Downscaling
&
Bias correction

Dr. V.R.DURAI

SCIENTIST-E

NWP DIVISION

vrdurai@gmail.com

भारत मौसम विज्ञान विभाग

INDIA METEOROLOGICAL DEPARTMENT

Beyond Interpolation to Post-Processed Grids

Deterministic NWP

Raw NWP model data is further processed beyond mere interpolation to pressure levels or to nearest neighbor points, to deal with model deficiencies of various kinds.

These include,

1. **Downscaling - (typically based on data sets such as MOS developed from high resolution surface data, and high resolution topography)**
2. **Bias correction - (typically based on recent model performance history)**

STATISTICAL GUIDANCE

Statistical Guidance Approaches

❖There are two Statistical Guidance approaches that are most commonly applied to NWP output for use as operational guidance,

1. Perfect Prog (PP) and
2. Model Output Statistics (MOS)

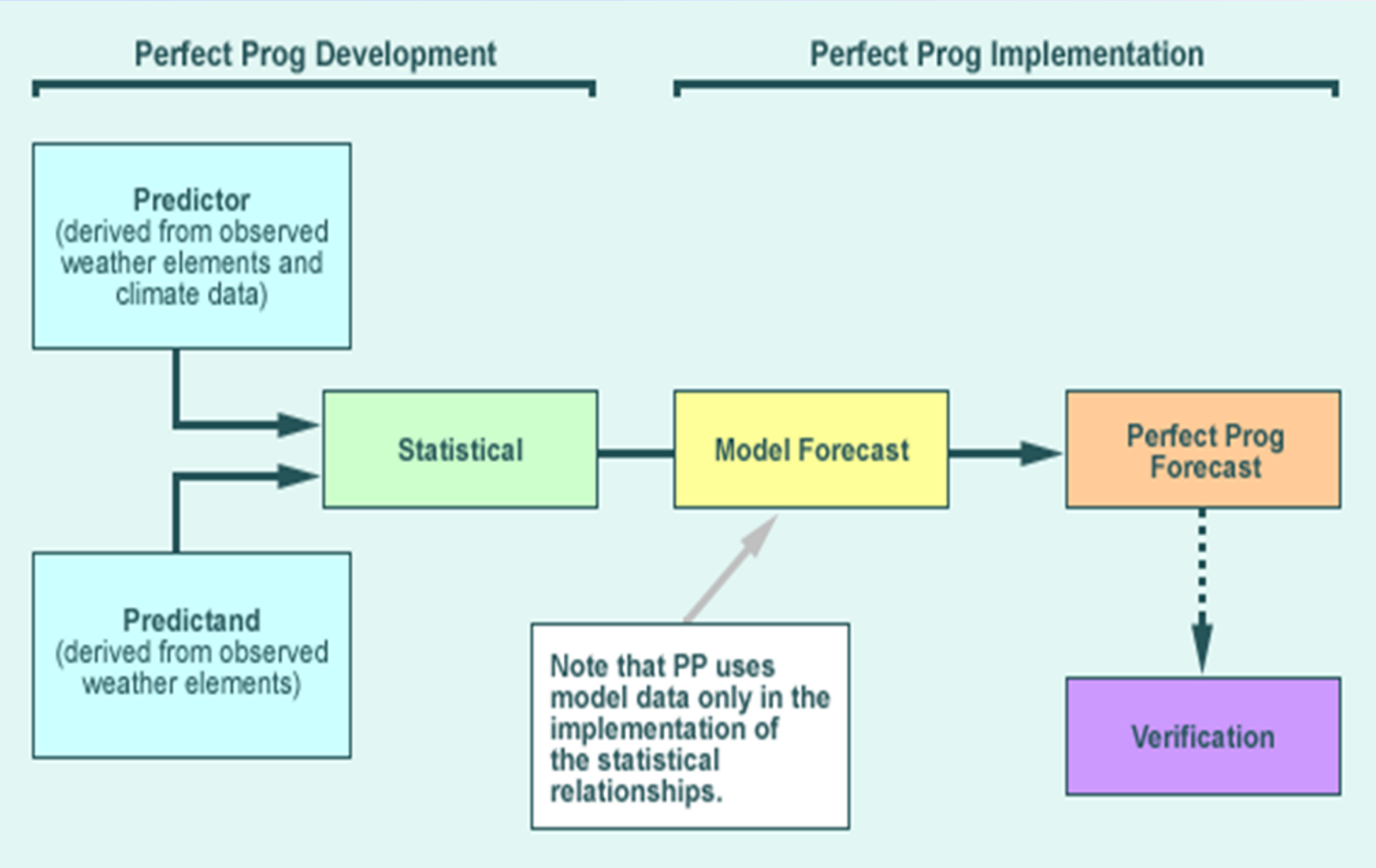
Perfect Prog (PP) & MOS statistical Approaches

- While both **MOS** and the perfect prog **(PP)** utilize similar statistical methods and mathematical procedures.

- The perfect prog (PP) statistical technique develops equations based on the relationship of co-existing observed weather elements (including climate data), which are then applied to raw model output to produce statistical guidance.

- The MOS technique develops relationship equations from both observed and model forecast weather elements, which are then applied to raw model output to produce statistical guidance.

Perfect Prog



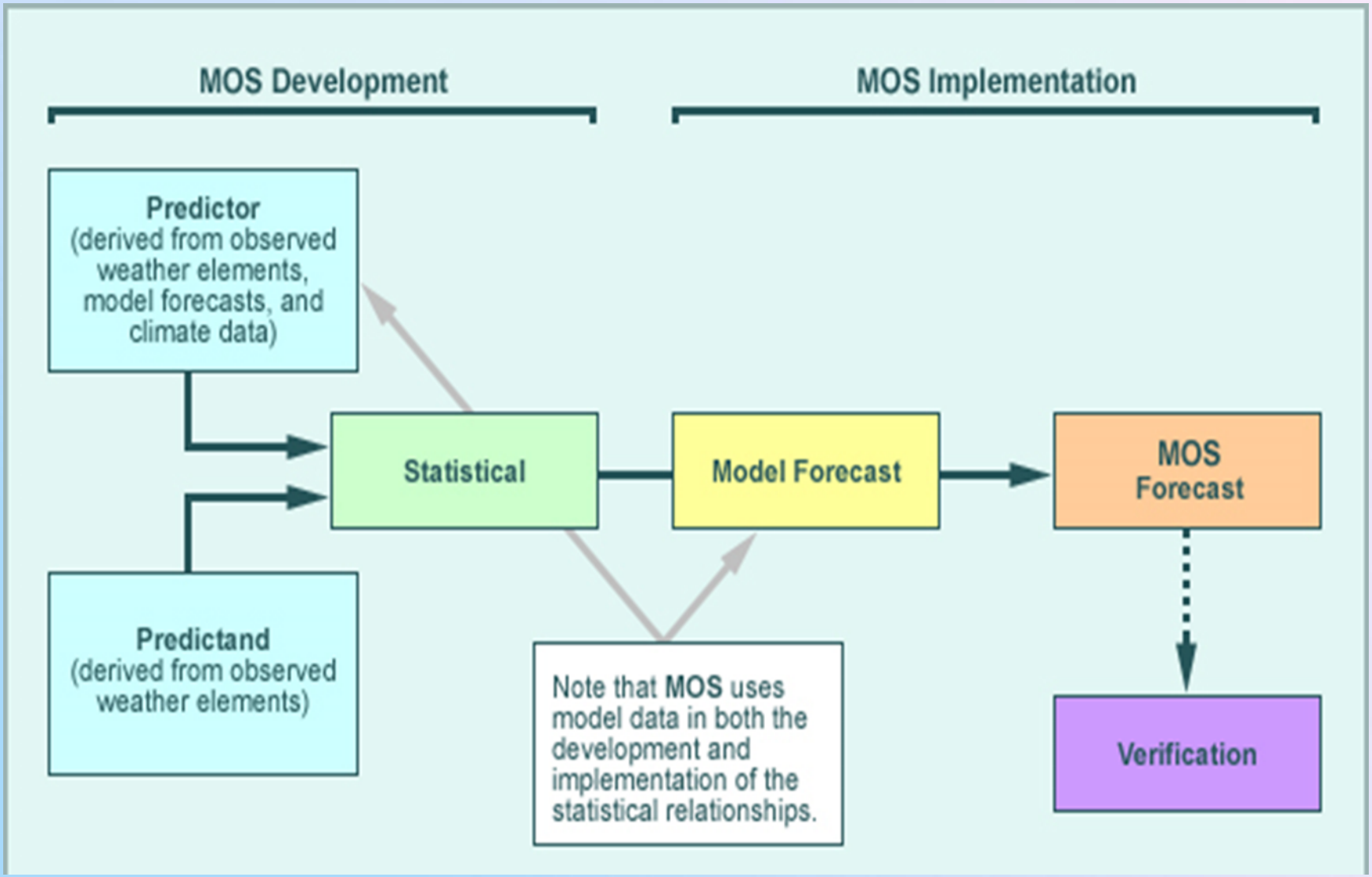
Advantages of PP Approach

- PP does not require dataset of historical nwp model data.
- PP guidance is likely to improve with improvements to the raw model forecasts.

Limitations of PP Approach

- Because historical model data are not used in the development of the statistical equations, systematic model errors cannot be accounted for in PP forecasts.

MOS


MOS Development
MOS Implementation
Predictor
(derived from observed weather elements, model forecasts, and climate data)
Predictand
(derived from observed weather elements)
Statistical
Model Forecast
MOS Forecast
Verification
Note that MOS uses model data in both the development and implementation of the statistical relationships.

Advantages of MOS

❖Because historical model data is used in development , MOS can account for systematic model errors and deteriorating model accuracy at greater forecast projections.

Limitations of MOS

❖MOS requires a developmental dataset of historical model data that is used as the predictor.

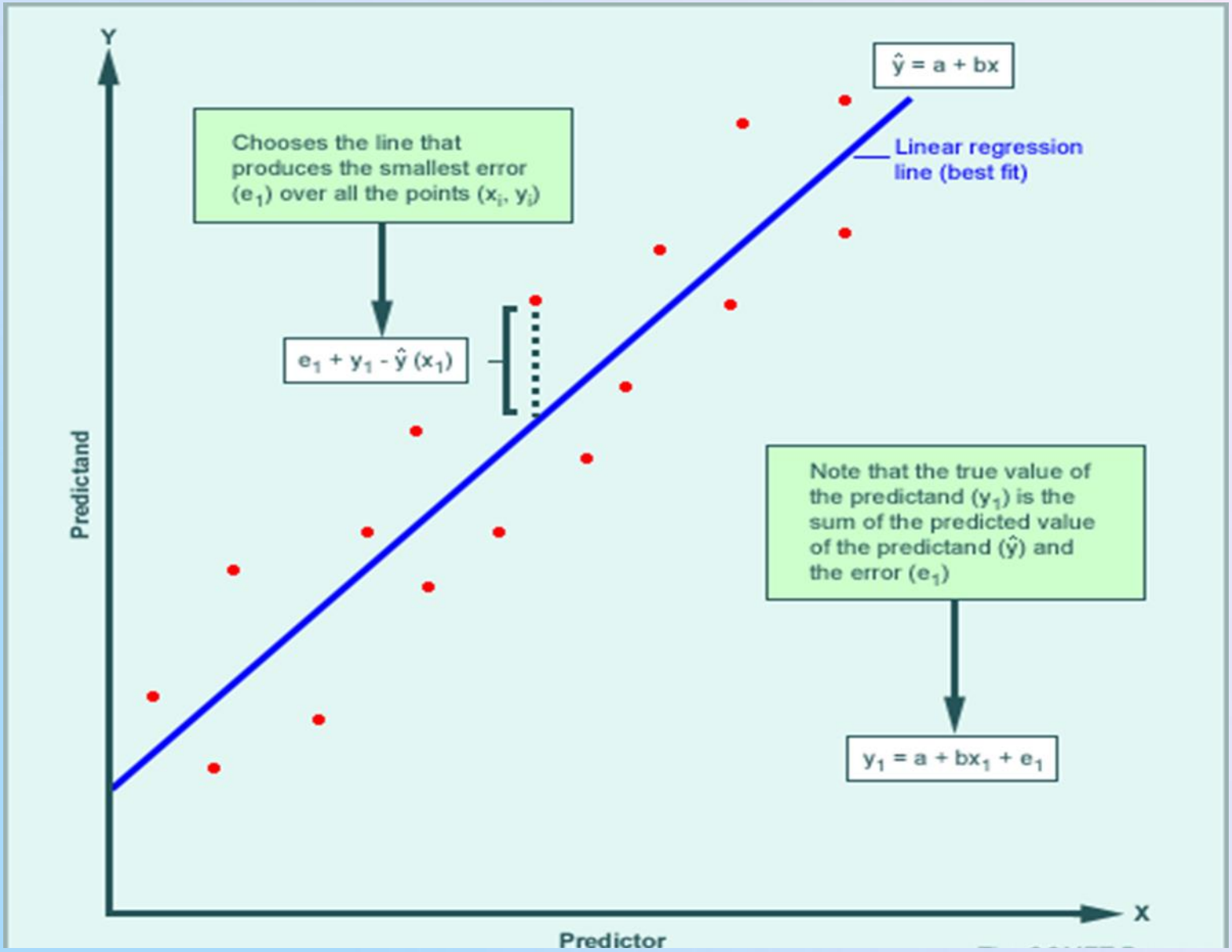
❖MOS can be inefficient, particularly in situations where the operational model suite is evolving (requires redevelopment of statistical relationships).

Statistical Techniques

- There are several statistical techniques that are used to develop the relationships defined in both the PP and MOS approaches to statistical post-processing of NWP output.

- The most common method is that of multiple linear regression

Linear Regression



Multiple Linear Regression

The **multiple linear regression process is iterative**. Individual predictors are selected and used with the predictand (T_{max}) to describe a linear relationship.

Potential predictors
1. 850 - 1000 hPa Δz
2. 500 - 1000 hPa Δz
3. 850 - 1000 hPa RH

I. Determination of regression equations
$\hat{T}_{max} = b_0 + b_1 x_1 + e$
(Click here for linear relationship)

II. Result is the selection of the best predictor
1. 850 - 1000 hPa Δz
2. 500 - 1000 hPa Δz
3. 850 - 1000 hPa RH

Predictand
1. T_{max}

III. Remaining predictors are examined with respect to the best predictor
$\hat{T}_{max} = b_0 + b_1 \Delta z_{850\text{-}1000} + b_2 RH_{850\text{-}1000} + e$
(Click here for 3-dimensional relationship)

Other Statistical Methods

There are several other techniques that may be used instead of multiple linear regression to generate statistical relationships between predictor and predictand.

These include

- Kalman Filters
- Neural Networks
- Artificial Intelligence

Bias correction Techniques for NWP Model outputs

Bias Correction methods

- In the bias correction, we estimate the systematic component of the NWP model forecast errors, based on the past forecast errors.
- The Model Output Statistics (MOS) approach is used to improve the NWP model output through bias correction.
- Drawback of MOS is that it requires a long training period of archived model data from an unchanged or static model.
- **Today, NWP modeling centers make frequent changes to numerical procedures, physics, and resolution of models.**
- To overcome this ever-changing NWP model base, we need to use a bias correction technique that update bias daily and depends on only the most recent past data for bias correction.

Bias Correction Methods

1. Linear Regression (LR)

$$Y = aX + b$$

The lagged Linear Regression (LR) method has been used in the past (e.g. Stensrud and Yussouf 2005), and it uses a least-squares line to model the trend in the bias of the forecasts over the training period at each grid.

where Y =Observation; X=Forecast,

a=measure of multiplicative bias
b=systematic bias

Computed from last 90 days forecasts and observations

Bias Correction Methods

2. Best *Easy systematic (BES)* mean

$$BIAS = (Q_1 + 2Q_2 + Q_3) / 4$$

Woodcock and Engel (2005) evaluated the usefulness of the best easy systematic mean statistics (BES) bias correction methodology for the bias correction of 2-m maximum and minimum temperature forecasts over Australia

- **Where Q1, Q2, and Q3 are the first, second, and third quartiles of forecast errors (F-O)**
- **Computed from daily forecast error of the previous 90 days.**
- **Not sensitive to extreme values.**

Bias Correction Methods

3. Running Mean Error (RME)

$$ME = \frac{1}{N} \sum_{k=1}^{N} [F - O]$$

$$where \ldots N = 1,2,\ldots 15.days$$

Stensrud and Skindlov (1996) showed that a simple bias correction method using the previous 7-day Running Mean bias correction improved the direct model forecasts of maximum temperature

Steed and Mass (2004) experimented with several different spatial techniques of applying bias removal to temperature forecasts from a mesoscale model. Their study showed that a bias removal method using a 2-week running bias had the least amount of error compared to periods of 1, 3, 4 and 6 weeks

The previous 15 days forecast errors are averaged together using equal weight

Bias Correction Methods

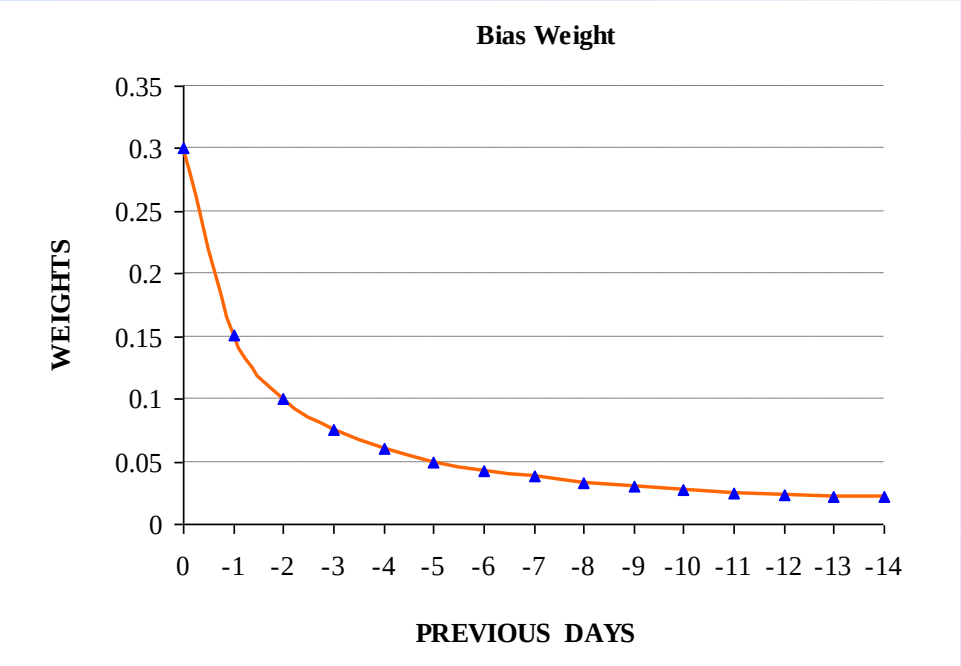
4. Nearest Neighborhood (NN)

$$wt(i) = \frac{w(i)}{\sum_{i=1}^{15} w(i)}$$

$$w(i) = \frac{1}{i+1}$$

where i=0,1, 2,3, …15 days

$$b_m = wt(i) * b(i)$$



In the Nearest Neighbor moving average;-all the previous forecast errors are averaged together using an exponentially increasing weighting so that the recent data has the largest weight.

