Assimilation of the GSMaP Precipitation Data with the SCALE-LETKF System

Cheng Da[1] and Guo-Yuan Lien[2]

[1] University of Maryland, College Park, USA

[2] RIKEN Advanced Institute for Computational Science, Kobe, Japan

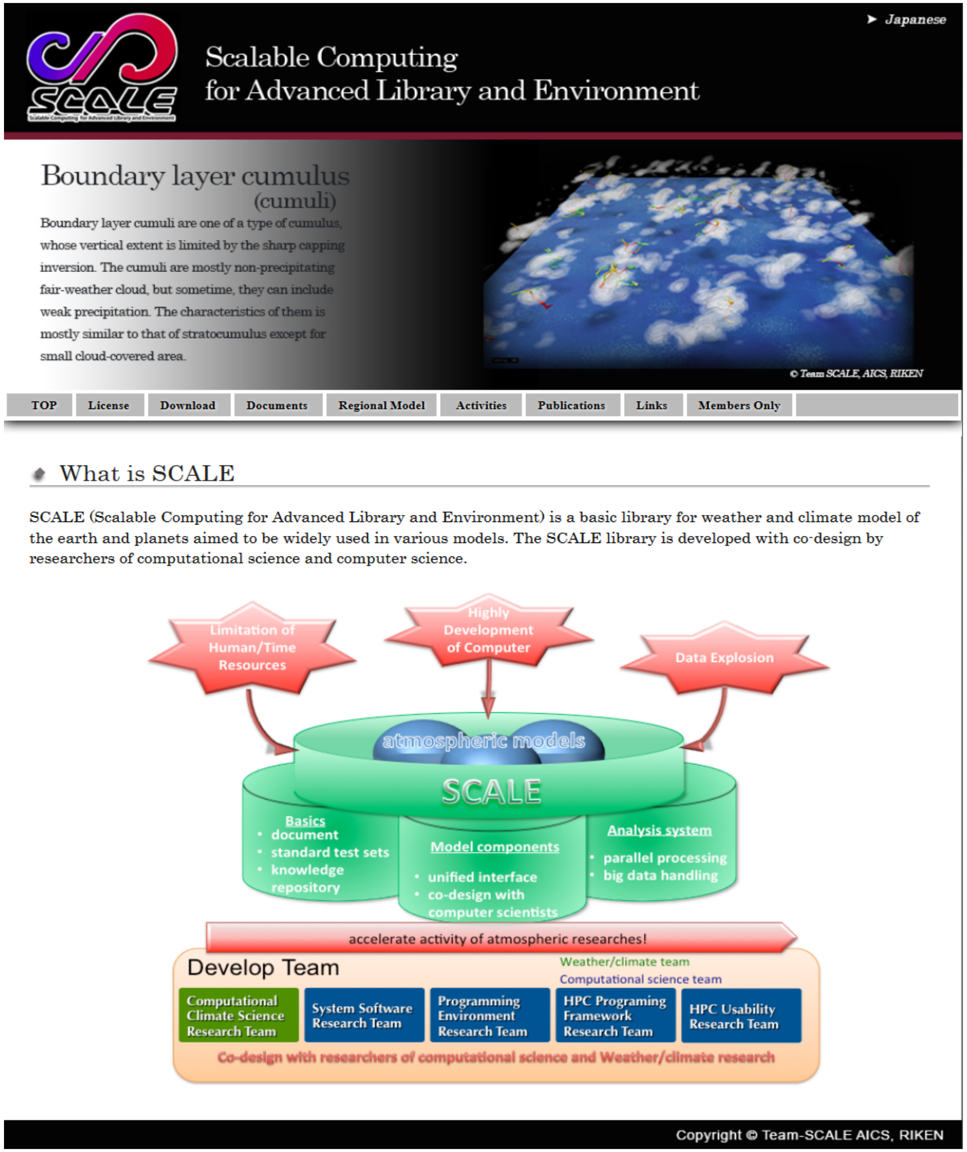
AICS internship report, Sep 26, 2017

- SCALE-LETKF

- Precipitation assimilation
  (Lien et al. 2013, 2016a,b, Kotsuki et al. 2017)

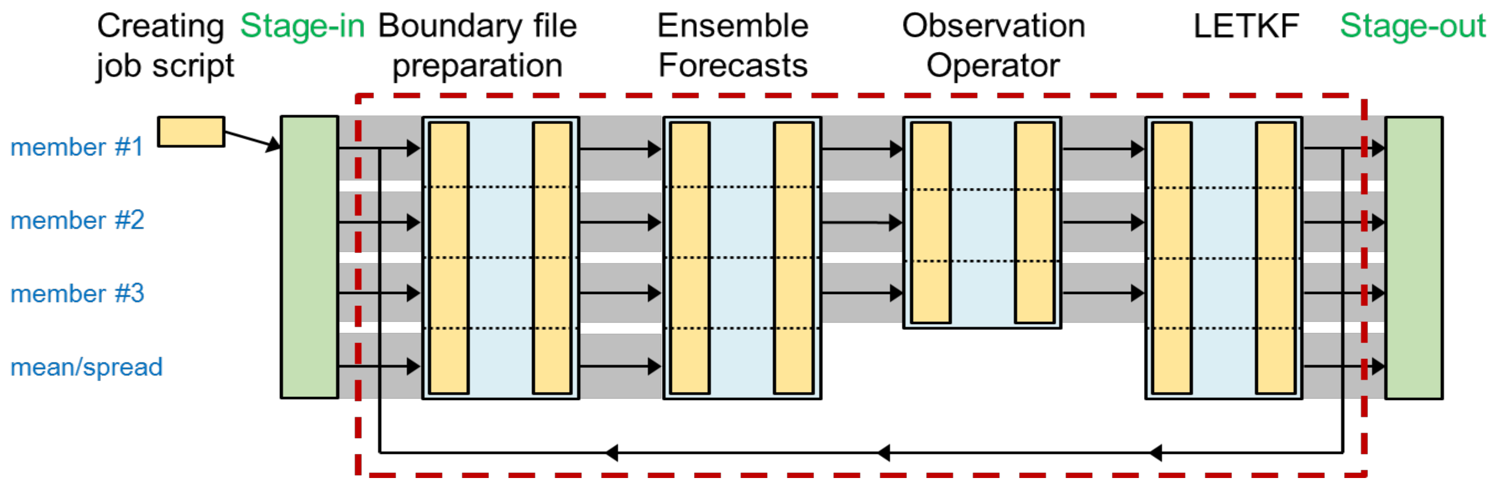
SCALE model (Nishizawa et al. 2015; Sato et al. 2015)

http://scale.aics.riken.jp/



- <u>S</u>calable <u>C</u>omputing for <u>A</u>dvanced <u>L</u>ibrary and <u>E</u>nvironment (SCALE).
- An open-source basic library for weather and climate simulation.
- Developed by Computational Climate Science Research Team.
- **SCALE-RM**: A regional NWP model based on SCALE.

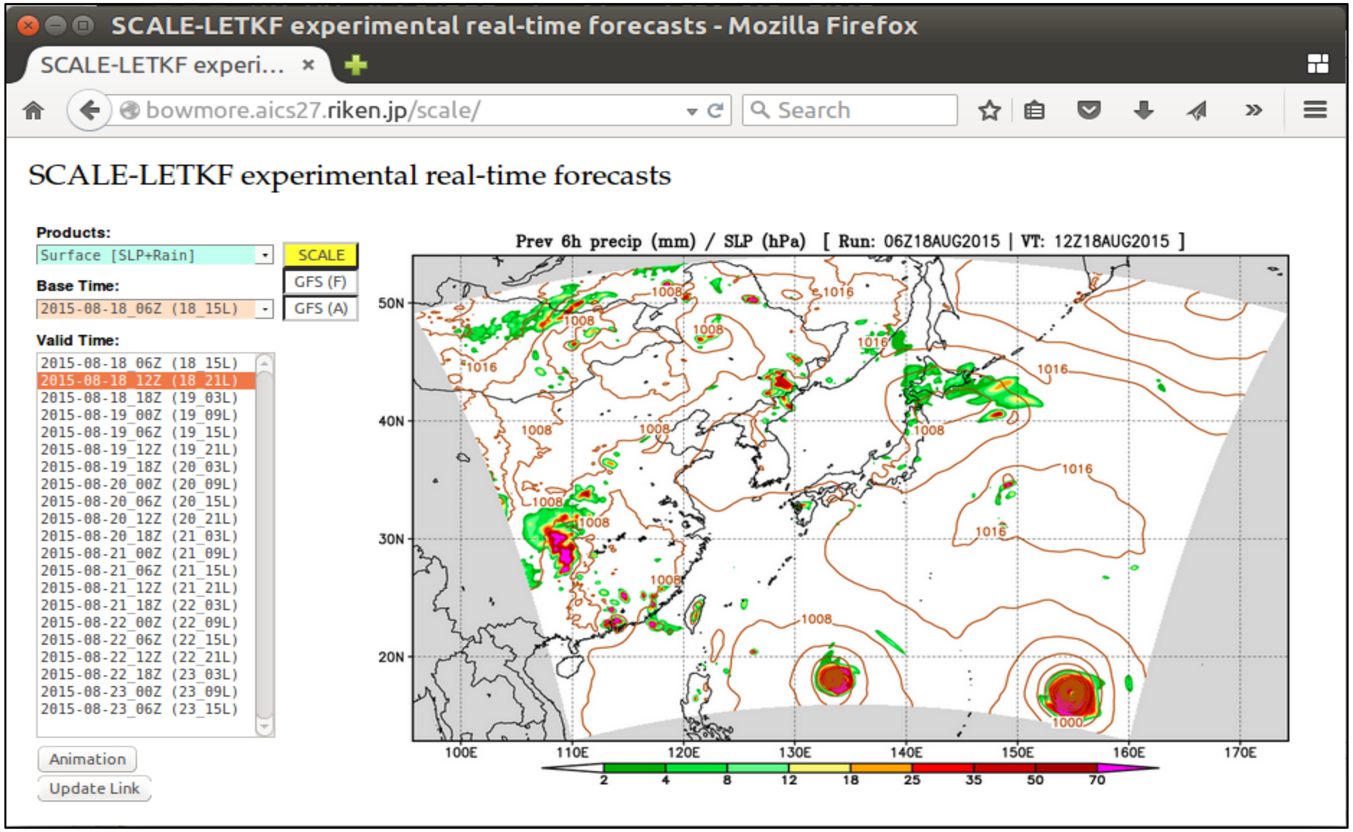
SCALE-LETKF (Lien et al. 2017)



https://github.com/gylien/scale-letkf

- A Local Ensemble Transform Kalman Filter (LETKF) data assimilation package with the SCALE-RM.
- The SCALE-LETKF is:
  - A **configurable** and **scalable** regional data assimilation system.
  - Featured for the LETKF scheme, which is not focused in other community data assimilation code (such as DART and Community GSI).
  - Tested on a very large supercomputer (i.e., the K computer), aware of the efficiency of high performance computing.

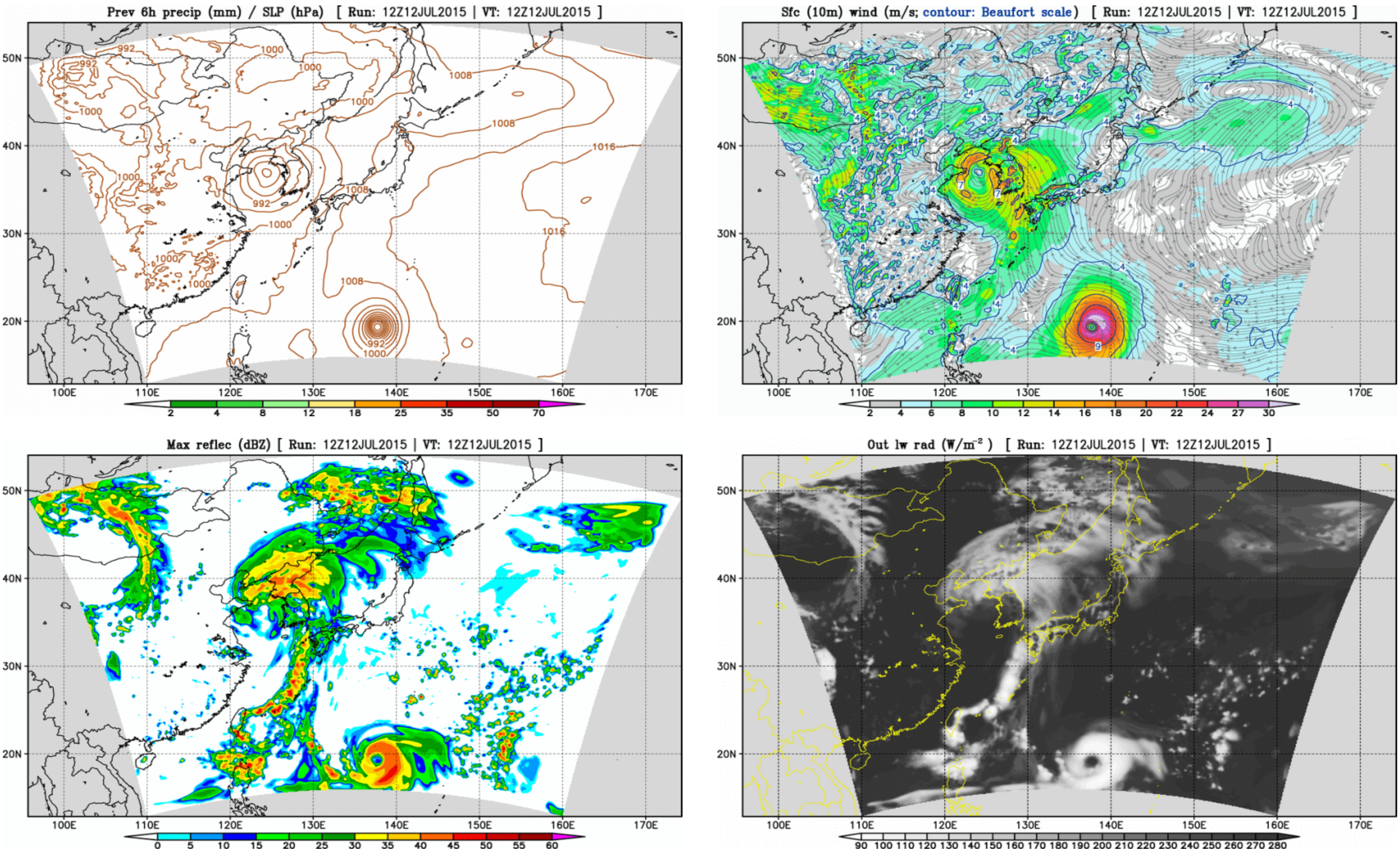
Experimental near-real-time SCALE-LETKF



- 18-km resolution; 5760 x 4320 km area; 50 members
- 6-hour cycle; 5-day deterministic forecasts
- Test the stability of the system.
- Provide important guidance on the performance of the model and the data assimilation settings.

