NMRF/TR/02/2018

सत्यमेव जयते

TECHNICAL REPORT

Implementation of Very High Resolution (12 km) Global Ensemble Prediction System at NCMRWF and its Initial Validation

Ashu Mamgain, Abhijit Sarkar, Anumeha Dube, Arulalan T., Paromita Chakraborty, John. P. George and E.N. Rajagopal

August 2018

National Centre for Medium Range Weather Forecasting
Ministry of Earth Sciences, Government of India
A-50, Sector-62, NOIDA-201309, INDIA

Implementation of Very High Resolution (12 km) Global Ensemble Prediction System at NCMRWF and its Initial Validation

Ashu Mamgain, Abhijit Sarkar, Anumeha Dube,
Arulalan T., Paromita Chakraborty, John. P. George
and E.N. Rajagopal

August 2018

National Centre for Medium Range Weather Forecasting
Ministry of Earth Sciences
A-50, Sector 62, NOIDA-201309, INDIA

Ministry of Earth Sciences
National Centre for Medium Range Weather Forecasting
Document Control Data Sheet

1	Name of the Institute	National Centre for Medium Range Weather Forecasting
2	Document Number	NMRF/TR/02/2018
3	Date of publication	August 2018
4	Title of the document	Implementation of Very High Resolution (12 km) Global Ensemble Prediction System at NCMRWF and its Initial Validation
5	Type of Document	Technical Report
6	No. of pages, figures and tables	21 pages, 8 figures and 4 tables
7	Number of References	15
8	Author (s)	Ashu Mamgain, Abhijit Sarkar, Anumeha Dube, Arulalan T., Paromita Chakraborty, John. P. George and E.N. Rajagopal
9	Originating Unit	NCMRWF
10	Abstract	There are increasing efforts towards the prediction of high impact weather systems using state-of-the-art numerical models. The global Ensemble Prediction System of NCMRWF (NEPS) of 12 km horizontal resolution has been operationally implemented to improve weather forecasts and services. The NEPS global model configuration is based on the recent version of UK Met Office Global and Regional Ensemble Prediction System (MOGREPS). The initial condition perturbations are generated by Ensemble Transform Kalman Filter (ETKF) method. The model uncertainties are taken care by the Stochastic Kinetic Energy Backscatter (SKEB) and Random Parameters (RP) schemes. The forecast perturbations obtained from 6 hour short forecasts of 22 ensemble members are updated by ETKF four times a day (00, 06, 12 and 18 UTC). Perturbations of surface parameters such as sea-surface temperature, soil moisture content and soil temperature are also included in the current NEPS. The NEPS aims to provide 10-day probabilistic forecasts using 23 ensemble members (22 perturbed + 1 control). The 12-km NEPS shows improvements in terms of forecast agreement among the members in comparison to previously operational 33-km NEPS. The ratio between root-mean-square error of ensemble mean and ensemble spread as a function of lead time has improved in both northern and southern hemispheres in the 12-km NEPS.
11	Security classification	Non-Secure
12	Distribution	Unrestricted Distribution
13	Key Words	Ensemble, Perturbation, Data Assimilation, NEPS

Table of Contents

	Page No.
Abstract	1
1. Introduction	2
2. Methodology	3
2.1 Brief Description of NEPS	3
2.2 Computational infrastructure	3
2.3 Rose suite/Cylc scheduler	6
3. Description of the Components of NEPS	7
3.1TrimObstore	7
3.2 OPS	7
3.3 Reconfiguration	8
3.4 ETKF Perturbations	8
3.5 SST, SMC and TSOIL Perturbations	9
3.6 Short forecasts	10
3.7 Long forecasts	11
4. Cold Start of NEPS	11
5. NEPS Forecast Products	12
6. Comparison of 12-km NEPS with 33-km NEPS	15
6.1 Spread-Skill Relationship	15
6.2 Ensemble Mean of a Heavy Rainfall Event	17
6.3 Probability of Rainfall Exceedance	17
7. Summary	19
Acknowledgments	20
References	20

Abstract

There are increasing efforts towards the prediction of high impact weather systems using state-of-the-art numerical models. The global Ensemble Prediction System of NCMRWF (NEPS) of 12 km horizontal resolution has been operationally implemented to improve weather forecasts and services. The NEPS global model configuration is based on the recent version of UK Met Office Global and Regional Ensemble Prediction System (MOGREPS). The initial condition perturbations are generated by Ensemble Transform Kalman Filter (ETKF) method. The model uncertainties are taken care by the Stochastic Kinetic Energy Backscatter (SKEB) and Random Parameters (RP) schemes. The forecast perturbations obtained from 6 hour short forecasts of 22 ensemble members are updated by ETKF four times a day (00, 06, 12 and 18 UTC). Perturbations of surface parameters such as sea-surface temperature, soil moisture content and soil temperature are also included in the current NEPS. The NEPS aims to provide 10-day probabilistic forecasts using 23 ensemble members (22 perturbed + 1 control). The 12-km NEPS shows improvements in terms of forecast agreement among the members in comparison to previously operational 33-km NEPS. The ratio between root-mean-square error of ensemble mean and ensemble spread as a function of lead time has improved in both northern and southern hemispheres in the 12-km NEPS.

1. Introduction

The forecast of the future state of the atmosphere using a single deterministic model is unlikely to exactly match the true state because of uncertainties in specifying the initial state and in representation of atmospheric processes in the model. The Ensemble Prediction Systems (EPS) are numerical weather prediction systems in which several scenarios of the same model with slightly different initial conditions are used to estimate the uncertainty in a weather forecast as well as the most likely outcome. One of the main goals of EPS is to provide the probabilistic forecast of details of extreme weather events which requires its resolution to be sufficiently high to resolve the small scale features of these events. Besides these, the location specific forecasts like ensemble meteogram or EPSgram and storm following meteograms of high resolution EPS should be closer to the reality.

The NCMRWF EPS (NEPS) has been upgraded to ~12 km horizontal resolution with 22 members in the newly acquired Mihir HPC. Previous version of NEPS with horizontal resolution of ~33 km and having 44 members was running in Bhaskara HPC. The details about pervious operational NEPS and related forecast products are given in Sarkar et. al. (2016). The recent version has been implemented operationally and providing products regularly from 1st June 2018. The new 12-km NEPS is expected to be more skillfull, especially in generating more accurate and area specific forecast of extreme weather events like rains, heat wave and cold wave, the track and the intensity of the cyclonic storms due to its very high horizontal resolution. However, predicting the short term/small scale features like thunderstorm events may still be challenging.

The HPC System (HPCS) resources available with Ministry of Earth Sciences (MoES) were augmented recently to 6.8 Peta Flops (PF) and the same has been installed at Indian Institute of Tropical Meteorology (IITM), Pune with 4.0 Peta Flops HPCS (Pratyush) and National Centre for Medium Range Weather Forecasting (NCMRWF), Noida with 2.8 Peta Flops HPCS (Mihir) in January 2018. Implementation of a 12-km global EPS for the first time in the world became possible due to above mentioned HPCS augmentation. Currently most of the leading NWP centres have their global EPS at horizontal resolutions in the range of 18-20 km (viz., ECMWF: 18 km, UK Met Office: 20 km).

