Seasonal Predictability Analysis of DEMETER Project Models and Statistical Downscaling

Avinash Chandra Pandey

Ravi P Shukla
Sandipan Mukherjee
K Banerjee Centre of Atmos & Ocean Studies
University of Allahabad

Targeted Training Activity 2008

Outline

- Scientific Motivation
- Objectives
- Experimental Plan
- Results & Discussions
- Feedback TTAs

Different Approaches for the Prediction of Seasonal Weather Phenomena

- Dynamical seasonal prediction is used for the predictability of seasonal mean atmospheric circulation and rainfall. For each of the models (NCAR, COLA, GSFC, GFDA, NCEP) SST forced variance (signal), internal dynamics variance (noise), total variance, correlation coefficient with observation and probability distribution was calculated.

  It was found that with specified SST boundary condition, all model showed that winter seasonal mean circulation over Pacific north American (PNA) region were highly predictable during the year of large tropical sea-surface temperature anomaly. The degree of reproducibility is highly variable from one model to other model and quantity such as signal to noise ratio vary significantly between the different AGCMs in PNA region. *"Dynamical Seasonal Prediction"* by *Shukla J et. al.(2000),*

- Canonical Correlation Analysis (CCA) for the prediction of ENSO events for different parts of Pacific Ocean using MSLP and SST as a predictor and it was found that *large region in the eastern equatorial Pacific has a highest overall predictability with the best result obtained for the winter. "Prediction of ENSO episode using Canonical Correlation Analysis " Barnston A G & Ropelewski C F*

- Wavelet banding and Linear regression (WBLR) was used as the prediction scheme. Pentad rainfall was predicted using four different predictors and wavelet banding was used to decompose the signal and then linear regression was used for prediction.

  The forecast compared with Artificial Neural Network (ANN) and Multivariate adaptive regression spline model (MARS) and it was found that WBLR was giving better results.

  "Prediction of Monsoon Rainfall and River Discharge on 15-30 Day Time span" by *Webster P J and Hoyos C (2004),*

- Multiple regression was used to determine coefficients from the multimodel forecasts and observation. The coefficients were the used in super ensemble technique. The super ensemble is superior in terms of forecasts to an ensemble mean.

  "Improved Weather and Seasonal Climate Forecasts from Multimodel Superensemble" by *Krishnamurti T N et al. (1999).*

Predictability analysis of the DEMETER model output for Seasonal Precipitation over Indian Region

In the present study we have performed the following analysis:

1. We have calculated the root mean square error (RMSE), correlation coefficient of individual model and ensemble mean and standard deviation of the observation.

2. Signal to noise ratio and Zonal mean distribution of theoretical limit of correlation skill for Models has been calculated for the individual model.

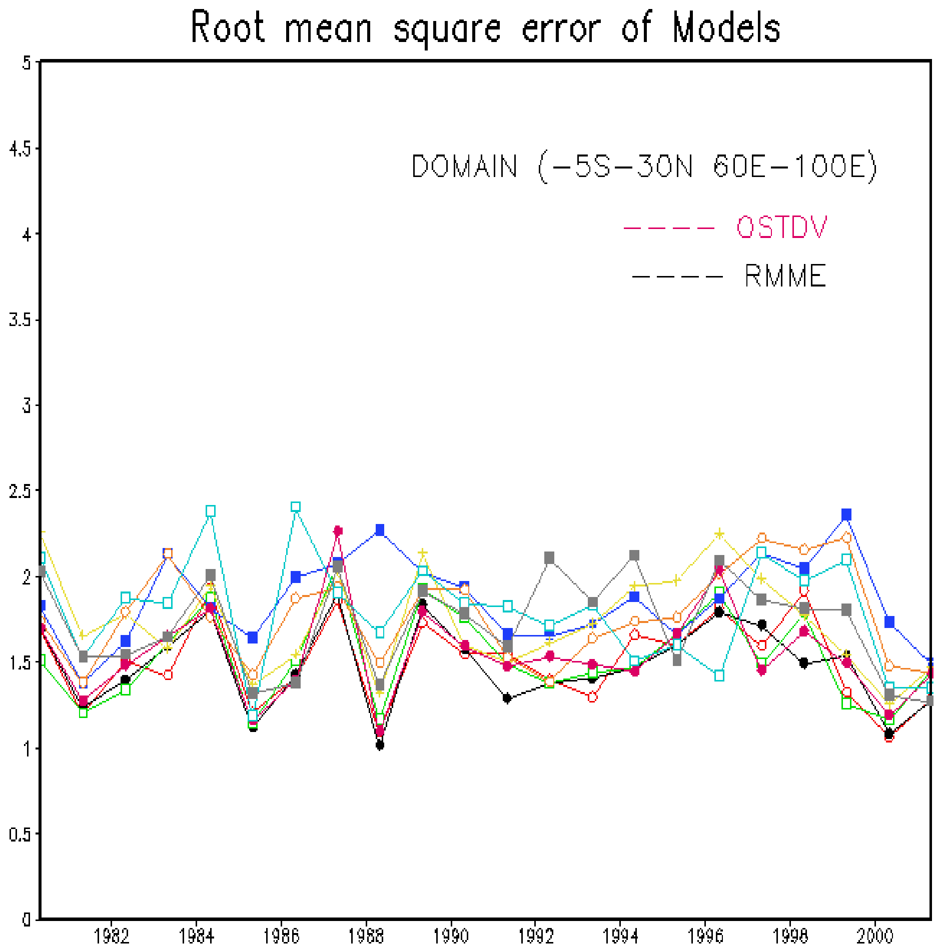
3. Empirical Orthogonal Function (EOF) analysis has been performed for the ensemble mean and observation data.

