Application of Particle Filter on Data Assimilation

Prashant Kumar
Scientist-SF
Atmospheric Sciences Division
Space Applications Centre, ISRO
prashant22@sac.isro.gov.in

Who uses Satellite Data ?

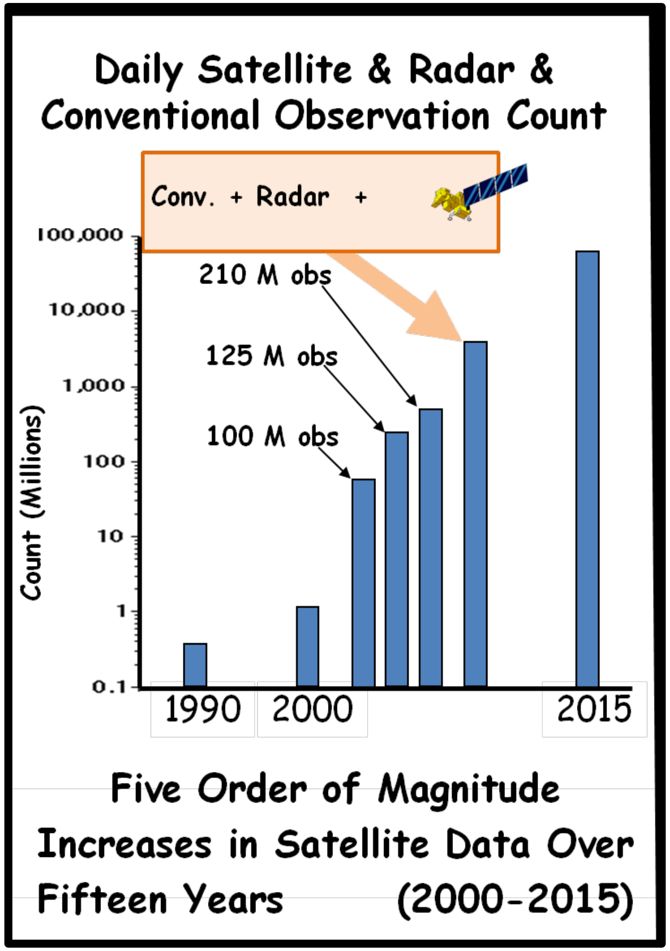
NWP models use most data from met-satellites through Assimilation

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

Who uses Satellite Data ?

NWP models use most data from met-satellites through Assimilation

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

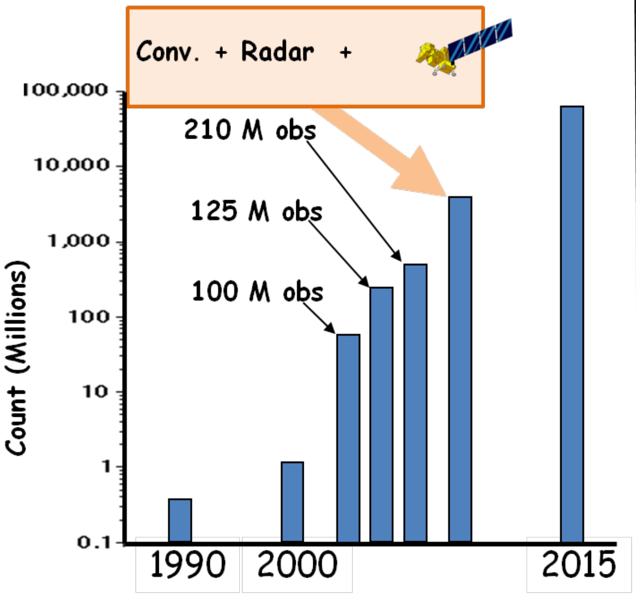


Who uses Satellite Data ?

NWP models use most data from met-satellites through Assimilation

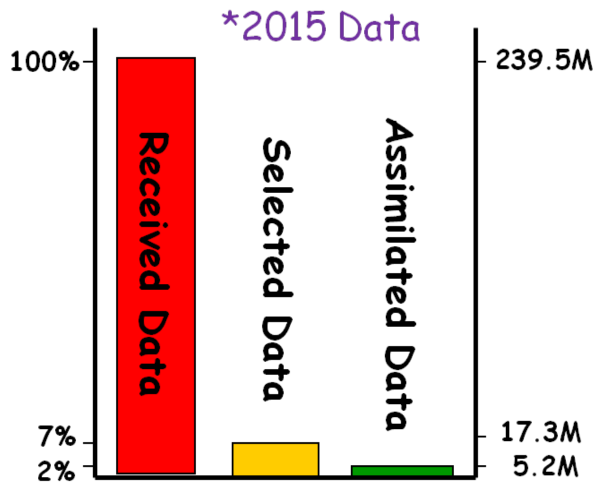
Today, more than 95% observations for weather prediction are provided by satellites

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



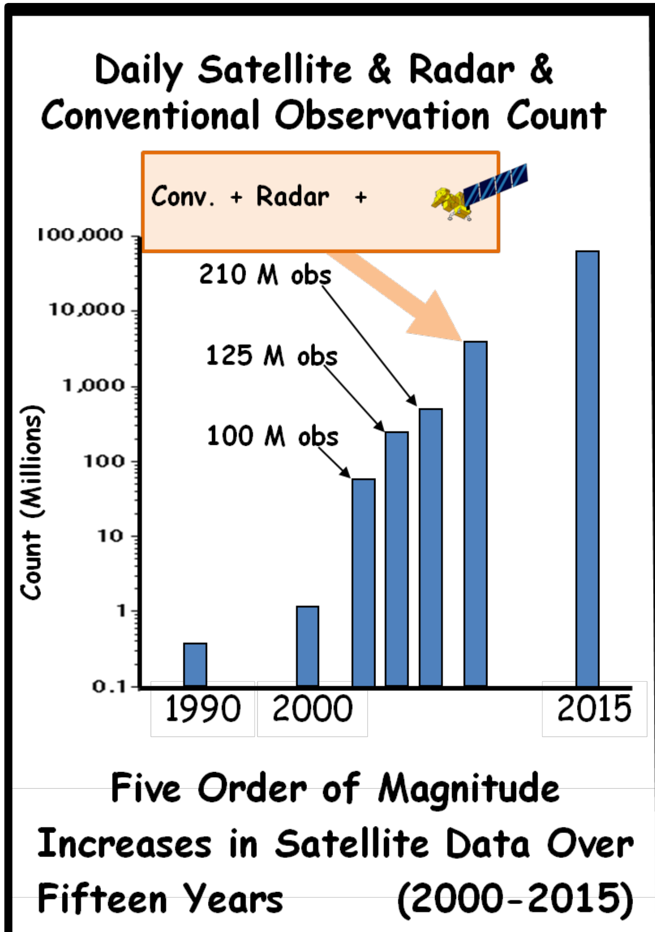
Received = All observations received operationally from providers
Selected = Observations selected as suitable for use
Assimilated = Observations actually used by models

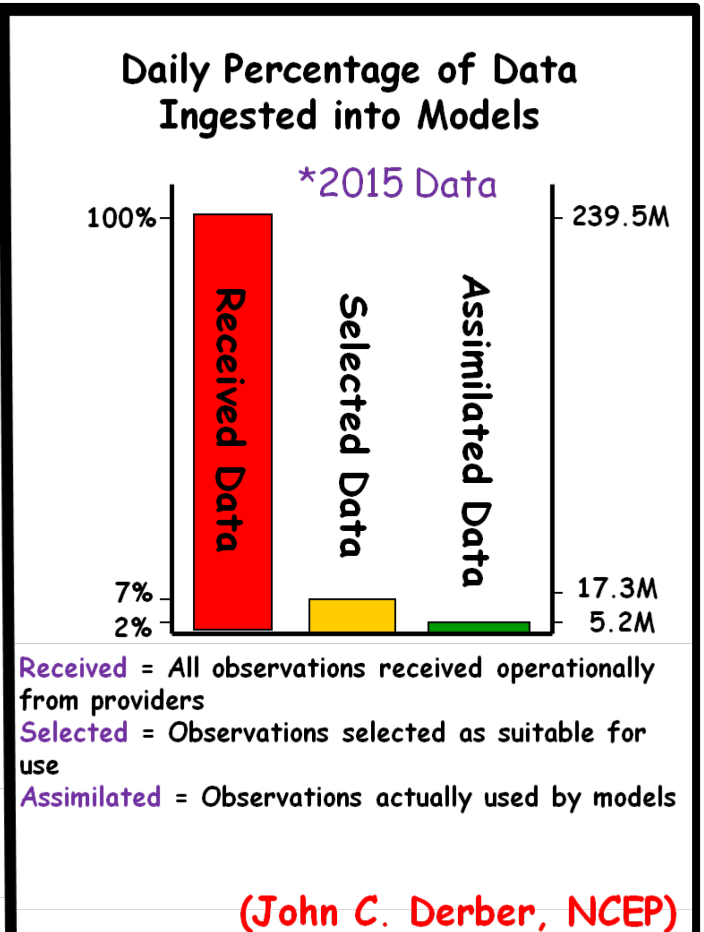
(John C. Derber, NCEP)

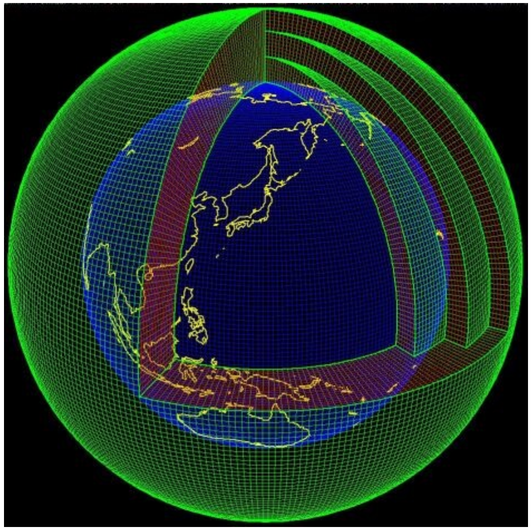
Who uses Satellite Data ?