2. Methodology

2.1 Brief Description of NEPS

NCMRWF global Ensemble Prediction System (NEPS) was upgraded from ~33 km (N400L70) to ~12 km (N1024L70) resolution. It is based on Unified Model version 10.8 (UM10.8) which is a part of latest 'Operational Parallel Suite' (PS40) developed at Met Office, UK. Apart from horizontal resolution, there is not much difference in dynamics or representation of physical processes in the new model compared to the previous version NEPS based on UM8.5. However, this time more number of observations and surface perturbations such as sea surface temperature, deep soil temperature and soil moisture content are included. A total of 23 ensemble members (22 perturbed forecasts + 1 control forecast) constitute this ensemble system. The 22 analysis perturbations of horizontal wind speed components (u and v), potential temperature (θ), specific humidity (q) and exner pressure (π) are generated by Ensemble Transfer Kalman Filter (ETKF) method from the forecast perturbations of previous cycles four times a day (00, 06, 12 and 18 UTC) at all 70 model vertical levels. Perturbations have also been added for deep soil temperature, soil moisture content on four model soil levels and sea surface temperature. These analysis perturbations are added to the reconfigured analysis obtained from the flow dependent, hybrid four-dimensional variational data assimilation system (hybrid-4DVar; Clayton et al., 2013) as a part of PS40 suite. A 10.5 days forecast of NEPS is routinely generated based on 00 UTC and 12 UTC initial conditions which include a control forecast starting from hybrid-4DVar analysis and 22 (11 from 12 UTC of previous day + 11 from 00 UTC of current day) ensemble members starting from perturbed initial conditions. Details of the configuration of the new 12-km NEPS is given in Table 1. The sequences of all the processes involved in the execution of the new 12-km NEPS are represented by the flow diagram in Figure 1.

2.2 Computational Infrastructure

NEPS is running operationally in Mihir HPCS at NCMRWF. Mihir HPCS is a Cray-XC40 Liquid Cooled System with 2320 nodes running with a peak performance of 2.8 PF and a total system memory of 290 TB. Number of nodes used by NEPS and the wall clock time taken by each component are given in Table 2. The 12-km NEPS uses 550 compute nodes for about 5.5 hours for the long forecasts (10.5 days) at 00 and 12 UTC respectively.

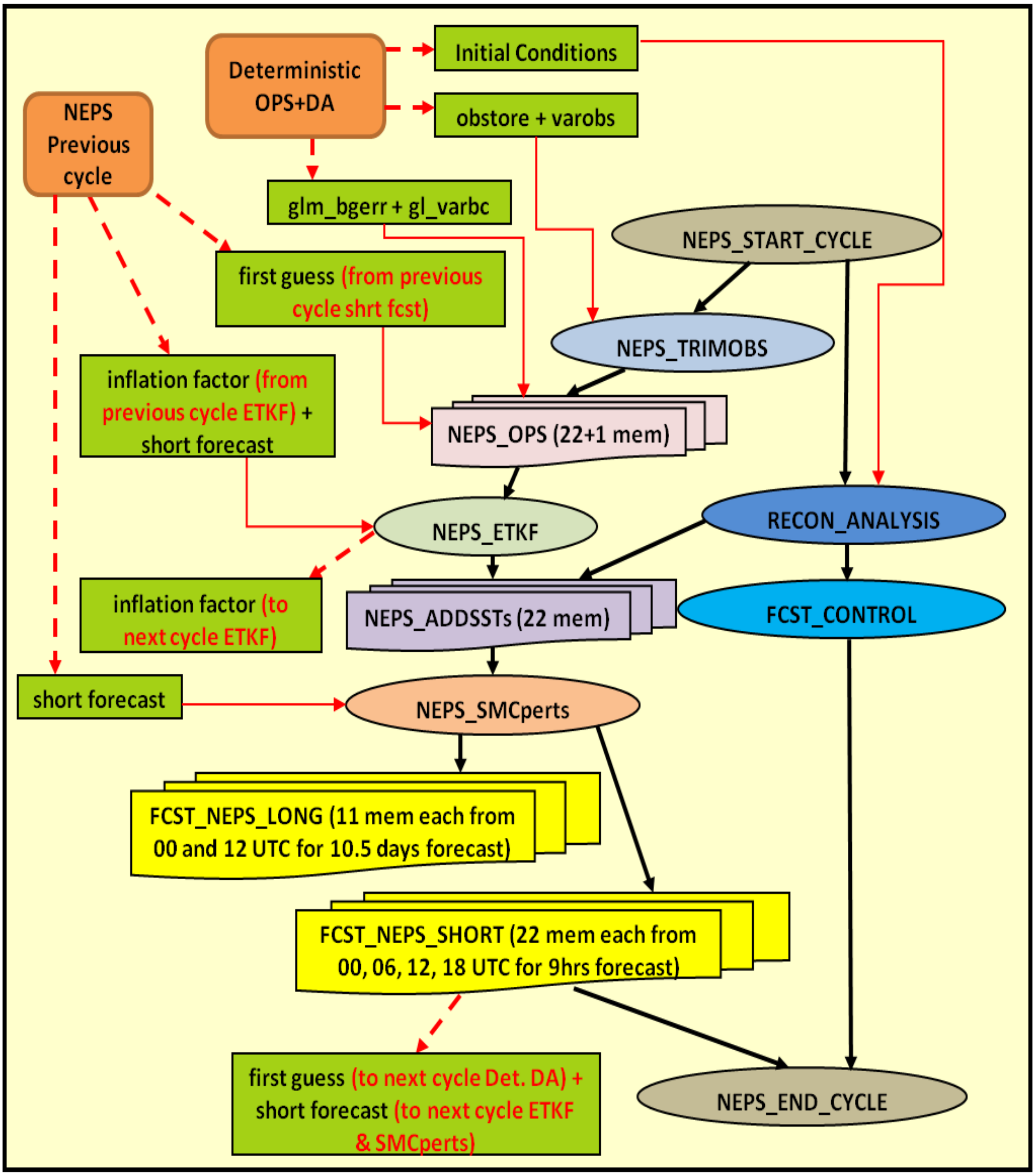

Figure 1: Flow chart of the NEPS (dashed red lines indicate output data from a task and solid red lines indicate input data to a task)

Table 1: Salient Features of the 12-km NEPS

Model Details	*Initial Condition and Perturbations*
Model: Unified Model; Version 10.8	**Initial condition:** Analysis from global deterministic Hybrid 4DVar atmospheric data assimilation (DA) system.
Domain: Global	**DA Resolution:** N320L70 (~40 km) with N144L70 Hessian based pre-conditioning
Resolution: 12 km, 70 Levels (~80 km model top)	**DA Method:** Hybrid incremental 4D-Var. Information on "errors of the day" is provided by NEPS forecast at every data assimilation cycle
Grid Points: 2048 x1536	**DA Cycles:** 4 analyses per day at 00, 06, 12 and 18 UTC. Observations within +/- 3 hrs from the cycle time is assimilated in the respective DA cycle
Time Step: 5 minutes	**Model Physics Perturbations:** Stochastic Kinetic Energy Backscatter (SKEB) and Random Parameters (RP) schemes
Parameterizations: Based on GA6.1 (Walters et al., Geosci. Model Dev., 10: 1487-1520, 2017)	**Initial Condition Perturbations:** Ensemble Transform Kalman Filter (ETKF) method.
Long Forecast length: 10.5 days (based on 00 and 12 UTC initial conditions)	**Surface Perturbations:** SST perturbations, Deep soil temperature and Soil moisture perturbations
Short Forecast length: 9 hours (based on 00, 06, 12 and 18 UTC initial conditions)	**SST data:** Updated at 12 UTC DA cycle with OSTIA based SST and sea-ice analysis
No. of Ensemble members: 22	**Snow Analysis data:** Satellite derived snow analysis. Updated at 12 UTC DA cycle
	Soil Moisture analysis Method: Extended Kalman Filter Analysis time: 00, 06, 12 and 18 UTC Observations assimilated: ASCAT soil wetness observations, Screen Temperature and Humidity (pseudo observations from 3D-Var screen analysis)

