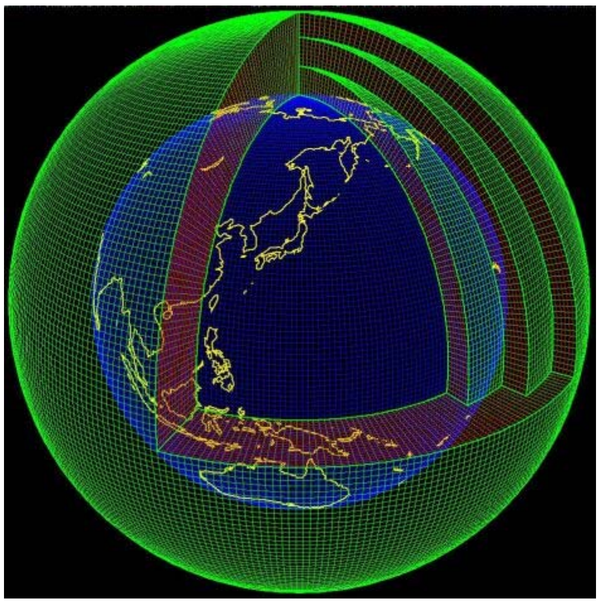
Assimilation of INSAT-3D Thermal Infrared Window Imager Observation using Particle Filter

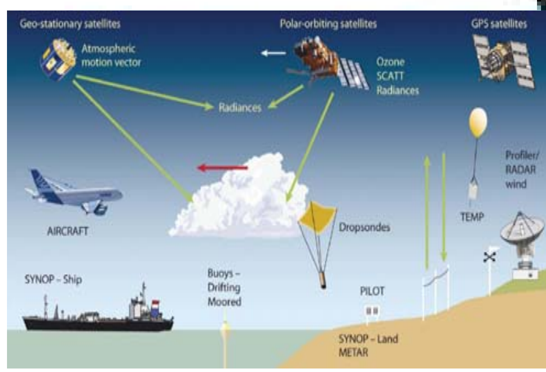
Prashant Kumar
(prashant22@sac.isro.gov.in)
Atmospheric Sciences Division
Atmospheric and Oceanic Sciences Group, EPSA
Space Applications Centre, ISRO, India

International conference on Ensemble Methods in Modelling and Data Assimilation (EMMDA)
24-26 February 2020

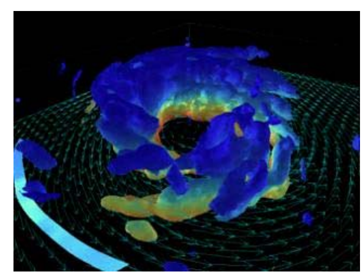
Who uses Satellite Data ?

NWP models use most data from meteorological/ocean satellites through assimilation





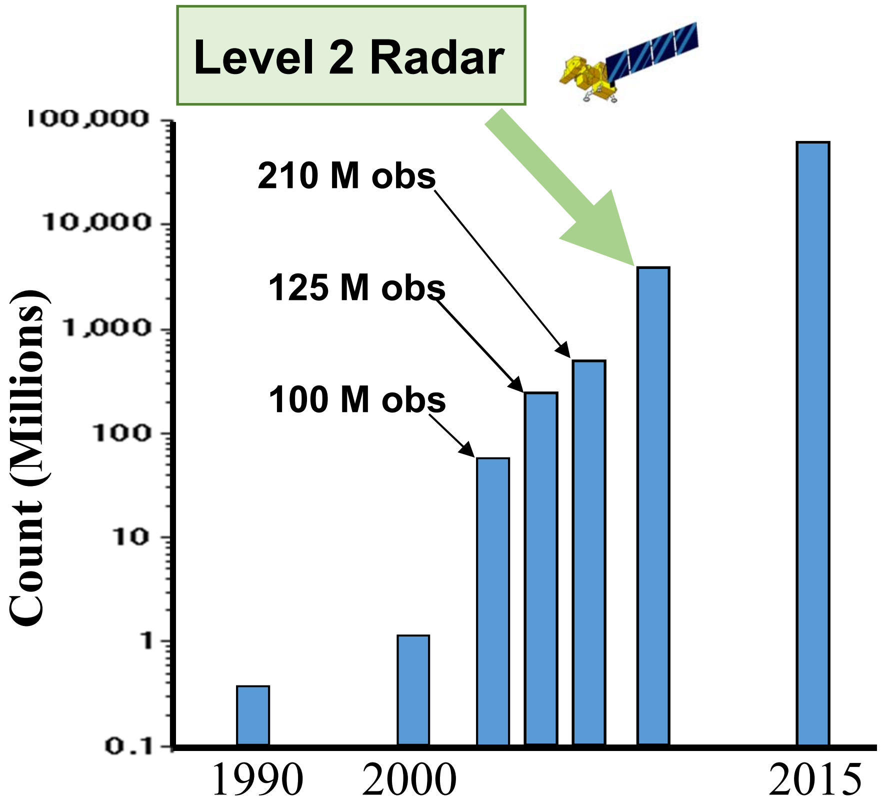
Today, more than 95% observations for weather prediction are provided by satellites !



THIS IS ONLY 5% OF WHAT SATELLITES OBSERVE !!

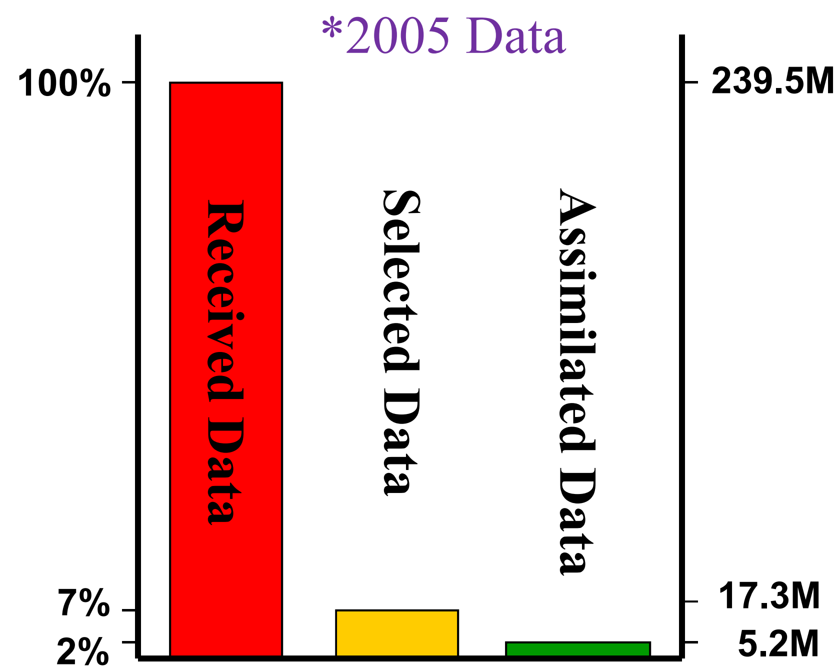
Data Ingest in NWP Model

Daily Satellite & Radar & Conventional Observation Count



Five Order of Magnitude Increases in Satellite Data Over Fifteen Years (2000-2015)

Daily Percentage of Data Ingested into Models



Received = **All observations received operationally from providers**

Selected = **Observations selected as suitable for use**

Assimilated = **Observations actually used by models**

(John C. Derber, NCEP)

Challenges in Data Assimilation ?

