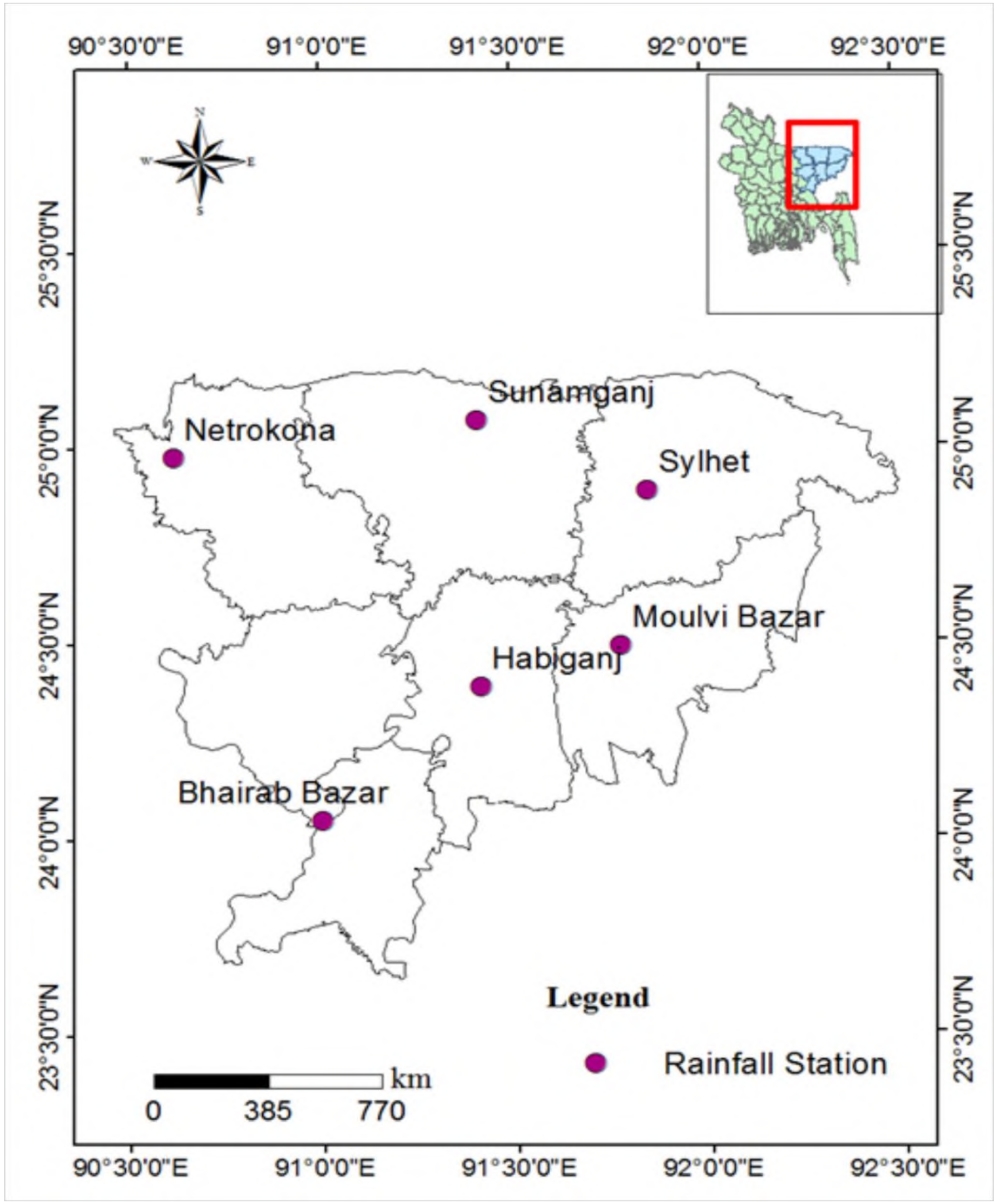
Bangladesh University of Engineering and Technology

*University of Hawaii

#University of Bergen, Norway

Rainfall is a major concern for NE Bangladesh

- NE Bangladesh is marked by extensive wetlands which are locally called as “haor.” The haors are bowl shaped low-lying flood plain. This area is also characterized by hills, forests, tea gardens and rubber gardens.
- The haors mostly dried up during December to May and Boro is cultivated in haors during this time. Boro rice during pre-monsoon (March to May) accounts for the majority of agricultural output, which contributes significantly to total rice production of the country. The haors remain under water from June to November so fisheries become important in livelihood earning.
- The pre-monsoon flash floods can destroy the seasonal Boro rice harvest. Knowledge of future rainfall variability during the pre-monsoon and monsoon (June-Sept) periods is vital for government policy planning.



Study area: 24N-25N, 90.62E-92.50E

Six rainfall stations with long-term daily data

Table 2: List of rainfall stations and their location.

Station Name	District	Location of Station (Latitude, Longitude)	Pre-monsoon mean rainfall (mm)	Monsoon mean rainfall (mm)	
Sylhet	Sylhet	24.90°N, 91.88°E	1085	2700	3,785 mm
Sunamganj	Sunamganj	25.06°N, 91.44°E	1095	4435	5,500 mm
Netrokona	Netrokona	24.98°N, 90.62°E	670	2485	
Moulvi Bazar	Moulvi Bazar	24.49°N, 91.70°E	740	1890	
Habiganj	Habiganj	24.39°N, 91.41°E	695	1525	
Bhairab Bazar	Kishoreganj	24.05°N, 91.00°E	580	1330	

- Gridded daily rainfall data from six regional climate models (RCMs) over CORDEX South Asia domain

- The World Climate Research Program (WCRP) initiated a framework named as the Coordinated Regional Climate Downscaling Experiment (CORDEX) to assess climate change impacts on regional to local scales

- A number of international climate modeling groups downscaled GCMs using various regional climate models (RCMs) with finer spatial resolution (10-50 km) under four different RCPs. We choose the grid which is closet to the rainfall station.