Data

❖In this study, the day-1 to day-5 maximun and minimum temperature forecast data from four operational global models for the period from 1 January to 30 September 2011 are used

1. European Centre for Medium-Range Weather Forecasts (ECMWF), England
2. The U.S. National Centers for Environmental Prediction (NCEP)'s Global Forecasting System (GFS).
3. Japan Meteorological Agency (JMA) and
4. Indian Meteorological Department (IMD)'s Global Forecasting System (GFS T382)

❖The daily maximum and minimum temperature observation through Global Telecommunication system (GTS) from IMD over land is used for the study

Verification of bias Corrected Forecast

Mean error

$$ME = \frac{1}{N}\sum_{i=1}^{N}(F_i - O_i)$$

Mean absolute error

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|F_i - O_i\right|$$

Root mean square error

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(F_i - O_i)^2}$$

Improvement Skill

$$SKILL\ (\%) = \frac{(MAE_{DMO} - MAE_{BIASCOR})}{MAE_{DMO}} \times 100$$

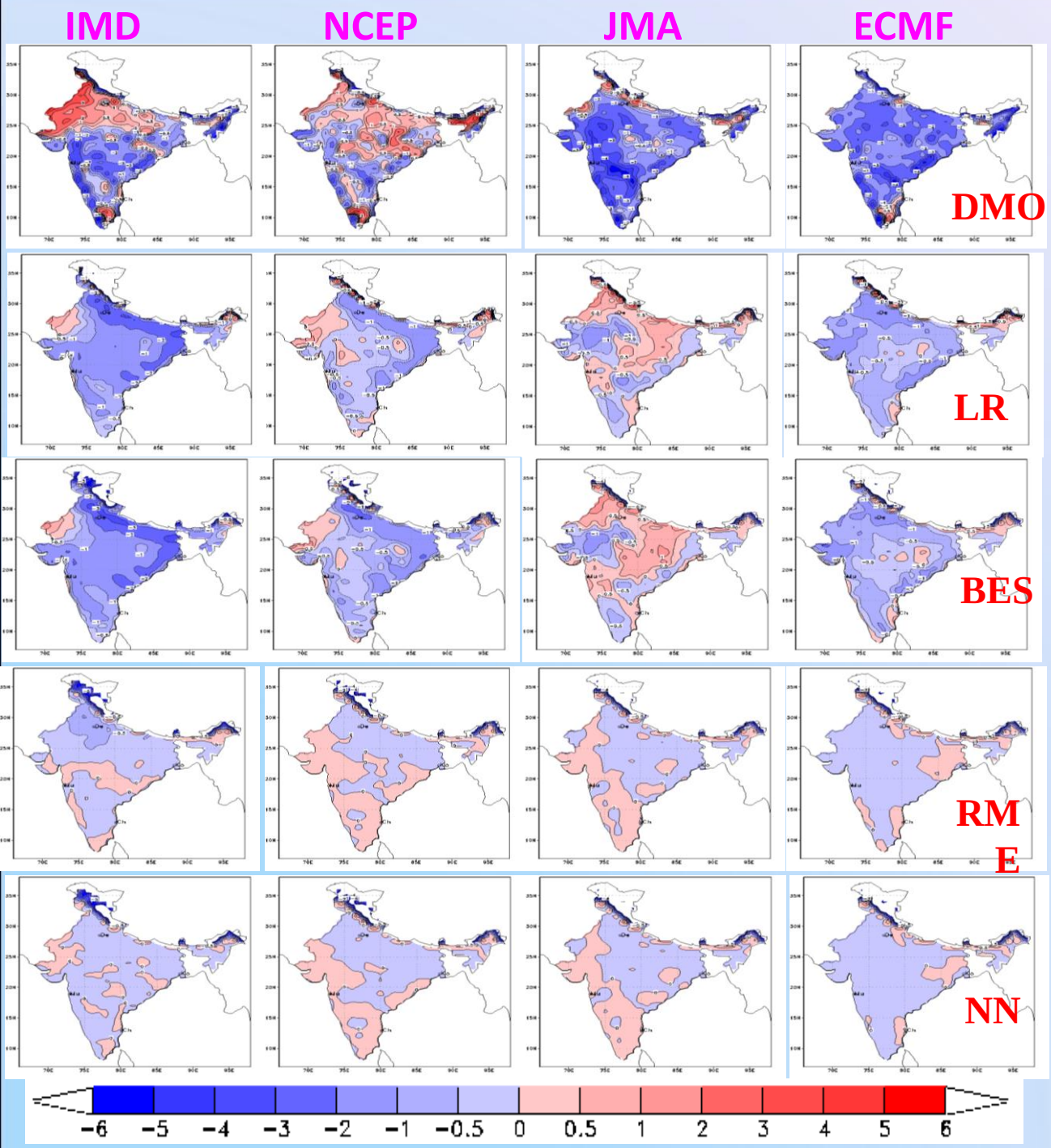
Result & Discussions

Spatial distribution of Error pattern (Tmax/Tmin) during 1 July-30 September 2011

a) Mean Error
b) Mean Absolute Error
c) Root mean square error

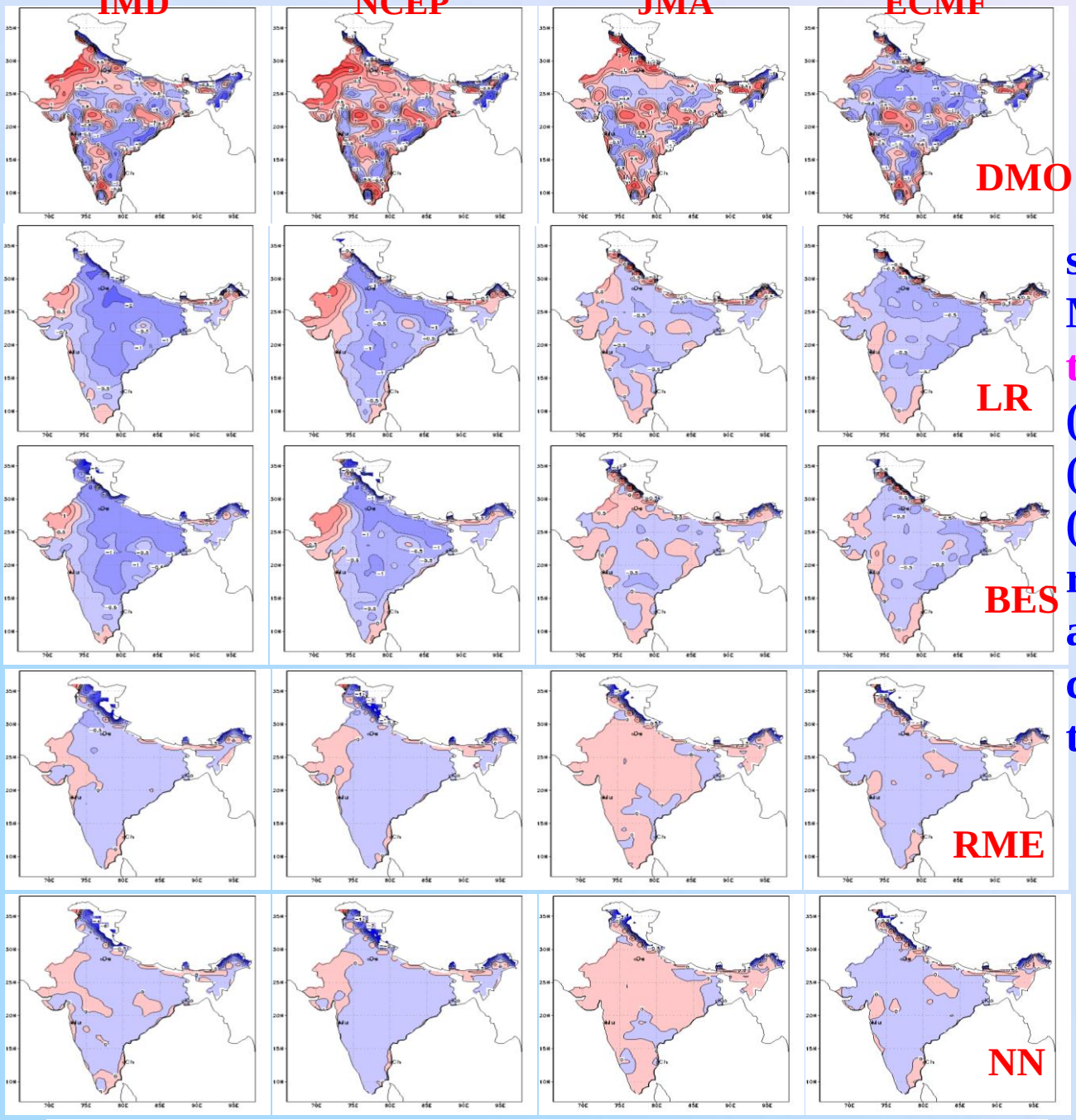
Forecast Skill over

a) NW India
b) East India
c) Central India



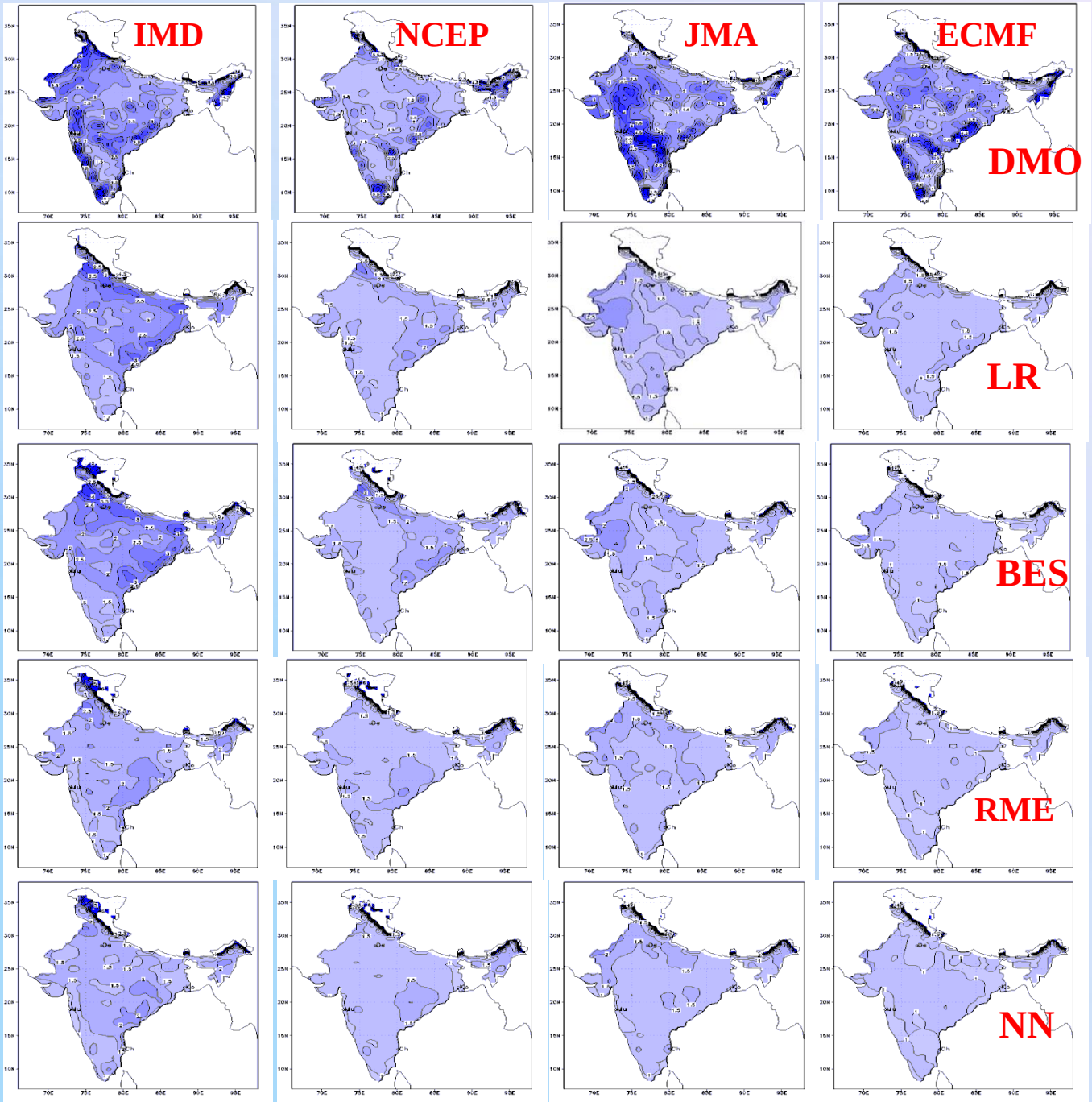
ME in Maximum temperature (C)

Spatial distribution of ME in day-1 maximum temperature forecast in deg C from DMO (first row) and LR (second row), BES (third row), RME (fourth row) and NN (last row) bias correction methods of all the NWP models.



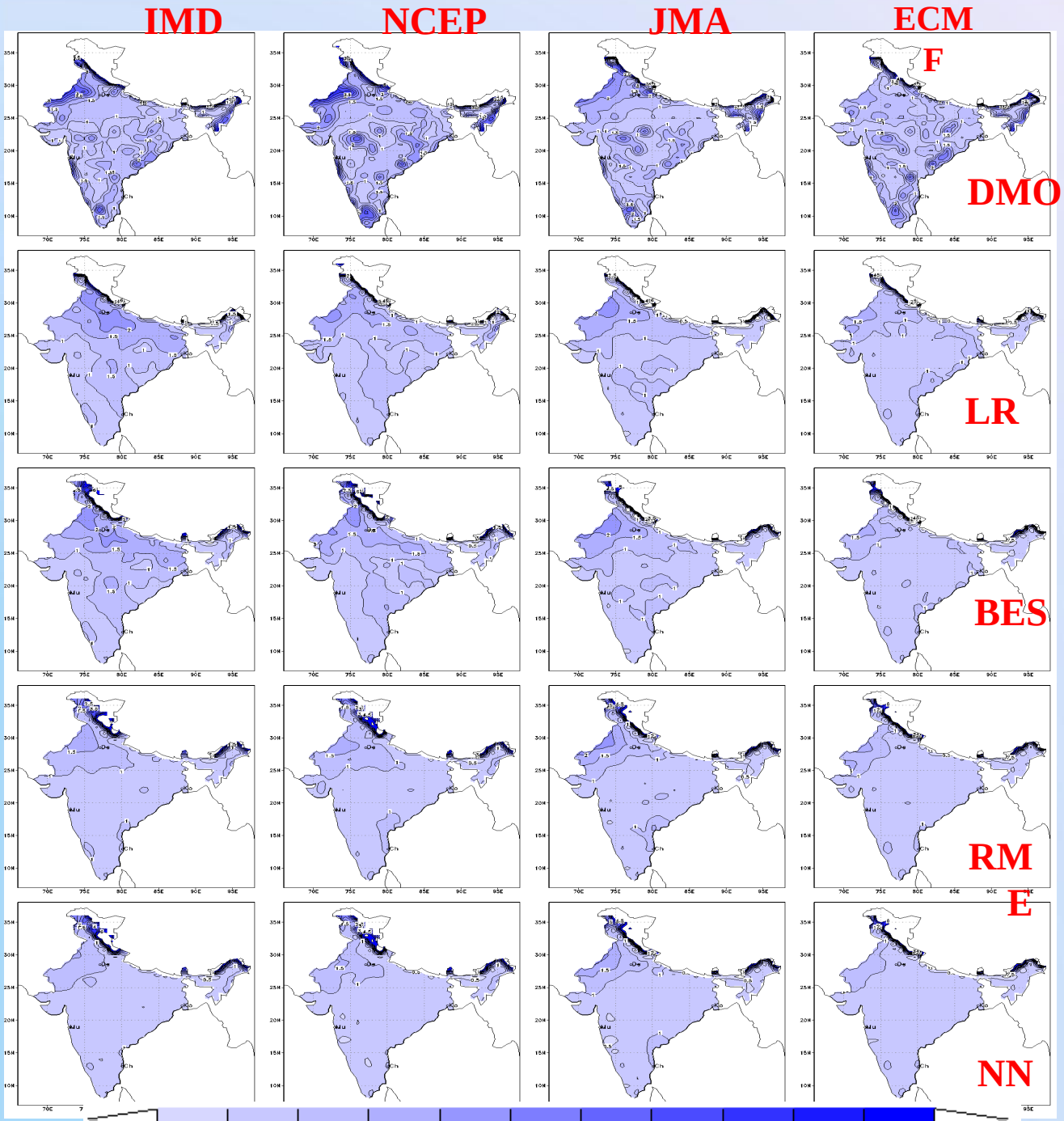
ME in minimum temperature (C)

spatial distribution of ME in day-3 minimum temperature forecast (deg C) from DMO (first row) and LR (second row), BES (third row), RM (fourth row) and NN (last row) bias correction methods of all the NWP models.