EPS Tasks	Start Time (IST)	Wall clock time (min.)	Number of Nodes	Input from Deterministic OPS+DA
TrimObstore	0800, 1700, 2200, 0530	~3	1	glm_obstore glm_varobs
OPS	After trimobstore is over	~10	2 (for each 23 member)	glm_bgerr gl_varbc
ETKF	After OPS is over	~7	1	
SST perturbations	After ETKF perturbations and atmanl from DA	~1	1	atmanl (initial condition)
SMC perturbations	After SST perturbations are over	~6	1	
Long forecast (11 +1 members) for 10.5 days	9:45, 23:50	~330/member	50 (for each member)	
Short forecast (22 +1 members) for 9 hours	10:00, 17:30, 00:00, 06:00	~15/member	45 (for each member)	

Table 2: Operational Node Usage/Timings in Mihir HPC by the components of 12-km NEPS

2.3 Rose suite/Cylc scheduler

Python based Rose/Cylc environment is used for managing and running operational jobs of the new NEPS. Rose is a framework for managing and running meteorological suites. Suite is a collection of scientific application softwares for a common purpose. Rose contains all the features required for configuration management of suites and their components. Cylc is a workflow engine for cycling systems (tools for managing the workflows required by the Rose) that drives task submission and monitoring. Cylc has all the key features required for both operational and research job scheduling including run, rerun, kill, poll, hold individual task or a family of tasks. NCMRWF uses Rosie database for suite management. Both Rose and Cylc are open source softwares managed under GitHub (https://github.com/cylc/cylc & https://github.com/metomi/rose).

3. Description of the Components of NEPS

3.1 TrimObstore

TrimObstore uses the ‘obstore’ and ‘varobs’ files for different types of observation. The description of the types of observations (obstore) that are used in NEPS is given in Table 3. The ‘varobs’ files used in TrimObstore are generated by the deterministic OPS but these files contain subsets of observations after a series of quality control steps and thinning.

Along with the ‘obstore’ files, deterministic OPS also reads the model background file from deterministic model short forecast and background error file from the deterministic OPS. The processing of observations is done in ‘extract and process’ task. The ‘extract’ part retrieves the observations from the ‘obstore’ files and the ‘process’ part carries out the jobs of quality control, thinning and rewriting the data in required formats. The observations and model data processed by the deterministic OPS are written in mainly three data structures. Those are (1) varobs; the quality controlled observations, (2) var.cx; the horizontally interpolated background fields at observation location for its use in OPS and (3) modelobs; the background fields exactly like observation fields at observation locations. Only the quality controlled and thinned observations from the ‘obstore’ files are written in ‘varobs’ files. Like the deterministic OPS, ensemble based OPS also needs to process ‘obstore’ files to generate ‘varobs’ and ‘modelobs’ files. To speed up this process by the OPS task of the NEPS, TrimObstore produces trimmed ‘obstore’ files which contain the same observations from original ‘obstore’ data at the location of corresponding ‘varobs’ data. These trimmed ‘obstore’ files are also used to update inflation factor by ETKF.

3.2 Observation Processing System (OPS)

The trimmed ‘obstore’ files from TrimObstore are used as input to the OPS task of the NEPS to generate ‘modelobs’ and ‘varobs’ files. The ‘modelobs’ files contain the model forecast of the observations. OPS carry out quality control of the observation including internal consistency checks, checks against model background and neighbouring observations. The processed observations are written in ‘varobs’ files. Each ‘obstore’ file has a corresponding ‘varobs’ file. The model background field (first guess) processing is also a part of the OPS ‘extract and process’ task. When the processing of first guess and observations are carried out in ‘extract and process’, resulting columns of model data are interpolated horizontally to the observation location for the data assimilation system as var_cx files. The ‘modelobs’ files also contain the first guess interpolated to observations' location. In ETKF, only 'modelobs' (not var_cx) files are used. Deterministic OPS also generates a background error file (glu.bgerr) using the 'Background Error Create' component. These geographically varying model errors are determined using model forecast tendency, model forecast gradient and background wind speed information taken from the first guess file from the model forecast. A detailed description of the deterministic OPS system implemented at NCMRWF is given in George et al. (2016). The calculation of transformation matrix in ETKF requires the model equivalent of each observation for every ensemble member. Successful completion of OPS task for each ensemble members provides these 'pseudo observations' in the form of 'modelobs' files as input to ETKF. The 'modelobs' from all the perturbed members and 'varobs' files from the control forecast are required in ETKF task.

3.3 Reconfiguration

At NCMRWF analysis data are prepared by Hybrid 4DVar DA system at 00, 06, 12 and 18 UTC using the deterministic UM model forecast as first guess. This analysis fields are further interpolated to the model resolution during reconfiguration process in order to generate a suitable initial condition to run the control and ensemble members of EPS. It is also verified that all the required initial conditions are present according to the model physics. If required, additional data can be added or updated from ancillary files in this step. However, in the present operational DA system in NCMRWF, full fields at model resolution (i.e., N1024) are present in the analysis file therefore we are skipping this step in present NEPS.

3.4 ETKF Perturbations

The inputs to the EPS forecast runs (both short forecast of 9 hours and long forecast of 10.5 days) are provided by Hybrid 4DVar DA system and the perturbations are generated by the ETKF method (Bowler et al., 2009). The control forecast does not need any input perturbation from ETKF. It uses only the reconfigured analysis at N1024L70 resolution as its initial condition. The magnitude and statistical error structure of the uncertainties associated with the analysis data are provided by the ETKF system. It generates ensemble perturbations by using information about the observation errors and the background perturbation structure. In the NEPS configuration, the ETKF cycles are running every six hours for all the 22 members. It updates the forecast perturbation matrix by multiplying it with a transformation matrix to generate analysis perturbations for wind components, potential temperature, specific humidity and exner pressure at all the model levels. ETKF uses the background 6-hr forecasts of previous cycle from each member to determine spread of the ensemble members. It compares this spread to the root mean square error of the ensemble mean with respect to the observation and then computes a region specific inflation factor which is multiplied with raw transformation matrix to improve ensemble spread. The analysis perturbations are added to the analysis data using the Incremental Analysis Update (IAU) scheme (Clayton, 2013) within the UM.

Table 3: Observations Assimilated in 12-km NEPS

Observation Type	Observation Description	Assimilated Variables
Aircraft	Upper-air wind and temperature from aircraft	u, v, T
AIRS	Atmospheric Infrared Sounder of MODIS	Tb
ATOVS	AMSU-A, AMSU-B/MHS, HIRS from NOAA-18 &19, MetOp-A&B	Tb
GOESClear	Cloud clear Imager radiances from GOES	Tb
GPSRO	Global Positioning System Radio Occultation observations from various satellites (including MT-ROSA)	Bending Angle
IASI	Infrared Atmospheric Sounding Interferometer from MetOp-A&B	Tb
Satwind	Atmospheric Motion Vectors from various geostationary and polar orbiting satellites (including INSAT-3D)	u, v
Scatwind	Advanced Scatterometer in MetOp-A & B, ScatSat-1, WindSat	u, v
SEVIRIClear	Cloud clear observations from SEVIRI of METEOSAT 8 &11	Tb
Sonde	Radiosonde, upper-air wind profile from pilot balloons, wind profilers, VAD winds from Indian DWR	u, v, T, q
Surface	Surface observations over Land and Ocean	u, v, T, q, Ps

3.5 SST, SMC and TSOIL perturbations

These steps are the additional components in the new operational 12-km NEPS as compared to the previous 33-km NEPS. Ensemble prediction near the surface is generally under-dispersive which results in overconfident forecasts of near-surface variables. One of the major reasons of underestimated ensemble spread near the surface is not accounting errors associated with the observations (Sætra et al., 2004). More practical reason behind underestimation of near-surface dispersion is that the identical lower-boundary initial conditions are used for all the ensemble members.

