Assimilation of the rain-gauges measurements using Particle Filter

Prashant Kumar[1]

[1]Indian Space Research Organisation

November 24, 2022

Abstract

The well recognized constraint of non-linear and non-Gaussian distribution of rainfall observation limits its assimilation in the high-dimensional numerical weather prediction (NWP) model. In this study, rain-gauges' observed rainfall from Indian Meteorological Department (IMD) over Indian landmass is assimilated in the Weather Research and Forecasting (WRF) model using particle filter. In the framework of imperfect weather model, particles (or ensembles) for rainfall predictions are created with various combinations of model physics (viz. cumulus parameterization, micro-physics, planetary boundary layer schemes). With the help of IMD observed rainfall, weights are provided to various particles using multiple hypotheses, and this is the step in which IMD rainfall observations are used for assimilation. Further, a resampling step is performed to generate new particle from high weight particle using stochastic kinetic-energy backscatter scheme (SKEBS) method in which dynamical variables are perturbed into the model physics. Results based on rainfall verification scores suggest that assimilation of the rain-gauges observed rainfall using particle filter improved prediction of rainfall over CNT runs (unweighted particle; without assimilation). Moreover, surface and vertical profile of temperature, water vapour mixing ratio (WVMR) and wind speed are also improved in 24 h forecasts.

Assimilation of the rain-gauges measurements using Particle Filter

Prashant Kumar

Atmospheric Sciences Division, Atmospheric and Oceanic Sciences Group, Space Applications Centre, ISRO, Ahmedabad, India

Abstract

The well recognized constraint of non-linear and non-Gaussian distribution of rainfall observation limits its assimilation in the high-dimensional numerical weather prediction (NWP) model. In this study, rain-gauges' observed rainfall from Indian Meteorological Department (IMD) over Indian landmass is assimilated in the Weather Research and Forecasting (WRF) model using particle filter. In the framework of imperfect weather model, particles (or ensembles) for rainfall predictions are created with various combinations of model physics (viz. cumulus parameterization, micro-physics, planetary boundary layer schemes). With the help of IMD observed rainfall, weights are provided to various particles using multiple hypotheses, and this is the step in which IMD rainfall observations are used for assimilation. Further, a resampling step is performed to generate new particle from high weight particle using stochastic kinetic-energy backscatter scheme (SKEBS) method in which dynamical variables are perturbed into the model physics. Results based on rainfall verification scores suggest that assimilation of the rain-gauges observed rainfall using particle filter improved prediction of rainfall over CNT runs (unweighted particle; without assimilation). Moreover, surface and vertical profile of temperature, water vapour mixing ratio (WVMR) and wind speed are also improved in 24 h forecasts.

Keywords: Rainfall; Particle filter; WRF model; SKEBS; Forecast.

1. Introduction

Accurate rainfall prediction from the numerical weather prediction (NWP) model is one of the most challenging concerns for weather modelling society. The rainfall assimilation has large impact on weather forecast mainly over tropics, in which moist convection plays a prominent role, and links directly or indirectly to humidity, cloud cover, latent heating, and the divergent component of the large-scale circulation (Marecal and Mahfouf, 2003; Hou.et.al, 2004). Various efforts have been attempted to assimilate rainfall information in the NWP model using nudging method, variational assimilation and Kalman filter in last decades (Kumar et al., 2014; Kumar and Varma, 2016; Kumar and Kishtawal, 2017 and references therein). It is well studied that assimilation of satellite derived rainfall in the NWP model helps to improve the analyses and subsequent short range weather forecasts (Donner, 1988; Krishnamurti et al., 1984, 1991, 1993; Heckley et al., 1990; Puri and Miller, 1990; Mathur et al., 1992; Kasahara et al., 1994; Manobianco et al., 1994; Van Tuyl, 1996; Peng and Chang, 1996; Treadon, 1996; Tsuyuki, 1997; Benedetti et al., 2005; Bauer et al., 2011; Lopez, 2011; Kumar and Varma, 2016; Kumar et al., 2014; Kumar and Kishtawal, 2017; Lien et al., 2013, 2016a,b). Kumar and Varma (2016) found that assimilation of satellite retrieved rainfall in the NWP model improved the forecast for unprecedented heavy rainfall, which is not able to predict from operational centres. Moreover, Kumar and Kishtawal (2017) showed that variational assimilation of both rain and no-rain information from satellite has positive impact on short range weather forecasts.

Errico et al. (2007) suggested that rainfall assimilation is more complex problem compared to assimilation of conventional or clear-sky satellite radiance. Due to non-linear and non-Gaussian characteristics of rainfall, still assimilation of rainfall in the NWP model is a challenging problem. Few studies (Bauer et al., 2011; Kumar and Varma, 2016; Kotsuki et al., 2017) are mentioned common difficulties in the rainfall assimilation mainly due to the strong nonlinearity of moist process and non-Gaussian characteristics of precipitation. Most of the previous studies based on variational method and ensemble Kalman filter (EnKF) assume/convert non-Gaussian distribution of rainfall to Gaussian error statistics which lead to suboptimal analysis (e.g., Van Leeuwen 2009, 2010; Posselt and Bishop 2012; Posselt et al. 2014). One well-known weakness of EnKF is that it commences the prediction and filtering probability distribution functions (PDF) to be Gaussian (Mattern et al., 2013; Ratheesh et al., 2016; Kumar and Shukla, 2019). Kumar et al. (2014) discussed importance of quality control on rainfall assimilation, and showed that with strict quality control generally difficult to improve forecasts beyond a few hours due to the non-Gaussian nature of rainfall data. Also, due to limitations of the realistic representation of the non-linear model physics as tangent linear model, numerical models are not able to assimilate rainfall precisely using 4D-Var.

The objective of this study is to assimilate the Indian Meteorological Department (IMD)'s rain-gauge observed rainfall in the Weather Research and Forecasting (WRF) model using particle filter, a non-linear filter which take care of non-Gaussian nature of rainfall observations. Details of the particle filter and design of experiment are given in section 2, and section 3 is discussing about data used. Results and discussions are provided in section 4, and are concluded in last section.

2. Particle Filter

Particle filter can be used in various fields of science e.g., meteorology, oceanography, etc. in which one wants to estimate the best state of a system from an imperfect model with noisy and inadequate data (Chorin et al., 2013). This method is computationally costly compared to traditional methods like optimal interpolation (OI), 3D/4D-Var and Ensemble Kalman filters (EnKF). These traditional methods are highly depending on accuracy of the first guess and accurate representation of the error covariances (Van Leeuwen, 2009). In particle filter, tangent linear and adjoint model of the non-linear model are not requisite, which have a large uncertainty over tropical region. Moreover, estimation of background error covariance matrices is also not needed which also contributed large errors in traditional data assimilation methods. However, the issue of flow dependent background error is resolved with some extent in EnKF method. Artificial tricks like covariance inflation and localisation are needed in EnKF to get good results in high dimensional systems (Van Leeuwen, 2009). Particle filter do not have these difficulties, and particles (or ensembles) are not adjusted which do not destroy the dynamical balances in the analysis. The issue in the particle filter implementation is that the particles are not modified, so that after a few analysis steps, only one particle has all the weights against observations, and all other particles are move away from the observations. It means that the statistical information in the ensemble becomes too low to be meaningful and known as filter degeneracy (Ades and van Leeuwen, 2015). Details of the mathematical formulation of particle filters are discussed in previous studies (Doucet et al. 2001; Maskell and Gordon 2001; chen 2003; Ristic et al. 2004; Van Leeuwen 2009, 2010; Ades and van Leeuwen 2015; Kumar and Shukla, 2019 and references therein).