NWP models use most data from met-satellites through Assimilation

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

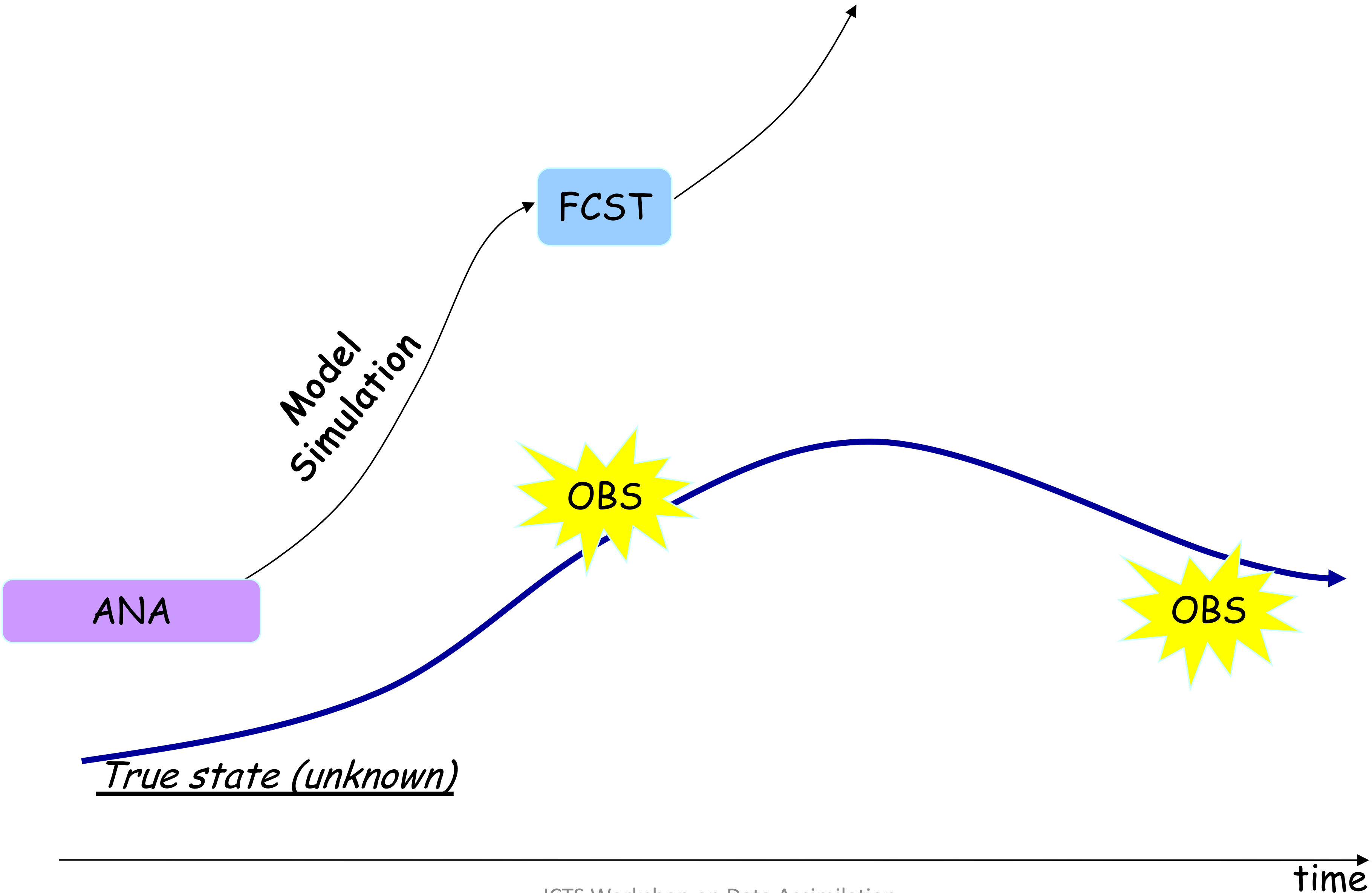






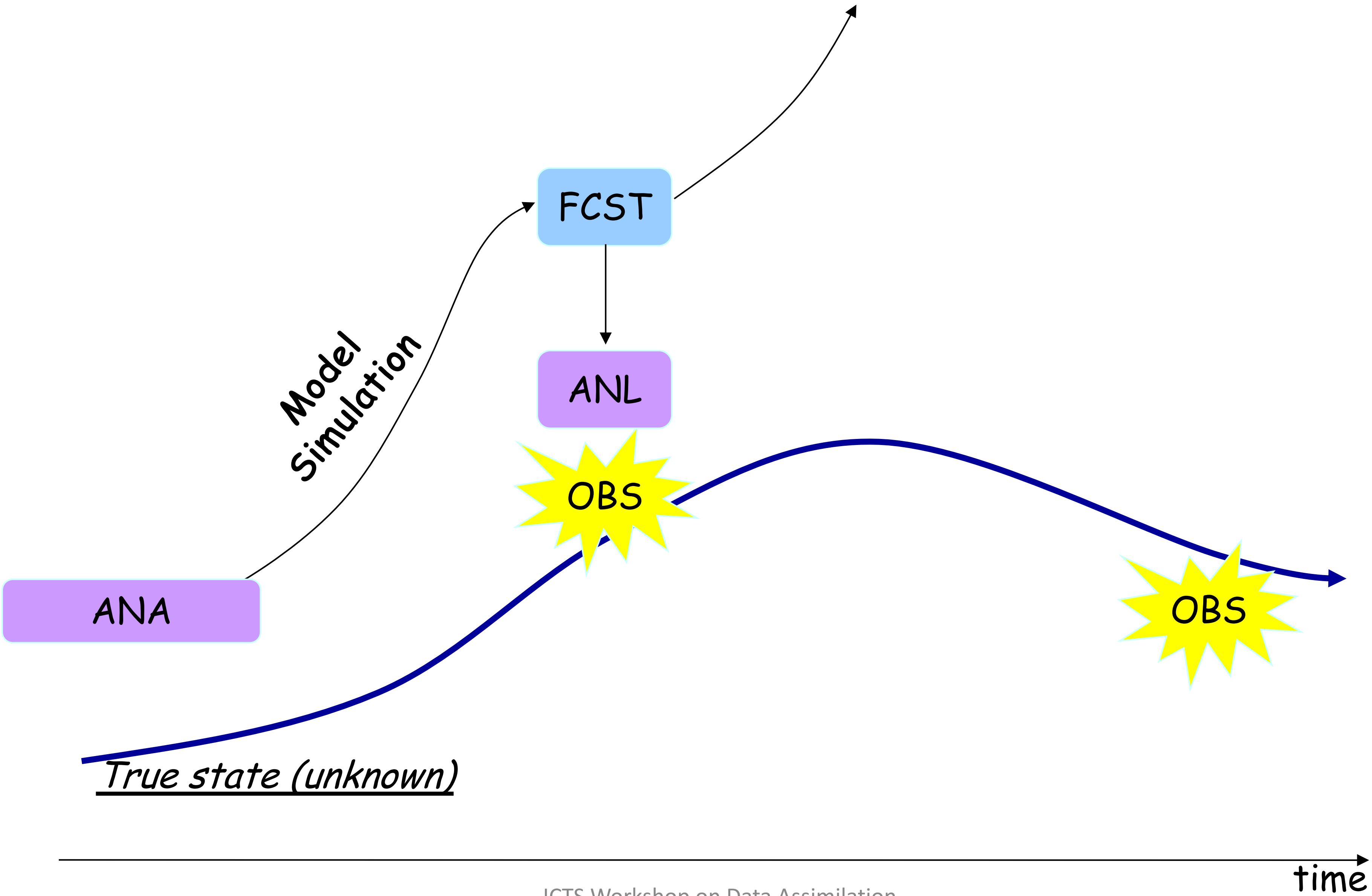
THIS IS ONLY 5-7% OF WHAT SATELLITES OBSERVE !!

