

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

Cheng Da¹ and Guo-Yuan Lien²

¹ University of Maryland, College Park, USA

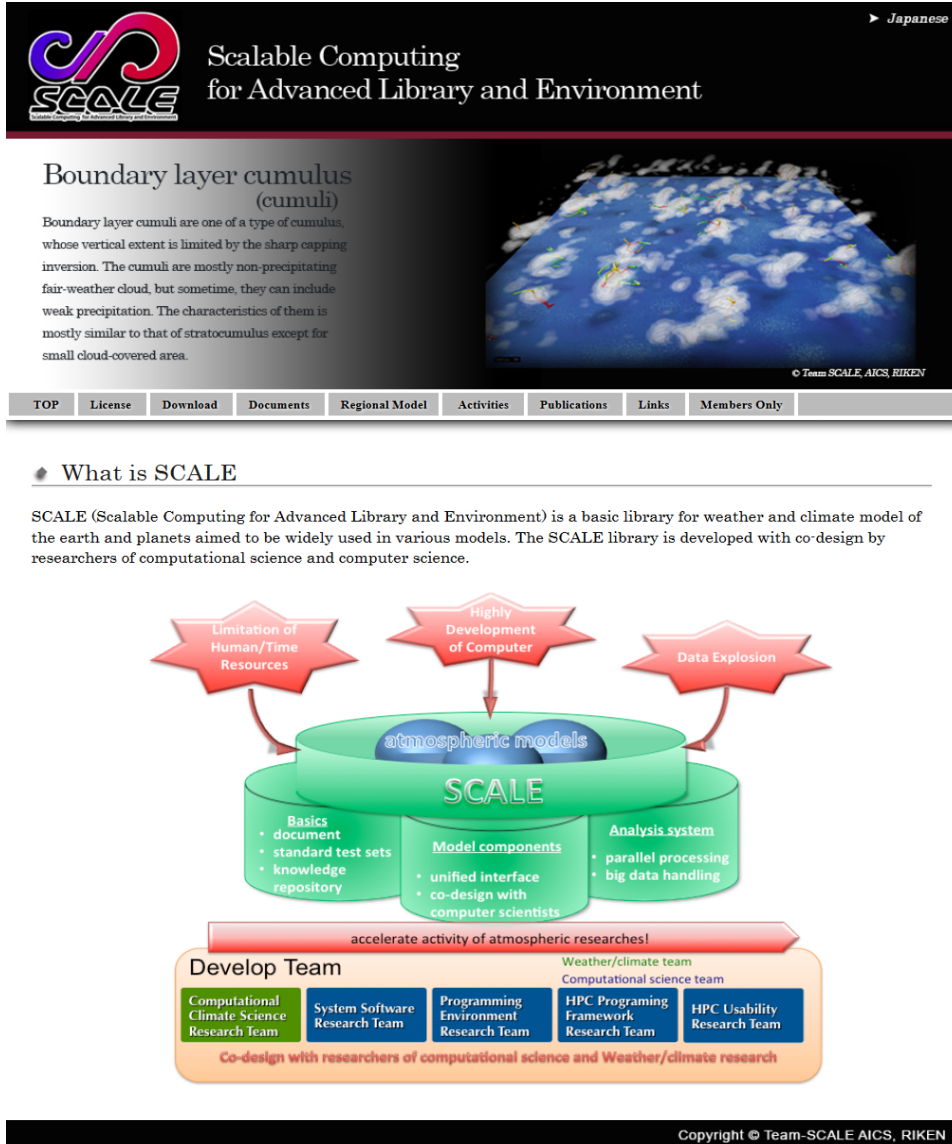
² RIKEN Advanced Institute for Computational Science, Kobe, Japan

➤ 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/>



The screenshot shows the SCALE website homepage. At the top left is the SCALE logo with the text "Scalable Computing for Advanced Library and Environment". Below this is a section titled "Boundary layer cumulus (cumuli)" with a descriptive paragraph and a 3D visualization of a cumulus cloud. A navigation menu includes links for "TOP", "License", "Download", "Documents", "Regional Model", "Activities", "Publications", "Links", and "Members Only". The main content area is titled "What is SCALE" and contains a paragraph describing the model. Below the text is a diagram showing the SCALE architecture. The diagram features a central green cylinder labeled "SCALE" with "atmospheric models" on top. Three red starburst shapes point to the top of the cylinder, labeled "Limitation of Human/Time Resources", "Highly Development of Computer", and "Data Explosion". The central cylinder is supported by three green cylindrical blocks: "Basics" (document, standard test sets, knowledge repository), "Model components" (unified interface, co-design with computer scientists), and "Analysis system" (parallel processing, big data handling). Below the diagram is a red arrow pointing right with the text "accelerate activity of atmospheric researches!". At the bottom, a "Develop Team" section lists five research teams: Computational Climate Science Research Team, System Software Research Team, Programming Environment Research Team, HPC Programming Framework Research Team, and HPC Usability Research Team. A note at the bottom of the diagram reads "Co-design with researchers of computational science and Weather/climate research".

Scalable Computing
for Advanced Library and Environment

Boundary layer cumulus (cumuli)
Boundary layer cumuli are one of a type of cumulus, whose vertical extent is limited by the sharp capping inversion. The cumuli are mostly non-precipitating fair-weather cloud, but sometime, they can include weak precipitation. The characteristics of them is mostly similar to that of stratocumulus except for small cloud-covered area.

TOP License Download Documents Regional Model Activities Publications Links Members Only

What is SCALE

SCALE (Scalable Computing for Advanced Library and Environment) is a basic library for weather and climate model of the earth and planets aimed to be widely used in various models. The SCALE library is developed with co-design by researchers of computational science and computer science.

Limitation of Human/Time Resources
Highly Development of Computer
Data Explosion

atmospheric models

SCALE

Basics
• document
• standard test sets
• knowledge repository

Model components
• unified interface
• co-design with computer scientists

Analysis system
• parallel processing
• big data handling

accelerate activity of atmospheric researches!

Develop Team

Weather/climate team
Computational science team

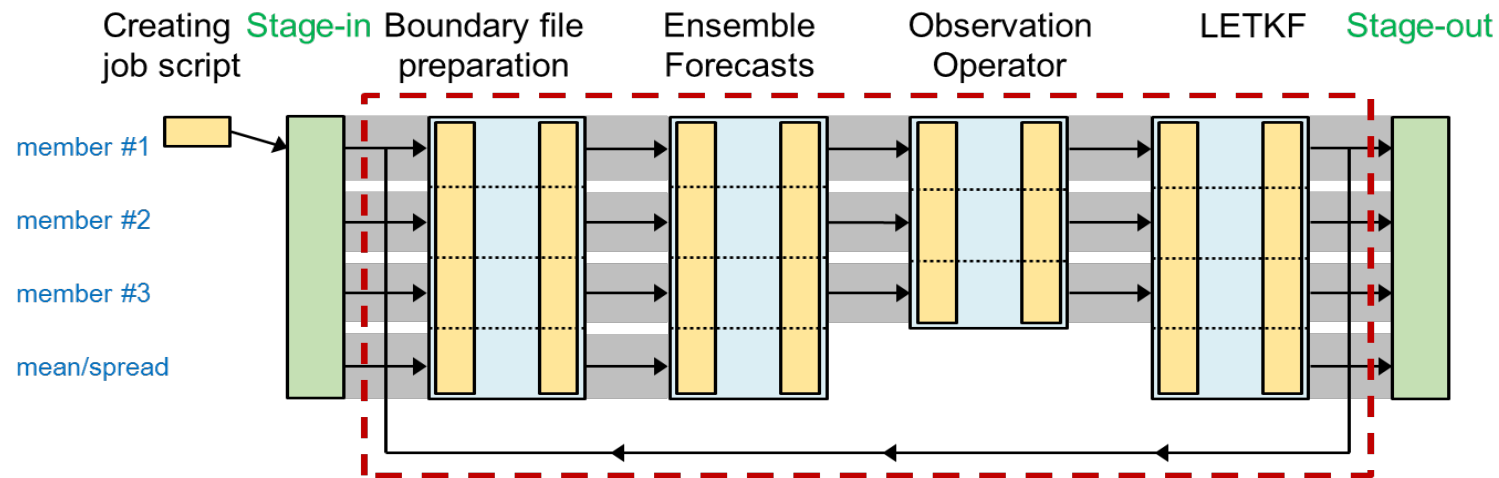
Computational Climate Science Research Team
System Software Research Team
Programming Environment Research Team
HPC Programming Framework Research Team
HPC Usability Research Team

Co-design with researchers of computational science and Weather/climate research

Copyright © Team-SCALE AICS, RIKEN

- Scalable Computing for Advanced Library and Environment (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.

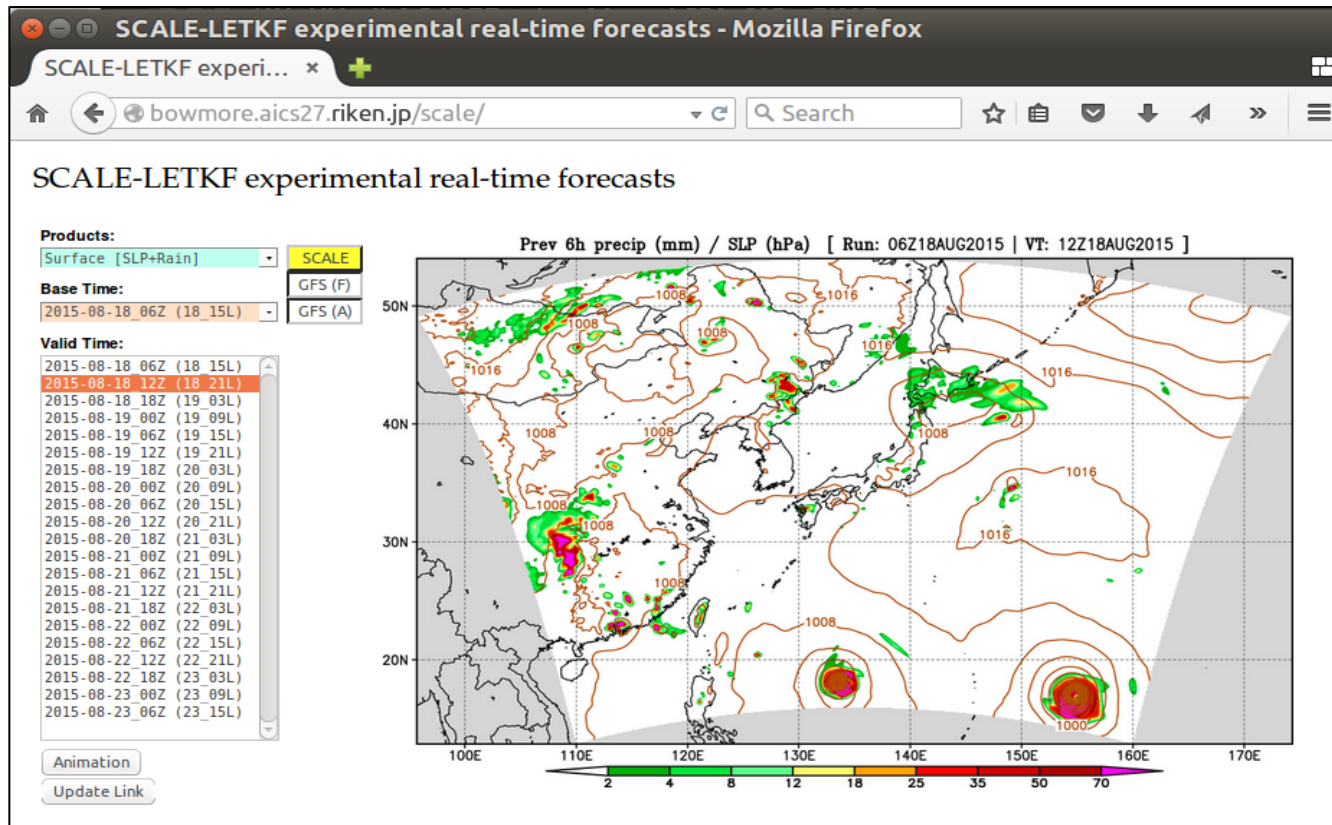
SCALE-LETKF (Lien et al. 2017)



<https://github.com/gyllien/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.

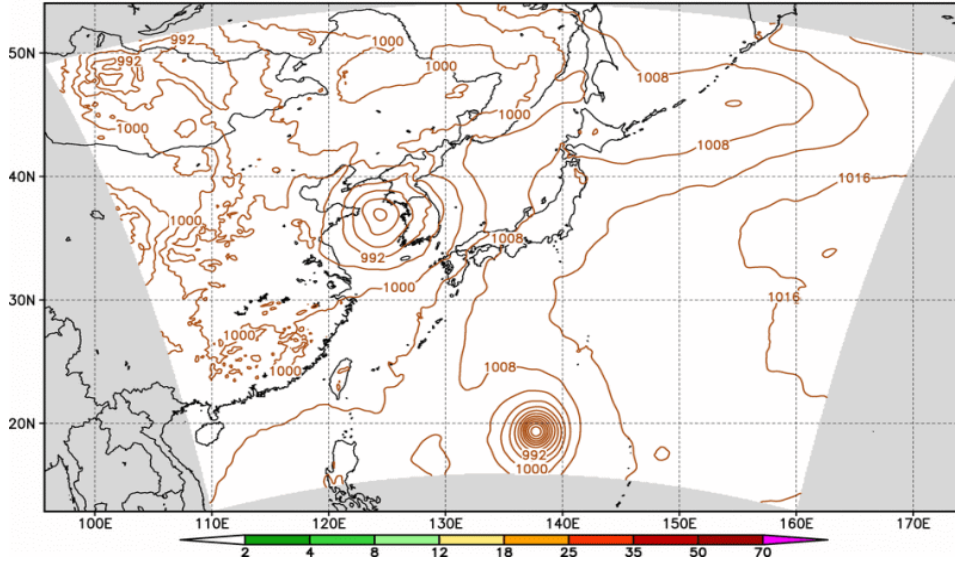
Experimental near-real-time SCALE-LETKF



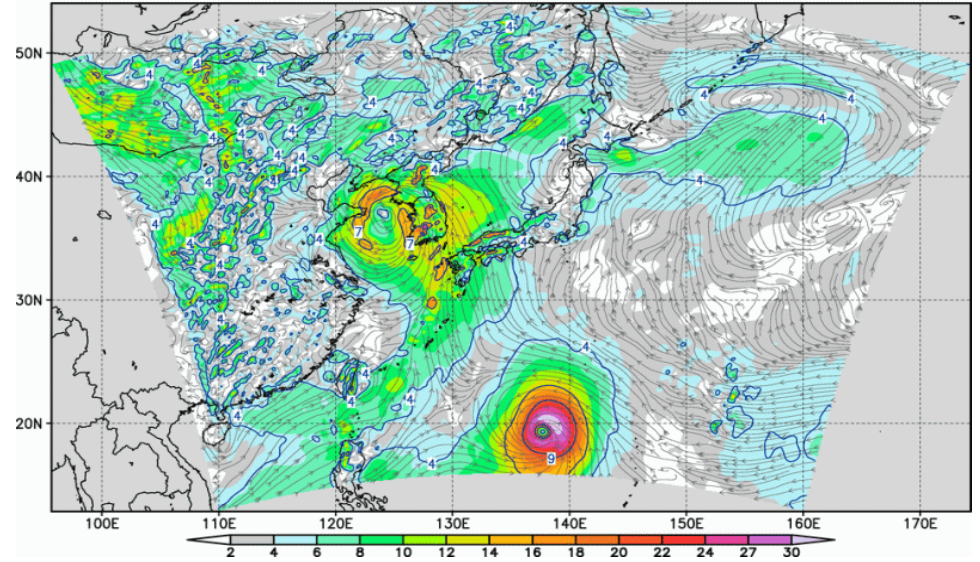
- 18-km resolution; 5760 x 4320 km area; 50 members
- 6-hour cycle; 5-day deterministic forecasts
- Test the **stability** of the system.
- Provide important guidance on the **performance of the model and the data assimilation settings**.