4. Lastly Principal component regression is used to improve the prediction skill of the individual model for the specific station (Allahabad)

Brief description of the seven ocean-atmosphere coupled models of DEMETER, SNU,CRU,NCEP

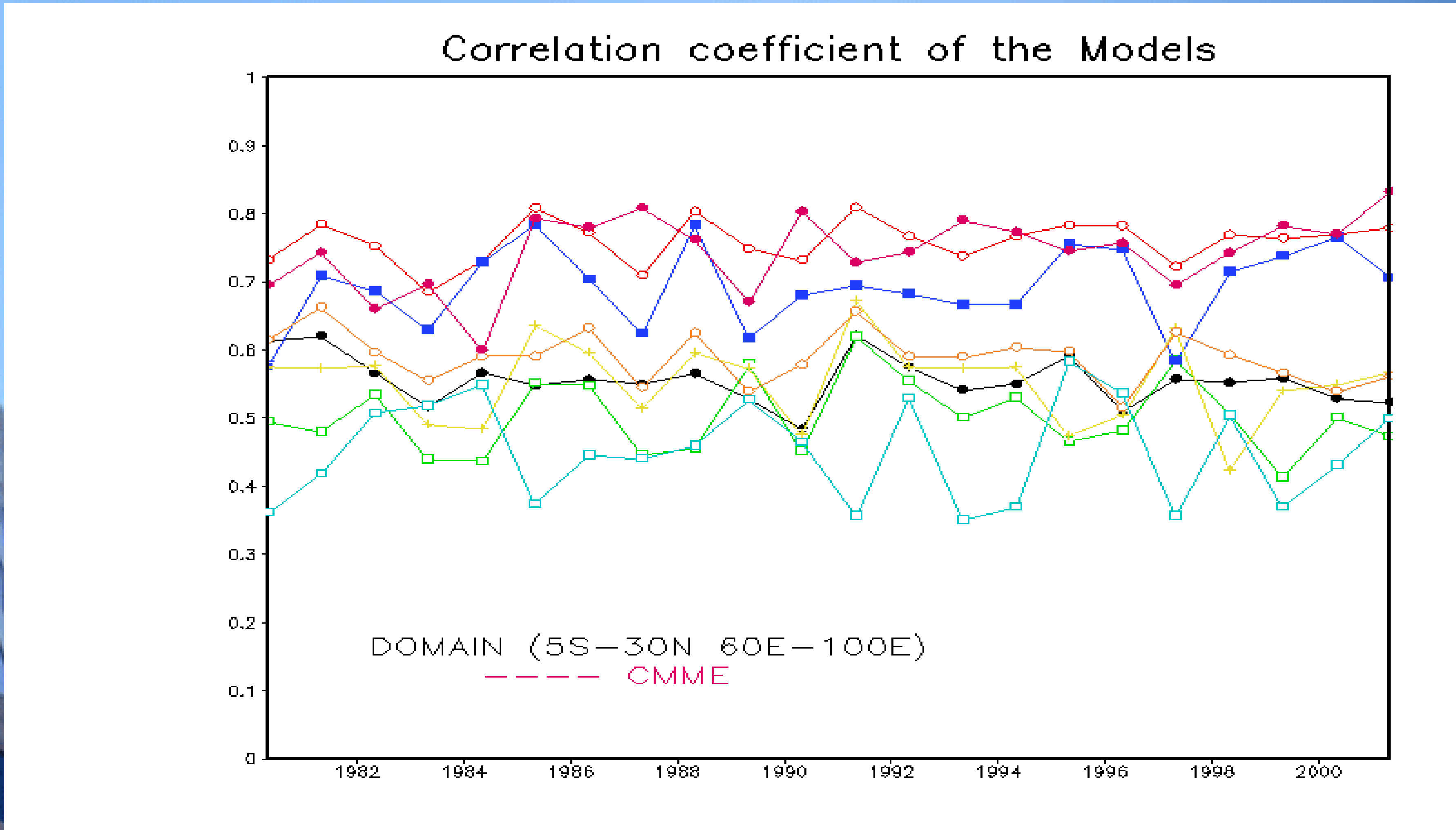
Institute	AGCM	Resolution	Atmosphere IC	OGCM	Resolution	Ocean IC
CERFACS (CERF)	ARPEGE	T63 31 levels	ERA-40	OPA 8.2	2.0° × 2.0° 31 levels	Ocean Analysis Forced by ERA-40
ECMWF (ECMW)	IFS	T95 40 levels	ERA-40	HOPE-E	1.4° × 0.3°-1.4° 29 levels	Ocean Analysis Forced by ERA-40
INGV	ECHAM-4	T42 19 levels	Coupled AMIP-type experiment	OPA 8.1	2.0° × 0.5°-1.5° 31 levels	Ocean Analysis Forced by ERA-40
LODYC (LODY)	IFS	T95 40 levels	ERA-40	OPA 8.2	2.0° × 2.0° 31 levels	Ocean Analysis Forced by ERA-40
Meteo-France (METF)	ARPEGE	T63 31 levels	ERA-40	OPA 8.0	182 GP × 152 GP 31 levels	Ocean Analysis Forced by ERA-40
Met Office (UKMO)	HadAM3	2.5° ×3.75° 19 levels	ERA-40	Glo Sea OGCM based on HadCM3	1.25° × 0.3°-1.25° 40 levels	Ocean Analysis Forced by ERA-40
MPI (MAXP)	ECHAM-5	T42 19 levels	Coupled run relaxed to observed SSTs	MPI-OMI	2.5° × 0.5°-2.5° 23 levels	Coupled run relaxed to observed SSTs

Root mean square error between the model produced precipitation (JJA) and CMAP observation for each model and MME of DEMETER project and standard deviation of observation (CMAP)



- The area chosen for analysis is: 5S – 30N, 60E – 100E for the period of 1980 – 2001.
- The RMSE of MME is less than the standard deviation of the observation data in comparison to the other individual models.
- In this case the MME has performed better than the other individual models.