Data Assimilation in NWP



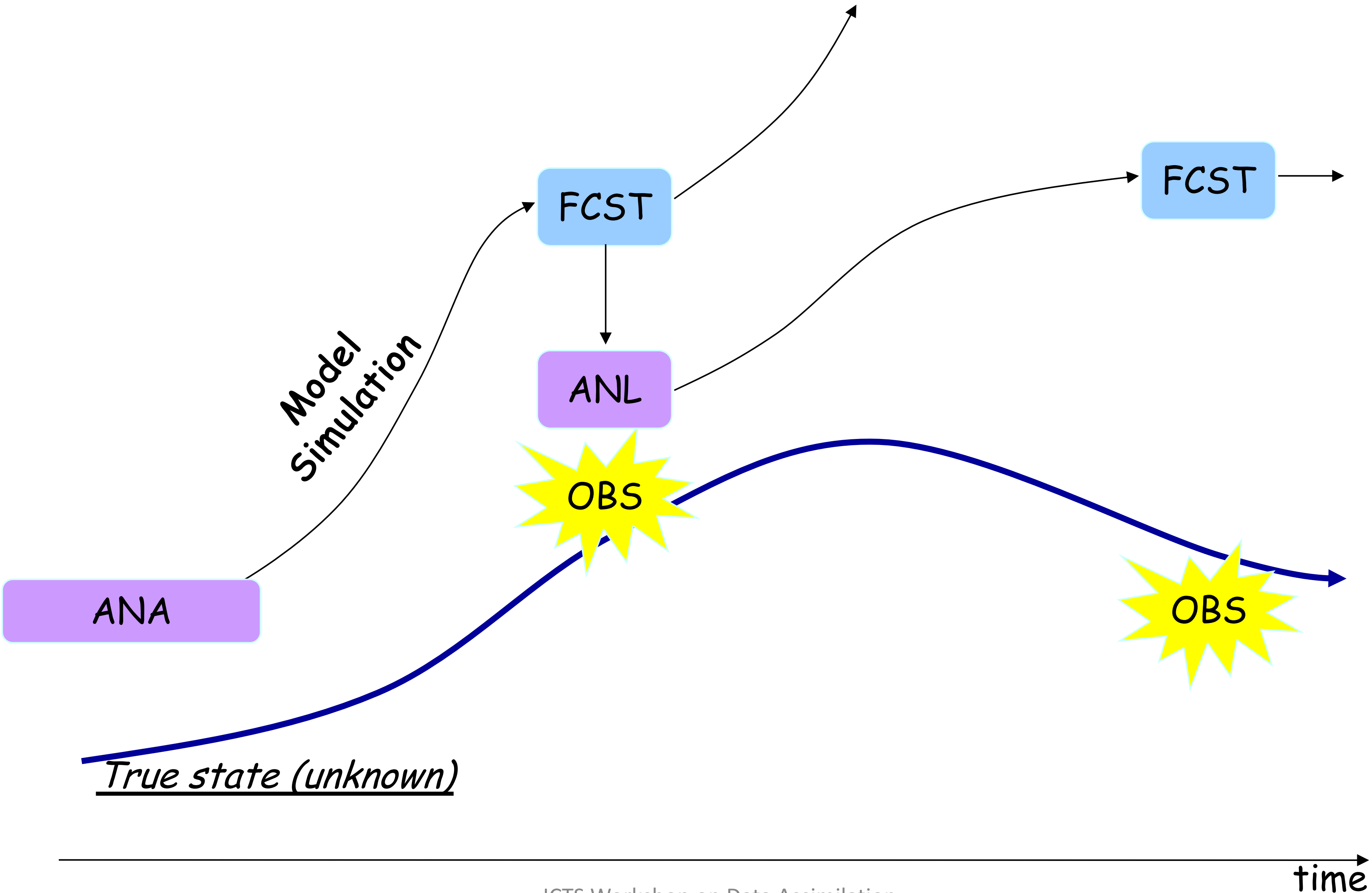
Source: Miyoshi, RIKEN

Data Assimilation in NWP



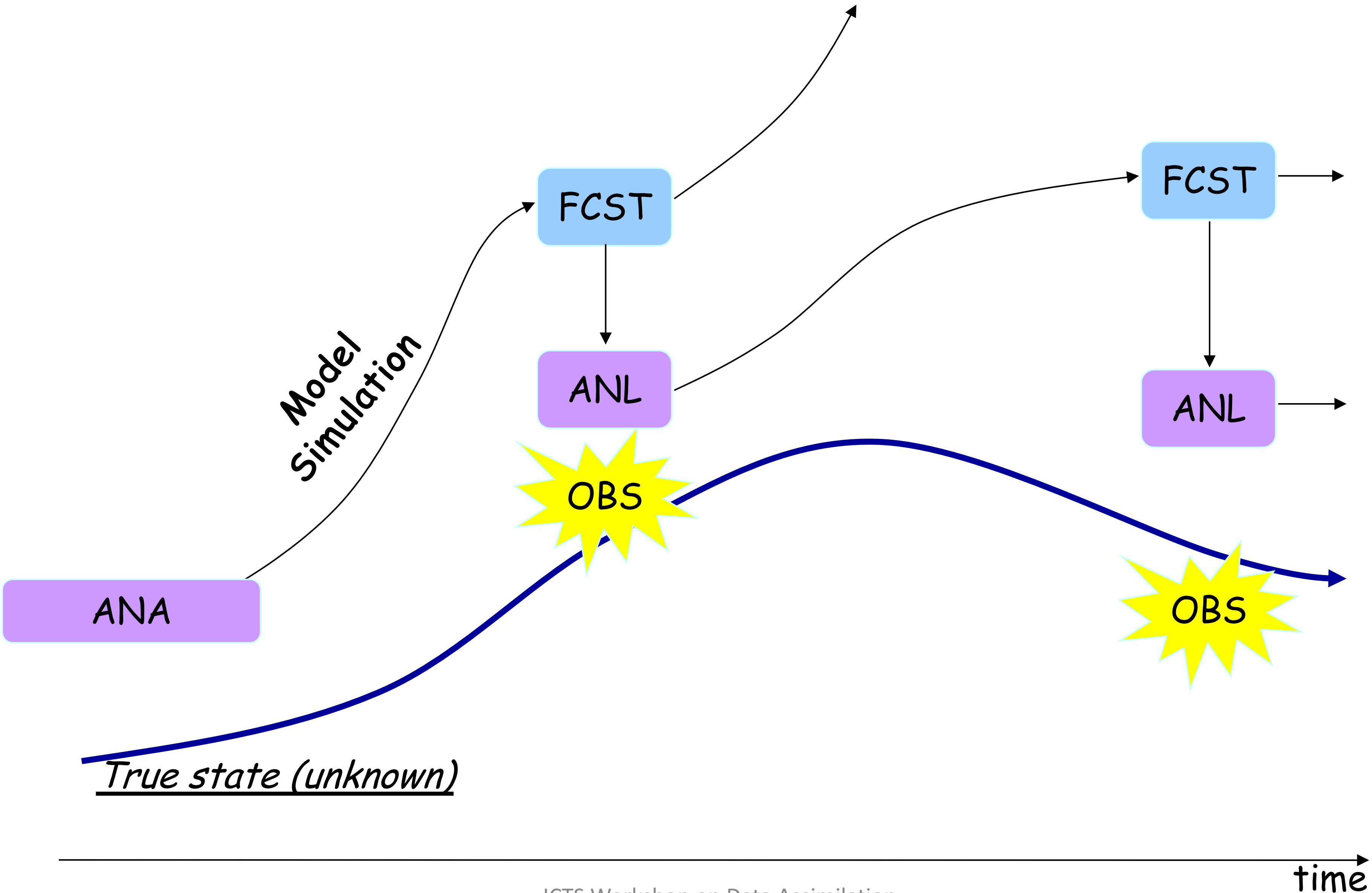
Source: Miyoshi, RIKEN

Data Assimilation in NWP



Source: Miyoshi, RIKEN

Data Assimilation in NWP

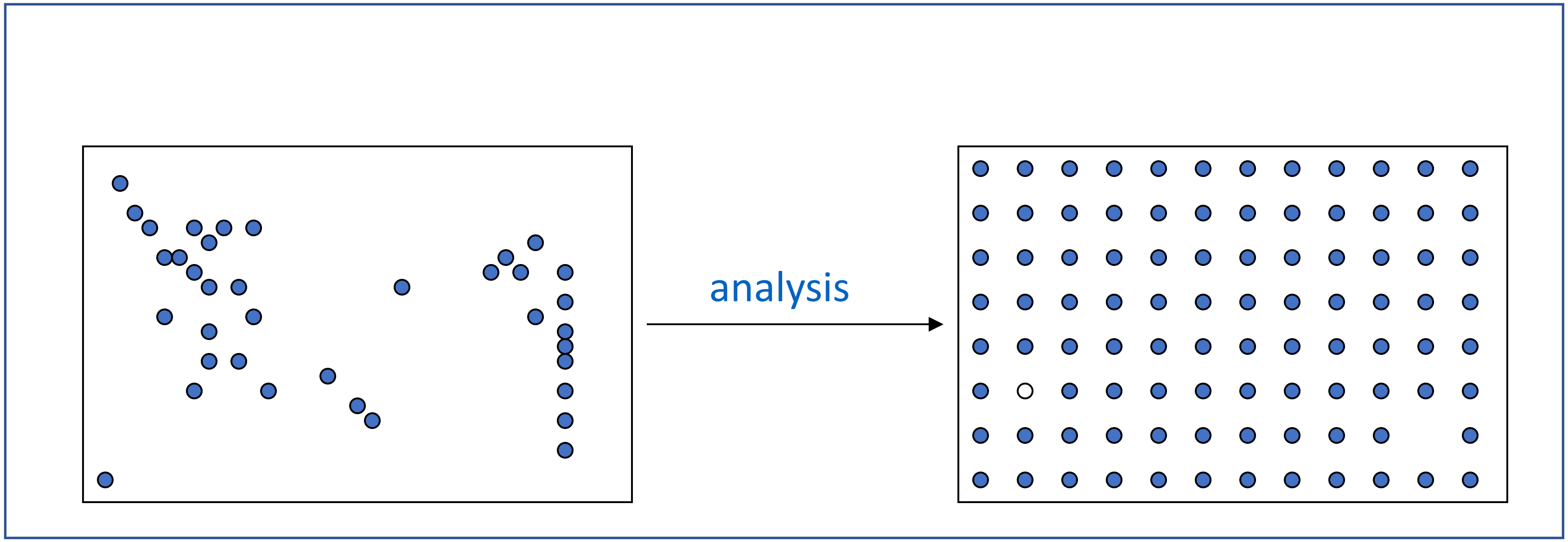


Source: Miyoshi, RIKEN

Aim of Data Assimilation

Too few and irregular (in space and time) observations, too many model grids.

Goal: To produce a regular, physically consistent best state of the atmosphere or system



(Daley, 1991)

Data Assimilation Techniques

✓ Objective Analysis

Cost of Computation

Data Assimilation Techniques

- ✓ Objective Analysis
- ✓ Optimum Interpolation (OI)

Cost of Computation

Data Assimilation Techniques

- ✓ Objective Analysis
- ✓ Optimum Interpolation (OI)
- ✓ Variation Method: 3 D-Var

  4 D-Var

Cost of Computation

Data Assimilation Techniques

- ✓ Objective Analysis
- ✓ Optimum Interpolation (OI)
- ✓ Variation Method: 3 D-Var

  4 D-Var
- ✓ Kalman Filter (e.g. EnKF, **LETKF**)

Hybrid Methods

(e.g. 3D/4DEnVar)

Cost of Computation

Data Assimilation Techniques

- ✓ Objective Analysis
- ✓ Optimum Interpolation (OI)
- ✓ Variation Method: 3 D-Var

  4 D-Var
- ✓ Kalman Filter (e.g. EnKF, **LETKF**)
- ✓ **Particle Filter**
- ✓ …..

Hybrid Methods

(e.g. 3D/4DEnVar)

Cost of Computation

Data Assimilation Techniques

- ✓ Objective Analysis
- ✓ Optimum Interpolation (OI)
- ✓ Variation Method: 3 D-Var
  4 D-Var
- ✓ Kalman Filter (e.g. EnKF, LETKF)
- ✓ Particle Filter
- ✓ .....

Hybrid Methods
(e.g. 3D/4DEnVar)

Cost of Computation

Modern Era: Beyond Linear & Gaussian Constraints.

Sources of errors

- Errors in the First Guess

- Errors in the model

  (Numerical Approx., Physics, etc.)

- Errors in Observations

Sources of errors

- ❑ Errors in the First Guess

- ❑ Errors in the model

  (Numerical Approx., Physics, etc.)

- ❑ Errors in Observations

- ✓ Limitations of Data Assimilation Techniques
- ✓ Errors can be random and/or systematic errors
- ✓ Intrinsic predictability limitations

Issues in schemes

Example 1: Add datatype with unrealistic low ***R***.

OSE/OSSE shows negative impact.

Adjoint sensitivity shows this datatype contributes most to forecast error reduction.

(CGMS, 2018)

Issues in schemes

Example 1: Add datatype with unrealistic low ***R***.

OSE/OSSE shows negative impact.

Adjoint sensitivity shows this datatype contributes most to forecast error reduction.

Adjoint sensitivity measures impact in this setup.

(CGMS, 2018)

Issues in schemes

Example 1: Add datatype with unrealistic low ***R***.

OSE/OSSE shows negative impact.

Adjoint sensitivity shows this datatype contributes most to forecast error reduction.

Adjoint sensitivity measures impact in this setup.

Example 2: Add two identical datatypes, first one, then other.

OSE show first has large impact, second small.

Adjoint sensitivity shows they have the same impact.

(CGMS, 2018)

Issues in schemes

Example 1: Add datatype with unrealistic low ***R***.

OSE/OSSE shows negative impact.

Adjoint sensitivity shows this datatype contributes most to forecast error reduction.

Adjoint sensitivity measures impact in this setup.

Example 2: Add two identical datatypes, first one, then other.

OSE show first has large impact, second small.

Adjoint sensitivity shows they have the same impact.

OSEs are sensitive to order of changes.

(CGMS, 2018)

Issues in schemes

Example 1: Add datatype with unrealistic low ***R***.

OSE/OSSE shows negative impact.

Adjoint sensitivity shows this datatype contributes most to forecast error reduction.

Adjoint sensitivity measures impact in this setup.

Example 2: Add two identical datatypes, first one, then other.

OSE show first has large impact, second small.

Adjoint sensitivity shows they have the same impact.

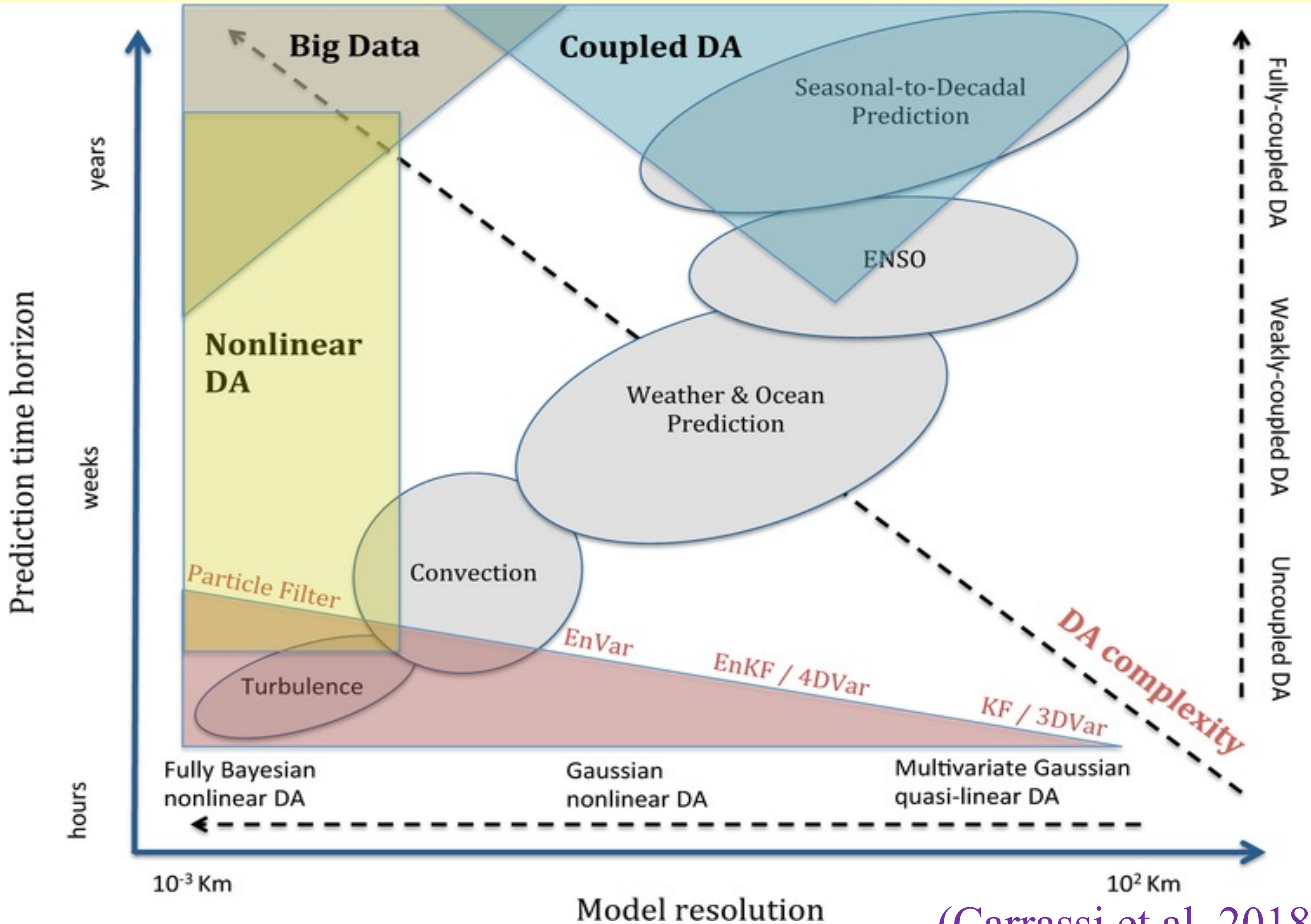
OSEs are sensitive to order of changes.

Both are correct!

But both are open to misinterpretation.

(CGMS, 2018)

DA method versus Resolution and Prediction Time



(Carrassi et al. 2018)

Conditional Probability Distribution:

Conditional Probability Distribution:

Background: $$\mathbf{P(x/x_b)} = \frac{1}{(2\Pi)^{n/2}\left|B\right|}\ e^{\frac{-1}{2}\left[\left(x-x_b\right)^T B^{-1}\left(x-x_b\right)\right]}$$

Conditional Probability Distribution:

Background: $\mathbf{P(x/x_b)} = \dfrac{1}{(2\Pi)^{n/2}|B|} e^{\frac{-1}{2}\left[(x-x_b)^T B^{-1}(x-x_b)\right]}$

Observations: $\mathbf{P(y/x)} = \dfrac{1}{(2\Pi)^{m/2}|R|} e^{\frac{-1}{2}\left[(y-Hx_t)^T R^{-1}(y-Hx_t)\right]}$

$$where \quad m < n$$

Since the above two distribution are mutually exclusive (independent), their joint Probability Distribution is the product of the two Gaussian distribution

Since the above two distribution are mutually exclusive (independent), their joint Probability Distribution is the product of the two Gaussian distribution

Joint Probability Distribution

$$L(x) \propto P(x/x_b) * P(y/x)$$

Since the above two distribution are mutually exclusive (independent), their joint Probability Distribution is the product of the two Gaussian distribution

Joint Probability Distribution

$$L(x) \propto P(x/x_b) * P(y/x)$$

Original problem is multiplication !

Since the above two distribution are mutually exclusive (independent), their joint Probability Distribution is the product of the two Gaussian distribution

Joint Probability Distribution

$$L(x) \propto P(x/x_b) * P(y/x)$$

Original problem is multiplication !

Bayes Theorem tells that data assimilation is a multiplication problem: given the prior and the likelihood, the solution is the point-wise multiplication of the two.