An example of 5-day forecasts

(Typhoon NANGKA; 12:00 UTC July 12, 2015)



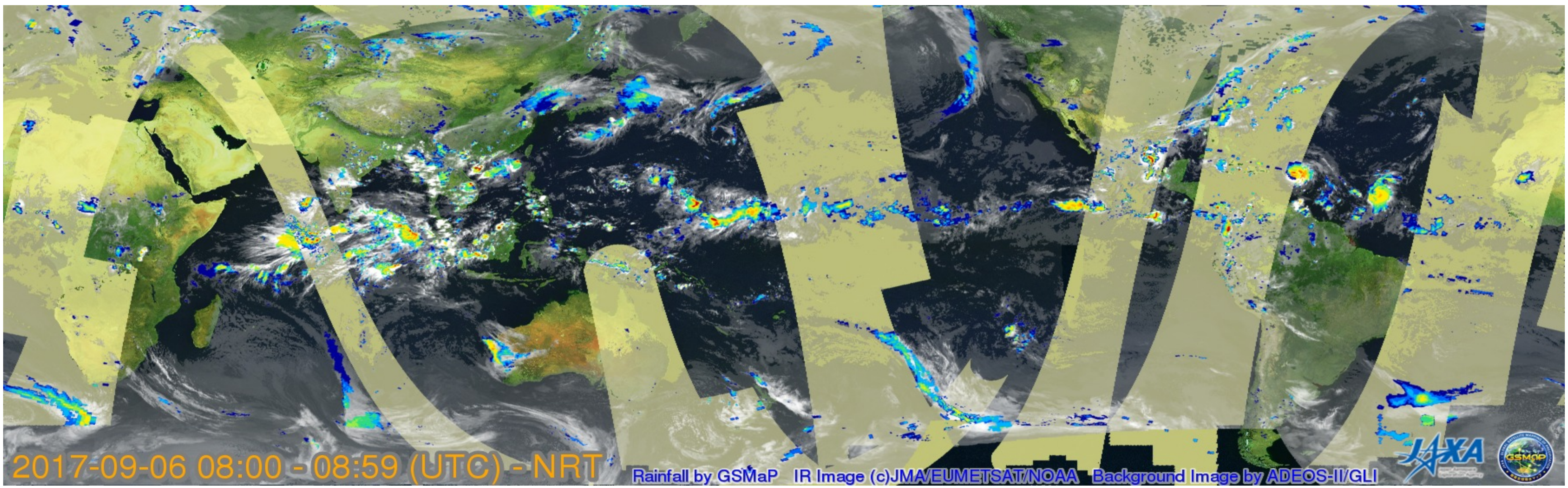
- SCALE-LETKF

- Precipitation assimilation
  (Lien et al. 2013, 2016a,b, Kotsuki et al. 2017)

Satellite precipitation estimates

- In recent years, several satellite-based near-real-time global precipitation estimates have been available:
  - TRMM Multisatellite Precipitation Analysis (TMPA) / Integrated Multi-satellitE Retrievals for GPM (IMERG)
  - Global Satellite Mapping of Precipitation (GSMaP)

http://sharaku.eorc.jaxa.jp/GSMaP/



Assimilation of precipitation data

- Past studies of precipitation assimilation show good analyses, but the model forgets about the changes soon after the assimilation stops.
  - The change in moisture is not an efficient way to update the dynamical variables that primarily determine the evolution of the forecast in NWP models.
  - Flow-dependent covariances between the precipitation variable and other dynamical variables are important.
    → EnKF
- More difficulties:
  - The non-Gaussianity of the precipitation variable
  - The large model and observation errors.

Gaussian transformation

$$F^G(\tilde{y}) = F(y) \quad \Rightarrow \quad \tilde{y} = {F^G}^{-1}[F(y)]$$

y : original variable (mm/6hr)

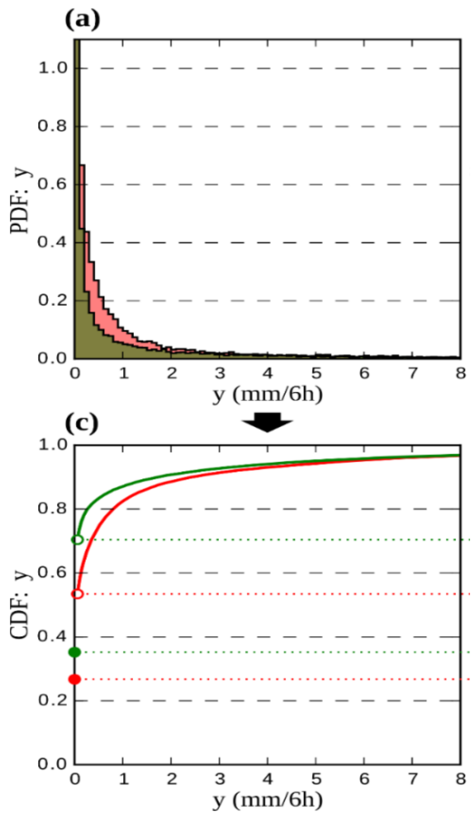
$F(\,)$: CDF of original variable

$\tilde{y}$: Transformed variable (sigma)

$F^G(\,)$: CDF of Gaussian distribution

PDF

CDF



Original variable

—: **Model**

—: **Obs.**

Step 0: Obtain PDF & CDF

(Lien et al. 2013, 2016b)

Gaussian transformation

$$F^G(\tilde{y}) = F(y) \quad \Rightarrow \quad \tilde{y} = {F^G}^{-1}[F(y)]$$

y : original variable (mm/6hr)

$F(\,)$: CDF of original variable

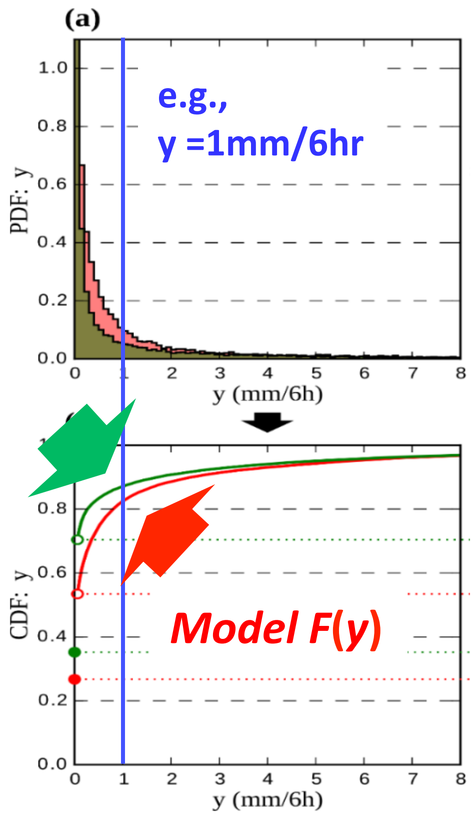
$\tilde{y}$: Transformed variable (sigma)

$F^G(\,)$: CDF of Gaussian distribution

PDF

Obs F(y)

CDF



Original variable

—: Model
—: Obs.

Step 0: Obtain PDF & CDF

Step 1: Compute $F(y)$

(Lien et al. 2013, 2016b)

Gaussian transformation

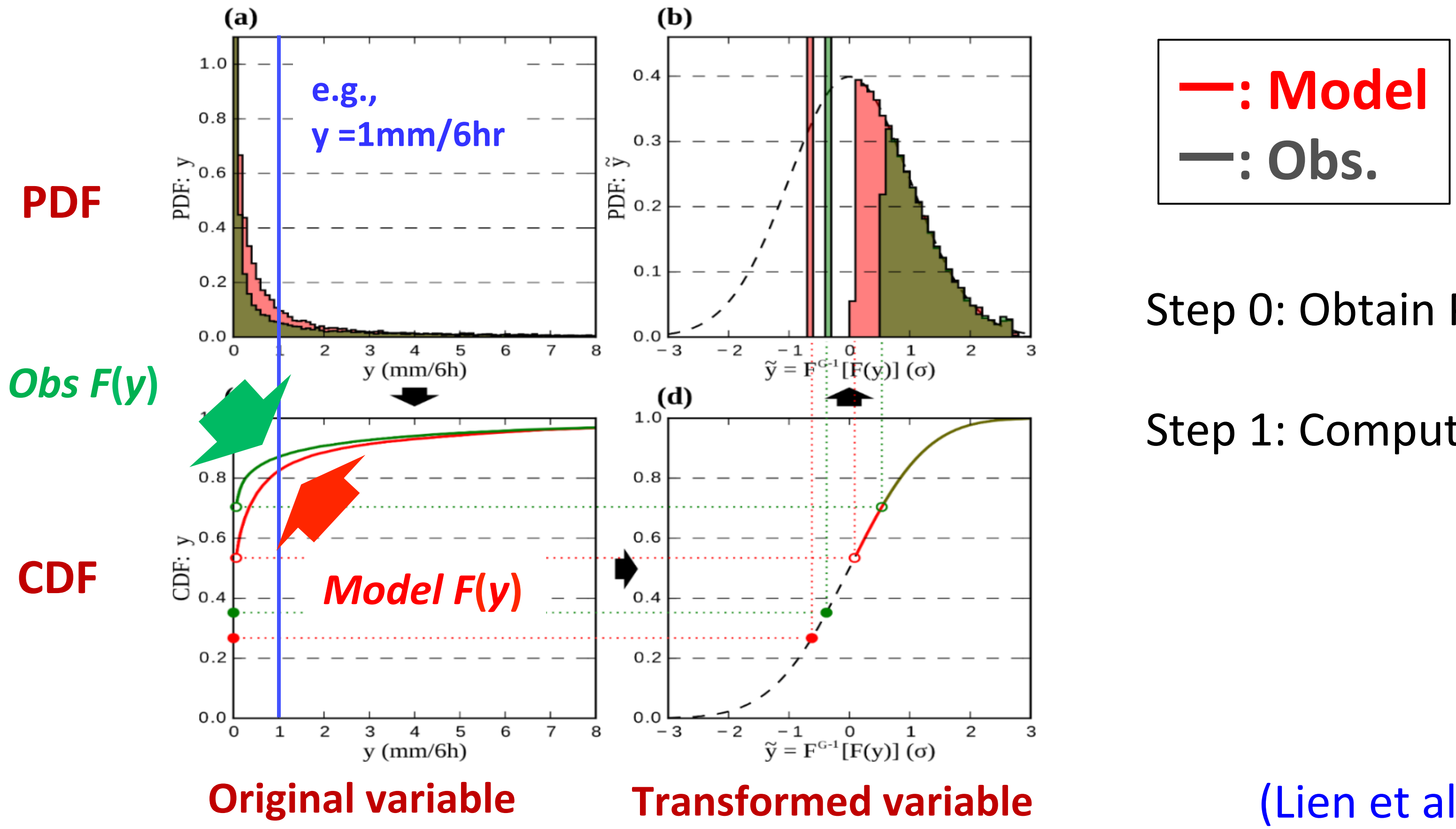
$$F^G(\tilde{y}) = F(y) \quad \Rightarrow \quad \tilde{y} = F^{G^{-1}}[F(y)]$$

y : original variable (mm/6hr)

$F(\,)$: CDF of original variable

$\tilde{y}$: Transformed variable (sigma)

$F^G(\,)$: CDF of Gaussian distribution



Step 0: Obtain PDF & CDF

Step 1: Compute $F(y)$

(Lien et al. 2013, 2016b)

Gaussian transformation

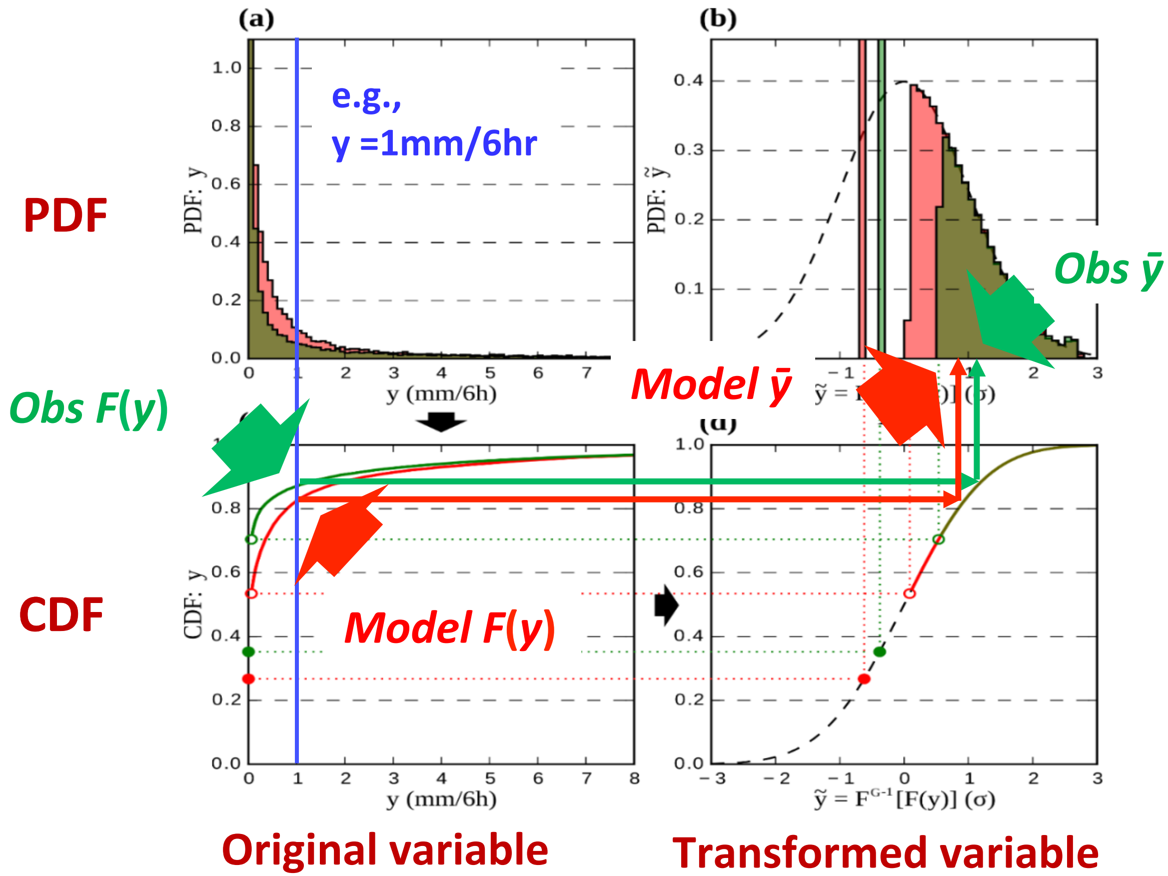
$$F^{G}(\tilde{y}) = F(y) \quad \Rightarrow \quad \tilde{y} = {F^{G}}^{-1}[F(y)]$$

y : original variable (mm/6hr)

$F(\,)$: CDF of original variable

$\tilde{y}$: Transformed variable (sigma)

$F^{G}(\,)$: CDF of Gaussian distribution



—: Model

—: Obs.

