"Big Data Assimilation"

Progress and Plans

Takemasa Miyoshi*

RIKEN Advanced Institute for Computational Science

*PI and presenting, Takemasa.Miyoshi@riken.jp

K. Kondo, K. Terasaki, T. Honda, J. Ruiz, G.-Y. Lien, M. Kunii, Y. Maejima, S. Otsuka, S. Satoh, T. Ushio, H. Tomita, Y. Ishikawa, K. Bessho, H. Seko

With many thanks to

JMA

UMD Weather-Chaos group

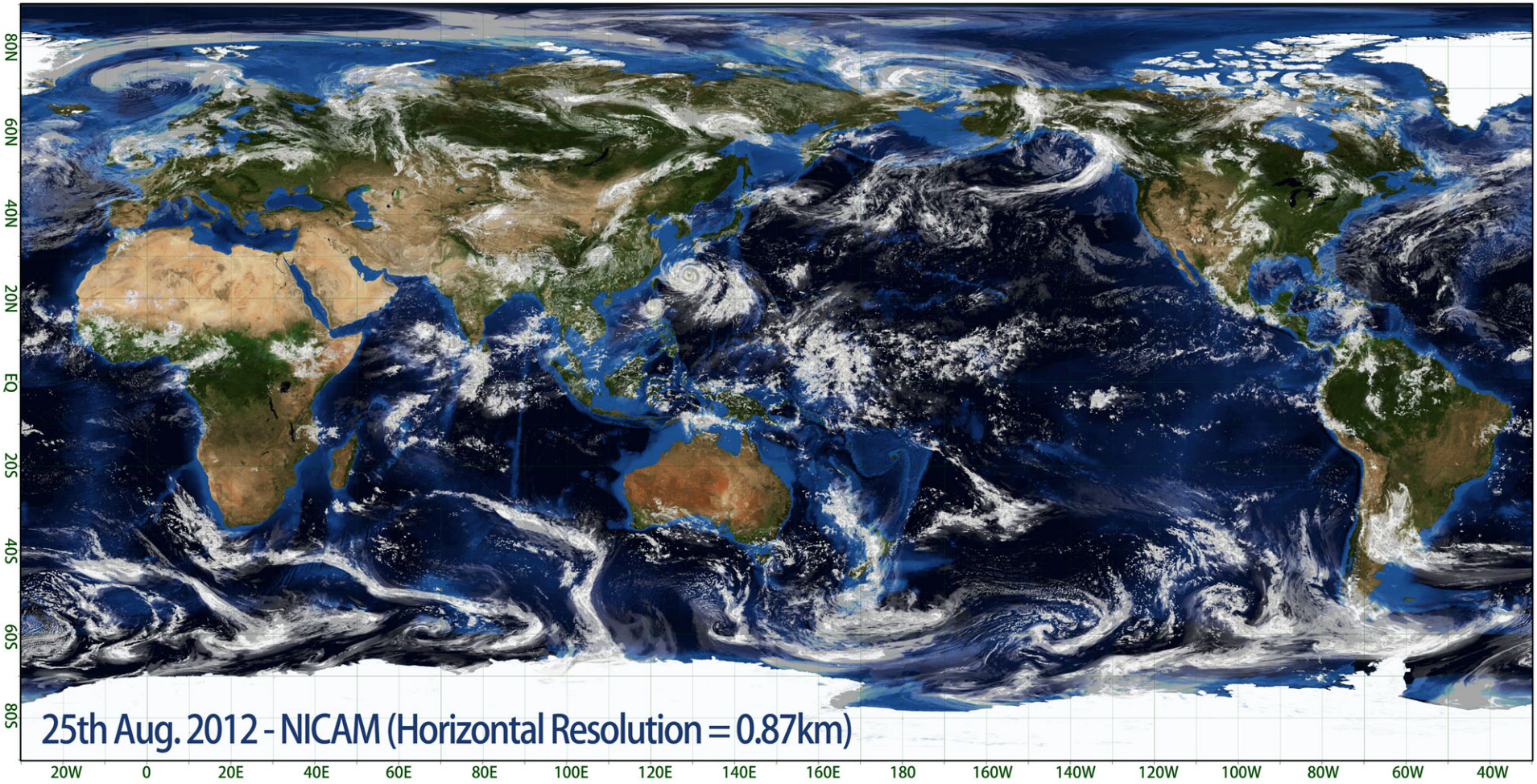
JST CREST "Big Data Assimilation" project

JAXA PMM "Ensemble Data Assimilation" project

Japan's FLAGSHIP 2020 project

RIKEN Data Assimilation Research Team

Global 870-m simulation (*Miyamoto et al. 2013*)



Visualized by Ryuji Yoshida

Computers keep advancing…

- With the "post-K" supercomputer (~2020), we can afford 100 members of the global 870-m simulation.

With the Post-K, we aim to run 1000-member global NICAM-LETKF at 3.5-km resolution

The Japanese 10-Peta-Flops K computer

Press Release

Previous Index Next

July 23, 2014

K computer runs largest ever ensemble simulation of global weather

Ensemble forecasting is a key part of weather forecasting today. Computers typically run multiple simulations, called ensembles, using slightly different initial conditions or assumptions, and then analyze them together to try to improve forecasts. Now, in research published in *Geophysical Research Letters*, using Japan's flagship 10-petaFLOPS K computer, researchers from the RIKEN Advanced Institute for Computational Science (AICS) have succeeded in running 10,240 parallel simulations of global weather, the largest number ever performed, using data assimilation to reduce the range of uncertainties.

The assimilation of the 10,240 ensemble data sets was made possible by a cross-disciplinary collaboration of data assimilation experts and eigenvalue solver scientists at RIKEN AICS. The "Local Ensemble Transform Kalman Filter" (LETKF), an already efficient system, was further improved by a factor of eight using the "EigenExa" high-performance eigenvalue solver software, making possible a three-week computation of data from the 10,240 ensembles for simulated global weather. By analyzing the 10,240 equally probable estimates of atmospheric states, the team discovered that faraway observations, even going beyond 10,000 kilometers in distance, may have an immediate impact on eventual state of the estimation. This finding suggests the need for further research on advanced methods that can make better use of faraway observations, as this could potentially lead to an improvement of weather forecasts.

News & Media
Press Center
Press Releases
2014
2013
2012
2011
2010
2009
2008
2007
2006
2005
News
Events & Symposiums
Publications
Videos

Press Release

Previous Index Next

July 23, 2014

K computer runs largest ever ensemble simulation of global weather

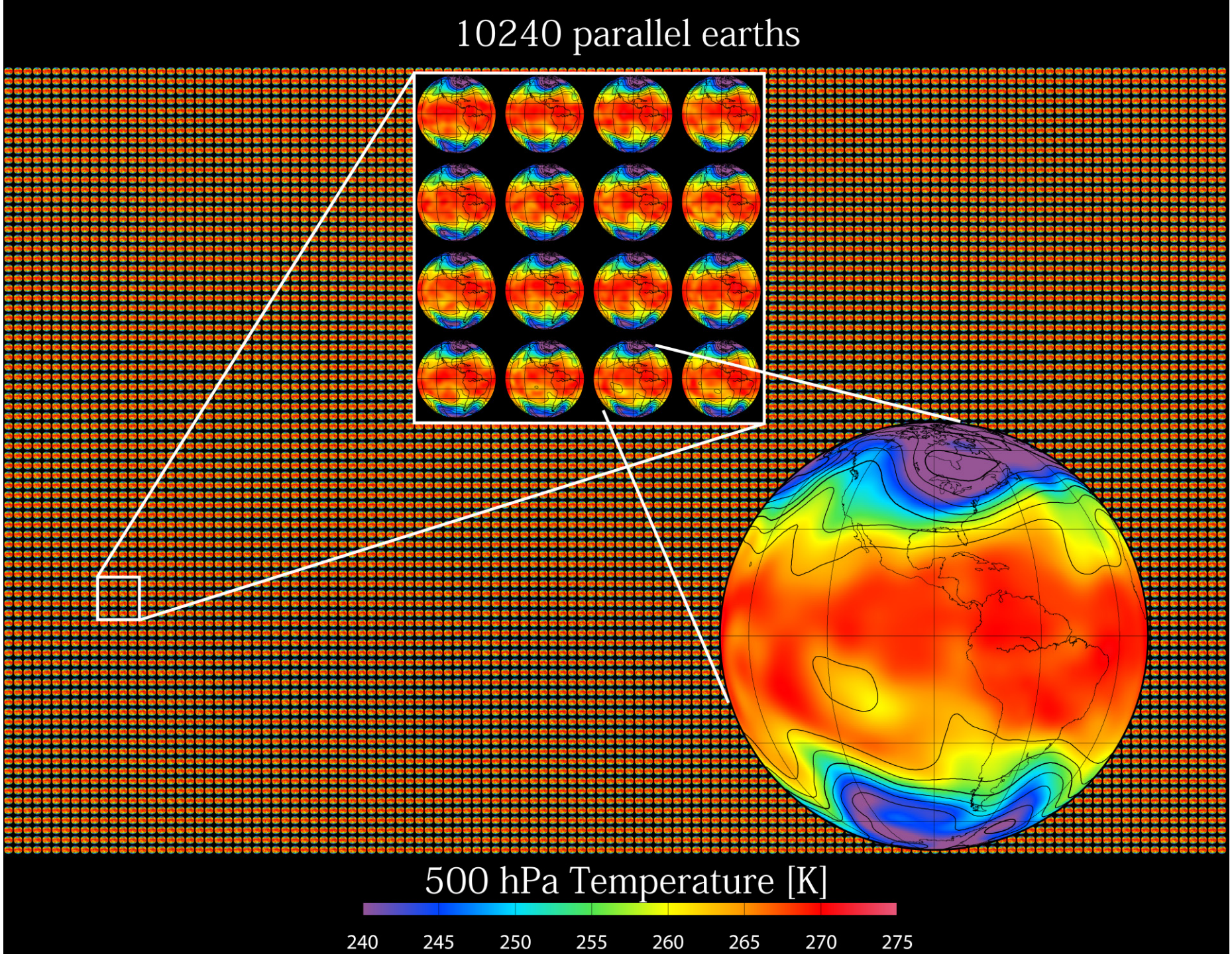
Ensemble forecasti
called ensembles, (
try to improve fore
10-petaFLOPS K co
have succeeded in
using data assimila(

A simulated study using the T30/L7 SPEEDY AGCM

(Miyoshi, Kondo, Imamura 2014)

The assimilation of the 10,240 ensemble data sets was made possible by a cross-disciplinary collaboration of data assimilation experts and eigenvalue solver scientists at RIKEN AICS. The "Local Ensemble Transform Kalman Filter" (LETKF), an already efficient system, was further improved by a factor of eight using the "EigenExa" high-performance eigenvalue solver software, making possible a three-week computation of data from the 10,240 ensembles for simulated global weather. By analyzing the 10,240 equally probable estimates of atmospheric states, the team discovered that faraway observations, even going beyond 10,000 kilometers in distance, may have an immediate impact on eventual state of the estimation. This finding suggests the need for further research on advanced methods that can make better use of faraway observations, as this could potentially lead to an improvement of weather forecasts.

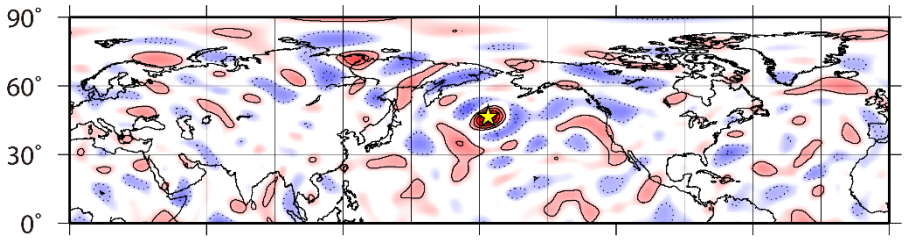
News & Media
Press Center
Press Releases
2014
2007
2006
2005
News
Events & Symposiums
Publications
Videos

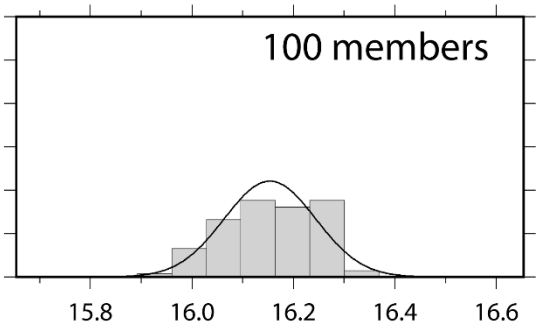

10240 parallel earths
500 hPa Temperature [K]
240
245
250
255
260
265
270
275

Advantage of large ensemble

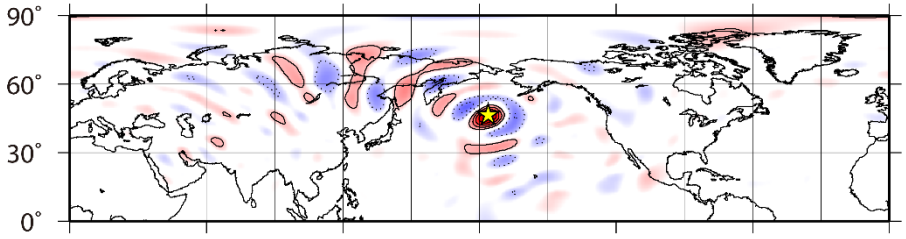
(Miyoshi, Kondo, Imamura 2014)

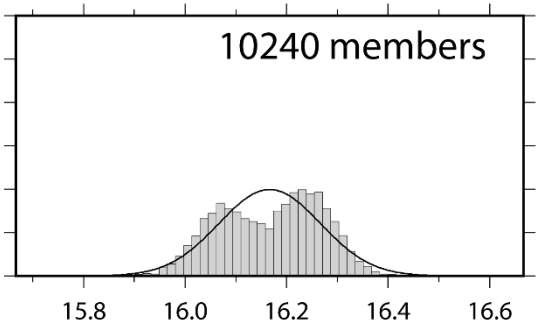
100 members





10240 members





Sampling noise reduced

High-precision probabilistic representation

Press Release

Previous Index Next

November 11, 2015

Largest ensemble simulation of global weather using real-world data

When performing numerical weather predictions, it is important that the simulation itself be accurate, but it is also key for real-world data, based on observations, to be accurately entered into the model. Typically, weather simulations work by having the computer conduct a number of simulations based on the current state, and then entering observational data into the simulation to nudge it in a way that puts it closer to the actual state. The problem of incorporating data in the simulation—data assimilation—has become increasingly complex with the large number of types of available data, such as satellite observations and measurements taken from ground stations. Typically, supercomputers today spend an approximately equal amount of time running the simulations and incorporating the real-world data.