Cost Function is defined as the M.L. estimate of the product of two exponential as,

$$J(x) = \frac{1}{2}(x - x_b)^T B^{-1}(x - x_b) + \frac{1}{2}(y - Hx)^T R^{-1}(y - Hx)$$

Minimizing the cost function:

$$\partial J / \partial x = 0$$

$$B^{-1}(x - x_b) - H^T R^{-1}(y - Hx) = 0$$

Cost Function is defined as the M.L. estimate of the product of two exponential as,

$$J(x) = \frac{1}{2}(x - x_b)^T B^{-1}(x - x_b) + \frac{1}{2}(y - Hx)^T R^{-1}(y - Hx)$$

Minimizing the cost function:

$$\partial J / \partial x = 0$$

$$B^{-1}(x - x_b) - H^T R^{-1}(y - Hx) = 0$$

But, constraints simplify it as Linear optimization problem

Drawbacks of OI, 3D-Var, 4D-Var

✓ Background error Covariances (B) update is omitted

✓ Ideally the B should be flow-dependent
(e.g. Error of the day)

Drawbacks of OI, 3D-Var, 4D-Var

- ✓ Background error Covariances (B) update is omitted
- ✓ Ideally the B should be flow-dependent (e.g. Error of the day)

Kalman Filter (Computationally Expensive)

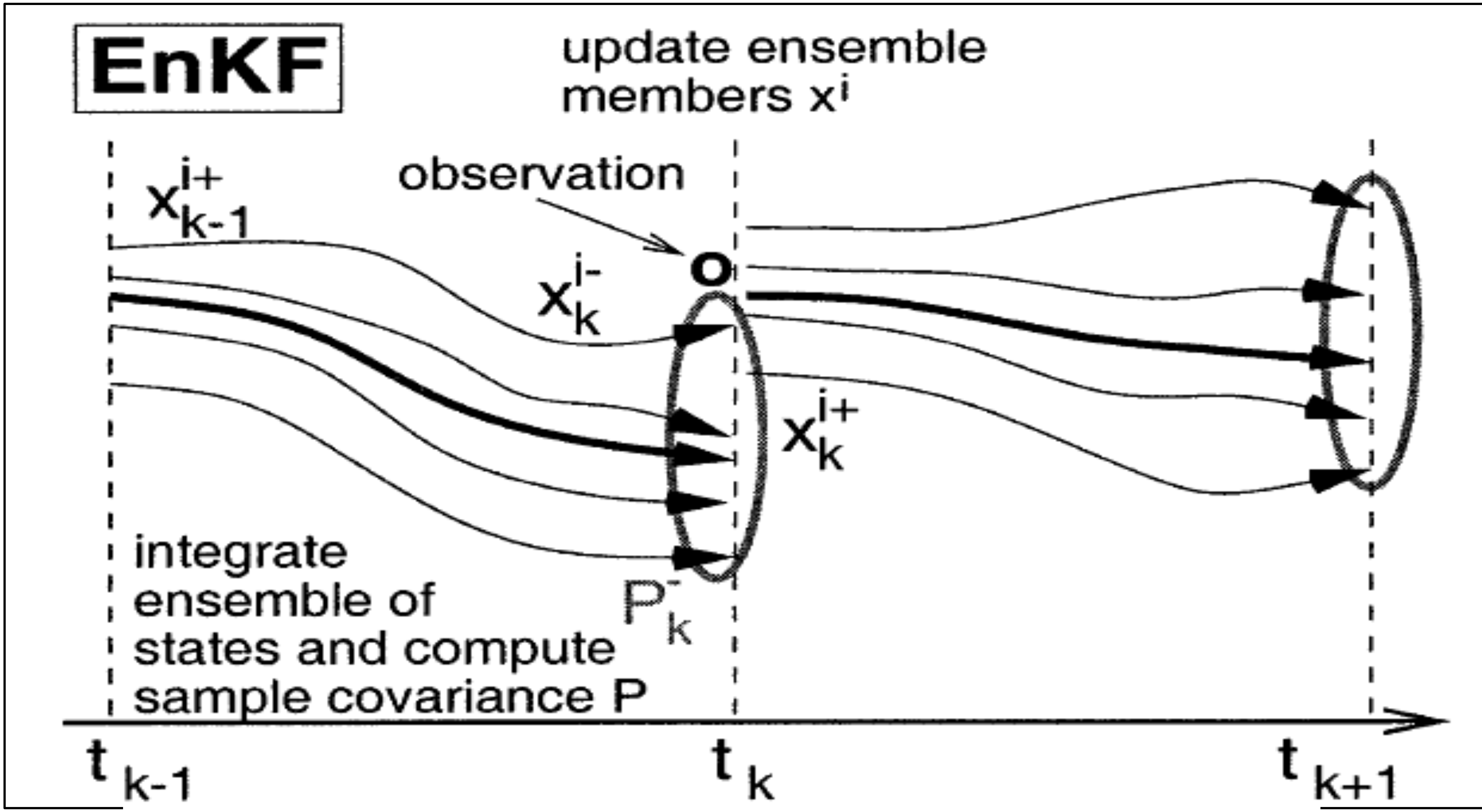
- ✓ Use Model equations to propagate B forward in time.

$$B \longrightarrow B(t)$$

	$\vec{u}$	$\vec{v}$	$\vec{\theta}$	$\vec{p}$	$\vec{q}$
$\vec{u}$	B_{uu}	B_{uv}	$B_{u\theta}$	B_{up}	B_{uq}
$\vec{v}$	B_{uv}	B_{vv}	$B_{v\theta}$	B_{vp}	B_{vq}
$\vec{\theta}$	$B_{u\theta}$	$B_{v\theta}$	$B_{\theta\theta}$	$B_{\theta p}$	$B_{\theta q}$
$\vec{p}$	B_{up}	B_{vp}	$B_{\theta p}$	B_{pp}	B_{pq}
$\vec{q}$	B_{uq}	B_{vq}	$B_{\theta q}$	B_{pq}	B_{qq}

- ✓ Analysis step as in OI

Ensemble Kalman Filter (Possible)



$$P^f_{k+1}(N) = E\left[e^f_{k+1}\left(e^f_{k+1}\right)^T\right] = \frac{1}{N-1}\sum_{i=1}^{N}\left[e^f_{k+1}(i)\left[e^f_{k+1}(i)\right]^T\right]$$

Better than EKF; No linearization here

Issues in Ensemble Kalman Filter

The success of the EnKF depends on size of the ensembles to give an accurate estimate of the sample mean and covariance.

Issues in Ensemble Kalman Filter

The success of the EnKF depends on size of the ensembles to give an accurate estimate of the sample mean and covariance.

But for large scale problems, ensemble under sampling cause major problems:

Issues in Ensemble Kalman Filter

The success of the EnKF depends on size of the ensembles to give an accurate estimate of the sample mean and covariance.

But for large scale problems, ensemble under sampling cause major problems:

- ✓ underestimated ensemble variance,
- ✓ filter divergence,
- ✓ errors in estimated correlations, in particular Spurious long-range correlations

Possible solutions

1. Use more ensemble members (see Miyoshi et al. 2014)
How many ensembles are sufficient ?

Possible solutions

1. Use more ensemble members (see Miyoshi et al. 2014)

How many ensembles are sufficient ?

Who Knows?

Possible solutions

1. Use more ensemble members (see Miyoshi et al. 2014)

How many ensembles are sufficient ?

Who Knows?

2. Localization

- Addresses problem of spurious correlations
- Increases the effective ensemble size

Possible solutions

1. Use more ensemble members (see Miyoshi et al. 2014)

How many ensembles are sufficient ?

Who Knows?

2. Localization

- Addresses problem of spurious correlations
- Increases the effective ensemble size

3. Inflation

- Addresses problem of filter divergence

Possible solutions

1. Use more ensemble members (see Miyoshi et al. 2014)
How many ensembles are sufficient ?
Who Knows?

2. Localization
- Addresses problem of spurious correlations
- Increases the effective ensemble size

3. Inflation
- Addresses problem of filter divergence

4. Combine ensemble with variational approaches

Possible solutions

1. Use more ensemble members (see Miyoshi et al. 2014)

How many ensembles are sufficient ?

Who Knows?

2. Localization

- Addresses problem of spurious correlations
- Increases the effective ensemble size

3. Inflation

- Addresses problem of filter divergence

4. Combine ensemble with variational approaches

- These are known as hybrid methods

(Operational Techniques)

Beyond Kalman filtering

✓ All EnKF methods (~ 30-100) using Gaussian assumption.

Beyond Kalman filtering

✓ All EnKF methods (~ 30-100) using Gaussian assumption.

Is this a valid assumption when the dynamical model is non-linear?

Beyond Kalman filtering

✓ All EnKF methods (~ 30-100) using Gaussian assumption.

Is this a valid assumption when the dynamical model is non-linear?

✓ The optimality of the Kalman filters is no longer preserved when applied to non-linear problems.

✓ This is a common issue for all ensemble filters whose analysis step is based on the equations of the Kalman filter.

Beyond Kalman filtering

✓ All EnKF methods (~ 30-100) using Gaussian assumption.

Is this a valid assumption when the dynamical model is non-linear?

✓ The optimality of the Kalman filters is no longer preserved when applied to non-linear problems.

✓ This is a common issue for all ensemble filters whose analysis step is based on the equations of the Kalman filter.

✓ **This is a very hot topic!**

And what if the distributions are non-Gaussian?

Fully non-Gaussian data-assimilation methods exist, but none has proven to work for large-scale systems.

And what if the distributions are non-Gaussian?

Fully non-Gaussian data-assimilation methods exist, but none has proven to work for large-scale systems.

A few do have potential among which the Particle Filter.

And what if the distributions are non-Gaussian?

Fully non-Gaussian data-assimilation methods exist, but none has proven to work for large-scale systems.

A few do have potential among which the Particle Filter.

It has been shown that the ‘curse of dimensionality’ has been cured….

How does Gaussianity comes into DA

Is non–Gaussianity relevant in DA

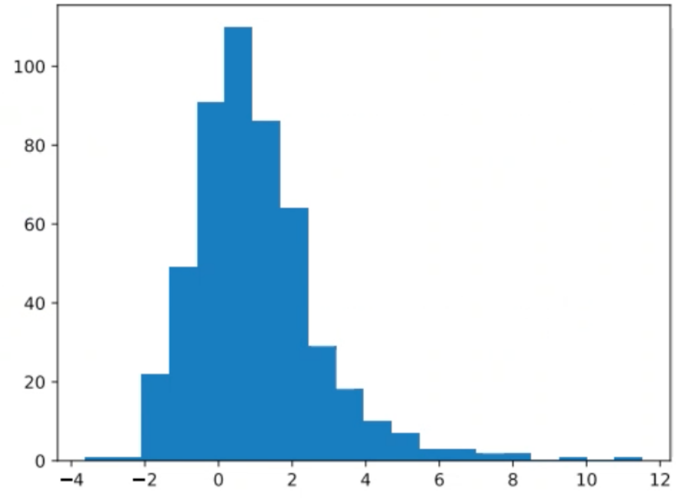
How does Gaussianity comes into DA

Reasons for non–Gaussianity: Nonlinear models

If X_n is Gaussian, then $X_{n+1} = \mathcal{M}(X_n) + R_{n+1}$ is generally not Gaussian as soon as $\mathcal{M}$ is nonlinear (whether R is Gaussian or not).

Example

$$X_{n+1} = X_n^2 + R_{n+1}, \qquad X_n \text{ and } R_{n+1} \text{ standard Gaussian.}$$



How does Gaussianity comes into DA

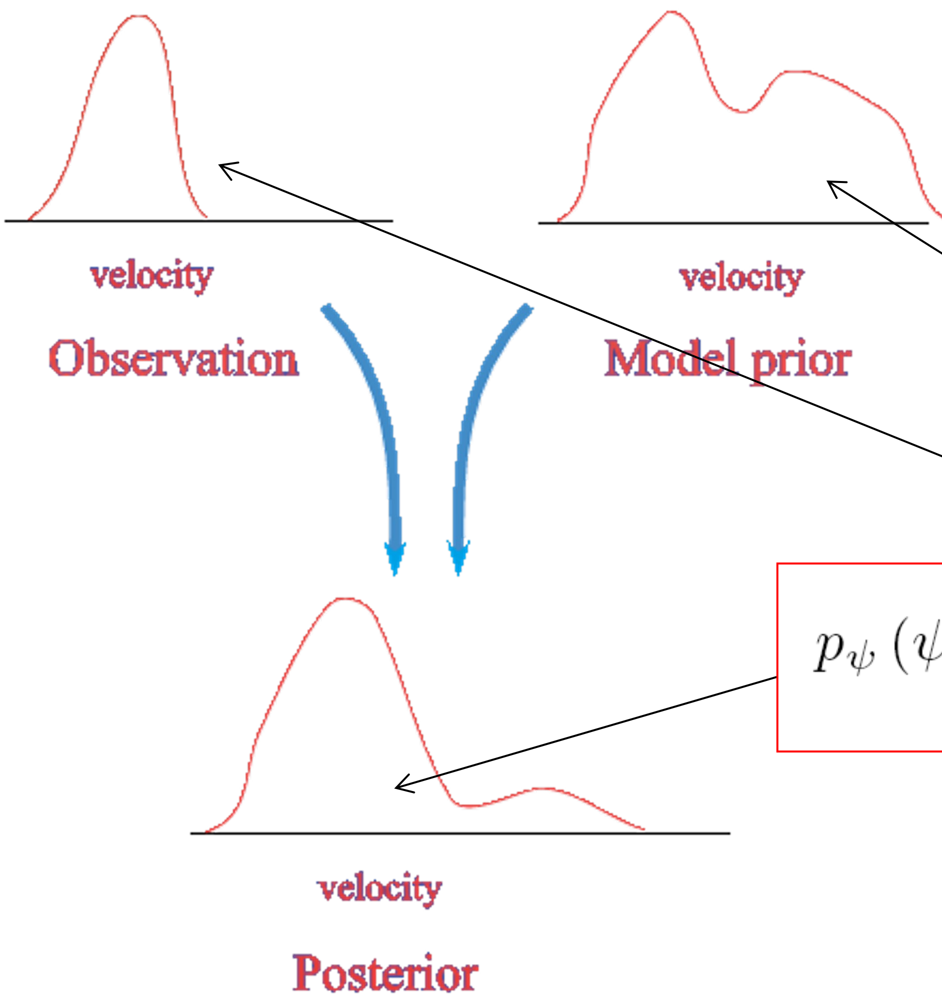
Reasons for non–Gaussianity: Nonlinear observations

If X_n is Gaussian and $Y_n = \mathcal{H}(X_n) + S_n$, then $X_n | Y_n$ is generally not Gaussian as soon as $\mathcal{H}$ is nonlinear (whether S is Gaussian or not).