Use 6 RCMs over CORDEX South Asia domain under RCP4.5 and 8.5 for 2041-2070 and 2071-2100 for projecting future rainfall variations

Table 1: List of RCMs and their driving models

RCM	Driving GCM	Experiment Name	Institute
ACCESS	ACCESS1-0	ACCESS-CSIRO-CCAM	CSIRO
CCSM4	CCSM4	CCSM4-CSIRO-CCAM	CSIRO
CNRM	CNRM-CM5	CNRM-CM5-CSIRO-CCAM	CSIRO
MPI	MPI-ESM-LR	MPI-ESM-LR-CSIRO-CCAM	CSIRO
MPI-REMO	MPI-M-MPI-ESM-LR	MPI-CSC-REMO2009	MPI-CSC
SMHI	ICHEC, EC-EARTH	ICHEC-EC-EARTH-SMHI-RCA4	SMHI

- Representative Concentration Pathways (RCPs) are greenhouse gas concentration trajectories adopted by IPCC in 2014

- RCP 8.5 means radiative forcing values in the year 2100 will be 8.5 W/m**2 higher than pre-industrial values. Under RCP8.5 scenario, global warming increase will be about 3.7C and its likely range is 2.6C to 4.8C by the end of this century – too warm!

The RCMs often show considerable systematic errors; use quantile mapping for bias correction

- Generally, the RCMs simulate too many rainfall events with low intensity compared to observed rainfall. This is known as “drizzle effect.”
- To adjust the wet-day frequency of RCM simulated rainfall according to observed rainfall, a cut-off threshold corresponding to wet days (>1 mm) is chosen before applying the quantile mapping method

- Quantile mapping bias correction is an empirical statistical technique that matches the quantile of a RCM simulated value to the observed value at the same quantile (e.g., Maurer and Pierce, 2014).

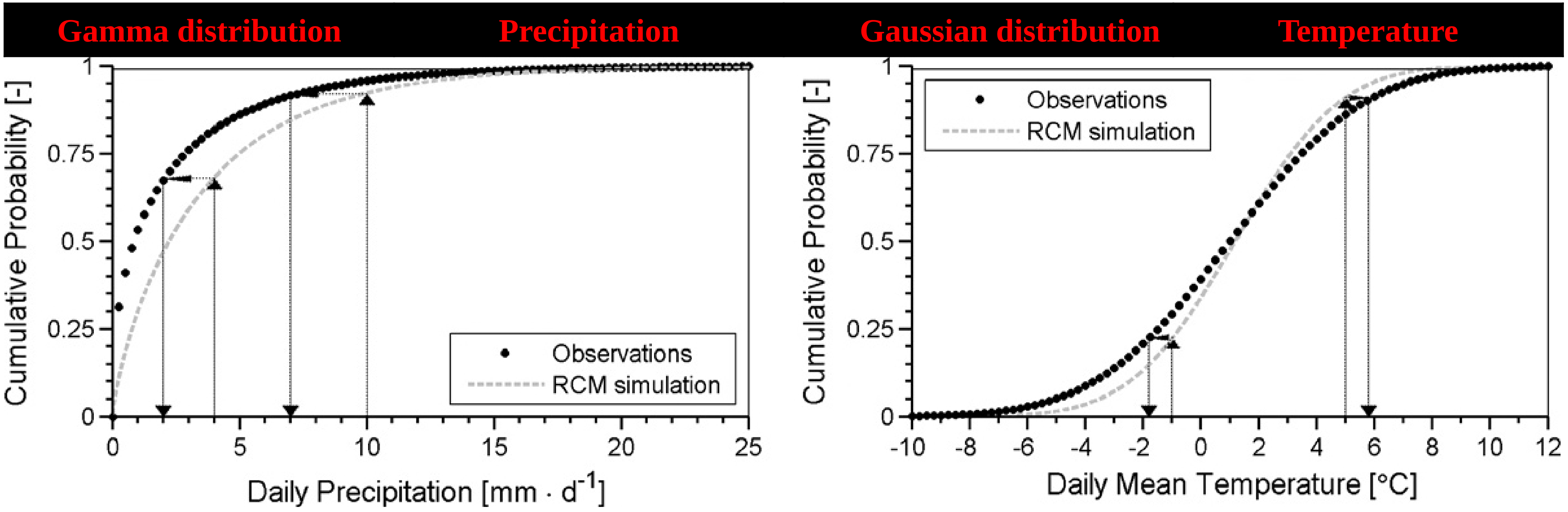
- The quantiles are determined by sorting RCM output and observations for the same historical base (i.e., reference or baseline) period, and constructing cumulative distribution functions (CDFs) for each.

- The gamma CDF of the observed and RCM reference rainfall for each month is build separately.
- The CDF of RCM reference simulation is mapped with CDF of the observations for generating the transfer function.
- This function (inverse gamma) is then used to correct the RCM scenario (i.e., future) period.

- Reference or baseline period: 1976-2005
- Future scenario periods: 2041-2070 and 2071-2099

Gamma inverse cumulative distribution function (rainfall)

- $\mathrm{x} = F^{-1}(F(sim, \alpha_{sim}, \beta_{sim}), \alpha_{obs}, \beta_{obs})$ α: shape, β: scale
- MATLAB gammaincinv



BMA produces a complete PDF of ensemble mean and quantifies the associated uncertainty of forecasts

- Given the training data Y^T and k climate models $(M_1 \ldots . M_k)$, the predictive PDF of a variable Y (i.e., rainfall) is given by:

$$P(Y|Y_1, Y_2, \ldots \ldots Y_k) = \sum_{k=1}^{k} P(Y|M_k) P(M_k|Y^T) \qquad (1)$$

where $P(M_k|Y^T)$ is the posterior probability of model M$_k$. $P(Y|M_k)$ is the forecast PDF based on model M_k alone.