An example of 5-day forecasts (Typhoon NANGKA; 12:00 UTC July 12, 2015)

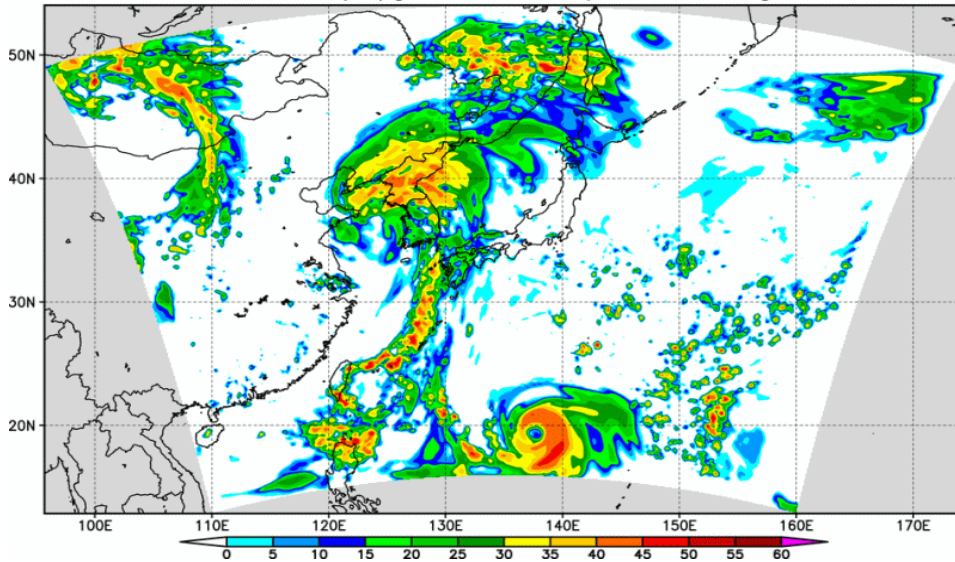
Prev 6h precip (mm) / SLP (hPa) [Run: 12Z12JUL2015 | VT: 12Z12JUL2015]



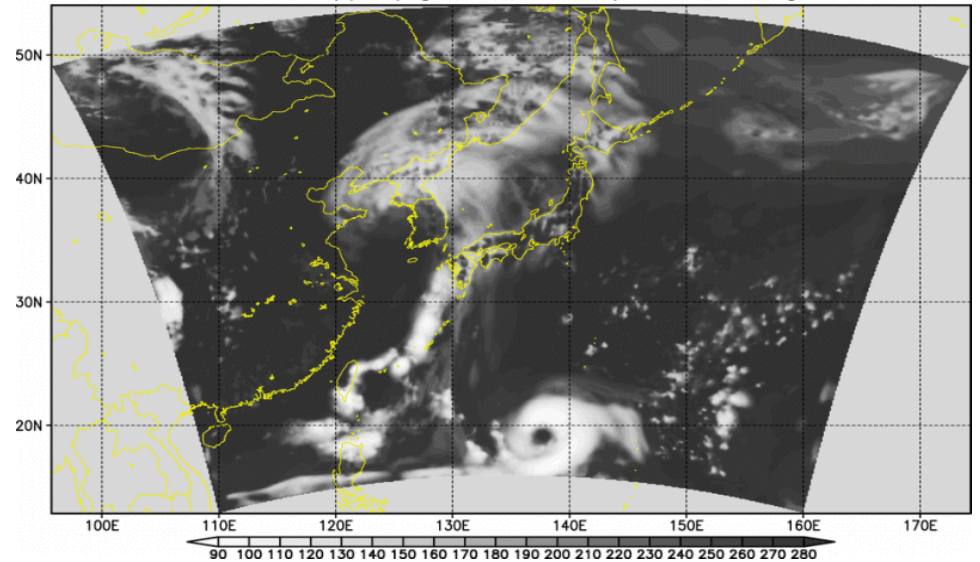
Sfc (10m) wind (m/s; contour: Beaufort scale) [Run: 12Z12JUL2015 | VT: 12Z12JUL2015]



Max reflex (dBZ) [Run: 12Z12JUL2015 | VT: 12Z12JUL2015]



Out lw rad (W/m^2) [Run: 12Z12JUL2015 | VT: 12Z12JUL2015]



➤ SCALE-LETKF

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

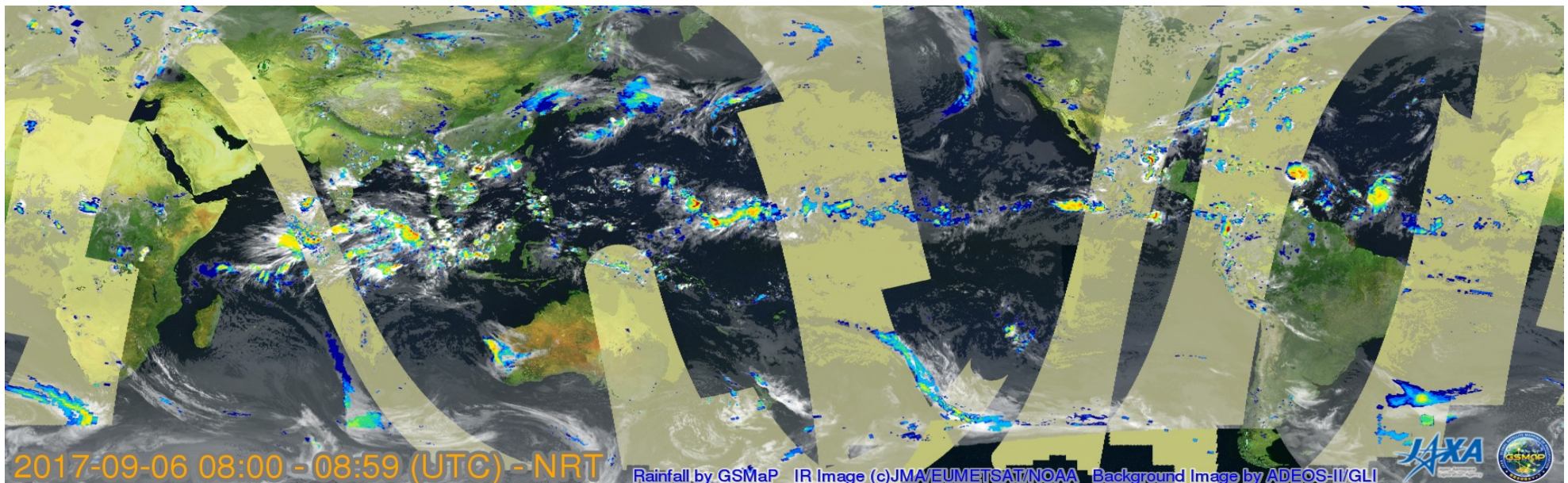
Satellite precipitation estimates

- 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)



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

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

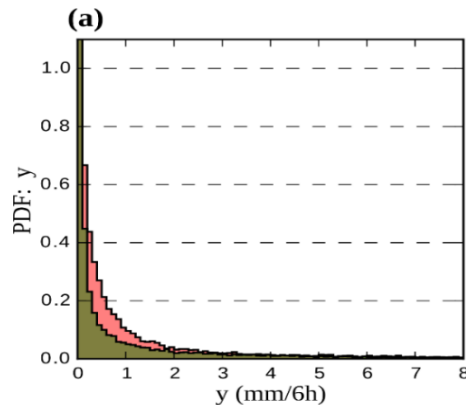
y : original variable (mm/6hr)

\tilde{y} : Transformed variable (sigma)

$F(\)$: CDF of original variable

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

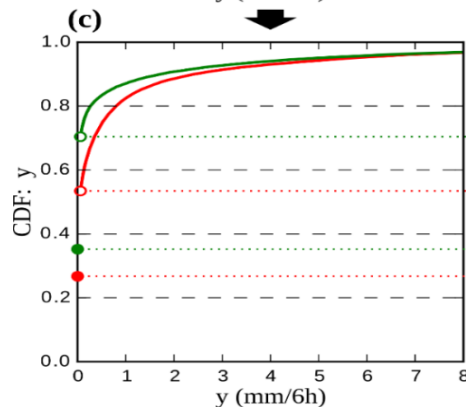
PDF



—: Model
—: Obs.

Step 0: Obtain PDF & CDF

CDF



Original variable

(Lien et al. 2013, 2016b)

Gaussian transformation

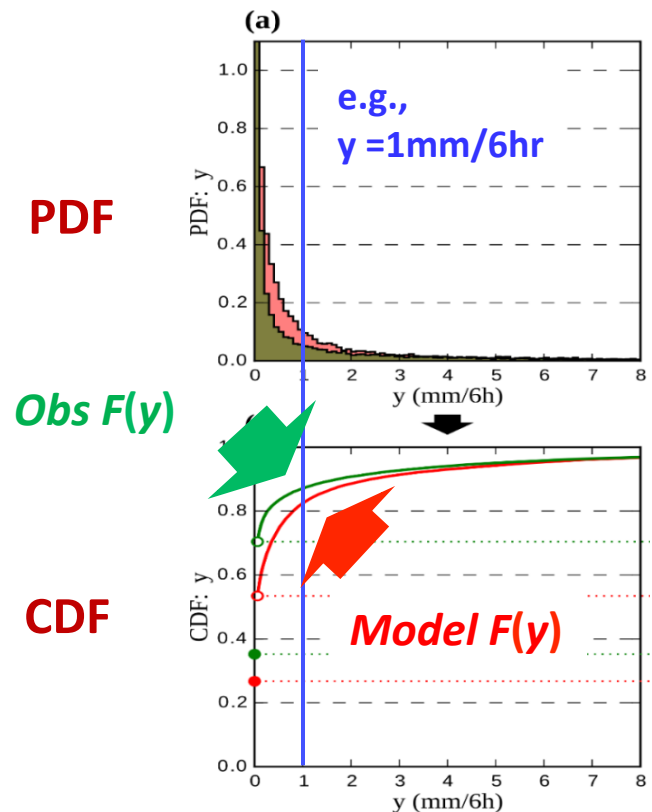
$$F^G(\tilde{y}) = F(y) \Rightarrow \tilde{y} = F^G{}^{-1}[F(y)]$$

y : original variable (mm/6hr)

\tilde{y} : Transformed variable (sigma)

$F(\)$: CDF of original variable

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



Original variable

—: Model
—: Obs.

Step 0: Obtain PDF & CDF

Step 1: Compute $F(y)$

(Lien et al. 2013, 2016b)

Gaussian transformation

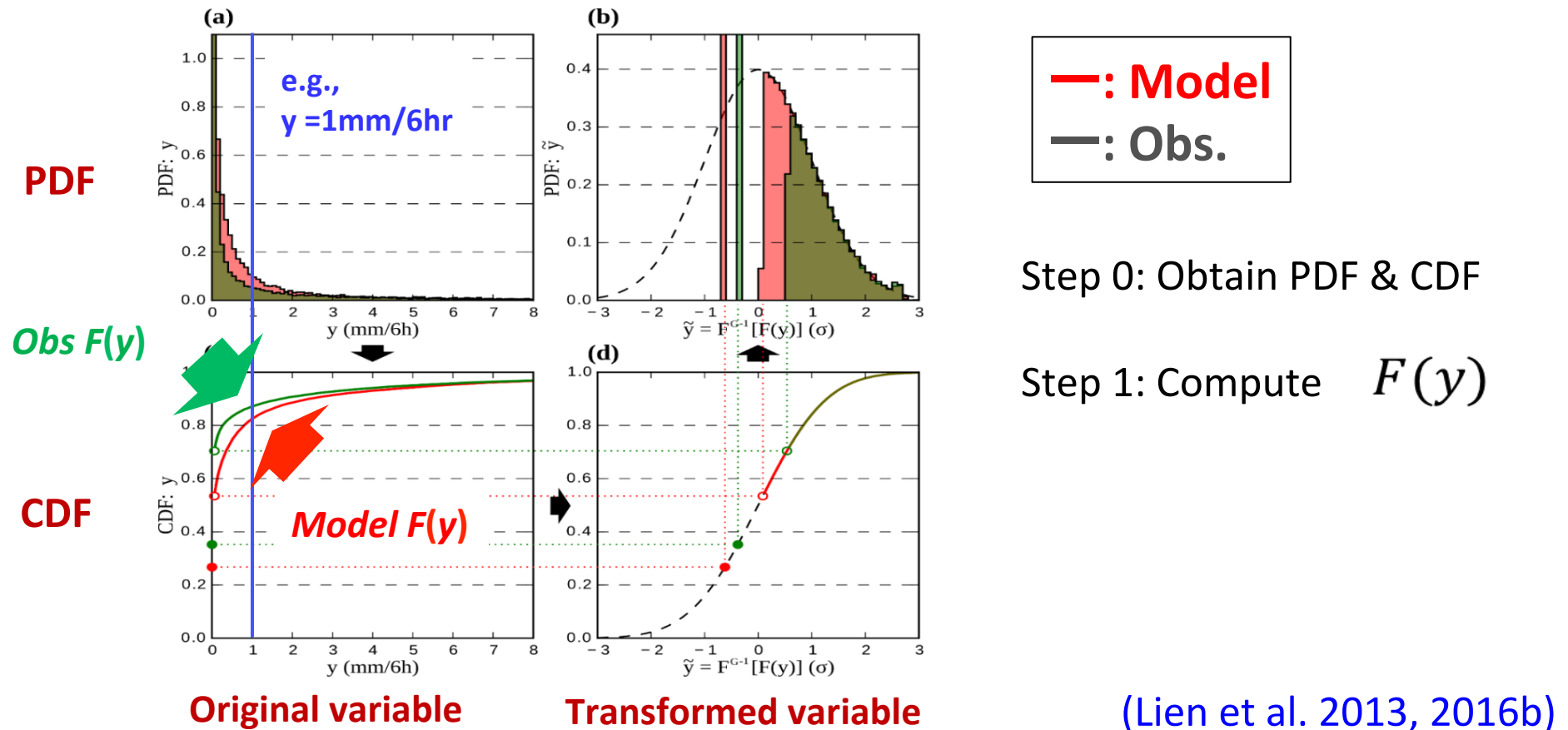
$$F^G(\tilde{y}) = F(y) \Rightarrow \tilde{y} = F^{G^{-1}}[F(y)]$$

y : original variable (mm/6hr)

\tilde{y} : Transformed variable (sigma)

$F(\)$: CDF of original variable

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



Gaussian transformation

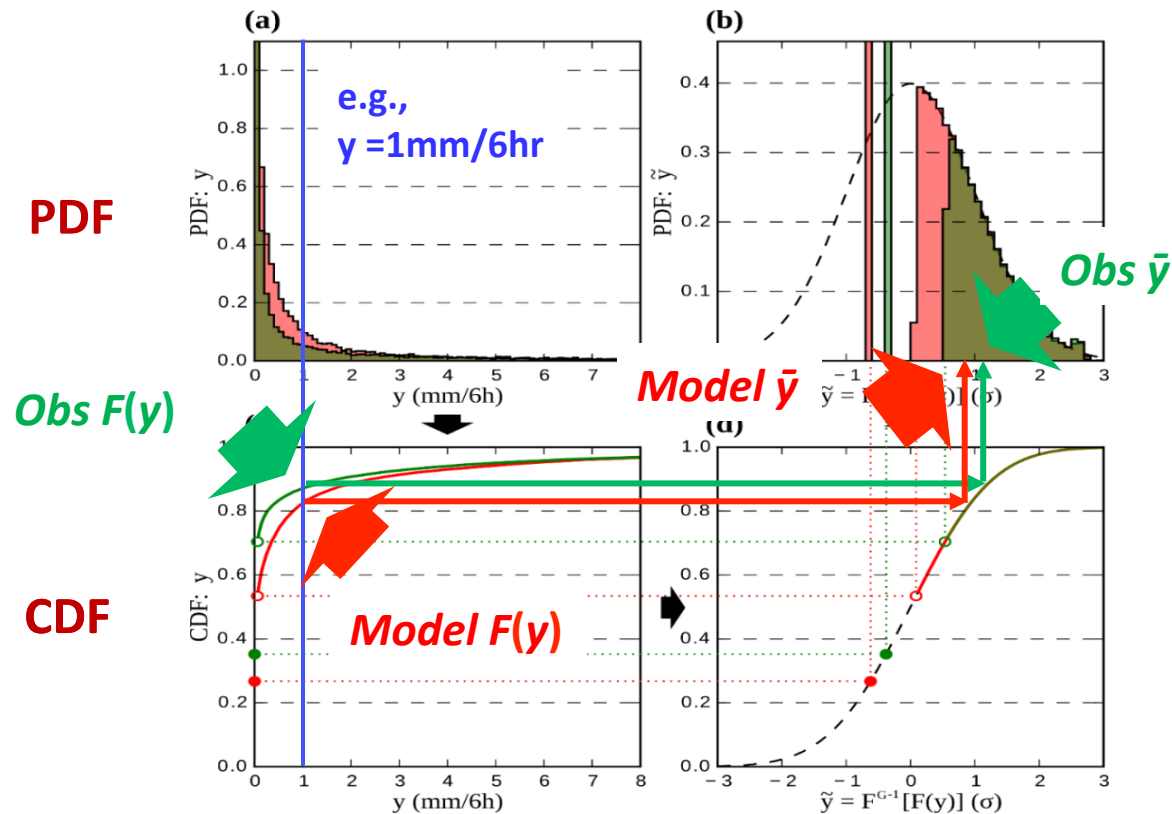
$$F^G(\tilde{y}) = F(y) \Rightarrow \tilde{y} = F^{G^{-1}}[F(y)]$$

y : original variable (mm/6hr)

\tilde{y} : Transformed variable (sigma)

$F(\)$: CDF of original variable

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



—: Model
—: Obs.