In present study, initially rainfall forecasts are simulated from the WRF model with total 90 different combinations of model physics viz. cumulus physics, micro-physics and planetary boundary layer (PBL) schemes (Figure 1). Nine different cumulus parameterization schemes available in the WRF model are selected named as new Kain-Fritsch (CP1), Betts-Miller-Janjic (CP2), Grell-Freitas (CP3), Simplied Arakawa Schubert (CP4), Grell-3D ensemble (CP5), Tiedtke (CP6), New SAS (CP14), Grell-Devenyi (CP93), and old Kain-Fritsch (CP99) schemes. Five different microphysics schemes available in the WRF model named as Lin (Purdue; MP2), WSM3 (WRF Single Moment; MP3), Eta (Ferrier; MP5), WSM6 (MP6) and Thompson (MP8) are used to generate particles. These micro-physics are selected based on complexity from simple WSM3 scheme which consider only rain and cloud as hydrometer to WSM6 or Thompson scheme in which other hydrometers e.g., ice, snow, graupel are also considered. The YSU (PBL1) and MYNN2 (PBL5) two different planetary boundary layer schemes are selected to generate particles. Details of these physics schemes are given in the WRF user guide (Skamarock et al., 2008). Each cumulus physics (9), micro-physics (5) and PBL (2) schemes are used to generate 10, 18 and 45 particles, respectively (Figure 1) on 01 August 2015. These choices of physics options are considering to perform reference experiments (unweighted particles or CNT runs). This is the forecast step of the algorithm.

In general, the particle filter considers a PDF of a state, and the PDF is approximating by particles consisting of large number of discrete samples (here choice of physics options) to represent and approximate posteriori by a weighted sample. The selected particles based on different physics options (*p*) represent a sample from its priori PDF, and are assumed to be of the form

$$x_{p,k} = f_{k-1}^{p}\left(x_{k-1}, v_{k-1}\right) \text{ for } k > 0 \tag{1}$$

Here, $x_{p,k}$ is the set of state vector with p different physics options to be estimated at time step k, and f_{k-1}^{p} is a known imperfect non-linear model (here WRF) with p different physics options, x_{k-1} is the best state taken from global model analysis having noise of v_{k-1} at time step $k-1$. The idea is to represent the prior pdf by a set of particles $x_{p,k}$, which are delta functions centred around state vectors. If one represents the prior pdf by a number of particles, like in the Ensemble Kalman Filter, so

$$p(x) = \sum_{p=1}^{N} \delta\left(x - x_{p,k}\right) \qquad (2)$$

Where, N is number of particles (which are 90 here). Then, from Bayes theorem

$$p(x/y) = \sum_{p=1}^{N} w_p \ \delta\left(x - x_{p,k}\right) \qquad (3)$$

in which weights w_p are given by

$$w_p = \frac{p\left(y/x_p\right)}{\sum_{q=1}^{N} p\left(y/x_q\right)} \qquad (4)$$

The density $p\left(y/x_p\right)$ is the probability density of the observations (y) given the model state x_p in forecast time step at which IMD observed rainfall is available over Indian landmass grids.

In the analysis step, observed rainfall from IMD are used to determine weight $\left(w_p\right)$ for each particles. This step involves weighting to each particle and subsequent weight-based resampling. The weights for each particle depend on IMD rainfall observation, and this is the step in which rainfall observation use for assimilation process. Assimilation in particle filtering amounts to sequential importance resampling (SIR) weighting of particles. The weights, in principle, should be proportional to likelihoods or conditional probabilities (Ratheesh et al., 2016). In this study, the likelihood is depending on multiple hypotheses (Dubuisson, 2015). The first hypothesis is the value of variance in rainfall forecast should be less, and the next hypothesis is the value of mean equitable threat score (ETS; section 4) should be high for different rainfall thresholds (like 5 mm/24 h, 10 mm/24 h, etc.). In first hypothesis, likelihood is inversely related to a suitable distance between model simulated rainfall and IMD observed rainfall. The distance is taken to be usual variance between simulations and observations for daily accumulated rainfall at observation grids. First, the distances d_p is computed between the accumulated rainfall from p particle at time k and IMD observed rainfall valid for same time period, with p varying from 1 to N, which is the total number of particles. Now, we calculate the raw weights as inverses of these distances. Intermediate weights are calculated by dividing the raw weights by the maximum weight and by raising the ratio to a power a, which is an adjustable parameter. The intermediate weights are then normalized so that their sum is unity. Finally, a median filter is used to select first 45 particles having higher weights. Next 45 particles are selected for which likelihood is directly related to a suitable mean ETS between models simulated rainfall and IMD observed rainfall for different rainfall thresholds. In this way, total numbers of particle (here 90 particles) are same for all time steps, whereas, few combination of model physics receive high weights (duplicate). In this process, observation uncertainty is not considering which may be a scope for future. Particles having large variance and less mean ETS values are rejected which contribute very little to the approximation of the target PDF.

Due to imperfect model physics (no scheme is perfect), diverse combination of model physics are matches well with IMD observed rainfall at different time steps, whereas, few combinations of model physics are completely rejected which are not performed well over tropical regions in this study (Figure 1). Moreover, rainfall prediction from the NWP model is a complicated process based on different meteorological parameters (e.g., temperature, moisture, winds, fluxes, cloud, etc.), surface characteristics (e.g., vegetation, roughness, albedo, land type, etc.) and model physics (e.g., cumulus, micro-physics, etc.). It is important to note that dimension of the NWP model is very high ($\sim 10^9$). So, a sub-space (here rainfall; dimension of $\sim 10^5$) from high-dimensional model space is used for particle filter implementation, and changes in sub-space modifies full state of the high-dimensional model.

Finally, in the resampling step, particles having higher weights are resampled at the observation time, whose distribution forms a weak approximation of the target PDF. In this step, new particles are generated from large weight particles (selected physics option) using stochastic kinetic-energy backscatter scheme (SKEBS; Berner et al., 2009, 2011). The advantage of SKEBS scheme is that it perturbs the dynamic state directly, and the perturb dynamical variables are then fed into the physical parameterizations (model physics). The SKEBS scheme is very different from perturbing the physical tendencies directly which can introduce inconsistencies between the physics and dynamics. So, the tendency of the model might be to readjust any such inconsistencies, possibly leading to erroneous phenomena (e.g., spurious gravity waves) (Berner et al., 2011). In this way, the total numbers of particles are same again at observation time step. The idea is to focus the particles towards high probability regions of the target PDF, so that the number of particles required for a good approximation of target PDF remains manageable within sub-space having very less dimension as compared to actual model space.