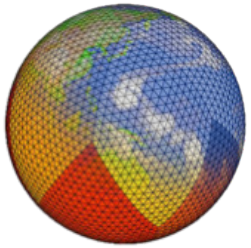
News & Media
Press Center
Press Releases
2015
2014
2013
2012
2011
2010
2009
2008

A real-world study using the NICAM

(Miyoshi, Kondo, Terasaki 2015)

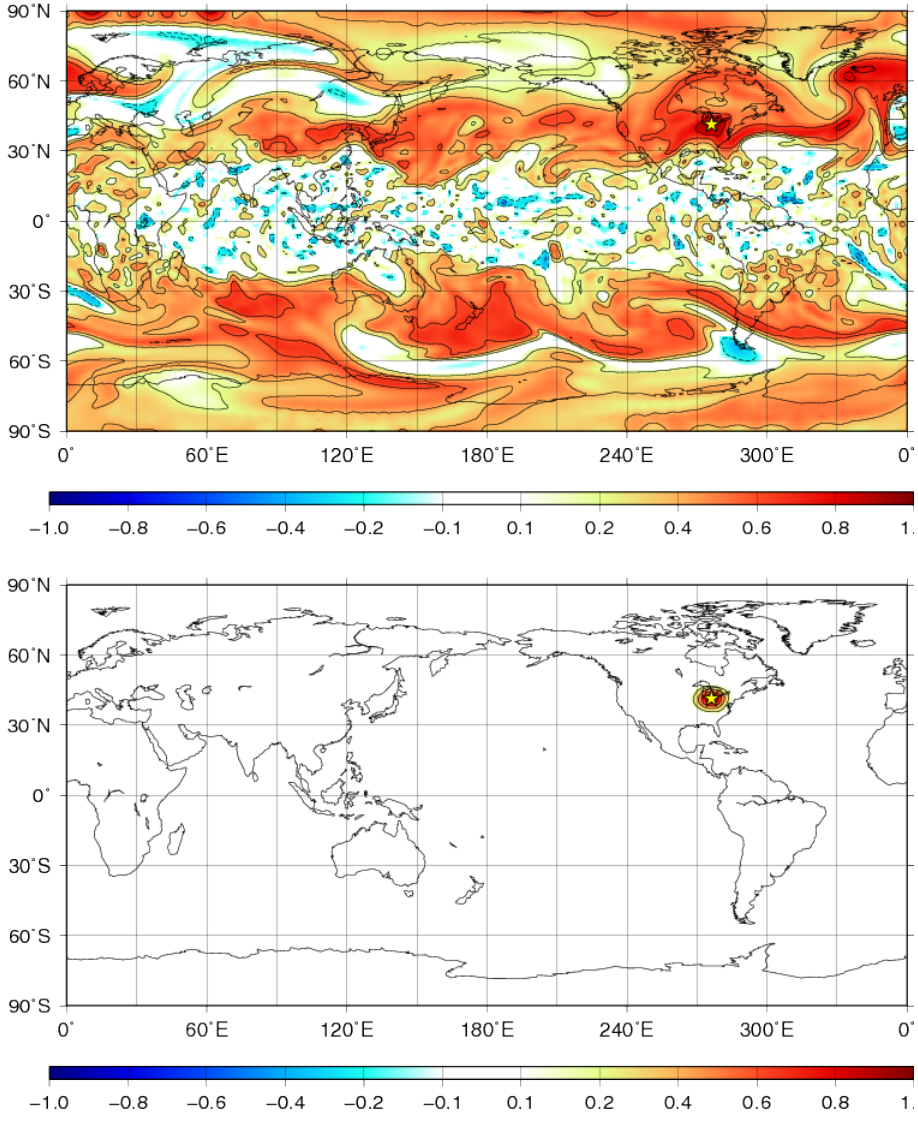
NICAM-LETKF

(Terasaki et al. 2015)



Correlation patterns (Q at ~100 hPa)

40 members *Kondo, Miyoshi (2015)*



Localized (σ=400km)

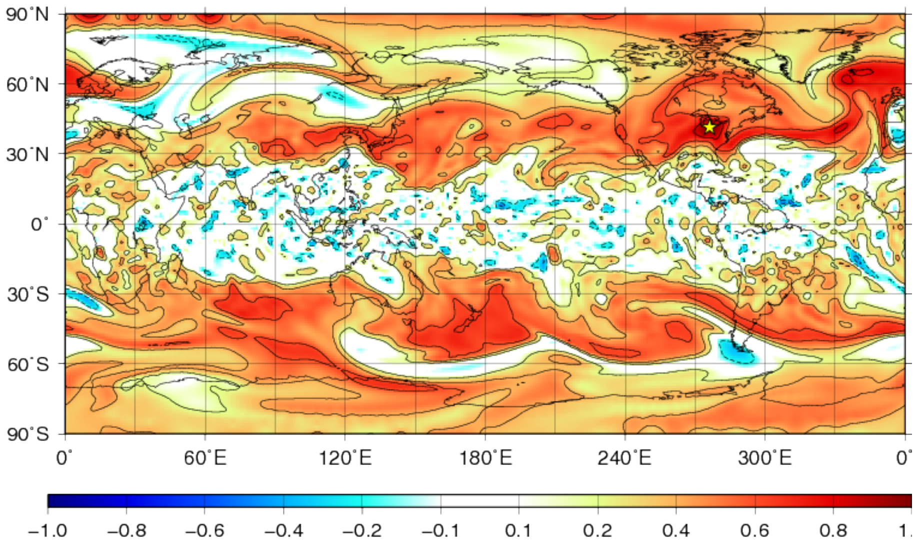
This is what we use for EnKF with 40 members.

11/8 00UTC after a week cycling

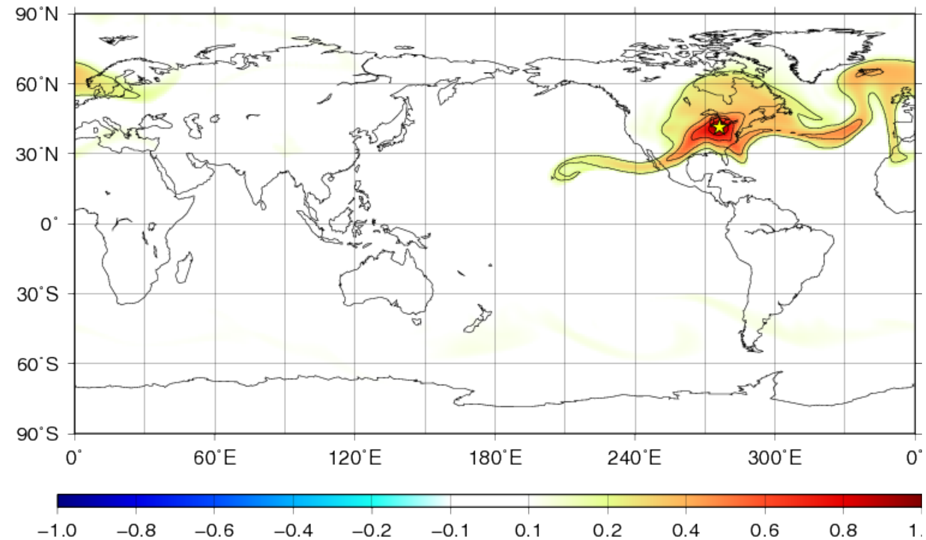
Correlation patterns (Q at ~100 hPa)

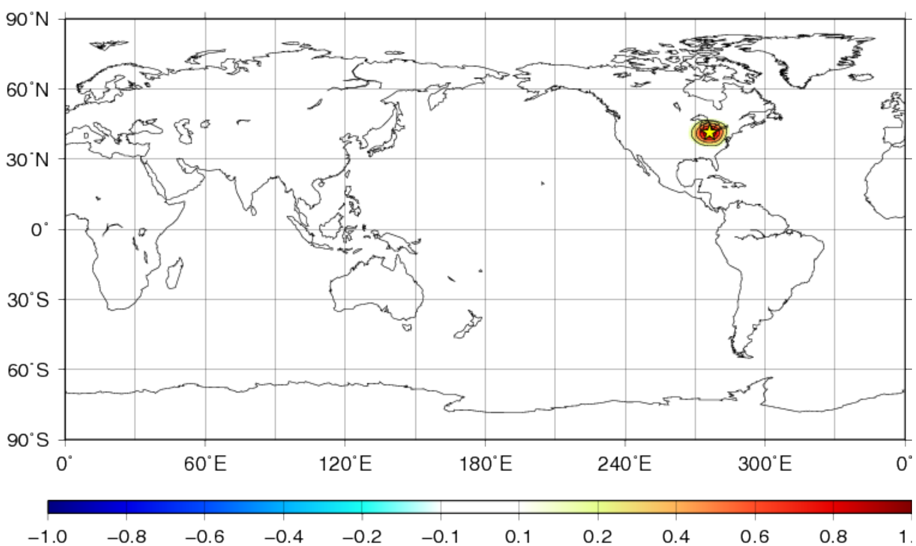
Kondo, Miyoshi (2015)

40 members



10240 members

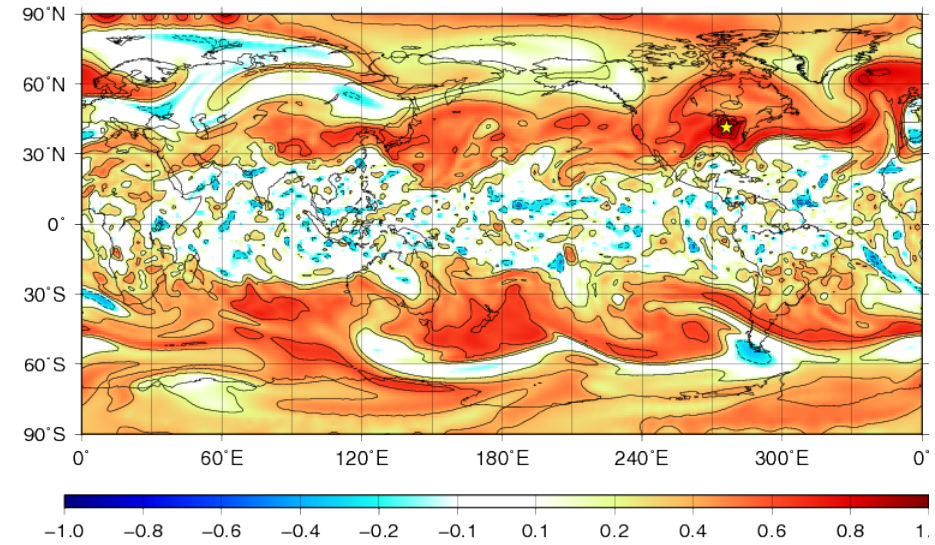


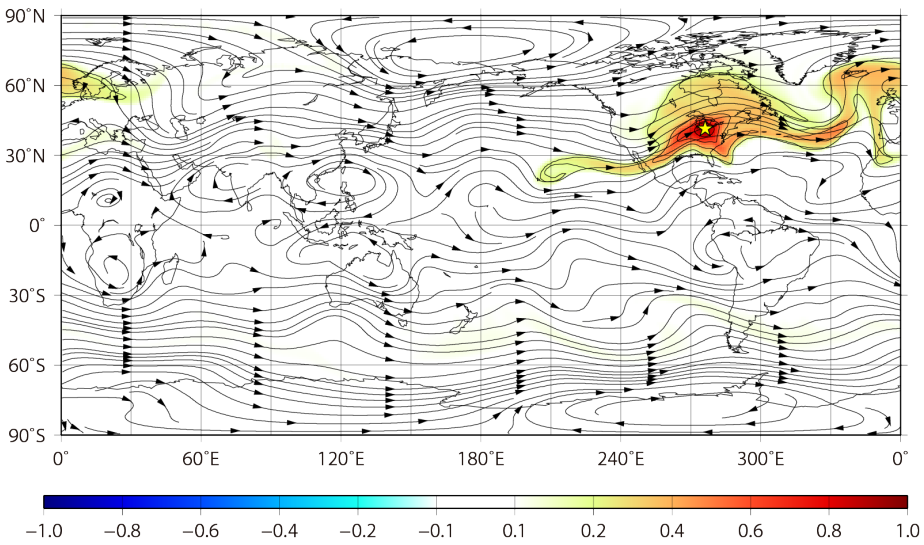


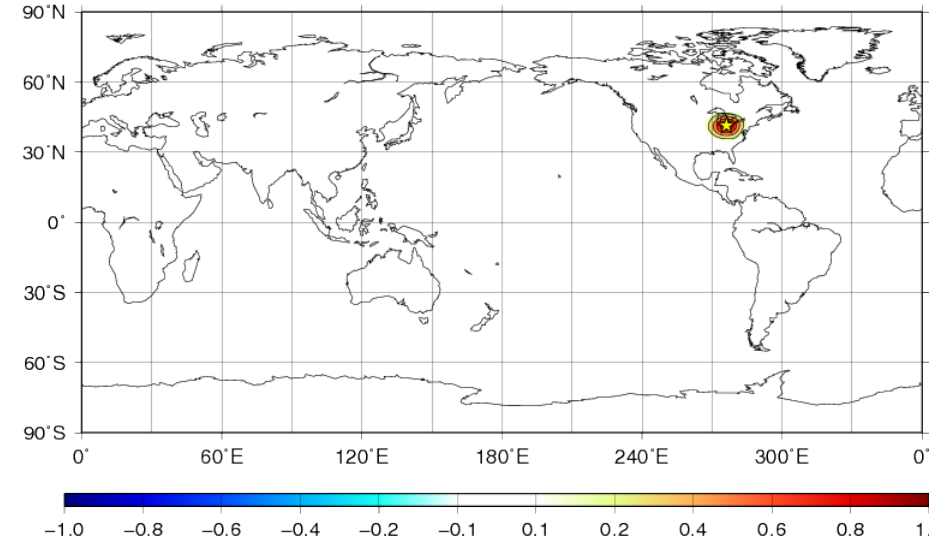
11/8 00UTC after a week cycling

Correlation patterns (Q at ~100 hPa)