Step 0: Obtain PDF & CDF

Step 1: Compute $F(y)$

Step 2: Compute $\tilde{y} = F^{G^{-1}}[F(y)]$

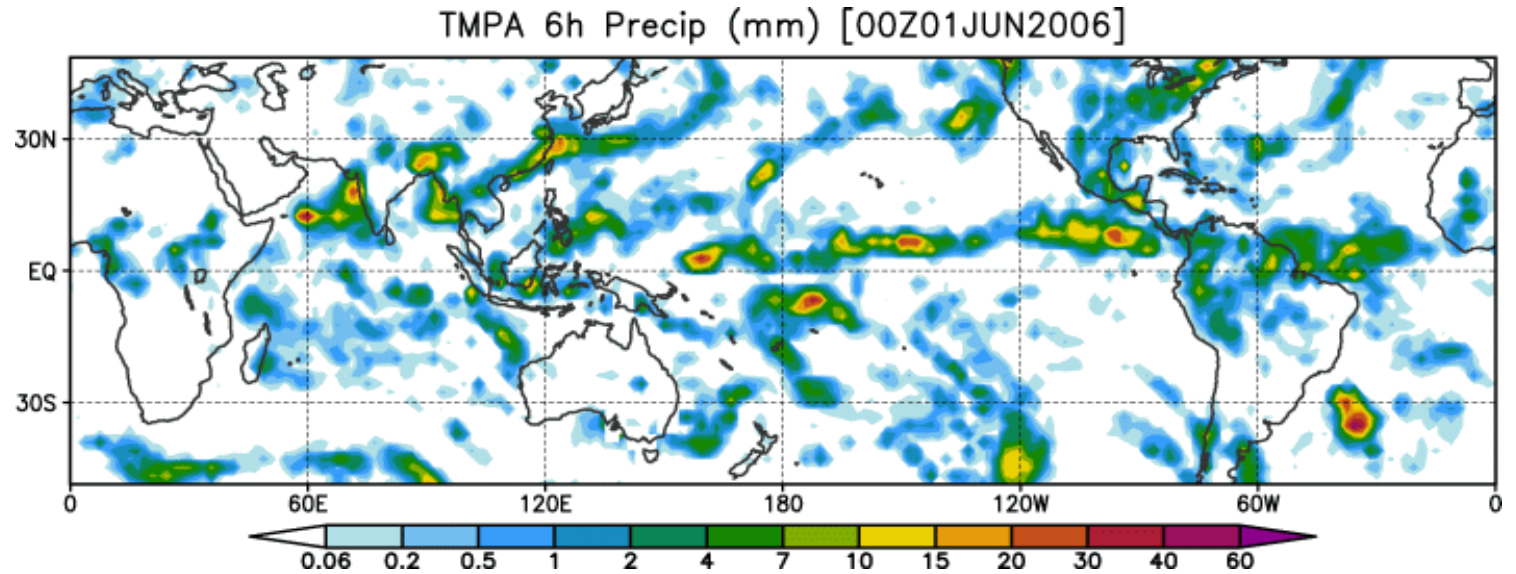
Original variable

Transformed variable

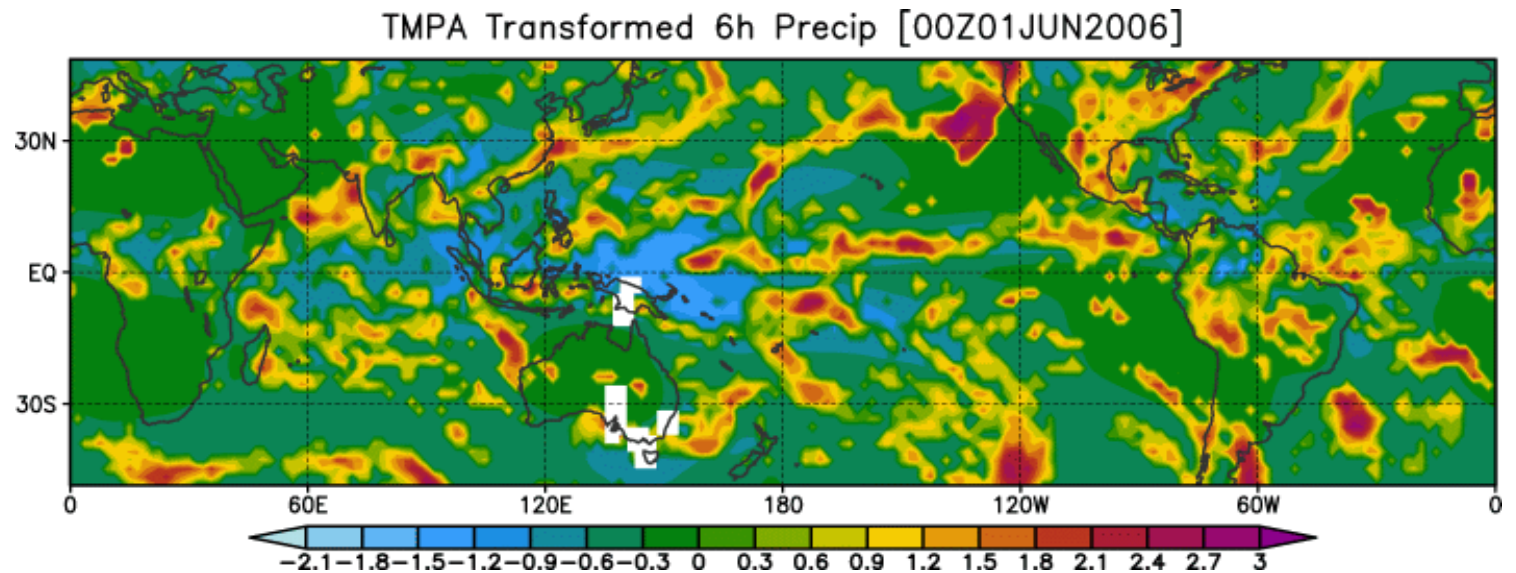
(Lien et al. 2013, 2016b)

Gaussian transformation

**Original
variable**



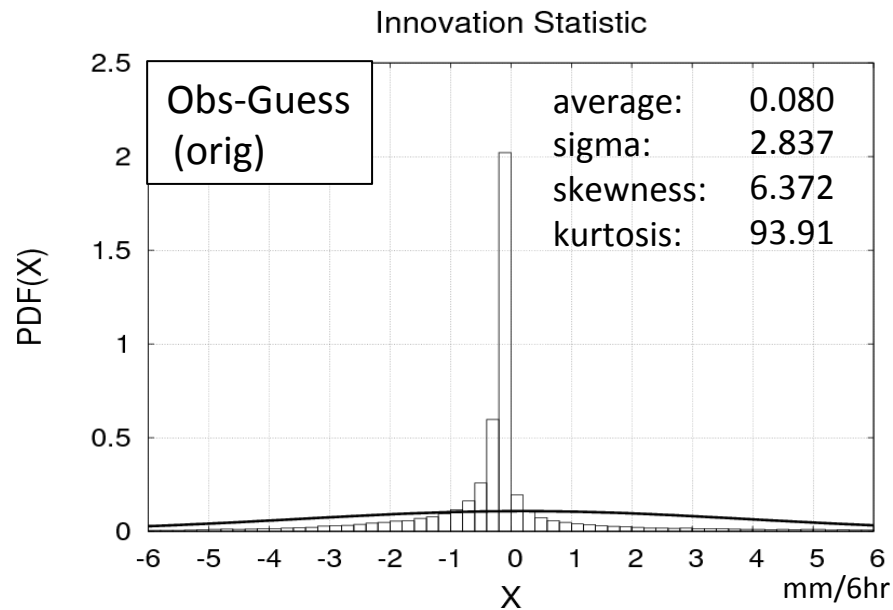
**Transformed
variable**



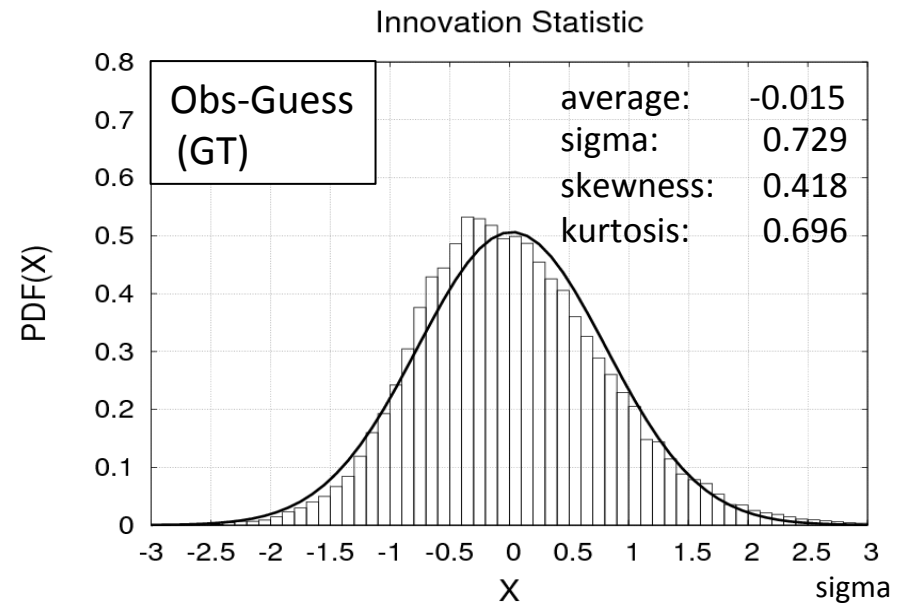
Gaussian transformation: Error distribution

(Kotsuki et al. 2017, JGR-A)

wo Gaussian-Transformation



w Gaussian-Transformation



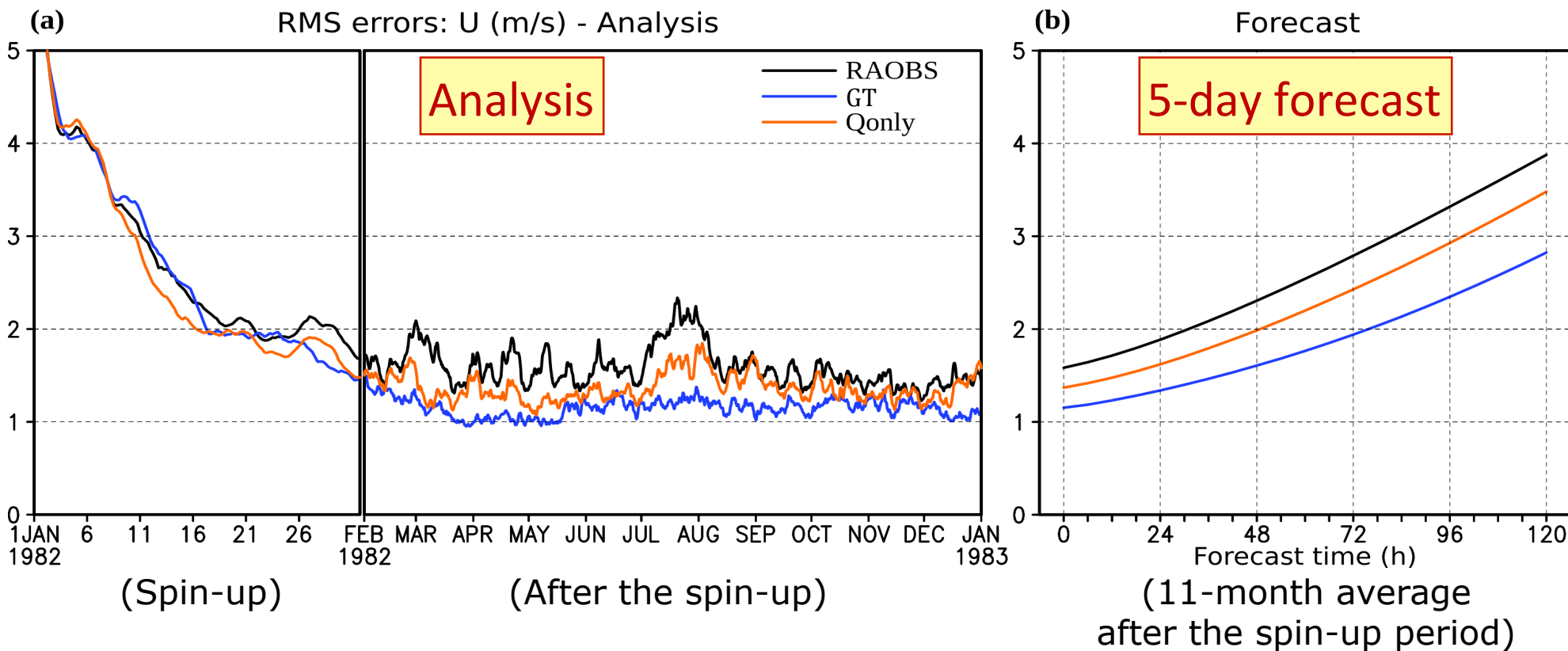
More Gaussian

Sampling period : 2014110100 - 2014110118

Construction of the precipitation CDF

- 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

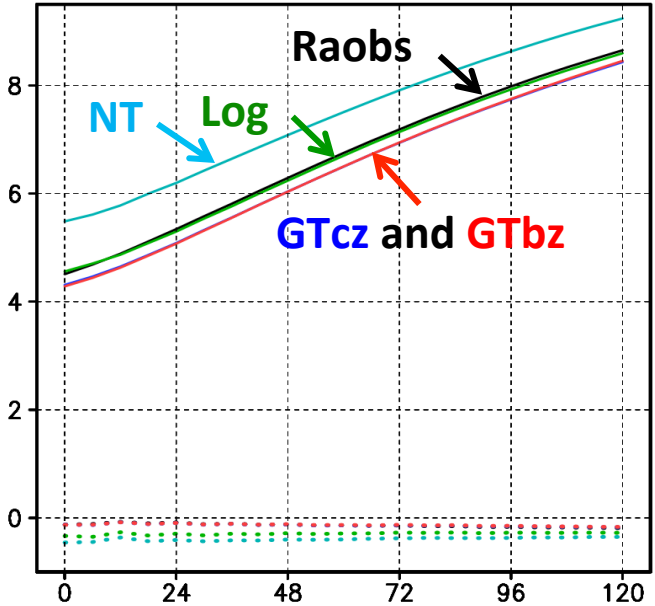
(Other variables show similar results)

Lien et al. 2016a, b:

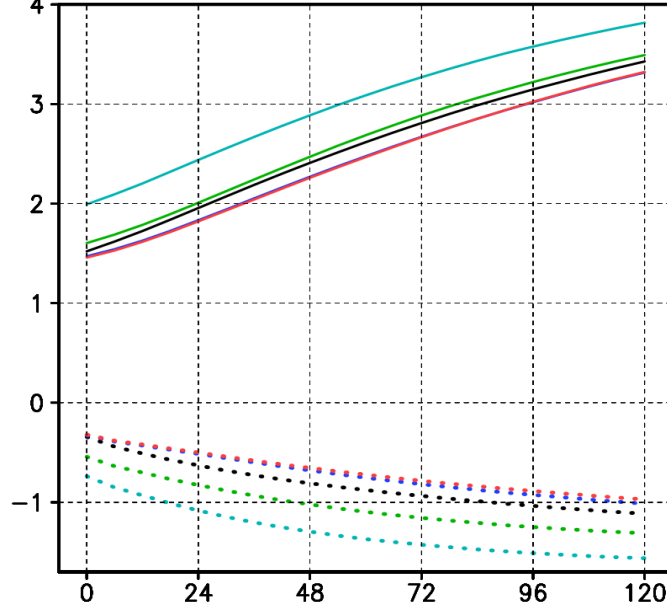
GFS model / TMPA observation

5-day forecasts

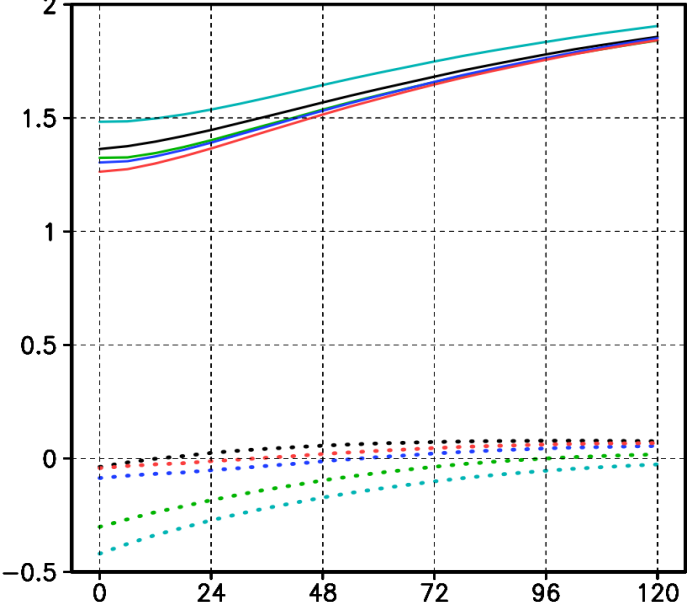
(a) RMSE/Bias [GL]: U (m/s) at 500hPa



(b) RMSE/Bias [GL]: T (K) at 500hPa



(c) RMSE/Bias [GL]: Q (g/kg) at 700hPa



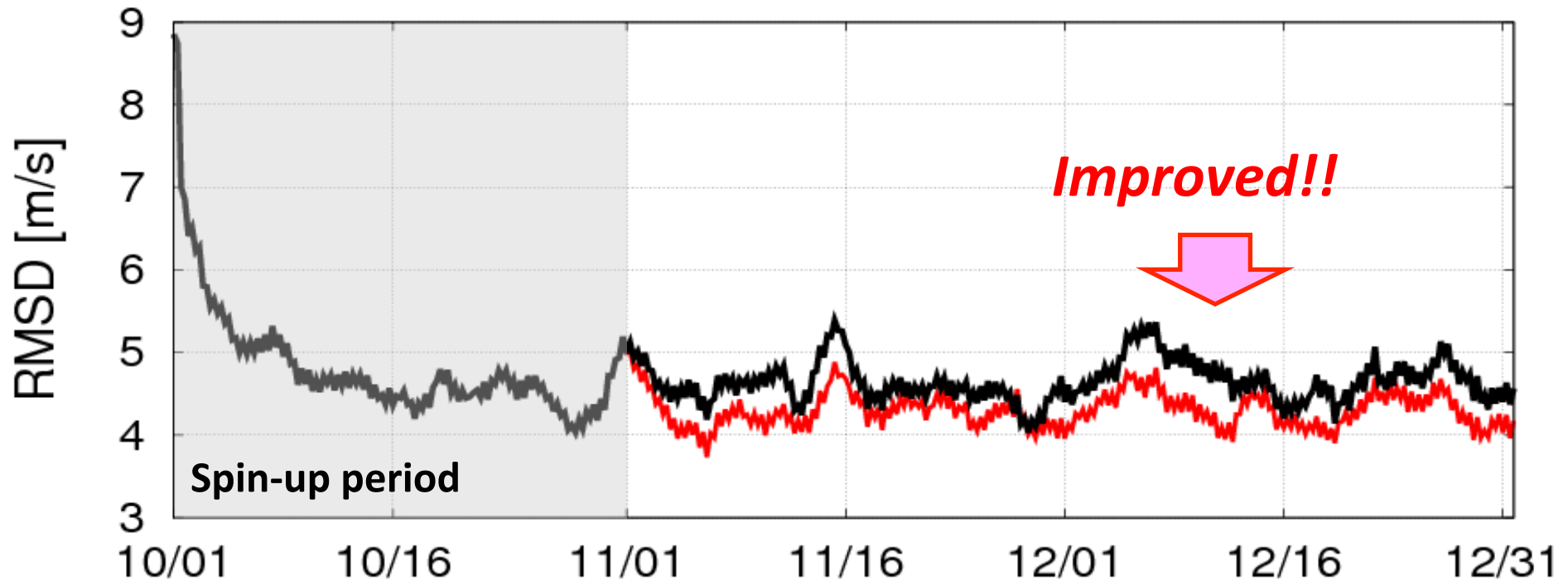
— RAOBS — NT — Log — GTcz — GTbz

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

U vs. ERA Interim (500 hPa, Global)



—: 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

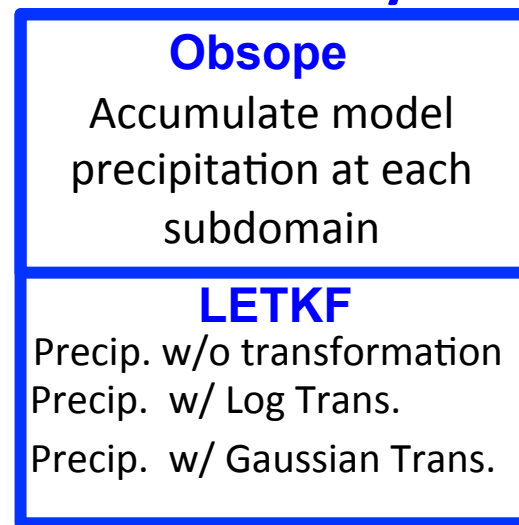
- The SCALE-LETKF system (*Lien et al., 2017*) doesn't have the precipitation DA component yet. So we need to implement it first.