2.1. Design of Experiment

In this study, two different set of experiments are performed with (EXP; weighted particles) and without (CNT; unweighted particles) assimilation of IMD observed rainfall using particle filter during August 1-9, 2015. For the comparison purpose, accumulated rainfall forecast predicted from the WRF model are interpolated to IMD observation grids using bilinear interpolation for the same time period (here 24 hours). The National Centers for Environmental Prediction (NCEP) Global Data Assimilation System (GDAS) global model analysis is used to prepare initial and lateral boundary condition for the WRF model. The WRF model simulations are performed at 25 km spatial resolution using different physics options and SKEBS perturbations. The details of the selected WRF model configurations are given in Kumar and Varma (2016). The NCEP GDAS analysis is the final analysis from NCEP which assimilated various kinds of observations available from ground and satellites, and also including late arriving observations. Here, the objective is to estimate target PDF from an imperfect model with different model physics and dynamic variable perturbation in the physical parameterization using the best initial state (assume here) available from the NCEP analysis. The selection of different model physics after rainfall assimilation during 1-9 August 2015 are shown in Figure 1, which shows that few model physics which are not very appropriate over this part of world are rejected, and provide the higher weight to model physics which are more suitable over this region. To avoid rapid filter degeneracy where it approaches to single model physics, a dynamic variable perturbation in model physics is included using SKEBS method. The choice of the NCEP GDAS analysis as input in CNT runs is mainly to assess the superiority of rainfall assimilation using particle filter.

3. Data used

In this study, the IMD observed rainfall, the TRMM (Tropical Rainfall Measuring Mission) 3B42 merged rainfall product, and the NCEP GDAS global model analysis are used in different stages. The NCEP GDAS global model analysis is used to create initial and lateral boundary conditions for the WRF model during 1-10 August 2015. Further, forecasts of surface and vertical profile of temperature, moisture and wind speed are compared with the NCEP GDAS analysis on 10 August 2015. The IMD observed rainfall is used majorly for assimilation by particle filter (EXP runs) during 1-9 August 2015, and TRMM 3B42 merged rainfall product are used to assess the skill of rainfall prediction on 10 August 2015. Details of these datasets are given below:

3.1. IMD Rainfall

Daily gridded rainfall data over the Indian landmass is available since January 1901 at a spatial resolution of 0.25° latitude/longitude. This data set is prepared from daily recorded information from about 7000 SRG (Surface Rain Gauge) stations well-spread across the country after incorporating the necessary quality control (Pai et al., 2014). The quality control test involves verification of the location information of the gauge station, eliminating the missing data, eliminating the coding errors, extreme value check, etc. These data are interpolated using Shepard interpolation method into a regular grid (Pai et al., 2014). The distribution of gauges over India is satisfactory in terms of number and regional distribution, except some small regions of Jammu and Kashmir (J&K) and extreme northwest parts of India.

3.2. NCEP GDAS analysis

The NCEP implemented operationally a series of numerical models for the generation of global model analyses and forecasts. One of the operational system is GDAS (Kanamitsu, 1989), which uses the spectral Medium Range Forecast (MRF) model. The GDAS analysis is the final run in a series of the NCEP operational model; therefore, it is also known as the **Final Run** at the NCEP which also includes the late arriving conventional and satellite observations. It is run four times a day, i.e., at 0000, 0600, 1200, and 1800 UTC. Model output at analysis time and a 6 hours forecast are available from the National Oceanic and Atmospheric Administration (NOAA) National Operational Model Archive & Distribution System (NOMADS; http://nomads.ncdc.noaa.gov/) server. After post-processing of the NCEP GDAS, data from spectral coefficient form converts to 1° latitude-longitude (360 by 181) grids, and from sigma levels to mandatory pressure levels. It uses three-dimensional variational data assimilation (3D-Var) method for data assimilation. Details of the GDAS analysis are described by Kalnay et al. (1996).

3.3. TRMM 3B42 Rainfall

The TRMM is a joint US-Japan satellite mission to monitor tropical and subtropical precipitation. It was launched in late November 1997 in to a near circular orbit approximately at 350 km altitude (raised to 403 km since 2001) at 35° inclinations from the equatorial plane. The complete description of sensor package of TRMM is given by Kummerow et al. (1998). The operational TRMM dataset used in the present study is TRMM 3B42, which is a merged product from Geostationary InfraRed (IR) and Microwave data (Huffman et al., 2003, 2007). The TRMM 3B42 estimates are produced in four stages; (1) the microwave precipitation estimates are calibrated and combined, (2) infrared precipitation estimates are created using the calibrated microwave precipitation, (3) the microwave and IR estimates are combined, and (4) rescaling to monthly data is applied. This rainfall product has been downloaded from TRMM Online Visualization and Archive System (TOVAS) at spatial resolution of 0.25° latitude/longitude.

4. Results and Discussions

The two different set of experiments are performed in this study during 1-10 August 2015. The collection of unweighted particles is considered as “CNT runs”, and collection of particles with sequential importance resampling is considered as “EXP runs”, in which IMD observed rainfall is used to select appropriate model physics and resample high weight particles using dynamic variable perturbations using SKEBS method. These particle filtering steps are performed during 1-9 August 2015, and selected model physics options on 9 August 2015 are used for forecast verification on 10 August 2015. The choice of the NCEP GDAS analysis as input is mainly to assess the superiority of rainfall assimilation using particle filter over CNT runs, where initial conditions are taken from the NCEP GDAS analysis which assimilated various kind of observations and one of the final analysis available from the NCEP. In verification step, the WRF model predicted accumulated rainfall, initialized from selected model physics are validate against accumulated rainfall from TRMM 3B42 rainfall valid for same time. The surface and vertical profile of temperature, moisture and wind speed forecasts are verified against NCEP final analysis.

The mean difference (Bias), RMSD, and rainfall verification scores are used as standard parameter for statistical evaluation. Various rainfall verification scores based on contingency table (Table 1) viz. ETS, extremal dependency score (EDS), probability of detection (POD), and false alarm rate (FAR) over a wide range of rainfall thresholds (1 mm/day to 80 mm/day) are used to measure the impact of rainfall assimilation on rainfall predictions for grid wise evaluation. The POD measures the fraction of observed events that were correctly diagnosed, and is sometimes called the "hit rate". The FAR gives the fraction of diagnosed events that were actually non-events. Perfect values for these scores are POD=1, and FAR=0. The ETS was formulated to account for the hits that would occur purely due to random chance. The ETS, though not a true skill score, is often interpreted that way since it has a value of 1 for perfect correspondence, and 0 for no skill. It penalizes misses and false alarms equally, and for this reason it is commonly used in the NWP rainfall verification. The new score EDS is used mainly for determining skill at higher value of rainfall thresholds. This score has advantage that it can converge to different values for different forecasting systems and furthermore, it does not explicitly depend upon the bias of the forecasting system (Stephenson et al. 2008).

Table 1: *Schematic 2 x 2 contingency table for the definition of scores. (a+b+c+d=n)*

	Observation ≥ Threshold	Observation < Threshold
Forecast ≥ Threshold	*a*=Hits	*b*=False alarms
Forecast < Threshold	*c*= Misses	*d*=Correct rejections

$$ETS = \frac{(a-e)}{(a+b+c-e)} \quad (5) \qquad EDS = 2 * \frac{log\frac{(a+c)}{n}}{log\frac{a}{n}} - 1 \quad (6)$$

$$POD = \frac{a}{a+c} \quad (7) \qquad FAR = \frac{b}{(a+b)} \quad (8)$$

where $e = \frac{(a+b)*(a+c)}{(a+b+c+d)}$ refers to the expected number of correct forecasts above a rain threshold with a random forecast.