- Linear estimation theory and Gaussian conditions
- Radiance Assimilation is restricted to clear-sky conditions / Now all-sky MW possible
- Selection of Uncorrelated Observations/Channels
- Limitations to use WV and CO2/O2 Absorption channels only
- **Error distribution of Hydrometeors like cloud, ice, etc.**
- Precise Inputs for Fast RT model in all-sky simulations
- **Utilization of high temporal/spatial resolution Observations (e.g. Visible channels)**

Impact studies using OSE/OSSE vs FSO
(Guidance for CGMS, Stephen English)

Strengths and weaknesses of FSOI and OSE approaches

- OSEs answer the question "what if I lose/add not have this data type?"
- FSOI answers the question "given the setup how much does this data type contribute to forecast error reduction?"

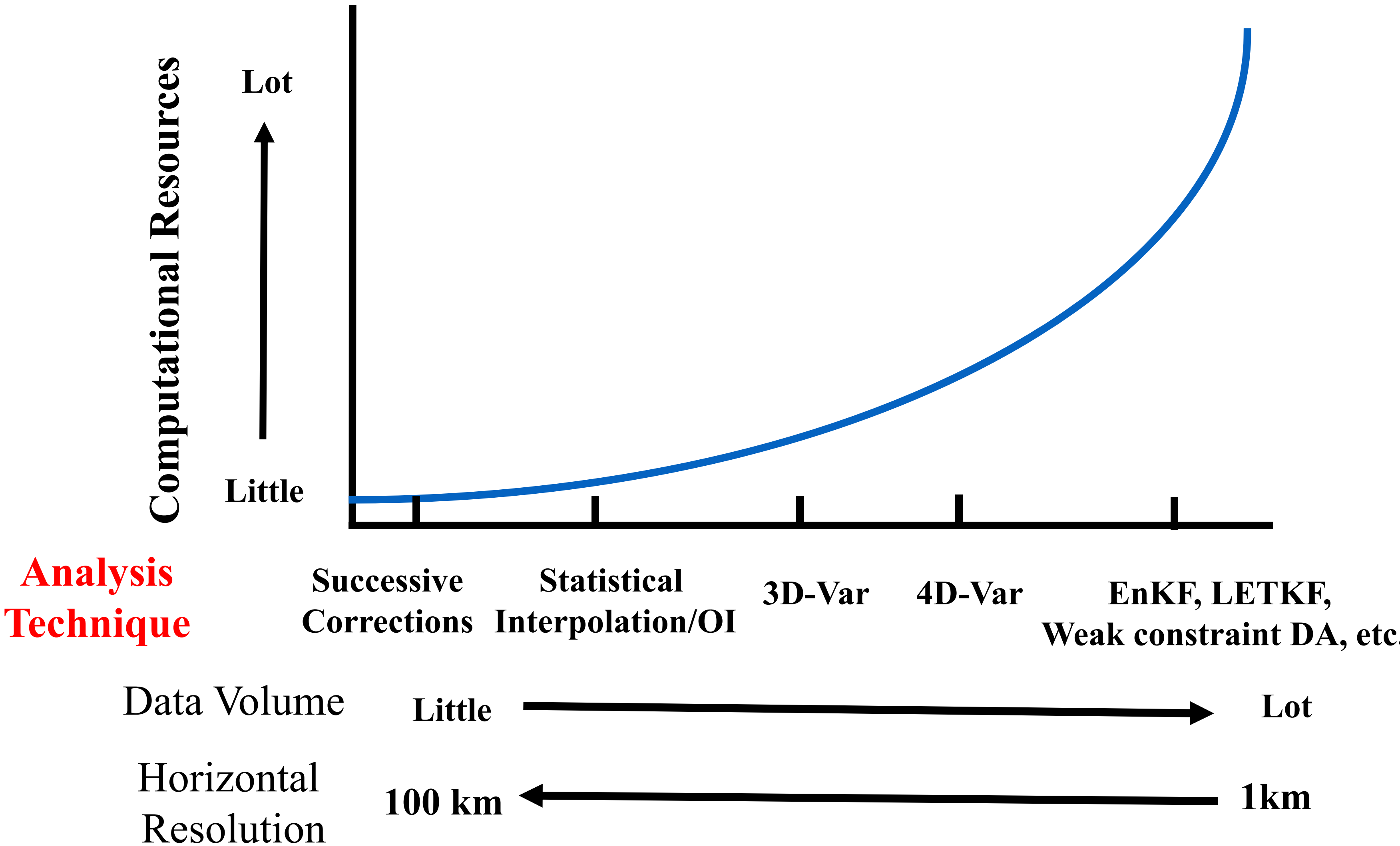
Example 1: add datatype with unrealistic low observation errors. OSE will show negative impact. FSOI will show this datatype contributes most to forecast error reduction.

Both are correct! But both are open to misinterpretation. ***FSOI measures impact in that setup.***

Example 2: add two identical datatypes, first one, then other. OSEs show first has large impact, second small impact. FSOI show they have the same impact.

Both are correct! But both are open to misinterpretation. ***OSEs are sensitive to order of changes.***

Computational Requirement



Example of NGFS GDAS (00Z 13 Feb 2020)

Type of Observations	No. of Observations (± 3 Hours)
Surface (BUFR, SYNOP,SHIP, BUOY, METAR, AWS,etc.)	53581
Sonde (TMP, PILOT, PROFILER)	4455
AIRCFT	172367
Humidity Sounder (MHS, SAPHIR, ATMS, SSMIS, MWHS)	1176864
Temperature Sounder (AMSUA, MWTS, ATMS)	209075
Multispectral IR Sounder (HIRS, GOES, INSAT)	318358
Hyperspectral IR Sounder (IASI, AIRS, CrIS)	553820
AVHRR Radiance	392648
Geo WV Radiance	174950
GPSRO	69328
SCAT Winds (SCATSat-1, ASCAT)	156652
AMVs (IR, VIS, WV)	909836
Winds LIDAR	-

Objective of the Study

- A big challenge in the satellite data assimilation is the effective use of InfraRed (IR) window channel radiances in the high-resolution weather model. (Generally not assimilated)

- A hybrid data-assimilation method is demonstrated in which 3DEnVar method is used to assimilate Reference observations (Conventional +Satellite), and particle filter method is used to assimilate TIR-1 T_B from INSAT-3D satellite.