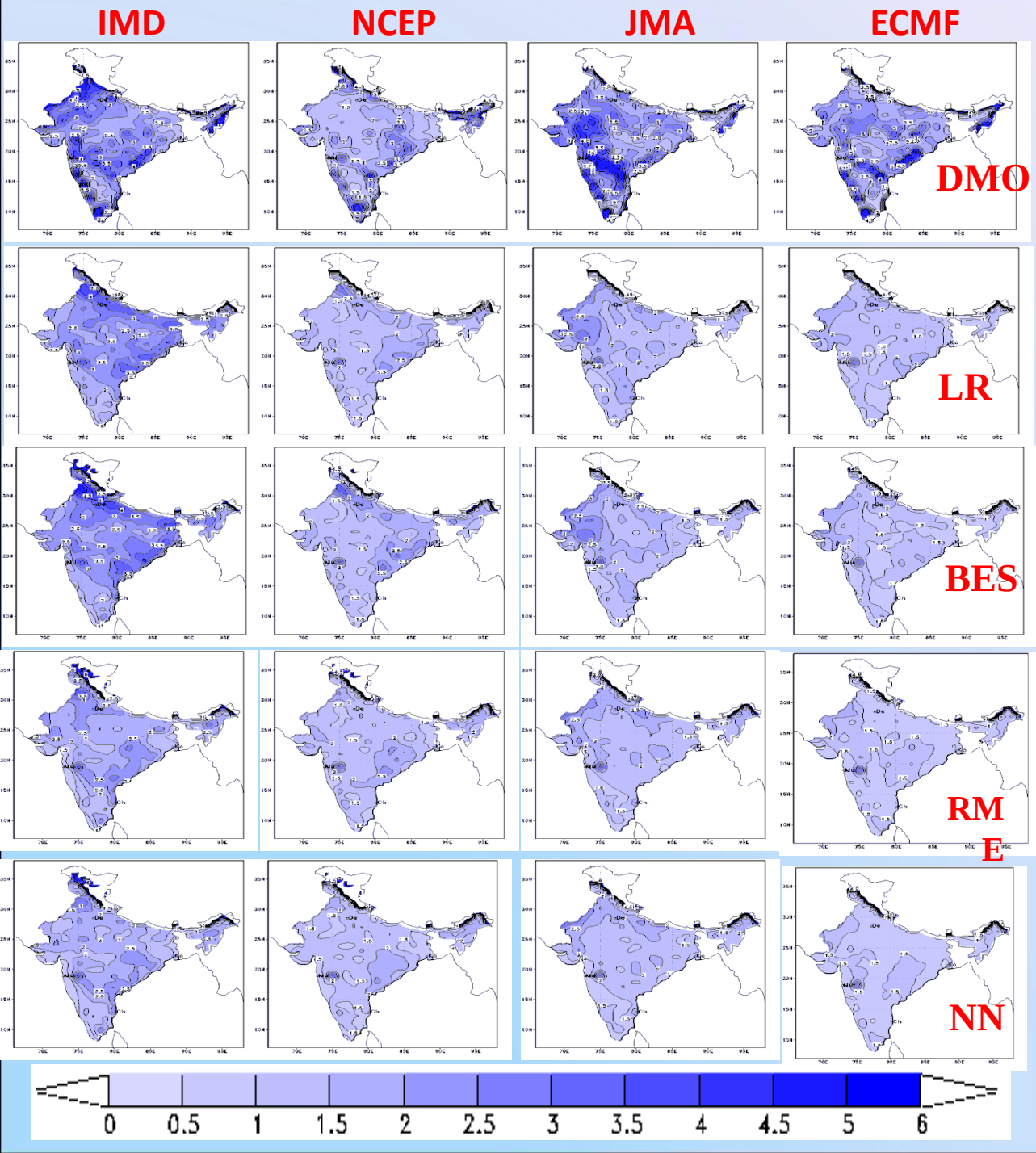
MAE in Maximum temperature (C)

Spatial distribution of MAE in day-1 maximum temperature forecast in deg C from DMO (first row) and LR (second row), BES (third row), RM (fourth row) and NN (last row) bias correction methods of all the NWP models.



MAE in minimum temperature (C)

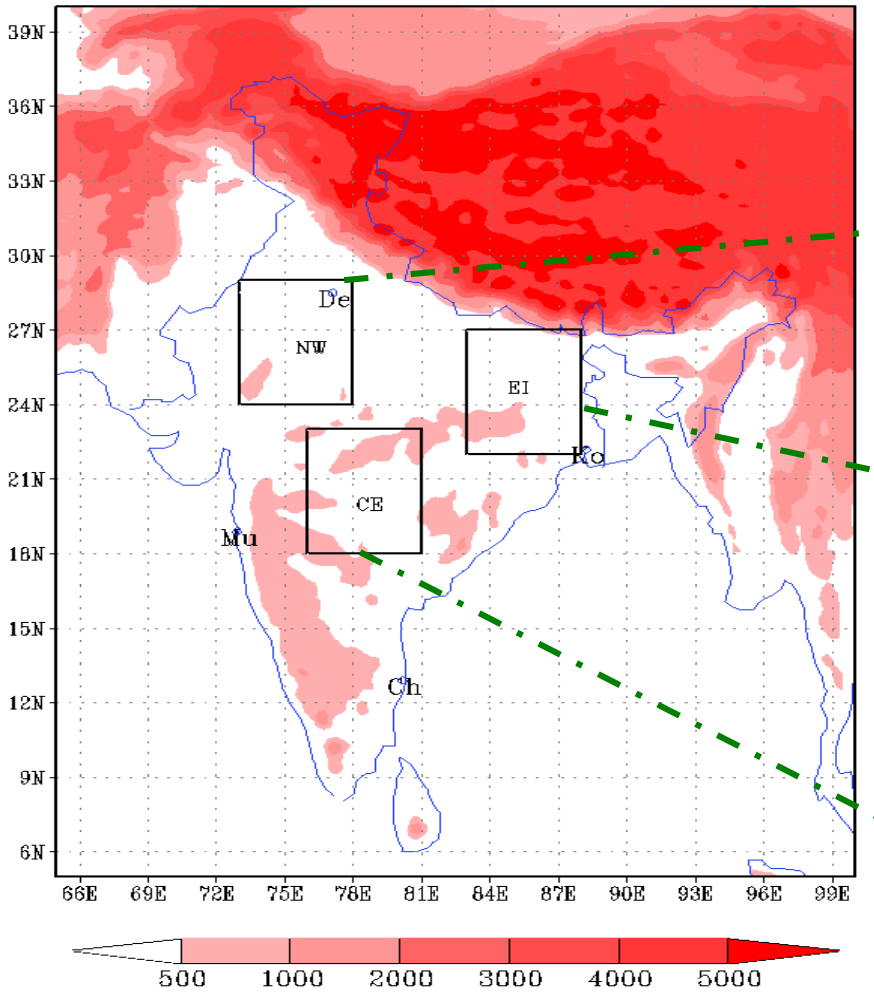
spatial distribution of MAE in day-3 minimum temperature (C) forecast from DMO (first row) and LR (second row), BES (third row), RM (fourth row) and NN (last row) bias correction methods of all the NWP models.



RMSE in Maximum temperature (C)

spatial distribution of RMSE in day-1 maximum temperature forecast in deg C from DMO (first row) and LR (second row), BES (third row), RME (fourth row) and NN (last row) bias correction methods of all the NWP models.

Topographic Map of India (altitude in meter)



Study Area

North West India:

Long: 73 -78E ; Lat :24-29 N

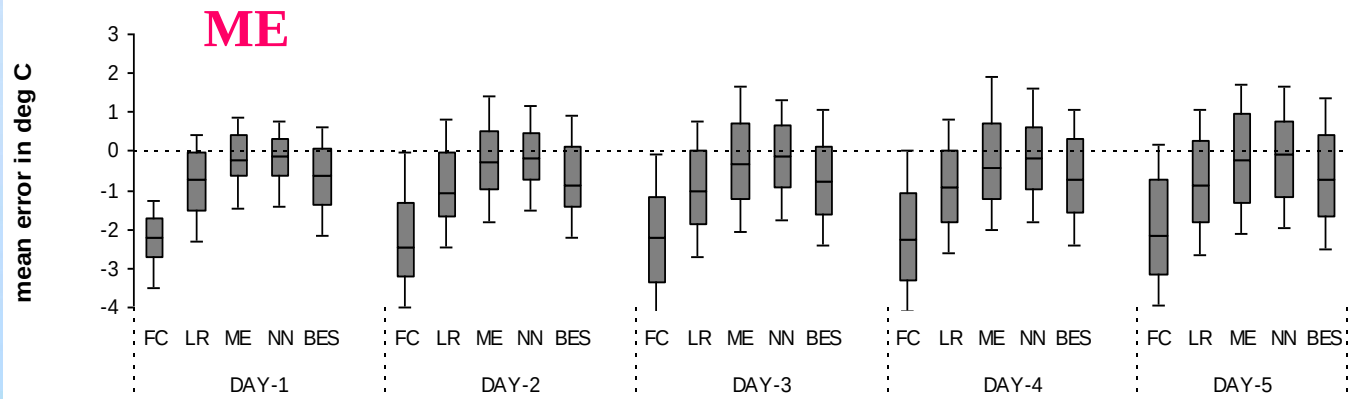
East India:

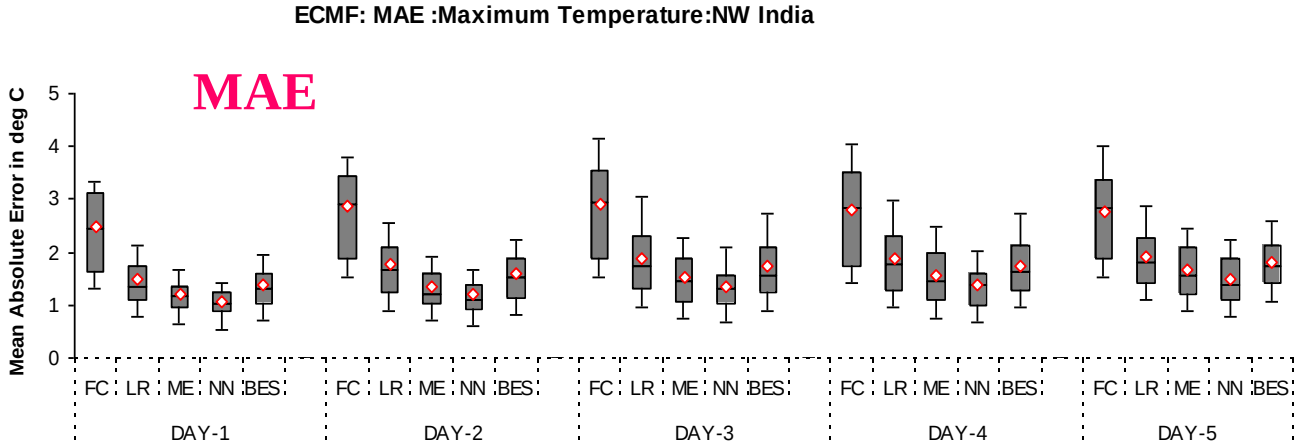
Long: 83 -88E ; Lat :22-27 N

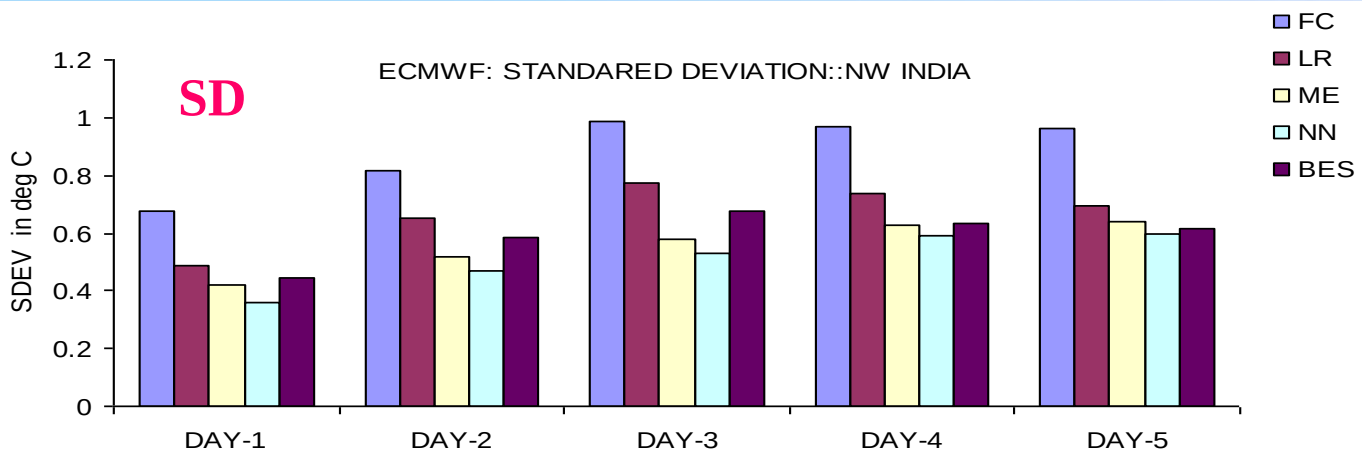
Central India:

Long: 77 -82E ; Lat :18-23 N

Maximum Temperature: NW India



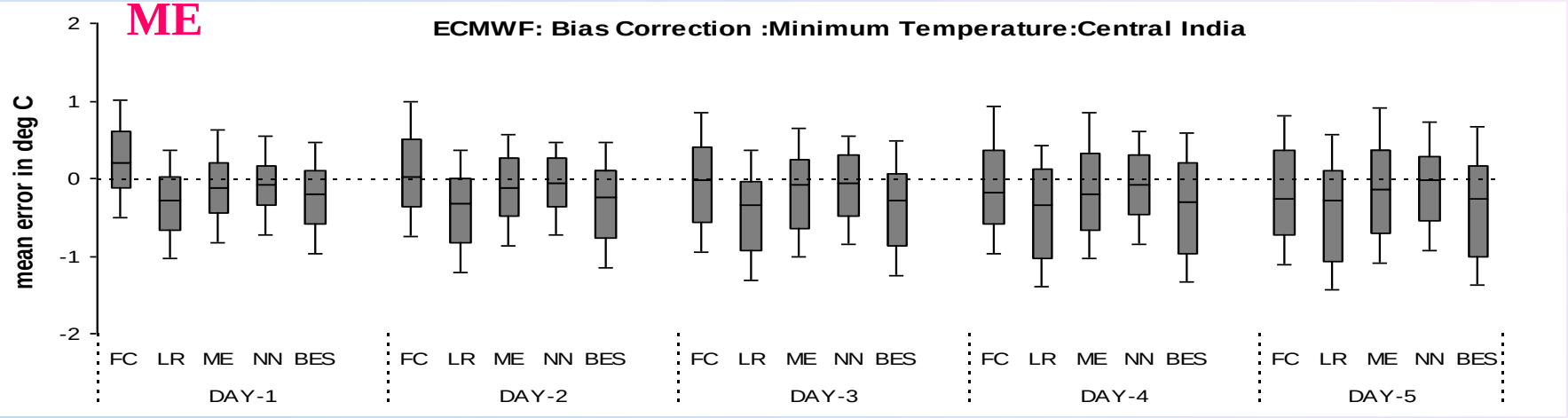


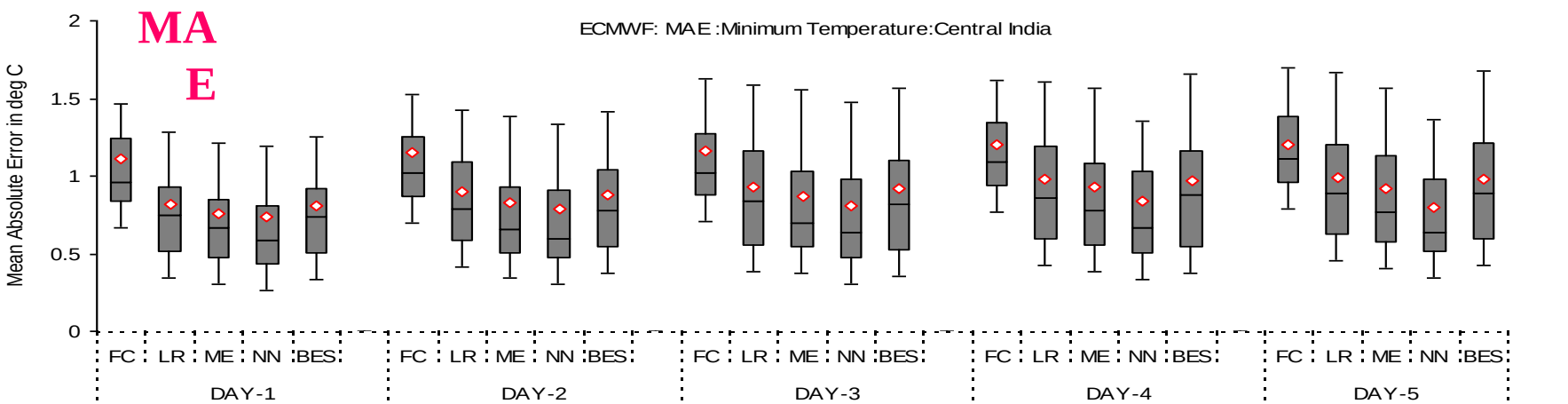


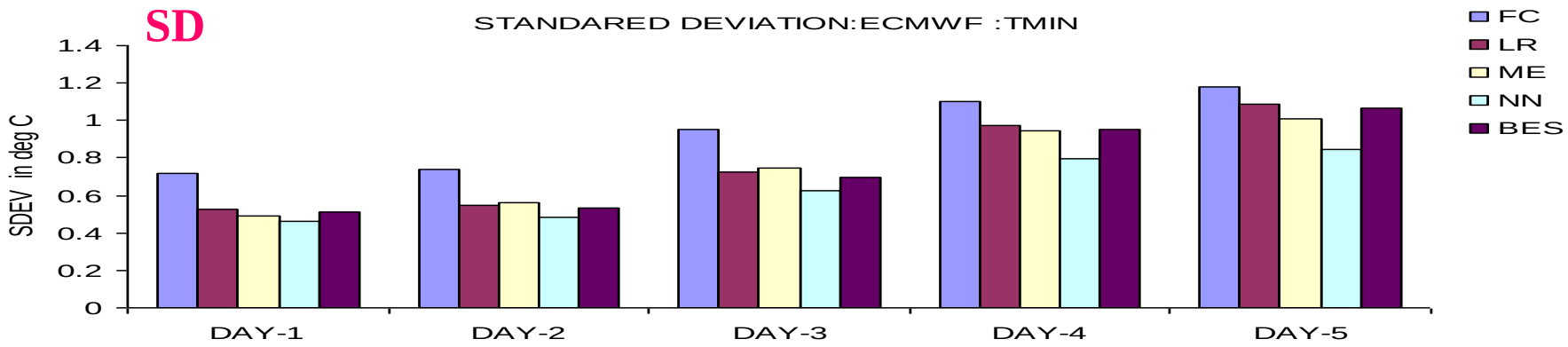
Box-and-whiskers plots of ECMWF model mean errors (top panel) and MAE (middle panel) and standard deviation (bottom panel) from DMO and bias corrected forecast.