40 members *Kondo, Miyoshi (2015)* **10240 members**





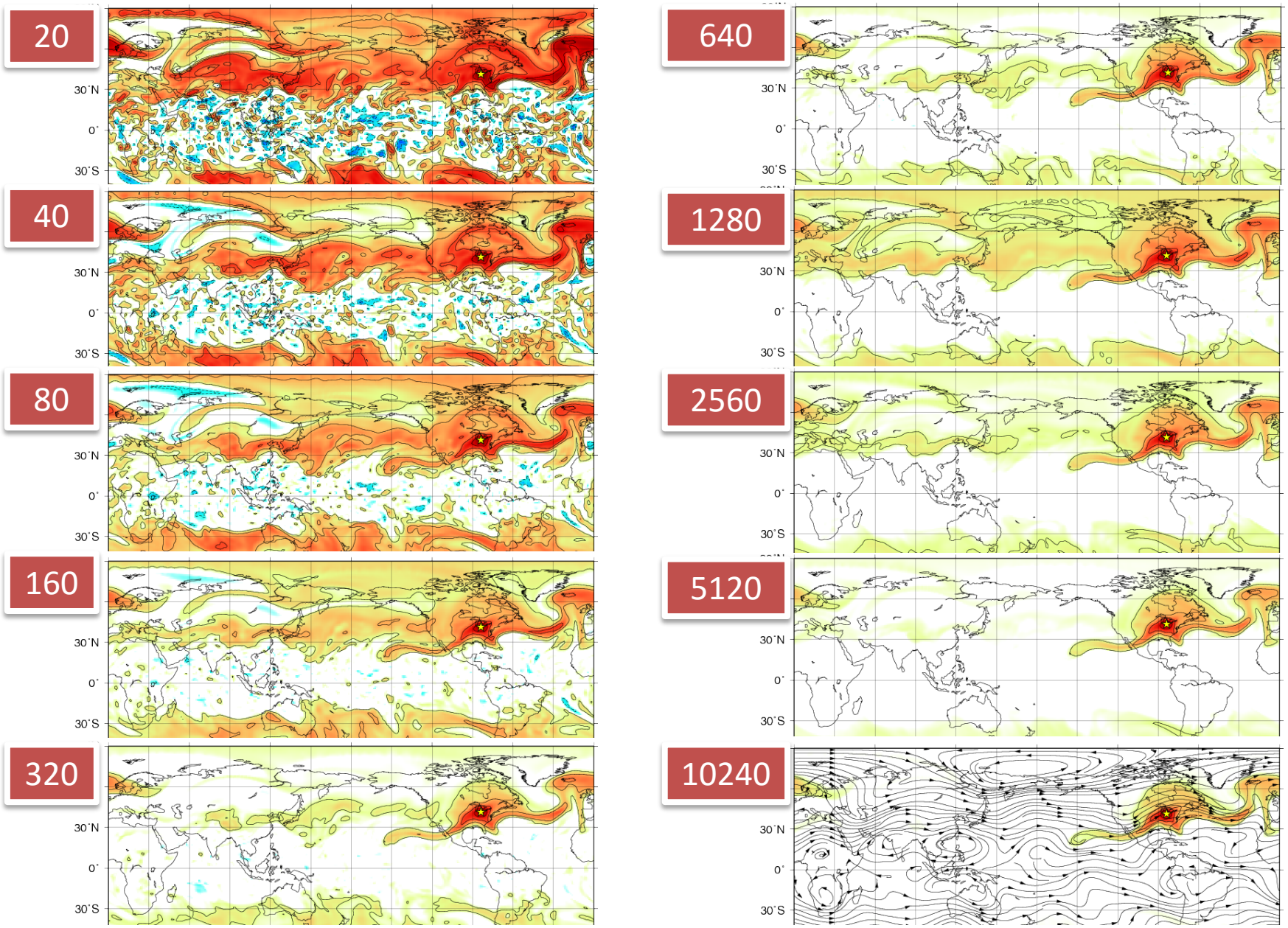


FLOW-DEPENDENT

11/8 00UTC after a week cycling

With subsets of 10240 samples

Kondo&Miyoshi (2015)



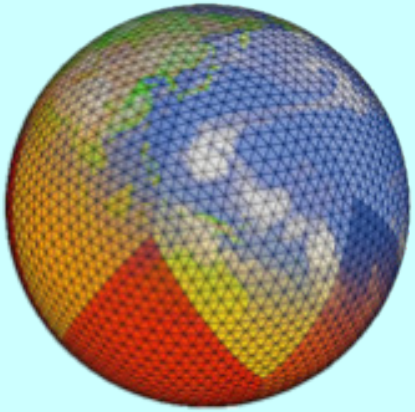
Sources of Big Data

Observations

Big Data

Advanced obs technology

Simulations



Big Data

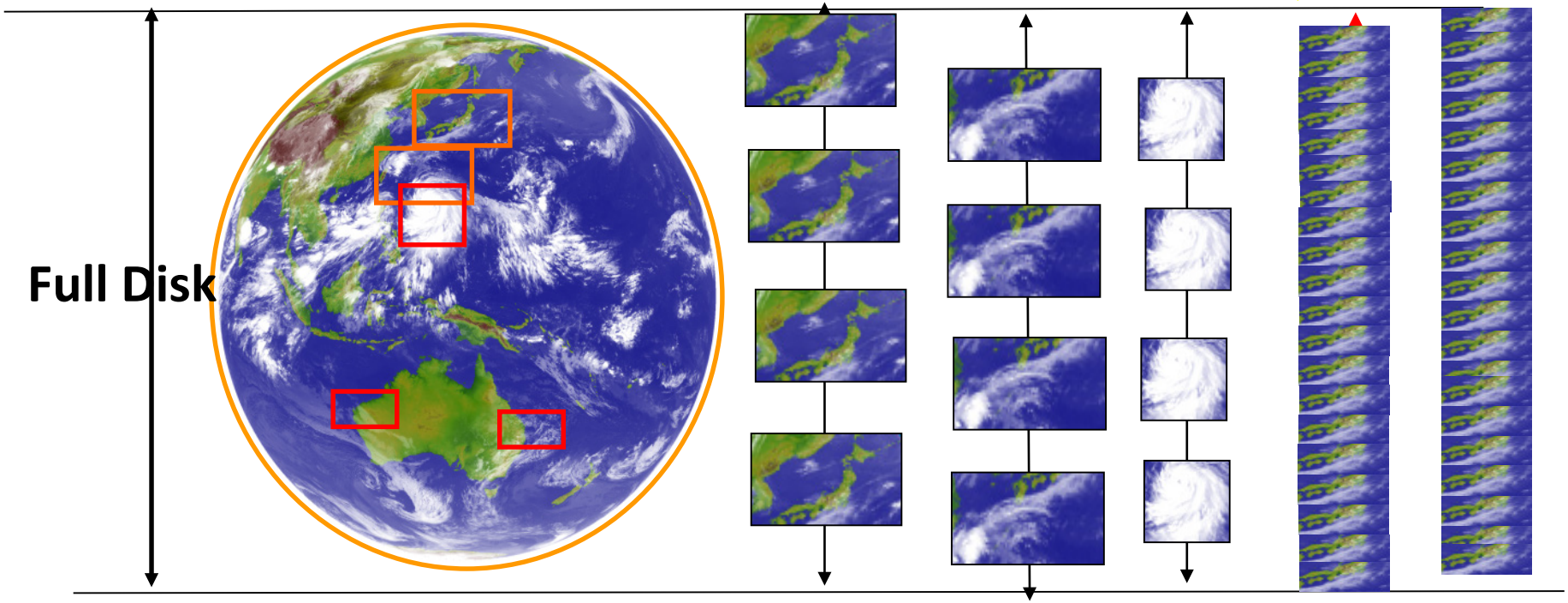
Powerful supercomputer

Next-generation geostationary satellite

Himawari-8 was launched successfully on 7 October 2014.
Himawari-9 will be launched in 2016.

Full operations started on 7/7/2015!!

Super Rapid Scan
every 30 seconds



50x more data!

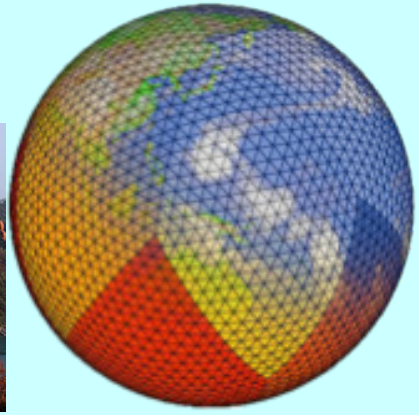
(Courtesy of JMA)