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
/
```

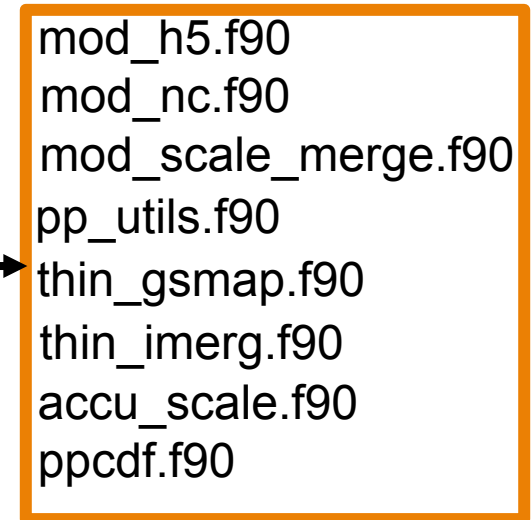
Precip. DA component:

SCALE-LETKF system



Precip. Preprocessing Utilities:

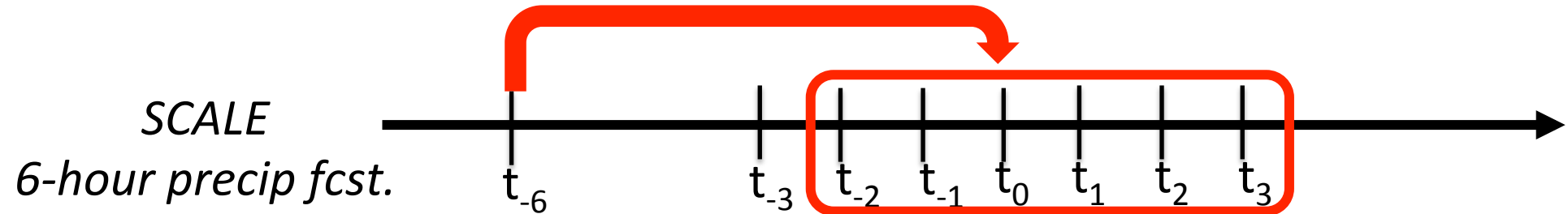
for both JAXA GSMaP &
NASA IMERG



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

Obs Operators / Obs

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

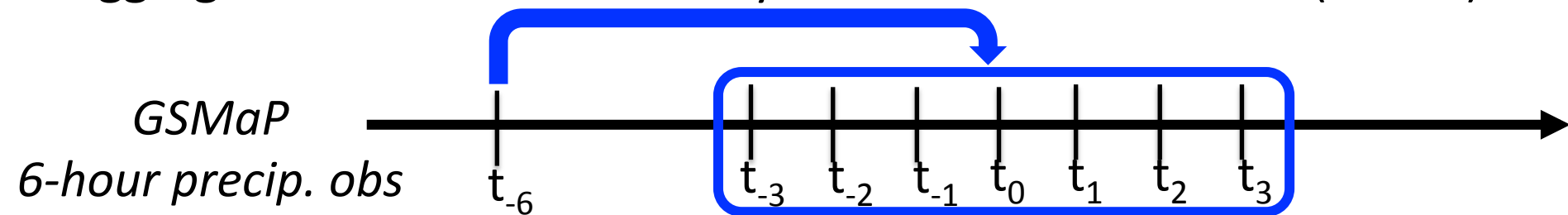


$$\text{Obs Operator } H(x): pp_{6hr}^{SCALE} = \sum_{i=-2}^3 pp_i$$

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

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

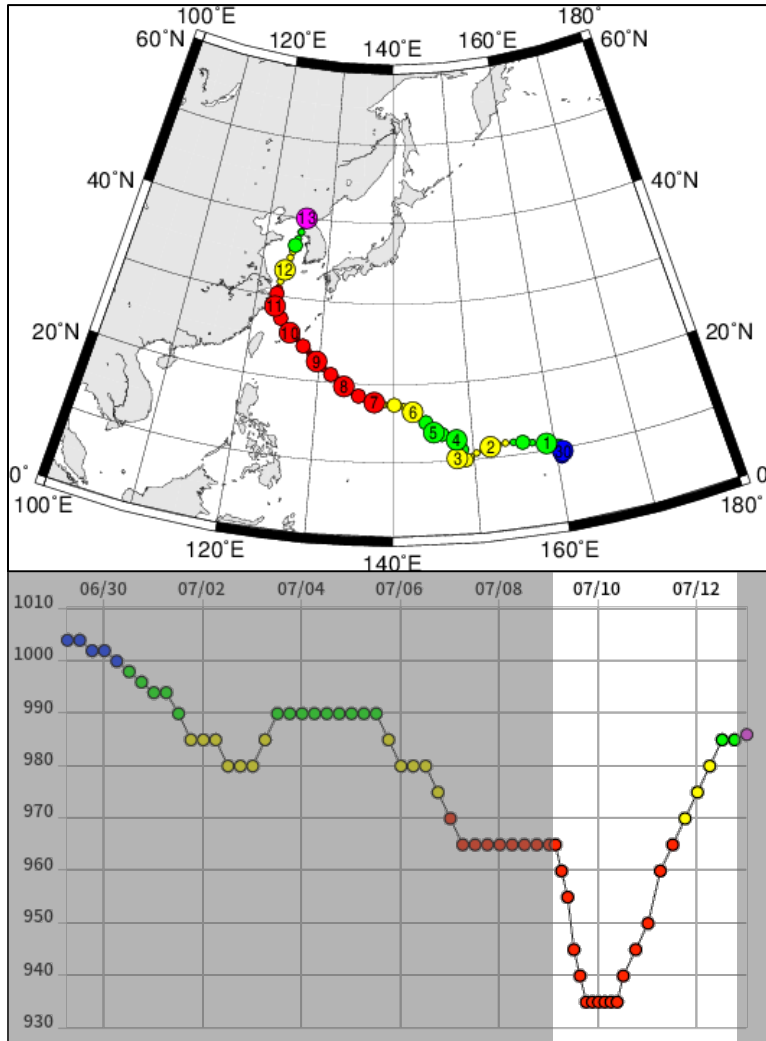
2. Aggregation: Accumulate hourly GSMaP data for 6 hours (7 slots)



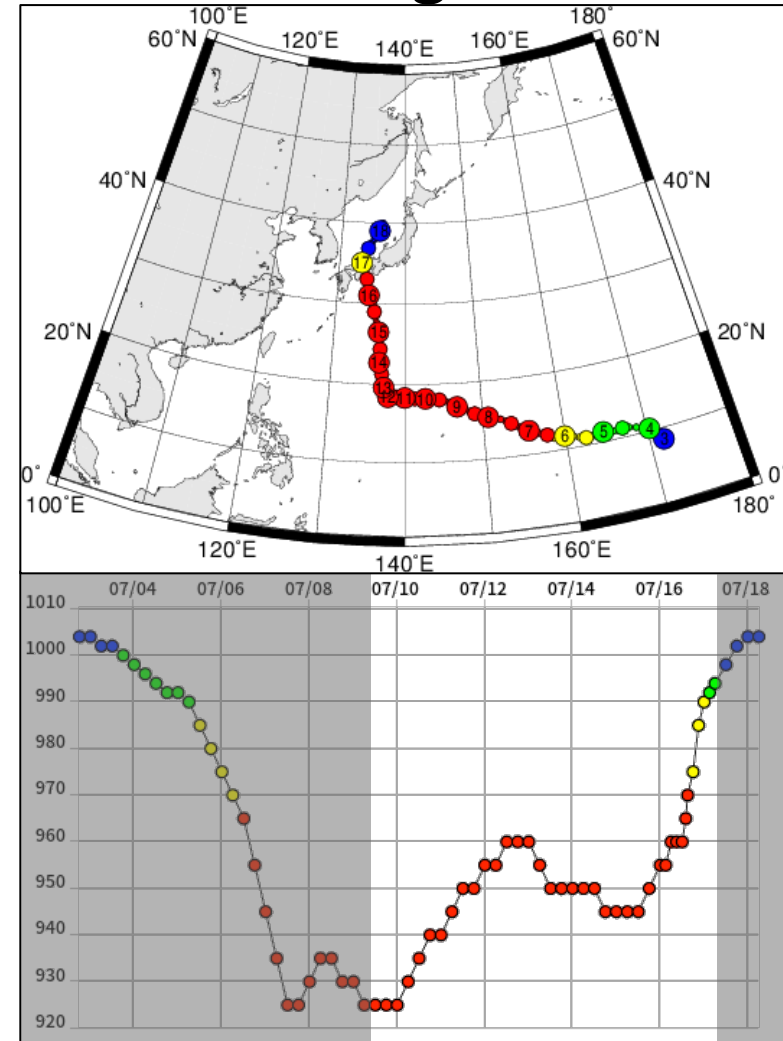
$$\text{Obs } y^o: pp_{6hr}^{GSMaP} = 0.5 \times pp_{-3} + \sum_{i=-2}^2 pp_i + 0.5 \times pp_3$$

Experiment Design: Overviews of Two TCs in 2015

Chan-hom



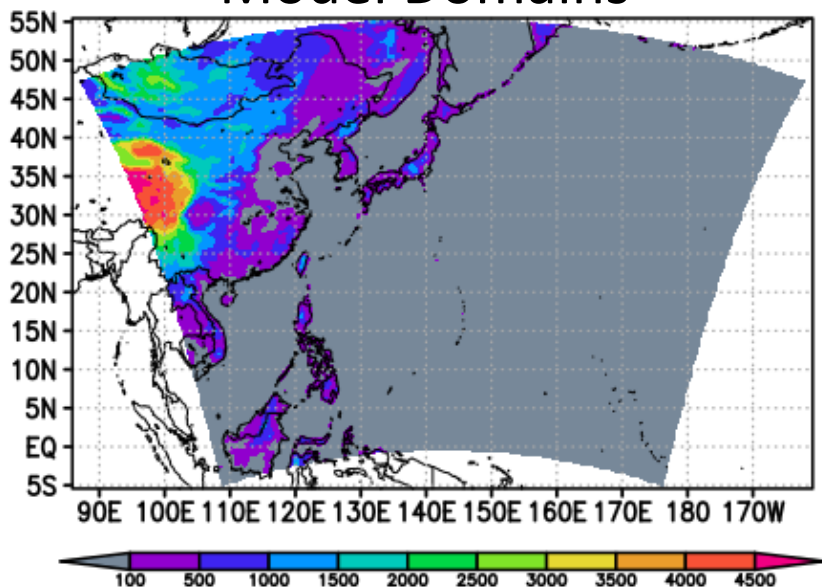
Nangka



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

Experiment Settings: NWP Model

Model Domains

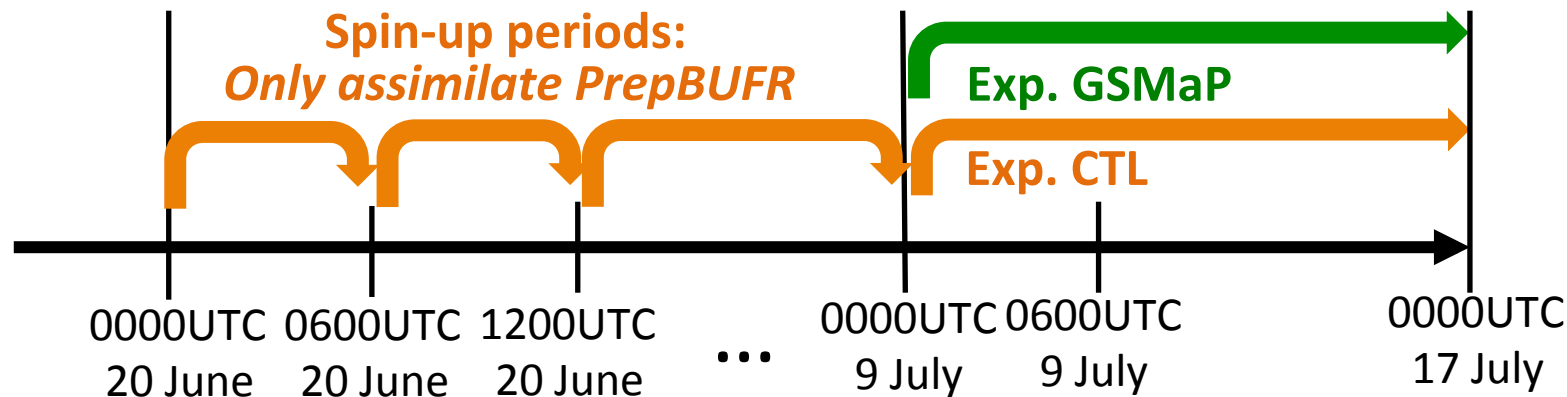


SCALE Settings	D01
Resolution	36km
Domain size	240×180 grids, 36 levels (~28km)
BC	GFS FNL every 6 hours
Microphysics	TOMITA08
Radiation	MSTRNX
Cumulus Para.	w/o and w/ KF

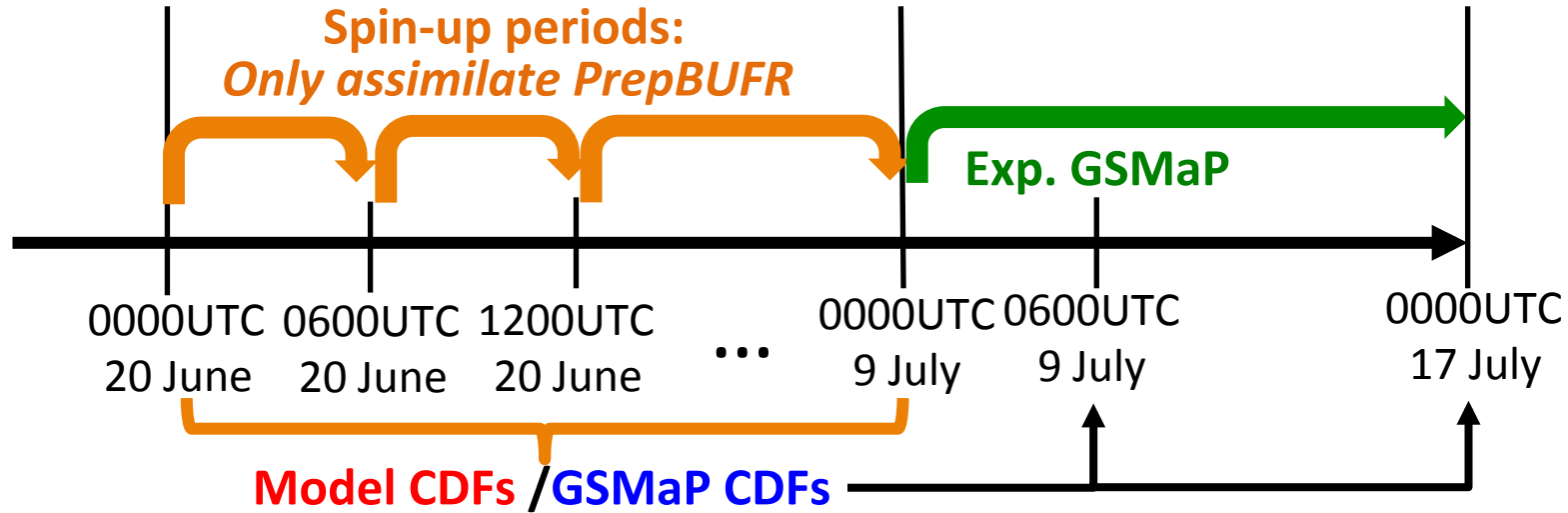
- Low resolution, in order to get some results within my internship period.