As a part of ocean-atmosphere interaction, generally SST has a strong impact on the forecast due to the large energy fluxes from the ocean into the atmosphere (Frankignoul, 1985). Atmospheric circulations are also sensitive towards soil moisture. A negative feedback mechanism between soil moisture and rainfall was noticed and documented by Dixon et. al. (2013). The SST and SMC perturbations in MOGREPS were added to better represent the uncertainties in the initial conditions at the surface. Tennant and Beare (2014) explained the methods of adding SST and SMC perturbations. The method used is briefly summarized below. Met Office Operational SST and Sea-Ice Analysis (OSTIA) data were used to generate statistics of daily mean SST state. Day-to-day SST changes were calculated over a period from 2006 to 2010. Power spectra of these daily difference fields were calculated and averaged over the four-year period as monthly means. The statistics obtained from the power spectrum analysis was used to set the amplitude of each spectral mode of a triangular spherical harmonic expansion. For each ensemble member, global SST anomaly pattern is generated using the Fast Fourier Transform. As these anomaly fields are having same power spectrum everywhere, those are multiplied by the monthly-mean of day-to-day SST change. This SST perturbation scheme targets to create larger perturbations (2-3^{0}C) at a region where day to day SST fluctuation is large and smaller perturbations (<0.5^{0}C) at a region under subtropical anticyclone.

Another scheme is added into the new 12-km NEPS to generate perturbations of SMC and TSOIL. UK Met Office started the soil moisture assimilation with simple nudging scheme which makes use of screen level analysis of temperature and humidity and later included surface soil wetness from the Advanced Scatterometer (ASCAT) on the Meteorological Operational (MetOp) satellite (Dharssi et al., 2011). In the 12-km NEPS both the SMC and TSOIL perturbations are generated using the vector breeding method. Short forecasts of 6 hours (previous cycle) from each ensemble member are used to extract the soil fields at each soil level in the model. It is the simpler way to develop perturbations which allows soil fields to evolve independently. Soil perturbations for each member is calculated by subtracting naturally evolved SMC and TSOIL from ensemble mean. After that some special checks for these differences (first guess perturbations) are done before applying those to the current model start file. These special checks (based on wilting, critical and saturation points) limit soil moisture perturbations to physically sensible bounds. Also, removal of SMC and TSOIL perturbations are done at points under snow and land ice fields. These checks are also to make sure that the soil perturbations sum up to zero in order to avoid systematic drift in the forecast. Separate sensitivity experiments are also planned to be carried out to analyse the impact of adding SST, SMC and TSOIL perturbations on near surface spread in NEPS.

3.6 Short Forecasts

The short forecasts are made daily in all the assimilation cycles (00, 06, 12 and 18 UTC) and the length of each short forecast in NEPS is 9 hours. Short forecasts of all the 22 perturbed members and the control member use the analysis obtained from the operational deterministic DA system and initial condition perturbations generated from the ETKF. The outputs from all the perturbed members of NEPS short forecast are used in operational Hybrid-4DVar DA system to make the background error covariance flow-dependent. Short forecast outputs are also used by the next cycle ETKF to generate analysis perturbations. These output files contain the forecast fields; wind, potential temperature, exner pressure and specific humidity. Short forecast background field file is used by the next cycle ensemble OPS to create 'varobs' and 'modelobs' files.

3.7 Long Forecasts

The model configuration and input requirements in long forecast are the same as those in short forecast except that it is integrated for 10.5 days. Long forecast runs start daily with initial conditions of 00 UTC and 12 UTC. Due to computational constraints, the 12-km NEPS long forecast uses only 11 perturbed ensemble members out of the available 22. We are considering first 11 members (Group1; 1-11) for long forecast starting from 00 UTC and the second 11 members (Group2; 12-22) for long forecast from 12 UTC. Probabilistic long forecast at 'T' hour are issued on the basis of 'T' hour forecast of Group1 and 'T+12' hours (started from 12 UTC of previous day) forecast of Group2 members. Finally, long forecasts of perturbed members from both the cycles and one control member from 00 UTC run are combined to form the 23 ensemble members long forecast which are used for our post-processing products. The deterministic model which is running at same resolution is used as the control member of NEPS. It is also to be noted that in the ETKF system each ensemble member is equally likely/skillfull apart from the control which is not perturbed.

4. Cold Start of NEPS

As a 12-km global EPS is being implemented for the first time in the world, ETKF perturbations were not available at this resolution from other centres (UKMO, BoM, etc.). Therefore, NEPS members were cold-started from the same initial condition. At the first short forecast cycle, all the 22 ensemble members are made to run from the same initial condition (which is the analysis of the deterministic model) but with perturbed model physics. The model physics is perturbed by Stochastic Kinetic Energy Backscatter (SKEB) and Random Parameter (RP) schemes (Tennant et al., 2011). Due to the perturbation in model physics this run produces 22 different model outputs and hence 22 forecast perturbations. In the next cycle ETKF applies the transformation matrix on these 22 forecast perturbations and multiplies with an inflation factor to generate 22 analysis perturbations. These analysis perturbations are added to the deterministic analysis by Incremental Analysis Update (IAU) method to generate 22 initial conditions for the next cycle. More details about model physics perturbation schemes in current version of NEPS are available at https://code.metoffice.gov.uk/doc/um/vn10.9/papers/umdp_081.pdf. The cold start of the short forecast runs was carried out in April 2018 and short forecasts from NEPS have been continuously running after that.

As a part of sensitivity experiments, we had earlier performed some spin-up experiments to check the number of ETKF cycles required to get realistic spread among the ensemble members if we cold-start the NEPS. These experiments were done using NEPS (N768) at a resolution of 17 km (for 11 and 22 members). Operational 33-km NEPS (N400) runs were used as a reference for comparing the ensemble spread. The results of these experiments indicated that after 15-16 ETKF cycles, 17-km NEPS (N768) ensemble spread becomes nearly equal to the operational 33-km NEPS (N400) ensemble spread (Figure 2). Ensemble spread is calculated using the analysis perturbations obtained from each ETKF cycle. Change in global average of spreads in specific humidity with time at near surface level (20 m) and at mid troposphere level (5600 m) are shown in Figures 2(a) and 2 (b) respectively.

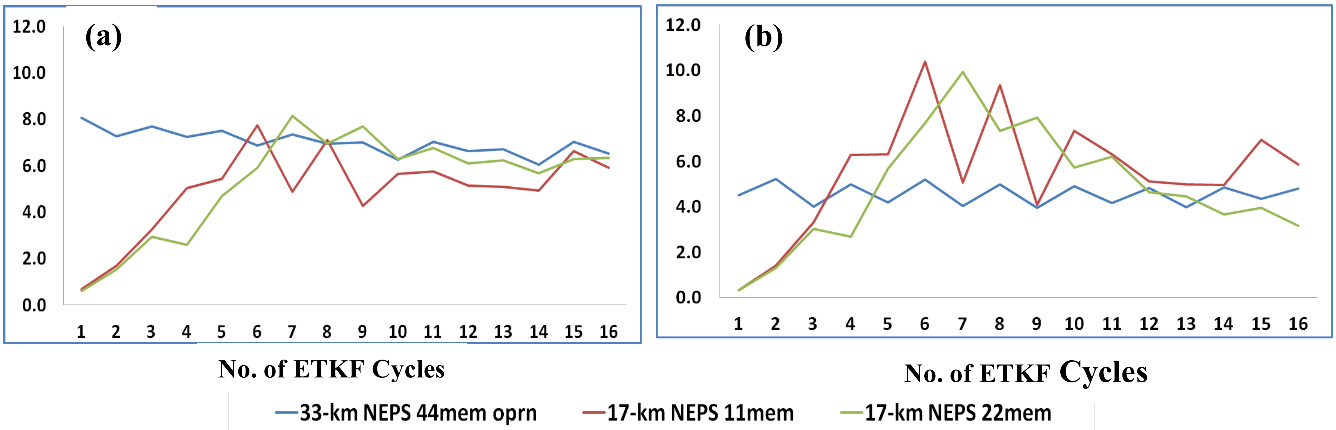

Figure 2: Ensemble spread in Specific Humidity (kg/kg) x10^{-4} at (a) 20 m and (b) 5600 m height