Example

$$Y_n = X_n^2 + S_n, \qquad X_n \text{ and } S_n \text{ standard normal.}$$

Data assimilation: general formulation



Bayes theorem:

$$p_\psi\left(\psi|d\right) = \frac{p_d\left(d|\psi\right)p_\psi\left(\psi\right)}{\int p_d\left(d|\psi\right)p_\psi\left(\psi\right)\;d\psi}$$

Peter Leeuwen

How do we get there? Particle filter?

$$p_\psi(\psi|d) = \frac{p_d(d|\psi)\, p_\psi(\psi)}{\int p_d(d|\psi)\, p_\psi(\psi)\ d\psi}$$

↓ Use ensemble

How do we get there? Particle filter?

$$p_\psi\left(\psi|d\right) = \frac{p_d\left(d|\psi\right)p_\psi\left(\psi\right)}{\int p_d\left(d|\psi\right)p_\psi\left(\psi\right)\;d\psi}$$

Use ensemble

$$p(\psi) = \sum_i \frac{1}{N}\delta(\psi - \psi_i)$$

How do we get there? Particle filter?

$$p_\psi(\psi|d) = \frac{p_d(d|\psi)\, p_\psi(\psi)}{\int p_d(d|\psi)\, p_\psi(\psi)\ d\psi}$$

Use ensemble

$$p(\psi) = \sum_i \frac{1}{N}\delta(\psi - \psi_i)$$

$$p_\psi(\psi|d) = \sum_i w_i \delta(\psi - \psi_i)$$

How do we get there? Particle filter?

$$p_\psi(\psi|d) = \frac{p_d(d|\psi)\, p_\psi(\psi)}{\int p_d(d|\psi)\, p_\psi(\psi)\; d\psi}$$

Use ensemble

$$p(\psi) = \sum_i \frac{1}{N}\delta(\psi - \psi_i)$$

$$p_\psi(\psi|d) = \sum_i w_i \delta(\psi - \psi_i)$$

with

$$w_i = \frac{p_d(d|\psi_i)}{\sum_i p_d(d|\psi_i)}$$

the weights.

How do we get there? Particle filter?

$$p_\psi\left(\psi|d\right) = \frac{p_d\left(d|\psi\right)p_\psi\left(\psi\right)}{\int p_d\left(d|\psi\right)p_\psi\left(\psi\right)\ d\psi}$$

Use ensemble

$$p(\psi) = \sum_i \frac{1}{N}\delta(\psi - \psi_i)$$

$$p_\psi(\psi|d) = \sum_i w_i \delta(\psi - \psi_i)$$

with

$$w_i = \frac{p_d\left(d|\psi_i\right)}{\sum_i p_d\left(d|\psi_i\right)}$$

the weights.

No Linear Assumption, Fully Non-linear,

No need of TL & AD Model,

FG is not change (Balence)

What are these weights?

- The weight w_i is the normalised value of the pdf of the observations given model state x_i .

What are these weights?

- The weight w_i is the normalised value of the pdf of the observations given model state x_i.
- For Gaussian distributed variables it is given by:

$$
\begin{aligned}
w_i &\propto p(y|x_i) \\
&\propto \exp\left[-\frac{1}{2}\left(y - H(x_i)\right) R^{-1} \left(y - H(x_i)\right)\right]
\end{aligned}
$$

No explicit need for state covariances

- 3DVar and 4DVar need a good error covariance of the prior state estimate: complicated

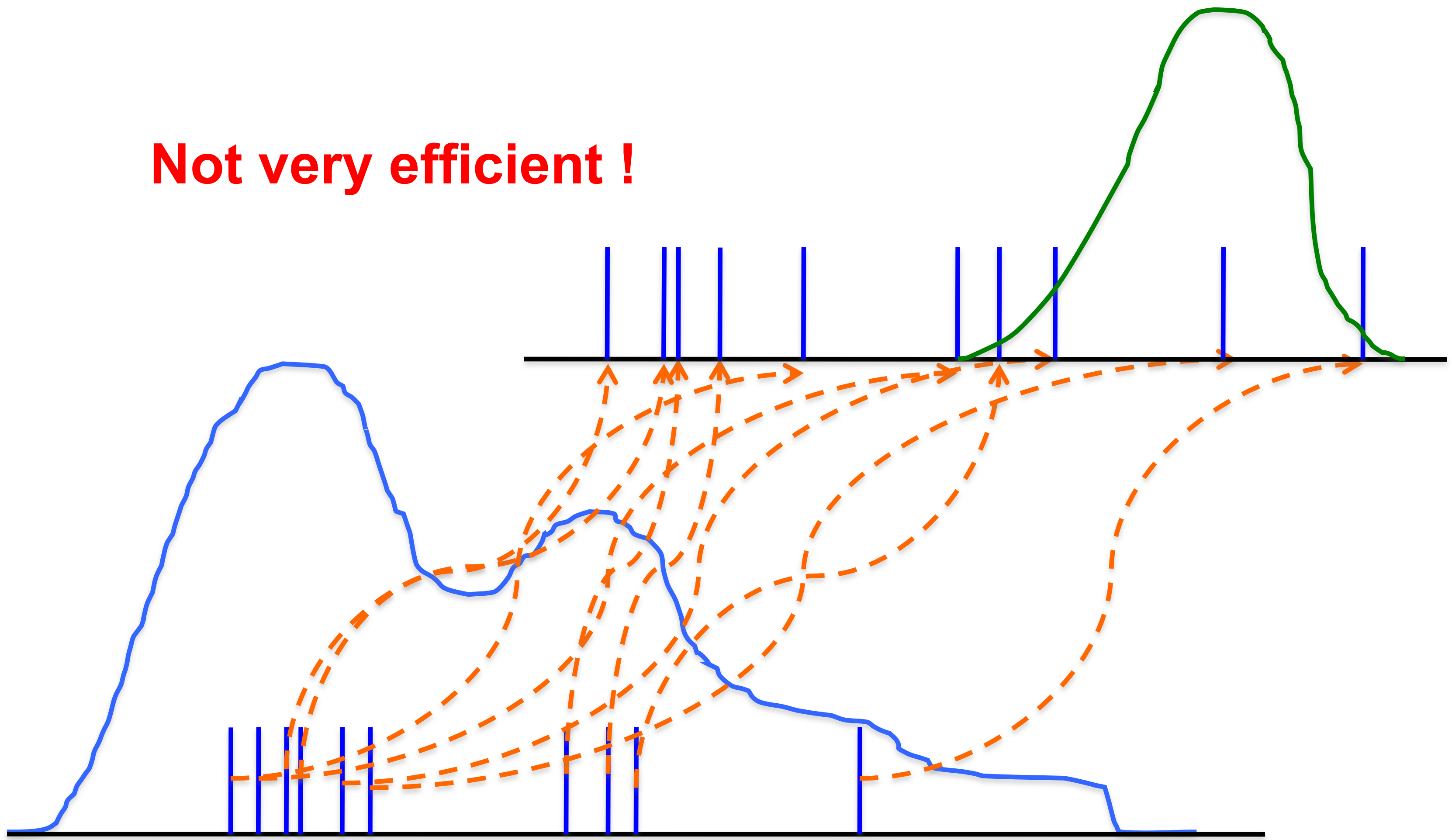
No explicit need for state covariances

- 3DVar and 4DVar need a good error covariance of the prior state estimate: complicated
- The performance of Ensemble Kalman filters relies on the quality of the sample covariance, forcing artificial inflation and localisation.

No explicit need for state covariances

- 3DVar and 4DVar need a good error covariance of the prior state estimate: complicated
- The performance of Ensemble Kalman filters relies on the quality of the sample covariance, forcing artificial inflation and localisation.
- Particle filter doesn't have this problem, but…

Standard Particle filter



The standard particle filter is degenerate for moderate ensemble size in moderate-dimensional systems.

Particle Filter degeneracy: resampling

- **With each new set of observations the old weights are multiplied with the new weights.**

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…

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…**
- **Solution:**

Resampling : duplicate high-weight particles, proposal density, etc.

However: degeneracy

- **For large-scale problems with lots of observations SIR method is still degenerate:**

However: degeneracy

- **For large-scale problems with lots of observations SIR method is still degenerate:**
- **Only a few particles get high weights; the other weights are negligibly small.**

However: degeneracy

- For large-scale problems with lots of observations SIR 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.

However: degeneracy

- For large-scale problems with lots of observations SIR method is still degenerate:

- Only a few particles get high weights; the other weights are negligibly small.

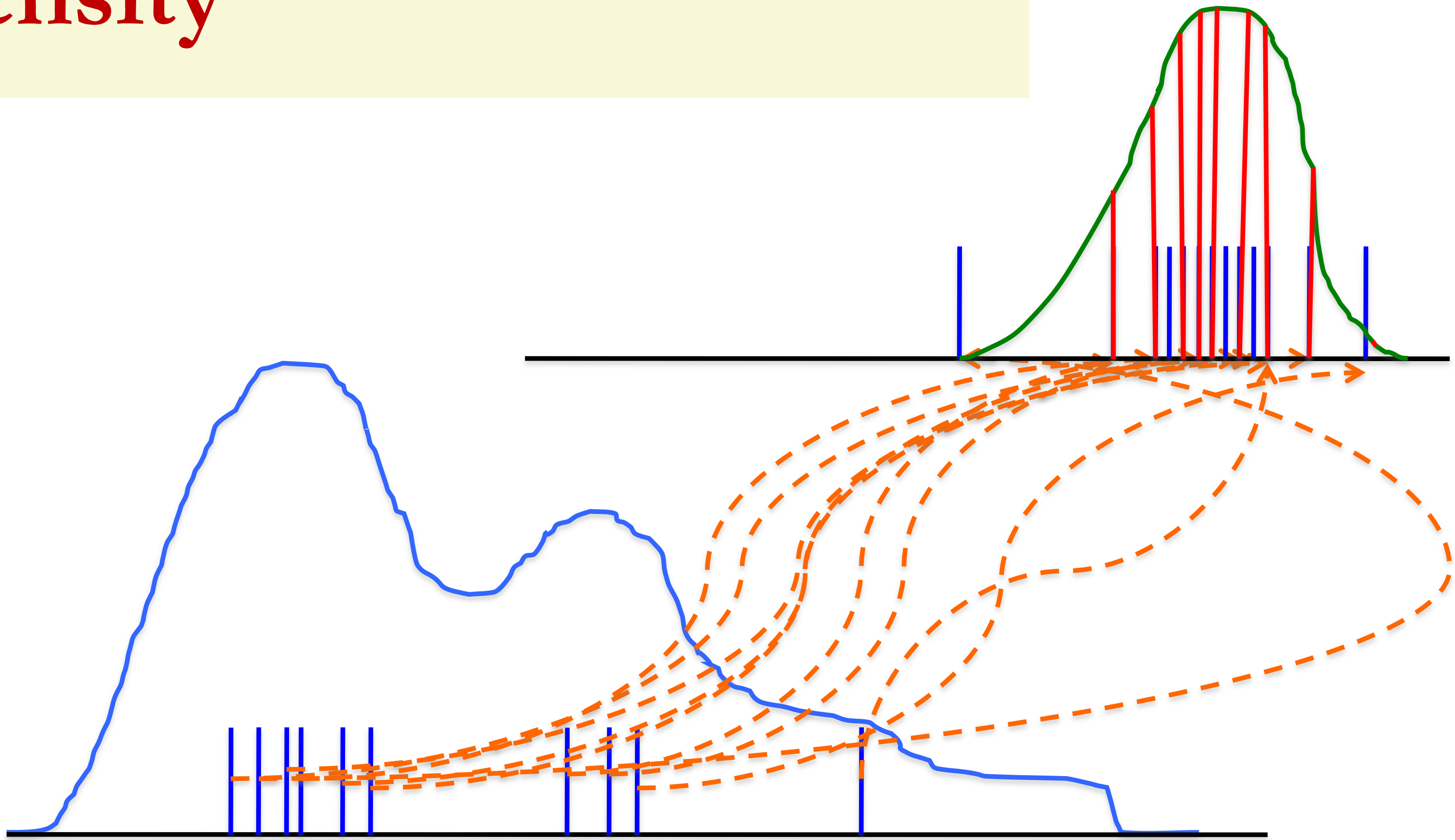
- However, we can enforce almost equal weight for all particles.