Kumar, P., & Shukla, M. V. (2019). *Journal of Geophysical Research: Atmospheres*, 124, 1897– 1911.

Methodology

The particle filter considers a probability density function (*pdf*) of a state, and the *pdf* is approximated by particles consisting of large number of discrete samples to approximate posteriori by a weighted samples. In this study, various choices of model physics are selected to generate initial particles.

These particles represent a sample from its priori *pdf*, and assume to be of the form

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

Here, $x_{m,k}$ is the set of state vector with m different model physics schemes to be estimated at time step k, and f_{k-1}^{m} is a known imperfect non-linear model with m different model physics options, x_{k-1} is 6-hours forecasts from past run having noise of v_{k-1} at time step $k-1$.

The idea is to represent the prior *pdf* by a set of particles $x_{m,k}$, which are delta functions center around state vectors. If one represents the prior *pdf* by a number of particles, like in the EnKF, so

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

where, N is number of particles (which are 92 here) with different model physics ($x_{m_i,k}$) at time step *k*. Then, from Bayes theorem

$$p(x \mid y) = \sum_{i=1}^{N} w_i \ \delta\left(x - x_{m_i,k}\right)$$

where, weights w_i are given by

$$w_i = \frac{p\left(y \mid x_{m_i}\right)}{\sum_{j=1}^{N} p\left(y \mid x_{m_j}\right)}$$

Particle Filter degeneracy: resampling

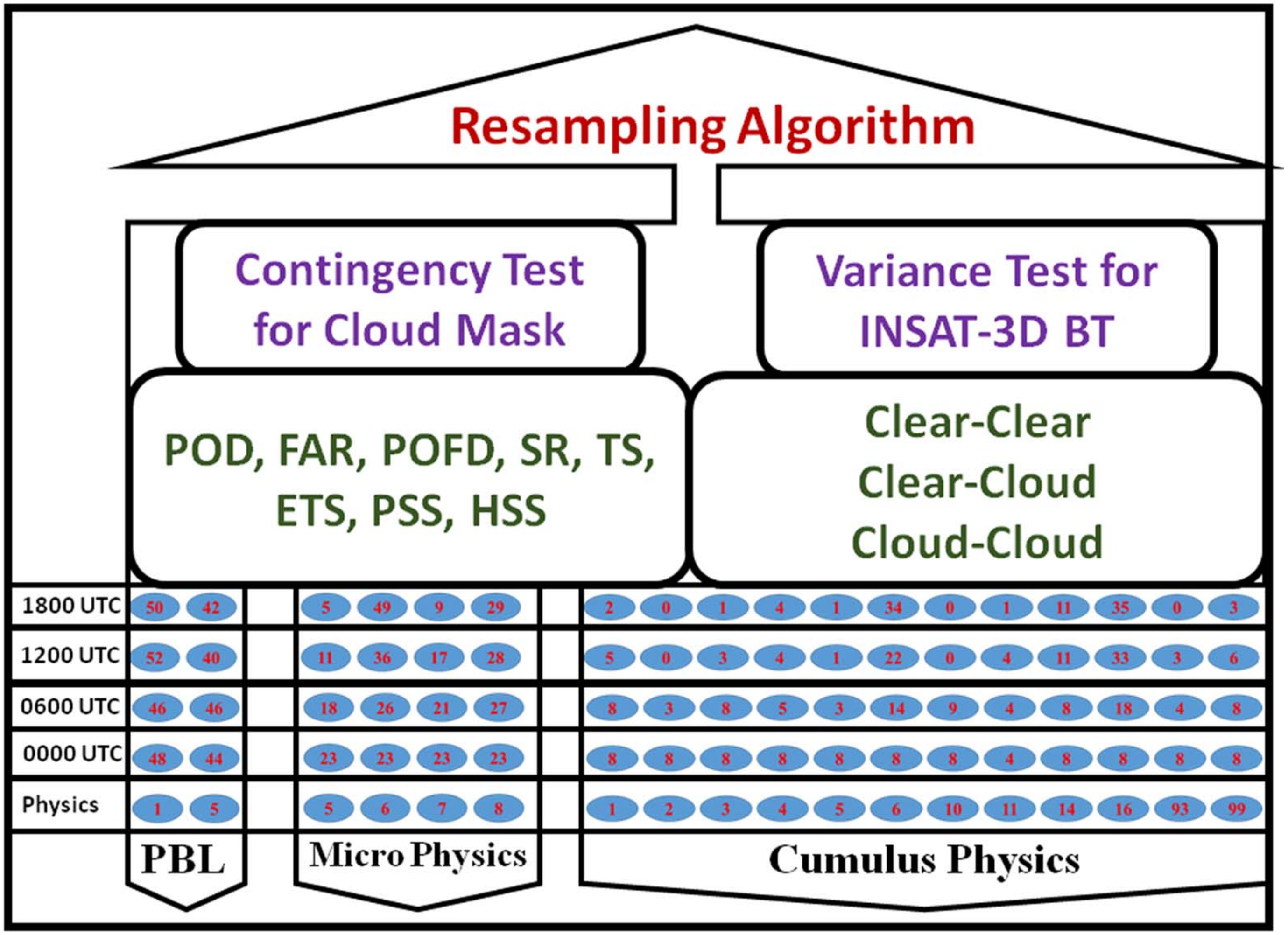
- With each new set of observations the old weights are multiplied with the new weights.
- Very soon only one particle has all the weight…

However: degeneracy

- For large-scale problems with lots of observations this method is still degenerate:
- Only a few particles get high weights; the other weights are negligibly small.
- However, we can enforce almost equal weight for all particles:

(Van Leeuwan, 2015)

Analysis Step



In the analysis step, INSAT-3D measured TIR-1 BT and cloud mask product are used to determine weight $\left(w_i\right)$ for each particle. This step involves weighting to each particle and subsequent weight-based resampling