Step 0: Obtain PDF & CDF

Step 1: Compute $F(y)$

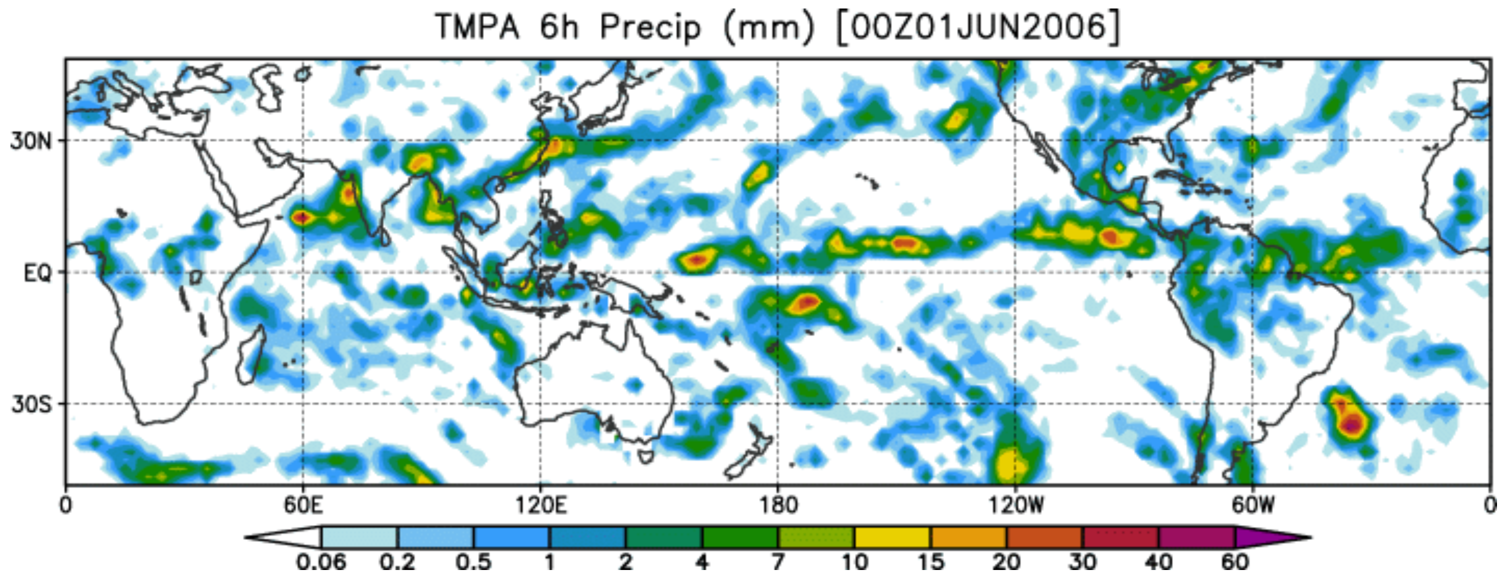
Step 2: Compute

$$\tilde{y} = {F^{G}}^{-1}[F(y)]$$

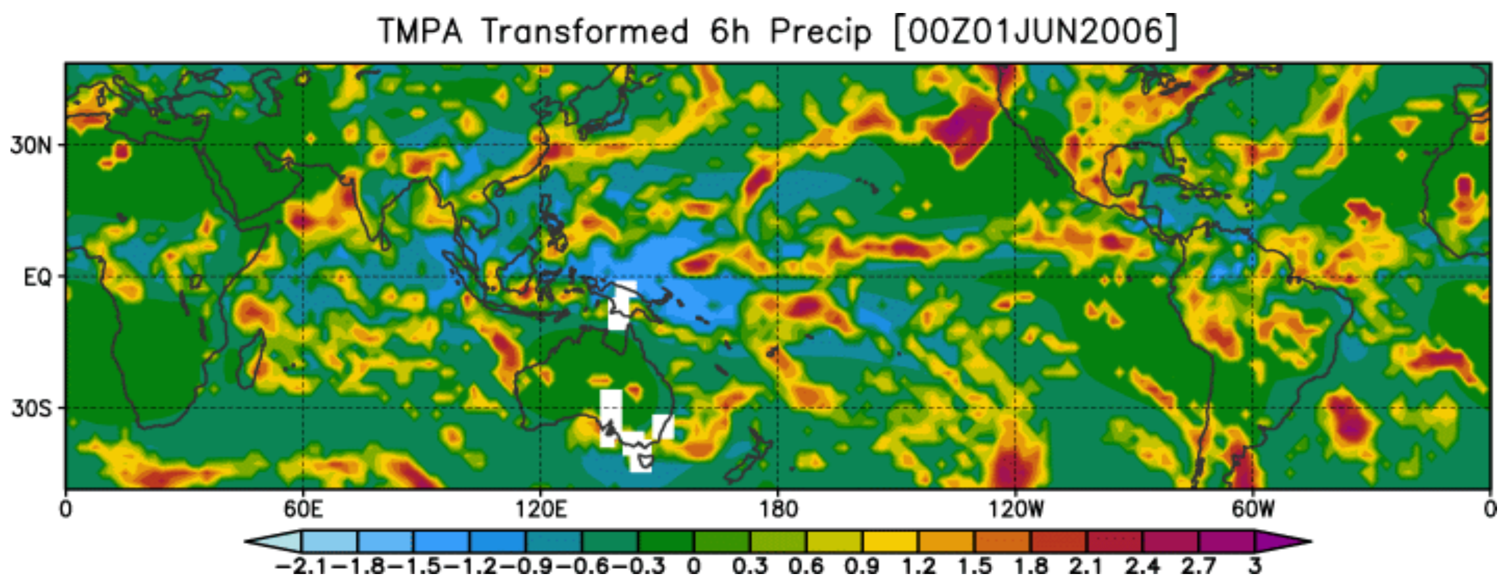
(Lien et al. 2013, 2016b)

Gaussian transformation

Original variable



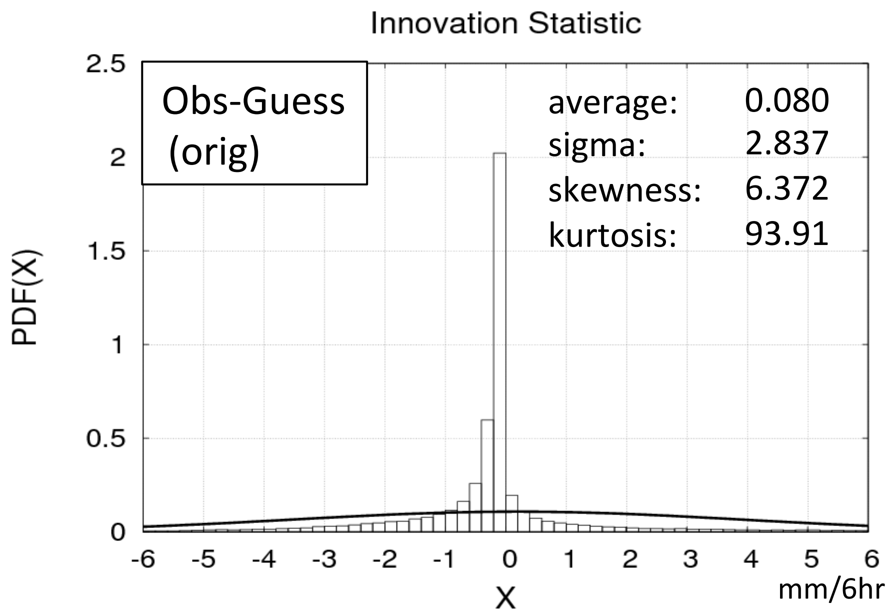
Transformed variable



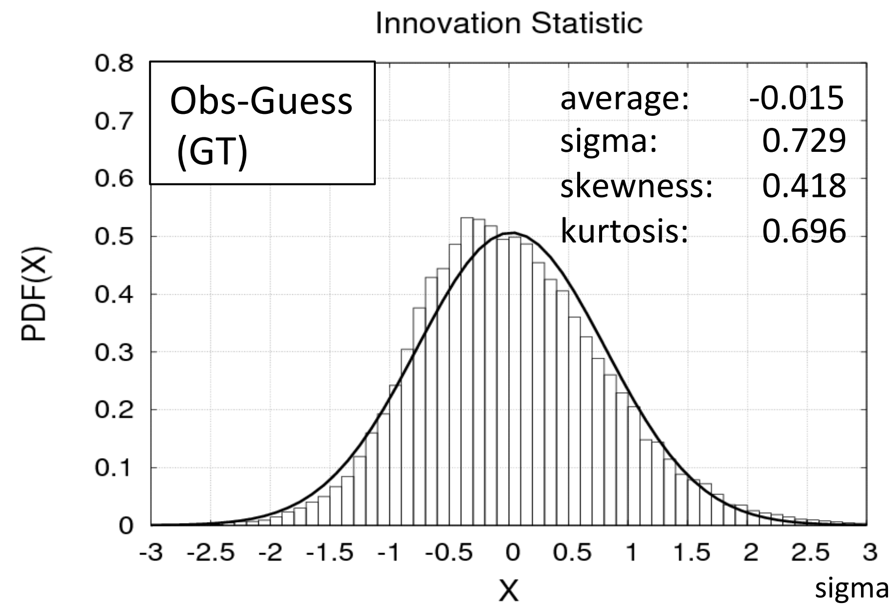
Gaussian transformation: Error distribution

(Kotsuki et al. 2017, JGR-A)

wo Gaussian-Transformation



w Gaussian-Transformation



More Gaussian

Sampling period : 2014110100 - 2014110118

Construction of the precipitation CDF

Data Assimilation Team

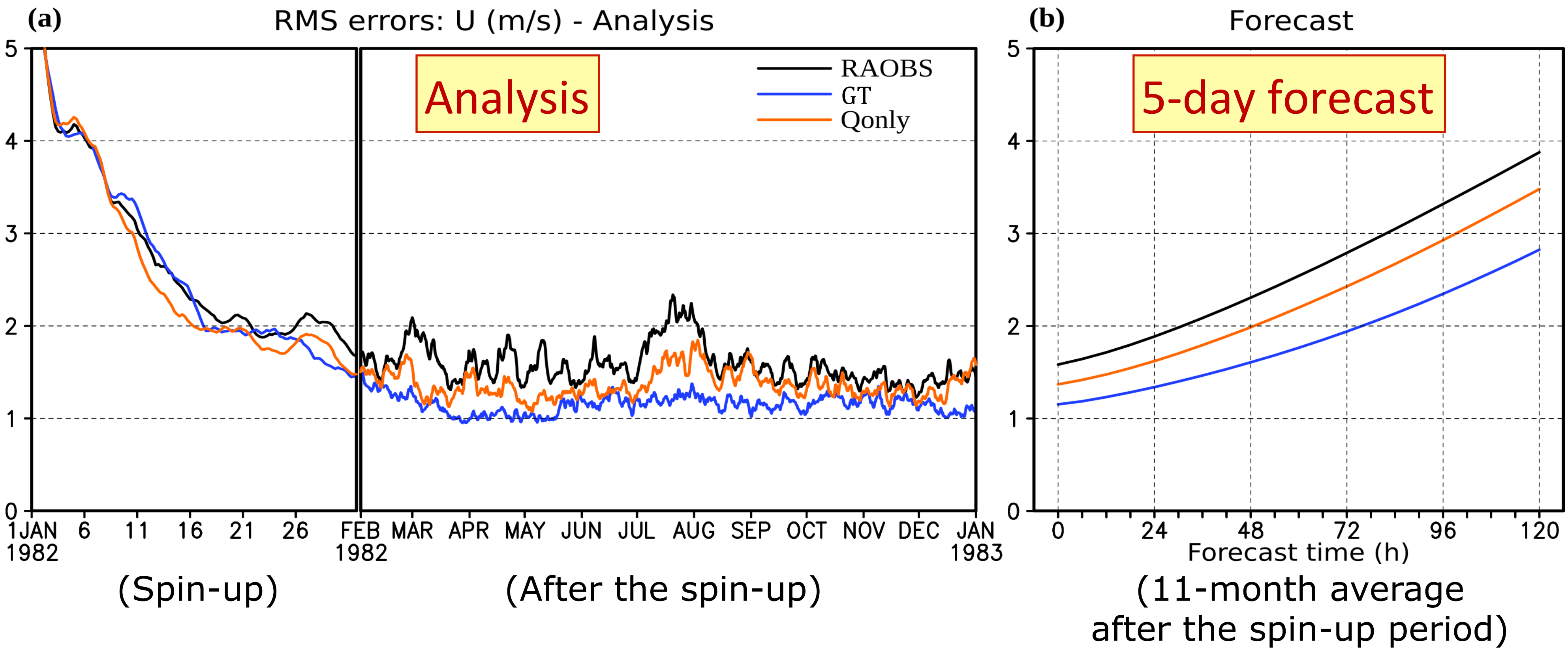
- The CDF of precipitation variables is empirically determined based on the model/observation climatology at each grid point.
  It requires a long period of model/observation data.

- Lien et al. 2013, 2016a,b: Using a fixed 1-year climatology

- Kotsuki et al. 2017: Using the past month data prior to the assimilation time

Lien et al. 2013:

SPEEDY model / idealized experiments



RAOBS: Assimilate rawinsonde observations

GT: Assimilate rawinsondes + uniformly distributed global precipitation

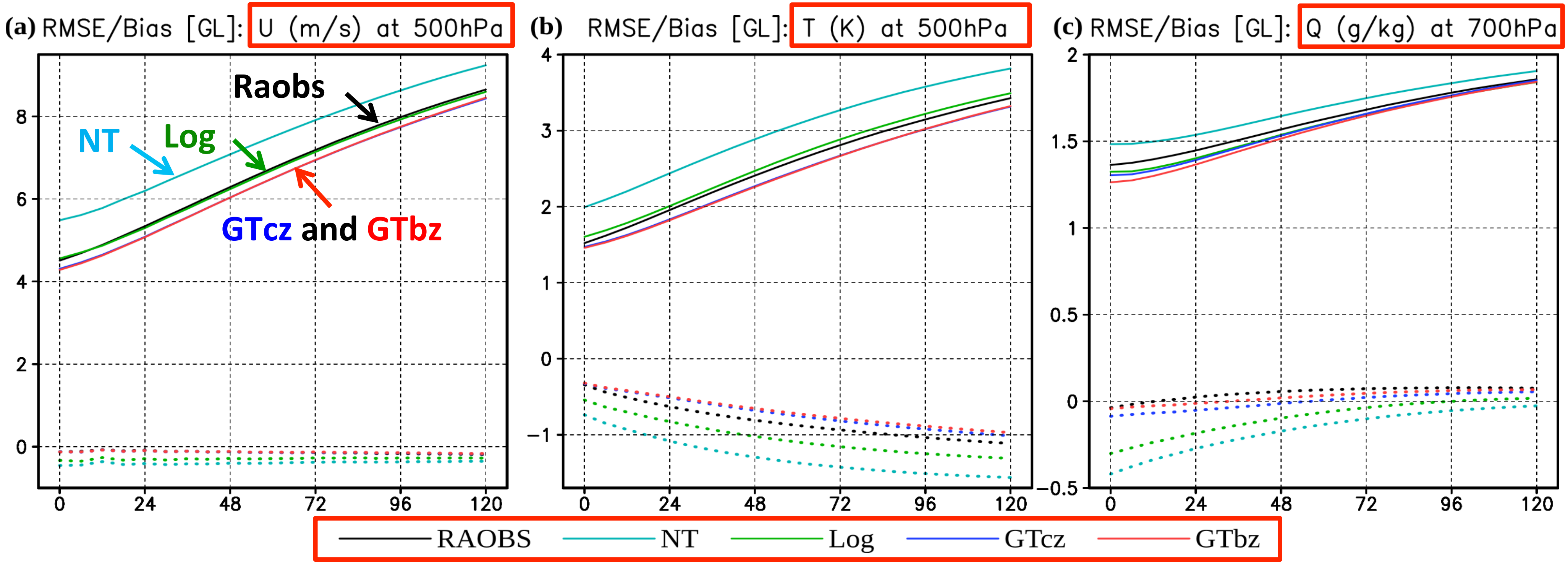
Qonly: Same as **GT**, but only update moisture field by precipitation assimilation