Figure 2 shows mean (line) and median (dash line) value of POD (eq. 7) and FAR (eq. 8) for **CNT** (blue) and **EXP** (red) runs. The POD for CNT runs are shown as light blue lines valid on 10 August 2015. The POD for EXP runs which assimilate IMD observed rainfall using particle filter are shown as light red lines. Figure 2a shows that slightly more mean POD is found in EXP compared to CNT for less rainfall threshold. This positive impact of rainfall assimilation is more for high rainfall threshold (> 35 mm/day). It suggests that rainfall assimilation using particle filter improve skill of rainfall forecast for heavy rainfall. It is also important to mention here that slightly less value of median POD is found for high rainfall threshold (> 40 mm/day), which indicate that forecasts from CNT show reduce skill for heavy rainfall. Moreover, median value of POD for EXP runs is slightly higher than mean POD for high rainfall threshold. It is also found that different model physics are predicting similar values of low rainfall (less spread for low rainfall threshold), whereas, this distribution is more for high rainfall threshold. Moreover, most of the particles are predicting better POD values after rainfall assimilation (light red lines) compared to CNT run (light blue lines). Similar to POD, mean and median value of FAR also show less number of false alarm in rainfall assimilation experiments (EXP) compared to CNT experiments. Better FAR score is seen for high rainfall threshold compared to low rainfall threshold which shows that all particles are able to predict low rainfall precisely, and large uncertainty are seen for heavy rainfall. Overall, these results show that assimilation of IMD observed rainfall using particle filter improve rainfall predictions for higher rainfall values. The accurate prediction of heavy rainfall has large societal benefits which obtained with the help of particle filter over CNT runs.

The ETS (eq. 5) is one of the most widely used skill score for rainfall verification, and EDS (eq. 6) score is normally used for high rainfall threshold (Stephenson et al. 2008). Figure 3 shows mean and median value of ETS and EDS rainfall verification score for CNT and EXP runs. Figure 3a shows that skill of rainfall prediction is improved for low and high rainfall threshold after rainfall assimilation. It is also seen that ETS predicted from the WRF model is ~0.35 for low rainfall threshold which represent a very high skill of prediction. This high skill score is mainly due to initialization of the WRF model from the NCEP final analysis (best state) which is generally not a situation in operational weather forecasts. Moreover, it is noticed from figure 3a that model predictions have more uncertainty for high rainfall values, and after rainfall assimilation skill of the rainfall forecasts are improved for high rainfall threshold (>40 mm/day). Similar to ETS, value of EDS rainfall score is also improved after rainfall assimilation. Around 0.2 value of EDS is found for >40 mm/day rainfall threshold, whereas, value of ETS is less than 0.1 for same rainfall threshold. It is important to note here that these improvements in rainfall prediction are over CNT runs which are performed using NCEP GDAS analysis as initial condition. Large number of satellite and conventional observations are assimilated in this analysis. So, observed advances in EXP runs after rainfall assimilation have noteworthy improvement over CNT runs.

Results discussed above support that assimilation of IMD observed rainfall using particle filter improved rainfall forecasts compared to CNT runs. The further interest is to evaluate the impact of assimilation in sub-space (here rainfall) on the prediction of other model sub-spaces (like temperature, moisture, winds) due to non-linear coupling of rainfall with these meteorological parameters. The WRF model predicted temperature, moisture and wind speed are verified against the NCEP GDAS analysis valid at same time. Figure 4 shows RMSD in 2-meter air temperature and water vapour mixing ratio (WVMR), and 10-meter wind speed for CNT and EXP runs. Figure 4a shows that assimilation of rainfall improved temperature forecasts for all forecast lengths (at 3-hour interval on 10 August 2015), except 9 hours' forecasts (a local maximum temperature is occurred at this time). Moreover, some particles having large RMSD in CNT runs are rejected in EXP runs after rainfall assimilation. Similar kind of positive impact can be seen in WVMR (Figure 4b) and wind speed (Figure 4c) forecasts. It is important to mention here that assimilation of IMD observed rainfall improved other basic meteorological parameters. These findings are similar to variational method in which due to multi-variate nature of data assimilation, assimilation of particular control parameter also modifies other control parameters (Kumar et al., 2014). Overall, we found that rainfall assimilation using particle filter improved surface temperature, WVMR, and wind speed forecasts. Further, we want to focus that these improvements are over CNT runs where the WRF model is initialized from the NCEP final analysis which assimilated all kind of observations including late arriving observations to prepare final analysis.

Further, vertical profiles of 24 h temperature, WVMR and wind speed forecasts valid on 10 August 2015 are also verified against NCEP GDAS analysis valid at same time (Figure 5). Results suggest that assimilation of IMD rainfall in EXP runs improve temperature profile

(Figure 5a) at different vertical levels compared to CNT runs. Slightly higher positive impact can be seen at upper levels (above 300 hPa) in EXP runs. Mixed impact is found in WVMR profile from surface to 900 hPa (Figure 5b), and depicts improve prediction of WVMR above 900 hPa in EXP runs. Assimilation of IMD rainfall also improves vertical profile of wind speed with maximum improvements at mid and upper vertical levels.

Overall, these preliminary results suggest that assimilation of rainfall using particle filter improve prediction of basic meteorological parameters (like temperature, moisture, and winds) at surface and vertically. These improvements in basic meteorological parameters are mainly due to rainfall which indirectly coupled with these basic parameters. Generally, in most of the previous rainfall assimilation studies (Kumar and Varma, 2016 and references therein), major objectives are to improve initial model states (like temperature, moisture, and winds) using rainfall observation either through indirect (like latent heat nudging, 1D+4D Var, etc.) or direct (4D Var, LETKF) assimilation of rainfall. But in this study, particle filter is used to select appropriate model physics with perturbation in dynamic variables in model physics using IMD rainfall observations, whereas, no changes are performed in the initial model state like traditional methods. Moreover, rainfall sub-space is indirectly coupled with other model sub-spaces (like temperature, moisture, and winds), so any modification in rainfall sub-space changes other sub-spaces also in forecasts. The another important point to note that less distribution is observed in EXP runs compared to CNT runs in short forecasts (mainly 3-hour forecasts; Figure 4). Since, all particles are initialized from same model state (here NCEP GDAS for initial state), the differences are mainly due to selection of model physics and dynamic variable perturbation in model physics using SKEBS. So, in short forecasts (3 hours) all selected particles are not able to represent true PDF, and it is the step where this filter may not be able to produce appropriate PDF. The possible solution may be to use local particle filter (Poterjoy, 2015) or Equivalent-weight particle filter (Ades and van Leeuwen 2015; Browne 2016) which consider proposal density to generate distribution of initial state in place of deterministic initial state opt in present study.

5. Conclusion

In this study, IMD observed rainfall is assimilated using particle filter. Two different set of experiments are performed with and without rainfall assimilation using different model physics options during August 1-10, 2015. Particle filter is implemented in rainfall sub-space (having less dimension) compared to full high dimensional model space with multiple hypothesis (based on less variance and large value of mean ETS) to produce new particles in resampling step. Rainfall is one of diagnostic parameters from the weather model which non-linearly depends on various parameters (like initial model state, terrestrial data, model physics, dynamical variables into parameterizations, etc.). Further, the dynamic variable perturbation through SKEBS method is used to generate new particles from high weight particles such that total number of particles (here 90) should be same. Ratheesh et al. (2016) also used this approach to change two dynamical parameters in guided particle filter to assimilate satellite measurements. The uses of SKEBS method to generate new particles in resampling step provide additional guidance to the particle towards future observations. Results based on different rainfall verification scores suggest that skill of rainfall forecast is improved with the assimilation of rainfall using particle filter compared to CNT runs. Moreover, rainfall assimilation also improves temperature, WVMR and wind speed forecasts at surface and different vertical levels. These results support that implementation of rainfall assimilation using particle filter, which consider non-linear and non-Gaussian distribution, improve prediction from the WRF model. In the case of the EnKF, the same configuration of physics parameterizations is kept for each ensemble member and all that changes at analysis time is the model state $x_{p,k}$ itself. In this study, particles are targeting towards best suited model physics with assimilation of rainfall.