Correlation coefficient between the model produced precipitation and CMAP observation for each model and MME of DEMETER project



Formulae used for calculation of signal to noise ratio of DEMETER project models

- The internal dynamics variance is given by:

$$\sigma^2_{noise} = \frac{1}{N(n-1)} \sum_{i>1}^{N} \sum_{j>1}^{n} (x_{ij} - \overline{x_i})^2$$

- The signal is given by:

$$\sigma^2_{signal} = \sigma^2_{EM} - \frac{1}{n}\sigma^2_{noise}$$

- The variance of ensemble mean is :

$$\sigma^2_{EM} = \frac{1}{N-1} \sum_{i>1}^{n} (x_i - \overline{\overline{x}})^2$$

['*Dynamical Seasonal Prediction*' by J Shukla et al.,2000.]

- From Rowell et al (1995), the total variance:

$$\sigma^2_{total} = \sigma^2_{noise} + \sigma^2_{signal}$$

- The theoretical limit of seasonal prediction correlation skill expressed in terms of signal to noise ratio (ρ): [From Kang I S and Shukla J, 2006]

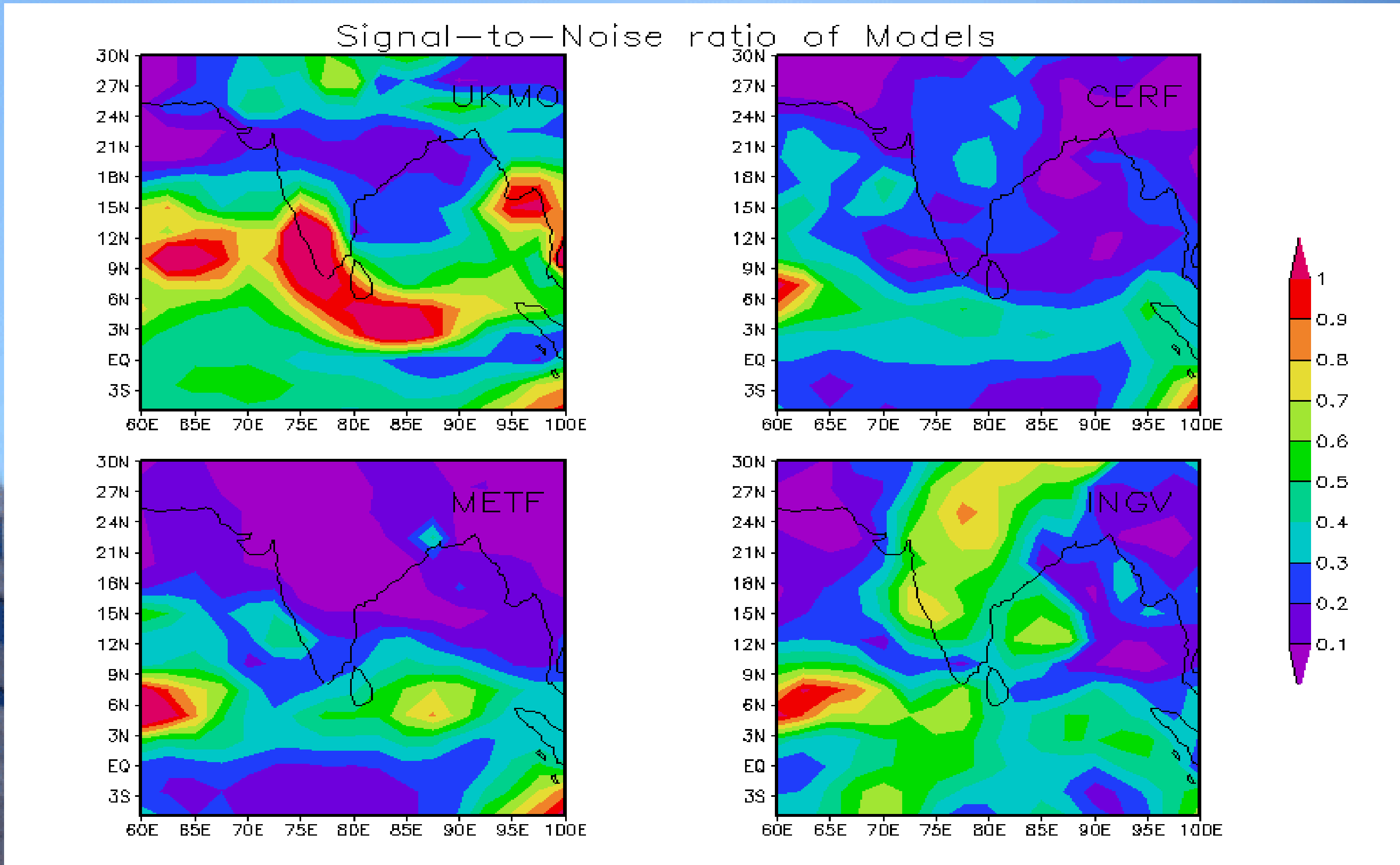
$$R_{limit} = \sqrt{\frac{\rho}{\rho + 1}}$$