(Other variables show similar results)

Lien et al. 2016a, b:

GFS model / TMPA observation

5-day forecasts

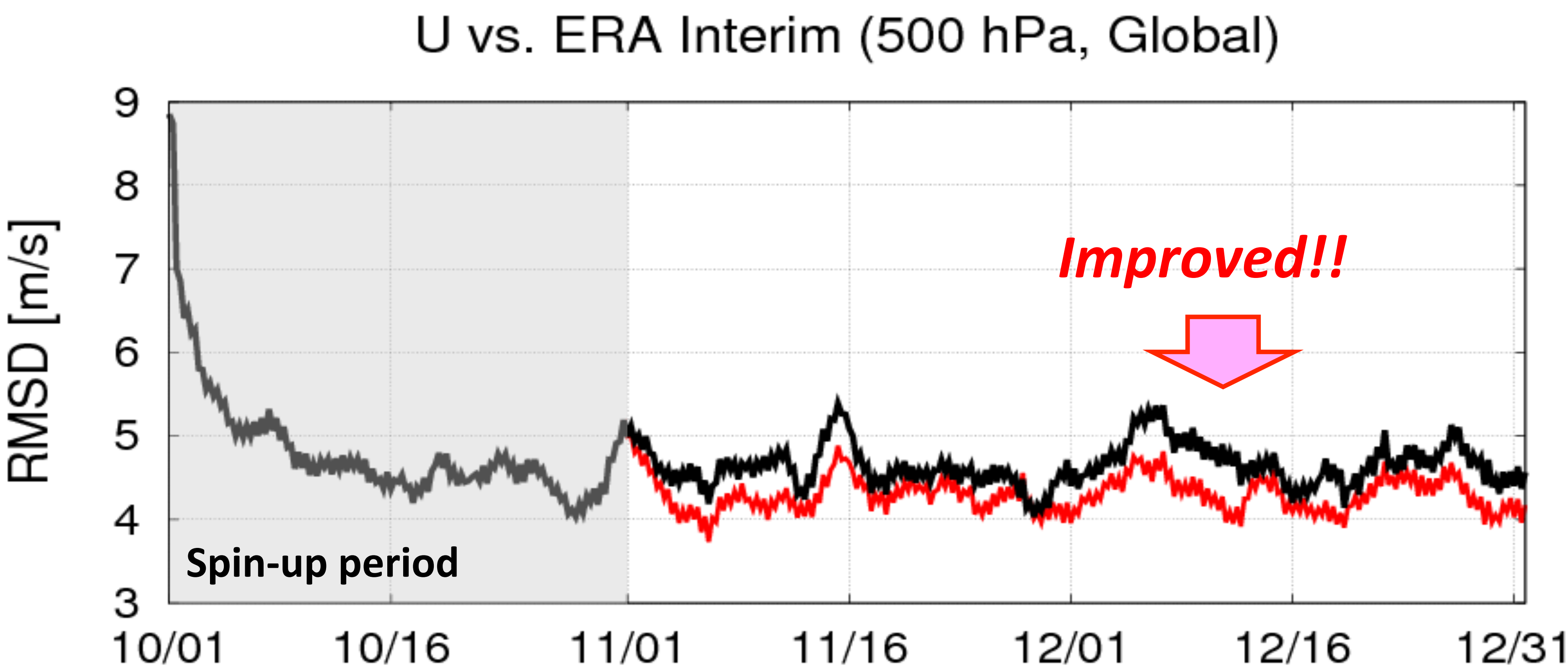


Global results Solid lines: RMS errors Dashed lines: Biases

- **No transformation (NT)** gives very bad results.
- **Logarithmic (Log)** transformation leads to marginal results.
  - Good for moisture, but bad for temperature.
- **Gaussian transformation (GTcz and GTbz)** lead to clear positive impacts.

Kotsuki et al. 2017:
NICAM model / GSMaP observation

Data Assimilation Team



—: CTRL: Radiosondes ONLY

—: TEST: Radiosondes + GSMaP/Gauge (every 5x5 grids)

Summary of the precipitation assimilation studies

- We successfully assimilated precipitation with global models, in both idealized OSSEs and a realistic model and observations, using the LETKF and the Gaussian transformation.
  - The impacts are seen in both analyses and 5-day forecasts.

- Gaussian transformation is beneficial to the precipitation assimilation:
  - Gaussian transformation based on the climatology does produce more Gaussian background error distribution of precipitation.
  - Applying Gaussian transformation separately to model/ observation precipitation can correct the bias.

Next… the intern study

- Precipitation assimilation in a regional mesoscale model?

- Impact of precipitation assimilation on typhoon analyses and forecasts?

Assimilation of the GSMaP Precipitation Data with the SCALE-LETKF System

Cheng Da

Department of Atmospheric & Oceanic Science
University of Maryland, College Park

Supervisor: Guo-Yuan Lien

Sponsored by the AICS HR development/AICS joint research program

Outline

1. **Motivation**
2. **Implementation of Precip. DA**
3. **Experiment Design**
   - ***3.1 Overview of Typhoon cases***
   - ***3.2 NWP & LETKF settings***
   - ***3.3 Precip. CDFs for Gaussian Transformation***
4. **Results**
   - ***4.1 Impact of GSMaP observations***
     - Adjustments of the SLP/hydrometeors
     - Impact on TC track & intensity forecast
   - ***4.2 Sensitivity to the QC schemes***
     - Total number of precipitating members
     - Assimilating ocean obs only
5. **Summary**

Motivation

- *Lien et al.* (2013) showed with the SPEEDY model that assimilating precipitation can improve both the **analysis** and medium-range **forecast** for almost **all variables**.
  - Apply the Gaussian transformation to the model & obs precipitation
  - Directly modify the dynamical variables through an EnKF system

- Consistent **positive** results with realistic global models and observations
  - ✓ GFS-LETKF/TMPA (*Lien et al.*, 2016b)
  - ✓ NICAM-LETKF/GSMaP (*Kotsuki et al*., 2017)

Question:

- Can we apply the same methodology to the mesoscale regional model?
- Specifically, can it provide additional benefits to the Typhoon forecast?

Implementation of Precip. DA component

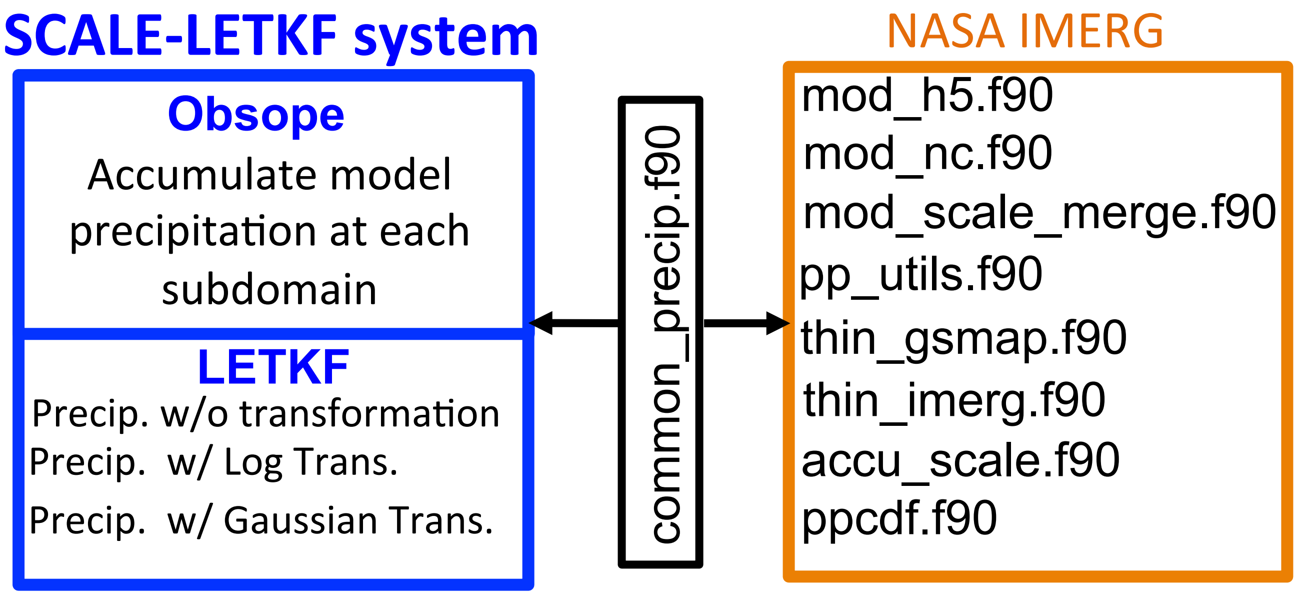
- The SCALE-LETKF system (*Lien et al.*, 2017) doesn't have the precipitation DA component yet. So we need to implement it first.

config.nml.letkf:

```
&PARAM_LETKF_PRECIP
        USE_PRECIP          = .true.,
        !--PPCDF_IN--
        NPPX               = 240,
        NPPY               = 180,
        NCDF               = 200,
        PPZERO_THRES        = 1.D-3,
        GAUSSTAIL_THRES      = 1.D-3,
        OPT_PPTRANS         = 3,
        OPT_PPOBSERR         = 3,
        LOG_TRANS_TINY       = 0.6D0,
        MIN_OBSERR_RAIN      = 0.3D0,
        OBSERR_RAIN_PERCENT  = 0.5D0,
        OBSERR_RAIN_LT       = 0.18D0,
        OBSERR_RAIN_GT       = 0.5D0,
        RAIN_SLOT_START      = 5,
        RAIN_SLOT_END        = 10,
        MIN_RAIN_MEMBER      = 75
        /
```

Precip. DA component:

Precip. Preprocessing Utilities: for both JAXA GSMaP & NASA IMERG

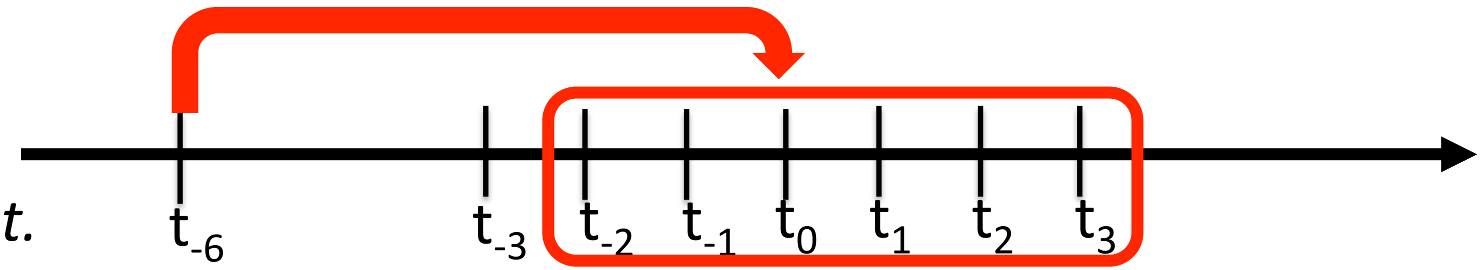


1. obs thinning
2. precip. CDF creation
3. Gross QC

Obs Operators / Obs

We assimilate the **accumulated 6-hr** precipitation:

SCALE
6-hour precip fcst.



***Obs Operator** H(x)*: $pp_{6hr}^{SCALE} = \sum_{i=-2}^{3} pp_i$

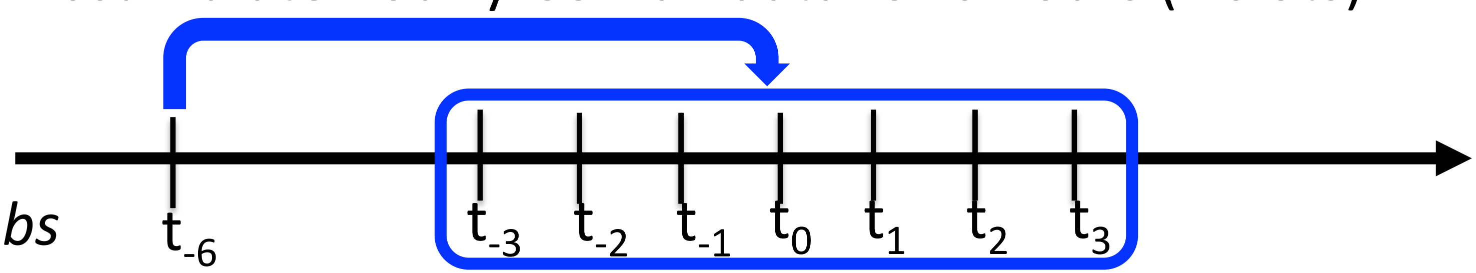
---

1. Thinning: keep only one GSMaP hourly obs per model grid

GSMaP has higher res. (10km) than model res. (36km)

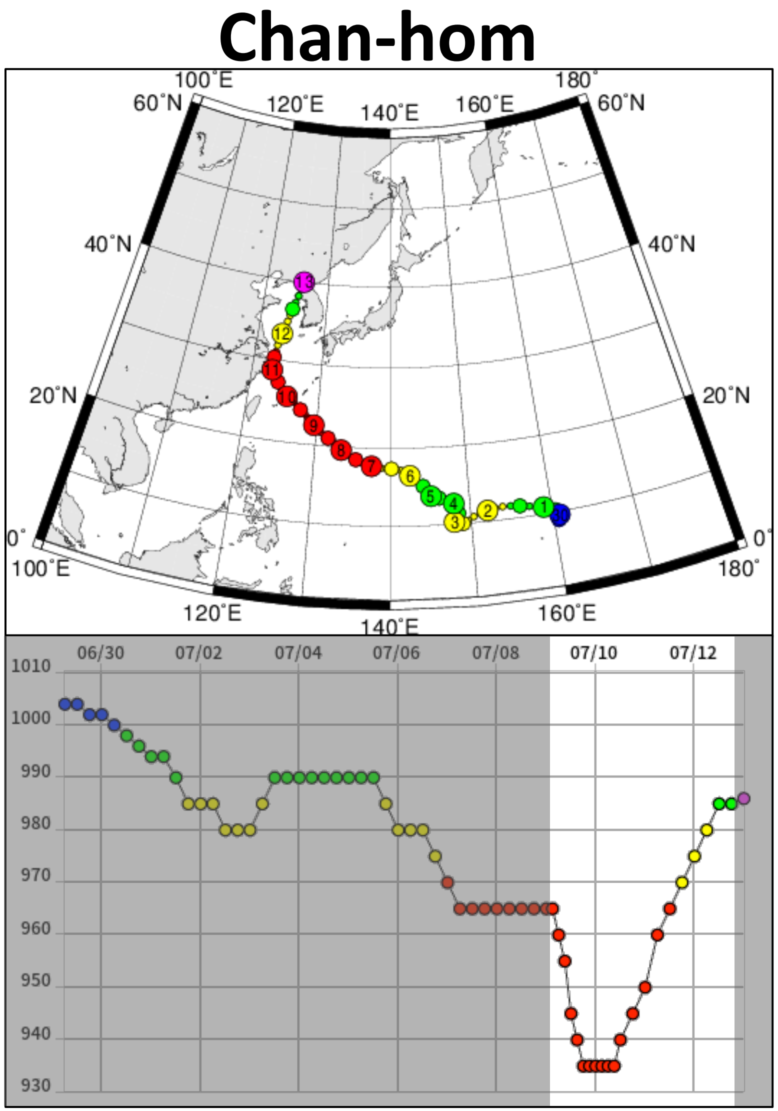
2. Aggregation: Accumulate hourly GSMaP data for 6 hours (7 slots)