In present study, all particles are initialized from same model state (here NCEP GDAS analysis), and the differences are mainly due to selection of model physics and dynamic variable perturbation in model physics using SKEBS. So, in short forecasts (3 h forecast; Figure 4) all selected particles are not able to represent true PDF, and it is the step where this filter may not be able to produce appropriate PDF. The possible solution may be to use local particle filter (Poterjoy, 2015) or Equivalent-weight particle filter (Ades and van Leeuwen 2015; Browne 2016) which consider proposal density to generate distribution of initial state in place of deterministic initial state. Since, objective of this study is to understand the role of model physics in an imperfect model using initial state (best) from the NCEP GDAS analysis. This work motivates to use Equivalent-weight particle filter proposed by Ades and van Leeuwen (2015) for a high-dimensional non-linear weather model to produce distribution of initial model states, and further select the appropriate model physics for imperfect (weak) model using particle filter and develop an "**efficient particle filter**" for the NWP model. This may be a scope for future research in fast developing field of non-linear data assimilation.

Acknowledgments

The author is thankful to the Director, Space Applications Centre, ISRO, India. Author acknowledges IMD rainfall observations available from http://www.imdpune.gov.in/. The NCEP global model analyses are downloaded from CISL-RDA (https://rda.ucar.edu/datasets/ds084.1/), and TRMM 3B42 merge rainfall product are available from https://gpm.nasa.gov/data-access/downloads/trmm.

References

Ades M, van Leeuwen PJ. 2015. The equivalent-weights particle filter in a high dimensional system. *Q. J. R. Meteorol. Soc.* **141(687)**: 484–503.

Bauer P, Auligne T, Bell W, Geer A, Guidard V, Heilliette S, Kazumori M, Kim MJ, Liu EHC, McNally AP, Macpherson B, Okamoto K, Renshaw R, Riishøjgaard LP. 2011. Satellite cloud and precipitation assimilation at operational NWP centres. *Q. J. R. Meteorol. Soc.* **137**: 1934–1951.

Benedetti A, Lopez P, Bauer P, Moreau E. 2005. Experimental use of TRMM precipitation radar observations in 1D + 4D − Var assimilation. *Q. J. R. Meteorol. Soc.* **131**: 2473–2495.

Berner J, Ha SY, Hacker JP, Fournier A, Snyder C. 2011. Model uncertainty in a mesoscale ensemble prediction system: Stochastic versus multiphysics representations. *Mon. Weather Rev.* **139(6)**: 1972-1995.

Berner J, Shutts GJ, Leutbecher M, Palmer TN. 2009. A spectral stochastic kinetic energy backscatter scheme and its impact on flow-dependent predictability in the ECMWF ensemble prediction system. *J. Atm. Sci.* **66(3)**: 603-626.

Browne PA. 2016. A comparison of the equivalent weights particle filter and the local ensemble transform Kalman filter in application to the barotropic vorticity equation. *Tellus A: Dynamic Meteorology and Oceanography* **68(1)**: 30466.

Chen Z. 2003. Bayesian Filtering: from Kalman filters to particle filters, and beyond, Report, McMaster University.

Chorin AJ, Morzfeld M, Tu X. 2013. A Survey of Implicit Particle Filters for Data Assimilation. In State-Space Models pp. 63-88 Springer New York.

Donner LJ. 1988. An initialization for cumulus convection in numerical weather prediction models. *Mon. Weather Rev.* **116**: 377–385.

Doucet A., De Freitas N, Gordon N. 2001. Sequential Monte-Carlo Methods in Practice, Springer-Verlag, New York.

Dubuisson S. 2015. Tracking with Particle Filter for High-dimensional Observation and State Spaces. John Wiley & Sons.

Errico RM, Bauer P, Mahfouf JF. 2007. Issues regarding the assimilation of cloud and precipitation data. *J. Atmos. Sci*. **64**: 3785–3798.

Huffman GJ, Adler RF, Stocker EF, Bolvin DT, Nelkin EJ, 2003. Analysis of TRMM 3-hourly multi-satellite precipitation estimates computed in both real and post-real time. 83rd AMS Annual Meeting. Paper P4.11 in 12th Conf. on Sat. Meteor. and Oceanog., 9-13 Feb. 2003, Long Beach, CA, 6 pp.

Huffman GJ, Coauthors, 2007. The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. J. Hydrometeor. 8: 38–55.

Heckley WA, Kelly G, Tiedtke M. 1990. On the use of satellite-derived heating rates for data assimilation within the tropics. *Mon. Weather Rev.* **118**: 1743–1757.

Hou AY, Zhang SQ, Reale O. 2004. Variational continuous assimilation of TMI and SSM/I rain rates: Impact on GEOS–3 hurricane analysis and forecast. *Mon. Weather Rev.* **132:** 2094–2109.

Kalnay E, Kanamitsu M, Kistler R, Collins W, Deaven D, Gandin L, Iredell M, Saha S, White G, Woollen J, Zhu Y, Leetmaa A, Reynolds R, Chelliah M, Ebisuzaki W, Higgins W, Janowiak J, Mo KC, Ropelewski C, Wang J, Jenne R, Joseph D. 1996. The NCEP/NCAR 40 year reanalysis project. *Bull. American Meteorol. Soc.* **77**: 437–471.

Kanamitsu M. 1989. Description of the NMC global data assimilation and forecast system. *Wea. Forecasting* **4**:335–342.

Kasahara A, Mizze AP, Donner LJ. 1994. Diabatic initialization for improvement in the tropical analysis of divergence and moisture using satellite radiometric imagery data. *Tellus A* **46**: 242–264.

Kotsuki S, Miyoshi T, Terasaki K, Lien GY, Kalnay E. 2017. Assimilating the global satellite mapping of precipitation data with the Nonhydrostatic Icosahedral Atmospheric Model (NICAM). *J. Geophys. Res. Atmos.* **122:** 631–650.

Krishnamurti TN, Ingles K, Cooke S, Kitade T, Pasch R. 1984. Details of low-latitude, medium-range numerical weather prediction using a global spectral model, Part 2: Effects of orography and physical initialization. *J. Meteorol. Soc. Japan.* **62**: 613–648.

Krishnamurti TN, Xue J, Bedi HS, Ingles K, Oosterhof O. 1991. Physical initialization for numerical weather prediction over the tropics. *Tellus A* **43**: 53–81.

Krishnamurti TN, Bedi HS, Ingles K. 1993. Physical initialization using SSM/I rain rates. *Tellus A* **45** 247–269.

Kumar P, Kishtawal CM. 2017. Importance of satellite-retrieved rain/no-rain information on short-range weather predictions. *International Journal of Remote Sensing* **38(13)**: 3851-3864.