$$\sum_{k=1}^{k} w_k = \sum_{k=1}^{k} P(M_k|Y^T) = 1$$

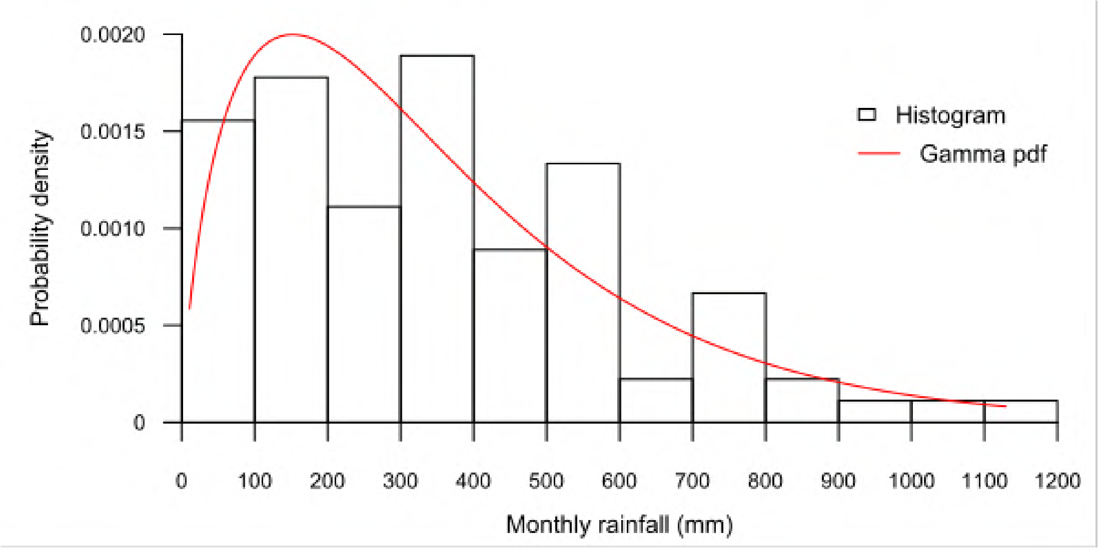
$$P(Y|Y_1, Y_2, \ldots \ldots Y_k) = \sum_{k=1}^{k} w_k P(Y|M_k) \qquad (2)$$

The BMA method assumes that the conditional PDF, $P(Y|M_k)$, of the individual model can be approximated by the normal distribution with mean $a_k + b_k\, M_k$ and standard deviation σ_k which is given by:

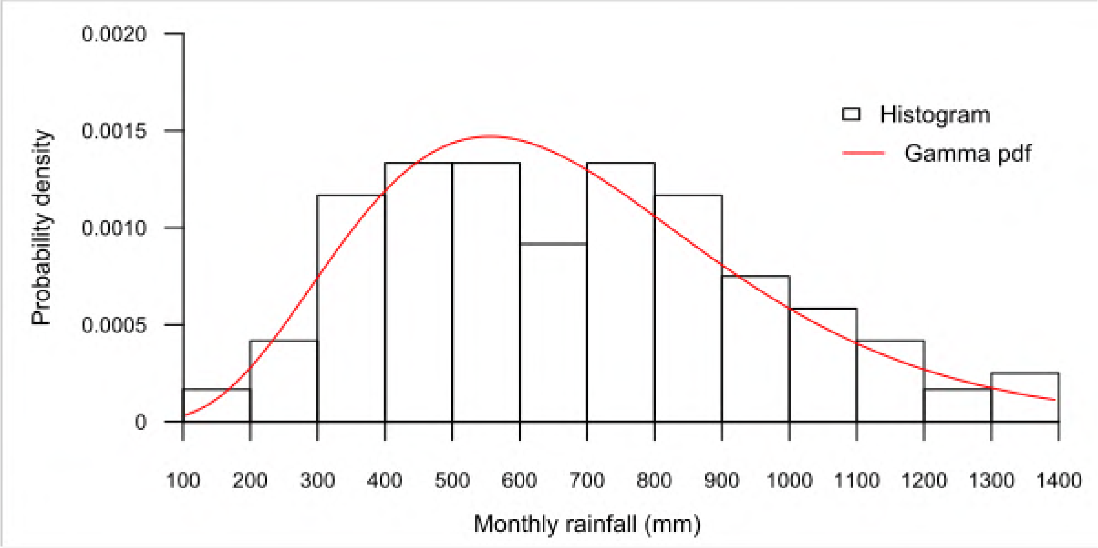
$$P(Y|M_k) \sim N(a_k + b_k\, M_k, \sigma_k^2) \qquad (3)$$

The values for a_k and b_k are estimated by simple linear regression of $P(Y|M_k)$ on M_k for each model.

But for non-Gaussian distributed data such as rainfall?



(a) Pre-monsoon



(b) Monsoon

Sylhet

- **The conditional PDF for gamma distribution with shape parameter α and scale parameter β can be given by:**

$$P(Y|M_K) \sim \frac{1}{\beta\Gamma(\alpha)} Y^{\alpha-1} \exp(-{}^{Y}/_{\beta}) \qquad (4)$$

for $Y > 0$. $P(Y|M_K) = 0$ for $Y \leq 0$. The mean of this distribution is $\mu = \alpha\beta$ and it's variance is $\sigma^2 = \alpha\beta^2$. The two gamma parameters can be expressed as a function of the mean and variance of the distribution as $\alpha_k = \mu_k^2/\sigma_k^2$ and $\beta_k = \sigma_k^2/\mu_k$. For simulated rainfall of a particular ensemble member, the mean and variance can be derived by the following relationship:

- $\mu_k = \overline{Y_k}$ (5)

- $\sigma_k^2 = c_0 Y_k + c_1$ (6)

where c_0 and c_1 are the coefficients.