5. NEPS Forecast Products

The products of the long forecast runs are generated according to the need of the user community and ensemble forecast verification techniques. A detailed list of some of the long forecast products are given in Table 4. These forecast products are made available in NCMRWF website under 'Monsoon Region' (http://www.ncmrwf.gov.in/product_main_mihir.php) and 'Indian Region' (http://www.ncmrwf.gov.in/product_main_ind_mihir.php). Some of these products are discussed in detail in the previous NEPS report by Sarkar et.al., (2016). In the previous operational 33-km NEPS based products, sub-divisional rainfall probabilities were calculated for the ranges 2-6, 6-11, 11-20 and >20 cm/day. In the new 12-km NEPS these products are based on probability of rainfall exceedance. The threshold values are selected as 2.5, 15.6, 65.5, 115 and 195 mm/day. Location specific forecasts in the form of ensemble meteograms or

EPSgrams are now issued for 660 districts in the country as well as for some major cities of neighbouring BIMSTEC (Bay of Bengal Initiative for Multi-Sectoral Technical and Economic Cooperation) countries. Time series plots of temperature at 1.5 m, relative humidity at 1.5 m, wind at 10 m, rainfall and mean sea level pressure at specific locations are included in EPSgrams. Full range of the distribution of the values of different parameters predicted by the ensemble members are displayed through box and whisker plots. The maximum, minimum, 75 percentile, 25 percentile and median are represented by the boxes and whiskers. One of the recent heavy rainfall events was recorded on 26th July 2018 in Delhi/NCR. Figure 3 shows EPSgram over Central Delhi district based on initial conditions of 00 UTC 22nd July 2018.

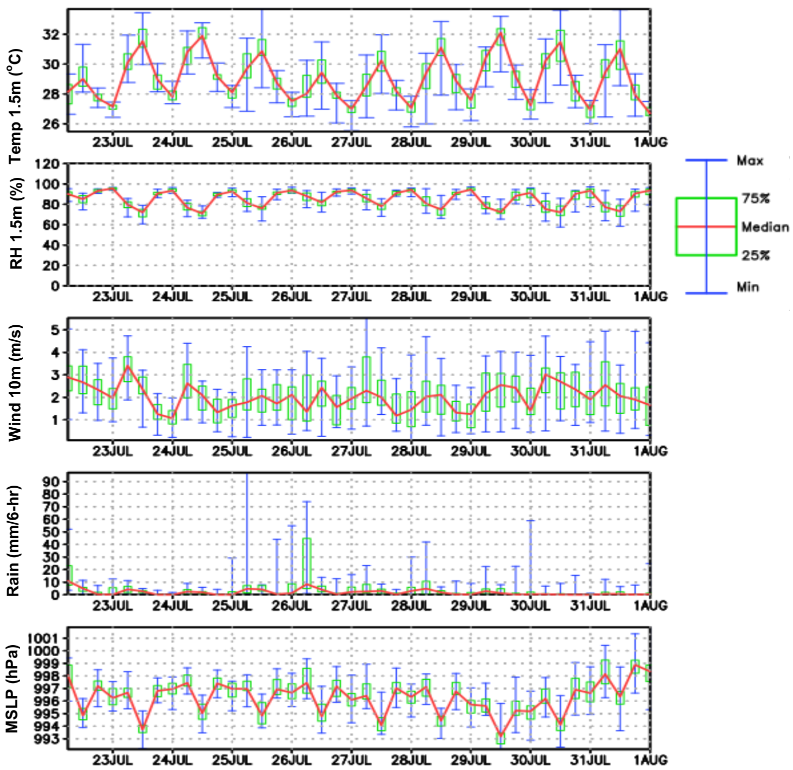

Figure 3: 10 days forecast EPSgram for Central Delhi District based on 00 UTC 22 July 2018 initial conditions, depicting temperature (°C) and Relative Humidity (%) at 1.5 m, wind speed (m/s) at 10 m, rainfall (mm/6hr) and MSLP (hPa).

NCMRWF is also contributing to ECMWF 'TIGGE - global ensemble forecast data' project since 1st Aug 2017. Earlier, datasets were provided from the 33-km NEPS as Version1 with 10 days forecast from 45 members (1 control + 44 perturbed) based on 00 UTC initial conditions. After the recent upgradation to 12-km NEPS, 10 days forecast outputs are provided to the TIGGE project for 12 members (1 control + 11 perturbed) at both 00 and 12 UTC respectively as Version2 from 1st July 2018.

Table 4: List of Products generated from the Operational 12-km NEPS

Products	**Variables Used**	**Levels**	**Brief Description**
Geo Potential Height	Geo-potential height	200, 500, 700, 850, 925 hPa	Daily instantaneous values of ensemble mean and spread at 00 UTC are provided.
Wind	u & v-wind	-do-	-do-
Mean Sea Level Pressure	MSLP	Mean sea level	-do-
Ensemble Stamp Plot	Accumulated precipitation, Streamlines and surface pressure	Rainfall at surface Winds at 850 hPa	Daily accumulated precipitation and instantaneous values of wind speed & direction at 00 UTC are provided for all the 22 perturbed members along with the control forecast
Temperature Probability	Max and Min temperatures in a day	1.5 m	Daily predicted probabilities of maximum and minimum temperatures exceeding some threshold values are generated. In summers, temperature probabilities greater than 40, 43 and 46^{o} C for maximum temperature and 23, 26 and 29^{o} C for minimum temperature are used. In winters threshold values are less than 15, 20 and 25^{o} C for maximum temperature and 5, 10 and 15^{o} C for minimum temperature.
Sub divisional Rainfall Probability	Accumulated precipitation	Surface	Daily predicted probabilities of occurrence of rainfall exceeding 2.5, 15.6, 65.5, 115 and 195 mm/day are issued.
Sub divisional Rainfall Departu	-do-	-do-	Daily rainfall departures are calculated with respect to 2007-2015 mean model climate values. Areas where rainfalls probabilities are more than 90th and 99th percentiles are also provided in these images
District wise EPSgrams	T at 1.5 m, RH at 1.5 m, Wind speed at10m , Accumulated precipitation and MSLP	1.5 m, 10 m, Mean sea level	6 hourly box-and-whisker's time series for the surface variables such as 1.5 m temperature, 1.5 m RH, 10 m wind speed, accumulated precipitation and MSLP are provided for 660 districts in the country as well as for some of the major cities in BIMSTEC countries.

6. Comparison of 12-km NEPS with 33-km NEPS

Increase in model resolution and ensemble size in practice improves the performance of EPS. However, in operational centers, an optimal configuration needs to be selected according to the available computational resources. The horizontal resolution of current version of NEPS (12 km) is around three times the previous version of NEPS (33 km), however the ensemble size has become half from 44 to 22. Here we are evaluating both the systems in order to ascertain the improvements in current operational 12-km NEPS.