Kumar P, Kishtawal CM, Pal PK. 2014. Impact of satellite rainfall assimilation on Weather Research and Forecasting model predictions over the Indian region. *J. Geophys. Res. Atmos.* **119(5):** 2017-2031.

Kumar P, Shukla MV. 2019. Assimilating INSAT-3D thermal infrared window imager observation with the particle filter: A case study for Vardah Cyclone. *J. Geophys. Res. Atmos.*, **124**, 1897– 1911.

Kumar P, Varma AK. 2016. Assimilation of INSAT-3D hydro-estimator method retrieved rainfall for short-range weather prediction. *Q. J. R. Meteorol. Soc.* DOI:10.1002/qj.2929

Kummerow C, Barnes W, Kozu T, Shiue J, Simpson J, 1998. The Tropical Rainfall Measuring Mission (TRMM) Sensor Package. J. Atmos. Oceanic Technol. 15: 809-817.

Lien GY, Kalnay E, Miyoshi T. 2013. Effective assimilation of global precipitation: Simulation experiments. *Tellus A* **65**: 19915.

Lien GY, Kalnay E, Miyoshi T, Huffman GJ. 2016a. Statistical properties of global precipitation in the NCEP GFS model and TMPA observations for data assimilation. *Mon. Weather Rev.* **144**: 663–679.

Lien GY, Miyoshi T, Kalnay E. 2016b. Assimilation of TRMM multisatellite precipitation analysis with a low-resolution NCEP global forecasting system. *Mon. Weather Rev.* **144**: 643–661.

Lopez P. 2011. Direct 4D-Var assimilation of NCEP stage IV radar and gauge precipitation data at ECMWF. *Mon. Weather Rev.* **139**: 2098–2116.

Manobianco J, Koch S, Karyampudi M, Negri AJ. 1994. The impact of assimilating satellite-derived precipitation rates on numerical simulations of the ERICA IOP 4 cyclone. *Mon. Weather Rev.* **122**: 341-365.

Marecal V, Mahfouf JF. 2003. Experiments on 4D-Var assimilation of rainfall data using an incremental formulation. *Q. J. R. Meteorol. Soc.* **129**: 3137–3160.

Maskell S, Gordon N. 2001. A Tutorial on Particle Filters for On-line Nonlinear/Non-Gaussian Bayesian Tracking. IEEE Transactions on Signal Processing **50(2)**: 174–188.

Mathur MB, Bedi HS, Krishnamurti TN, Kanamitsu M, Woolen JS. 1992. Use of satellite derived rainfall for improving tropical forecasts. *Mon. Weather Rev.* **120**: 2540-2560.

Mattern JP, Dowd M, Fennel K. 2013. Particle Filter-Based Data Assimilation for a ThreeDimensional Biological Ocean Model and Satellite Observations. *J. Geophys. Res.* **118**: 2746–2760.

Pai DS, Sridhar L, Rajeevan M, Sreejith OP, Satbhai NS, Mukhopadhyay B. 2014. Development of a new high spatial resolution (0.25° × 0.25°) long period (1901–2010) daily gridded rainfall data set over India and its comparison with existing data sets over the region. *Mausam* **65**: 1–18.

Peng MS, Chang SW. 1996. Impact of SSMI/I-retrieved rainfall rates on numerical prediction of a tropical cyclone. *Mon. Weather Rev.* **124**: 1181-1198.

Poterjoy J. 2015. A localized particle filter for high-dimensional nonlinear systems. *Mon. Weather Rev.* **144**: 59-76.

Posselt DJ, Bishop CH. 2012. Nonlinear parameter estimation: comparison of an ensemble Kalman smoother with a Markov chain Monte Carlo algorithm. *Mon. Weather Rev.* **140(6)** 1957–1974.

Posselt DJ, Hodyss D, Bishop CH. 2014. Errors in Ensemble Kalman Smoother Estimates of Cloud Microphysical Parameters. *Mon. Weather Rev.* **142**: 1631-1654.

Puri K, Miller MJ. 1990. The use of satellite data in the specification of convective heating for diabatic initialization and moisture adjustment in numerical weather prediction models. *Mon. Weather Rev.* **118**: 67–93.

Ratheesh S, Chakraborty A, Sharma R, Basu S. 2016. Assimilation of satellite chlorophyll measurements into a coupled biophysical model of the Indian Ocean with a guided particle filter. *Remote Sensing Letters* **7(5)**: 446-455.

Ristic B, Arulampalam S, Gordon N. 2004. Beyond the Kalman filter: Particle Filters for Tracking Applications. London, UK: Artech House.

Skamarock WC, Klemp JB, Dudhia J, Gill DO, Barker DM, Duda MG, Huang XY, Wang W, Powers JG. 2008. A description of the advanced research WRF version 3, Tech. Note NCAR/TN-475 STR. pp. 113, Mesoscale and Microscale Meteorology Division, National Center of Atmospheric Research, June 2008.

Stephenson D, Casati B, Ferro CAT, Wilson CA. 2008. The extreme dependency score: a non-vanishing measure for forecasts of rare events. *Met. App.* **15**: 41–50.

Treadon RE. 1996. Physical initialization in the NMC global assimilation system. *Meteorol. Atmos. Phys.* **60**: 57–86.

Tsuyuki T.1997. Variational data assimilation in the Tropics using precipitation data. Part III: Assimilation of SSM/I precipitation rates. *Mon. Weather Rev.* **125**: 1447–1464.

Van Leeuwen PJ. 2009. Particle Filtering in Geophysical Systems. *Mon. Weather Rev.* **137**: 4089–4114.

Van Leeuwen PJ. 2010. Nonlinear Data Assimilation in Geosciences: An Extremely Efficient Particle Filter. *Q. J. R. Meteorol. Soc.* **136**: 1991–1999.

Van Tuyl AH. 1996. Physical initialization with the Arakawa-Schubert scheme in the navy's operations global forecast model. *Meteorol. Atmos. Phys.* **60**: 47-55.

Sequential Importance Resampling for Model Physics

Model Physics	PBL1	PBL5	MP2	MP3	MP5	MP6	MP8	CP1	CP2	CP3	CP4	CP5	CP6	CP14	CP93	CP99
01 August 2015	45	45	18	18	18	18	18	10	10	10	10	10	10	10	10	10
02 August 2015	40	50	15	20	18	17	20	17	00	09	11	14	10	16	04	09
03 August 2015	38	52	12	22	19	15	22	19	00	00	13	8	10	28	04	08
04 August 2015	42	48	10	20	21	11	28	29	00	00	17	00	10	34	00	00
05 August 2015	28	62	08	25	17	11	29	49	00	00	18	00	10	13	00	00
06 August 2015	20	70	05	41	15	09	20	57	00	00	17	00	05	11	00	00
07 August 2015	15	75	04	46	15	07	18	57	00	00	17	00	05	11	00	00
08 August 2015	26	64	05	33	17	06	29	42	00	00	32	00	05	11	00	00
09 August 2015	47	43	05	08	20	07	50	15	00	00	59	00	05	11	00	00
	PBL		**Micro Physics**					**Cumulus Physics**								

Figure 1: Schematic of rainfall assimilation using particle filter to estimate target PDF from an imperfect model using initial state from the NCEP analysis in the WRF model using different model physics and dynamic variable perturbation in the physical parametrization during 1-9 August 2015. The numbers in the boxes represent the number of particles using that particular scheme.