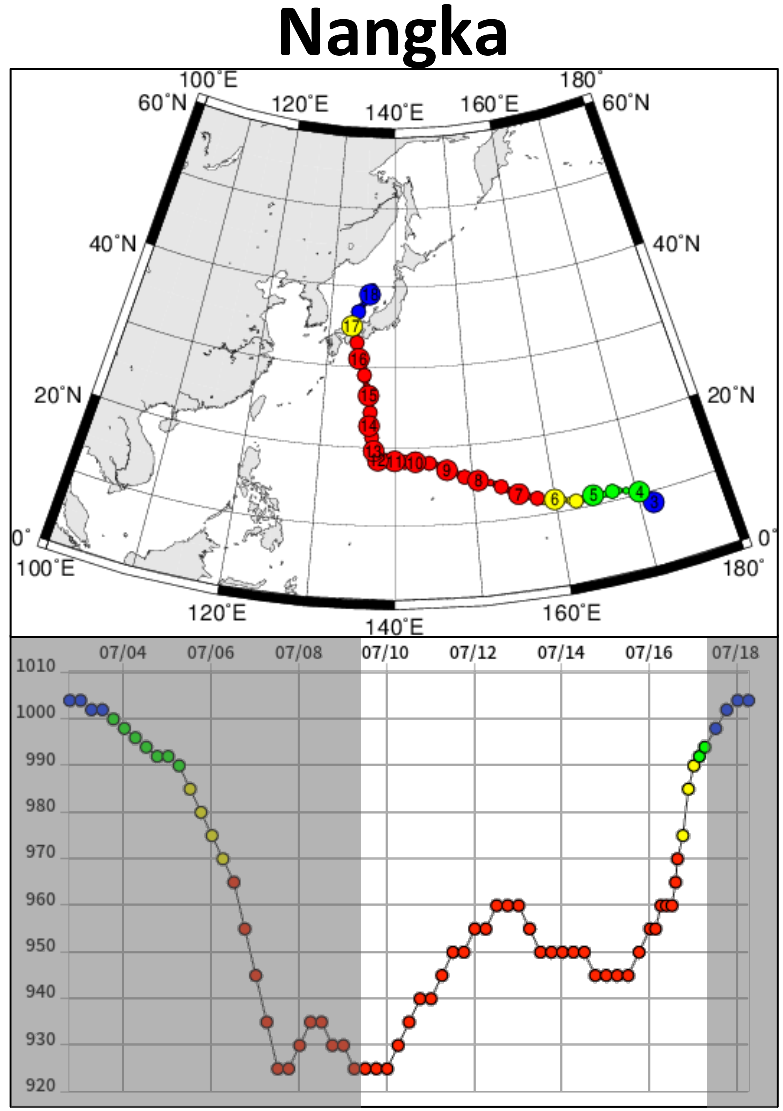
GSMaP
6-hour precip. obs



Obs y^o^: $pp_{6hr}^{GSMaP} = 0.5 \times pp_{-3} + \sum_{i=-2}^{2} pp_i + 0.5 \times pp_3$

Experiment Design: Overviews of Two TCs in 2015



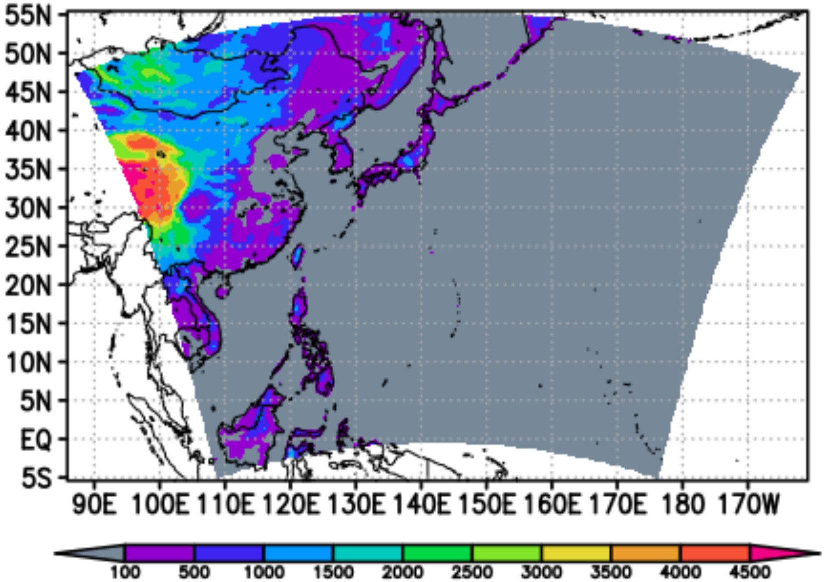


6hr DA cycles start from 0600UTC, July 9 to 0000UTC, July 17

Figures from http://agora.ex.nii.ac.jp

Experiment Settings: NWP Model

Model Domains

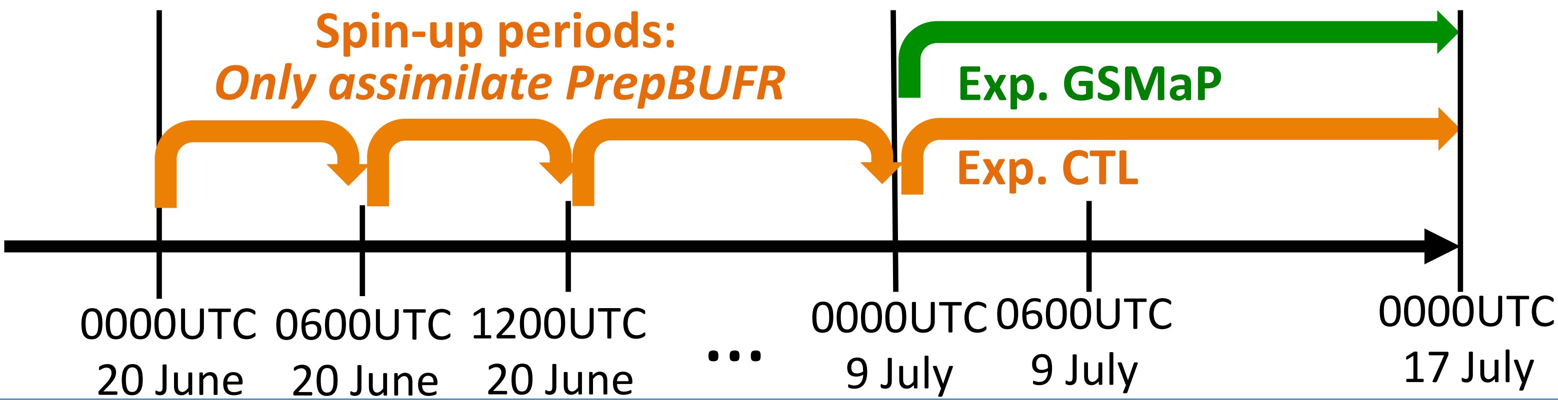


SCALE Settings	D01
Resolution	36km
Domain size	240×180 grids, 36 levels (~28km)
BC	GFS FNL every 6 hours
Microphysics	TOMITA08
Radiation	MSTRNX
Cumulus Para.	w/o and w/ KF

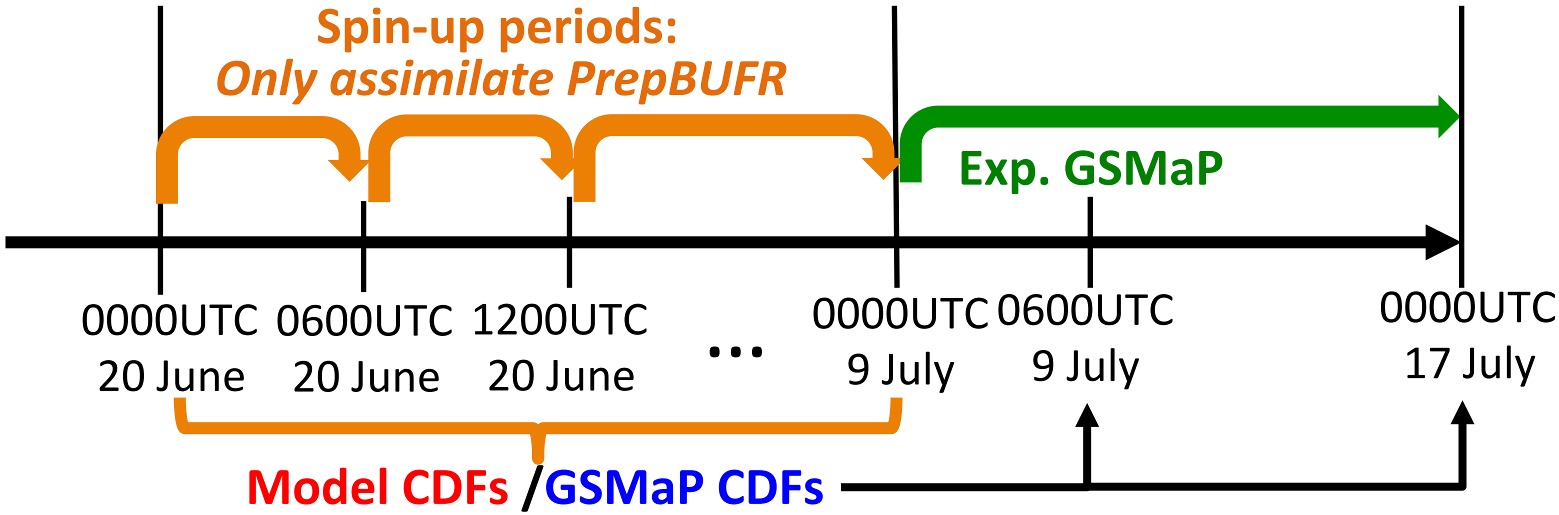
- Low resolution, in order to get some results within my internship period.
- We at first followed configurations of the SCALE NRT system (w/o KF). However, the precipitation fields w/o KF at 36km resolution is not similar to the GSMaP obs.

Experiment Settings: LETKF

LETKF settings	CTL	GSMaP
Ensemble size	100	
Inflation	Multiplicative inflation (1.25) RTPP (0.8)	
Vertical localization scale	0.3log(pressure)	
Horizontal localization scale	400km (PrepBUFR)/250km (GSMaP)	
Obs assimilated.	PrepBUFR	PrepBUFR + GSMaP
Threshold for precipitation		≥ $0.001 mm (6hr)^{-1}$
Precipitation CDFs		JUN 20 ~ JUL 6
Height of precipitation obs.		850mb
QC by min. number of precip. members		75/100 (75%)

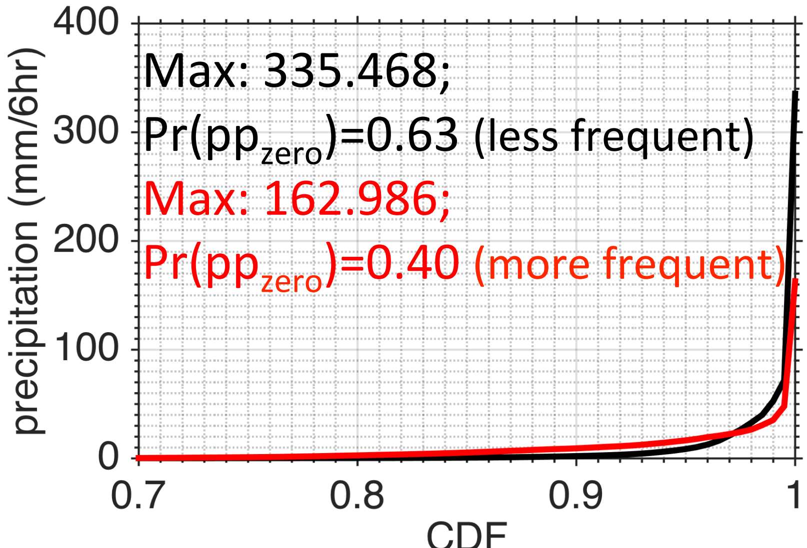


Construction of the Precipition CDFs



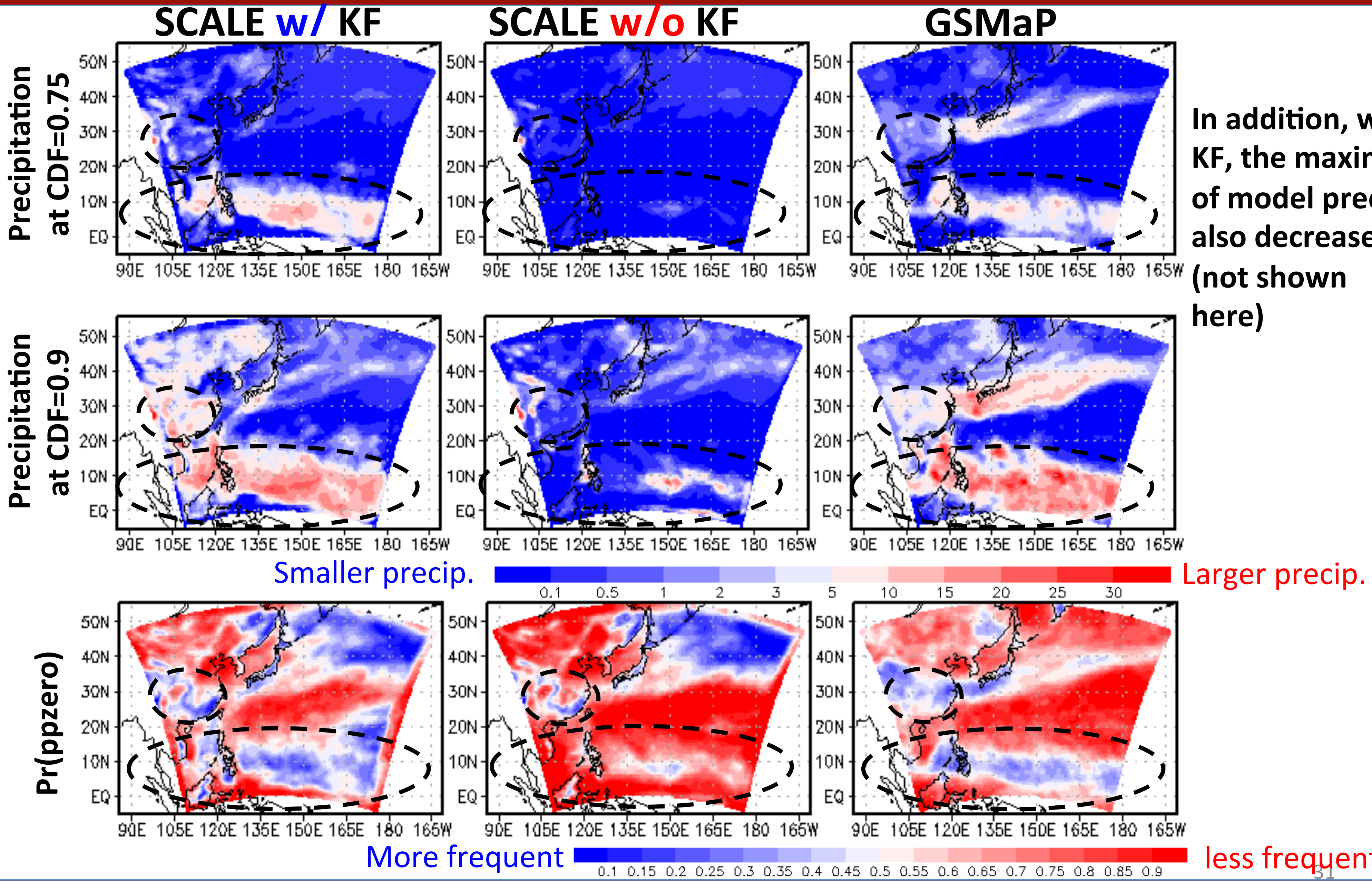
An example:

CDF for GSMaP & SCALE at one model grid



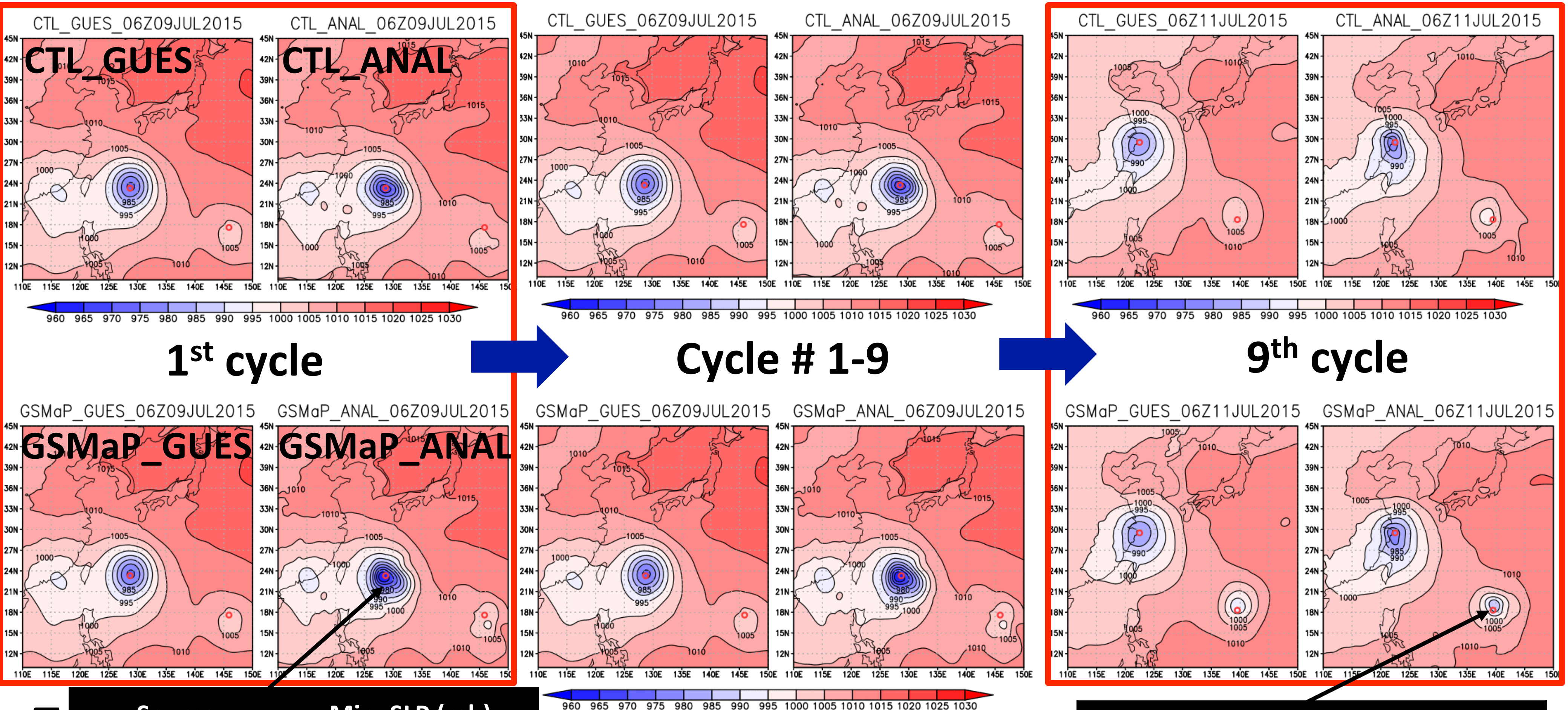
- If fixed at the same CDF value, we can get the snapshots of the model & GSMaP precipitation values for the entire domain.

Comparison of the Precipitation CDFs



In addition, with KF, the maxima of model precip. also decreases (not shown here)

Impact of GSMaP: Adjustments of the SLP



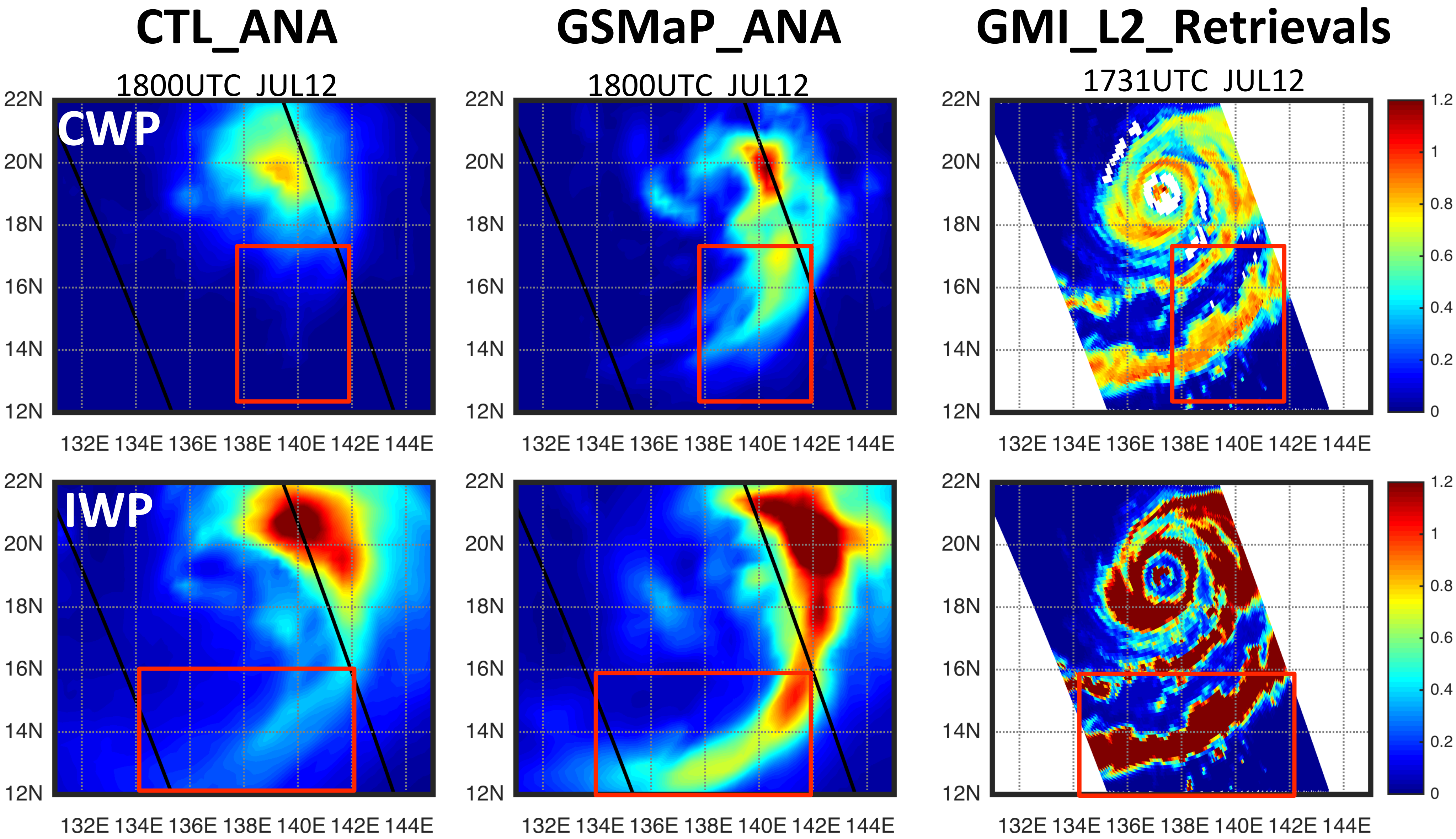
CHAN-HOM

Source	Min. SLP (mb)
Background	973.8
CTL_ANA	967.4
GSMaP_ANA	960.2
JMA Best Track	960

NANGKA

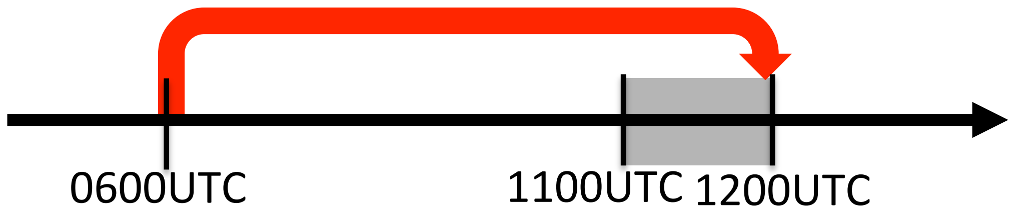
Source	Min. SLP (mb)
CTL_ANA	997.8
GSMaP_ANA	985.7
JMA Best Track	945

Adjustments of the Hydrometeors

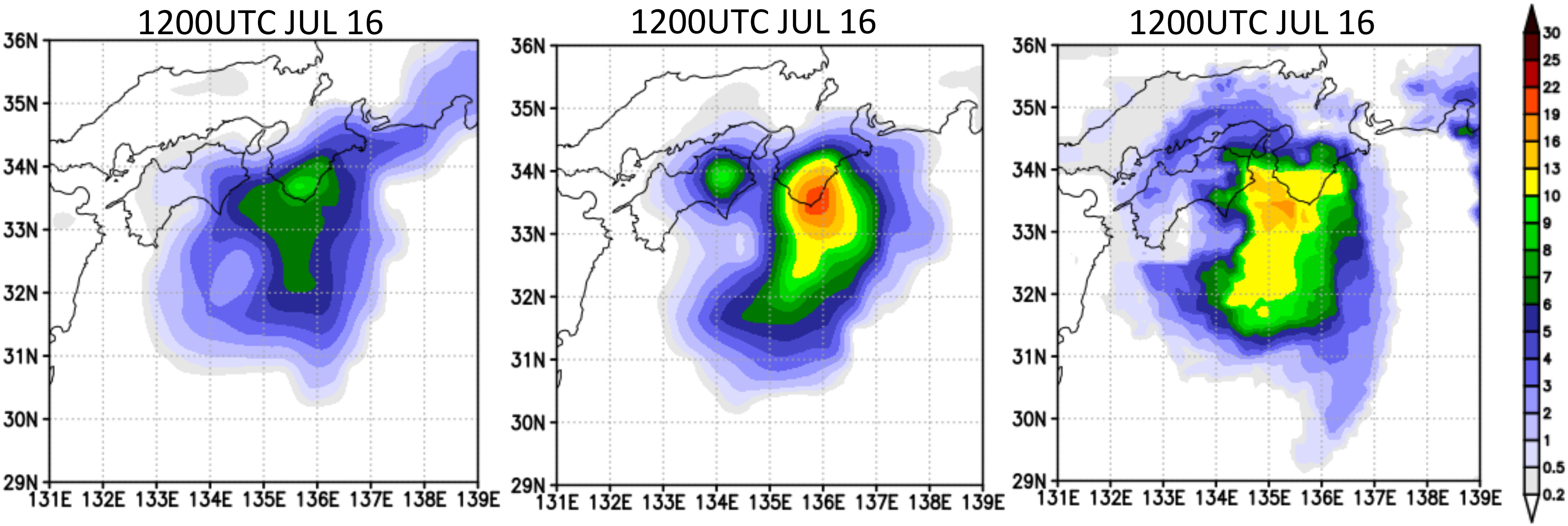


Accumulated 1-hr Precipitation Forecast

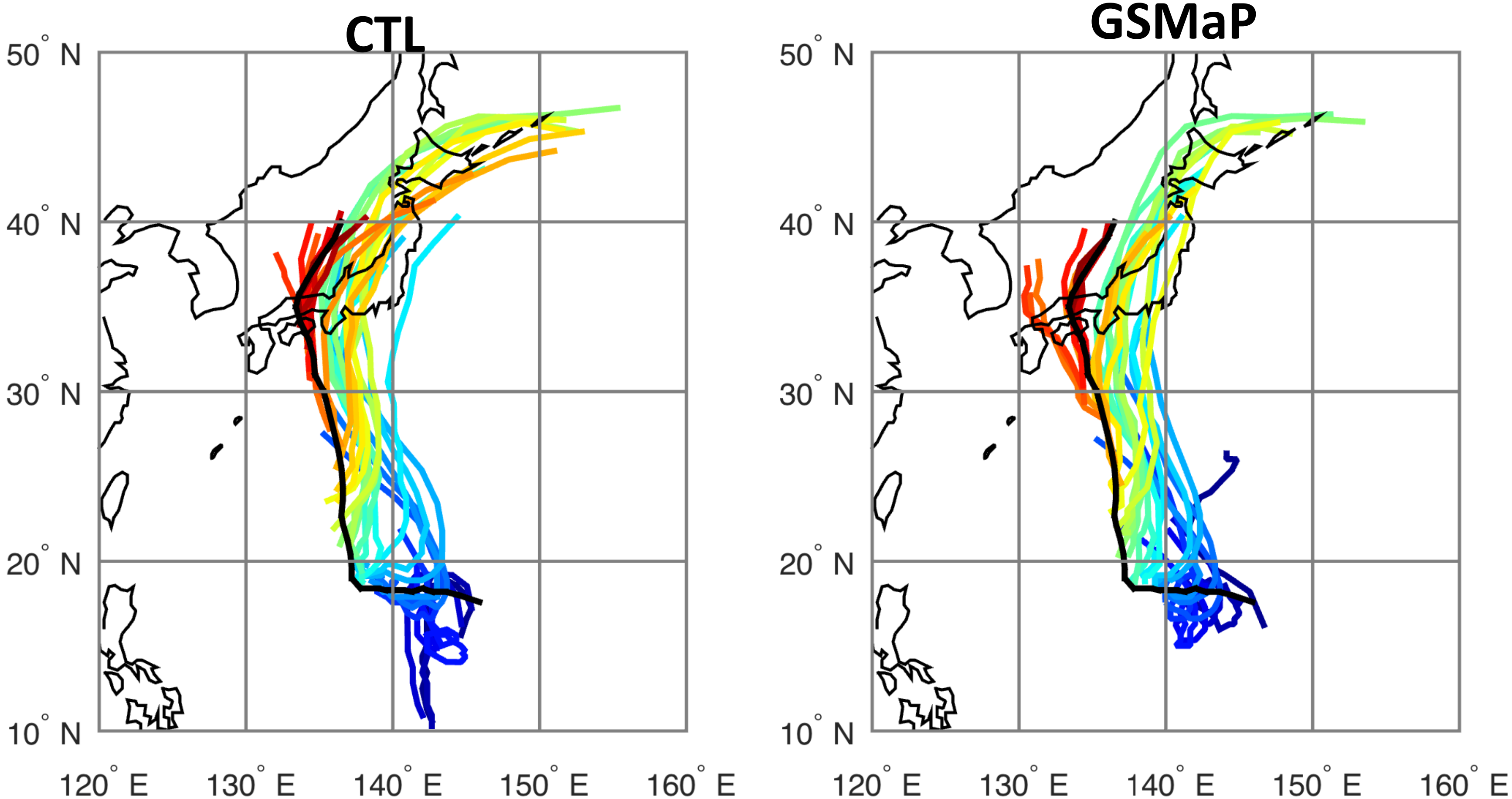
- 6-hr forecast initialized at 0600UTC JUL 16, and validated at 1200UTC, before Nangka made the landfall (1400UTC)