- We at first followed configurations of the SCALE NRT system (w/o KF). However, the precipitation fields w/o KF at 36km resolution is not similar to the GSMaP obs.

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		$\geq 0.001\text{mm (6hr)}^{-1}$
Precipitation CDFs		JUN 20 ~ JUL 6
Height of precipitation obs.		850mb
QC by min. number of precip. members		75/100 (75%)

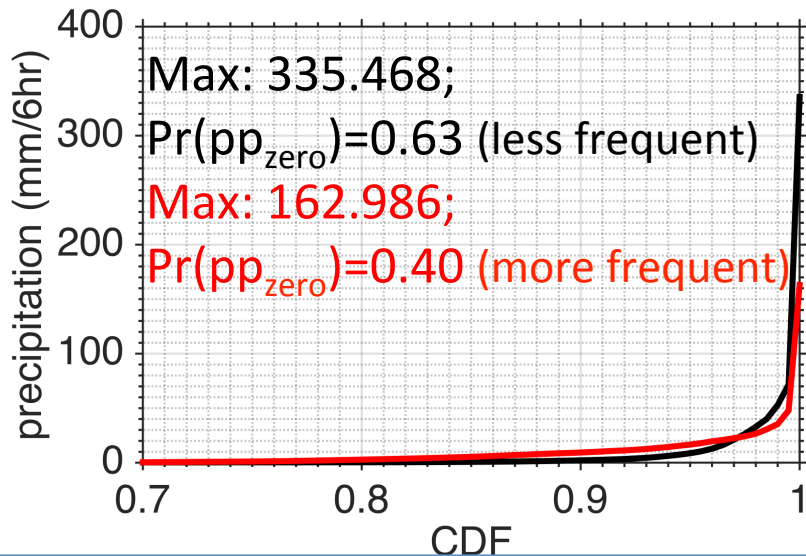


Construction of the Precipitation CDFs



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.

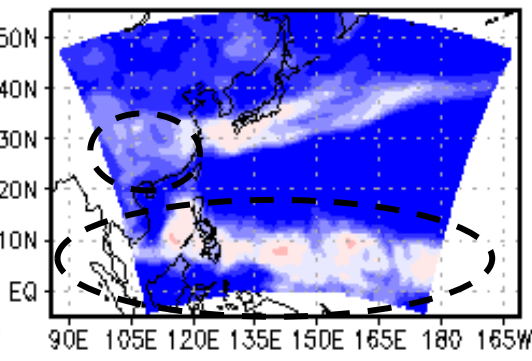
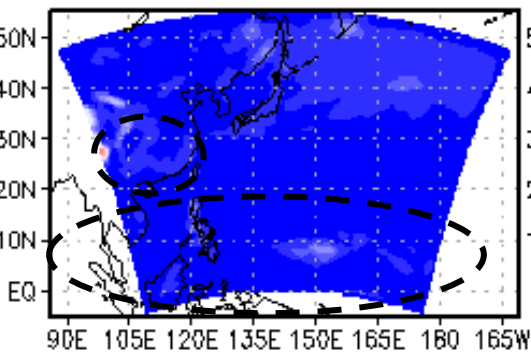
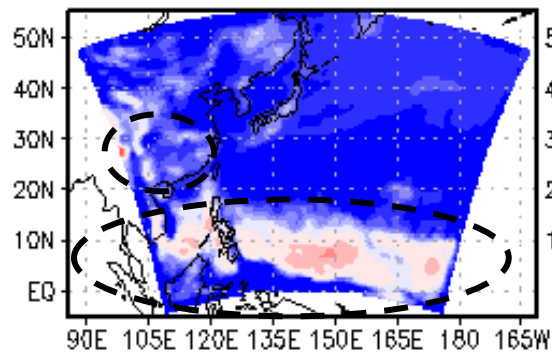
Comparison of the Precipitation CDFs

SCALE w/ KF

SCALE w/o KF

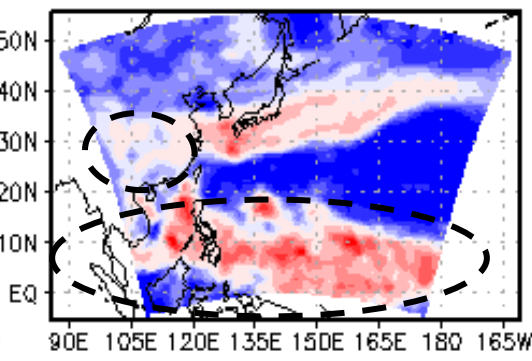
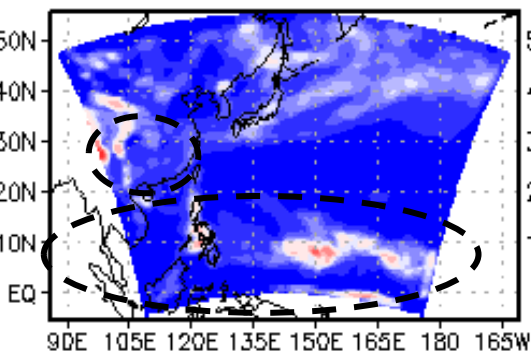
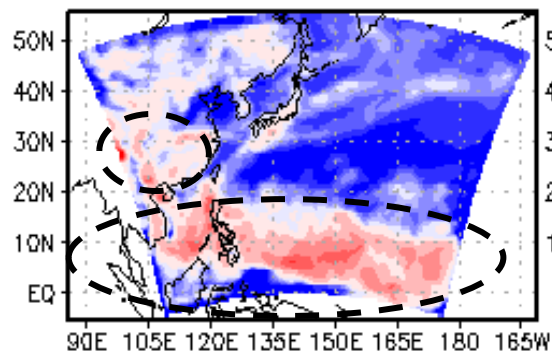
GSMaP

Precipitation
at CDF=0.75

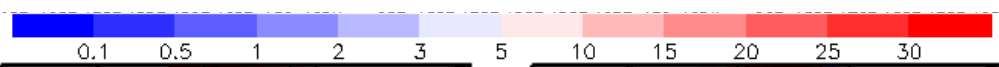


In addition, with KF, the maxima of model precip. also decreases (not shown here)

Precipitation
at CDF=0.9

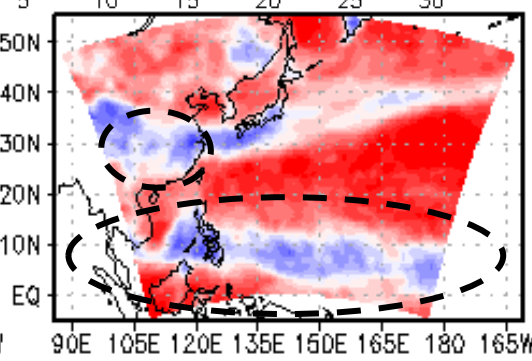
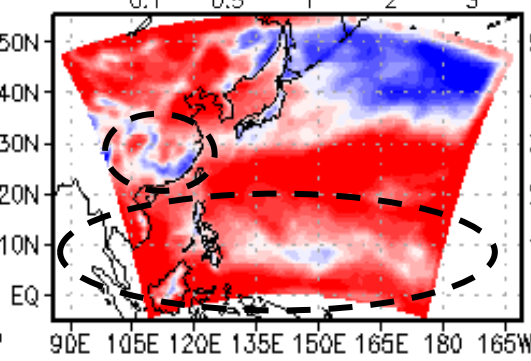
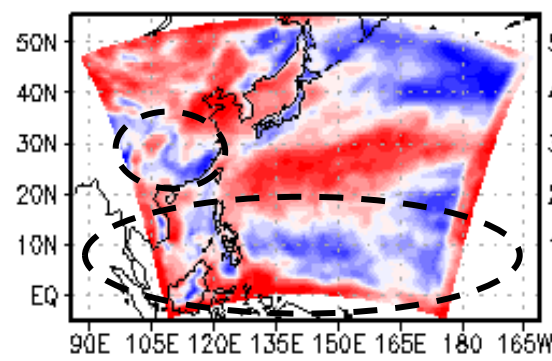


Smaller precip.

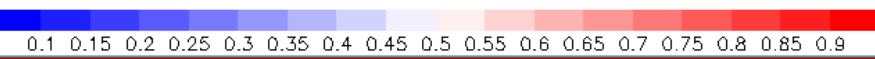


Larger precip.

Pr(ppzero)

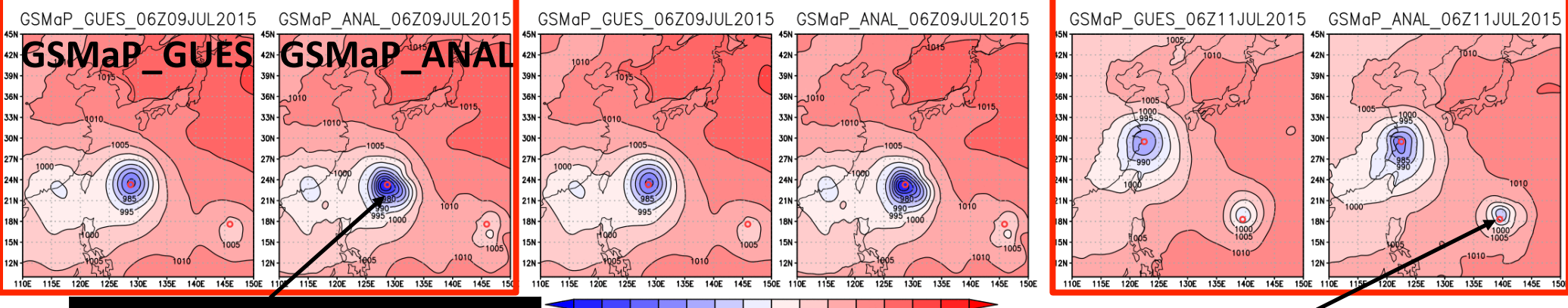
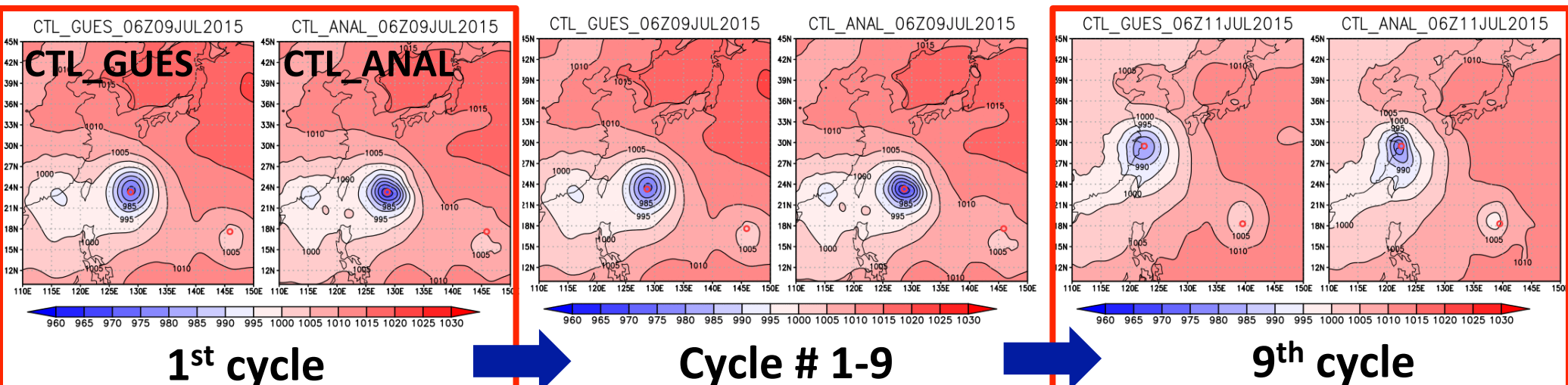


More frequent



less frequent