- Thus BMA multimodel ensemble mean is a conditional expectation which is defined as:

$$\bar{Y} = E\left[Y|M_{1,} \dots \dots M_k\right] \tag{7}$$

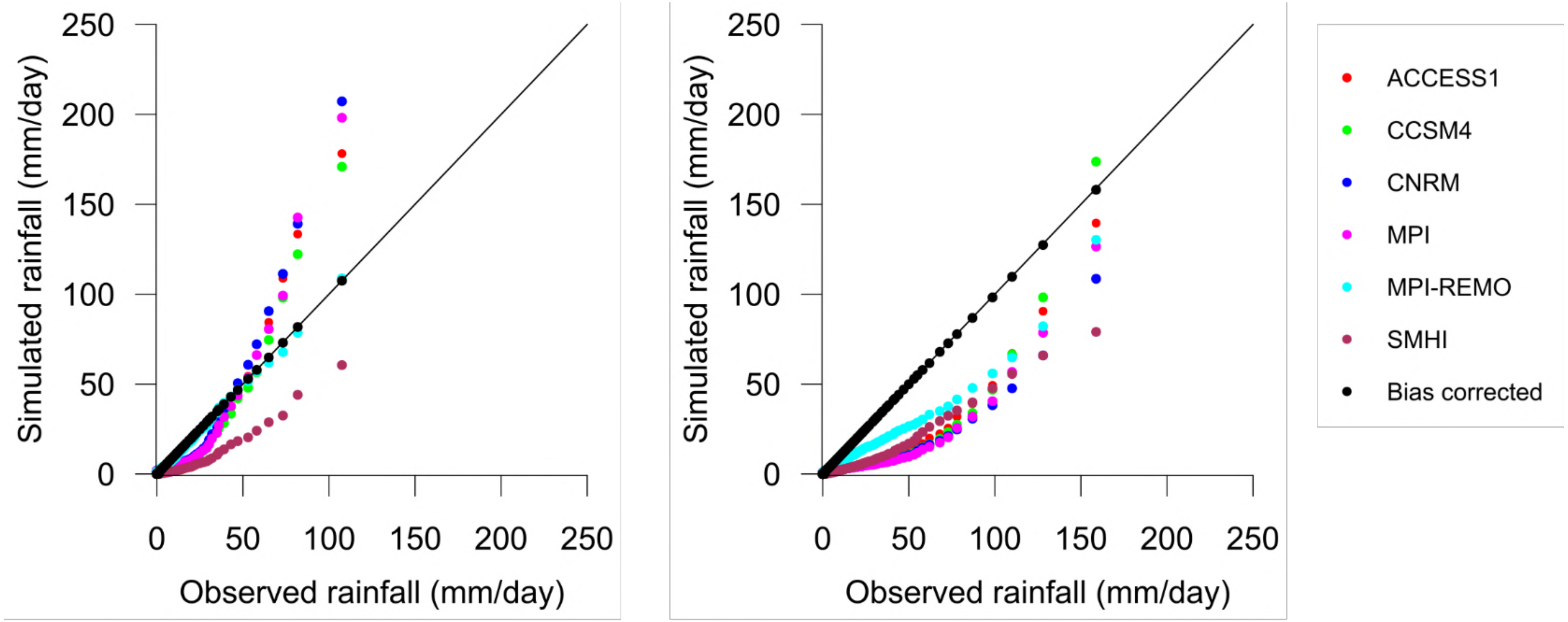
- The values of w_k and σ_k are estimated by the maximum likelihood function (ML) from simulated data set for the training period. The log-likelihood function, $\mathcal{L}$ for the BMA multimodel ensemble mean in equation (7) can be given as:

$$\mathcal{L}\left(w_1, \dots., w_k, \sigma^2 \middle| M_1, \dots M_{k,}\ P(Y|M_K)\right) = \sum_{n=1}^{N} \log\left(\sum_{k=1}^{k} w_k\ P(Y_n|M_{kn})\right) \tag{8}$$

where N is the total number of measurements in the training dataset.

- We optimized the ML function using the Differential Evolution Adaptive Metropolis (DREAM) Markov Chain Monte Carlo (MCMC) algorithm in MATLAB for estimating the BMA weights and variance (Vrugt et. al. 2008, 2015 and 2016).
- The DREAM scheme is capable of running multiple chains simultaneously for searching the global optimal solution.

Bias correction on daily rainfall after modifying wet-day frequency of RCM simulation - Sylhet



(a) Pre-monsoon

(b) Monsoon

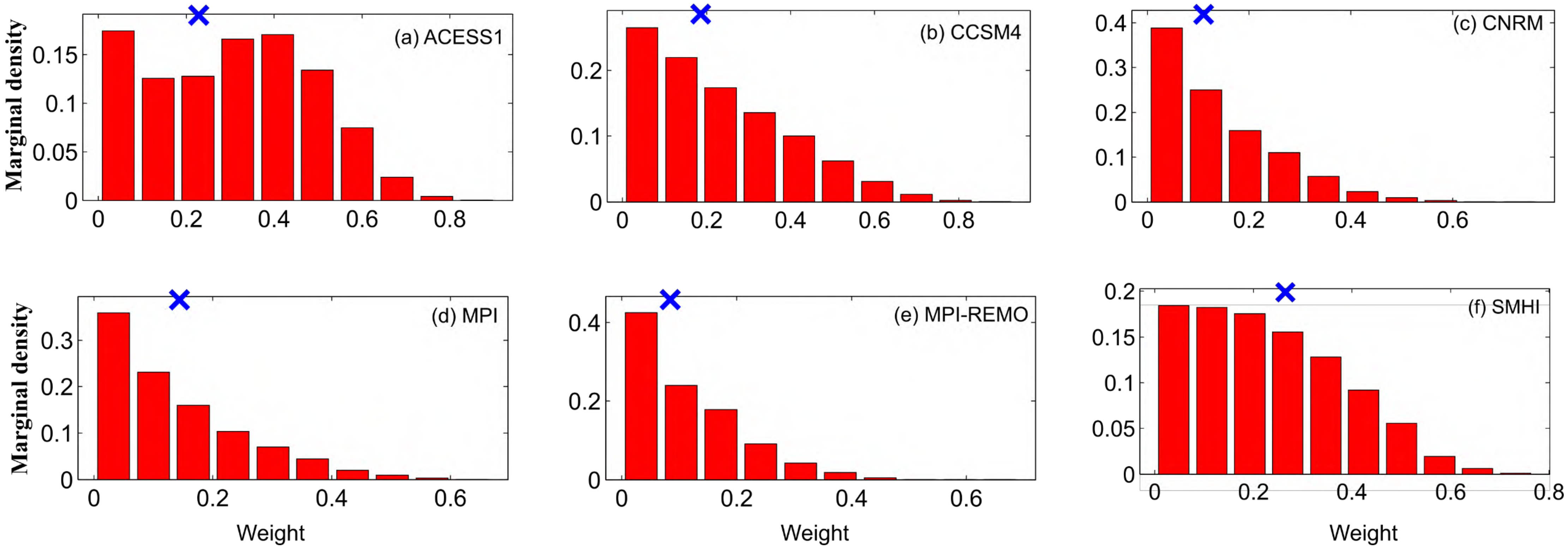
Most models overestimate the observed rainfall at high intensity (>50 mm/d) but underestimate low rainfall intensity. The worst is SMHI.