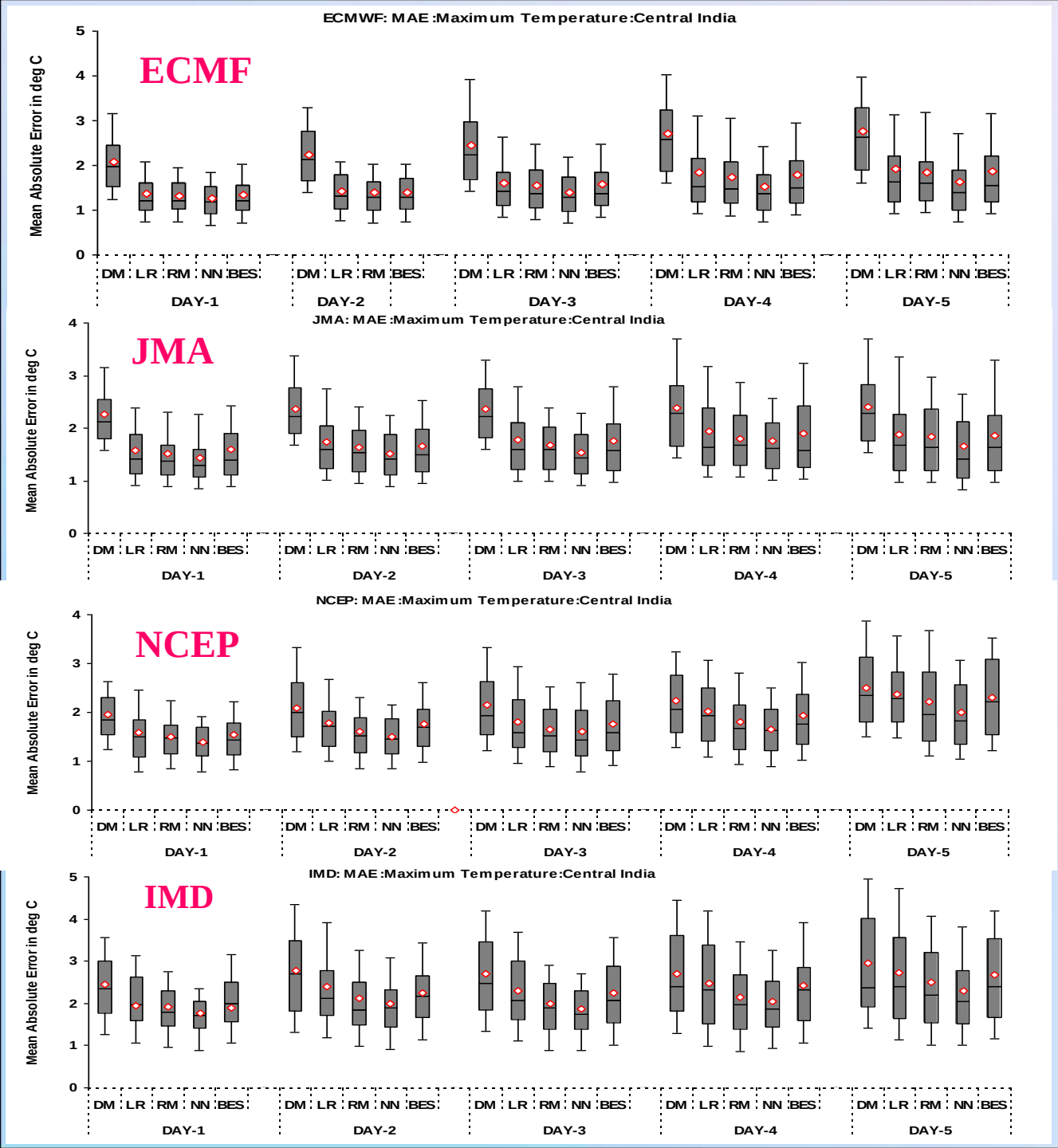
- **The boxes show the median and 25th and 75th percentiles, while the whiskers indicate below the 25th percentile and above the 75th percentile**
- **Mean, Median and SDEV is less for NN**

Minimum Temperature: Central India



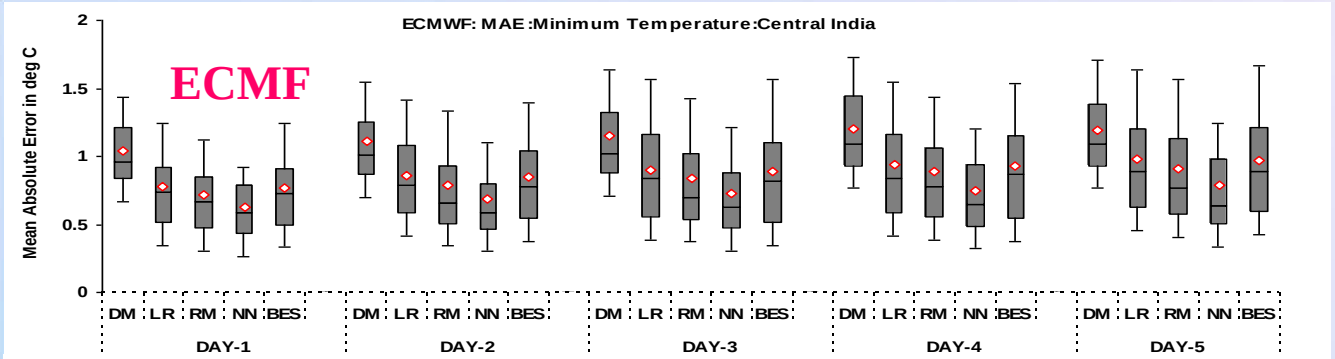


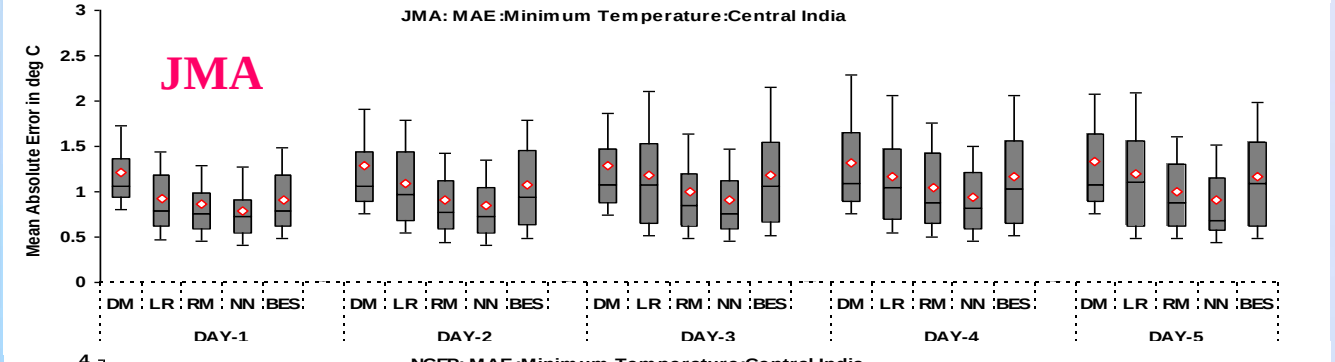


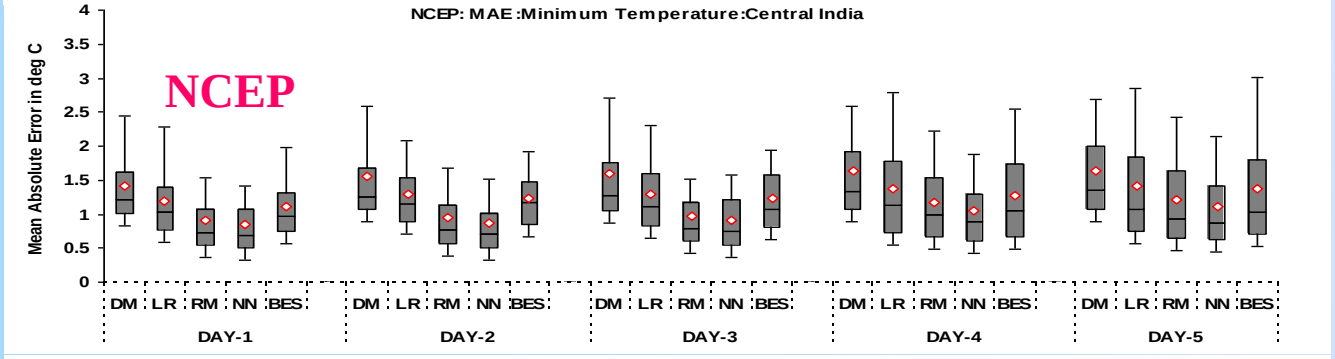


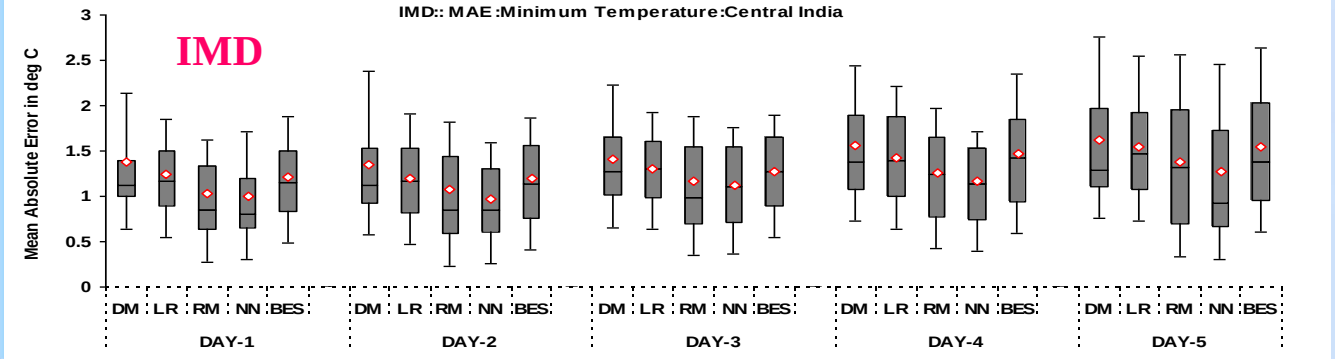
Maximum temperature

- Box-and-whiskers plots for ECMWF, JMA, NCEP and IMD model MAE (C) in maximum Temperature over central India
- The boxes show the median and 25th and 75th percentiles, while the whiskers indicate below the 25th percentile and above the 75th percentile
- Mean, Median and SDEV are less for NN BC method





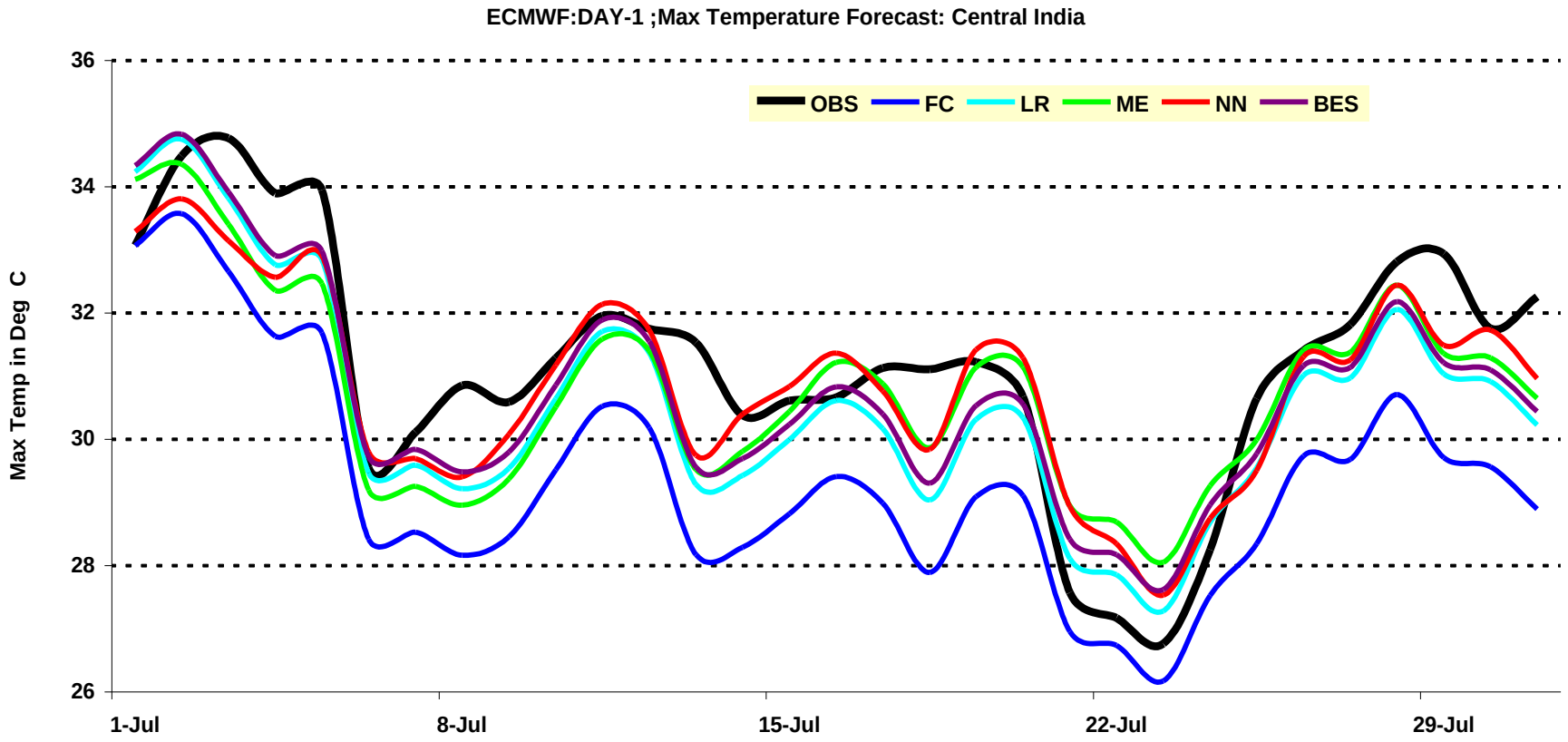




Minimum temperature

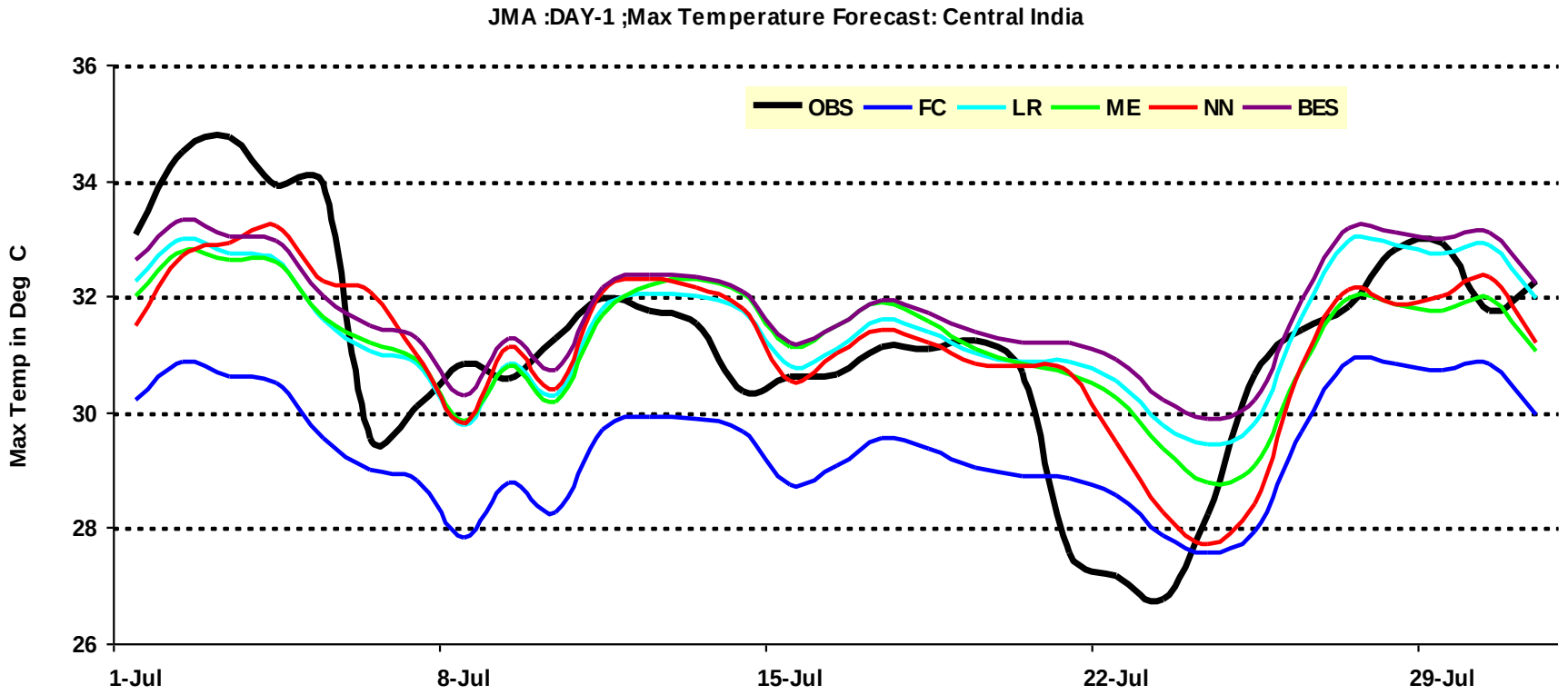
- Box-and-whiskers plots for ECMWF, JMA, NCEP and IMD model MAE (C) in minimum Temperature over central India.
- Mean and Median are less for NN BC method in all NWP models

24 hour maximum temperature Forecast over central India :ECMF



- The 24 hour maximum temperature forecasts over Central India using DMO and Bias Corrected ECMF during July 2011.
- The NN bias corrected forecast is reasonably very close to the observed temperature during the period.