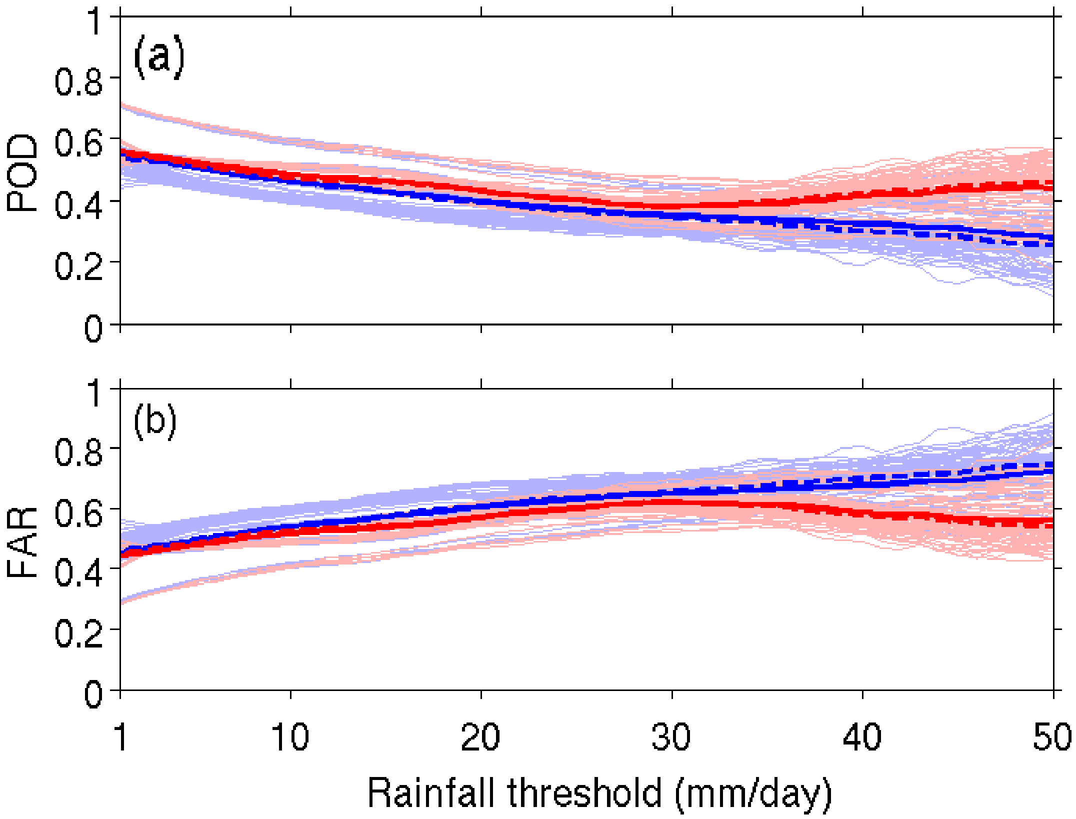

Figure 2: (a) POD and (b) FAR verification scores for the CNT (blue) and EXP (red) runs predicted daily accumulated rainfall at different rainfall thresholds valid on 10 August 2015. Individual particles are shown by light blue and red lines for CNT and EXP run, respectively. Mean and median are plotted using dark line and dark dash lines respectively.

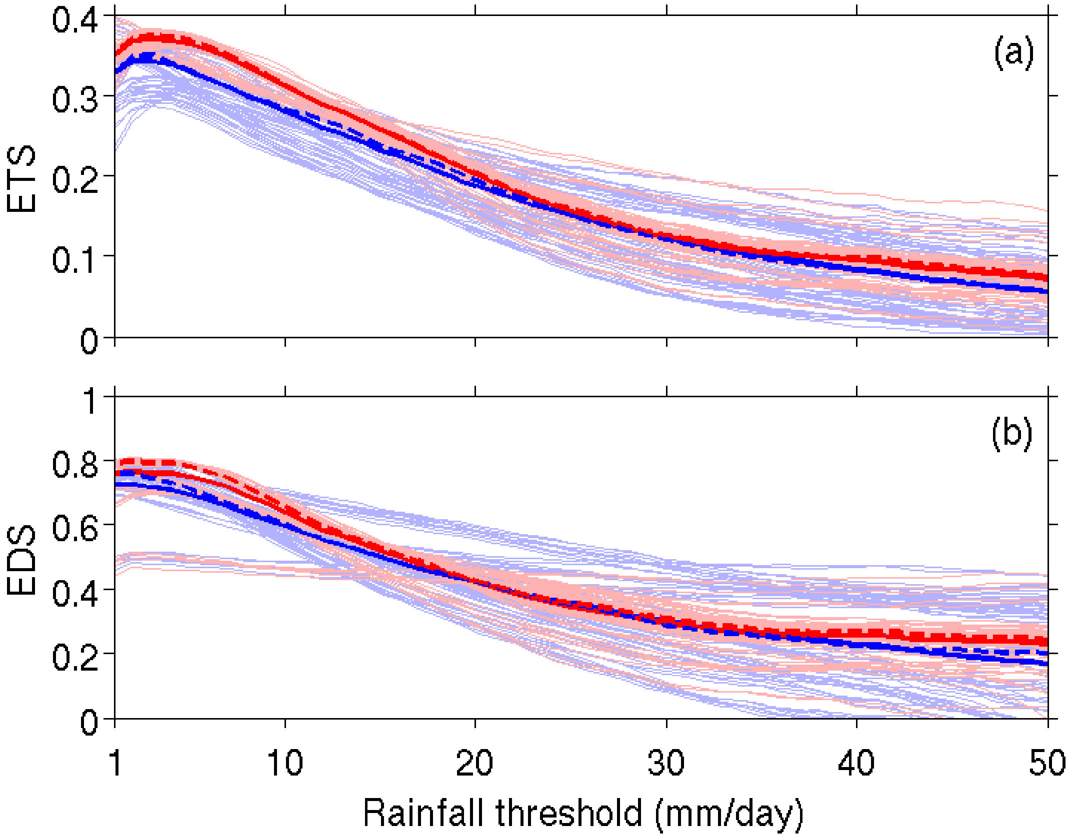

Figure 3: (a) ETS and (b) EDS verification scores for the CNT (blue) and EXP (red) runs predicted daily accumulated rainfall at different rainfall thresholds valid on 10 August 2015. Individual particles are shown by light blue and red lines for CNT and EXP run, respectively. Mean and median are plotted using dark line and dark dash lines respectively.