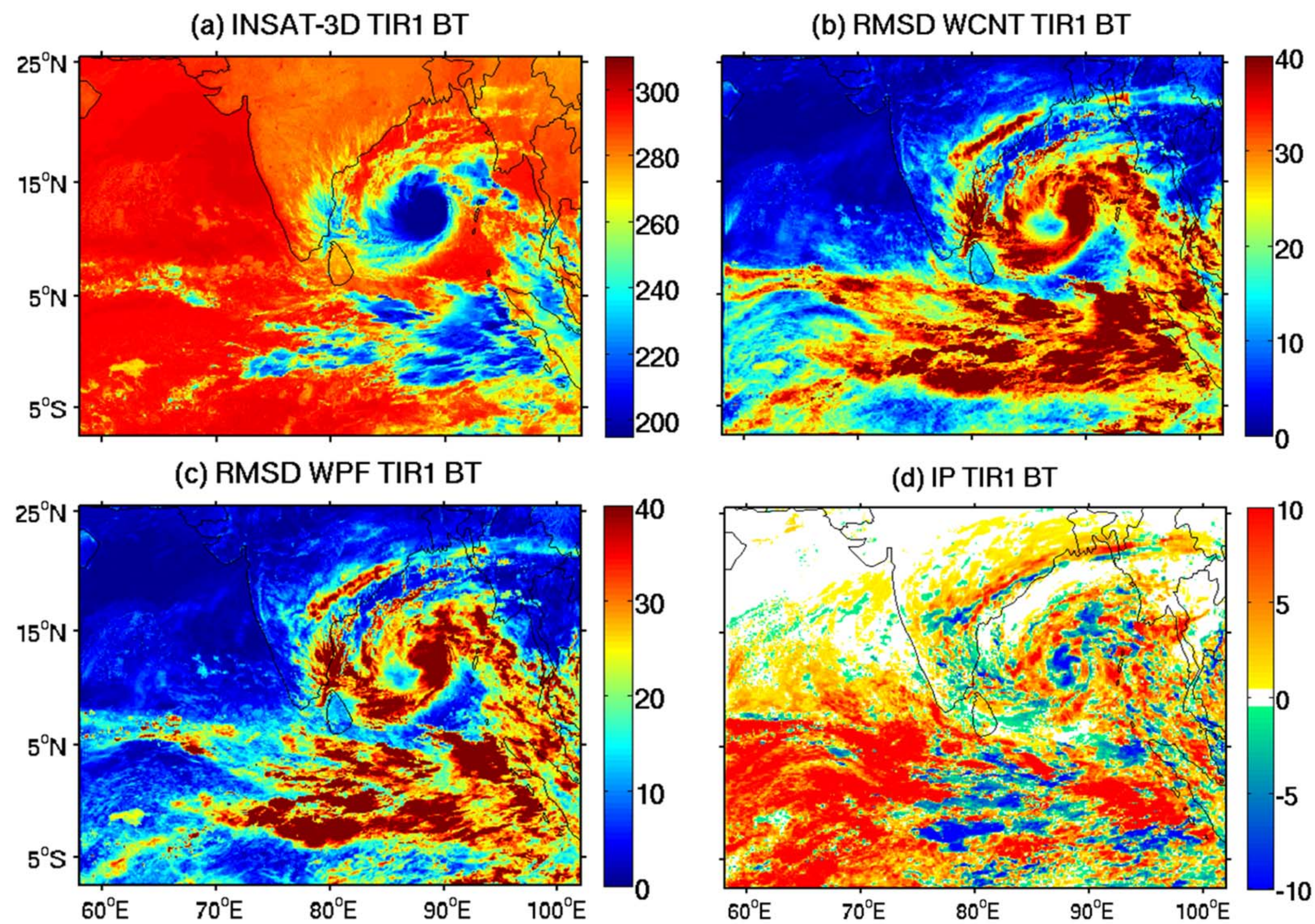
In brief, following steps are used to calculate weights: (a) initially raw weights are calculated using $w_{raw} = \frac{1}{variance}$ for TIR-1 BT, or $w_{raw} = NSC$ for cloud-mask; (b) these raw weights are used to calculate intermediate weights $\widetilde{w}_{raw} = \frac{w_{raw}}{\max(w_{raw})}$; and the final weights after normalization are given as $w_{raw}^{f} = \frac{\widetilde{w}_{raw}}{\sum \widetilde{w}_{raw}}$.

In this process, particles having high variance and less NSC skill scores are rejected which contribute very little to the approximation of the target *pdf*.