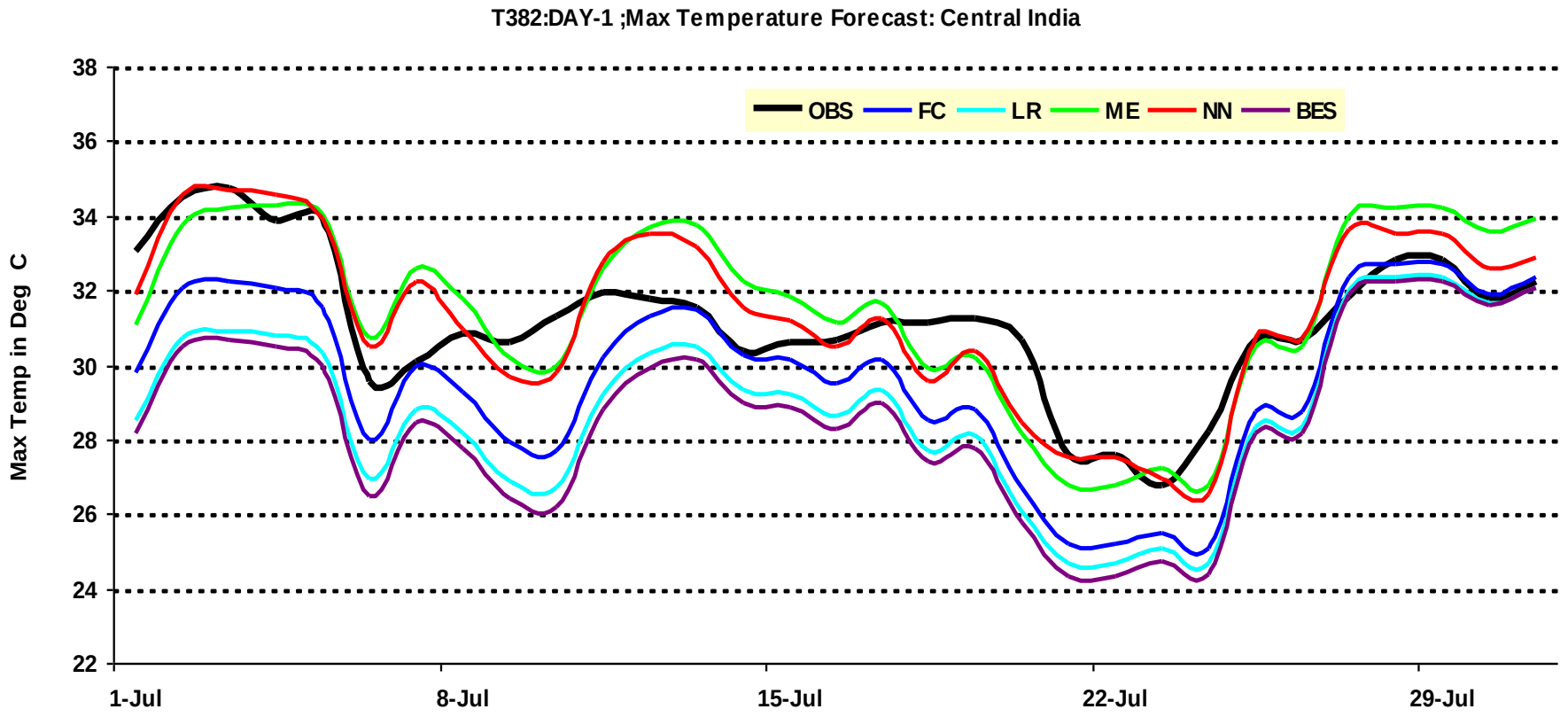
24 hour maximum temperature Forecast over central India :JMA



❖The 24 hour maximum temperature forecasts over Central India using DMO and Bias Corrected JMA during July 2011.

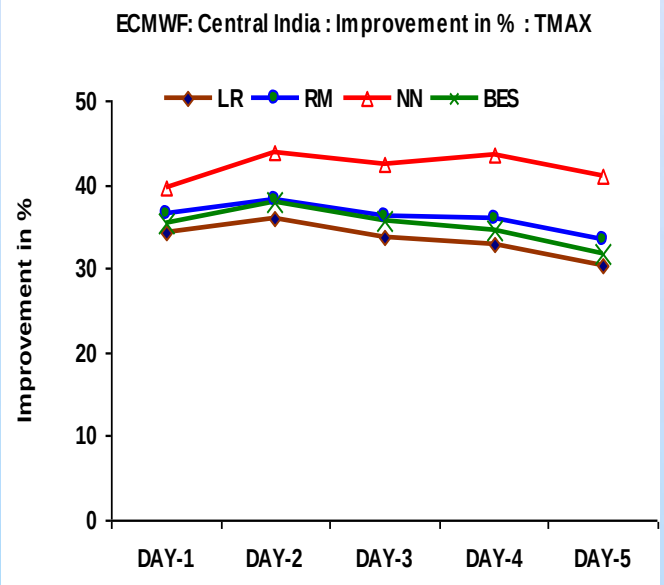
❖The NN bias corrected forecast is reasonably very close to the observed Max temperature during the period.

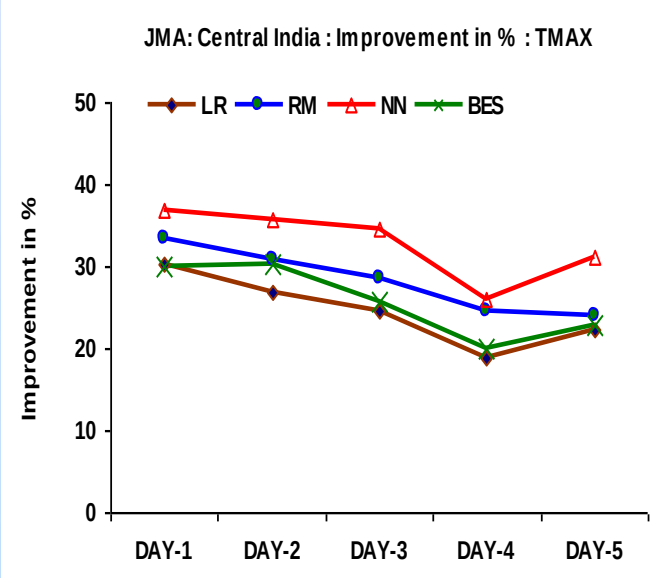
24 hour maximum temperature Forecast over central India: IMD GFS

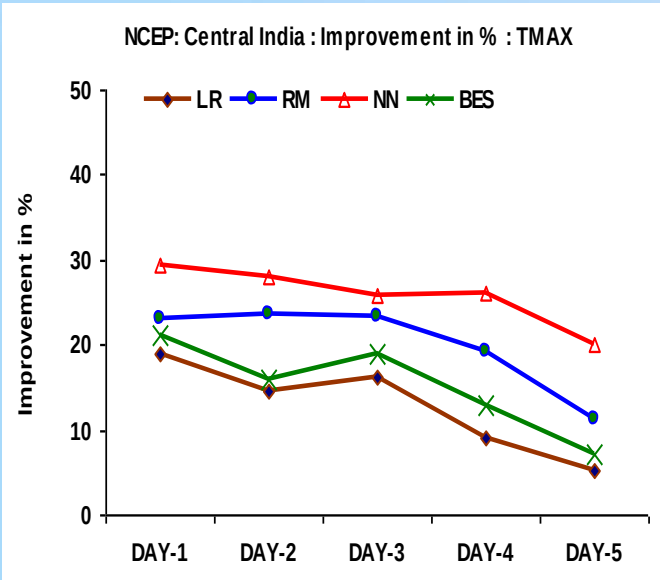


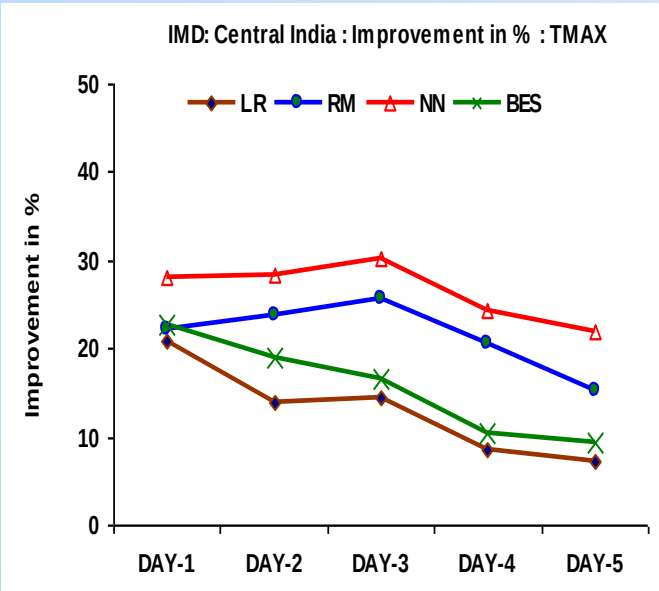
- The 24 hour maximum temperature forecasts over Central India using DMO and Bias Corrected IMD GFS during July 2011.
- The NN bias corrected forecast is reasonably very close to the observed Max temperature during the period.

Maximum Temperature Forecast improvement in % (MAE) : Central India





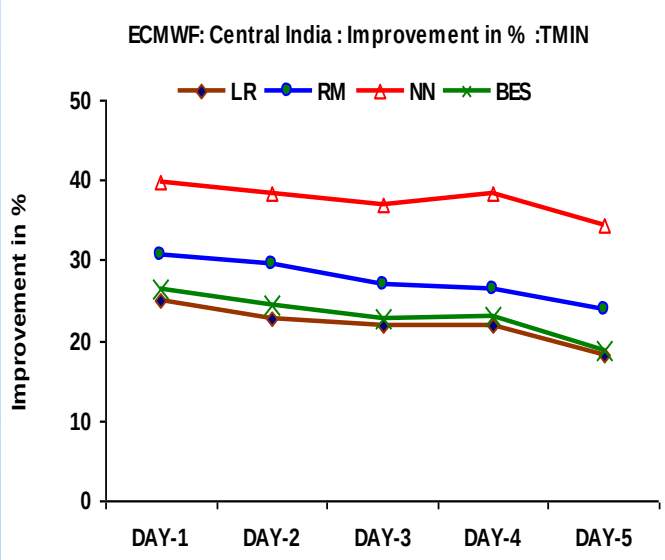


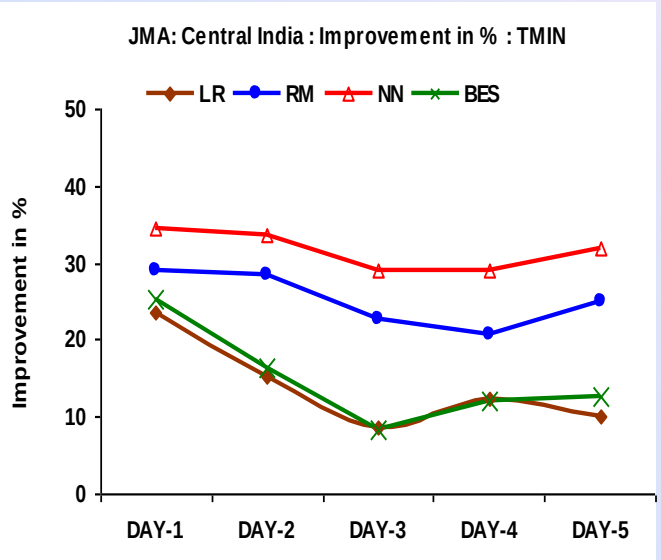


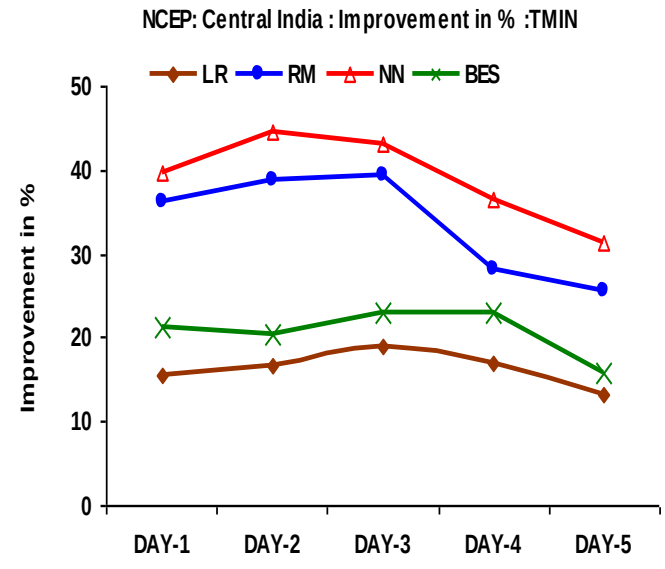
All the four bias correction methods have significant improvement in forecasting maximum temperature from 25 to 45% over the DMO.

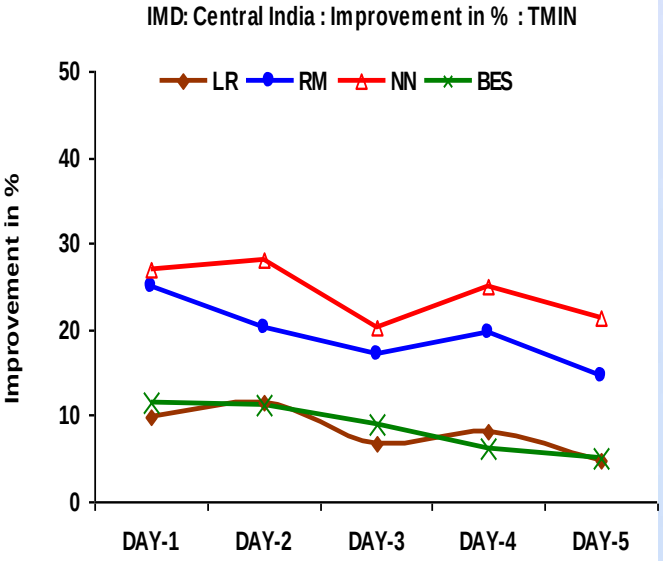
There is a significant difference in improvement between bias correction methods but NN method performs best

Minimum Temperature Forecast improvement in % (MAE) : Central India









All the four bias correction methods have significant improvement in forecasting minimum temperature from 10 to 40% over the DMO forecasts.

On the whole, the NN BC method shows higher skill (30-40 %) compared to other three bias correction methods for minimum temperature forecasts

Summary

- All the four bias correction methods for maximum and minimum temperatures forecasts have smaller ME, MAE and RMSE values over Indian regions for all day-1 to day-5 as compared to DMO.

- More improvement is noticed in the NN bias correction method as compared to the other three BC methods.

MODEL OUTPUT STATISTICS(MOS): Part-II

Location Specific Forecasting of Maximum and Minimum Temperature using *NN* Bias corrected output of GFS T574

IMD GFS WEATHER FORECAST FOR MAJOR CITIES

ISSUED ON: 01-09-2014 ;VALID FOR THE NEXT 7 DAYS

CITY NAME: NEW-DELHI

Date	Temperature (Deg. C)		Humidity (%)		Total Cloud (Okta)	Rainfall (mm)
	Maximum	Minimum	Maximum	Minimum		
02/09/2014	34	25	79	48	5	0
03/09/2014	35	25	74	51	6	0
04/09/2014	36	26	85	50	8	0
05/09/2014	32	24	93	69	7	46
06/09/2014	32	24	91	65	6	4
07/09/2014	32	25	92	60	5	20
08/09/2014	34	25	83	55	2	0

Objective

The main objective of this study is to verify the skill of Direct model output (DMO) and Bias Corrected (BC) GFS T574L64 model forecast for location specific forecast of maximum and minimum temperature over India

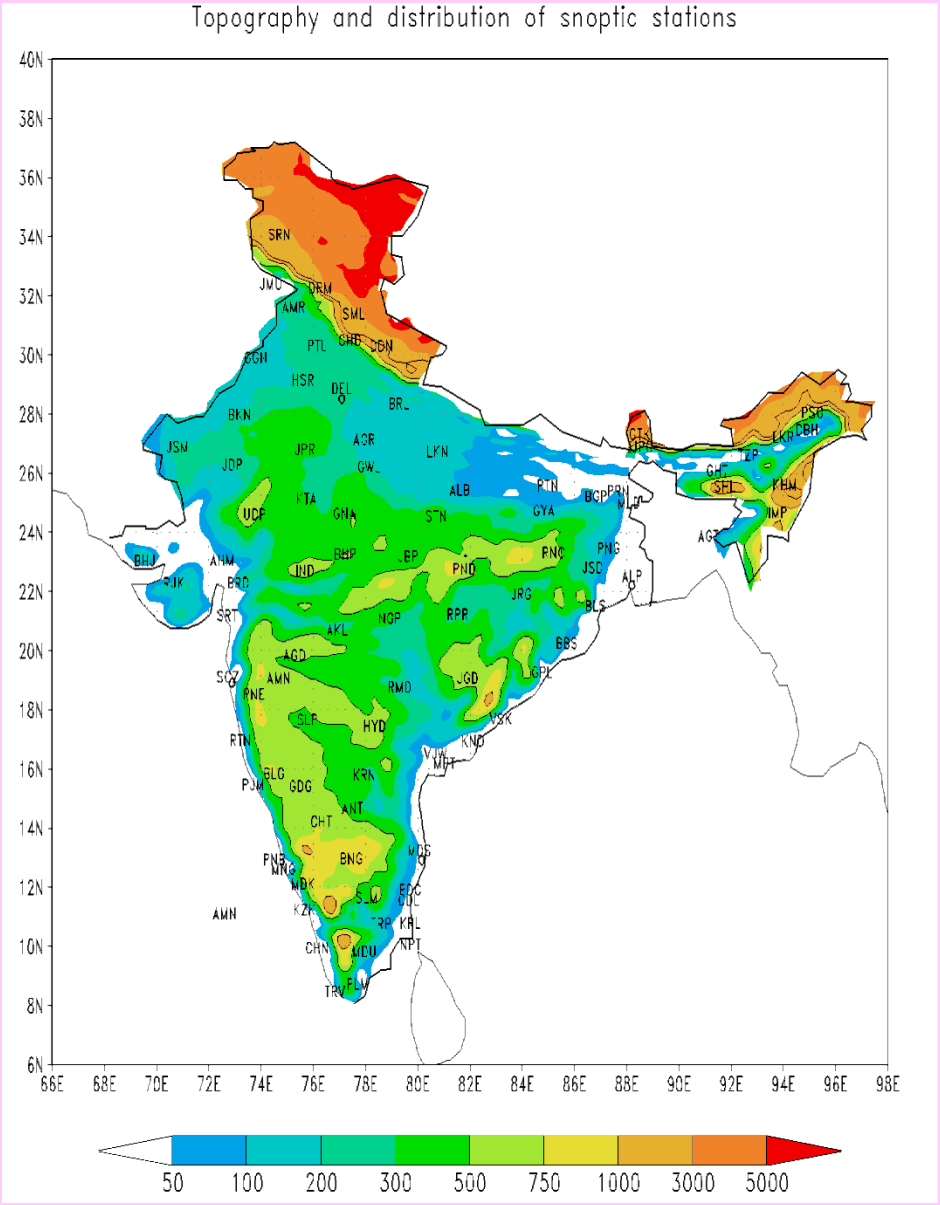
Data Source

❖**The output from Global Forecasting System (GFS) T574L64 operational at India Meteorological Department (IMD), New Delhi is used for obtaining location specific quantitative forecast of maximum and minimum temperatures for 100 synoptic stations, representing different geographical regions of India in the medium range (24 to 120 hr) time scale during may 2012 –Feb 2013.**