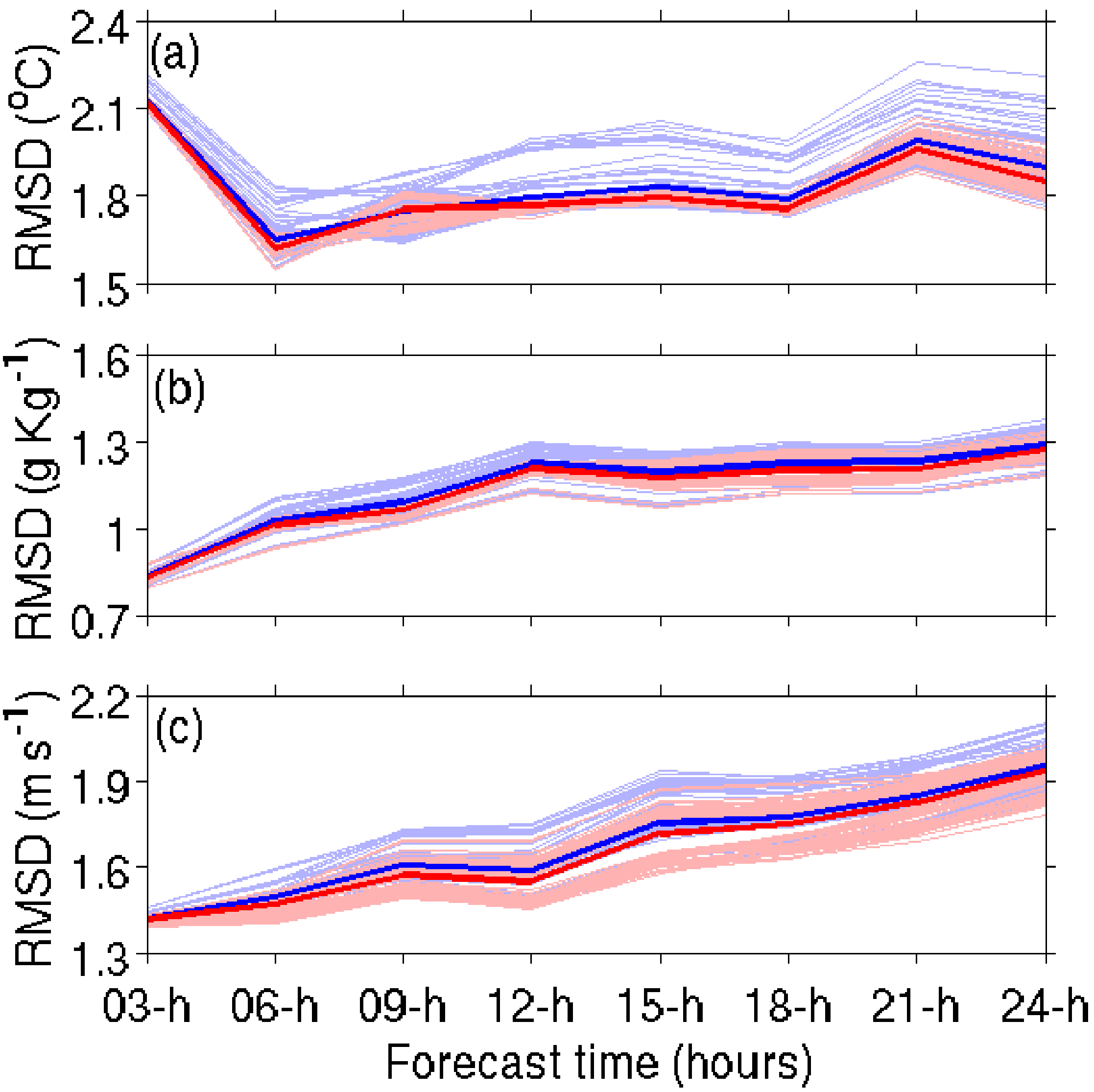

Figure 4: Temporal distribution of RMSD in surface (a) temperature, (b) WVMR, and (c) wind speed forecasts at 3 hours' interval from CNT (blue) and EXP (red) run against NCEP final analysis. Individual particles are shown by light blue and red lines for CNT and EXP run, respectively. Mean is plotted by dark line.

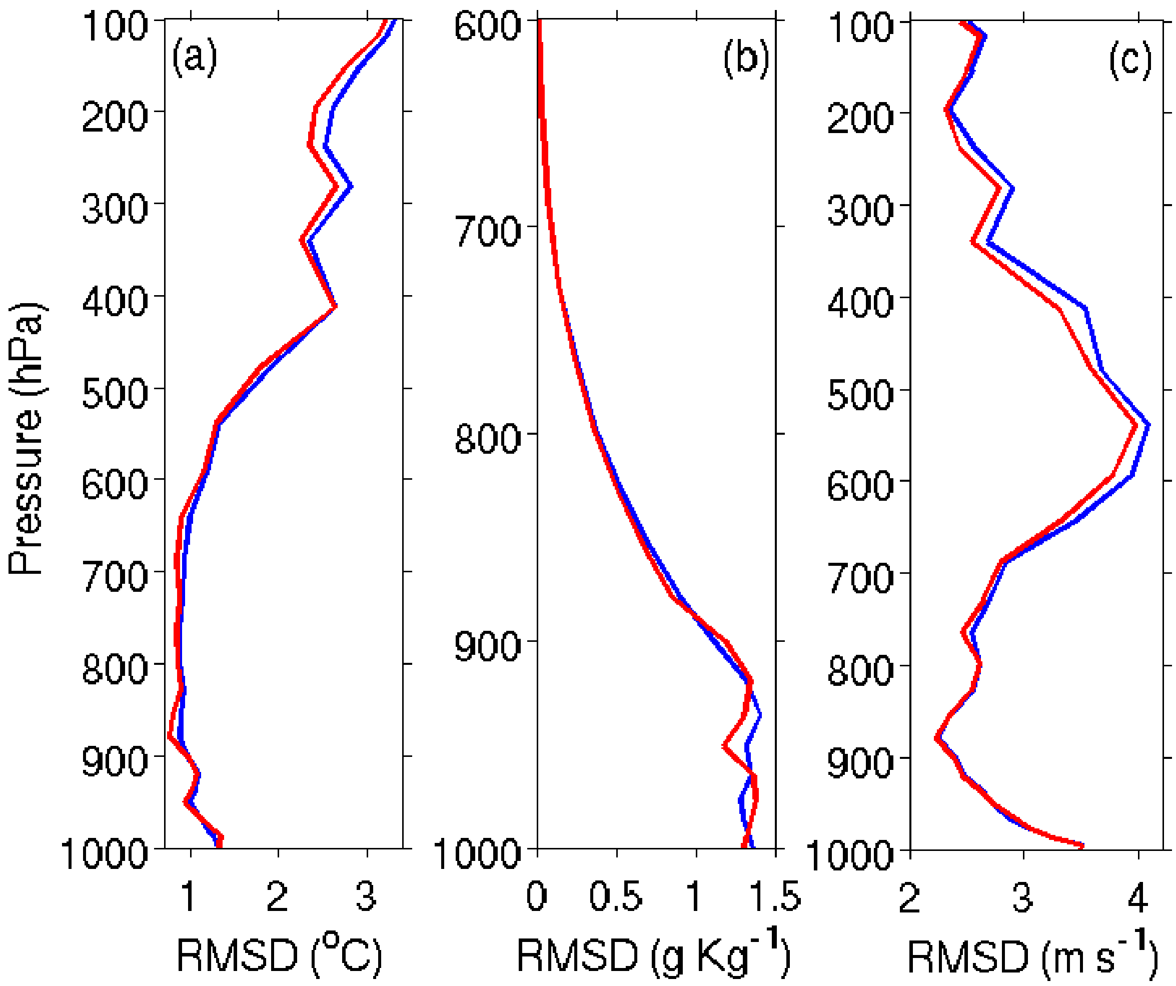

Figure 5: Vertical profile of RMSD in 24 hour (a) temperature, (b) WVMR, and (c) wind speed forecasts from CNT (blue) and EXP (red) run against NCEP final analysis.