6.1 Spread-Skill Relationship

Comparing ensemble spread with root mean square (RMS) error is a general practice while evaluating EPS (Johnson and Bowler, 2009; Palmer et al., 2006; Toth et al., 2003). RMS error of ensemble mean (RMSE) and ensemble spread (SPREAD) as a function of lead time are calculated for two-week period (10 June – 23 June 2018) for both the 12-km and 33-km NEPS. Area averaged geo-potential height at 500 hPa (Figure 4) and MSLP (Figure 5) over both northern and southern hemispheres are selected to verify the forecasts. RMSE is a measure of difference between forecast and analysis whereas SPREAD measures the deviation of ensemble members from ensemble mean. If all the uncertainties associated with the initial conditions and model errors are perfectly represented by the EPS, RMSE and SPREAD will be equal (Palmer et al., 2006). It means verifying analysis is statistically indistinguishable from the ensemble members in perfect EPS (Toth et al., 2003). In Figures 4 and 5, the SPREAD is closer to the RMSE in current operational 12-km NEPS compared to the previously operational 33-km NEPS in both the hemispheres. The results are similar for both 500 hPa geo-potential height (Figure 4) and MSLP (Figure 5). Major improvements are noticed over southern hemisphere as indicated in Figures 4c & 4d and Figures 5c & 5d. Another noticeable difference is a slight over dispersion in 12-km NEPS in Day 1 and a slight under dispersion in Day 2 forecasts. On the contrary there is always an over dispersion in 33-km NEPS.

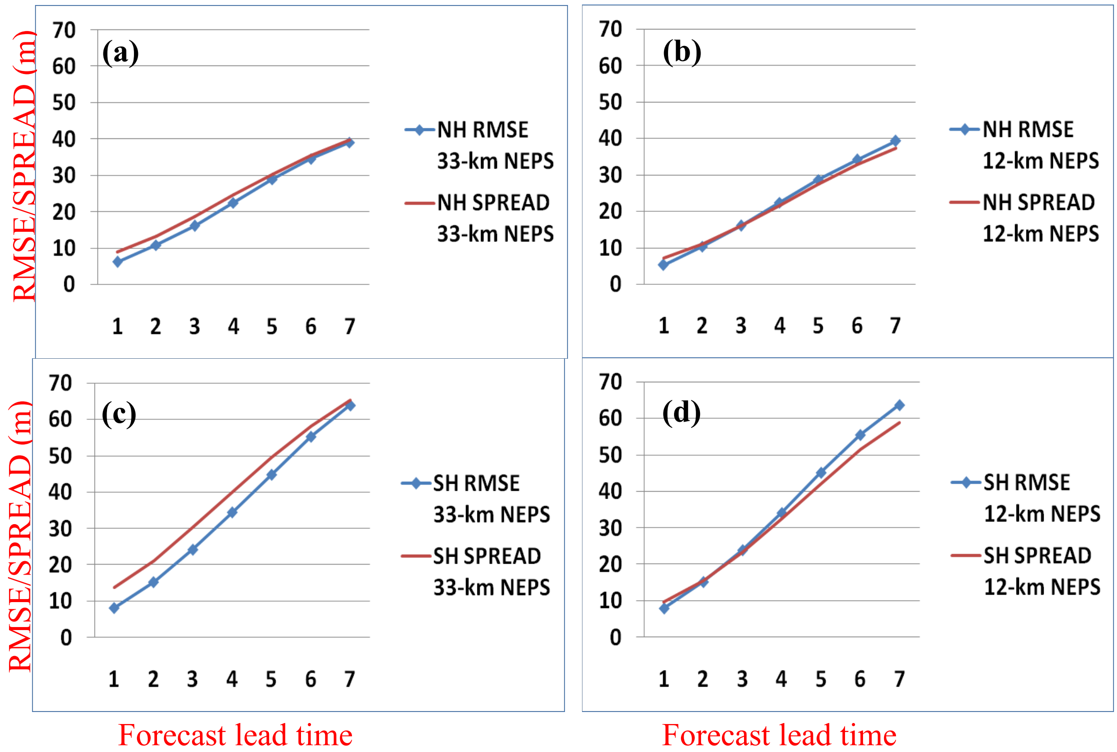

Figure 4: *RMSE and SPREAD of the (a) 33-km NEPS and (b) 12-km NEPS in Northern Hemisphere and (c) 33-km NEPS and (d) 12-km NEPS in Southern Hemisphere as a function of forecast lead time for 500 hPa Geo-potential height.*

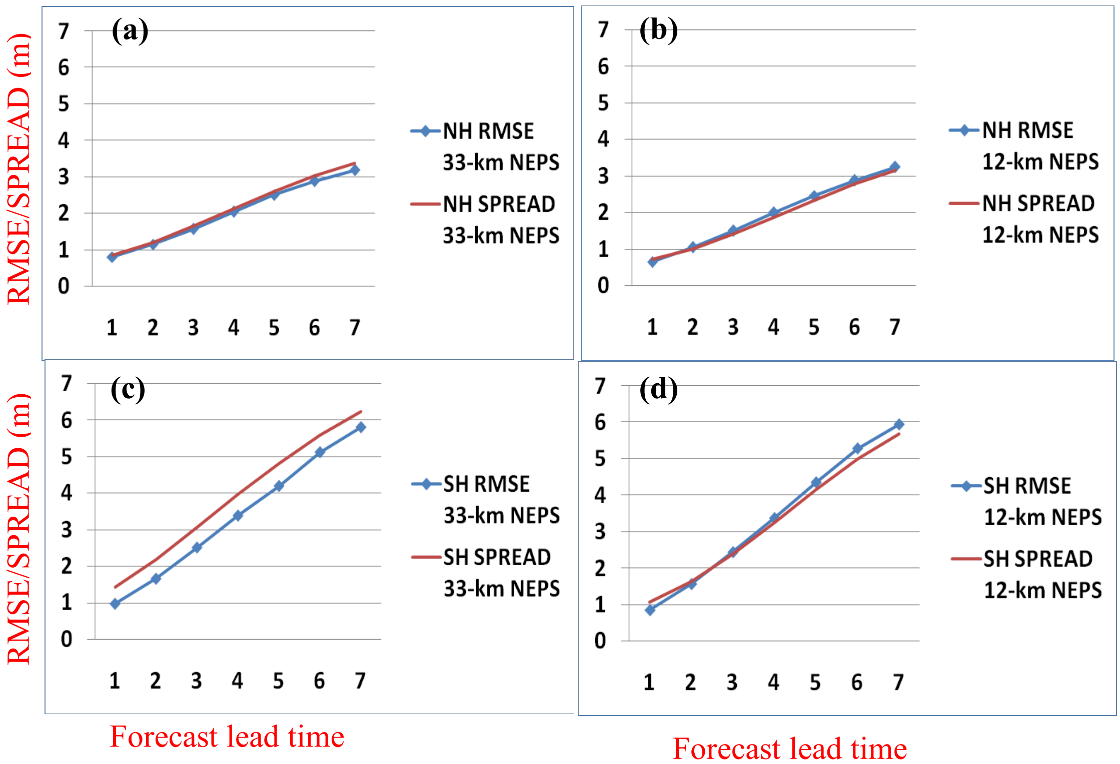

Figure 5: *RMSE and SPREAD of the (a) 33-km NEPS and (b) 12-km NEPS in Northern Hemisphere and (c) 33-km NEPS and (d) 12-km NEPS in Southern Hemisphere as a function of forecast lead time for MSLP.*

6.2 Ensemble Mean of a Heavy Rainfall Event

Day 5 forecasts of ensemble mean rainfall from both the operational 12-km NEPS and the previously operational 33-km NEPS have been compared with respect to the gridded satellite-gauge merged rainfall product at ~25 km resolution in Figure 6. These rainfall forecasts are based on 00 UTC 5th June 2018 initial conditions. Heavy rainfalls over many areas in West Coast were reported on 10th June 2018. Although both the systems have captured the locations of heavy rainfall, the 12-km NEPS has predicted rainfall amounts between 16-32cm/day. The locations of maximum rainfall predicted by both the systems are slightly shifted towards the east compared to the observations. In case of 12-km NEPS, the spatial spread of rainfall near the West Coast over the sea is less which agrees well with the observations. It indicates the better agreement among the members of 12-km NEPS in predicting heavy rainfall.