Equal-Weight Particle Filter

Frequent Filter Restart &

Local Particle Filter

Particle filter with proposal transition density



Particle Filter Assimilation of INSAT-3D TIR Channel All-Sky TB

JGR Atmospheres

Research Article | Free Access

Assimilating INSAT-3D Thermal Infrared Window Imager Observation With the Particle Filter: A Case Study for Vardah Cyclone

Prashant Kumar, Munn V. Shukla

First published: 28 January 2019 | https://doi.org/10.1029/2018JD028827

A big challenge in the satellite data assimilation is the effective use of InfraRed (IR) window channel radiances in the weather model.

A big challenge in the satellite data assimilation is the effective use of InfraRed (IR) window channel radiances in the weather model.

Generally not use for assimilation in NWP model

A big challenge in the satellite data assimilation is the effective use of InfraRed (IR) window channel radiances in the weather model.

Generally not use for assimilation in NWP model

A hybrid data-assimilation method is designed for very severe cyclonic storm "Vardah,"

A big challenge in the satellite data assimilation is the effective use of InfraRed (IR) window channel radiances in the weather model.

Generally not use for assimilation in NWP model

A hybrid data-assimilation method is designed for very severe cyclonic storm "Vardah,"

3D-Var method is used to assimilate control observations (Conv., Satellites, etc.)

A big challenge in the satellite data assimilation is the effective use of InfraRed (IR) window channel radiances in the weather model.

Generally not use for assimilation in NWP model

A hybrid data-assimilation method is designed for very severe cyclonic storm "Vardah,"

3D-Var method is used to assimilate control observations (Conv., Satellites, etc.)

Particle filter method is used to assimilate all-sky Thermal IR observations from INSAT-3D satellite.

Various particles (or ensembles) are prepared using diverse WRF model physics.

Various particles (or ensembles) are prepared using diverse WRF model physics.

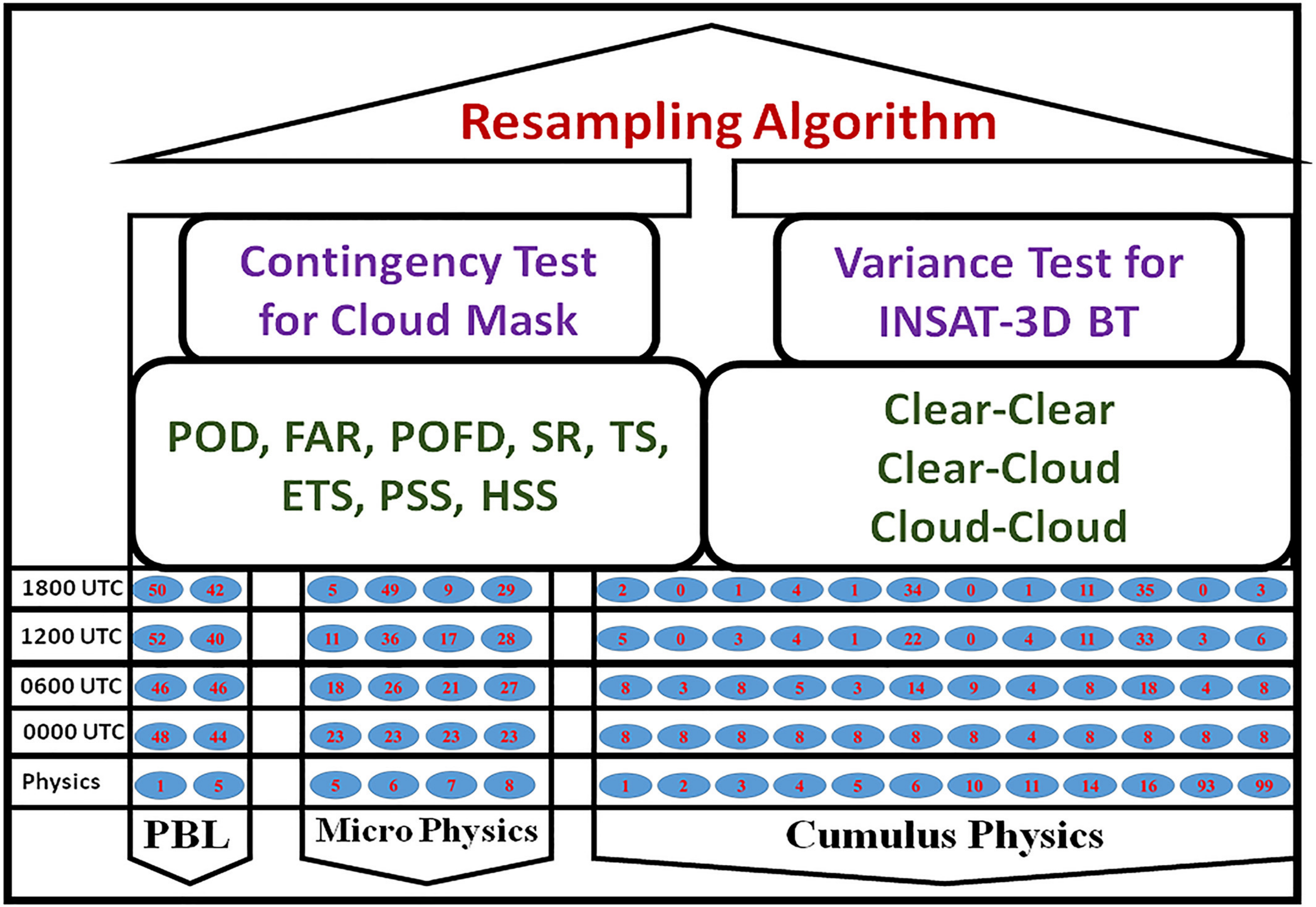
To implement particle filter, INSAT-3D thermal IR window channel 1 (TIR-1; center wavelength 11 μm) measured brightness temperature (BT) and cloud mask product are used to assign appropriate weights to different particles to reduce model uncertainties.

Various particles (or ensembles) are prepared using diverse WRF model physics.

To implement particle filter, INSAT-3D thermal IR window channel 1 (TIR-1; center wavelength 11 μm) measured brightness temperature (BT) and cloud mask product are used to assign appropriate weights to different particles to reduce model uncertainties.

This step is followed by **resampling step** in which new particles are generated from high weight particles using stochastic kinetic-energy backscatter scheme (SKEBS) method in which dynamical variables are perturbed into the model physics.

Schematic of INSAT-3D measured thermal infrared-1 BT and cloud-mask product assimilation using particle filter.



Resampling Step

- Particles having higher weights are resampled at the observation time.

Resampling Step

- Particles having higher weights are resampled at the observation time.
- In this step, new particles are generated from large weight particles using stochastic kinetic-energy backscatter scheme (SKEBS) to avoid rapid filter degeneracy.

Resampling Step

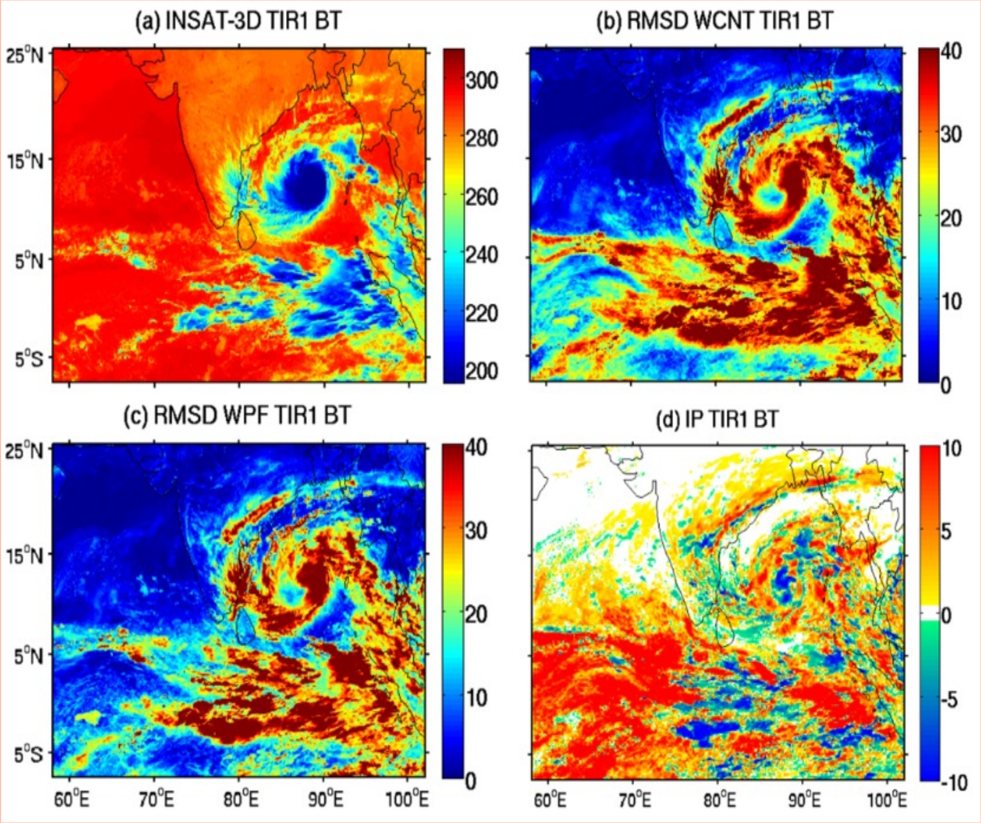
- Particles having higher weights are resampled at the observation time.
- In this step, new particles are generated from large weight particles using stochastic kinetic-energy backscatter scheme (SKEBS) to avoid rapid filter degeneracy.
- The advantage of SKEBS scheme is that it perturbs the dynamic state directly, and perturb dynamical variables feed into the physical parameterizations (model physics).

Resampling Step

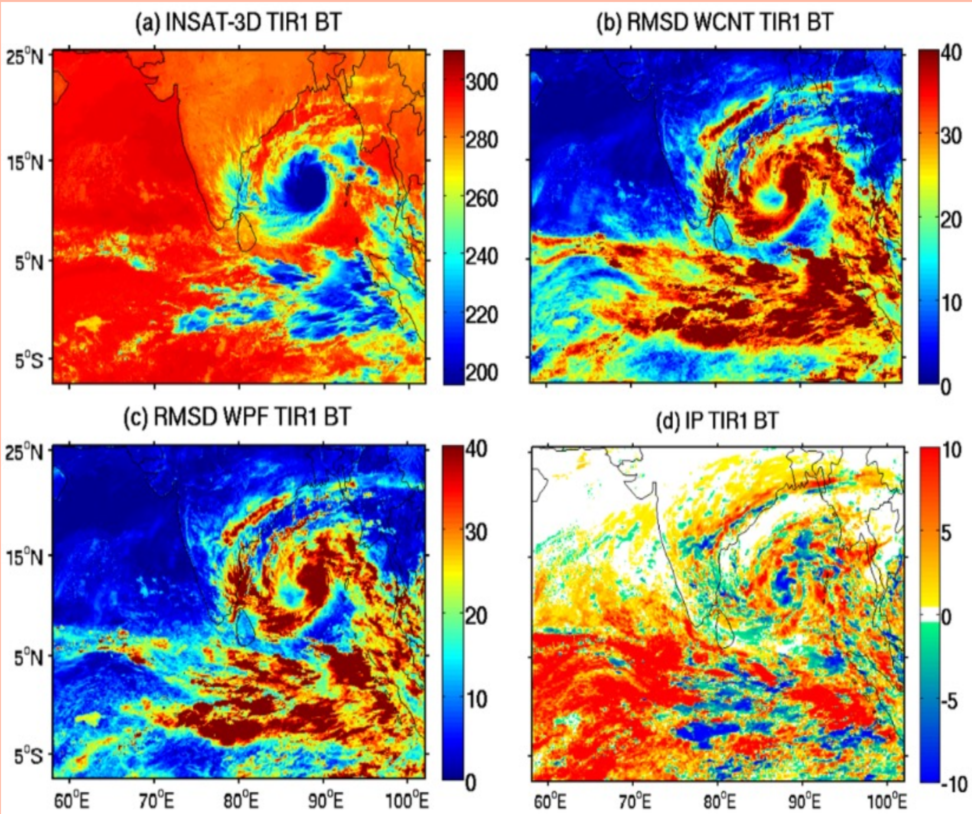
- Particles having higher weights are resampled at the observation time.
- In this step, new particles are generated from large weight particles using stochastic kinetic-energy backscatter scheme (SKEBS) to avoid rapid filter degeneracy.
- The advantage of SKEBS scheme is that it perturbs the dynamic state directly, and perturb dynamical variables feed into the physical parameterizations (model physics).
- In this way, the total numbers of particles are again same (92 here) at observation time step.