All 6 RCMs underestimate the seasonal rainfall
Same problem for other 5 stations

Table 3: Seasonal rainfall for Sylhet station before and after bias correction

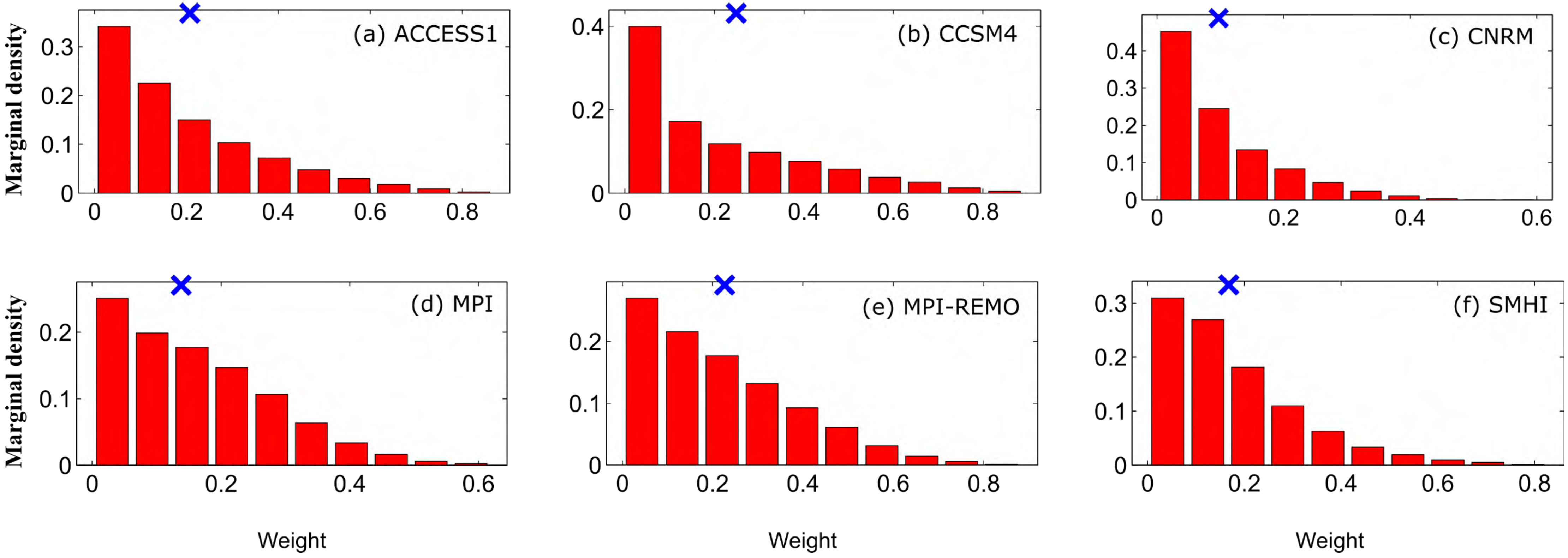
		Observe d	ACCES S	CCSM 4	CNR M	MPI	MPI-REM O	SMH I	
Pre-monsoon	Before bias correction	1087	1116	1050	1239	113 4	1088	386	
	After Bias correction	1087	1090	1072	1079	108 1	1059	1076	28 mm
Monsoon	Before bias correction	2733	1171	1168	976	104 9	1760	982	
	After Bias correction	2733	2707	2700	2710	271 9	2712	2712	33 mm