- For a seasonal mean climate variable *xij*(rainfall) for N years (*i* = 1,2,…….,N) and *n* ensemble members (*j* = 1,2,…..,*n*). Ensemble mean and climatological (ensemble) mean are defined as:

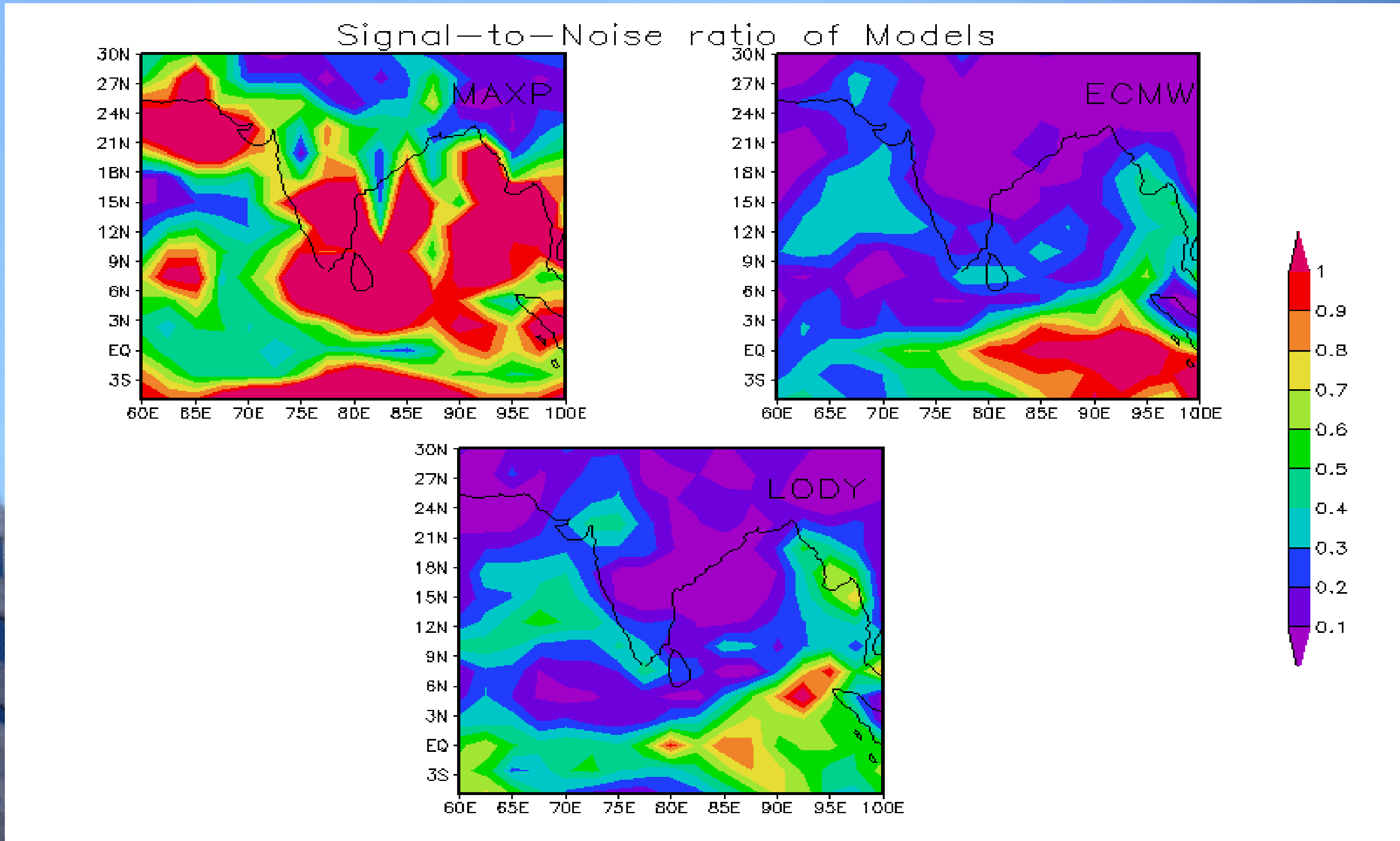
$$\overline{x_i} = \frac{1}{n} \sum_{j>1}^{n} x_{ij} \qquad \overline{\overline{x}} = \frac{1}{Nn} \sum_{i>1}^{N} \sum_{j>1}^{n} x_{ij}$$