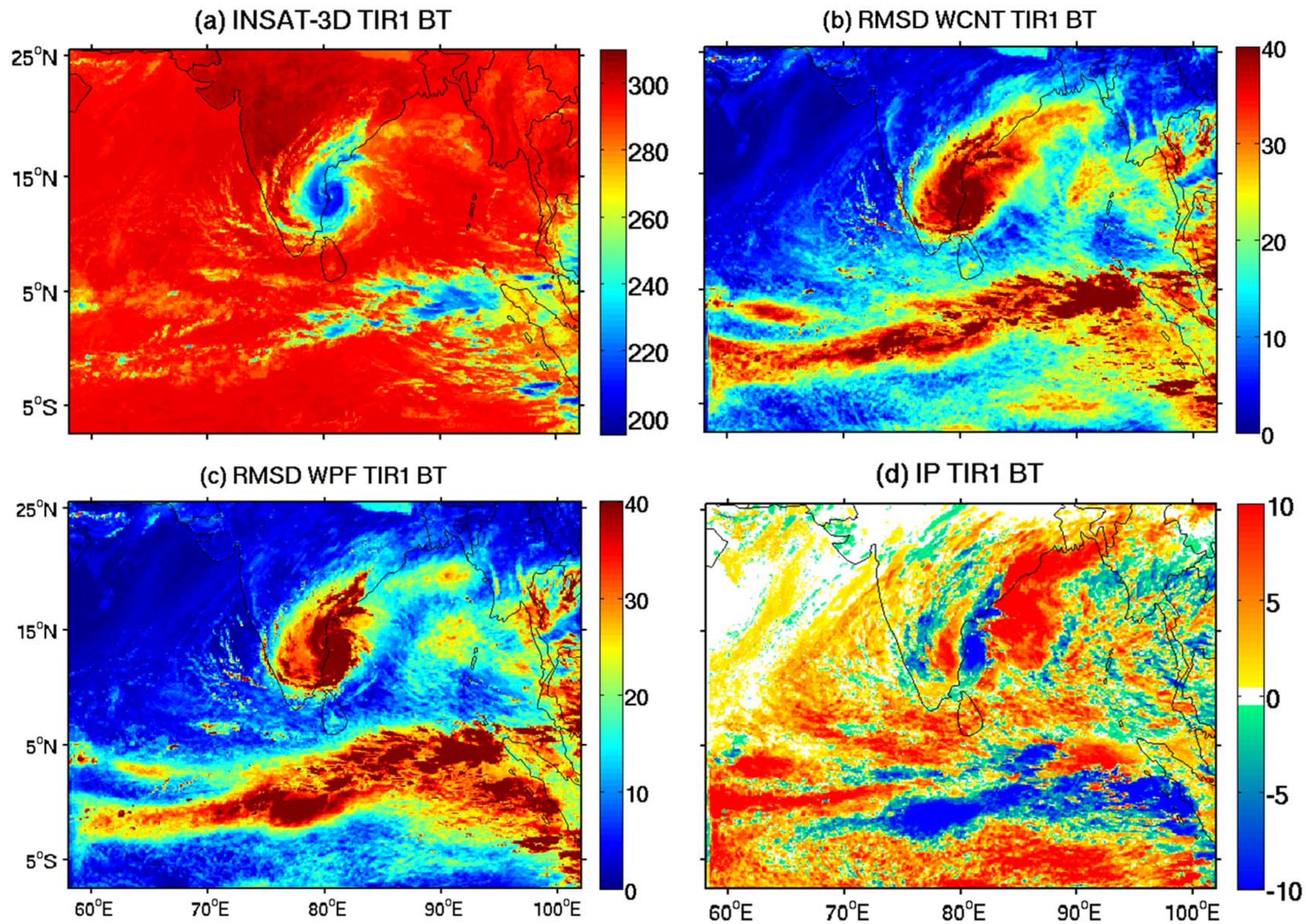
Resampling Step

particles having higher weights are resampled at the observation time, whose distribution forms a weak approximation of the target *pdf*.

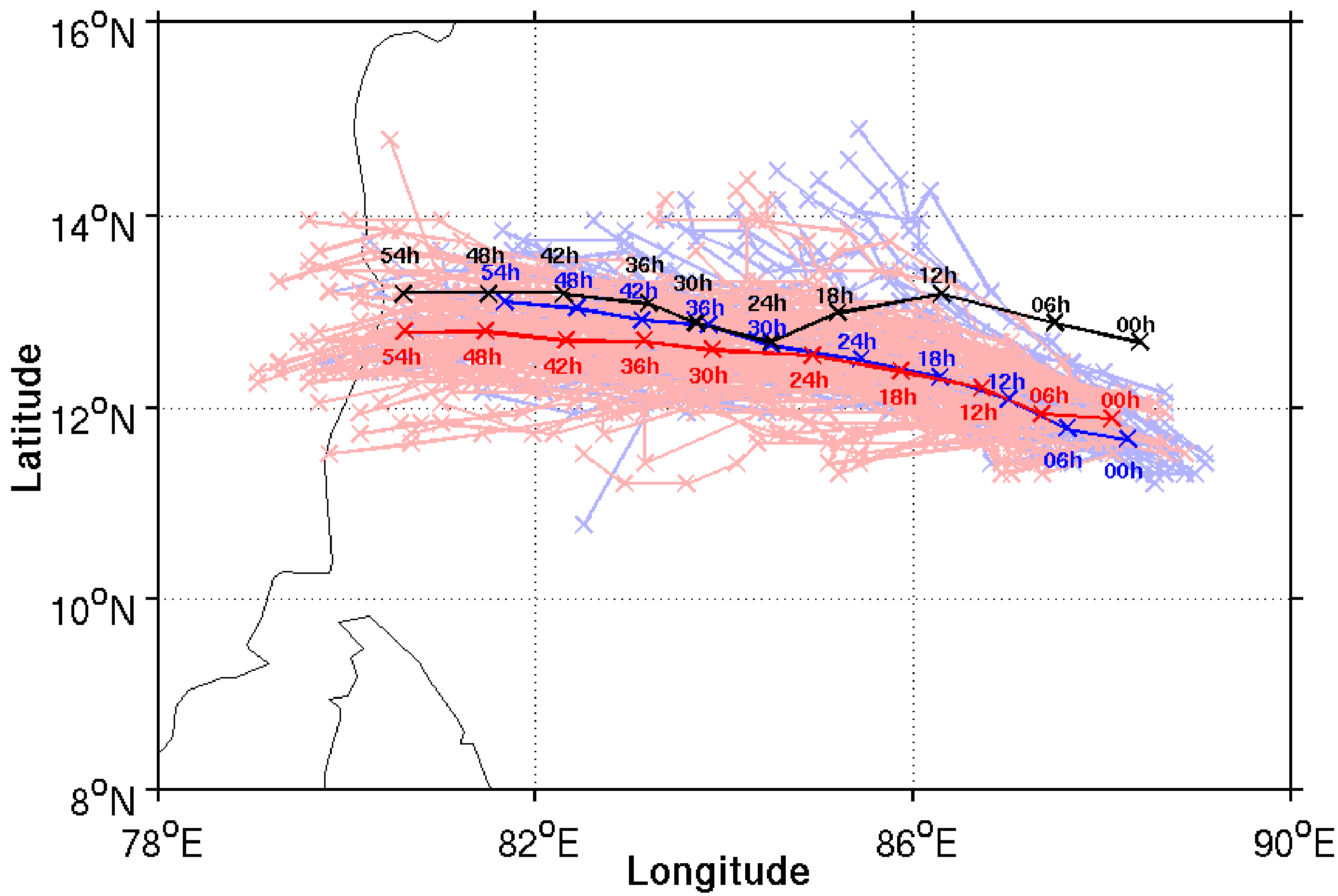
new particles are generated from large weight particles (with same physics options) using stochastic kinetic-energy backscatter scheme (SKEBS; *Berner et al. 2009, 2011*) to avoid rapid filter degeneracy