Posterior PDF of the BMA weights for each RCMs Pre-monsoon - Sylhet



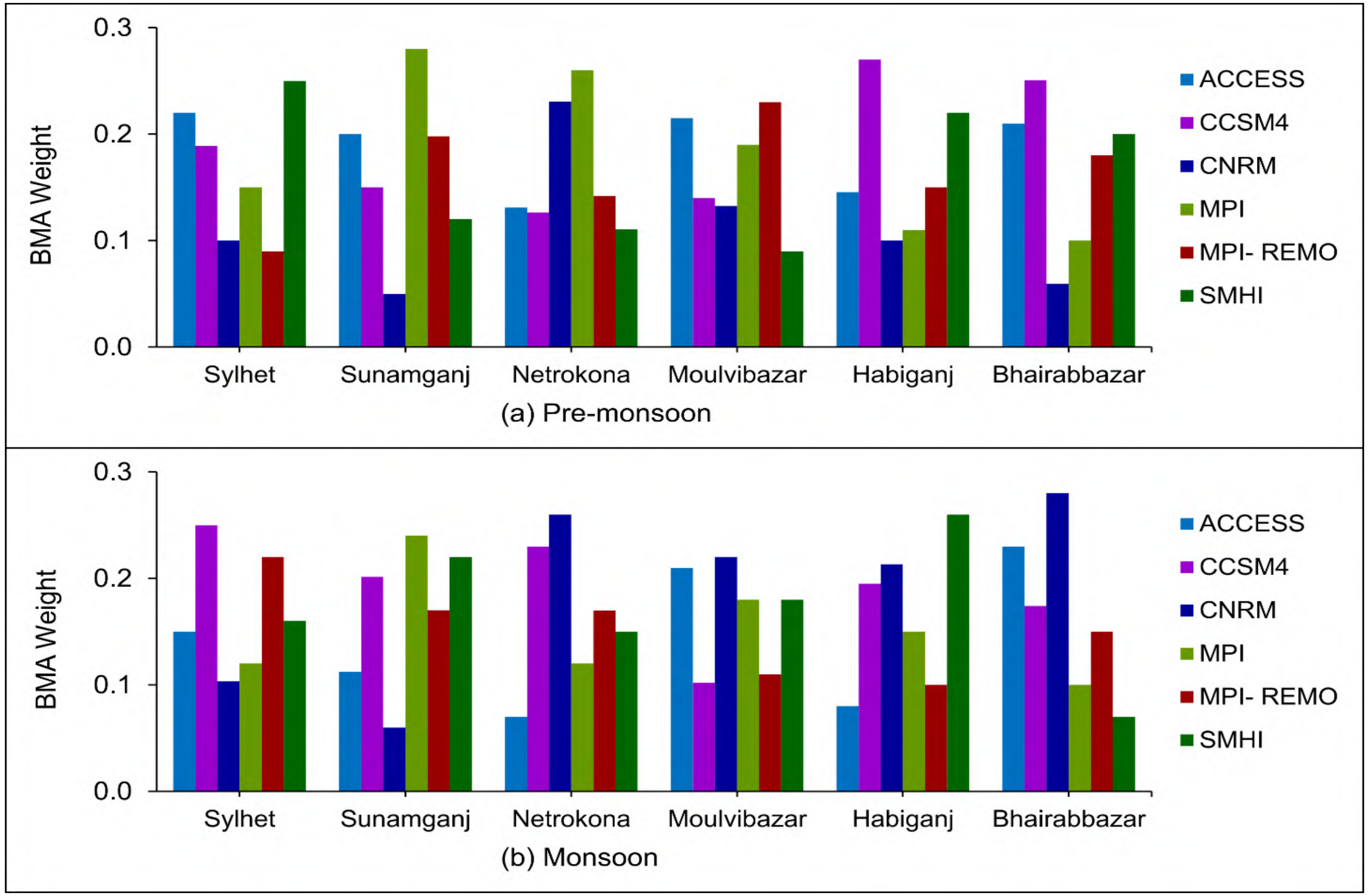
X for optimal weight derived from the MCMC algorithm; weights are between 0.1 and 0.2, except for SMHI (~0.3)

Monsoon - Sylhet



The weights from 6 RCMs appear to be closer to each other, all between 0.1 and 0.2

Optimal BMA weights for all six stations



RCMs show different combinations of BMA weights at different stations

There is no particular RCM in consistently capturing higher BMA weights for all stations

After bias correction, there is no single best or worst RCM

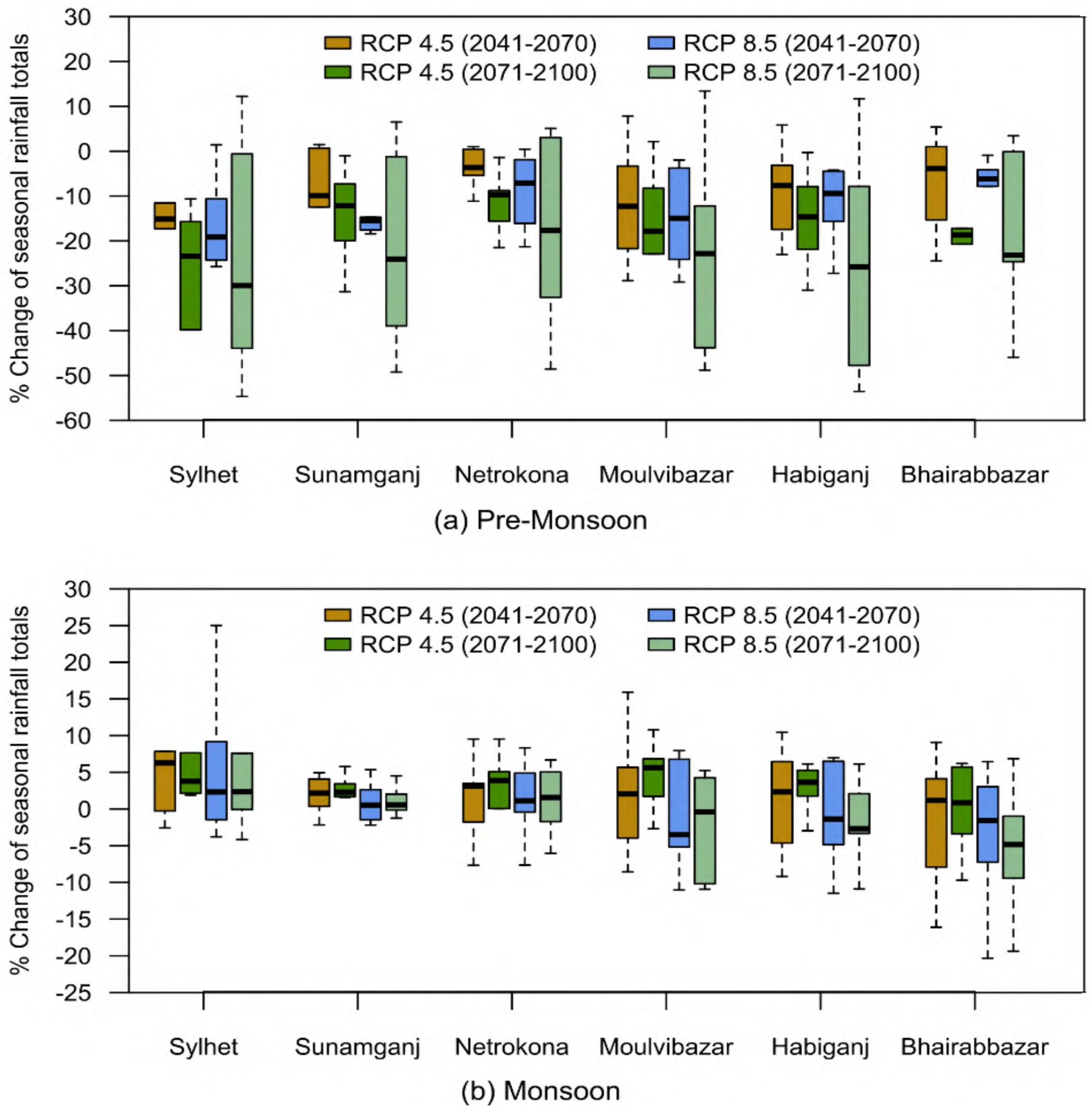
Multimodel ensemble mean – BMA and simple arithmetic ensemble mean (AEM)

Table 4: Normalised Root Mean Square Error (NRMSE) for different RCMs, Arithmetic Ensemble Mean (AEM) and BMA during historical period (1976-2005).