Big Data Assimilation

Observations

Data Assimilation

Simulations



Big Data

Advanced obs technology

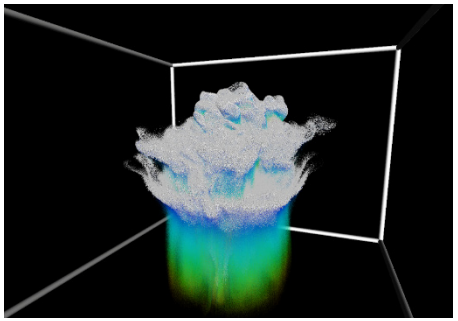
Big Data

Powerful supercomputer

Toward next 20 years of DA

BDA
BIG DATA ASSIMILATION

"Big Data Assimilation" Era



Courtesy of S. Shima

Exploding data

Enabling effective use

Big Data

Big Data

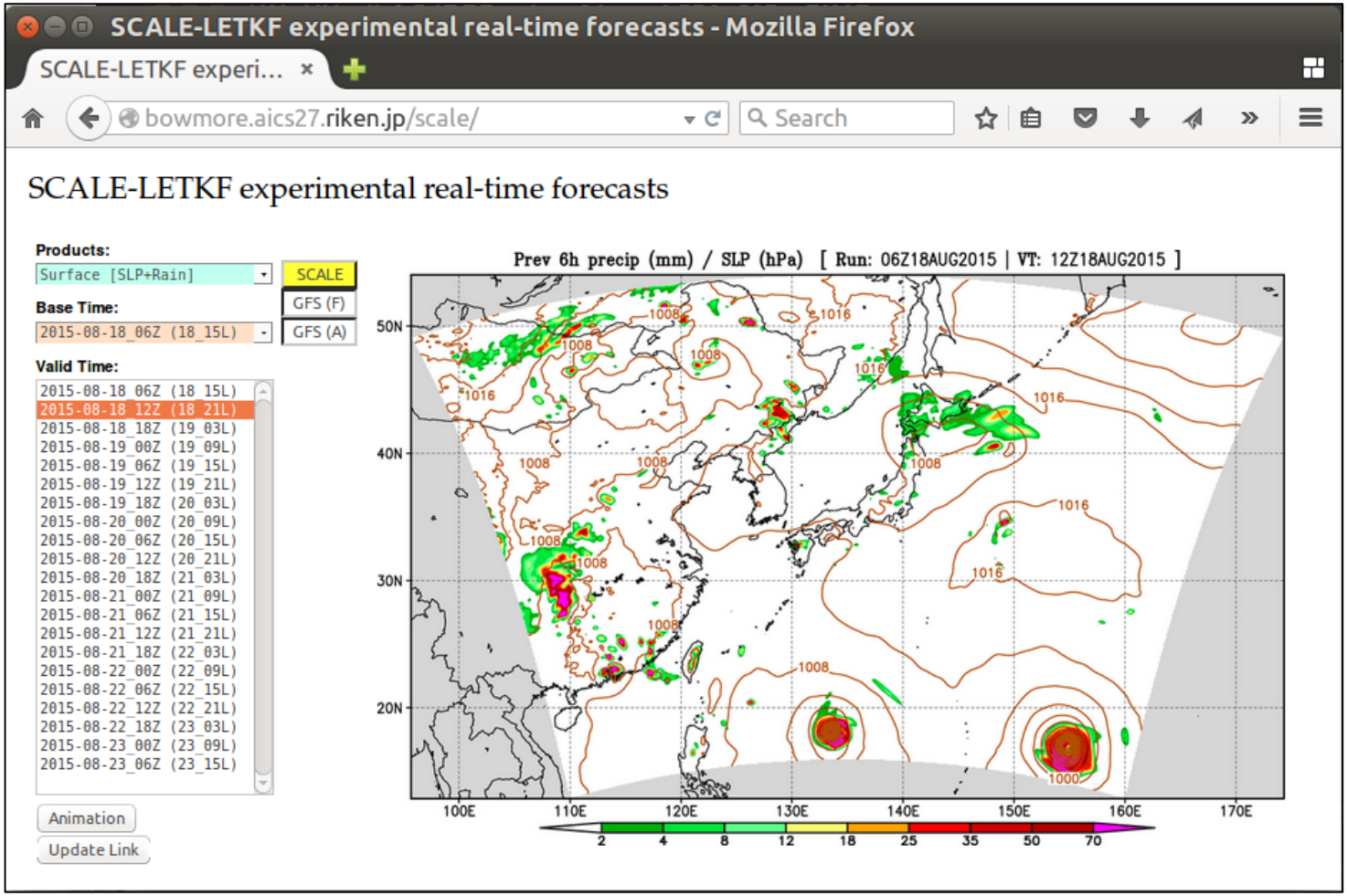
High-resolution simulation

High-resolution obs

More computational power

Advanced obs technology

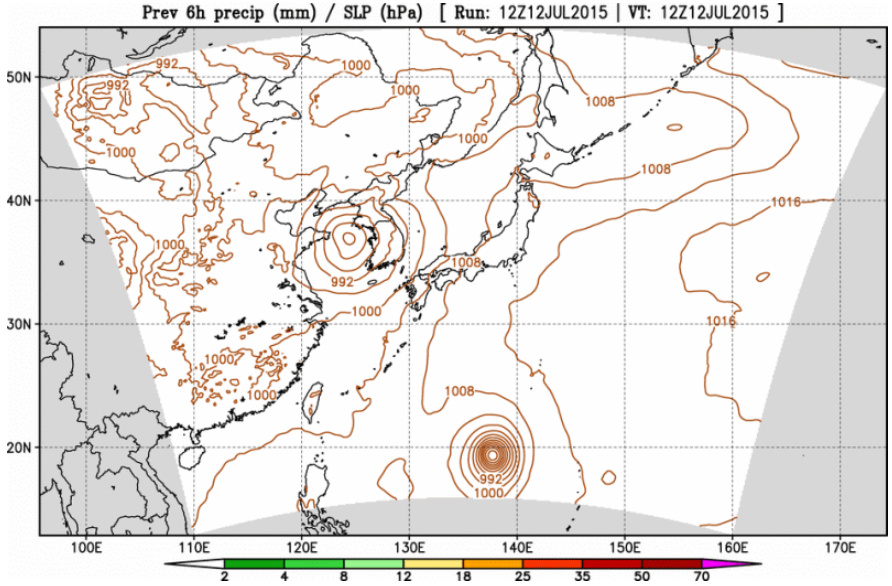
Near-real-time SCALE-LETKF

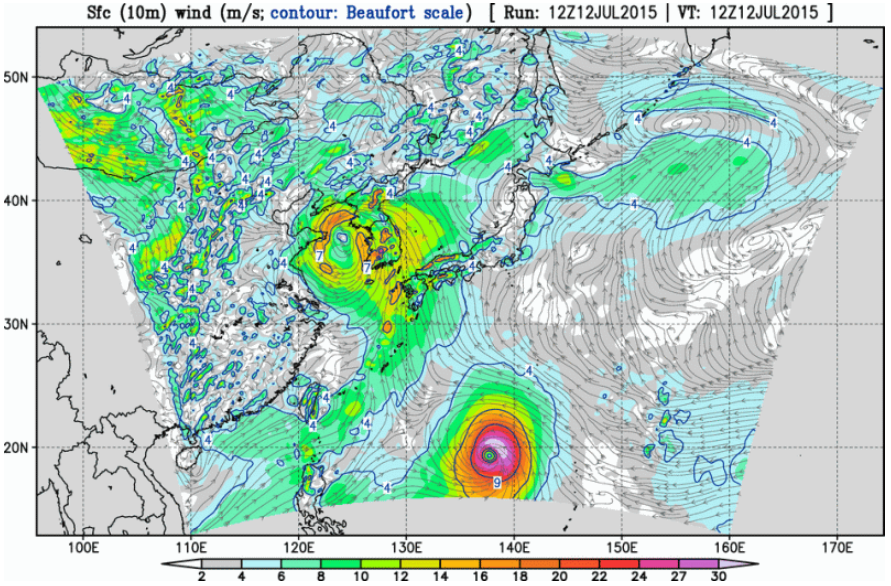


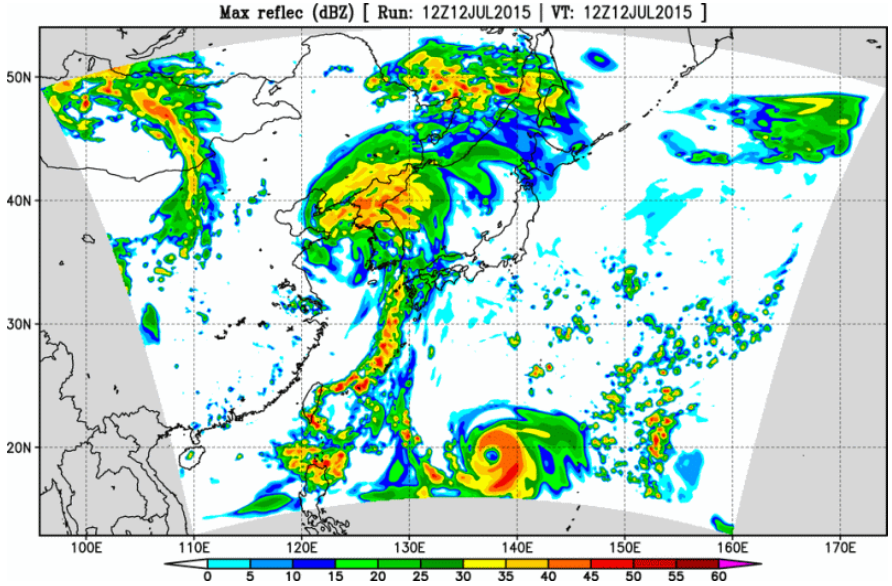
(Lien et al. 2016)

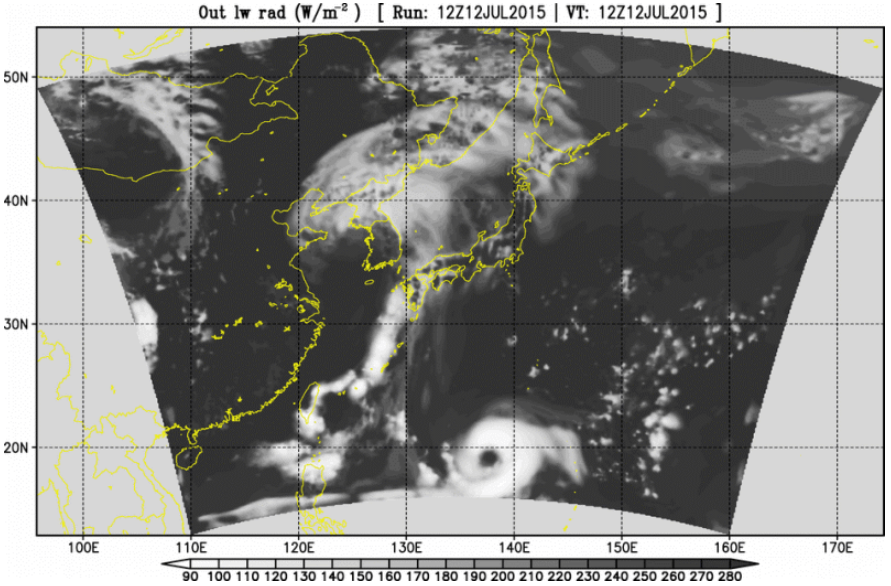
5 day forecast of Typhoon NANGKA (201511) stating at 12:00 UTC July 12

(Lien et al. 2016)



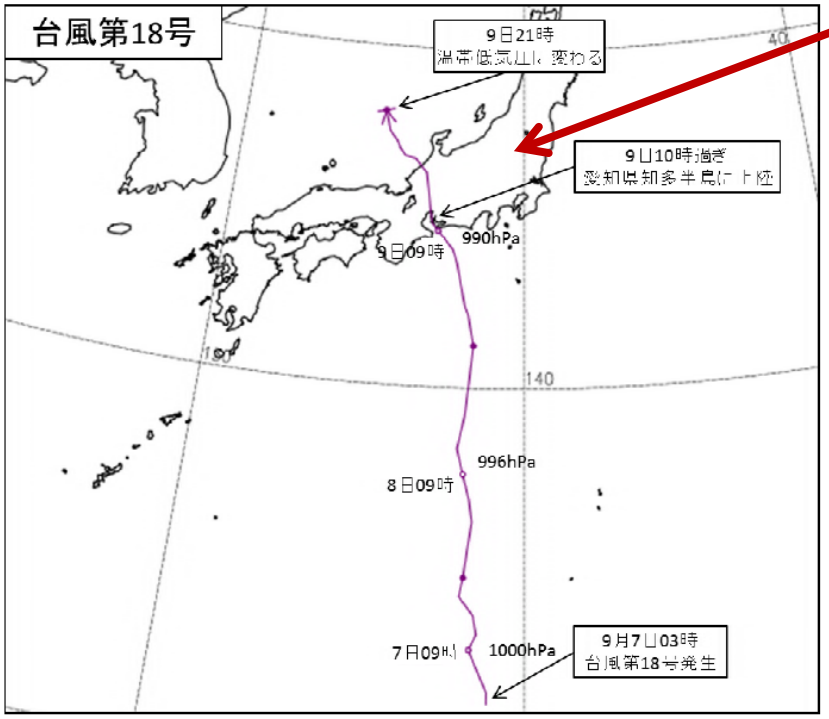






Severe disaster case in Sep. 2015

Track of Typhoon Etau
(Adopted from JMA report)



経路上の○印は傍に記した日の 9 時、●印は 21 時の位置を示す。

~ 600 mm accumulated rainfall in 2 days

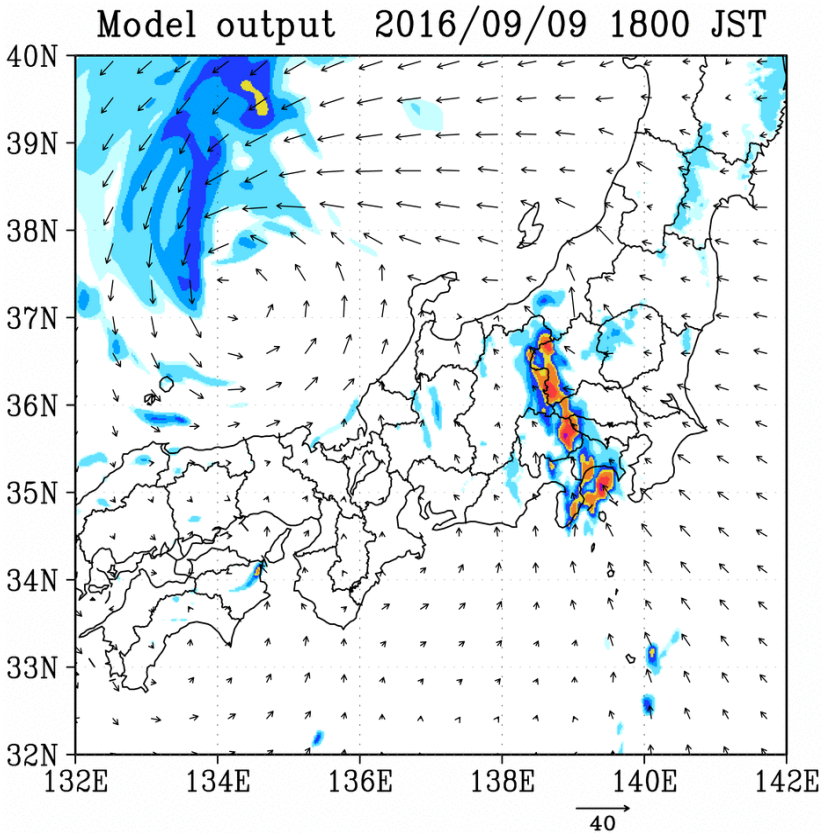
Flooding of Kinugawa River
Asahi News – September 10, 2015.

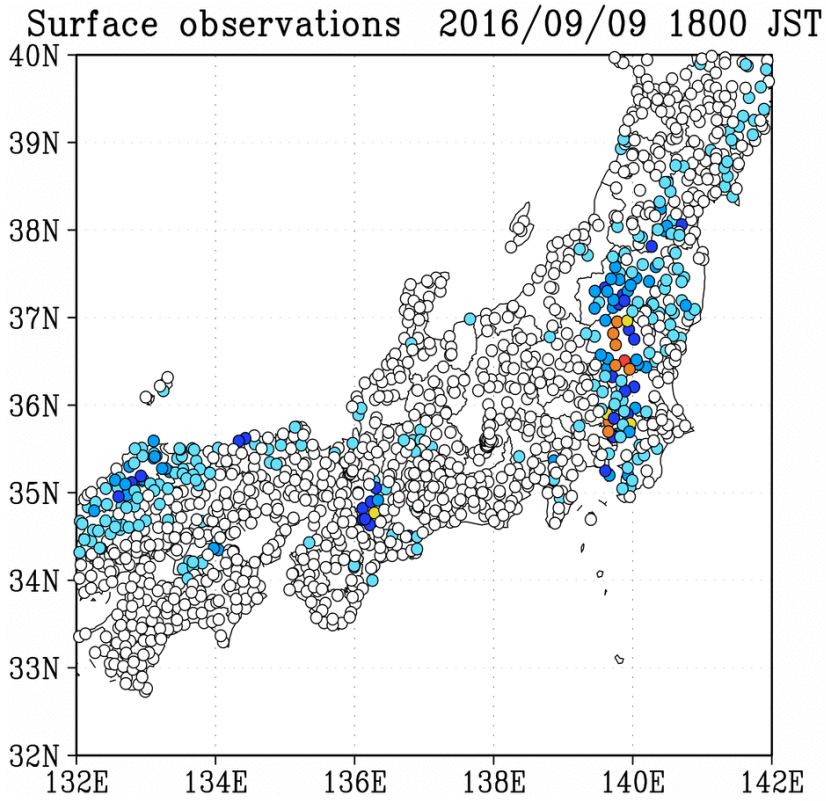
9/9~9/10, 2015

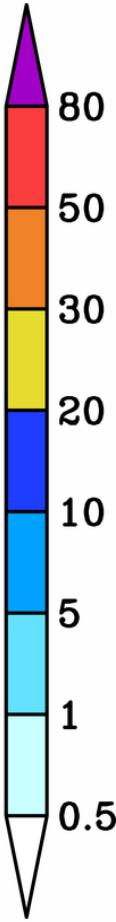
[mm/h]