CTL_6HR_FCST **GSMaP_6HR_FCST** **GSMaP_OBS**



TC Track Forecast Error

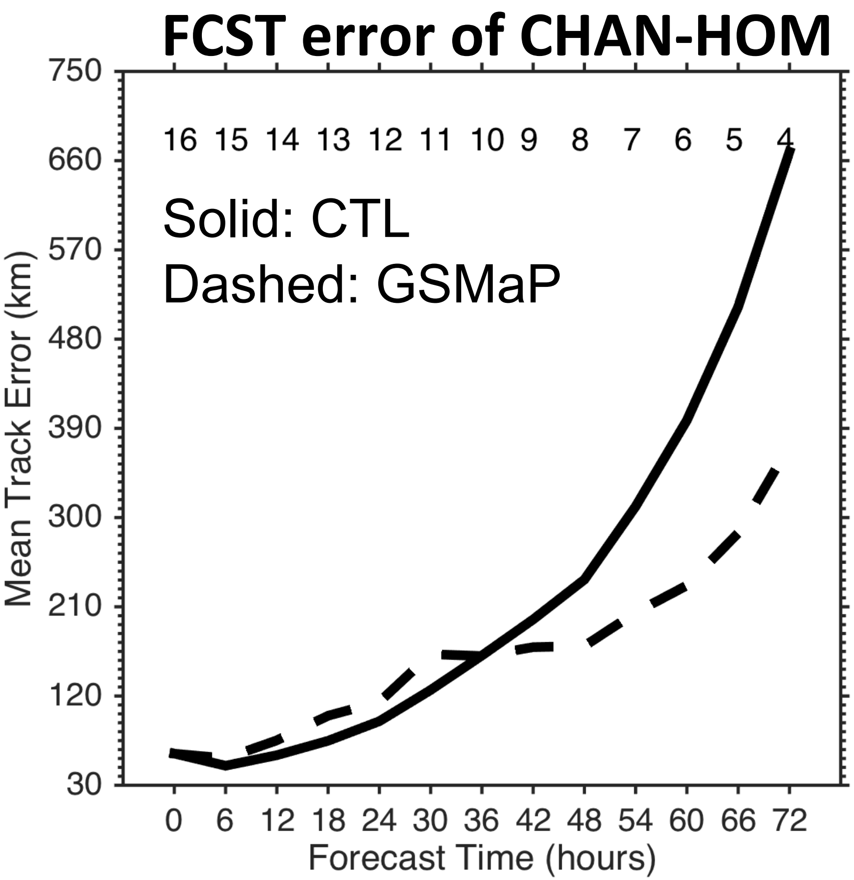


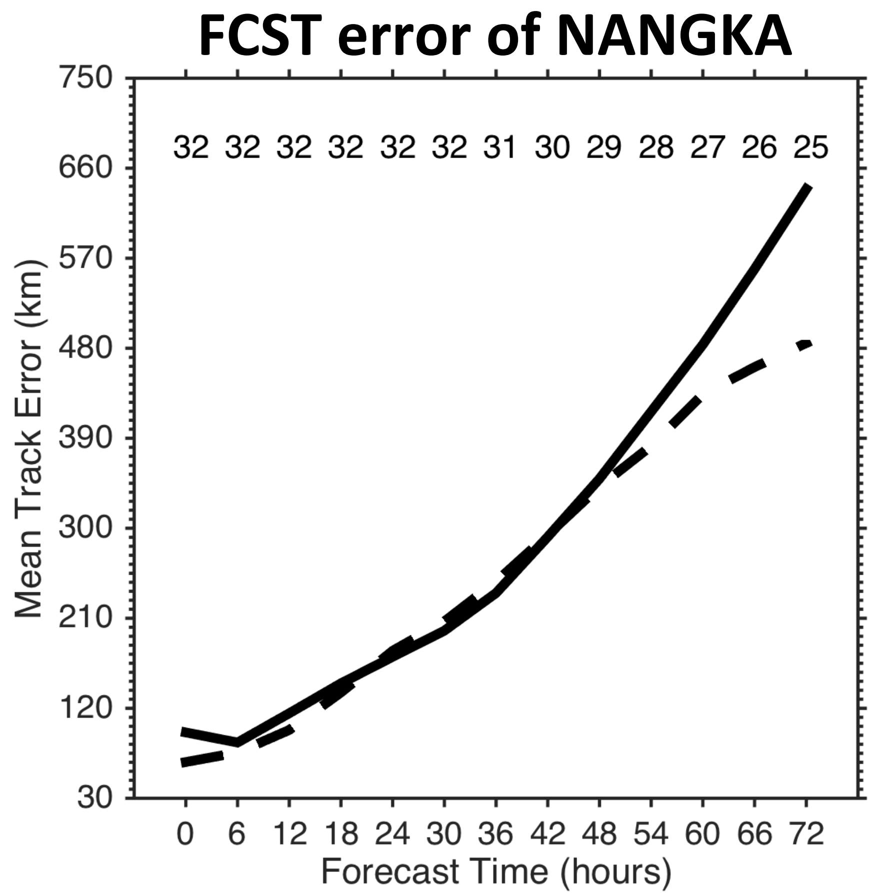
Thick black line: Best track

Color lines: 120-hr forecasts initialized at different time (warmer-> later INIT)

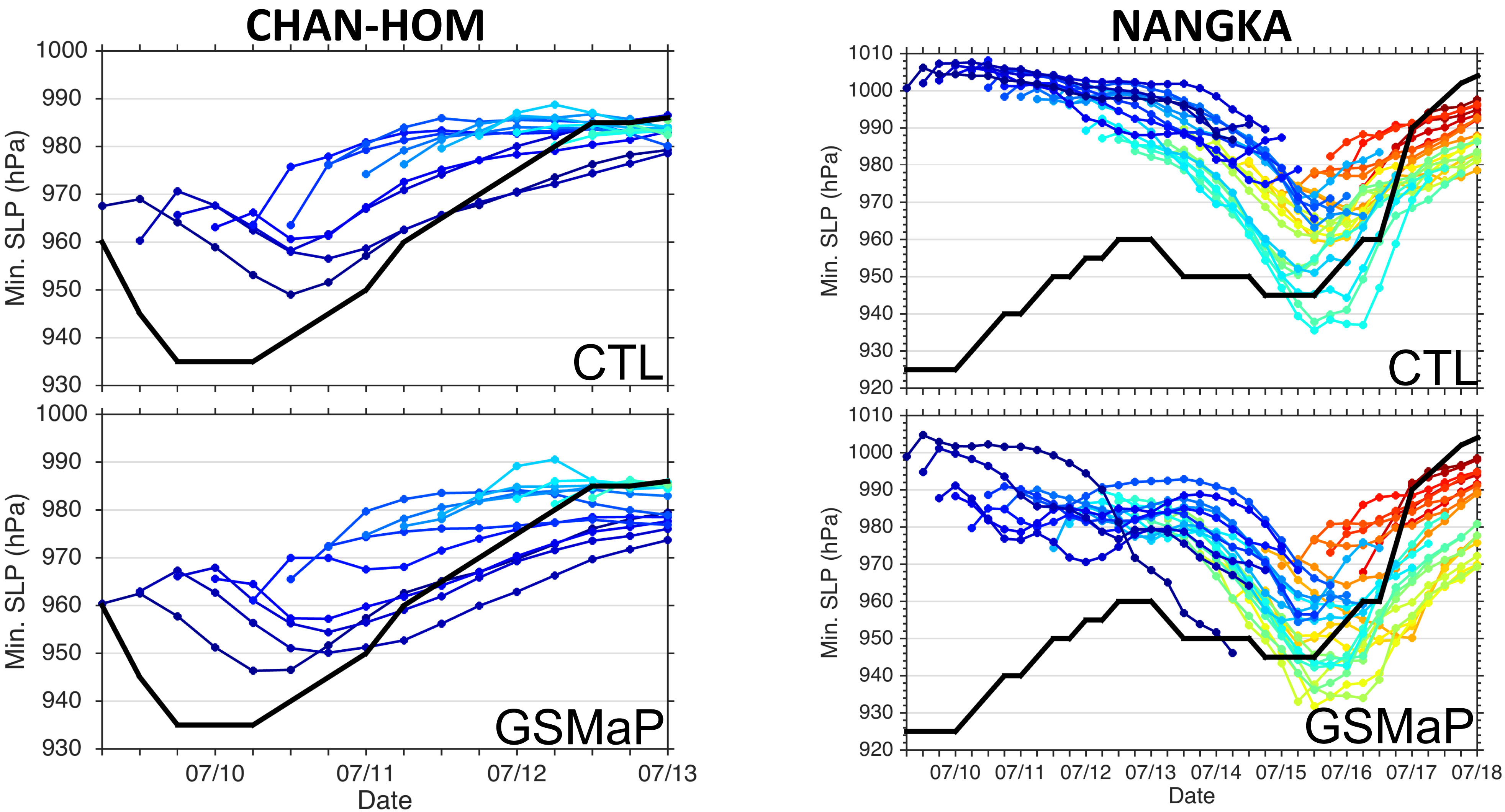
For the first several cycles, Nangka moves southward in the Exp. CTL, which does not occur in Exp. GSMaP.

Averaged TC Track Forecast Error





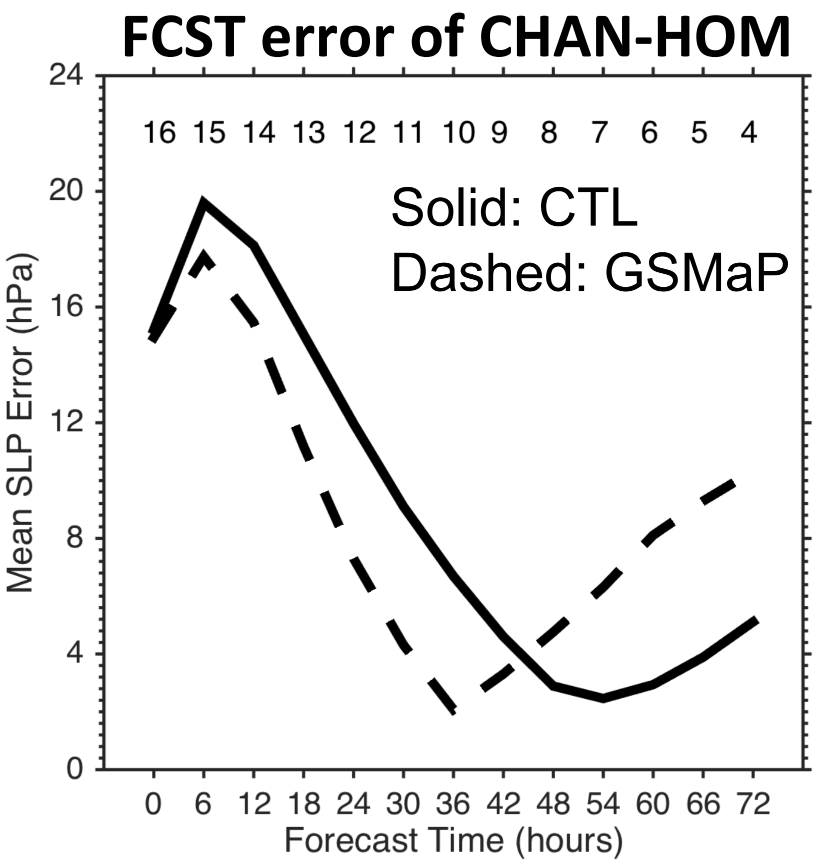
120-hr Intensity Forecast Initialized at Different Time

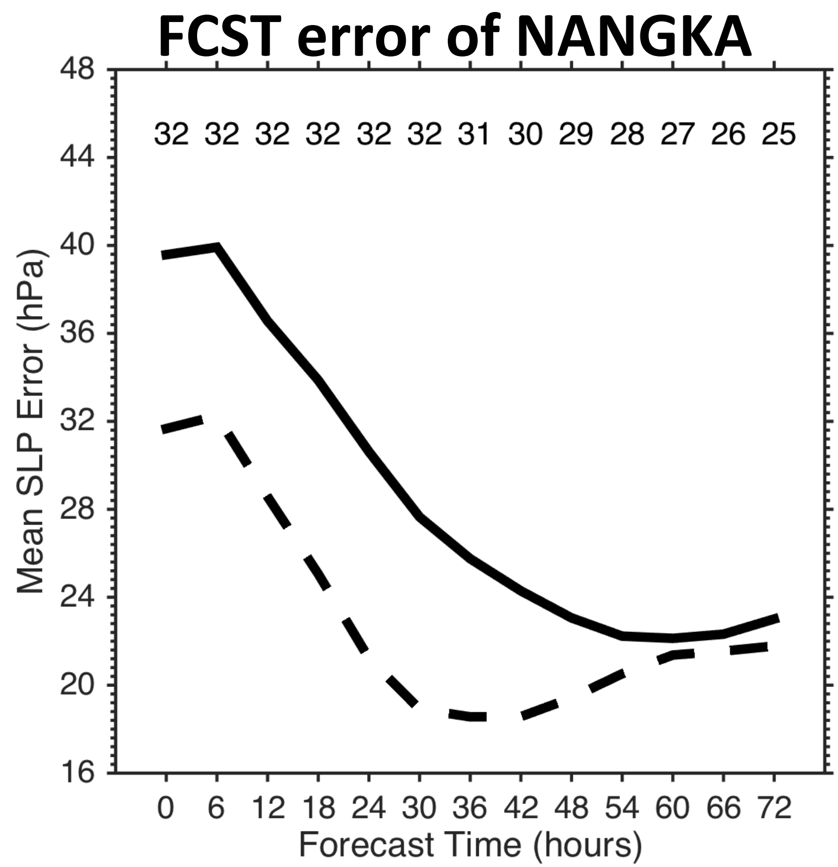


Thick black line: Best track

Color lines: 120-hr forecasts initialized at different time

Averaged TC Intensity Forecast Error





Sensitivity to the QC schemes

Number of Precipitating members (GSMaP_QC50):

1. Assimilation of GSMaP helps to intensify Nangka. However, not *strong* enough.

2. We assimilate GSMaP when at least 75% members precipitate (~0.75 in *Lien et al.*, 2016b; ~0.72 in *Kotsuki et al.*, 2017). The ensemble size in those studies: ~40.