	Station	ACCESS	CCSM4	CNRM	MPI	MPI-REMO	SMHI	AEM	BMA
Pre-Monsoon	Sylhet	1.38	1.35	1.31	1.37	1.33	1.33	1.25	1.04
	Sunamganj	1.34	1.32	1.30	1.31	1.33	1.28	1.23	1.11
	Netrokona	1.30	1.32	1.29	1.31	1.31	1.31	1.13	1.07
	Moulvibazar	1.31	1.27	1.32	1.34	1.33	1.26	1.17	1.02
	Habiganj	1.37	1.34	1.39	1.39	1.36	1.29	1.15	1.07
	Bhairabbazar	1.37	1.35	1.31	1.37	1.39	1.30	1.19	1.09
Monsoon	Sylhet	1.36	1.40	1.44	1.44	1.39	1.42	1.26	1.07
	Sunamganj	1.39	1.40	1.40	1.43	1.40	1.39	1.29	1.11
	Netrokona	1.37	1.36	1.40	1.37	1.39	1.39	1.23	1.13
	Moulvibazar	1.40	1.36	1.47	1.37	1.43	1.36	1.28	1.09
	Habiganj	1.40	1.37	1.40	1.37	1.44	1.39	1.33	1.15
	Bhairabbazar	1.43	1.38	1.42	1.43	1.46	1.43	1.25	1.04

NRMSE: RMSE is normalized by the standard deviation of observed data

For all 6 stations, the average percentage of decrease in NRMSE from AEM to BMA is 10% during the pre-monsoon and 22% during the monsoon season.

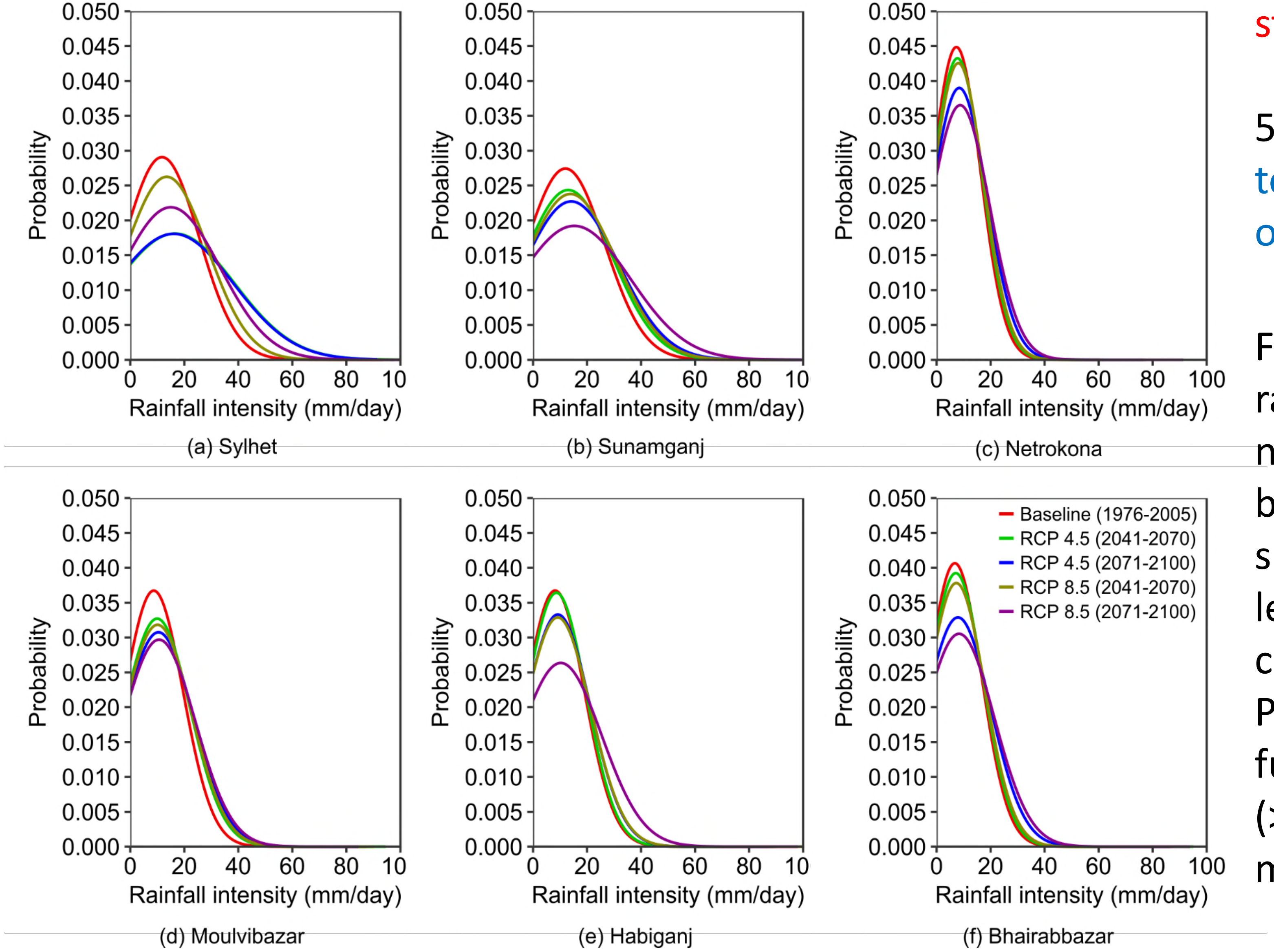


Projection of future seasonal rainfall change in % for 6 stations for the pre-monsoon (top) and monsoon (bottom) relative to the baseline (1976-2005)