Initial time: 9/8 21JST

NTT Docomo gauge network





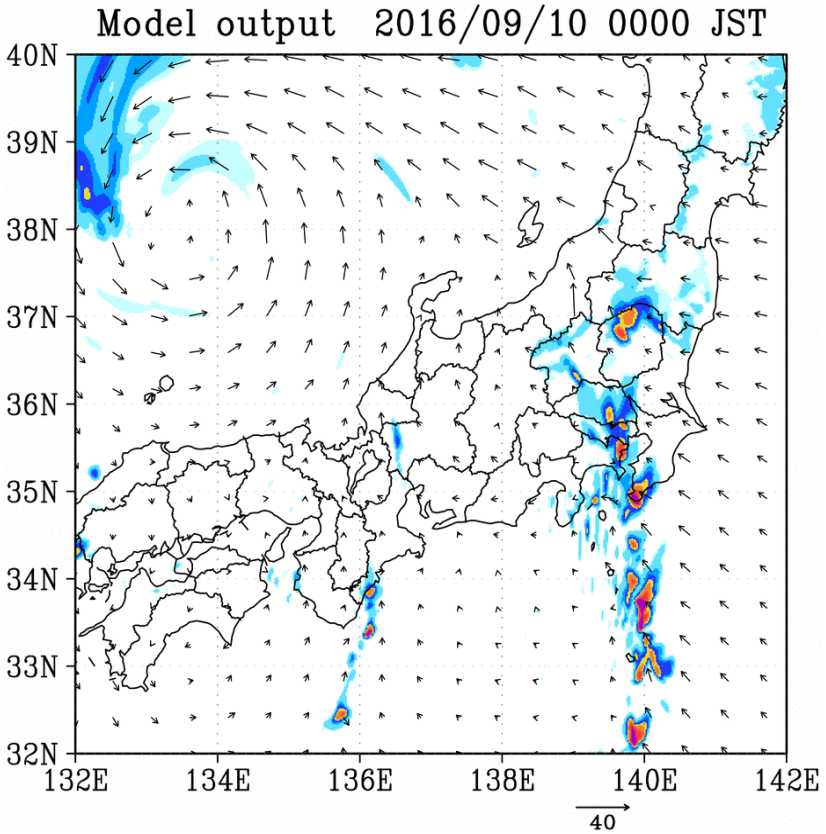


Shifted to the west

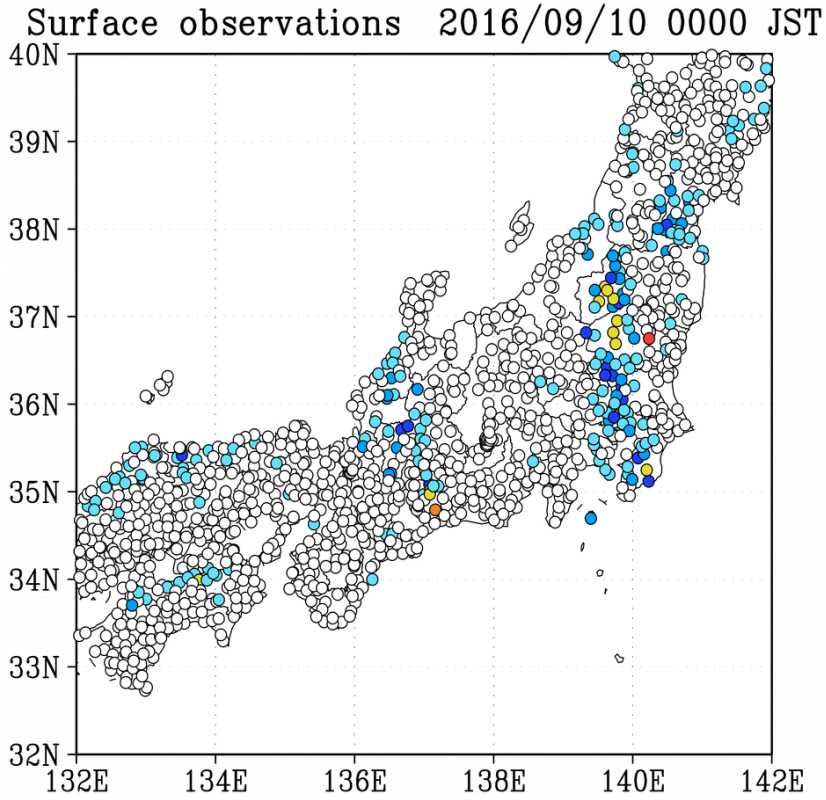
9/9~9/10, 2015

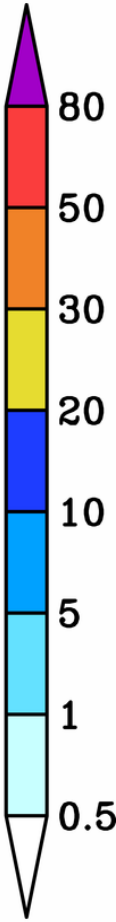
[mm/h]

Initial time: 9/8 21JST



NTT Docomo gauge network





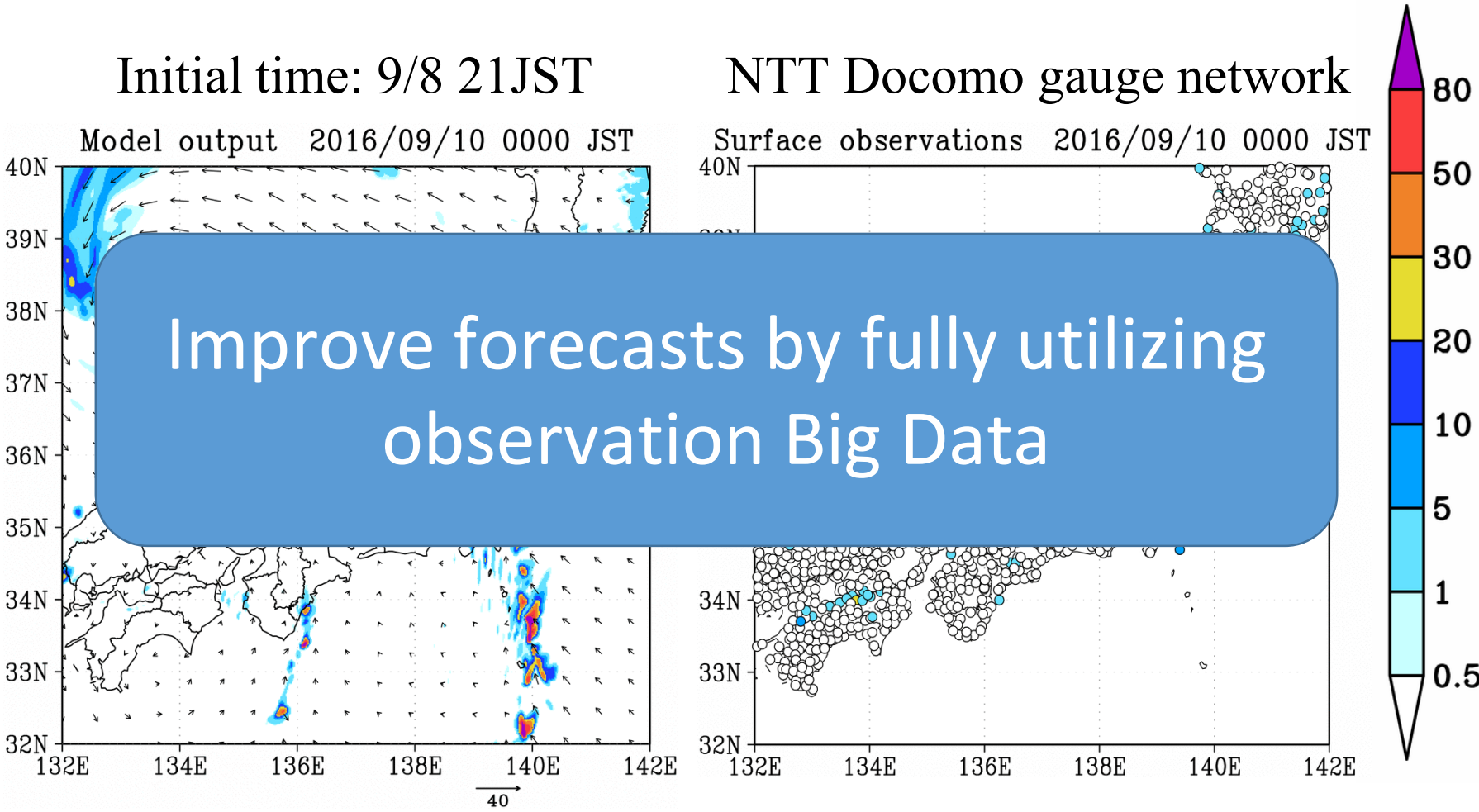
Overall good location, but heavy rain is missing!

9/9~9/10, 2015

[mm/h]

Initial time: 9/8 21JST

NTT Docomo gauge network



Improve forecasts by fully utilizing observation Big Data

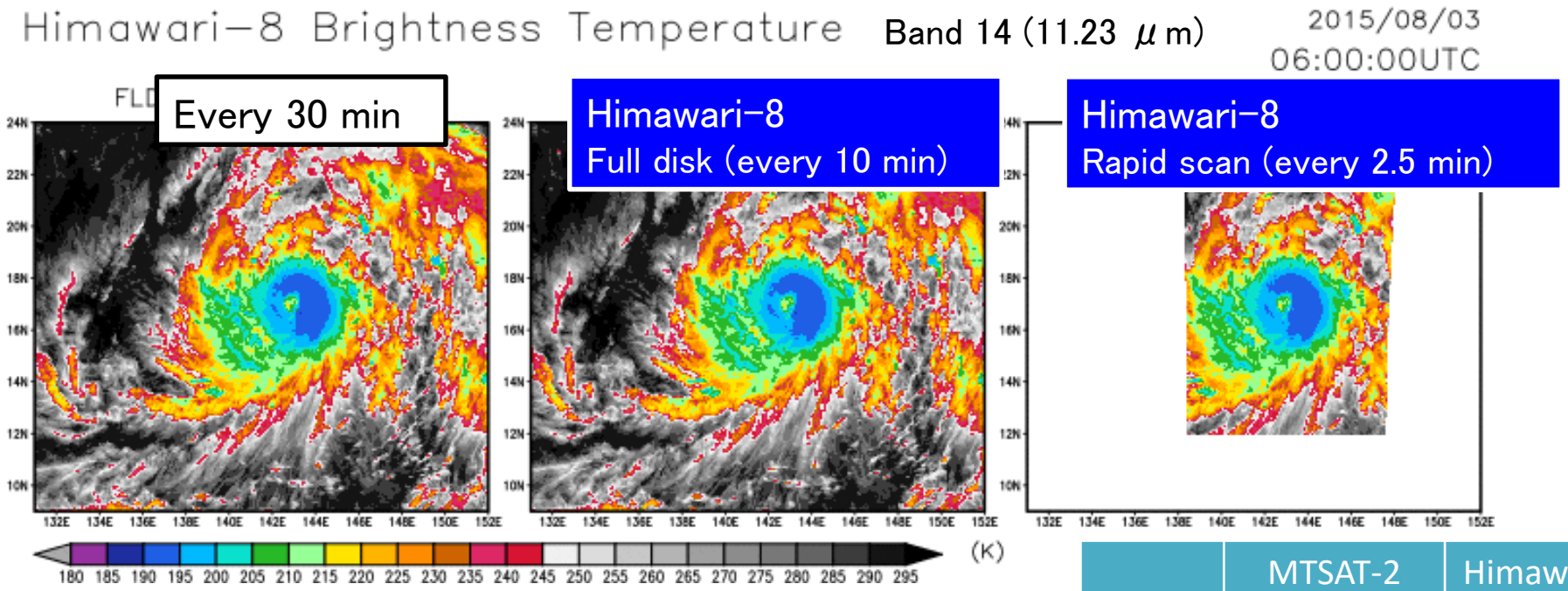
Overall good location, but heavy rain is missing!

Honda et al. (2016)

Himawari-8: A new generation satellite

Providing observation "Big Data" Similar to GOES-R

- High-spatiotemporal resolution radiance obs in 16 bands.