The following figure shows the signal to noise ratio of Models

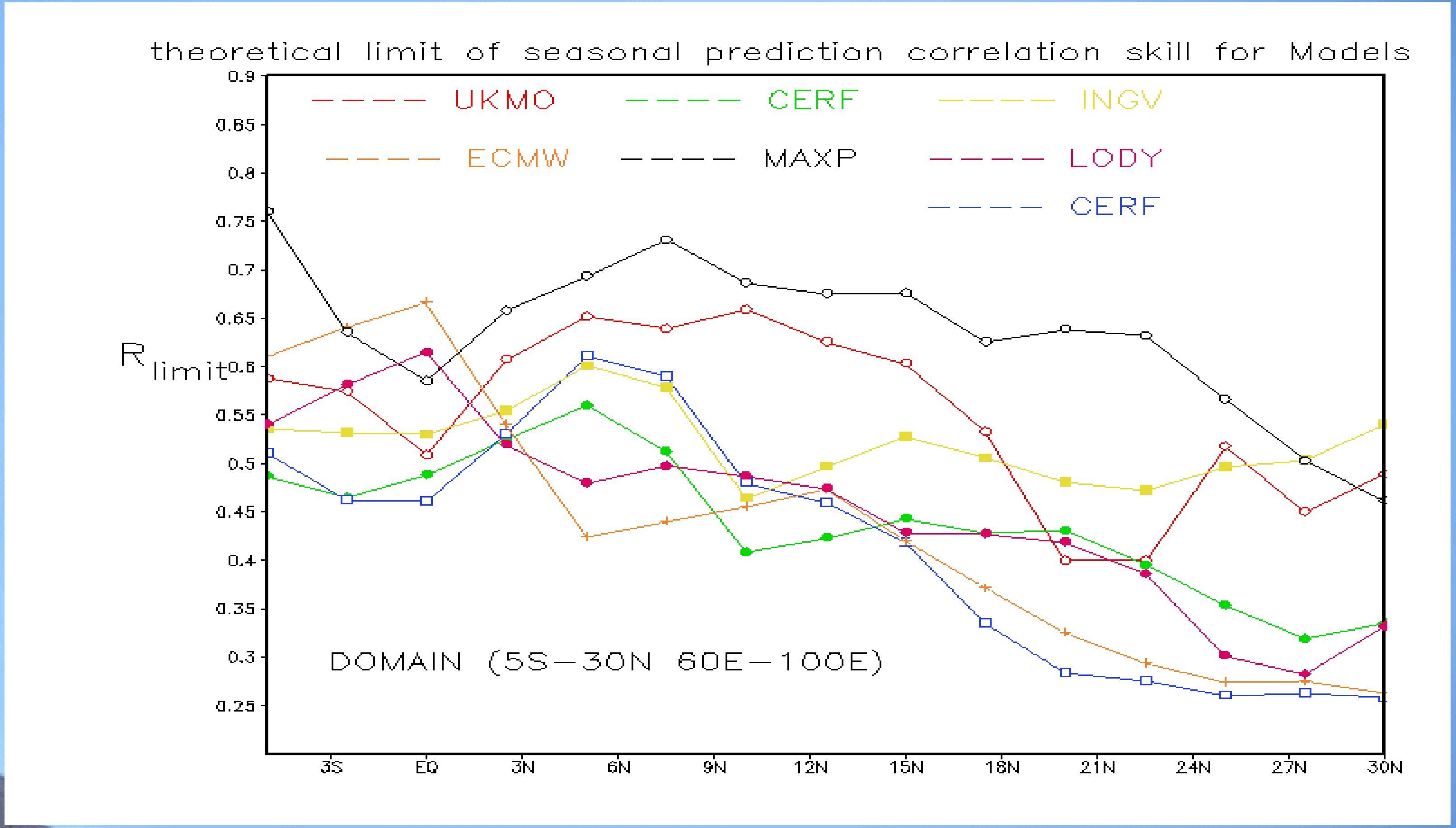
(UKMO, CERF, METF, INGV)



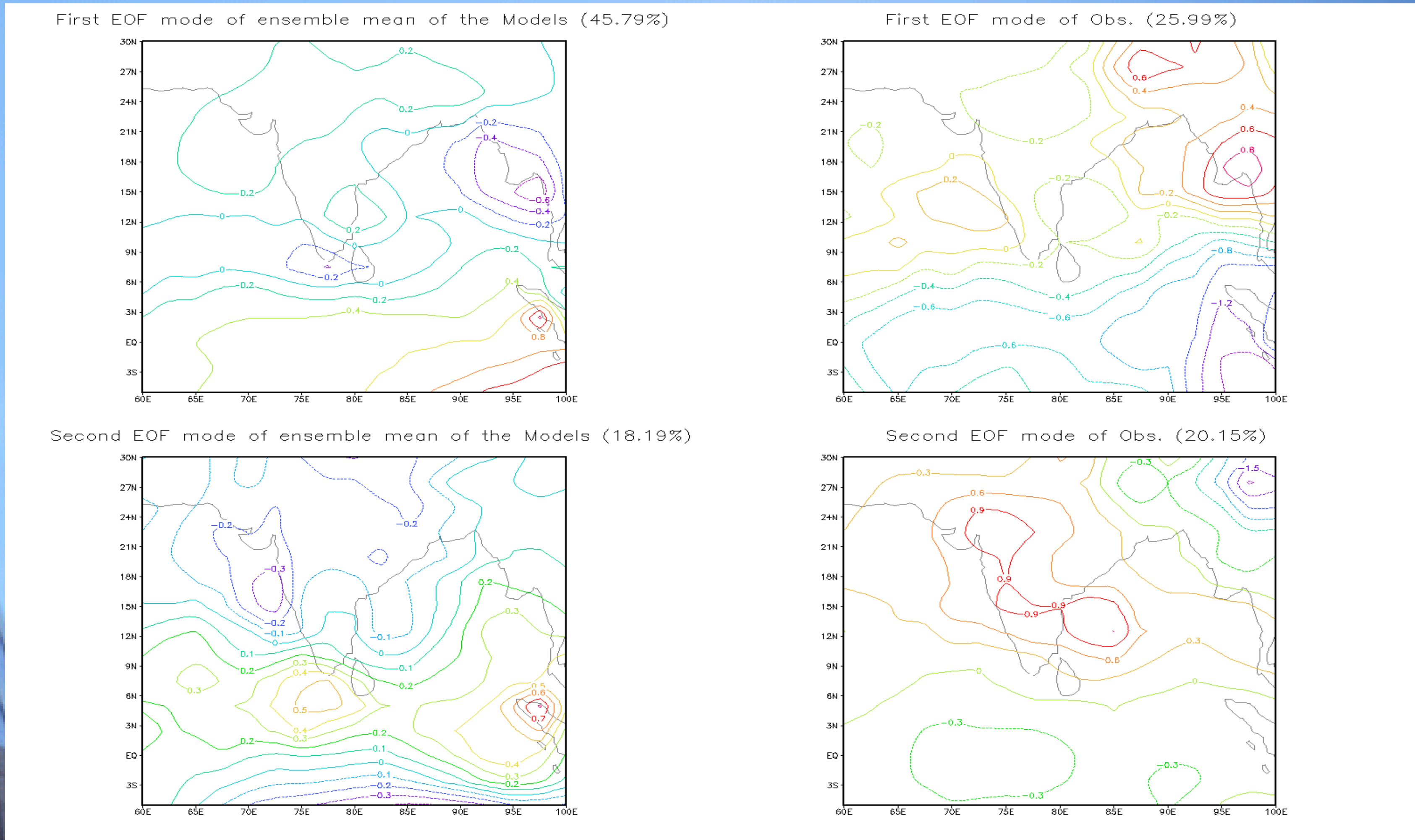
signal to noise ratio of Models
(MAXP, ECMW, LODY)