North West India (NW)	CODE	East and North East India (ENE)	CODE	Central India (CI)	CODE	Southern Peninsular India (SP)	CODE
SRINAGAR	SRN	PASIGHAT	PSG	GWALIOR	GWL	RAMAGUNDAM	RMD
JAMMU	JMU	GANGTOK	GTK	GUNA	GNA	HYDERABAD	HYD
DHARMSALA	DRM	N-LAKHIMPUR	LKR	SATNA	STN	VISHAKHAPATNAM	VSK
AMRITSAR	AMR	MOHANBARI	DBH	BHUJ	BHJ	VIJAYAWADA	VJW
SHIMLA	SML	JALPAIGURI	JPG	AHMADABAD	AHM	MACHILIPATNAM	MPT
PATIALA	PTL	GAUHATI	GHT	BHOPAL	BHP	KAKINADA	KND
AMBALA	AMB	TEZPUR	TZP	JABALPUR	JBP	BELGAUM	BLG
CHANDIGARH	CHD	PATNA	PTN	RAJKOT	RJK	GADAG	GDG
DEHRADUN	DDN	BHAGALPUR	BGP	BARODA	BRD	KURNOOL	KRN
GANGANAGAR	GGN	PURNEA	PRN	INDORE	IND	CHITRADURGA	CHT
HISSAR	HSR	MALDA	MLD	PENDRA	PND	ANANTAPUR	ANT
BIKANER	BKN	SHILLONG	SHL	SURAT	SRT	MADRAS	MDS
DELHI	SFD	KOHIMA	KHM	NAGPUR	NGP	MANGALORE	MNG
BAREILLY	BRL	GAYA	GYA	RAIPUR	RPR	PANAMBUR	PNB
AGRA	AGR	IMPHAL	IMP	JHARSUGUDA	JRG	MADIKERI	MDK
JAISALMER	JSM	RANCHI	RNC	BALASORE	BLS	BANGALORE	BNG
JODHPUR	JDP	PANAGARH	PNG	AKOLA WEST	AKL	AMINI	AMN
JAIPUR	JPR	AGARTALA	AGT	BHUBANESWAR	BBS	KOZHIKODE	KZK
LUCKNOW	LKN	JAMSHEDPUR	JSD	BOMBAY	SCZ	COIMBATORE	CMB
KOTA	KTA	CALCUTTA	ALP	AHMADNAGAR	AMN	SALEM	SLM
ALLAHABAD	ALB			AURANGABAD	AGD	CUDDALORE	CDL
UDAIPUR	UDP			JAGDALPUR	JGD	PONDICHERRY	PDC
				GOPALPUR	GPL	TIRUCHCHIRAPALLI	TRP
				POONA	PNE	KARAIKAL	KRL
				RATNAGIRI	RTN	NAGAPPATTINAM	NPT
				SHOLAPUR	SLP	COCHIN(IN-NAVY)	CHN
				GOA(PANJIM)	PJM	MADURAI	MDU
						PALAYAMKOTTAI	PLM
						THIRUVANANTHAPURM	TRV

Table.1 Meteorological Stations selected for the location specific study

Meteorological subdivision of India ; Topography and distribution of stations



3.Bias Correction

The Nearest Neighborhood (NN) statistical bias correction has been applied at station level in the 24 to 120 hour GFS location specific forecast of maximum and minimum temperature for 100 synoptic stations over India during summer (May –August 2012) and winter (November 2012 –February 2013) season

The bias $b_k(t)$ for each station (k) and each lead-time (t) is defined as

$$b_k(t) = f_k(t) - O_k(t)$$

The systematic bias B_k (t) at each station is computed daily by applying the weight coefficient wt_k (t) at each forecast hour as,

$$B_k(t) = Wt_k(t) * b_k(t)$$

The new bias-corrected model forecast Fk (t) will be generated by applying the bias Bk(t) to current direct forecasts fk(t) at each station

$$F_k(t) = f_k(t) - B_k(t)$$

Verification Methods

skill score (SS)

In addition to ME,MAE and RMSE error, we consider a skill score (SS) defined in terms of mean-squared error (MSE).

$$SS = 1 - \frac{MSE_f}{MSE_c}$$

where

$$MSE_f = \frac{1}{N}\sum_{i=1}^{N}(F_i - O_i)^2$$

forecast

$$MSE_c = \frac{1}{N}\sum_{i=1}^{N}(\bar{O}i - O_i)^2$$

climatology

observed daily climatology 1981 -2005 is used. (IMD)

A detailed description of mean squared skill score (SS) is provided by WMO (2002).

Result and Discussions

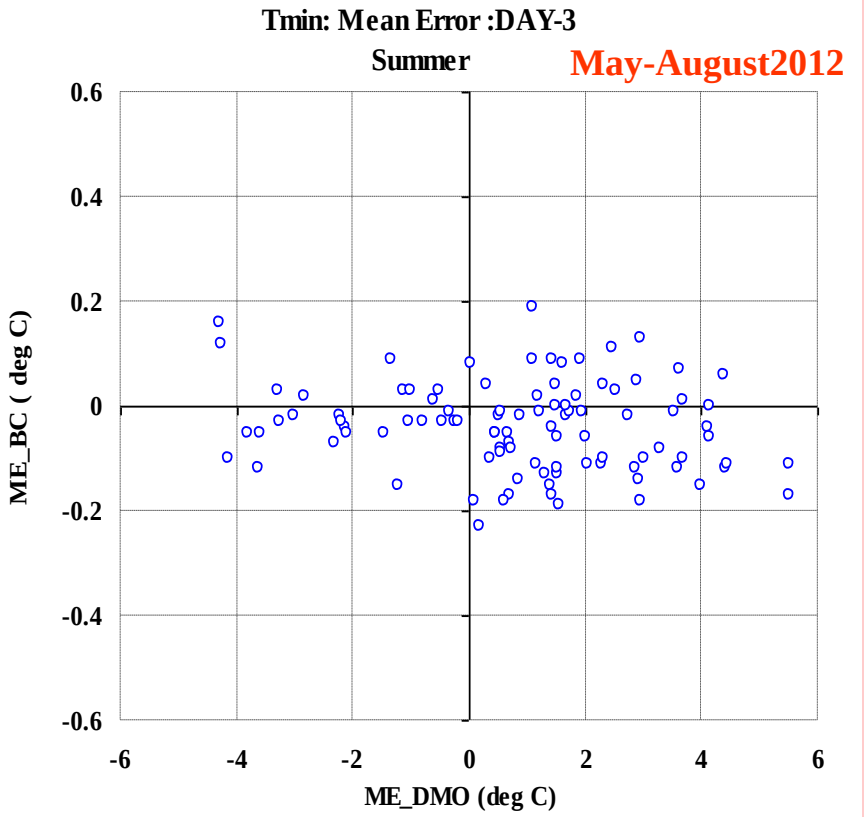
A quantitative inter-comparison of error statistics between Direct model output (DMO) and Bias Corrected (BC) minimum and maximum temperature forecast for 100 stations over India are discussed in this section.

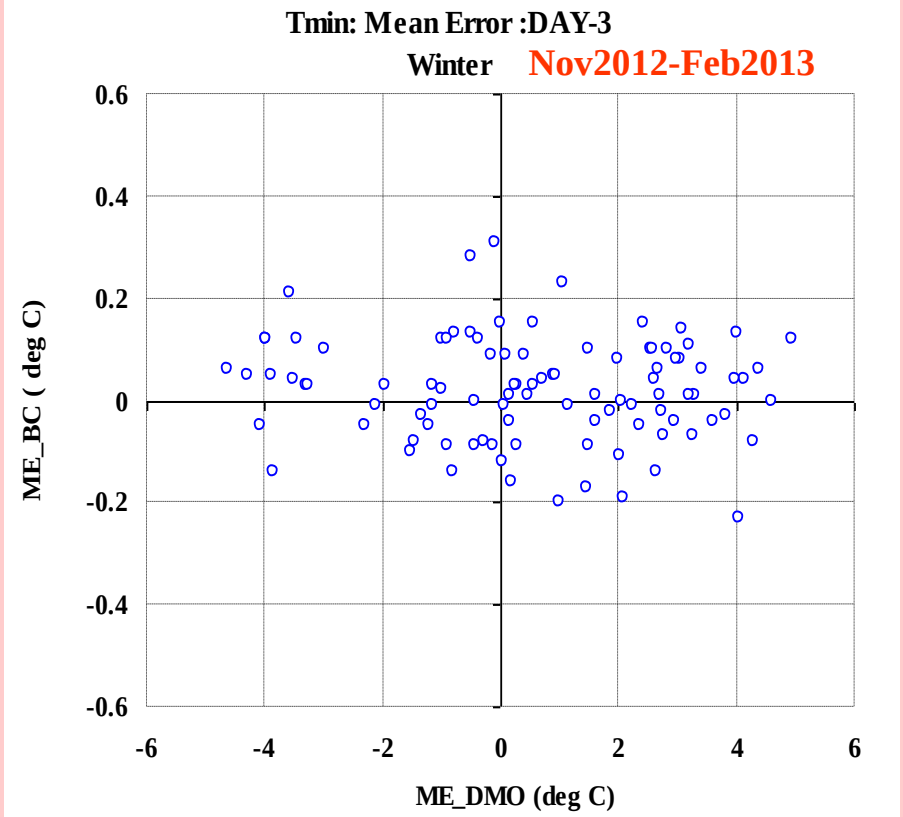
Verification of

1) **Minimum Temperatures**
2) **Maximum temperature at each meteorological stations during summer (May –August 2012) and winter (November 2012 –February 2013) seasons**

Mean Error

Verification of Minimum temperature



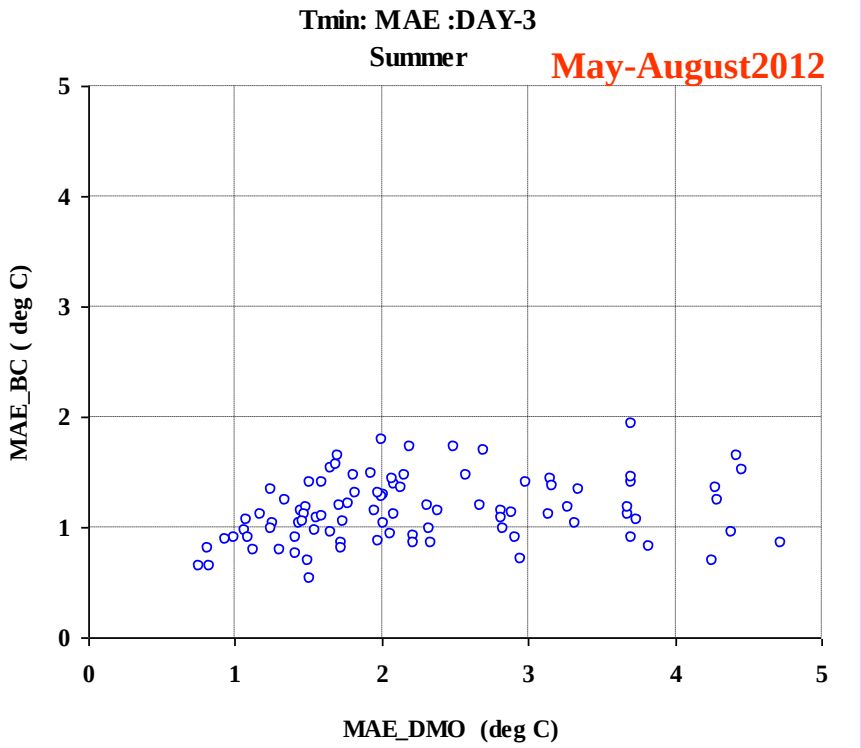


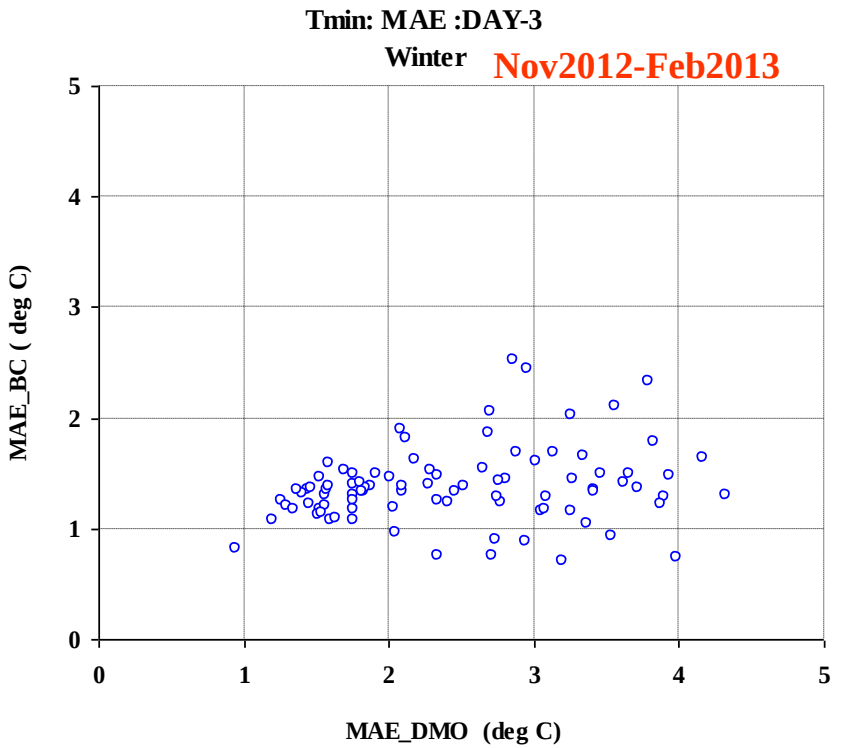
- The scatter diagram shows that the mean Error is in the order of -0.2 to +0.2 deg C for BC forecast, while it is in the order of -4 to +4 deg C for DMO in both summer and winter season for most of the stations (out of 100).

- The comparison also shows that the BC min temperature forecast produces bias values almost close to zero for most of the stations in summer as well as in winter.

Mean Absolute Error (MAE)

Verification of Minimum temperature



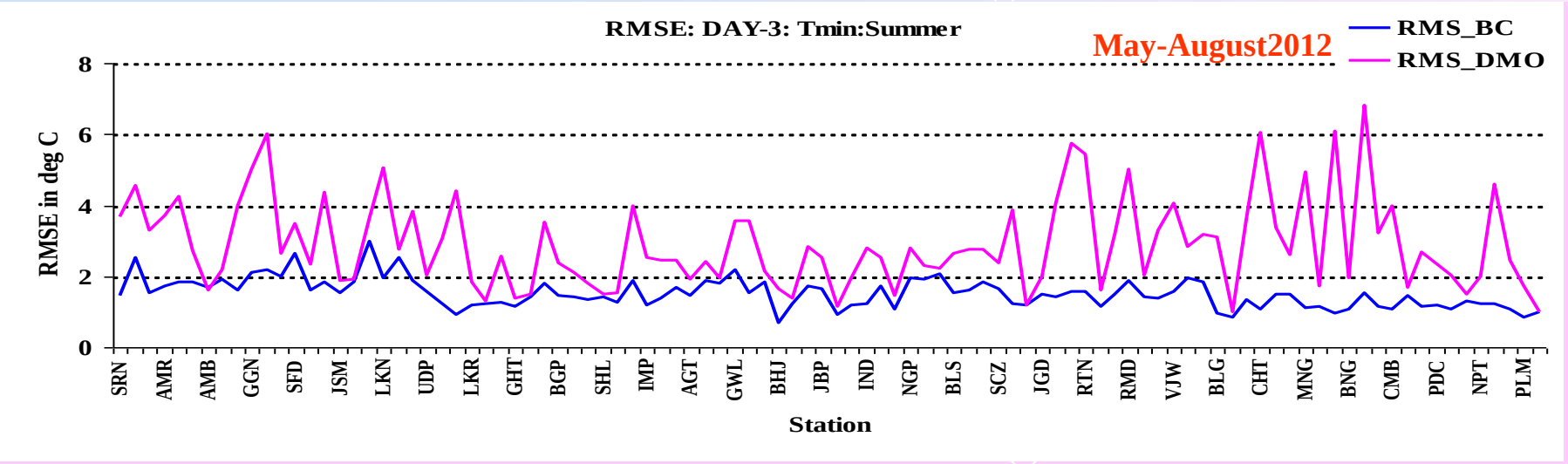


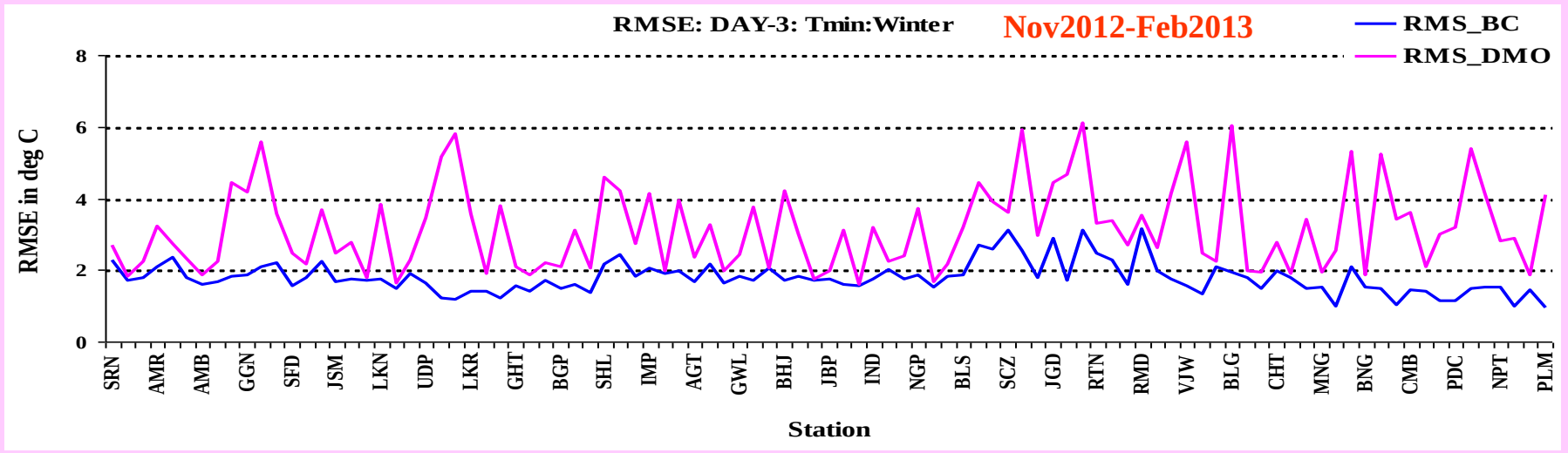
- The BC minimum temperature forecast shows MAE below 1.5 ◦C during summer and below 2 deg C during winter season for most of the station, while DMO MAE is below 2.5 ◦C in summer and below 3 ◦C in winter for most of the stations.
- We observe also that only a few stations have a DMO MAE as low as that using the BC forecast