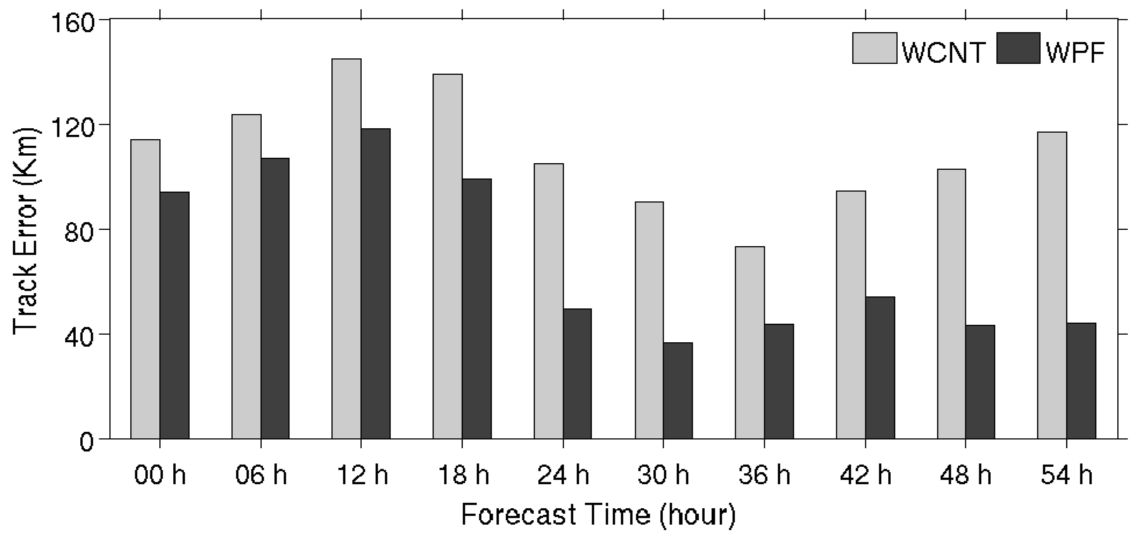
Spatial distribution of (a) TIR-1 BT (K) measured from INSAT-3D satellite, root-mean-square-difference (RMSD) in (b) WCNT and (c) WPF simulated TIR-1 BT analyses against INSAT-3D measured TIR-1 BT, and (d) improvement parameter (IP; in K) for TIR-1 BT. In which positive (negative) values show improvement (degradation) in the WCNT and WPF analyses valid at 0000 UTC 10 December 2016.



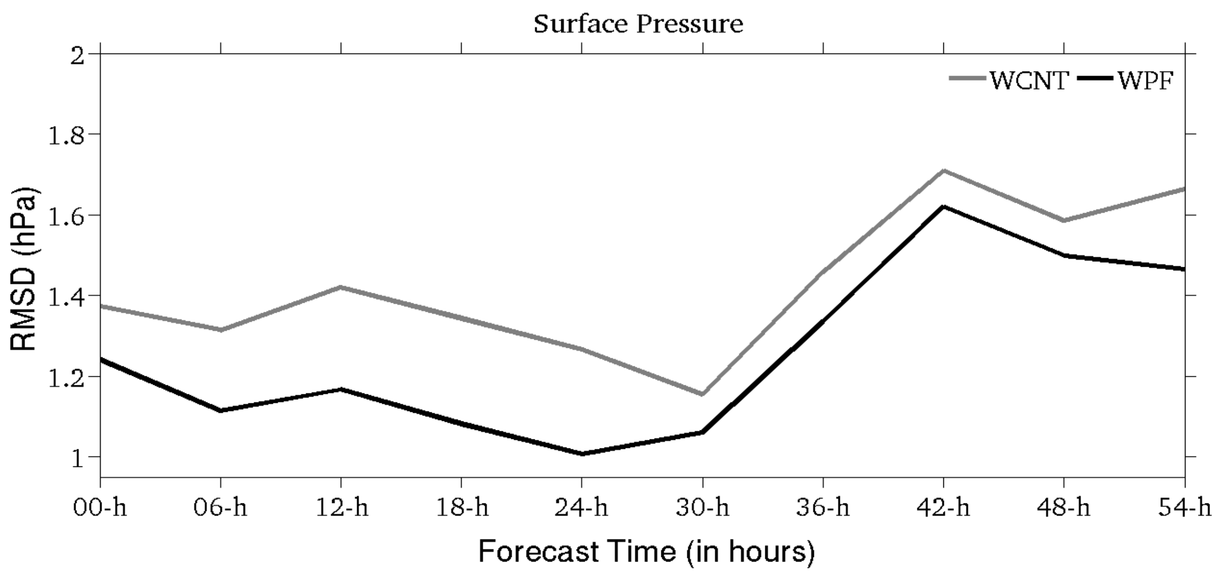
Spatial distribution of (a) TIR-1 BT (K) measured from INSAT-3D satellite, RMSD in (b) WCNT and (c) WPF predicted TIR-1 BT against INSAT-3D measured TIR-1 BT, and (d) improvement parameter (IP; in K) for TIR-1 BT valid at 0600 UTC 12 December 2016.



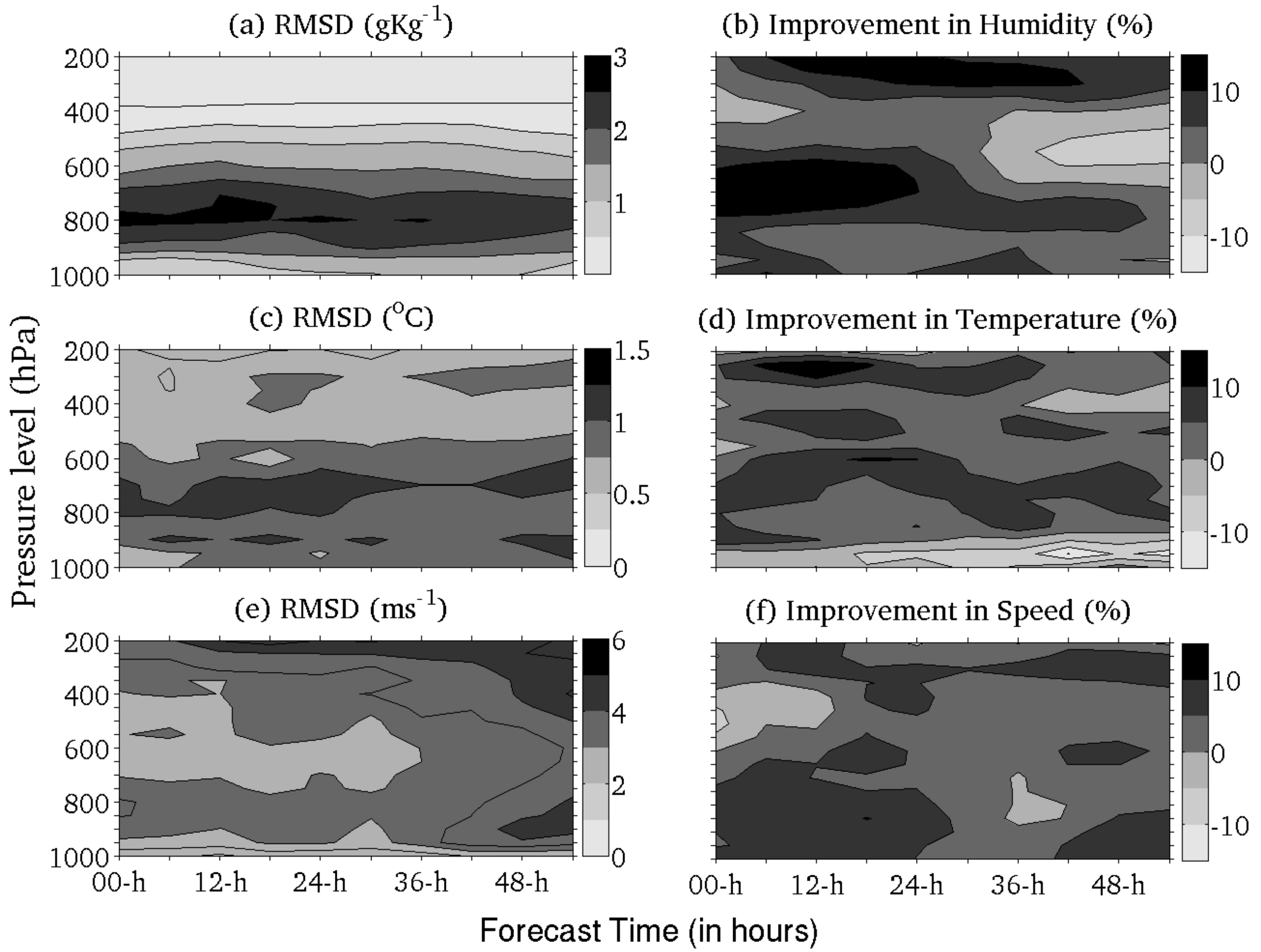
Track of the storm center for 54 h period starting from 0000 UTC 10 December 2016 and ending at 0600 UTC 12 December 2016. The light blue and light red lines show simulated cyclone tracks from different particles of WCNT and WPF experiments, respectively. The bold blue, red and black lines show mean track from WCNT, WPF experiments, and IMD observed best track, respectively.