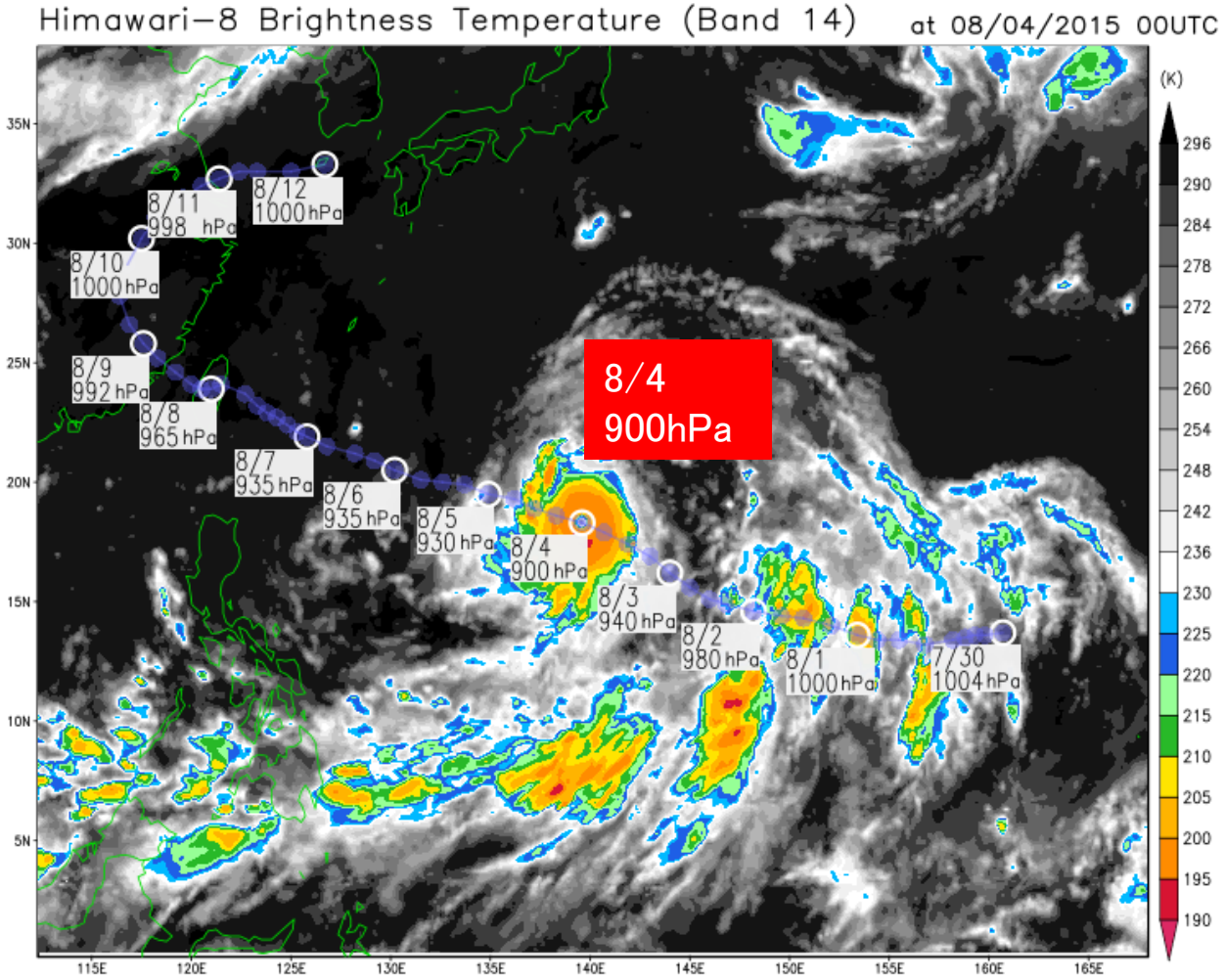
	MTSAT-2		Himawari-8	
VIS	1	5	3	16
NIR	0		3	
IR	4		10	

Number of bands

Honda et al. (2016)

Typhoon Soudelor (2015)

- The strongest northwestern Pacific TC in 2015.
- Himawari-8 observed successfully!



Juan, PR

Band 9

Before DA

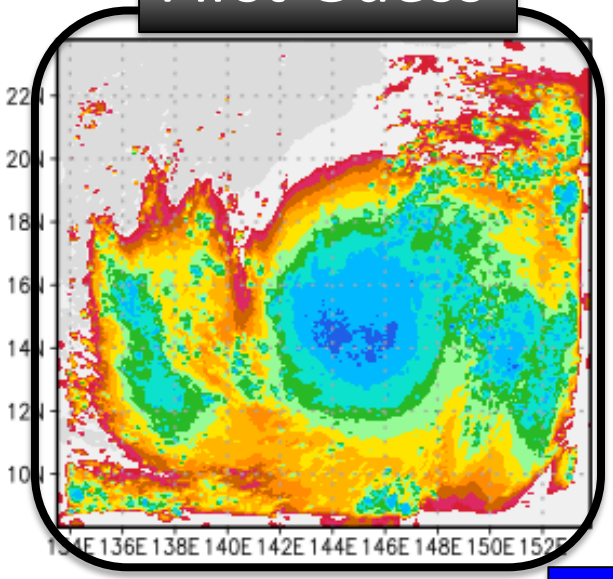
12:10z02Aug2015
1st cycle

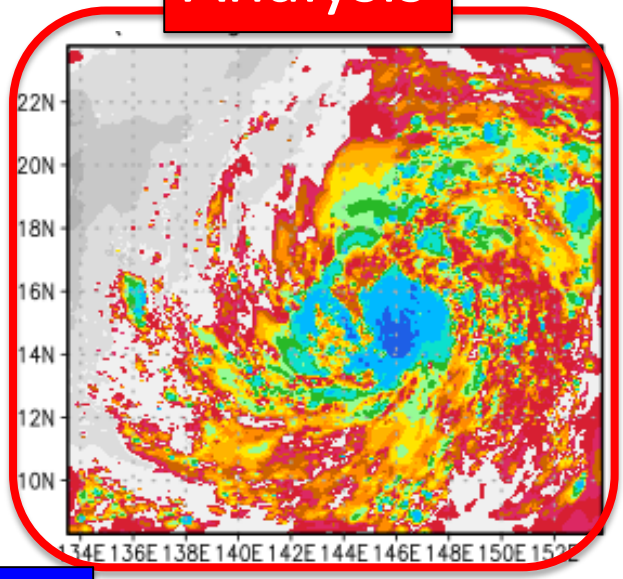
After DA

First Guess

LETKF

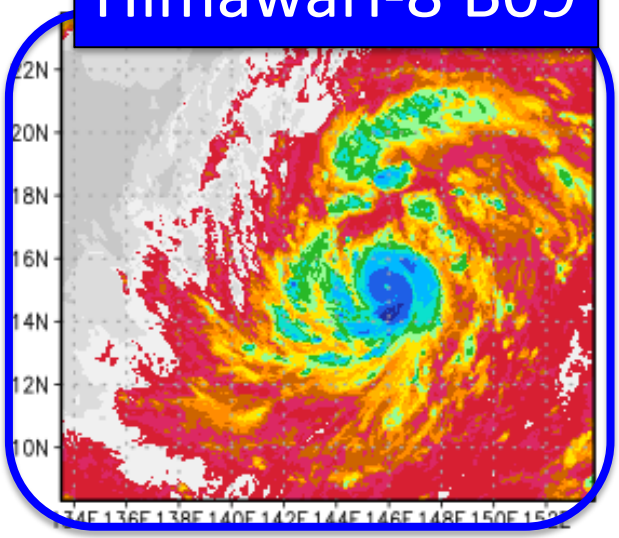
Analysis





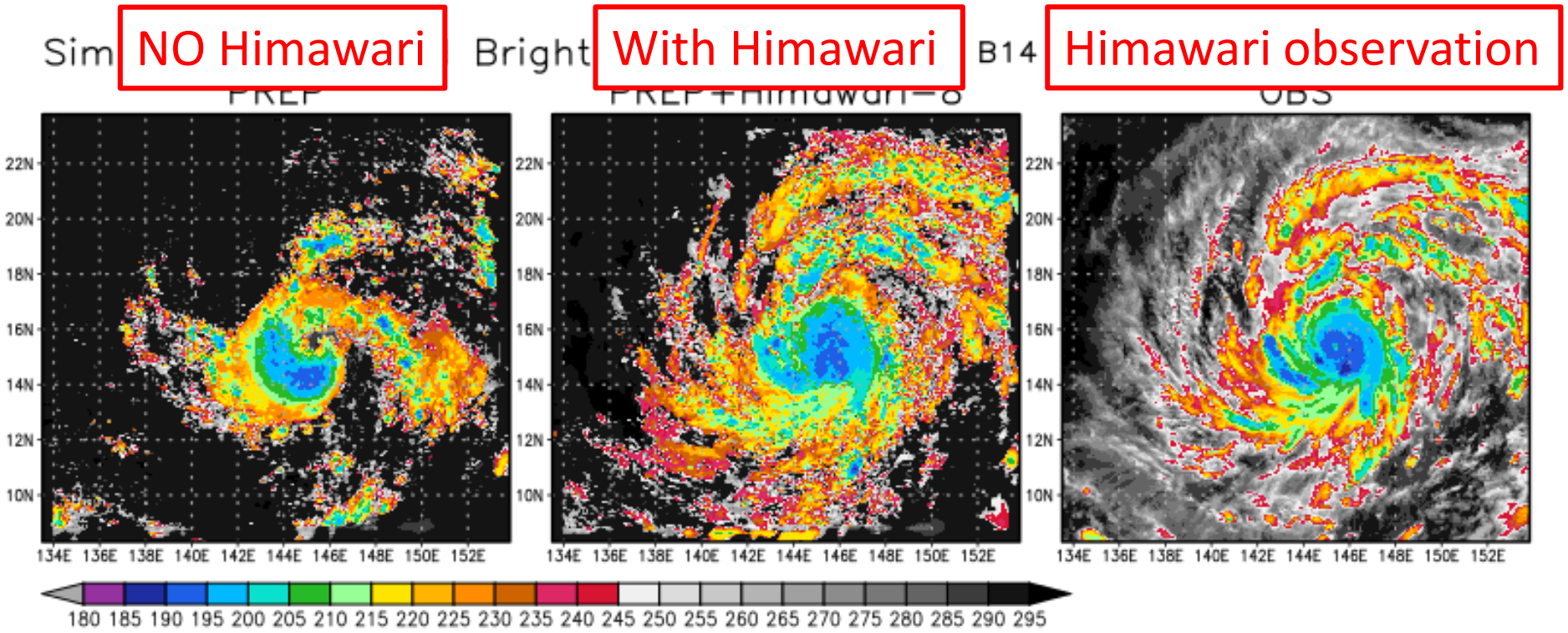
OBS

Himawari-8 B09



(Honda et al. 2016)

Band 14 Not assimilated



(Honda et al. 2016)

Honda et al. (2016)

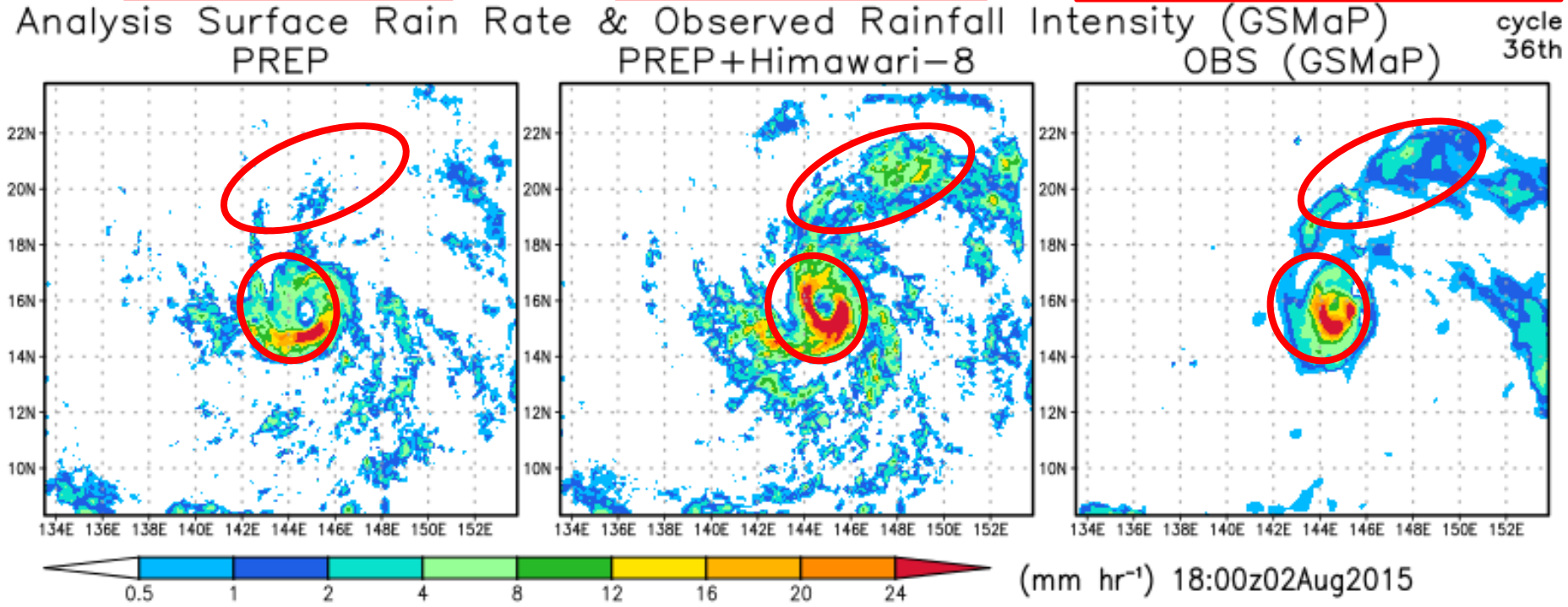
Precipitation patterns

not used

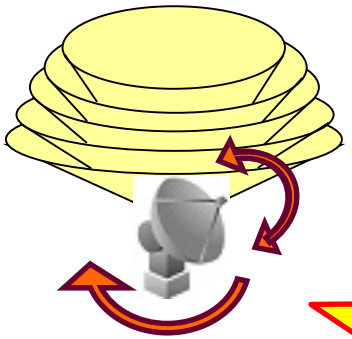
NO Himawari

With Himawari

GSMaP Observation



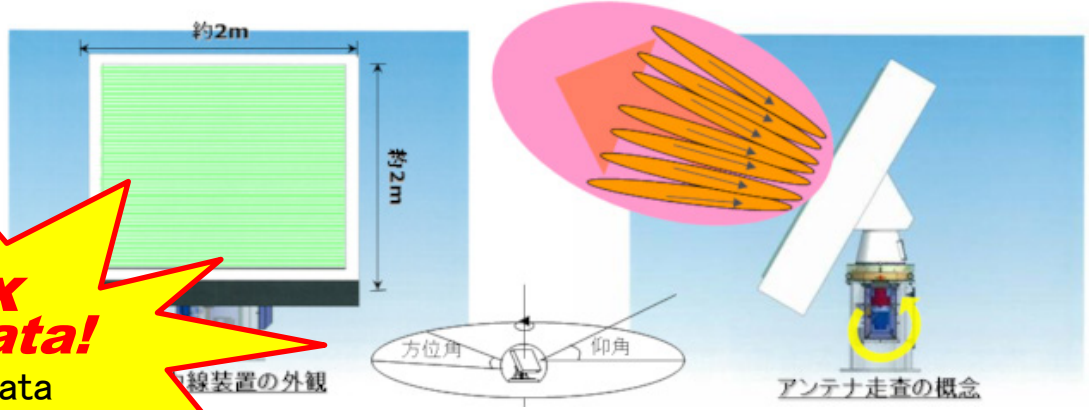
Phased Array Weather Radar (PAWR)