RMSE

Verification of Minimum temperature

(a)

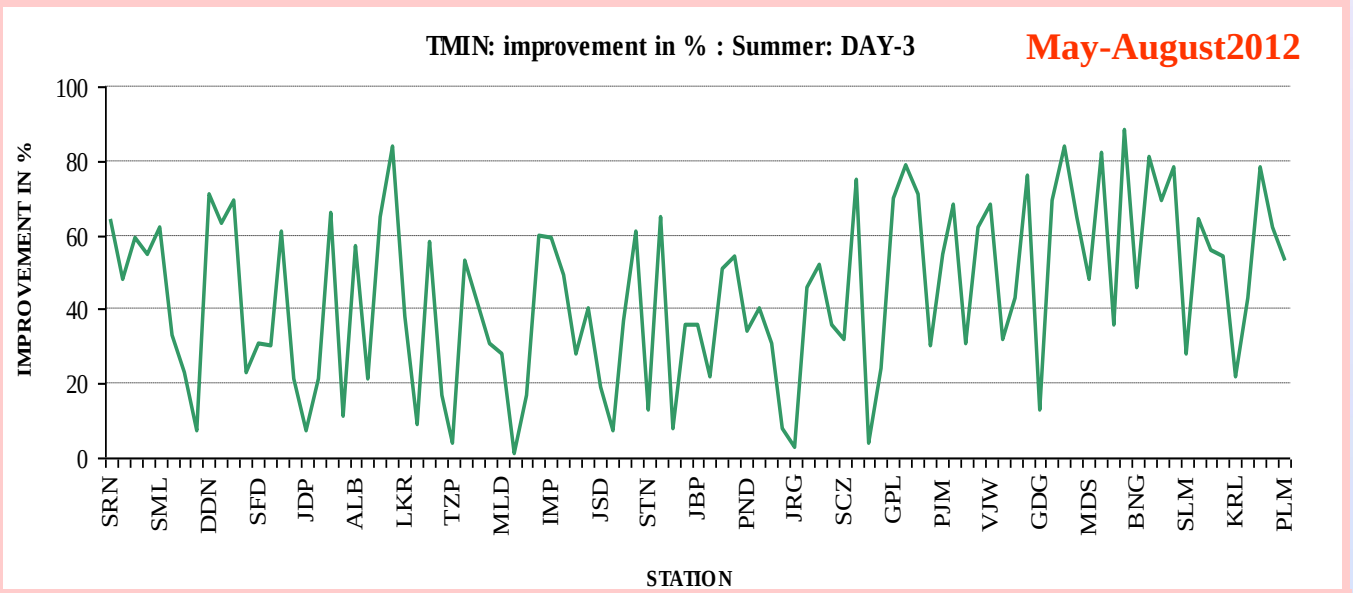


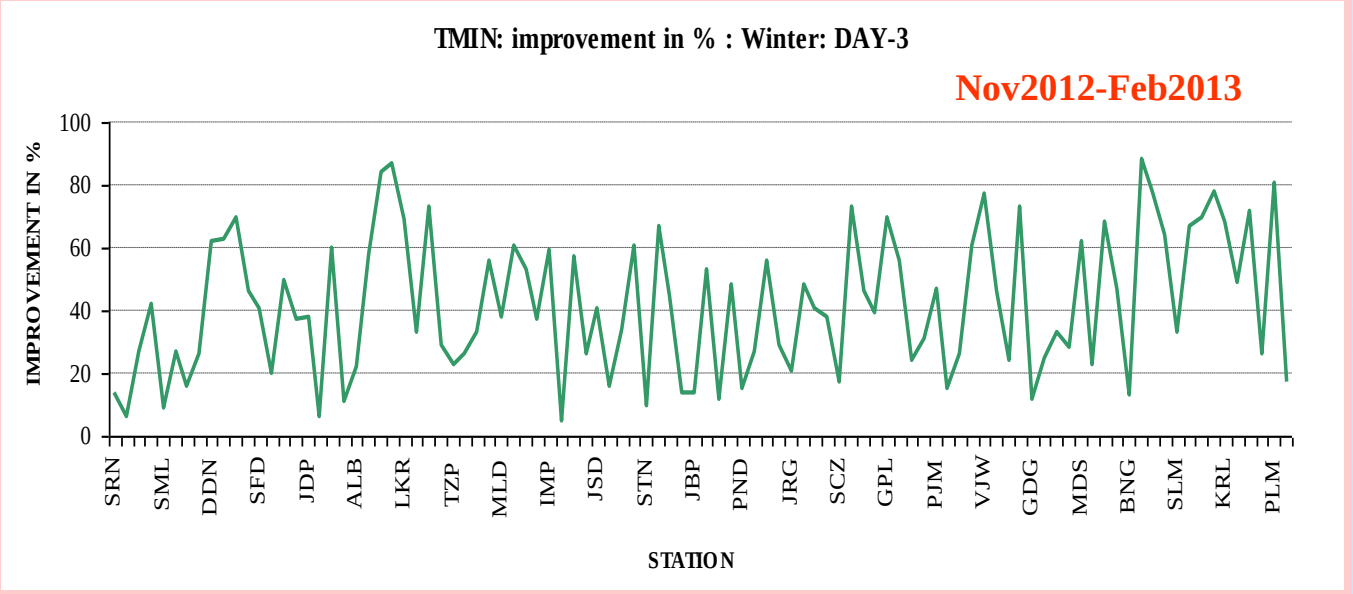


The day-3 RMSE in minimum temperature forecast for all 100 stations indicates that the DMO RMSE is higher than BC RMSE in all the stations during both summer and winter seasons.

Improvement in % (MAE)

Verification of Minimum temperature

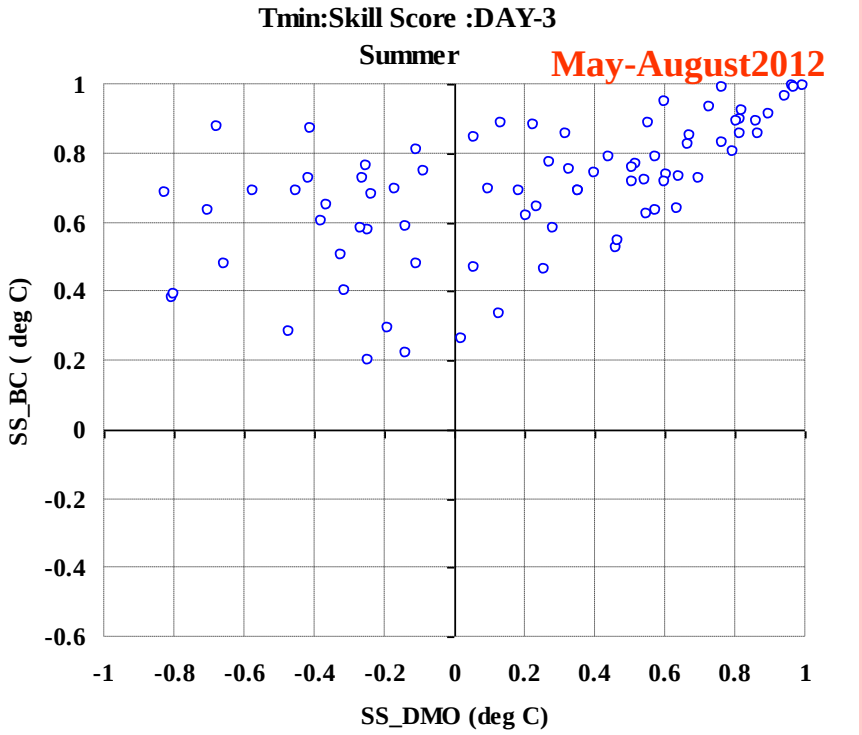


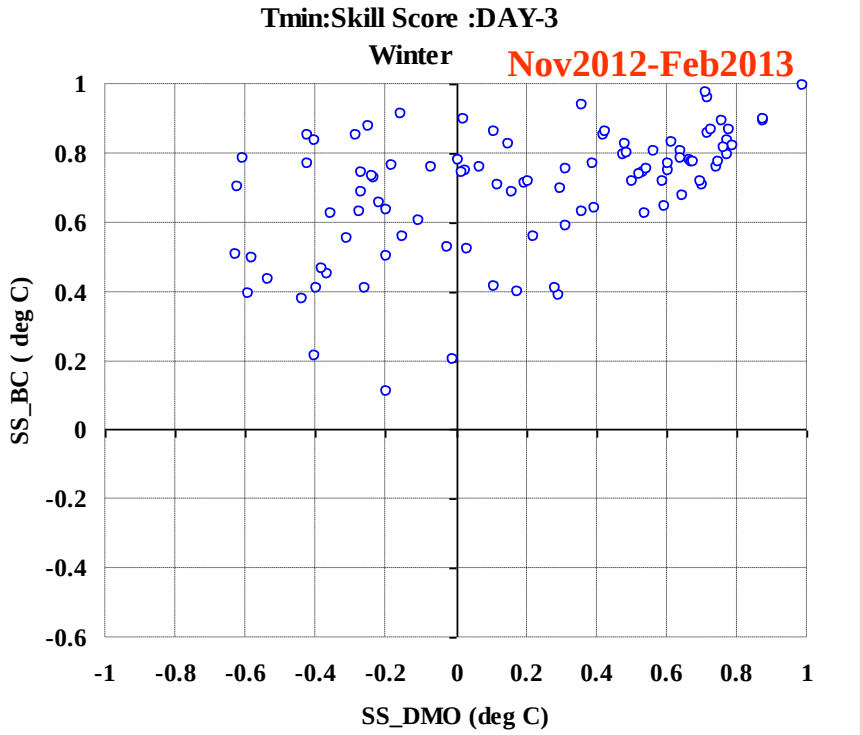


- The MAE skill score ranges from 0 to 100 with value of zero indicating no improvement skill and a value of 100 is for perfect forecasting skill.
- We observed a significant reduction of MAE in BC forecast as compared to DMO in all the stations during both summer and winter seasons

Skill Score (SS)

Verification of Minimum temperature

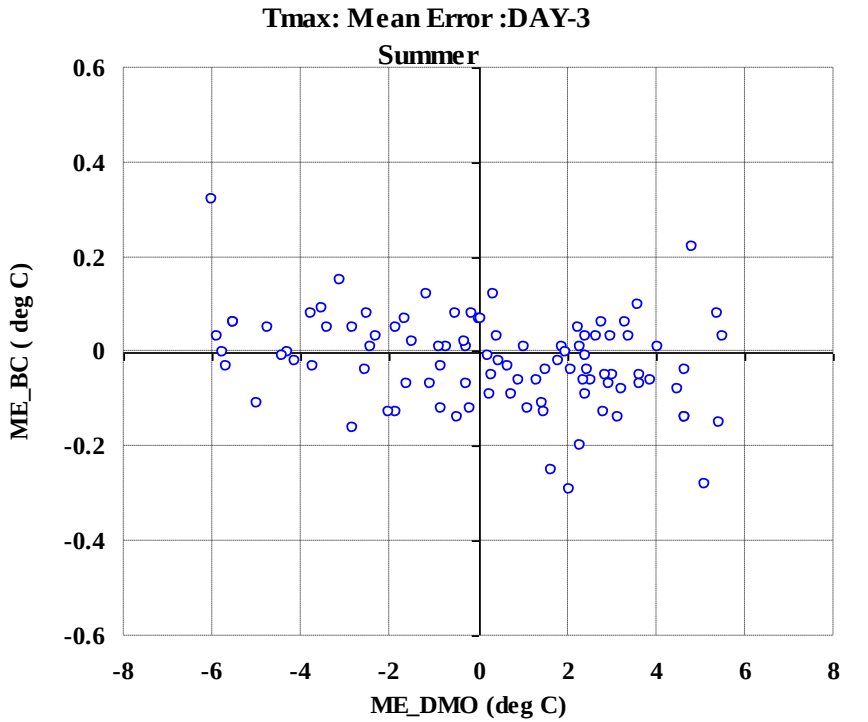


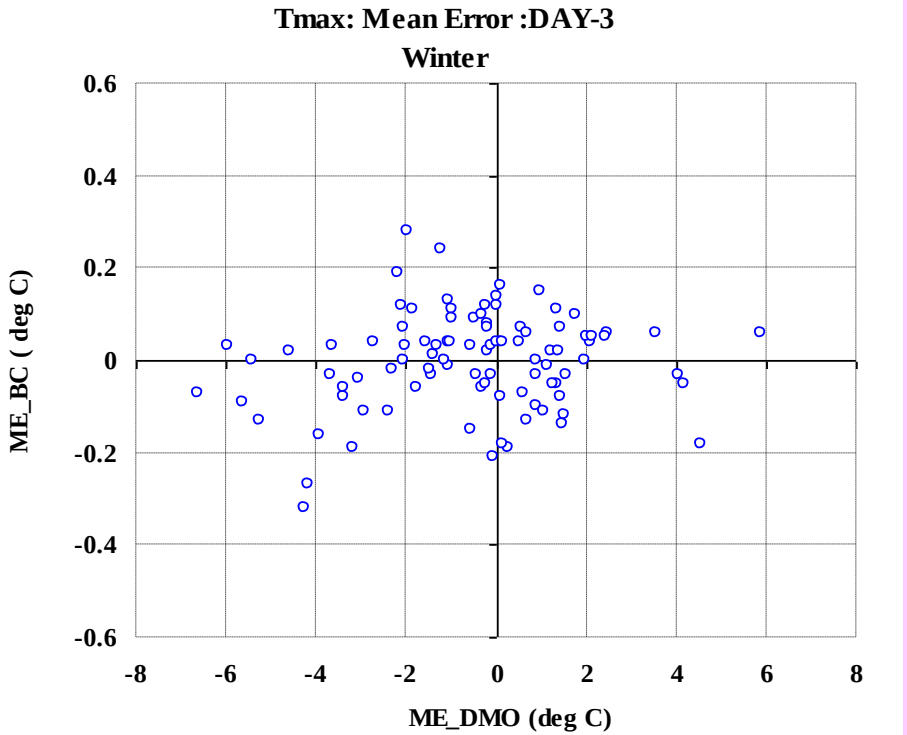


- The MSE skill score for BC forecast is better than both DMO and climatology reference forecast and its skill score values are greater than 0.6 for most of the stations during both summer and winter.
- The DMO skill score value of less than or equal to zero is observed in some of the stations in both summer and winter seasons.
- It is seen that the BC forecast skill scores for minimum temperatures is reasonably high for all the station in both the seasons.