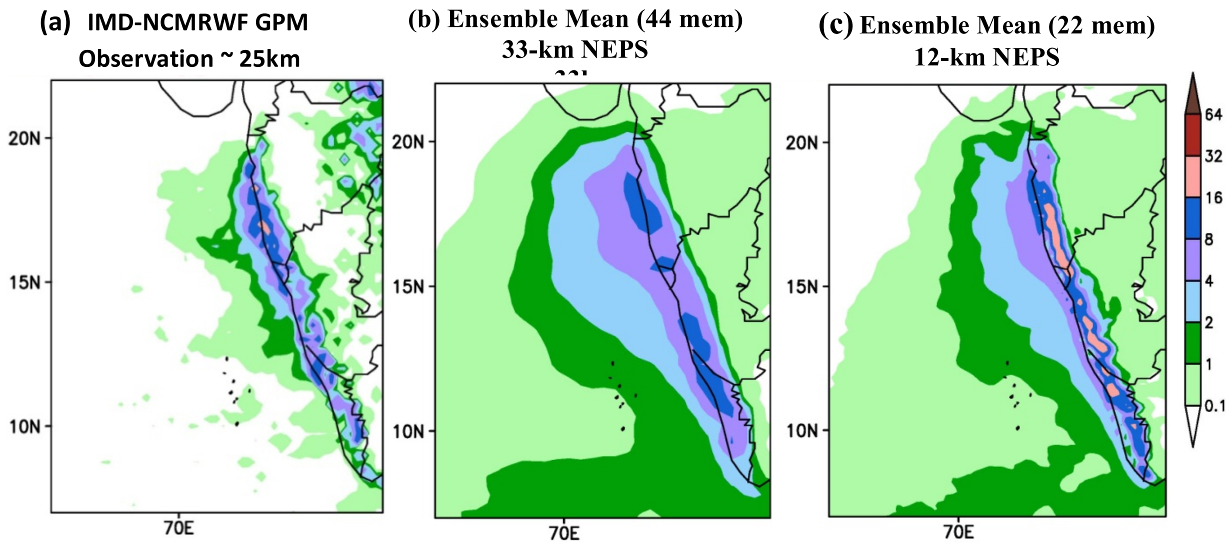

***Figure 6:** IMD-NCMRWF (GPM SAT+Gauge) Rainfall (cm/day) observations for 10th June 2018 are shown in (a). Ensemble mean Day 5 forecast valid for 10th June from (b) 33-km NEPS and (c) 12-km NEPS*

6.3 Probability of rainfall exceedance

Probabilistic quantitative precipitation forecast for day 3 is shown in Figures 7 and 8 from 33- km and 12-km NEPS, respectively. Probability of rainfall exceeding the threshold values 2.5, 15.6, 65.5, 115 and 195 mm/day predicted by both the systems are compared. The striking difference can be noticed in Figures 7 and 8 for threshold categories of 15.6 and 65.5 mm/day rainfall. In 12-km NEPS, the probability of getting rainfall more than 15.6 mm/day in Rajasthan and Madhya Pradesh regions is more than 75% in many districts whereas, in 33-km NEPS, probability is mainly lying between 25-50%, with only a few districts showing 50-75% probability. Additionally, according to 12-km NEPS forecasts there is some probability (25-50%) of getting rainfall more than 65.5 mm/day in some districts in the southern Rajasthan (Figure 8d). This feature is absent in the 33-km NEPS.

33-km NEPS

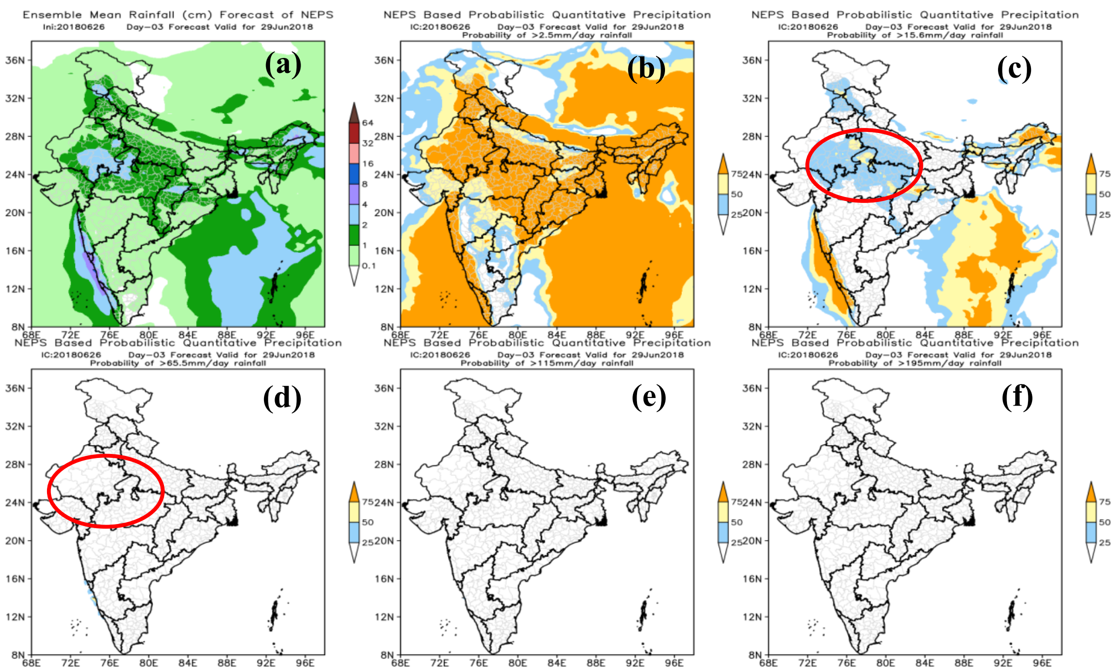

Figure 7: *Day 3 forecast of 33km NEPS probability of exceedances based on 00 UTC 26 June 2018 initial conditions have been shown. (a) ensemble mean, probability of getting rainfall greater than (b) 2.5, (c) 15.6, (d) 65.5, (e) 115, (f) 195 mm/day.*