4 scenarios; near future (2041-2070); far future (2071-2100)

A decrease in pre-monsoon rainfall for all 6 stations; larger interquartile range and spread for RCP8.5 for far future (larger uncertainty)

Smaller changes during the monsoon season, also smaller spread for the first 3 wetter stations

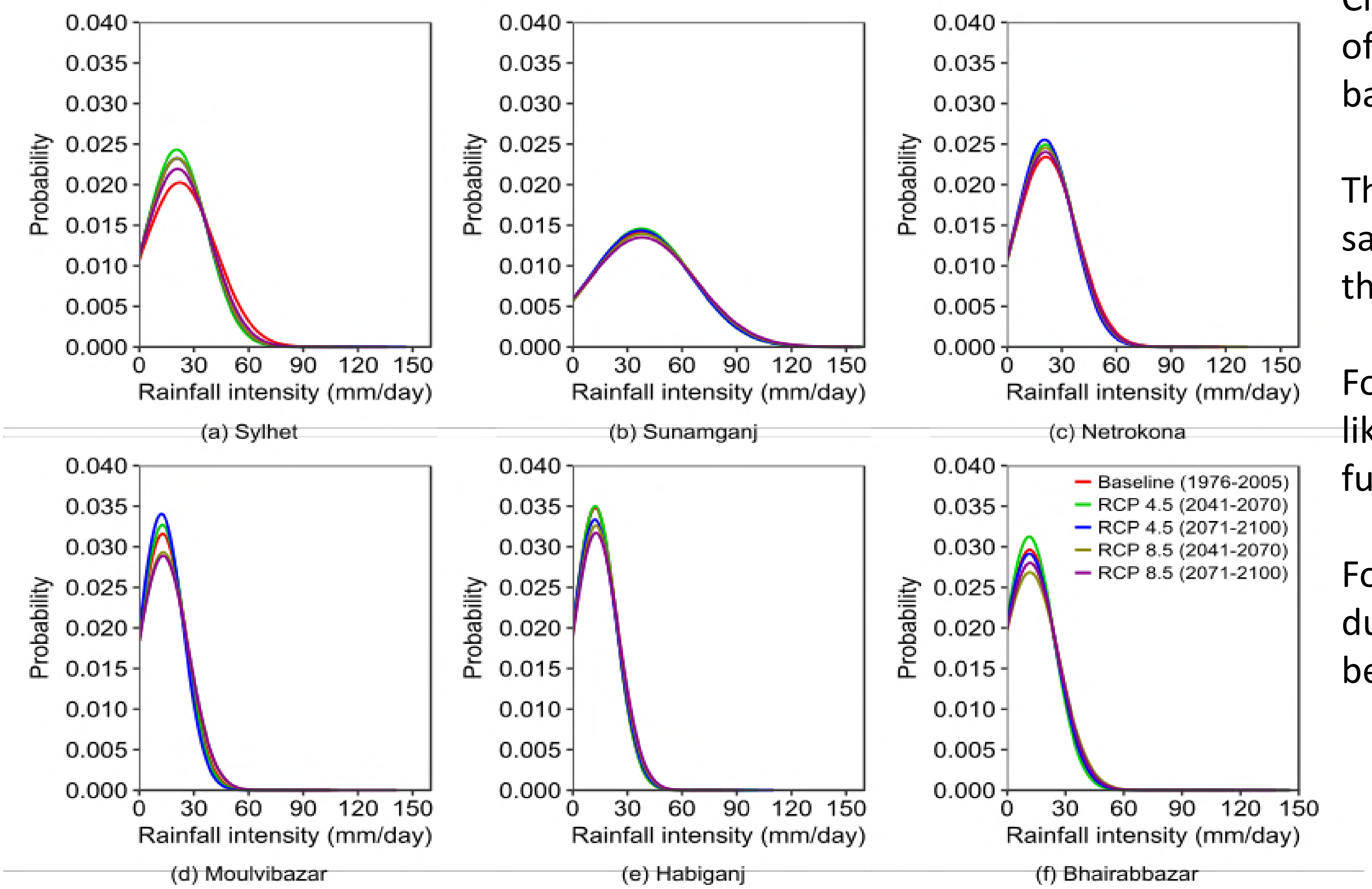


PDF of rainfall intensity for 6 stations during pre-monsoon

5 PDFs in each panel SDII- seasonal total rainfall divided by the number of wet days in a season (mm/day)

For Sylhet and Sunamganj where rainfall is high, the peak intensity is near 10 to 20 mm/day during the baseline. This peak is expected to shift towards higher intensity but less frequent from the current climate to the far future. The entire PDF shifts more to the right into the future: extreme rainfall intensity (>60 mm/day and above) may occur more often in the future.

Rainfall intensity - Monsoon



Change in the peak value and the shift of PDF is not so clear from the baseline to the future.

The shape of PDF is essentially the same for 3 stations (b, c, and e) from the current climate into the future.

For Sylhet, the peak (20-30 mm/d) is likely to occur more often in the future.

For Moulvibazar, the peak distribution during the baseline lies somewhere between two far future scenarios

Summary

- Quantile mapping is used to correct bias in rainfall for six RCMs over the CORDEX South Asia domain

- A BMA method is applied to optimally assign weights for RCMs based on their predictive skill during the training period. The BMA shows the lowest NRMSE compared to each individual model and the AEM during both the pre-monsoon and monsoon seasons.

- A decrease in pre-monsoon rainfall for all 6 stations in the future but uncertainty is also larger in the far future. Smaller changes during monsoon.

- The PDF of rainfall intensity during the pre-monsoon appears to shift to the right in the future.