3-dim measurement using
a parabolic antenna (150 m,
15 EL angles in 5 min)

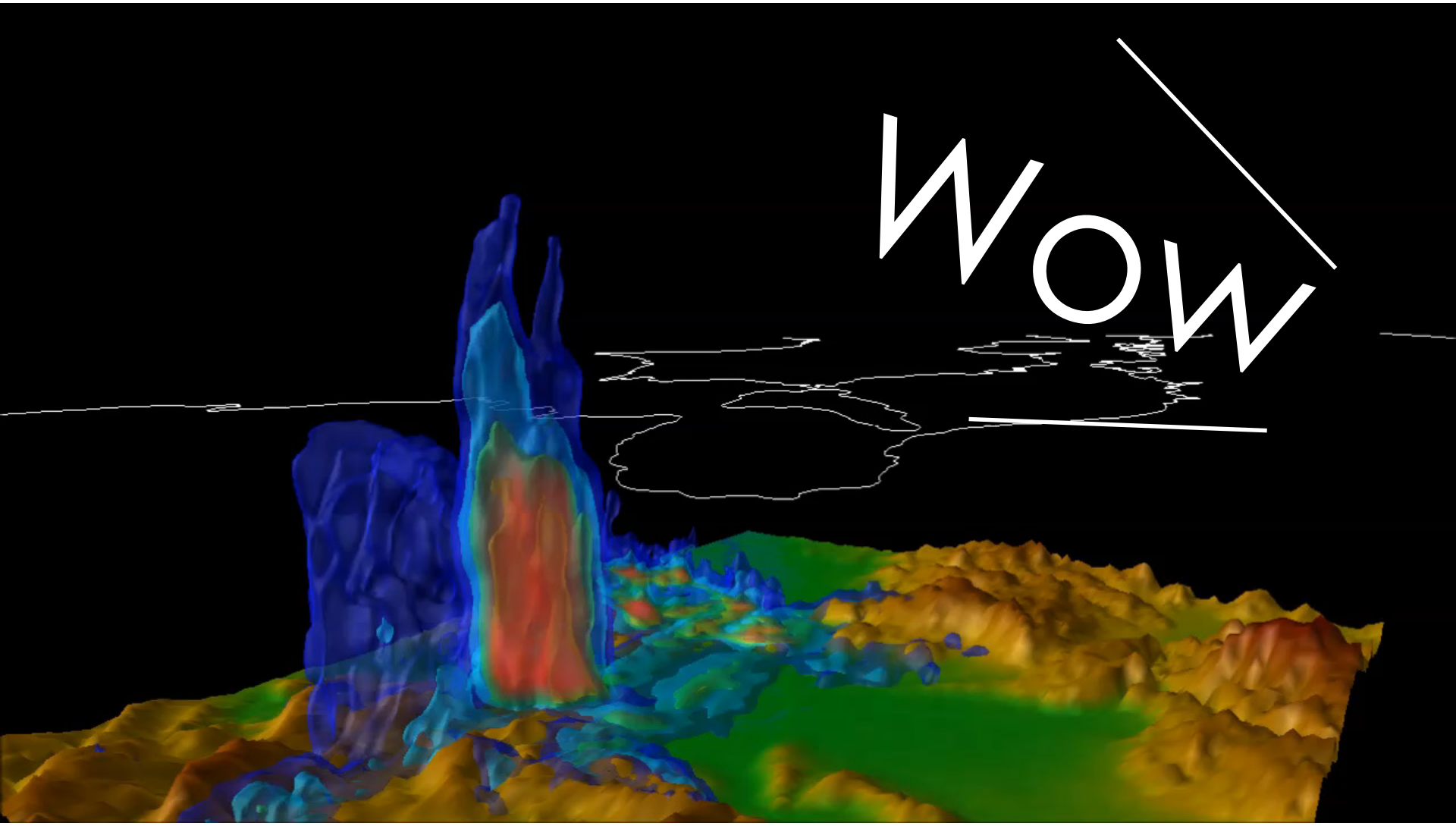
100x more data!

10x more data
in a 1/10 period



3-dim measurement using a phased array antenna
(100 m, 100 EL angles in 30 sec)

Phased Array Radar (every 30 sec.)

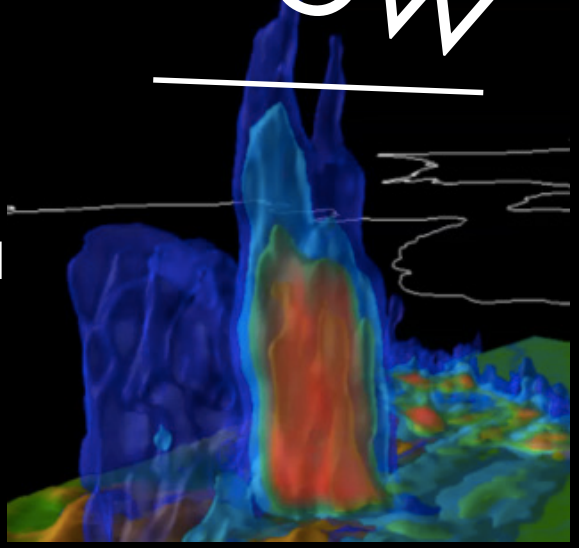


(Courtesy of NICT)

Wow

+

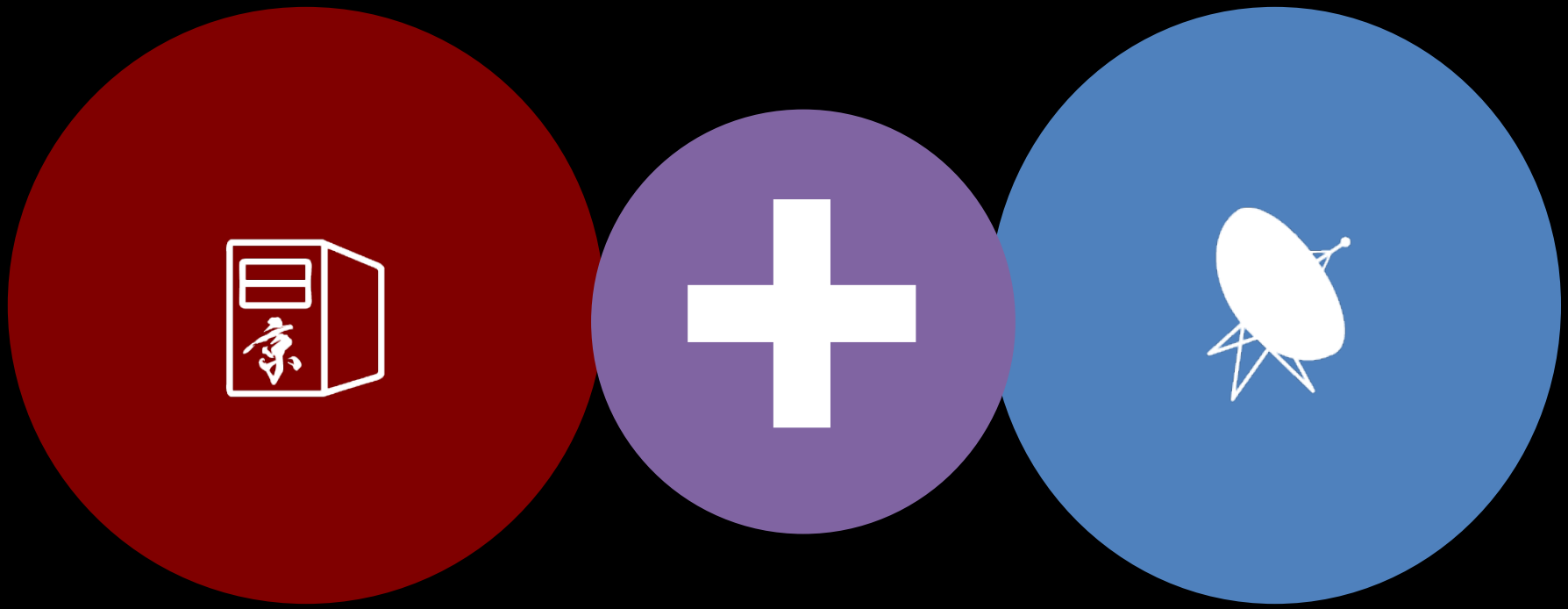
Wow

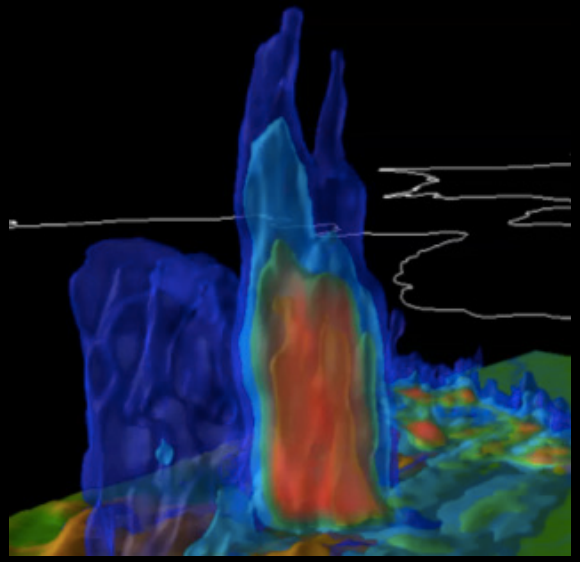


=

?

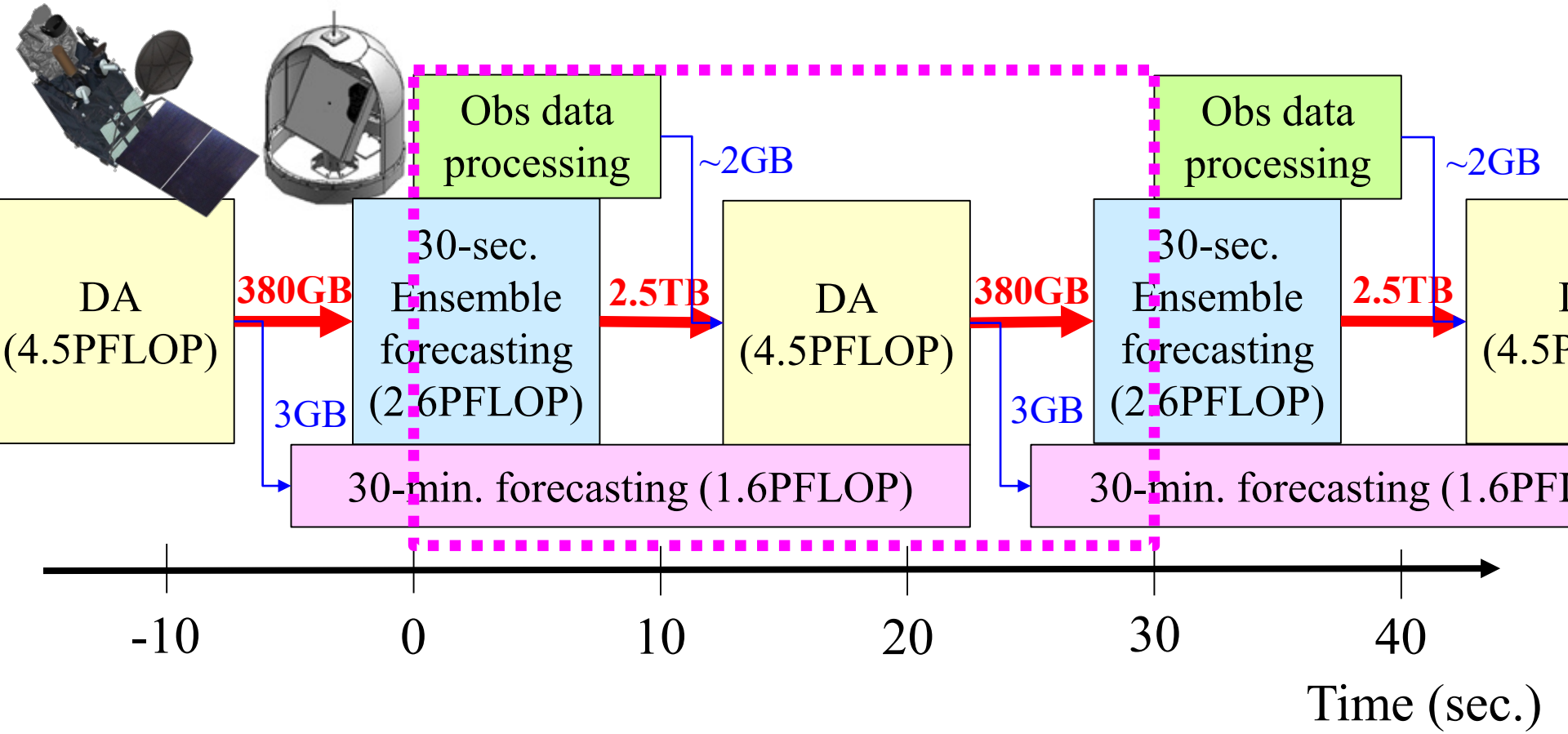
Data Assimilation





= ~~Sudden~~ heavy rain

Revolutionary super-rapid 30-sec. cycle



120 times more rapid than hourly update cycles

A case selected for the first offline study

讀賣新聞

2013年(平成25年)
7月14日日曜日

猛暑のち局地的豪雨

日本列島は13日、西日本を中心に引き続き猛暑となったが、午後には上空の寒気の影響で大気の状態が不安定になり、近畿や中国では局地的に落雷や激しい雨に見舞われた。

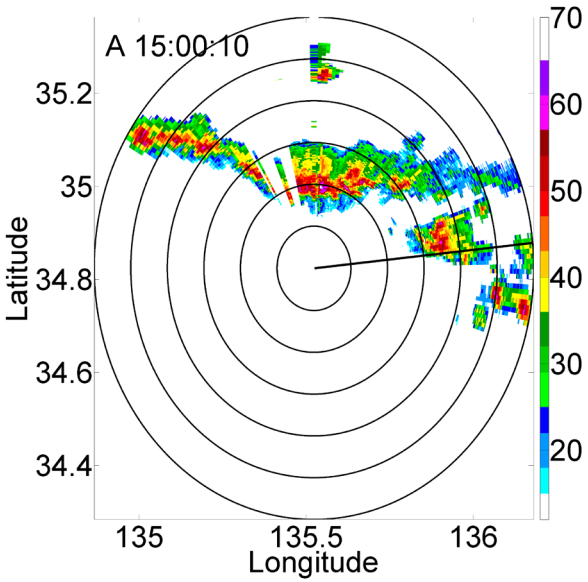
大阪管区気象台によると、1時間の降水量では、京都府京田辺市で56・5ミリ（午後4時27分まで）、広島県北広島町で50・5ミリ（午後5時半まで）を記録した。京都府宇治、城陽両市で床上浸水1軒、床下浸水22軒の被害が出たほか、京都市内で一時、約3300軒が停電した。