Resampling Step

- Particles having higher weights are resampled at the observation time.
- In this step, new particles are generated from large weight particles using stochastic kinetic-energy backscatter scheme (SKEBS) to avoid rapid filter degeneracy.
- The advantage of SKEBS scheme is that it perturbs the dynamic state directly, and perturb dynamical variables feed into the physical parameterizations (model physics).
- In this way, the total numbers of particles are again same (92 here) at observation time step.
- The idea is to focus the particles toward high probability regions of the target pdf, so that the number of particles required for a good approximation of target pdf remains manageable with fewer dimensions as compared to actual model space.

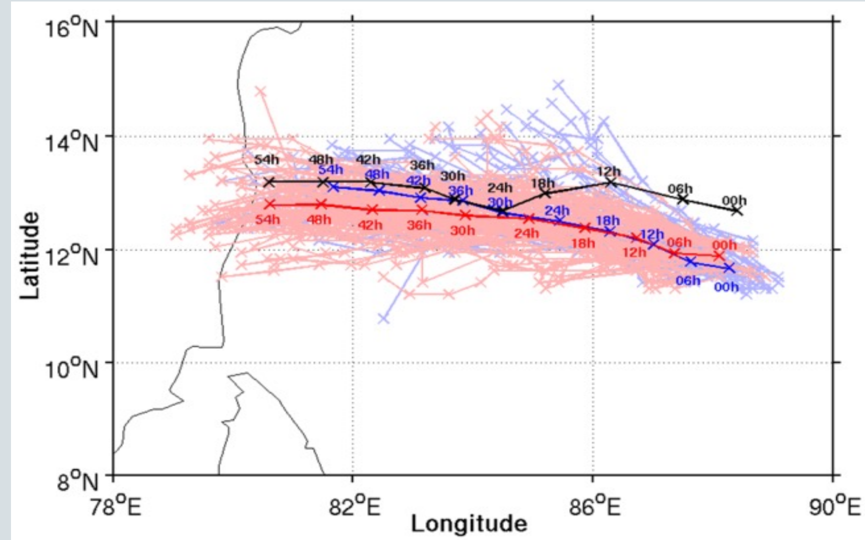


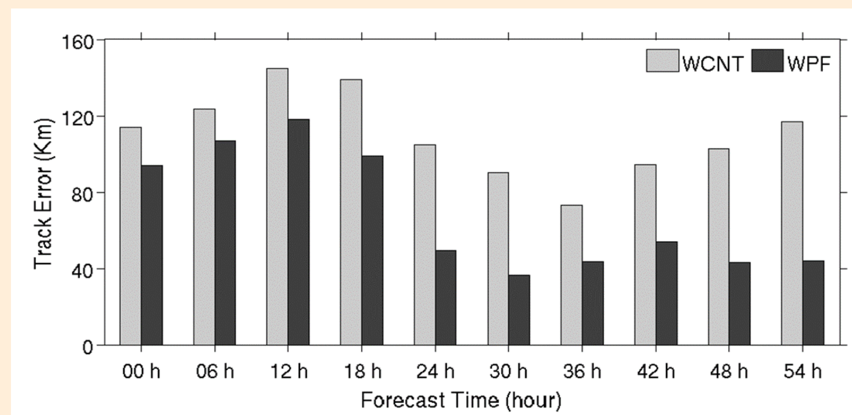
Simulated TC Landfall Error is better than IMD (14.7N, 80.0E) & SCORPIO (15.2N, 80.0E) predicted operational track forecasts from 00 UTC 10Dec2016.



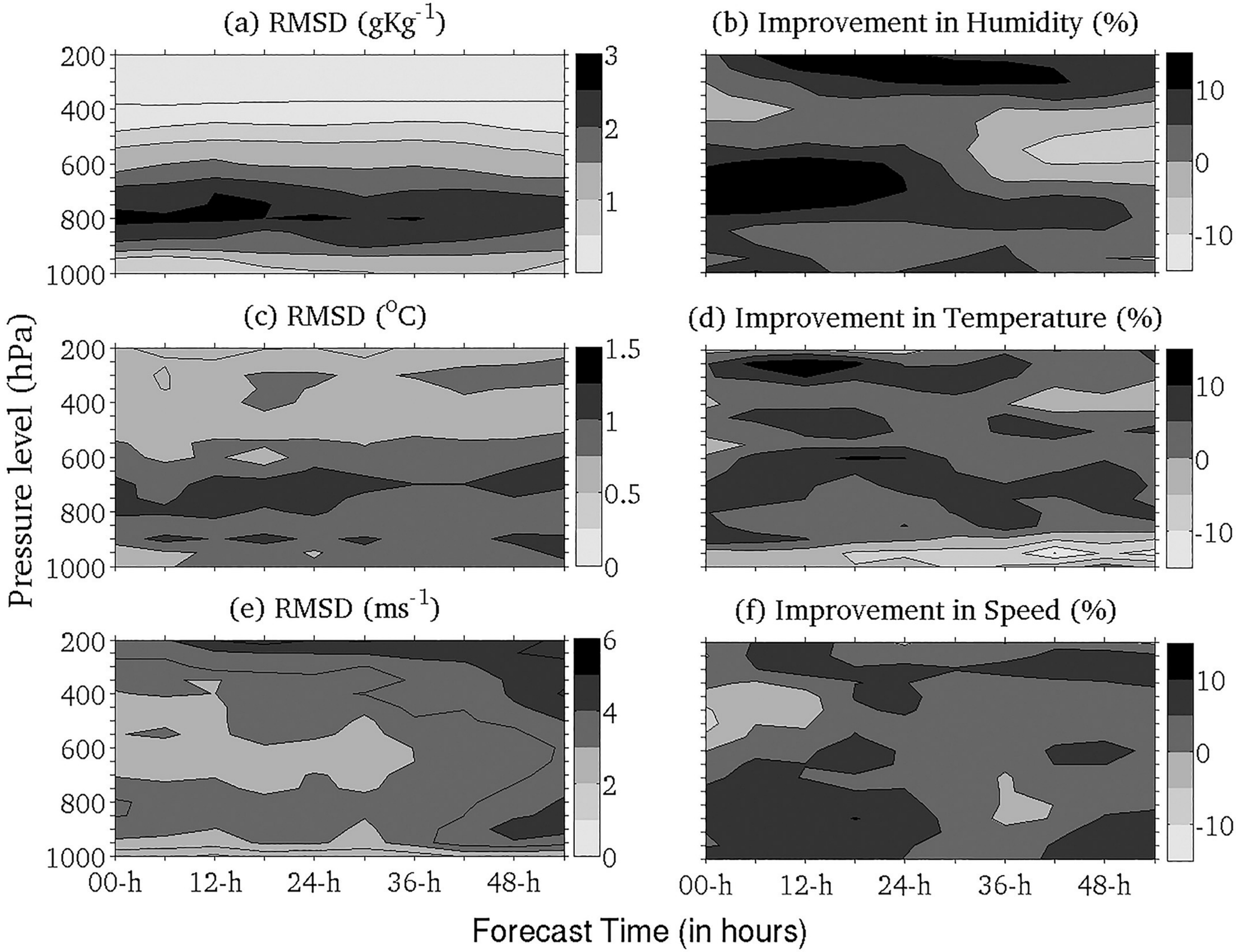
Simulated TC Landfall Error is better than IMD (14.7N, 80.0E) & SCORPIO (15.2N, 80.0E) predicted operational track forecasts from 00 UTC 10Dec2016.

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).



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 in (b) humidity, (d) temperature, and (f) wind speed in WPF experiments over WCNT experiments.

Rainfall Assimilation using PF: Motivation

Earth and Space Science

Research Article | Open Access

Assimilation of the Rain Gauge Measurements Using Particle Filter

Prashant Kumar

First published: 23 September 2020 | https://doi.org/10.1029/2020EA001212 | Citations: 3

Rainfall Assimilation using PF: Motivation

- ✓ Rainfall forecast from the NWP model is one of the most crucial and least accurate parameter compared to other parameters, e.g. temperature and humidity.
- ✓ Improving initial condition in precipitating regions is important for advancing the skill of the NWP models.

Rainfall Assimilation using PF: Motivation

- ✓ Rainfall forecast from the NWP model is one of the most crucial and least accurate parameter compared to other parameters, e.g. temperature and humidity.
- ✓ Improving initial condition in precipitating regions is important for advancing the skill of the NWP models.
- ✓ Assimilation of cloud and precipitation data in the NWP model is very preliminary, and limited to use of clear-sky data only (e.g. IR radiance assimilation).

Rainfall Assimilation using PF: Motivation

- ✓ Rainfall forecast from the NWP model is one of the most crucial and least accurate parameter compared to other parameters, e.g. temperature and humidity.

- ✓ Improving initial condition in precipitating regions is important for advancing the skill of the NWP models.

- ✓ Assimilation of cloud and precipitation data in the NWP model is very preliminary, and limited to use of clear-sky data only (e.g. IR radiance assimilation).

- ✓ Rainfall Assimilation is a direct way to use precipitation observations in the NWP model.

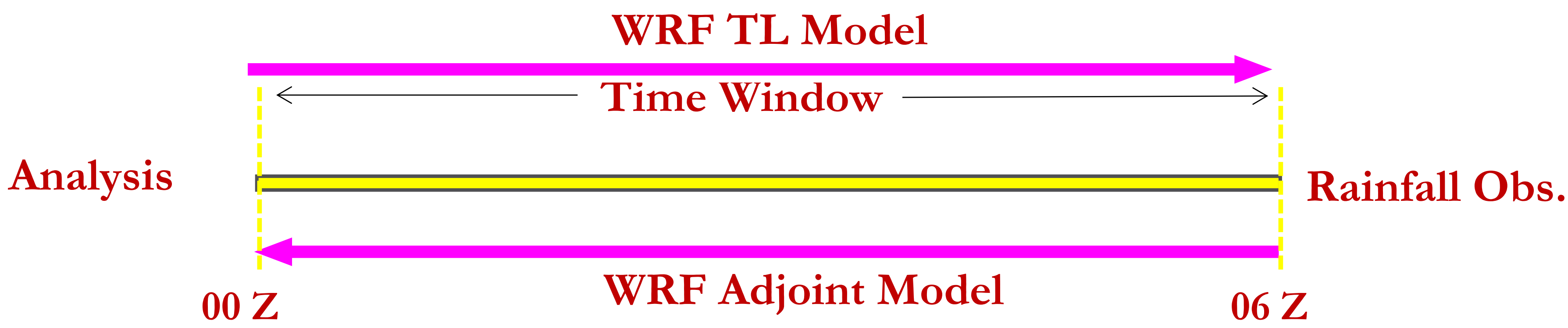
Previous Work on Rainfall Assimilation

1) Assimilation of JAXA GSMaP & TRMM 3B42 Rainfall using 4D-Var: Need of appropriate QC (JGR, 2014)

2) Assimilation of INSAT-3D HE Rainfall: Impact of real-time rainfall and sensitivity study for Heavy rainfall event, Need of appropriate first guess (QJRMS, 2016)

3) Objective: PF Assimilation of IMD observed Rainfall

Assimilation using 4D-Var Method



4D-Var Cost Function

$$J(x_0) = \frac{1}{2}\left(x_0 - x_0^b\right)^T \left[P_0^b\right]^{-1}\left(x_0 - x_0^b\right) + \frac{1}{2}\sum_{t=1}^{n}\left(y_t - H_t x_t\right)^T \left[R_t\right]^{-1}\left(y_t - H_t x_t\right)$$

3D-Var → 4D-Var:

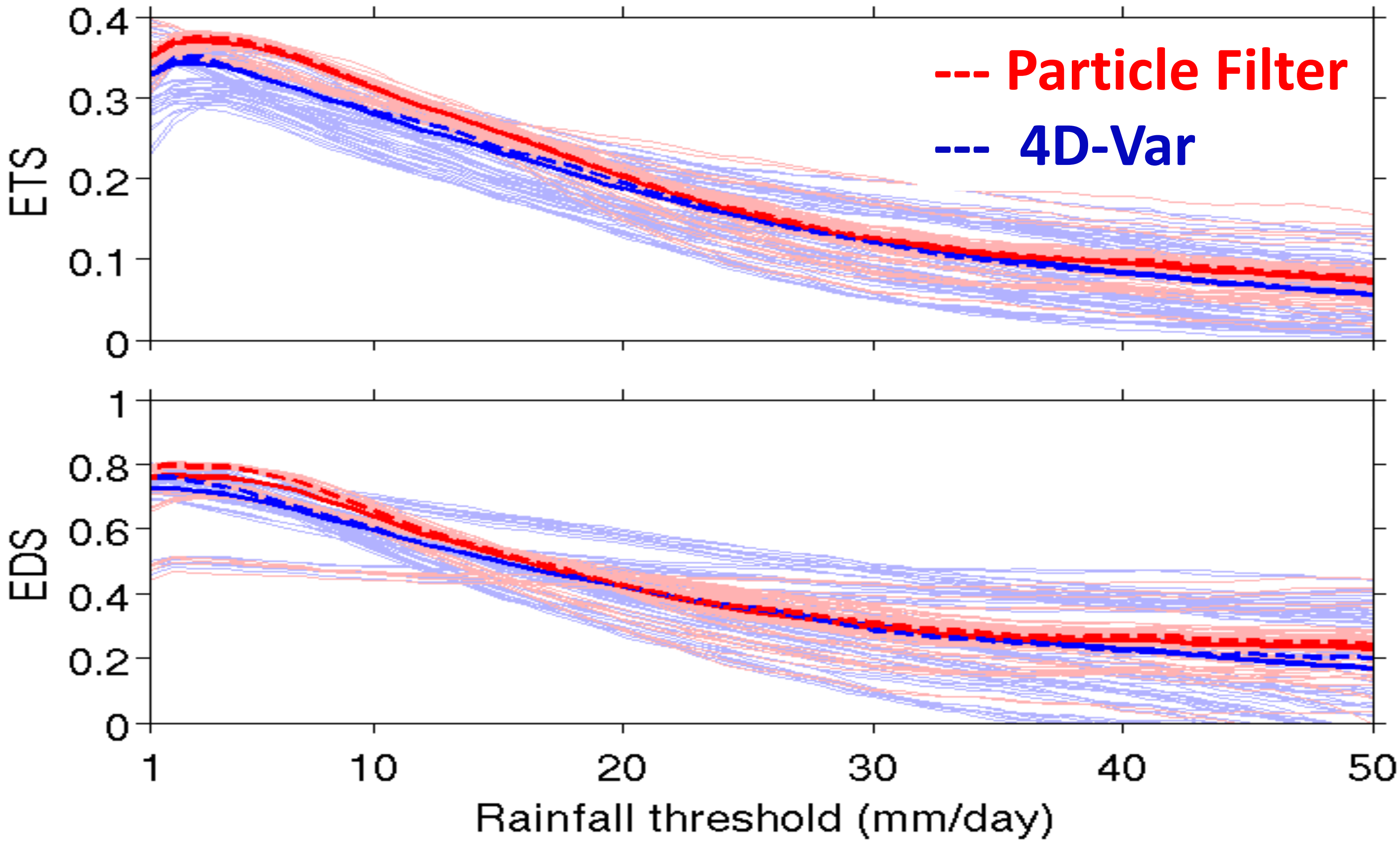
$$where\ \ \mathbf{x}_t = M\left(\mathbf{x}_0\right)$$

$H \rightarrow HM$; $\mathrm{H}^T \rightarrow \mathrm{M}^T\mathrm{H}^T$

The solution of 4D-Var is

$$\boldsymbol{x}^a = \boldsymbol{x}^b + \mathbf{B}M^T H^T\left[H\left(M\mathbf{B}M^T\right)H^T + \mathbf{R}\right]^{-1}\left[\boldsymbol{y} - HM\boldsymbol{x}^b\right]$$

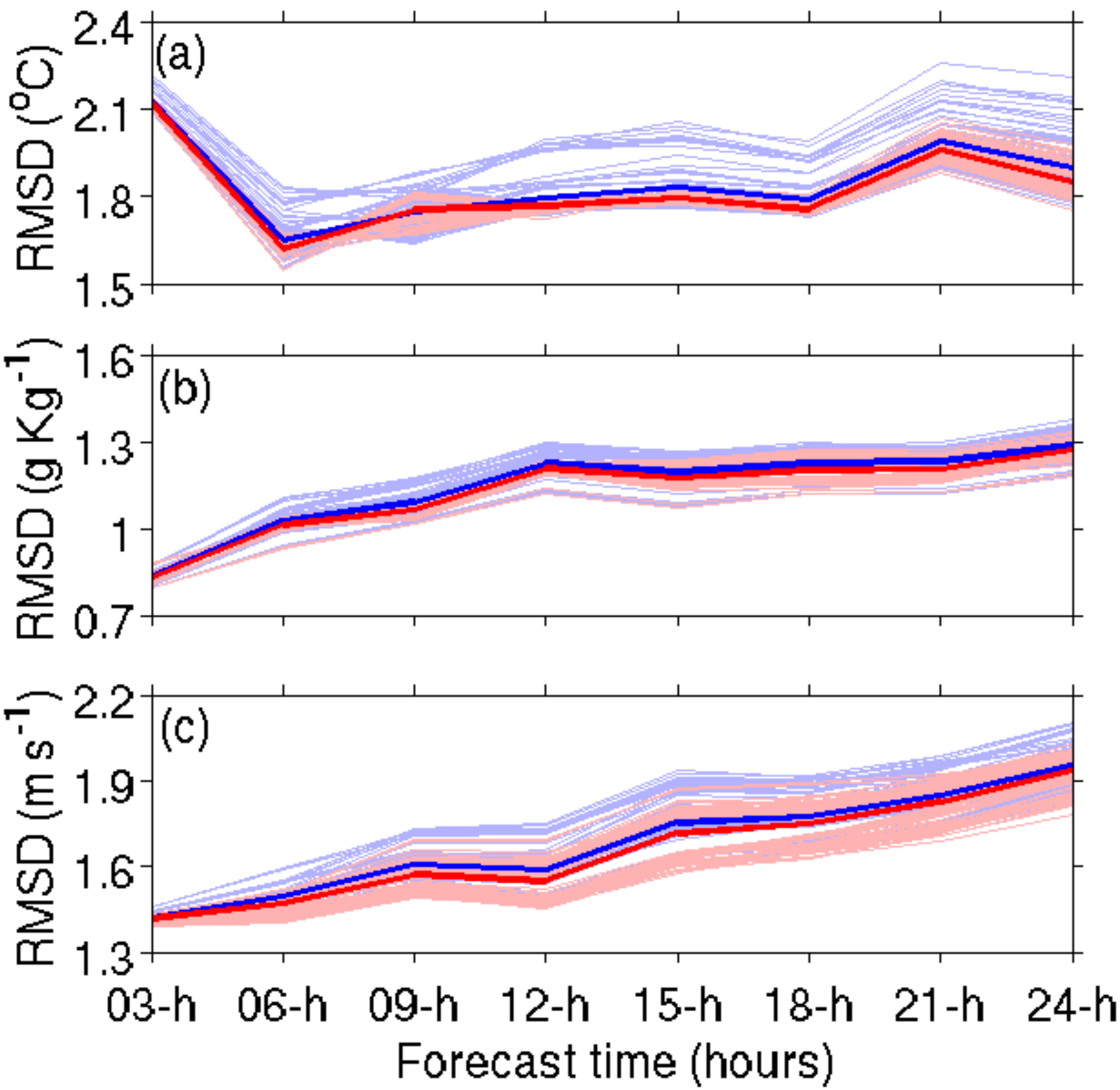
IMD Rainfall Assimilation using Particle Filter



Mean and median are plotted using dark line and dark dash lines respectively.

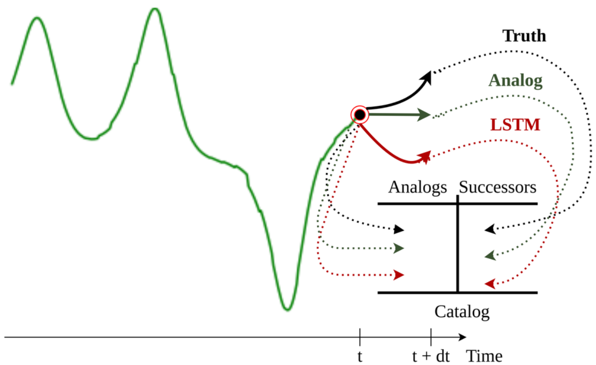
Particle Filter Assimilation (Cont …)

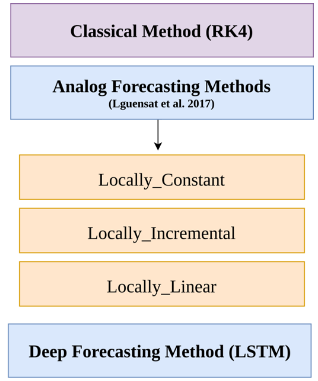
Individual particles are shown by light blue and red lines. Mean is plotted by dark line.

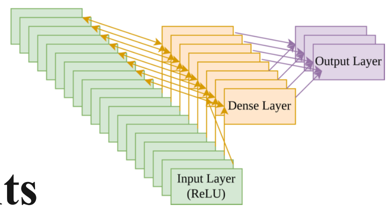


The Deep Data Assimilation for Geophysical Model

Design of Analog and Deep Forecasting Operators

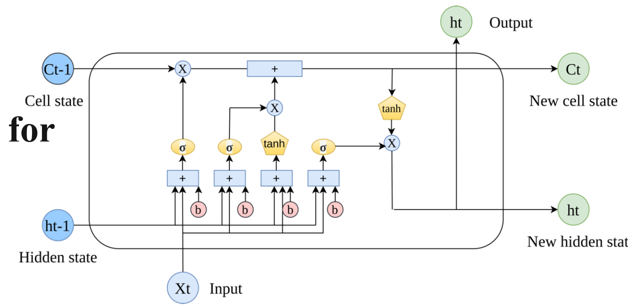






A Study towards ML developed Low-Cost NWP Ensembles

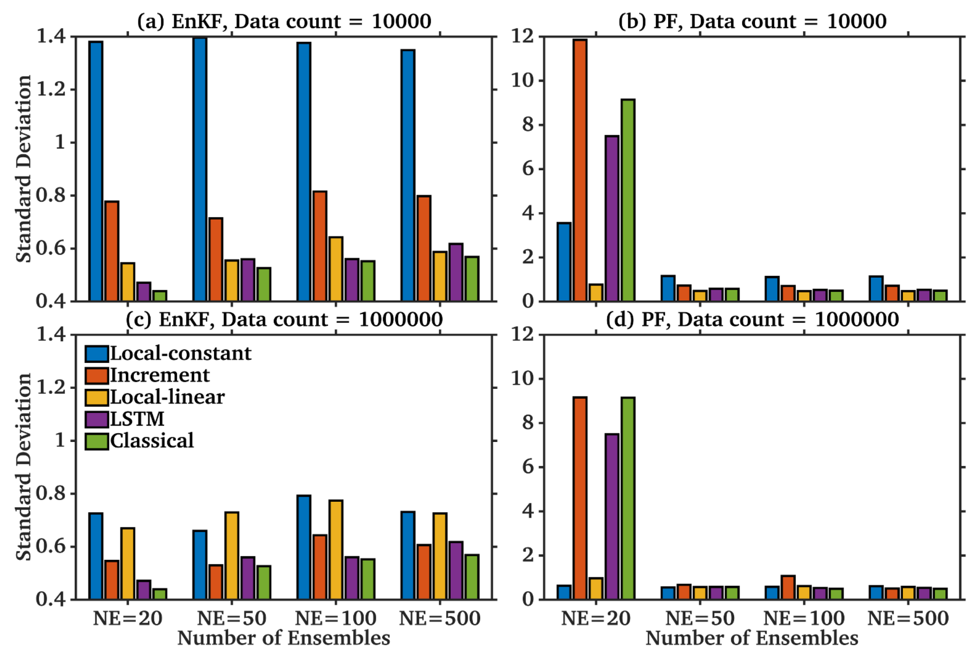
LSTM architecture for Deep Forecasting Operator



Highlights

- A DeepDA method has been proposed in which Deep Learning based sampling of the model dynamics is combined with ensemble-based assimilation techniques.
- The DeepDA method is a non-linear extension of the Analog Data Assimilation (AnDA) method (Lguensat, MWR, 2017).
- Results suggested that the proposed DeepDA is highly computationally efficient with sufficient skill against AnDA method.

Standard deviation with different enseble members for EnKF and PF filters with small and large catalog.



Which method is right for you?

	Var	KF	EnKF	EnVar	Hybrid	PF
Non-Gaussian	✗	✗	✗	✗	✗	✓
Large system	✓	✗	✓	✓	✓	✓
Need info on analysis error	✗	✓	✓	✗	✗	✓
TLM/ adjoint needed	✓	✓	✗	(✓✗)	(✓✗)	✗
Model expensive to run (no more than 50-100 runs)	✓	✓	✓	✓	✓	✗
Easily parallelizable	✗	✗	✓	✗	✗	✓

NCEO, UK
(7-10 May 2024)

Which problem you need to Attack?

Maths **Physics (Science)** **Forecasting** **Decision Support**

Thanking you