12-km NEPS

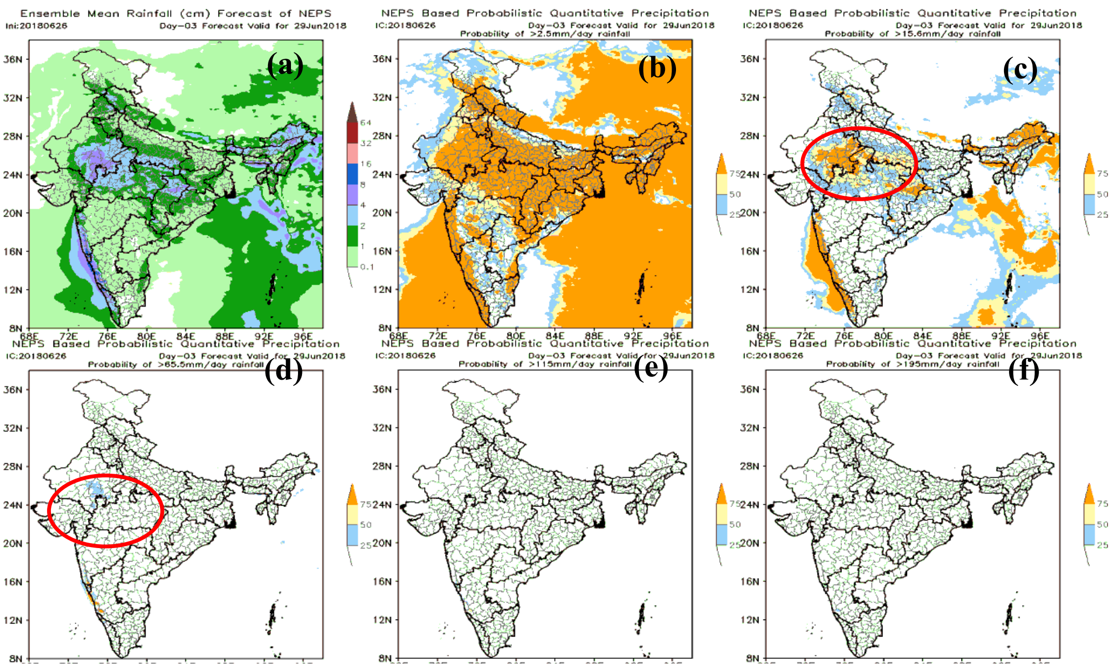

Figure 8: *Description is same as in Figure 7 but for 12-km NEPS*

7. Summary

Improvement in forecast accuracy of high impact weather events is the main priority of operational weather forecasting centres. It is well understood that weather forecasts are generally not perfect from deterministic model and hence require ensemble forecasting techniques to quantify uncertainties in the forecast. Ideally, ensemble prediction system should represent uncertainties in initial conditions as well as how those uncertainties evolve during the forecast. At the end of the forecast period the spread of the ensemble should represent the uncertainties in the forecast values. The current operational 12-km NEPS is based on UM10.8 version of latest UK Met Office weather forecast suite, PS40. The initial condition perturbations are generated by ETKF method. The model uncertainties are taken care by the SKEB and Random parameters schemes. In the current 12-km NEPS, surface perturbations (namely, sea-surface temperature, soil moisture content and soil temperature) are also provided. The control forecast run starts with analysis data of deterministic forecasting system and 22 ensemble members start from different perturbed initial conditions. The forecast perturbations are obtained from 6 hourly short forecasts of 22 members, which are updated by the ETKF four times a day (00, 06, 12 and 18 UTC).

NEPS operational products are based on 10 day long forecasts from 22 members. The products are generated everyday based on the combination of 11 members from 00 UTC and 11 members from 12 UTC (previous cycle) initial conditions. Some of the forecast products such as Geo-potential height, Winds, Mean Sea Level Pressure, Ensemble Stamp plots, Temperature Probability, Sub divisional Rainfall Probability, Sub divisional Rainfall Departure and District-wise EPSgrams are provided on daily basis through NCMRWF website. In the current operational 12-km NEPS, location specific forecasts in the form of ensemble meteogram or EPSgram are issued for 660 districts in the country and some major cities of neighbouring BIMSTEC countries. As compared to the previously operational 33-km NEPS, the 'Sub- divisional Rainfall Probability' product is now based on the probability of exceedance method. The threshold values selected are 2.5, 15.6, 65.5, 115 and 195 mm/day.

As compared to the 33-km NEPS, the 12-km NEPS shows improvements in prediction of high intensity rainfall areas. The SPREAD and RMSE relation is better in both the hemispheres for the 12-km NEPS. Also, the probabilities of predicting rainfall events are increased in 12-km NEPS. Further studies will be carried out by using more verification tools to verify skills of this high resolution 12-NEPS.

Acknowledgments

The authors are thankful to Secretary, MoES for his vision and constant encouragement during the 12-km NEPS implementation. We would also like to acknowledge technical support provided by Cray HPC Team especially to Ms. Shivali Gangwar, Mr. Puneet Singh and Mr Virender Kumar. The authors gratefully acknowledge use of PS40 global forecast suite made available under UM Partnership on Met Office Shared Repository Service (MOSRS).

References

Bowler N.E., Arribas A., Beare S. E., Mylne K.R., Shutts G. J. (2009): The local ETKF and SKEB: Upgrades to the MOGREPS short-range ensemble prediction system. *Quart. J. Roy. Meteor. Soc., 135:* 767–776

Bowler N.E., Arribas A., Mylne K.R., Robertson K.B., Beare S.E. (2008): The MOGREPS short-range ensemble prediction system. *Quart. J. Roy. Meteor. Soc., 134*: 703–722

Clayton A.M., Lorenc A.C., Barker D.M. (2013): Operational implementation of a hybrid ensemble/4D-Var global data assimilation system at the Met Office. *Quart. J. Roy. Meteor. Soc., 139*: 1445–1461

Dharssi I., Bovis K.J., Macpherson B., Jones C.P. (2011): Operational assimilation of ASCAT surface soil wetness at the Met Office, *Hydrol. Earth Syst. Sci., 15*: 2729-2746

Dixon N.S., Parker D.J., Taylor C.M., Garcia-Carreras L., Harris P.P., Marsham J.H., Polcher J., Woolley A. (2013): The effect of background wind on mesoscale circulations above variable soil moisture in the Sahel. *Quart. J. Roy. Meteor. Soc., 139*: 1009–1024

Frankignoul C. (1985): Sea surface temperature anomalies, planetary waves, and air–sea feedback in the middle latitudes. *Rev. Geophys. 23*: 357–390

George J.P., Rani S. I., Jayakumar A., Mohandas S., Mallick S., Rakhi R., Sreevathsa M.N.R., Rajagopal E.N. (2016): NCUM Data Assimilation System*, NMRF/TR/1/2016*, 20p.

Houtekamer P.L., Mitchell H.L. (1998): Data Assimilation Using an Ensemble Kalman Filter Technique, *Mon. Wea. Rev., 126*: 796-811

Johnson C., N. Bowler, (2009): On the reliability and calibration of ensemble forecasts. *Mon. Wea. Rev., 137*: 1717–1720.

Palmer T., Buizza R., Hagedorn R., Lawrence A., Leutbecher M., Smith L. (2006): Ensemble prediction:A pedagogical perspective, *ECMWF Newsletter, No. 106, ECMWF, Reading, United Kingdom*, 10–17

Sætra Ø., Hersbach H., Bidlot J-R., Richardson D.S. (2004): Effects of observation errors on the statistics for ensemble spread and reliability. *Mon. Wea. Rev., 132*: 1487–1501

Sarkar A., Chakraborty P., John P.G., Rajagopal E.N. (2016): Implementation of Unified Model based Ensemble Prediction System at NCMRWF (NEPS), *NMRF/TR/02/2016*, 26p

Tennant W., Beare S. (2014): New schemes to perturb sea-surface temperature and soil moisture content in MOGREPS. *Quart. J. Roy. Meteor. Soc., 140*: 1150–1160

Tennant W.J., Shutts G.J., Arribas A., Thompson S.A. (2011): Using a Stochastic Kinetic Energy Backscatter Scheme to Improve MOGREPS Probabilistic Forecast Skill. *Mon. Wea. Rev., 139*: 1190-1206

Toth Z., Talagrand O., Candille G., Zhu Y. (2003): "Probability and Ensemble Forecasts." In book of: Forecast Verification: *A practitioner's guide in atmospheric science. Ed.: I. T. Jolliffe and D. B. Stephenson. Wiley*: 137–163