Six-hourly track error in the simulated cyclone track (in km). The mean track error is the minimum for WPF runs.



RMSD in surface pressure analysis and forecasts in the WCNT and WPF experiments when compared with ECMWF analysis.



Vertical structure of RMSD in (a) humidity, (c) temperature, and (e) wind speed in WCNT experiments when compared with ECMWF analyses, and vertical distribution of improvement parameter $\left(\frac{RMSD_{WCNT} - RMSD_{WPF}}{RMSD_{WCNT}} x\ 100\right)$ in (b) humidity, (d) temperature, and (f) wind speed in WPF experiments over WCNT experiments.

Conclusion

- The particle filter is only used to select the most appropriate model physics.

- The INSAT-3D window channel data are used in particle filter with multiple criteria and sequential importance resampling to select the best suitable model physics, which approaches to target *pdf* and help to reduce uncertainties due to model physics. **This is the step in which INSAT-3D data are used for assimilation.**

- Results show that WPF runs are better simulated TIR-1 BT compared to WCNT runs in analyses as well as at landfall time. Moreover, **storm's track prediction is also improved with the help of INSAT-3D data assimilation using particle filter**, and leads to positive impact in storm intensity prediction in short range forecasts.

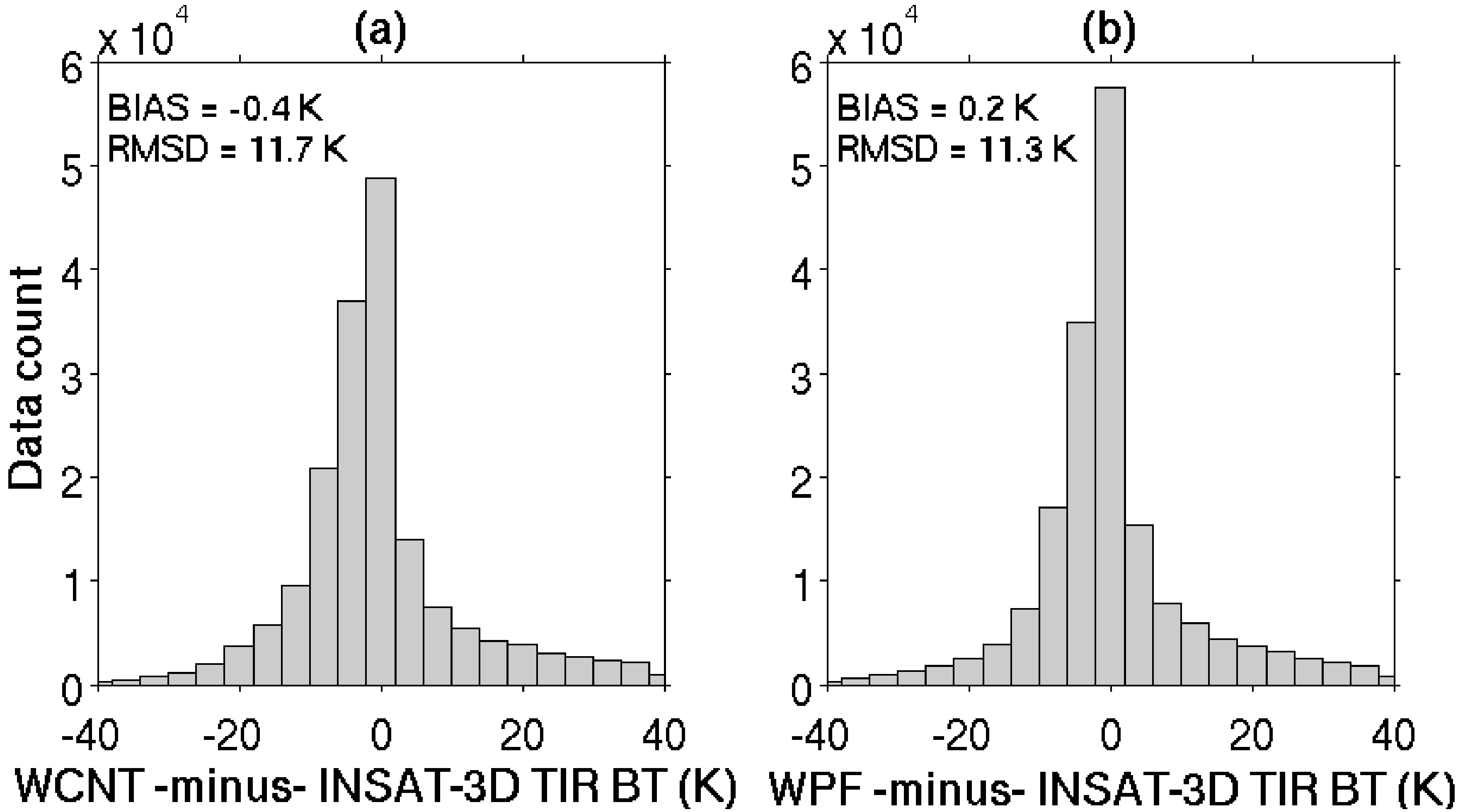
- The Vertical structure of humidity, temperature, wind speed and surface pressure against ECMWF analysis also demonstrate superiority of WPF experiments over WCNT experiments.