Mean Error

Verification of Maximum temperature



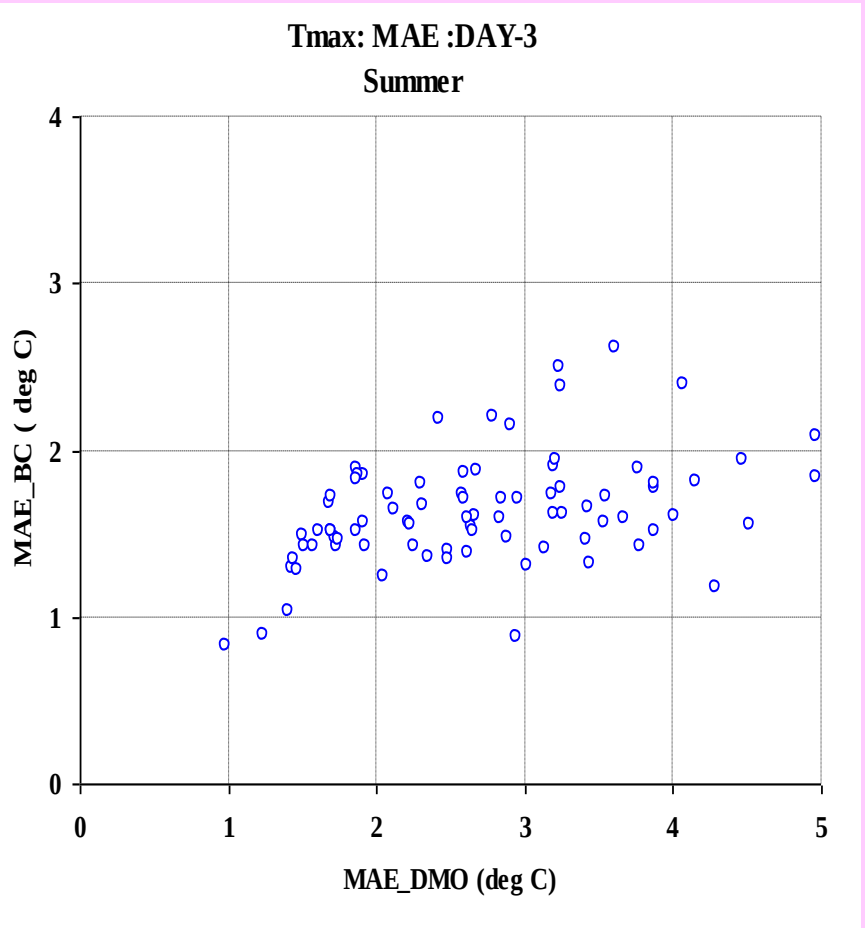


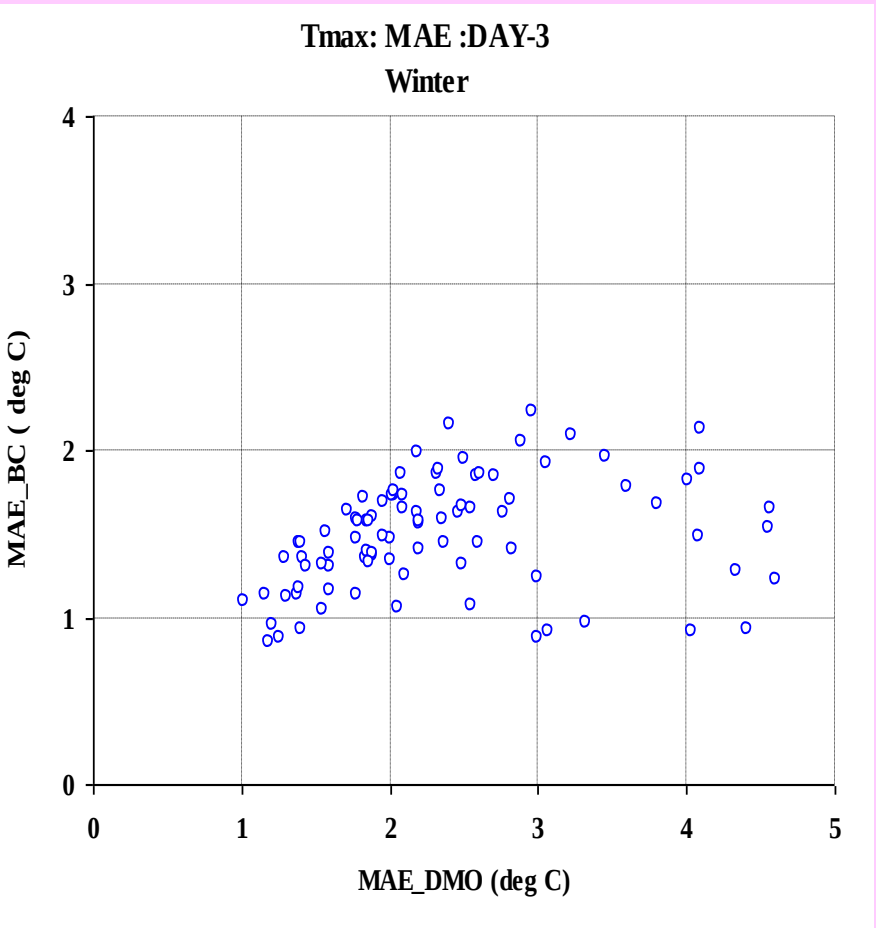
❖**The DMO maximum temperature mean error is between -6.0 ◦C and +4 0 ◦C in summer while it is between -4.0 ◦C and +2 0 ◦C in winter season.**

❖ **It is also seen that the BC forecast produces ME values close to zero for most of the stations in both summer and winter seasons.**

Mean Absolute Error (MAE)

Verification of Maximum temperature

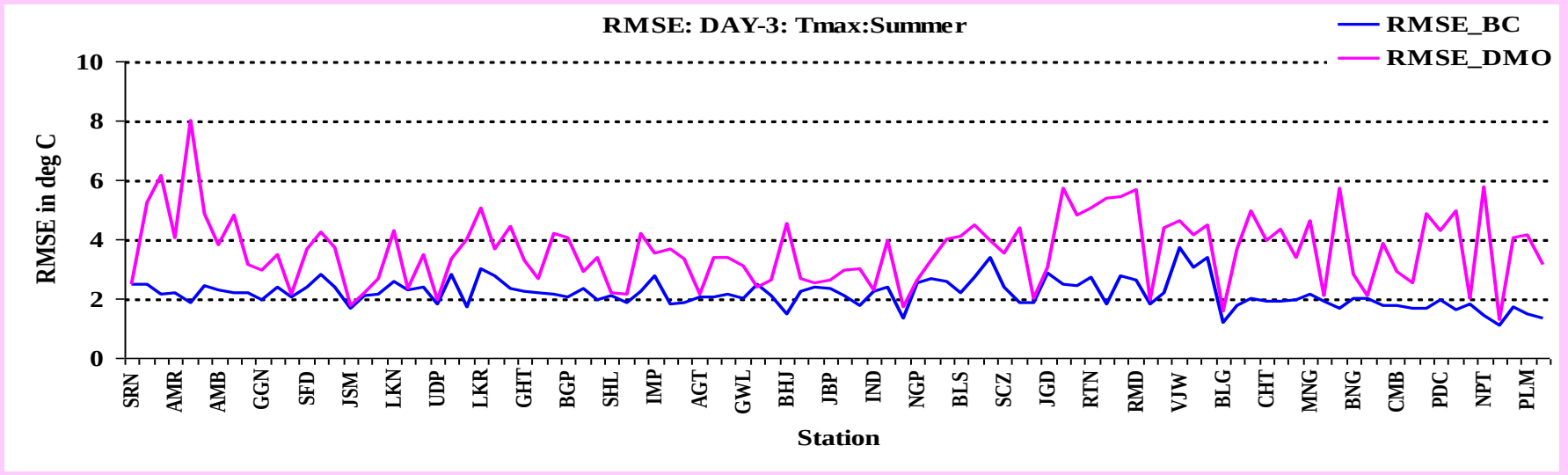


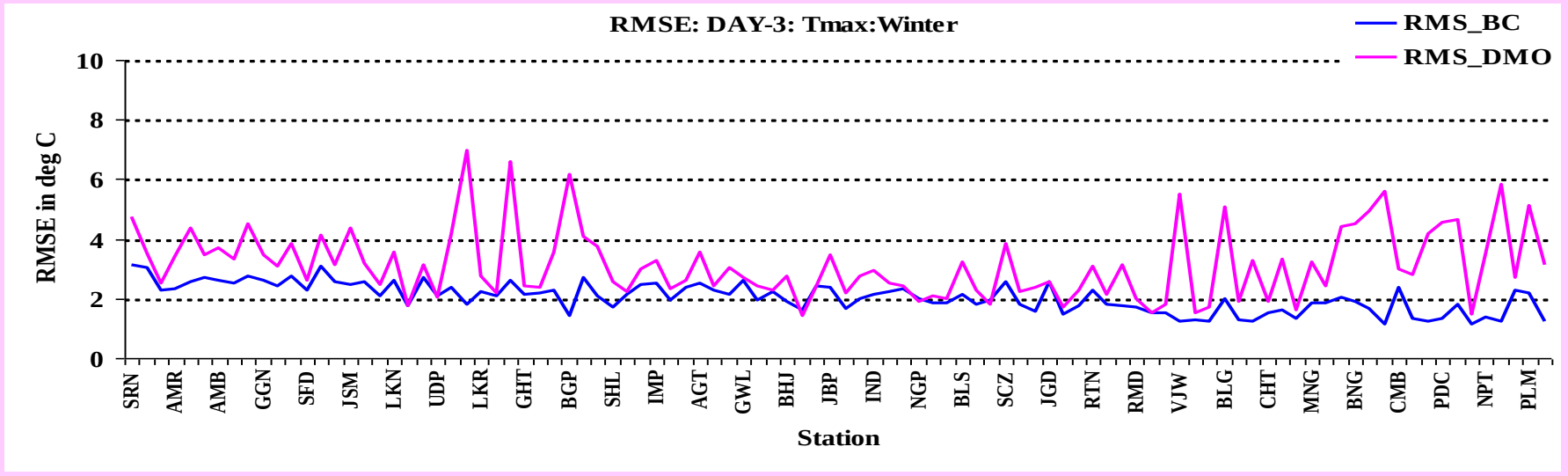


- The scatter diagram shows that the MAE ranged from 1.5 ◦C to 4 ◦C for DMO and from 1 ◦C to 2 ◦C for BC at most of the stations in both summer and winter.
- The MAE in winter season is lesser than that in summer season.

RMSE

Verification of Maximum temperature

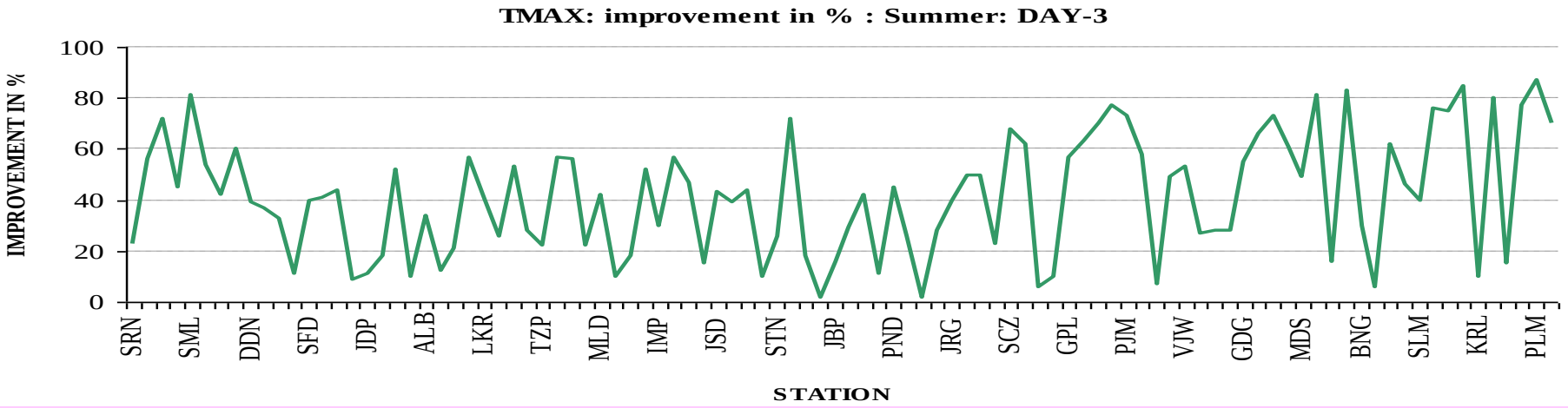


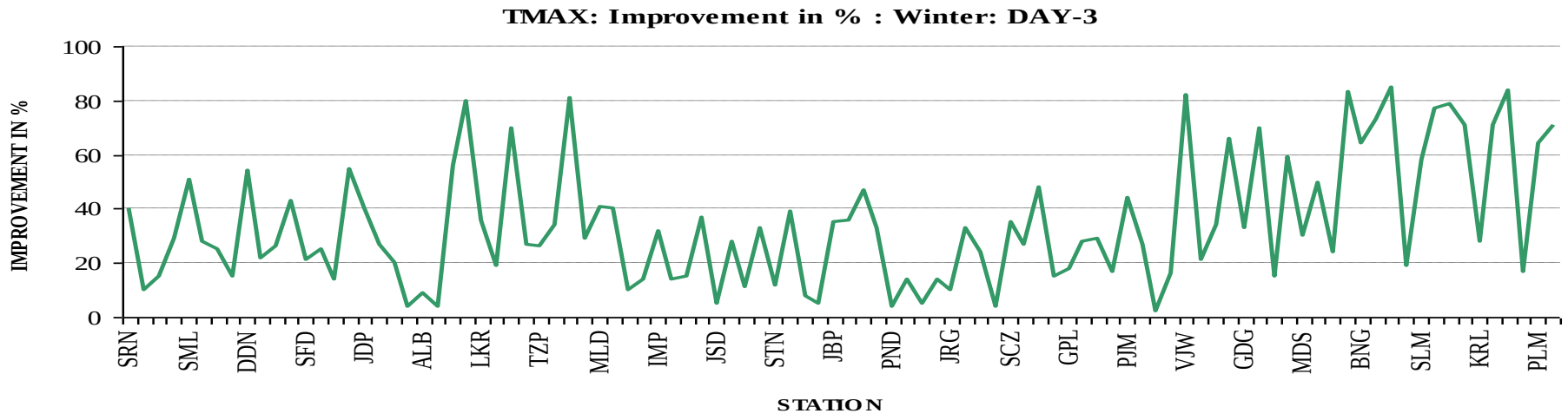


The DMO RMSE value varies between 3 ◦C and 4 ◦C, while the BC RMSE values ranges between 1.5 ◦C and 2.5 ◦C.

Improvement in % (MAE)

Verification of Maximum temperature



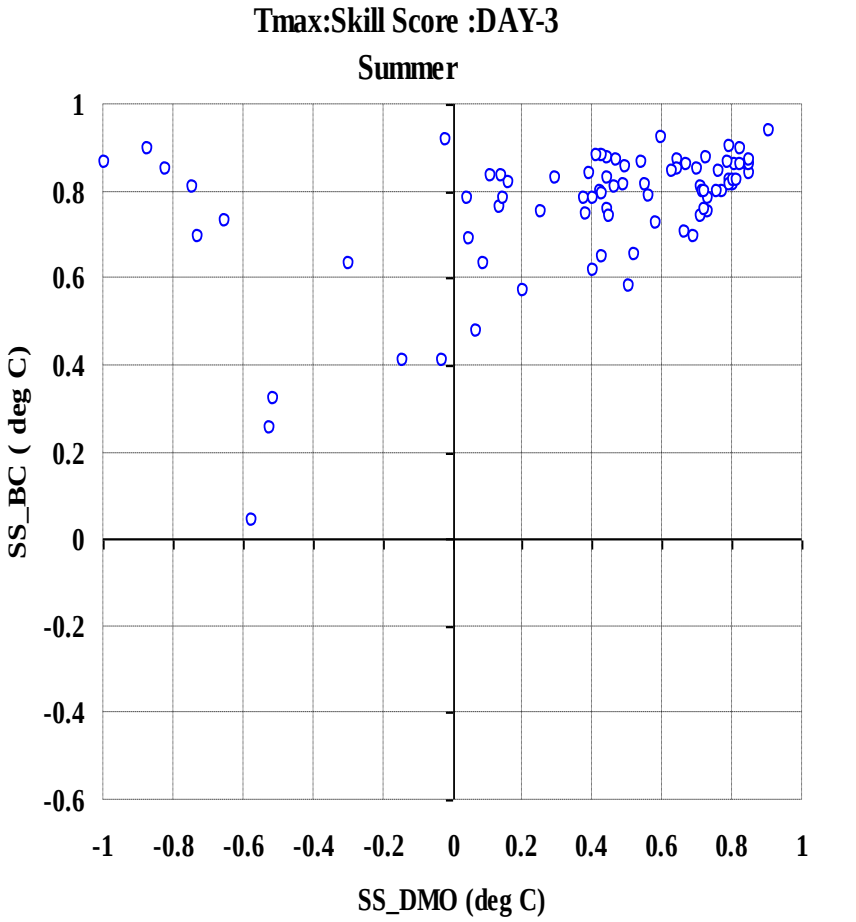


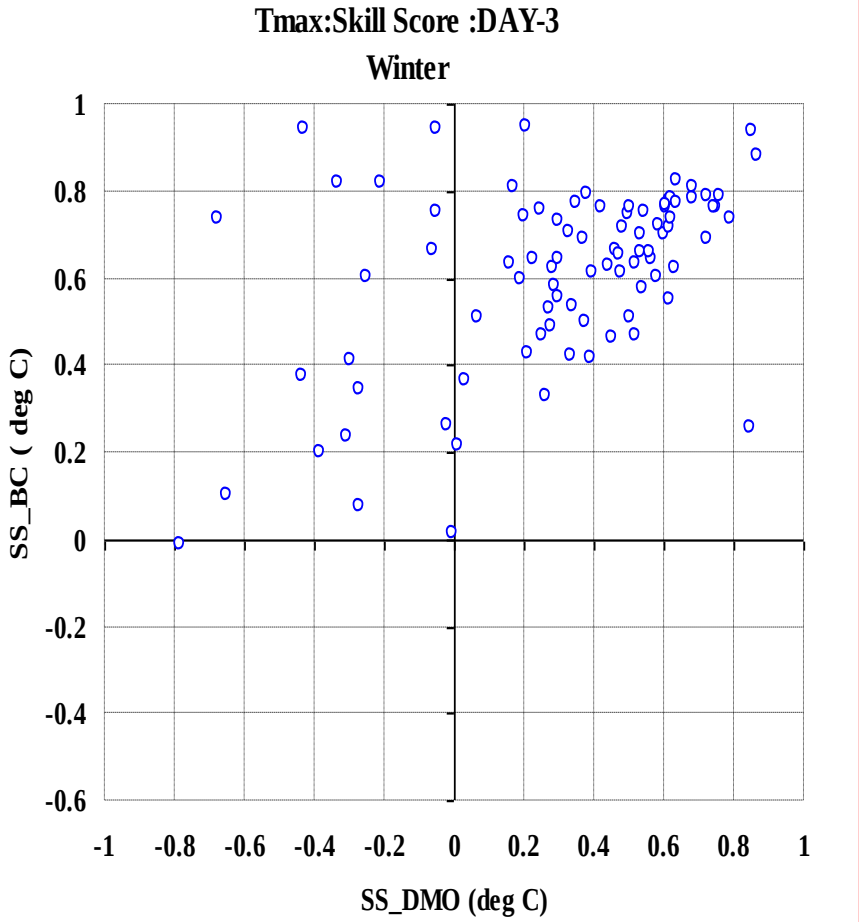
❖**The BC forecast has significant improvement in forecasting maximum temperature, from 20 to 30% over the DMO forecasts.**

❖ **More improvement (more than 50 %) is noticed over SP India regions in both summer and winter season.**

Skill Score (SS)

Verification of Maximum temperature





The maximum temperature forecasts skill score for most of the stations are varied from 0.7 to 0.9 for BC and from -0.4 to 0.8 for DMO in summer , while it is varied from 0.5 to 0.8 for BC and from -0.4 to 0.6 for DMO in winter

Conclusion

- The statistical Bias Corrected GFS forecast shows significant reduction (35 to 50%) of error in minimum and maximum temperature as compared to the DMO over most of the stations during both summer and winter seasons.

- This study also indicates that the BC GFS forecast improves over the GFS DMO reasonably and hence can be used for location specific forecast in real time


THANKS