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

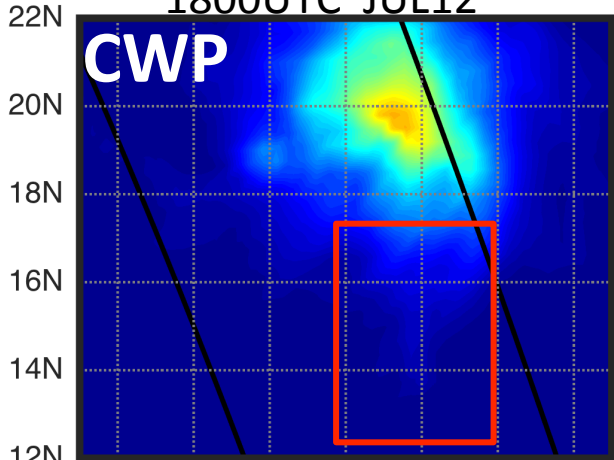
Source	Min. SLP (mb)
CTL_ANA	997.8
GSMaP_ANA	985.7
JMA Best Track	945

Adjustments of the Hydrometeors

CTL_ANA

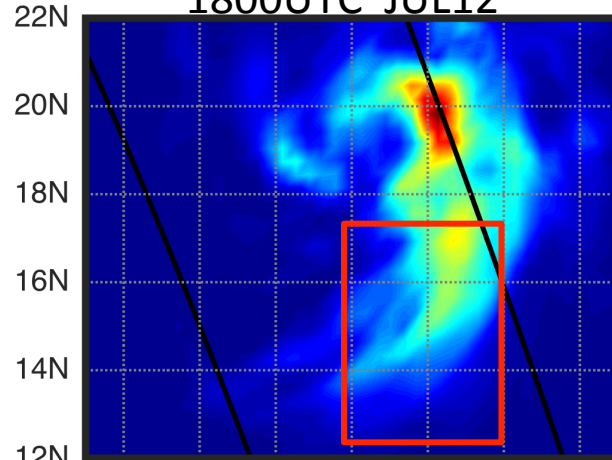
1800UTC JUL12

CWP



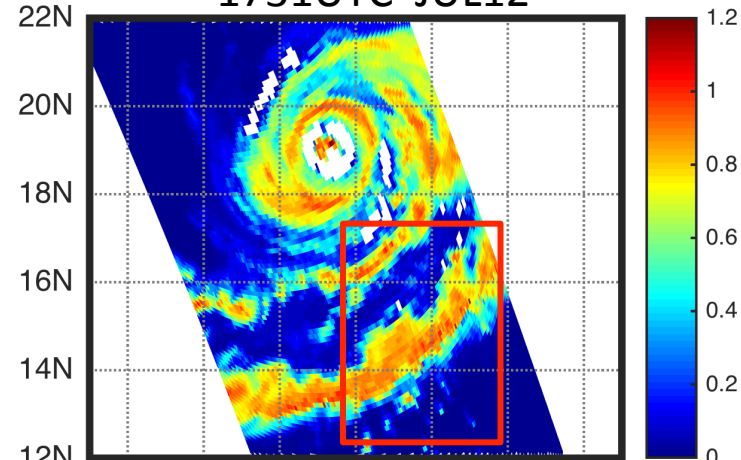
GSMaP_ANA

1800UTC JUL12



GMI_L2_Retrievals

1731UTC JUL12

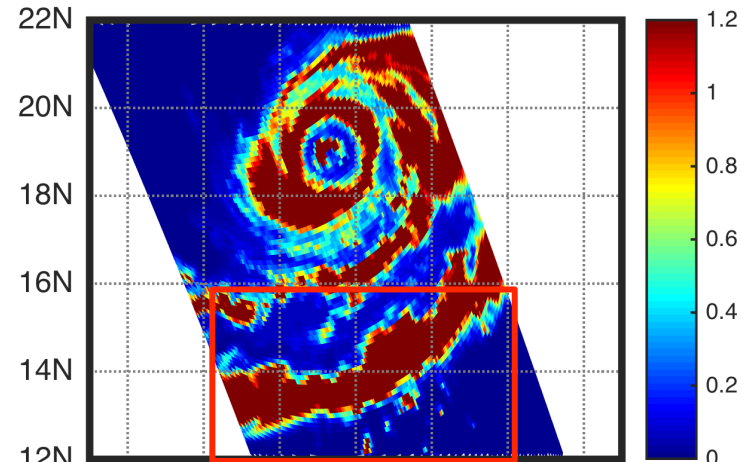
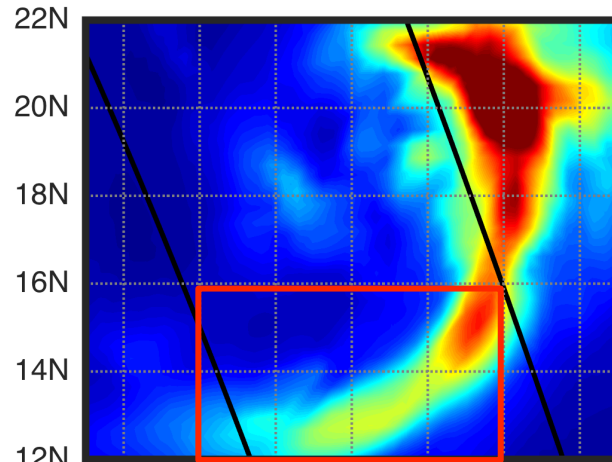
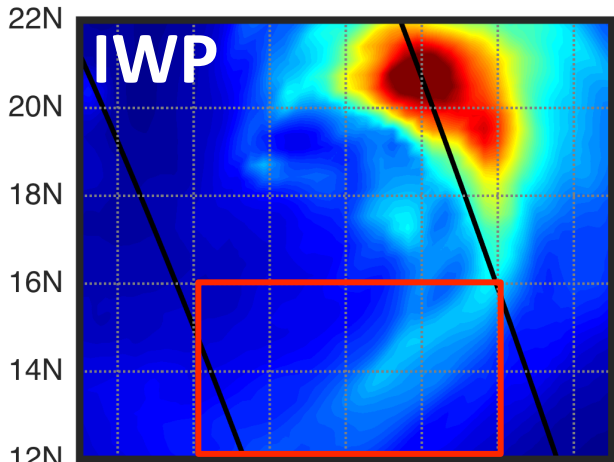


132E 134E 136E 138E 140E 142E 144E

132E 134E 136E 138E 140E 142E 144E

132E 134E 136E 138E 140E 142E 144E

IWP



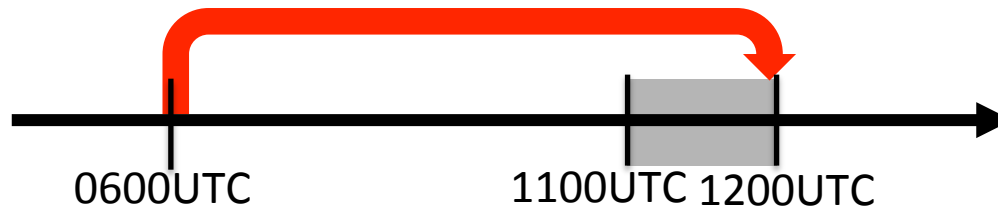
132E 134E 136E 138E 140E 142E 144E

132E 134E 136E 138E 140E 142E 144E

132E 134E 136E 138E 140E 142E 144E

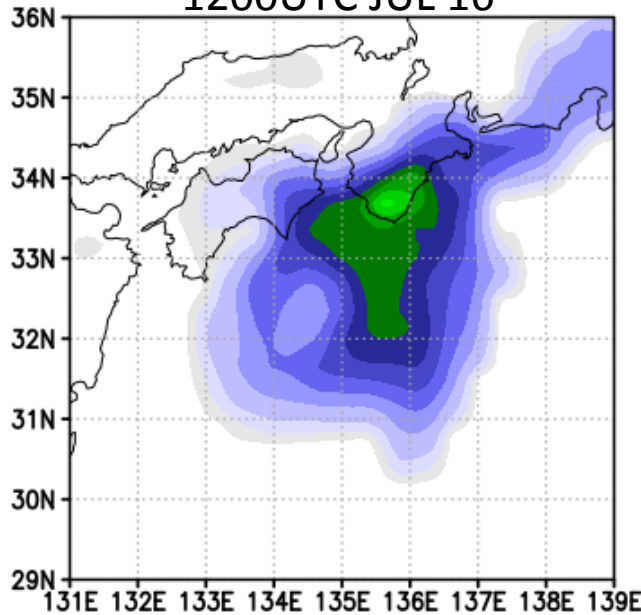
Accumulated 1-hr Precipitation Forecast

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



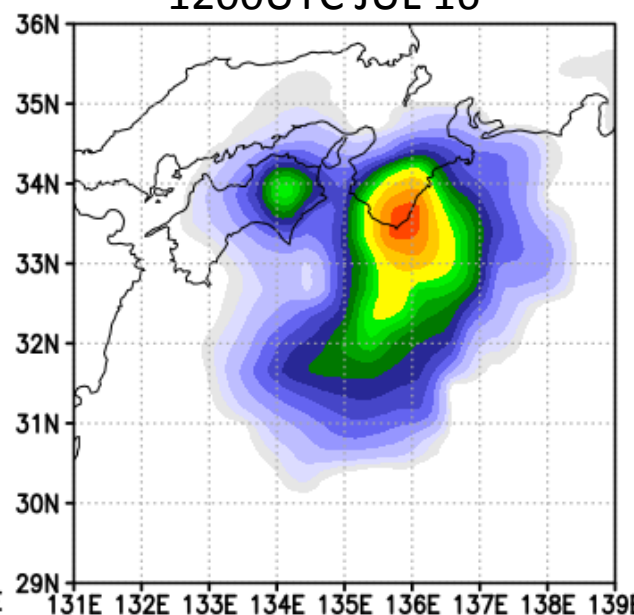
CTL_6HR_FCST

1200UTC JUL 16



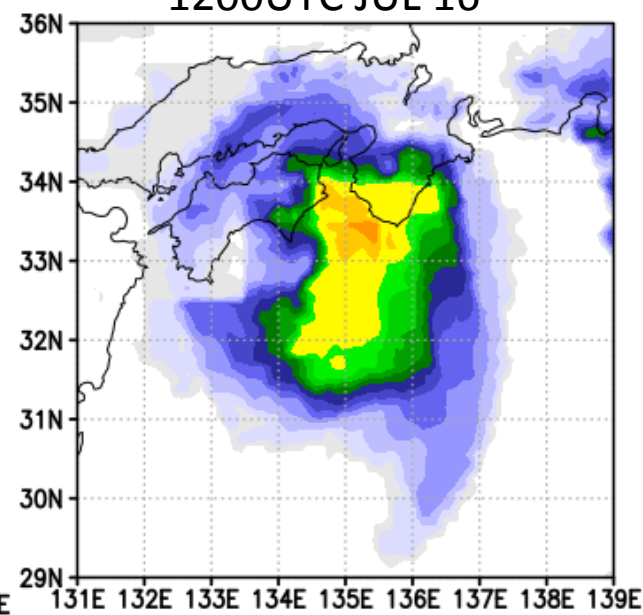
GSMaP_6HR_FCST

1200UTC JUL 16

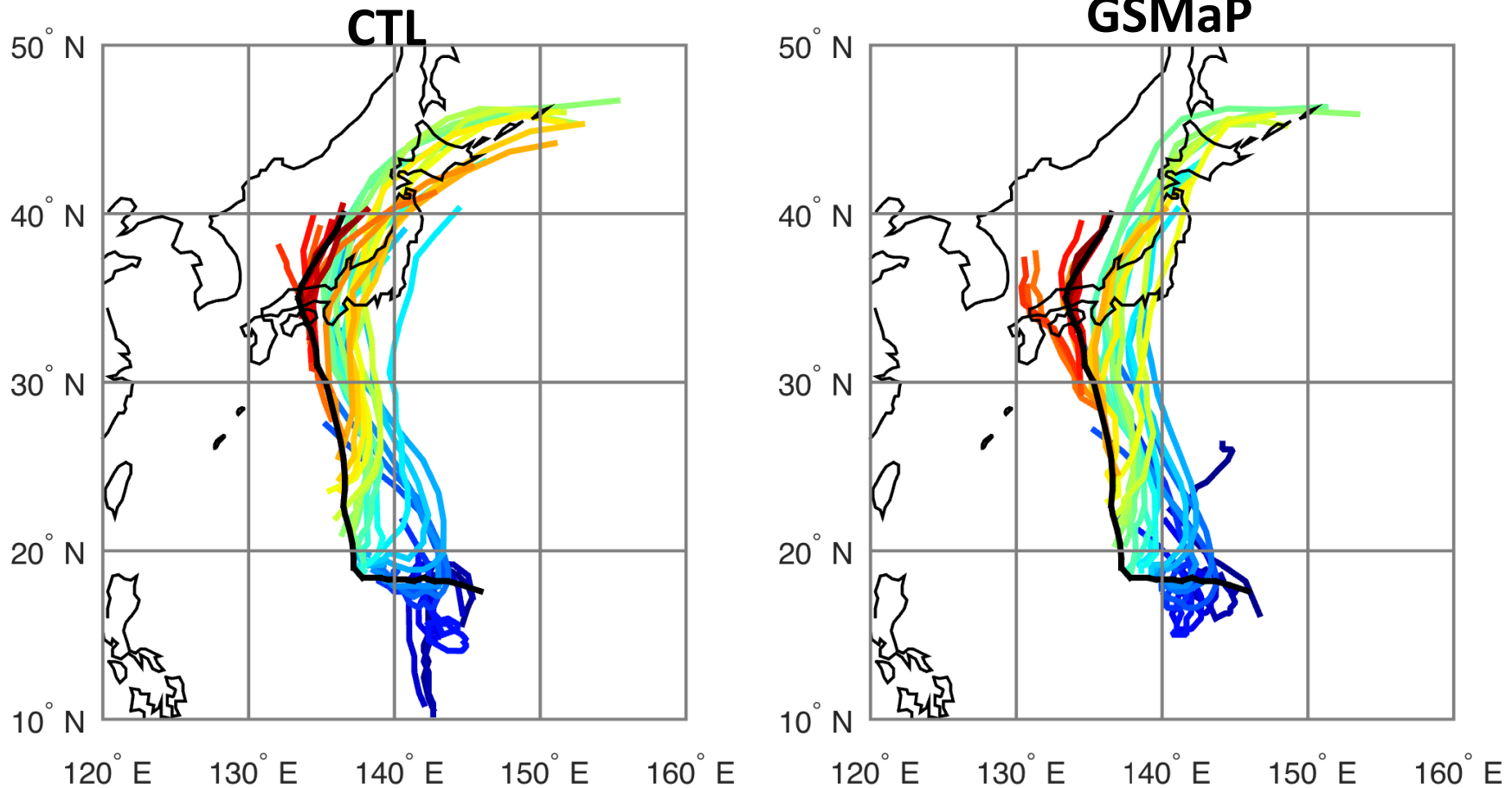


GSMaP_OBS

1200UTC JUL 16



TC Track Forecast Error



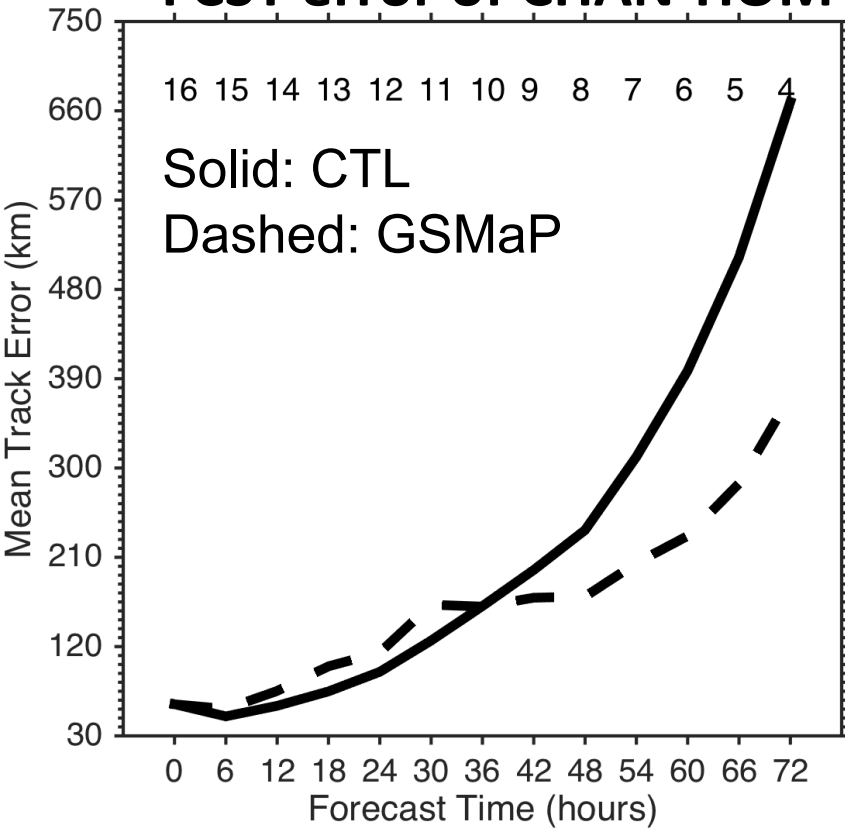
Thick black line: Best track

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

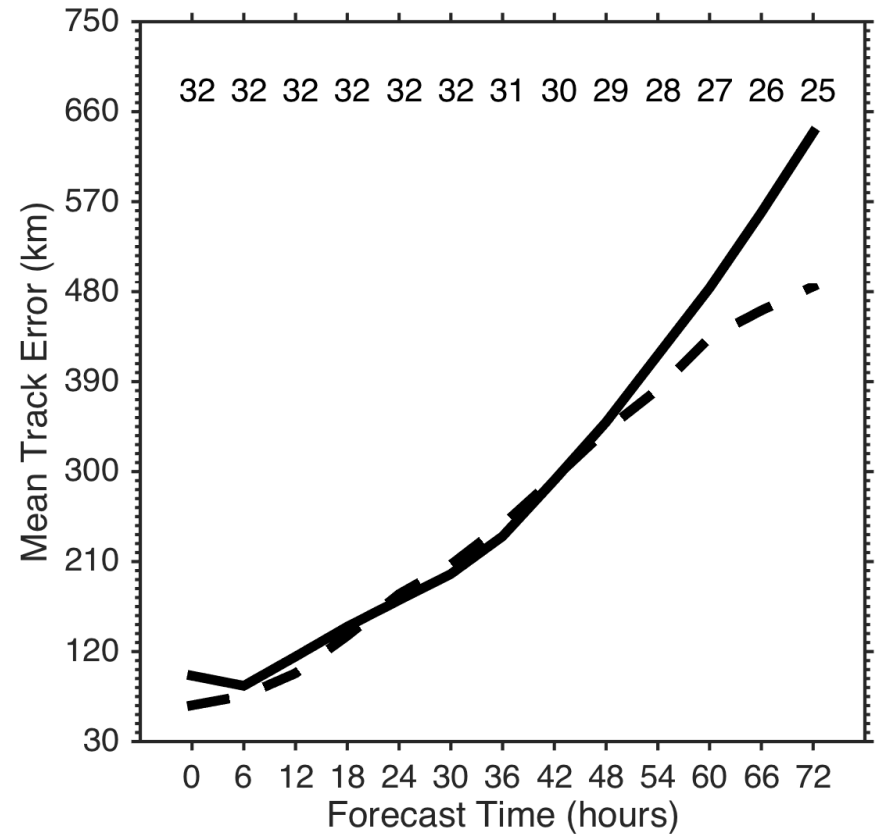
For the first several cycles, Nangka moves southward in the Exp. CTL, which does not occur in Exp. GSMaP.