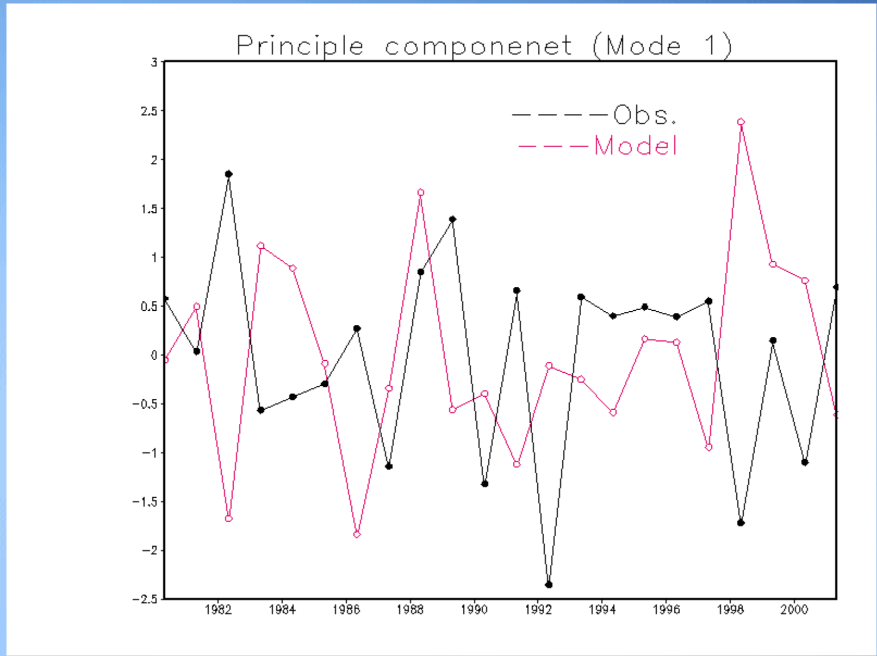
Zonal mean distribution of theoretical limit of correlation skill for Models

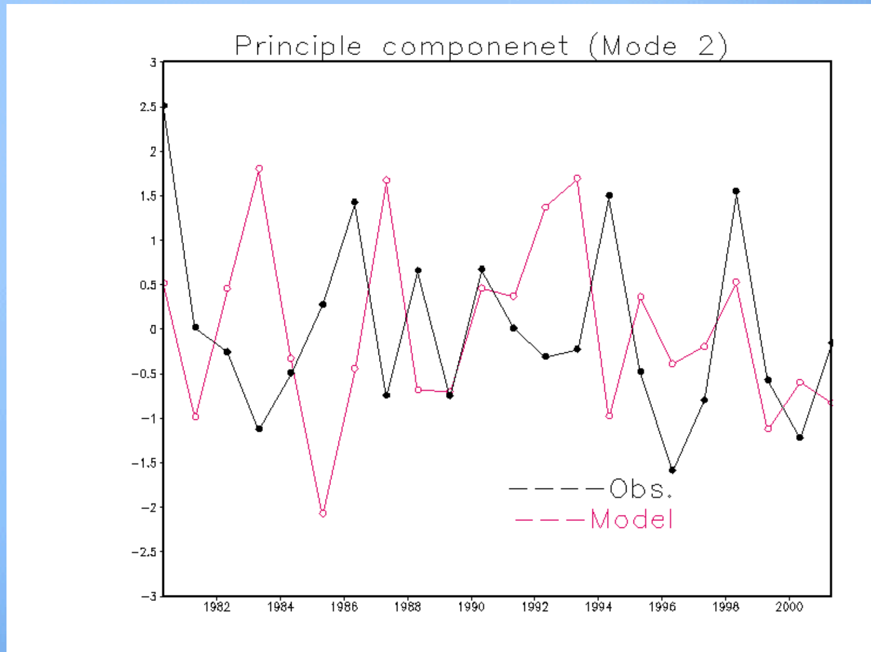


EOF modes of the observed and simulated ensemble mean precipitation



Principal component of mode-1 and mode-2 for observed and ensemble mean of models





Different statistical forecast skills are performed with the following formulas:

- **Hit (H):** Forecast to occur and did occur
- **Miss (M):** Forecast not to occur but did occur
- **False alarm (F):** Forecast to occur but did not occur
- **Correct negative (N):** Event forecast not to occur and did not occur.