3. But we have 100 members.

=> ***Will loosing this QC to 50 further Improve the forecast?***

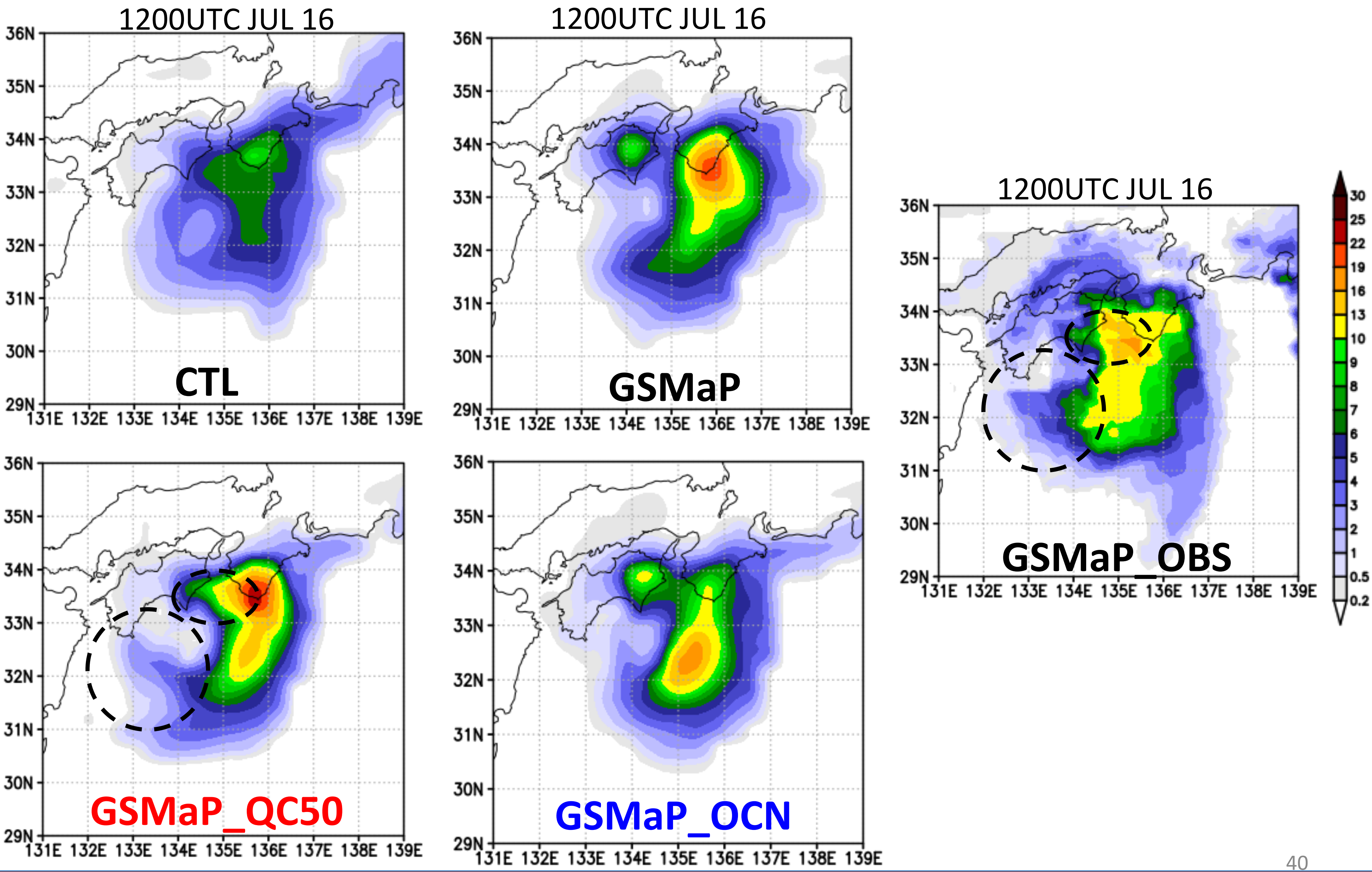
GSMaP over land (GSMaP_OCN):

1. **From the perspective of obs**: Lv3 GSMaP relies on Lv2 MW retrievals. However, the quality of MW precipitation over land is worse than those over ocean, since precipitation referred from the IWP from High-Frequency MW channels.

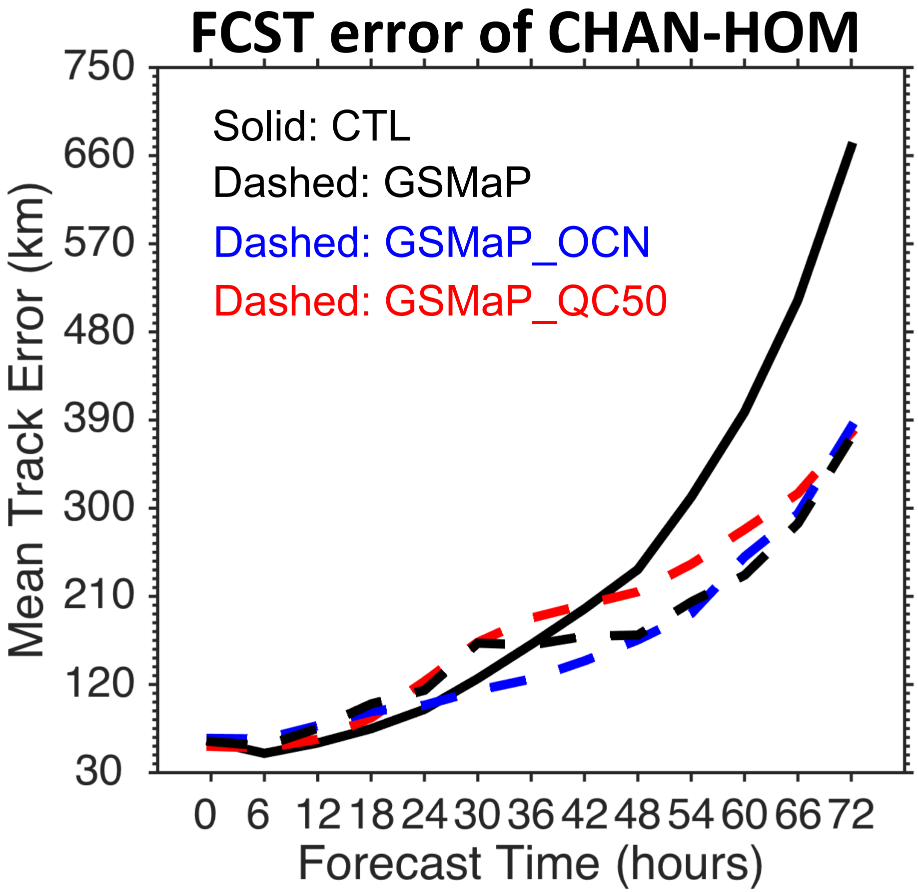
2. **From the perspective of model**: *Lien et al*. 2016a show that model simulation over land is not strongly correlated to the obs.

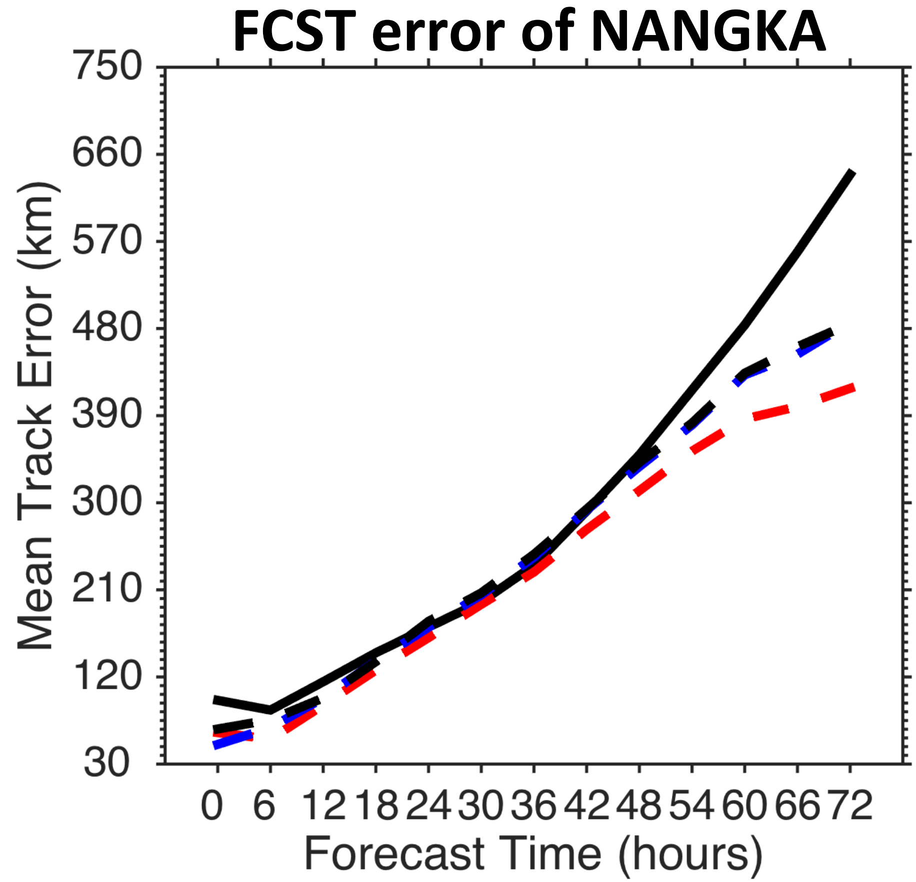
=> ***What if we only assimilate obs over ocean?***

Accumulated 1-hr Precipitation Forecast

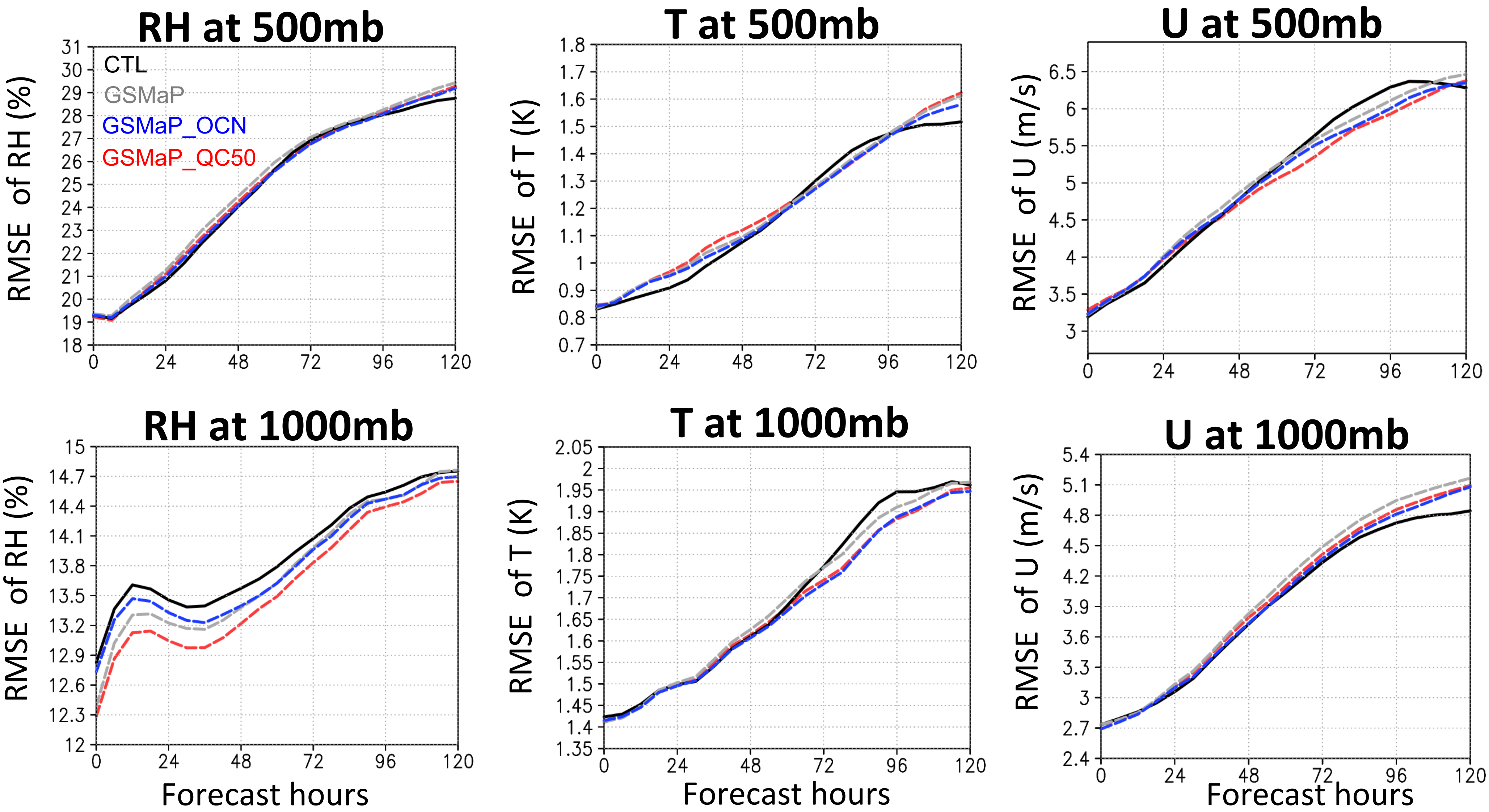


Sensitivity to the QC schemes





5-day Forecast RMSE of the Entire Domain



Summary

- This study investigates the impact of precipitation data within a regional NWP model. The precipitation DA component is implemented within the SCALE-LETKF system.
- We conduct two case studies on TC Chan-hom & Nangka in 2015. In both cases, 6-hr accumulated GSMaP data are assimilated with the Gaussian Transformation. The **preliminary** results show that:
  - ✓ The SLP/hydrometeor analysis with precip. DA is more close to obs.
  - ✓ 3-day typhoon track forecast error are decreased.
  - ✓ 5-day RH forecast at surface are improved.
  - ✓ By loosing the QC of precipitating members, and removing land obs, we could further improve the TC track forecasts.

Limitations

Current experiment settings:

- The resolution of the current model simulation is low.
  - 36km->18km
- Both cases does not start from the genesis phase.
  - New experiment starting before the genesis
- The obs might still be too dense.
  - Check the ensemble spread
- Too aggressive QC of the GSMaPs over land.
  - Further refine QCs

Methodology:

- Preparation of the CDFs requires a long period.
- Model-grid-based CDF might be not well suited for the TC applications.