-

"Mr Derber , may I go home? I can't assimilate any more data today."

Thanks for your time.

Results



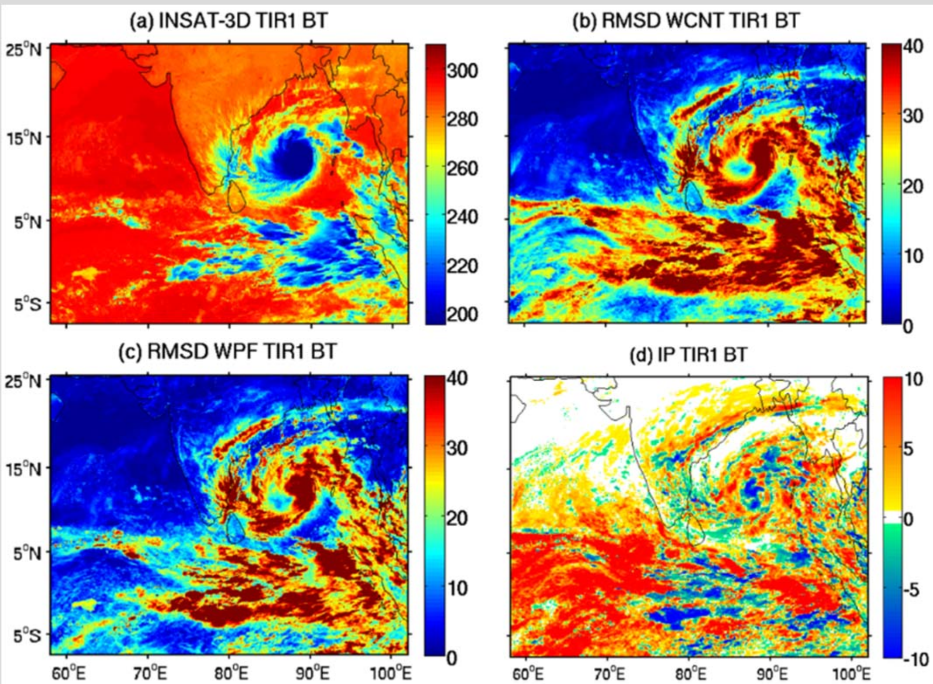
The distribution of (a) first-guess departure $\left(\overline{WCNT} - obs\right)$ and (b) analysis departure $\left(\overline{WPF} - obs\right)$ for TIR-1 brightness temperature (BT; in K) on 0000 UTC 10 December 2016, where $\overline{WCNT}$ and $\overline{WPF}$ represent mean BT for WCNT and WPF runs.

Key elements in interpretation of OSE and FSOI include:

- **Is there statistical significance testing?** If not, disregard the results;
- **Long experiments are needed to achieve statistical significance**, approaching
a year for a single forecast per day for a 0.5% change in RMS error;
- **Case studies** can illustrate results proven in long experiments, but prove nothing
by themselves;
- **Be clear on the type of study**: what is the baseline? Data addition or data denial?
- **Ask do the tests use the full system** e.g. feedback on background error – Sandy
case satellite data denial – retuning background error restored 50% of lost impact;
- **Verification against analysis can create misleading results** due to model biases
and error correlation between forecast and analysis;
- **Verification against observations can create misleading results** due to limited
geographical sampling;
- **Deterministic versus Ensemble forecasts.**

All-Sky Radiance Assimilation using Particle Filter

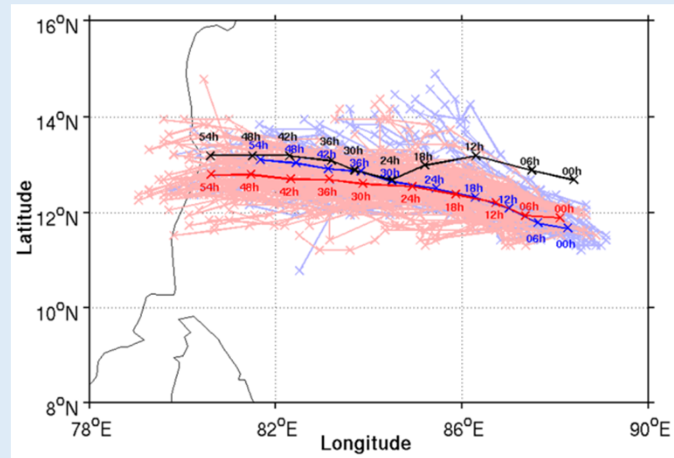
Target: Assimilation of cloud/rain influence observations, non-linear filter, non-Gaussian data ingestion in NWP

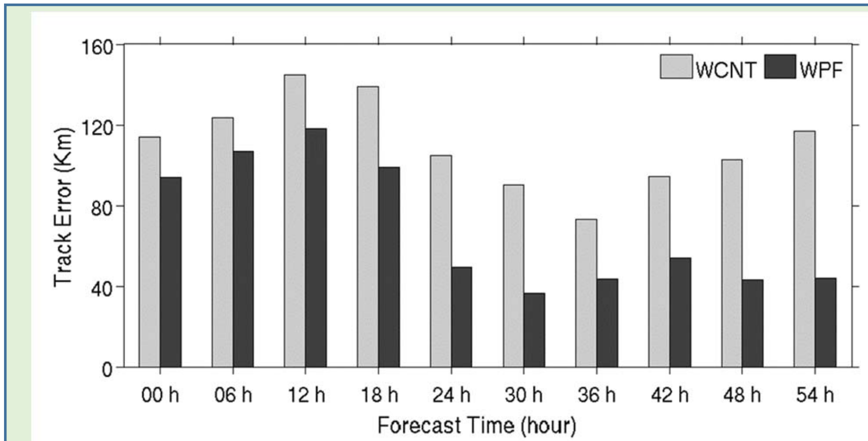


Simulated TC Landfall Error is better than IMD (14.7N, 80.0E) & SCORPIO (15.2N, 80.0E) predicted operational track forecasts from 00 UTC 10Dec2016.

(JGR-Atmosphere, 2019)

Track of the storm center from WCNT (blue line), WPF (red line) experiments along with IMD observed best track (black line)





Six-hourly track errors in the simulated cyclone track (in kilometers).