- *False alarm ratio* (Far) = F/(H+F) ;

- *Bias* = (H+F) / (H+M);

- *Accuracy* = (H+N) / Total;   Total = (H + M + N +F);

- *Probability of false detection (Pofd)* = F / (N + F);

- *Probability of detection (*Pod) = H / (H + M);

- *Threat score (Ts)* = H / (H + M +F);

- *Heidke skill score (Hss)* = ((H + N) – (correct random)) / (Total - (correct random))
  correct random = [(H+M)(H+F) + (N+M)(N+F)] / Total;

- *Equitable threat score (Ets)* = (H – H random) / (H +M +F – H random)
  H random = [(H+M)(H+F)] / Total;

Models	Accuracy	Bias	POD	FAR	POFD	TS	ETS	HSS
CERF	0.63	1.40	0.40	0.71	0.29	0.20	0.49	0.09
ECMW	0.59	1.60	0.40	0.75	0.35	0.18	0.20	0.39
INGV	0.41	1.60	0.0	1.0	0.47	0.0	-0.16	-0.38
LODY	0.50	1.60	0.20	0.87	0.41	0.08	-0.08	-0.17
MAXP	0.50	1.60	0.20	0.87	0.41	0.08	-0.08	-0.17
METF	0.72	1.40	0.60	0.57	0.23	0.33	0.19	0.31
UKMO	0.68	1.60	0.60	0.62	0.29	0.30	0.14	0.25

Categorical statistics computed for Total Rainfall obtained from *CMAP observations for Allahabad* from individual models.

Categorical statistics computed for Total Rainfall obtained from *Station Data observations for Allahabad* from individual models.

Models	Accuracy	Bias	POD	FAR	POFD	TS	ETS	HSS
CERF	0.45	1.0	0.14	0.86	0.40	.08	-0.11	-0.26
ECMW	0.59	1.14	0.43	0.62	0.33	0.25	0.05	0.09
INGV	0.59	1.14	0.43	0.62	0.33	0.25	0.05	0.09
LODY	0.36	1.0	0.0	1.0	0.47	0.0	-0.18	-0.47
MAXP	0.50	1.14	0.28	0.75	0.40	0.15	-0.05	-0.11
METF	0.54	1.0	0.28	0.71	0.33	0.16	-0.02	-0.05
UKMO	0.59	1.14	0.43	0.62	0.33	0.25	0.05	0.09

Correlation coefficient between models produced precipitation and Observations

(CMAP and Station Data) for each model of DEMETER project

Station Name		CERF	ECMW	INGV	LODY	MAXP	METF	UKMO
Allahabad	CMAP Obs.	0.17	0.14	0.06	0.02	0.23	-0.02	0.44
	Station Obs.	0.44	0.18	-0.06	0.22	0.06	0.42	0.51
Allahabad (after downscaling) Region of predictor (20-30N and 75-90E)	CMAP Obs.	0.32	-0.02	0.46	0.13	0.06	0.36	0.04
	Station Obs.	0.63	0.13	0.42	0.11	0.31	0.58	0.36
Allahabad (after downscaling) Region of predictor (5S-30N and 60-100E)	CMAP Obs.	0.18	0.20	0.83	0.25	0.45	0.60	0.30
	Station Obs.	0.22	0.45	0.21	0.32	0.52	0.56	0.53

Models	Accuracy	Bias	POD	FAR	POFD	TS	ETS	HSS
CERF	0.77	1.20	0.60	0.5	0.18	0.38	0.25	0.40
ECMW	0.77	0.0	0.00	NA	0.0	0.0	0.0	0.0
INGV	0.82	0.60	0.4	0.33	0.06	0.33	0.24	0.39
LODY	0.73	0.60	0.20	0.67	0.12	0.14	0.05	0.09
MAXP	0.45	1.40	0.0	1.0	0.41	0.0	-0.15	-0.36
METF	0.59	1.20	0.2	0.83	0.29	0.10	-0.04	-0.08
UKMO	0.72	0.60	0.2	0.66	0.11	0.14	0.05	0.09

Categorical statistics computed for Total Rainfall obtained from *downscaled observations for Allahabad* from individual models. The predictor parameter is MSLP of the region *(20-30N and 75-90E)* for the month of May.

Categorical statistics computed for Total Rainfall obtained from *downscaled CMAP observations for Allahabad* from individual models. The predictor parameter is MSLP of the region *(20-30N and 75-90E)* for the month of May.