Averaged TC Track Forecast Error

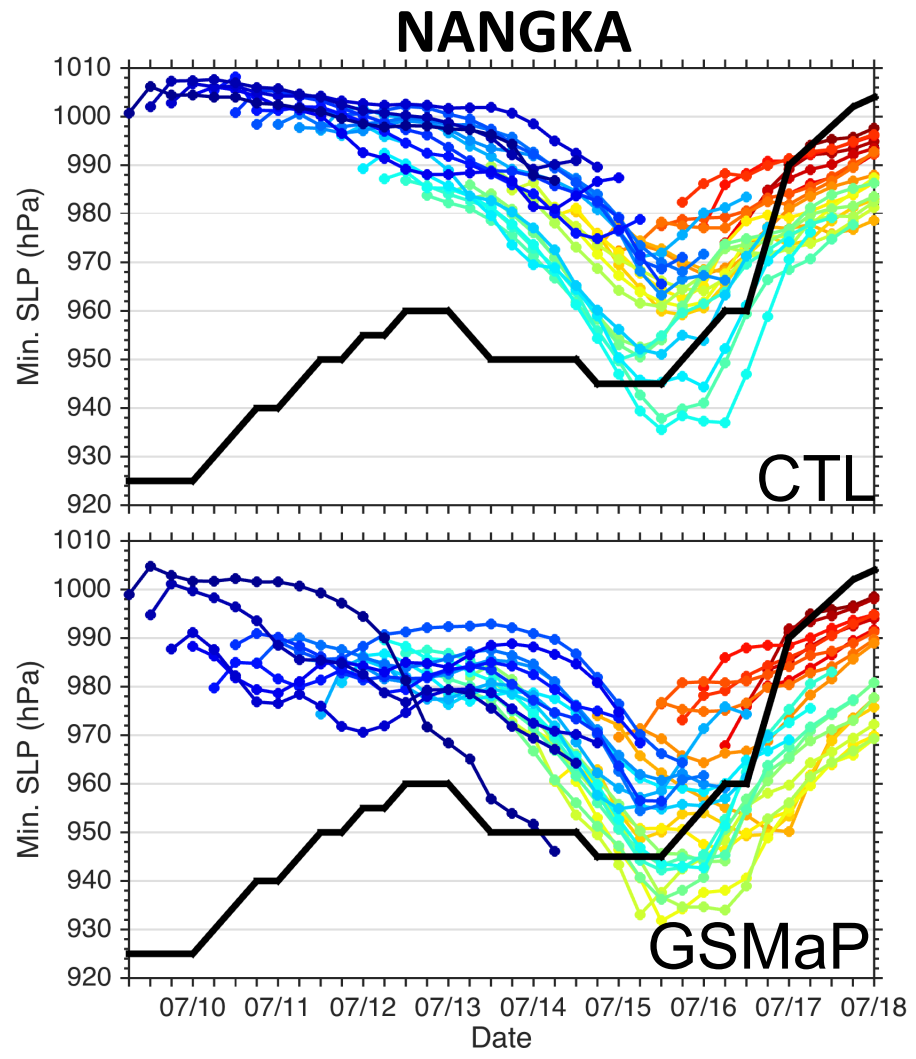
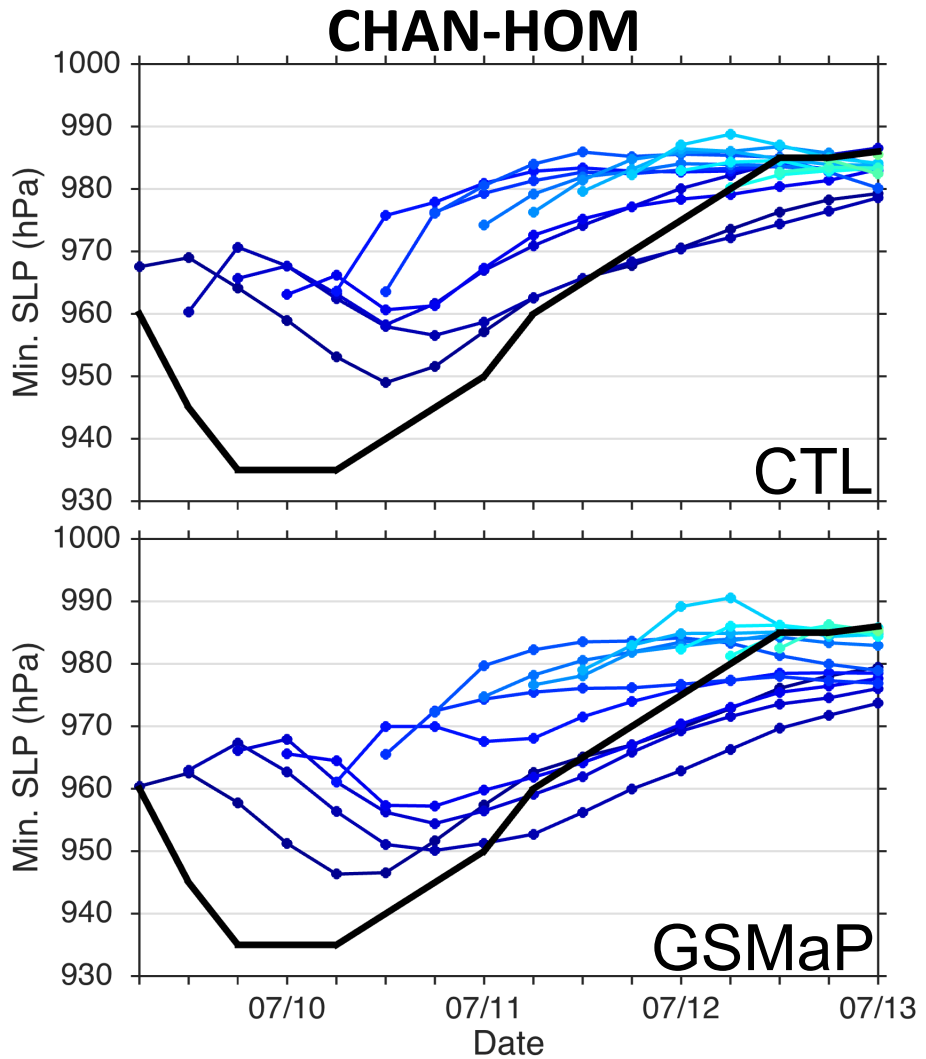
FCST error of CHAN-HOM



FCST error of NANGKA



120-hr Intensity Forecast Initialized at Different Time

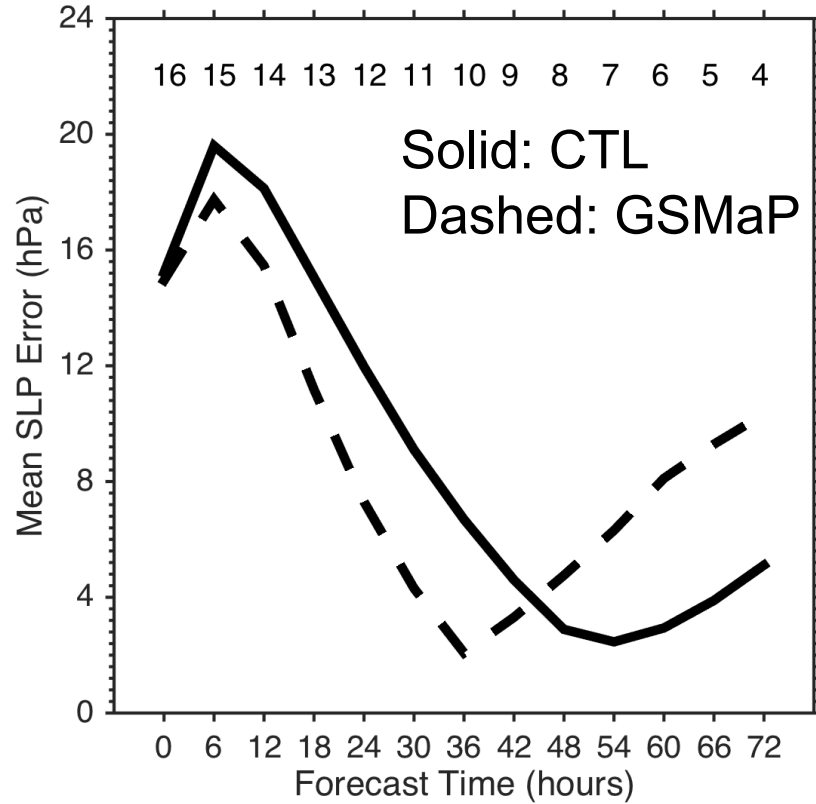


Thick black line: Best track

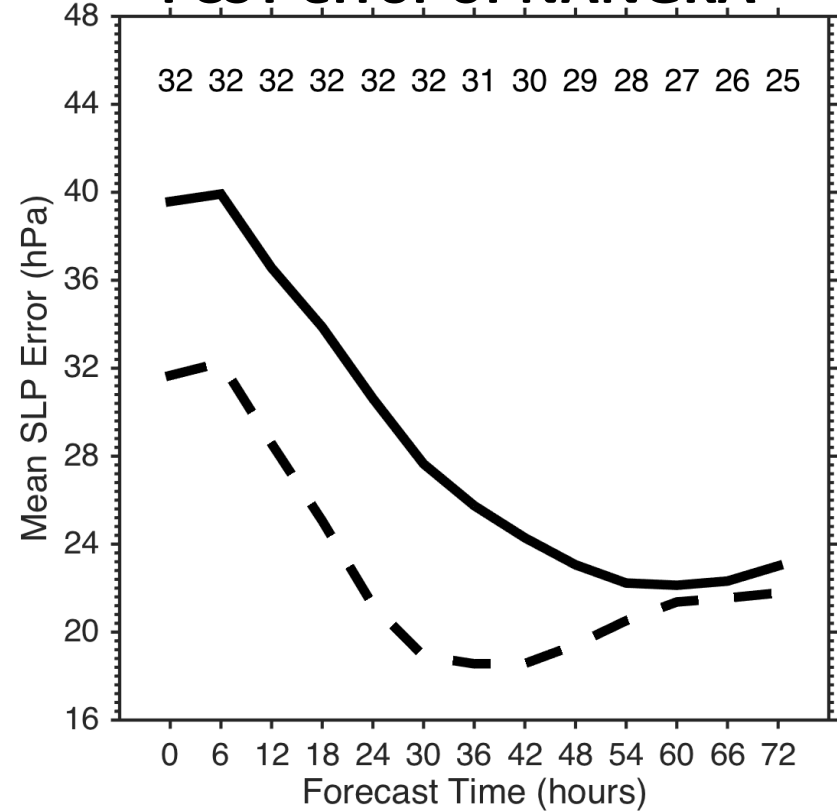
Color lines: 120-hr forecasts initialized at different time

Averaged TC Intensity Forecast Error

FCST error of CHAN-HOM



FCST error of NANGKA



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.

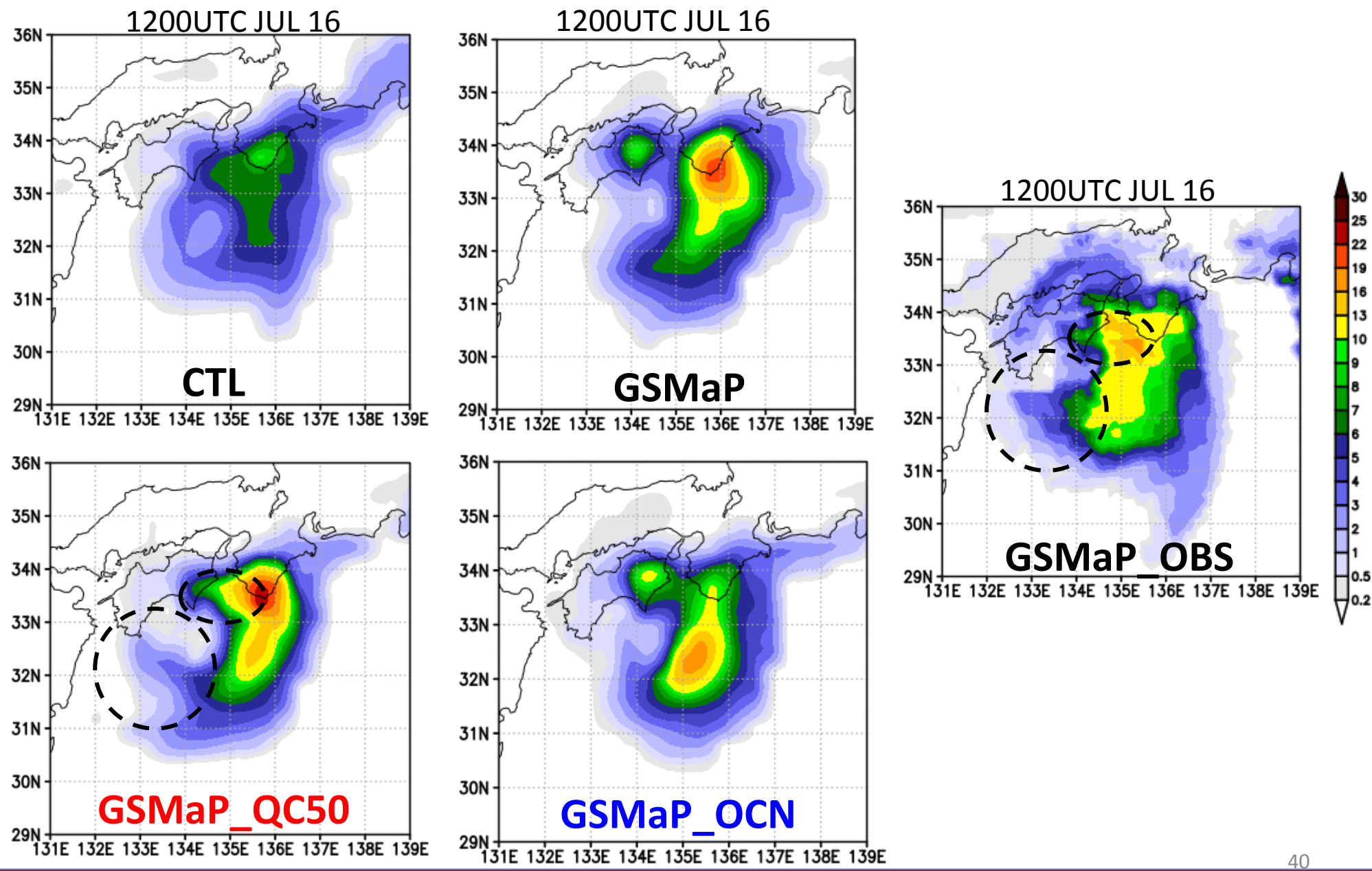
=> Will loosing this QC to 50 further Improve the forecast?

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.

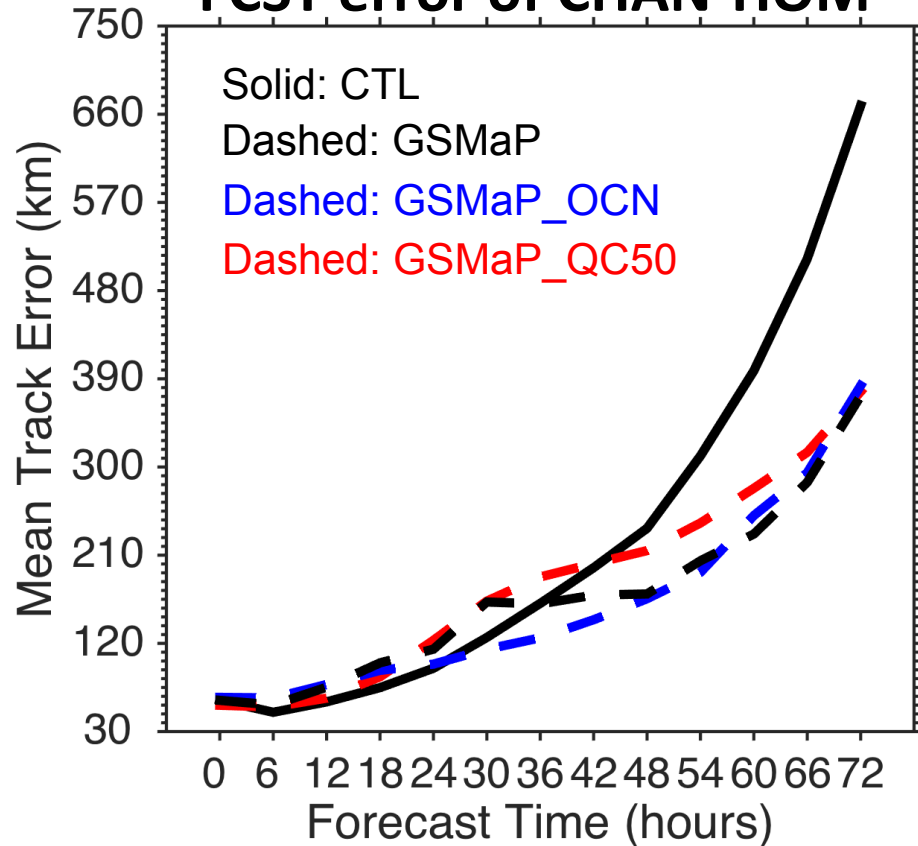
=> What if we only assimilate obs over ocean?

Accumulated 1-hr Precipitation Forecast

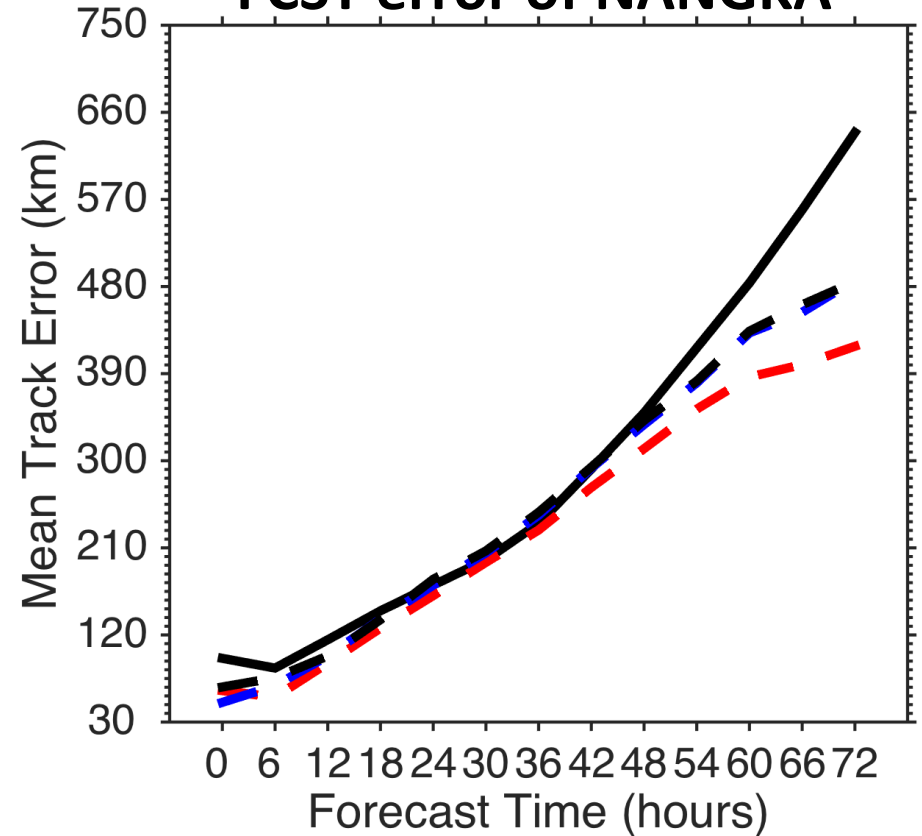


Sensitivity to the QC schemes

FCST error of CHAN-HOM

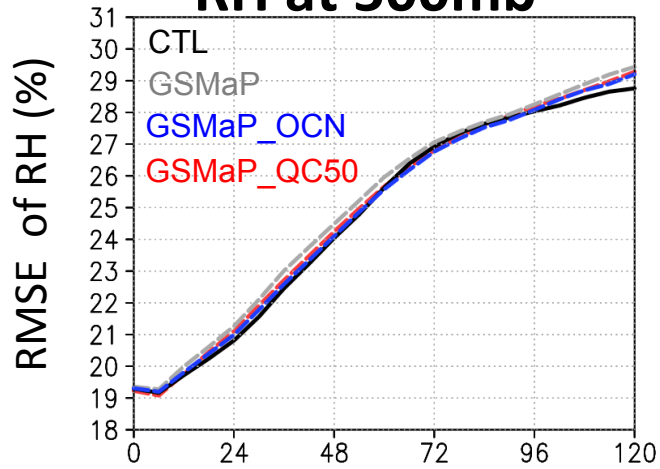


FCST error of NANGKA

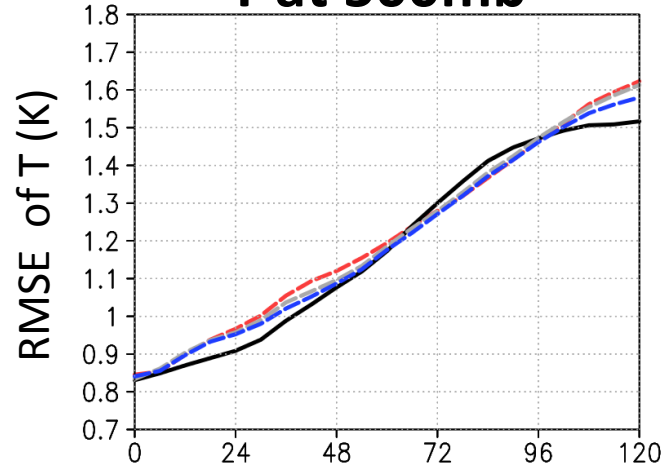


5-day Forecast RMSE of the Entire Domain

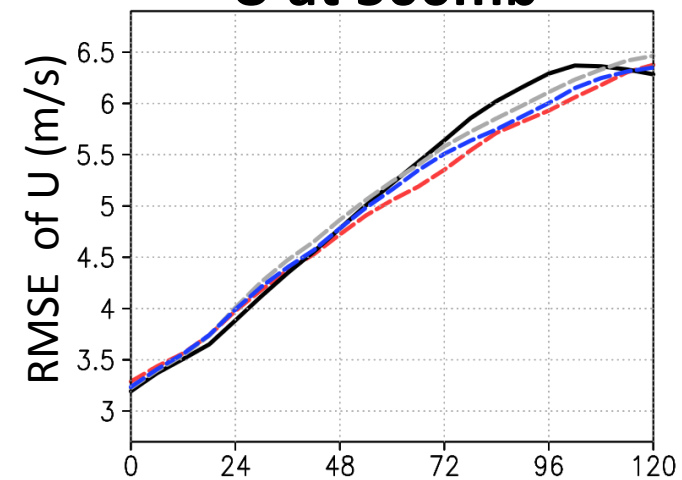
RH at 500mb



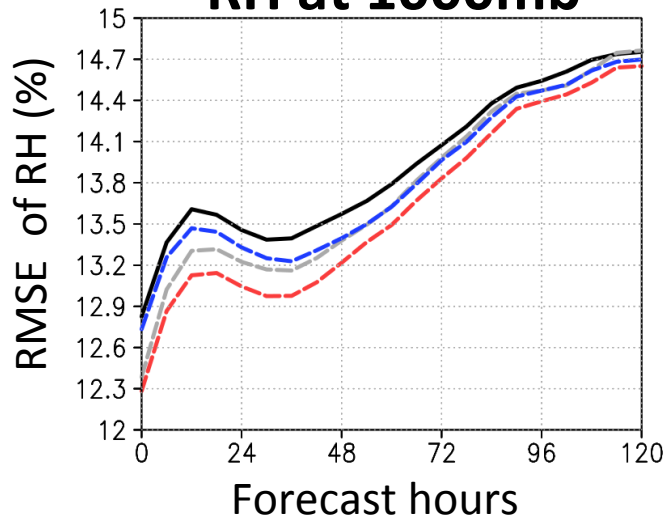
T at 500mb



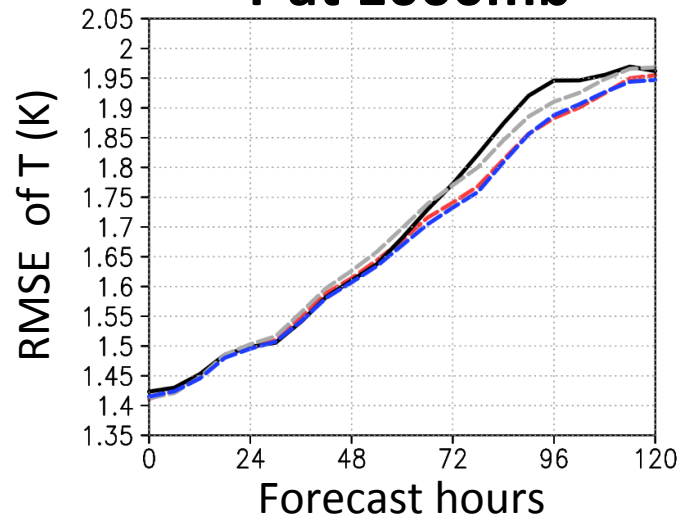
U at 500mb



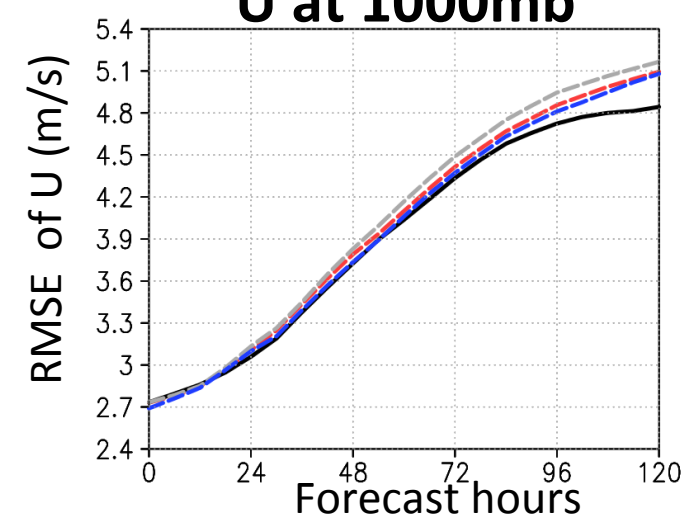
RH at 1000mb



T at 1000mb



U at 1000mb



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.