Models	Accuracy	Bias	POD	FAR	POFD	TS	ETS	HSS
CERF	0.54	1.0	0.28	0.71	0.33	0.16	-0.02	-0.04
ECMW	0.68	0	0.0	NA	0.0	0.0	0.0	0.0
INGV	0.63	0.71	0.28	0.60	0.20	0.20	0.05	0.09
LODY	0.32	1.14	0	1	0.53	0.0	-0.20	-0.51
MAXP	0.36	1.57	0.28	0.81	0.60	0.12	-0.12	-0.27
METF	0.59	0.85	0.28	0.66	0.26	0.18	0.01	0.02
UKMO	0.63	0.14	0.0	1.0	0.06	0.0	-0.04	-0.08

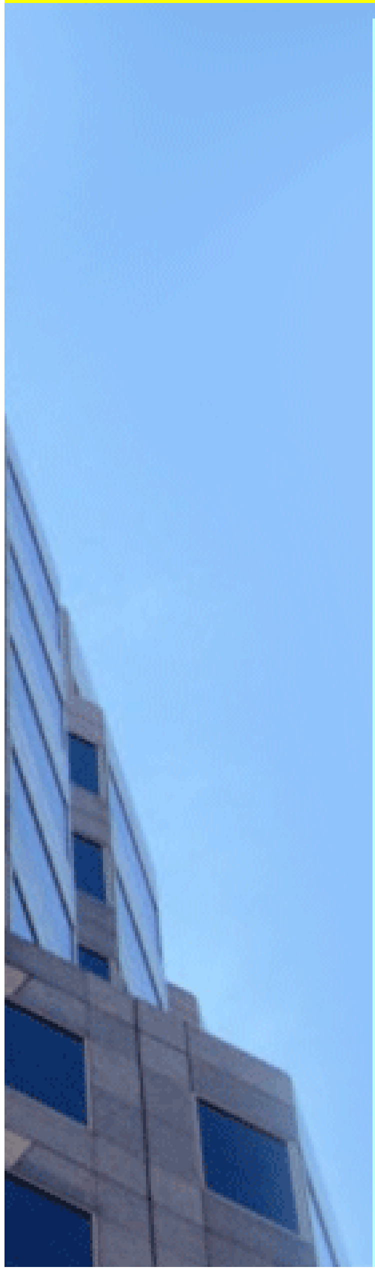
Models	Accuracy	Bias	POD	FAR	POFD	TS	ETS	HSS	CC
CERF	0.73	1	0.40	0.60	0.18	0.25	0.13	0.22	0.22
ECMW	0.64	1	0.20	0.80	0.24	0.11	-0.02	-0.04	0.45
INGV	0.68	0.80	0.20	0.75	0.18	0.01	0.01	0.03	0.21
LODY	0.46	1.4	0.0	1.0	0.41	0.0	-0.15	-0.36	0.32
METFA	0.73	0.60	0.20	0.67	0.12	0.14	0.05	0.09	0.56
MAXP	0.68	0.40	0.0	1	0.12	0.0	-0.07	-0.15	0.52
UKMO	0.73	1.4	0.60	0.57	0.24	0.33	0.19	0.32	0.53

Categorical statistics computed for Total Rainfall obtained from *downscaled observations for Allahabad* from individual models. The predictor parameter is MSLP of the region *(5S-30N and 60-100E)* for the month of May.

Categorical statistics computed for Total Rainfall obtained from *downscaled CMAP observations for Allahabad* from individual models. The predictor parameter is MSLP of the region *(5S-30N and 60-100E)* for the month of May

Models	Accuracy	Bias	POD	FAR	POFD	TS	ETS	HSS	CC
CERF	0.54	0.43	0.0	1.0	0.20	0.0	-0.10	-0.24	0.18
ECMW	0.54	0.71	0.14	0.80	0.27	0.09	-0.06	-0.13	0.20
INGV	0.82	0.71	0.57	0.20	0.07	0.5	0.37	0.54	0.83
LODY	0.45	0.71	0.0	1.0	0.33	0.0	-0.15	-0.36	0.25
METFA	0.54	1.0	0.29	0.71	0.33	0.17	-0.02	-0.04	0.60
MAXP	0.68	0.28	0.14	0.50	0.06	0.12	0.05	0.09	0.45
UKMO	0.68	0.57	0.28	0.50	0.13	0.22	0.09	0.17	0.30

CONCLUSION



- The RMSE of the seasonal precipitation (JJA) of the MME is lesser than the other individual models and the correlation coefficient is higher than the models indicating that the MME is providing better result.

- The signal to noise ratio of the model MAXP and UKMO is higher than the other models. Zonal mean distribution of theoretical limit of correlation skill for different models indicating that the predictability of all the models are relatively higher near the equator but it rapidly drops near 30N. In this case also the MAXP and the UKMO is better than the others.

- The first EOF of the MME is explaining 45% of the total variance and first EOF of the observed precipitation is explaining 25.99% of the total variance.

- To get a better result statistical down scaling (Climate Prediction Tool) is applied for two different predictor domain. Principal component regression of the predictor and the predictand is providing better results than individual models.

Feedback TTAs

Helped in Nucleating a Nascent Centre as an Offshoot of Physics Department

Research & Student Progression

PhDs

Suneet Dwivedi now at John Hopkins, USA

Mathematical Modeling of Atmosphere; Predictability and Feasibility

***Shailendra Rai now at COLA, USA**

“Simulation studies of Southern Indian Ocean”

***Vivek Kumar Pandey now at IISc, Bangalore**

Numerical Experiments in Ocean Transport Phenomenon

***Anshu Prakash Mishra now with CWC, Govt of India**

“ Numerical Study of Interaction Processes of Southern Indian Ocean”

Publications in international & National Journals: 30

Publications in Proceedings: 10

Annual reports:15

Thanks

K Banerjee Centre of Atmos. & Ocean Studies, University of Allahabad