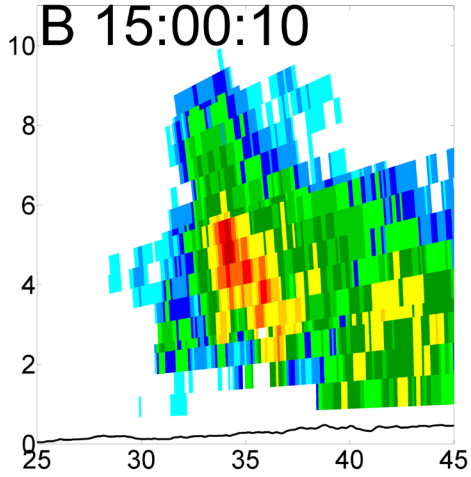
祇園祭の山鉾が並ぶ同市中心部では、午後2時過ぎから雷を伴う強い雨が降った。観光客らはアーケードの下で雨宿りするなか、芸舞妓が足元を気遣いながら道を急いでいた＝写真＝。（宇那木健一撮影）

The top page of Yomiuri newspaper
on 14 July, 2013

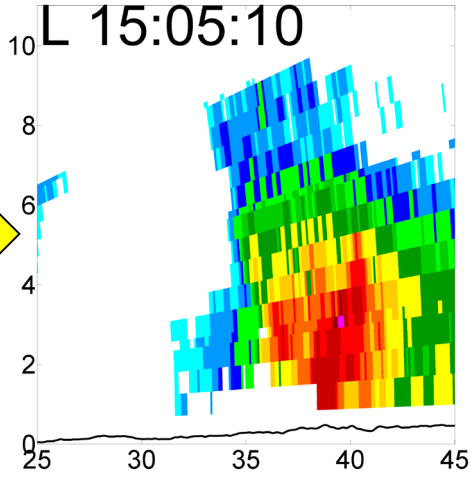
30-sec. and 5-min. evolutions

(Miyoshi et al. 2016)



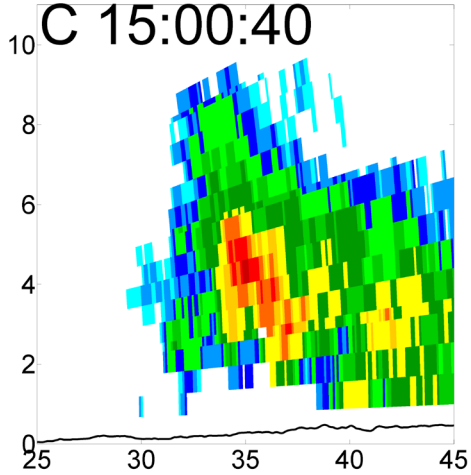


5 min



nonlinear evolution

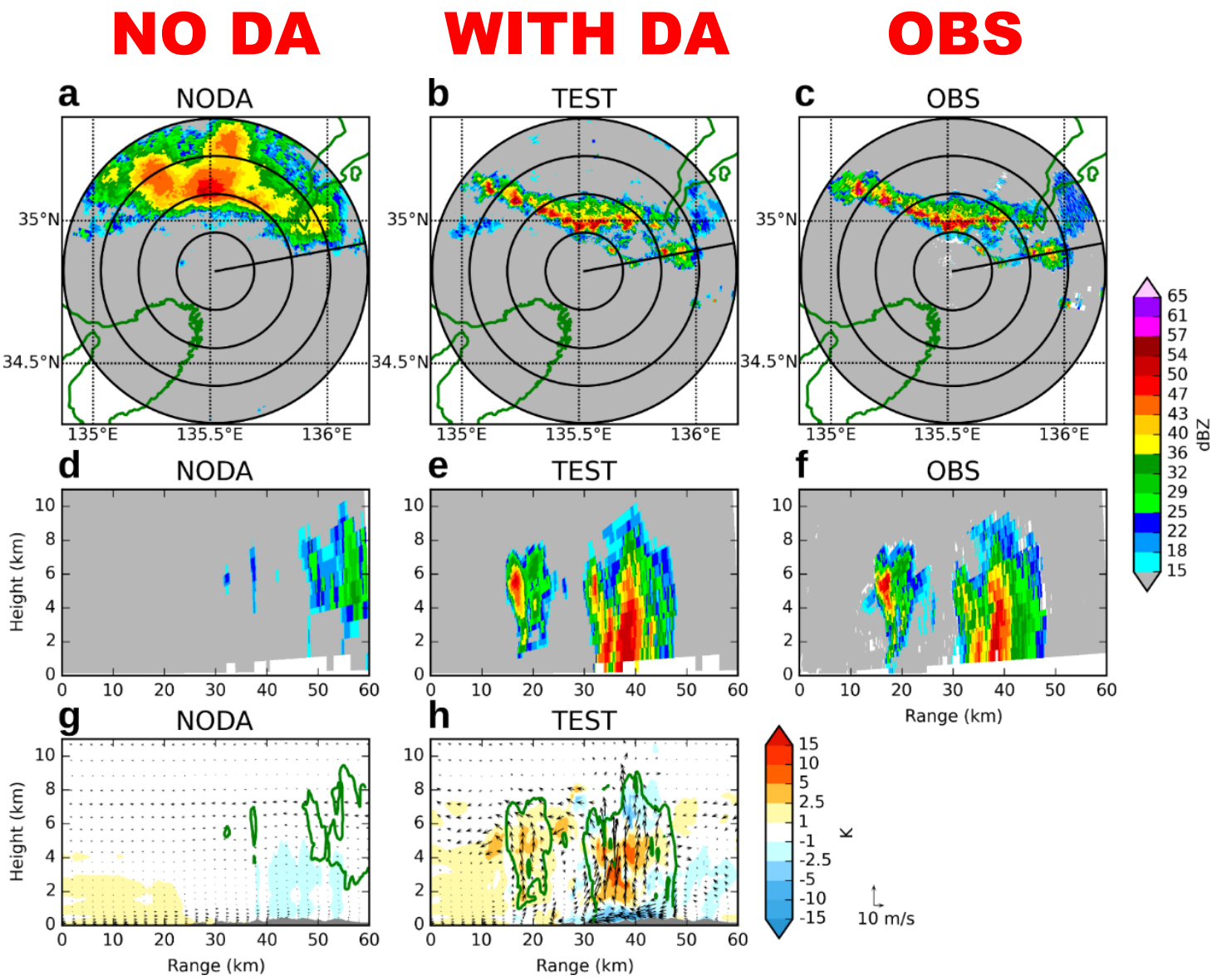
30 sec



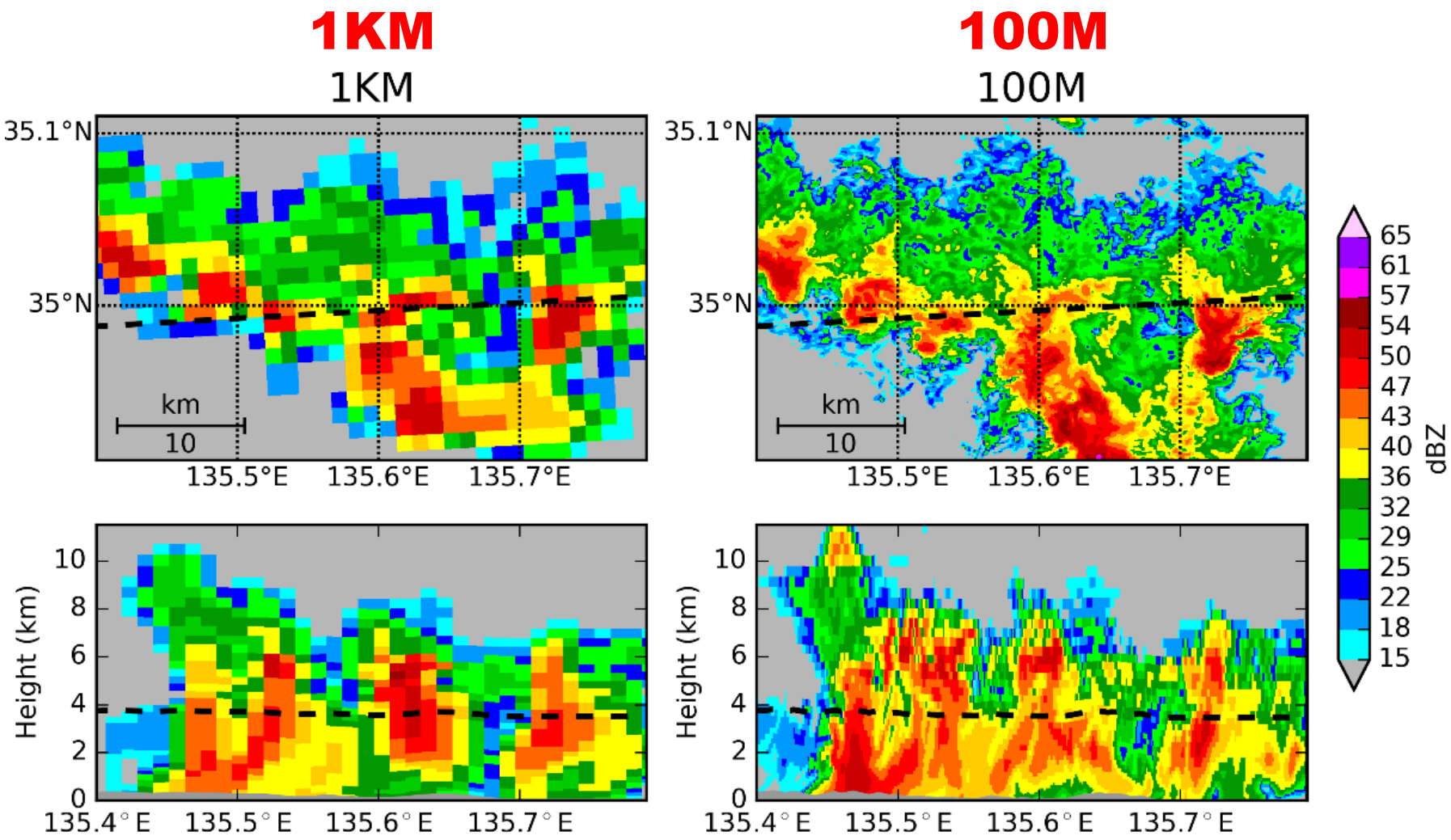
linear evolution

First results

(Miyoshi et al. 2016)



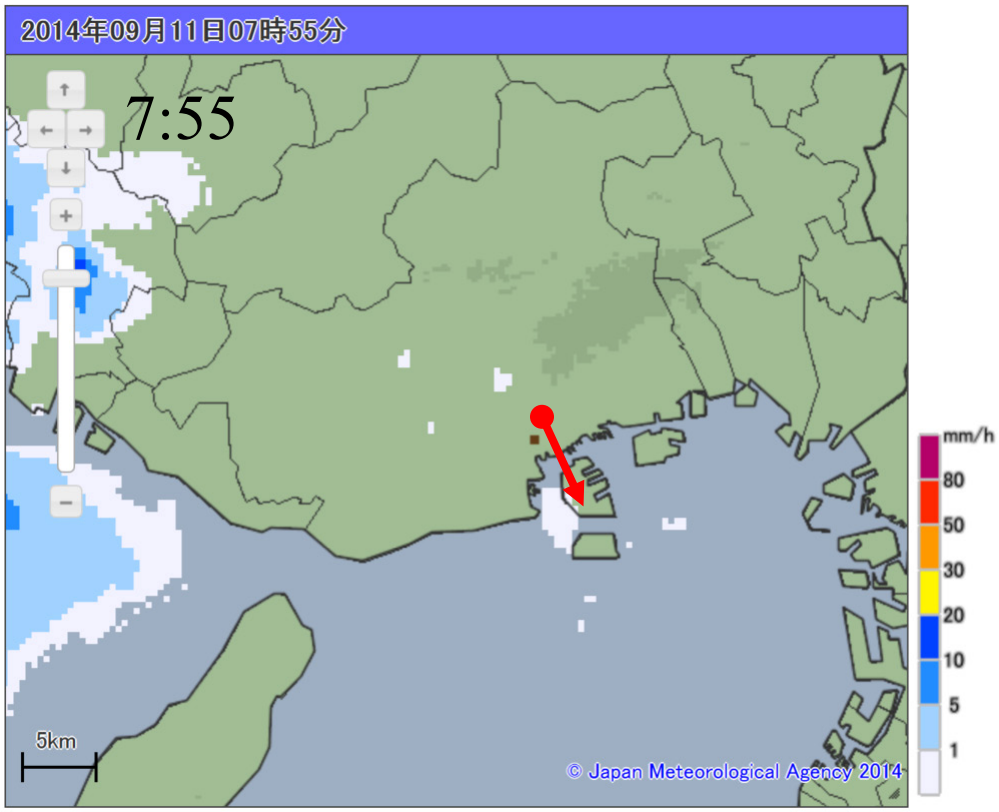
Advantage of 100-m simulation



(Miyoshi et al. 2016)

9/11/2014 morning, biked to office

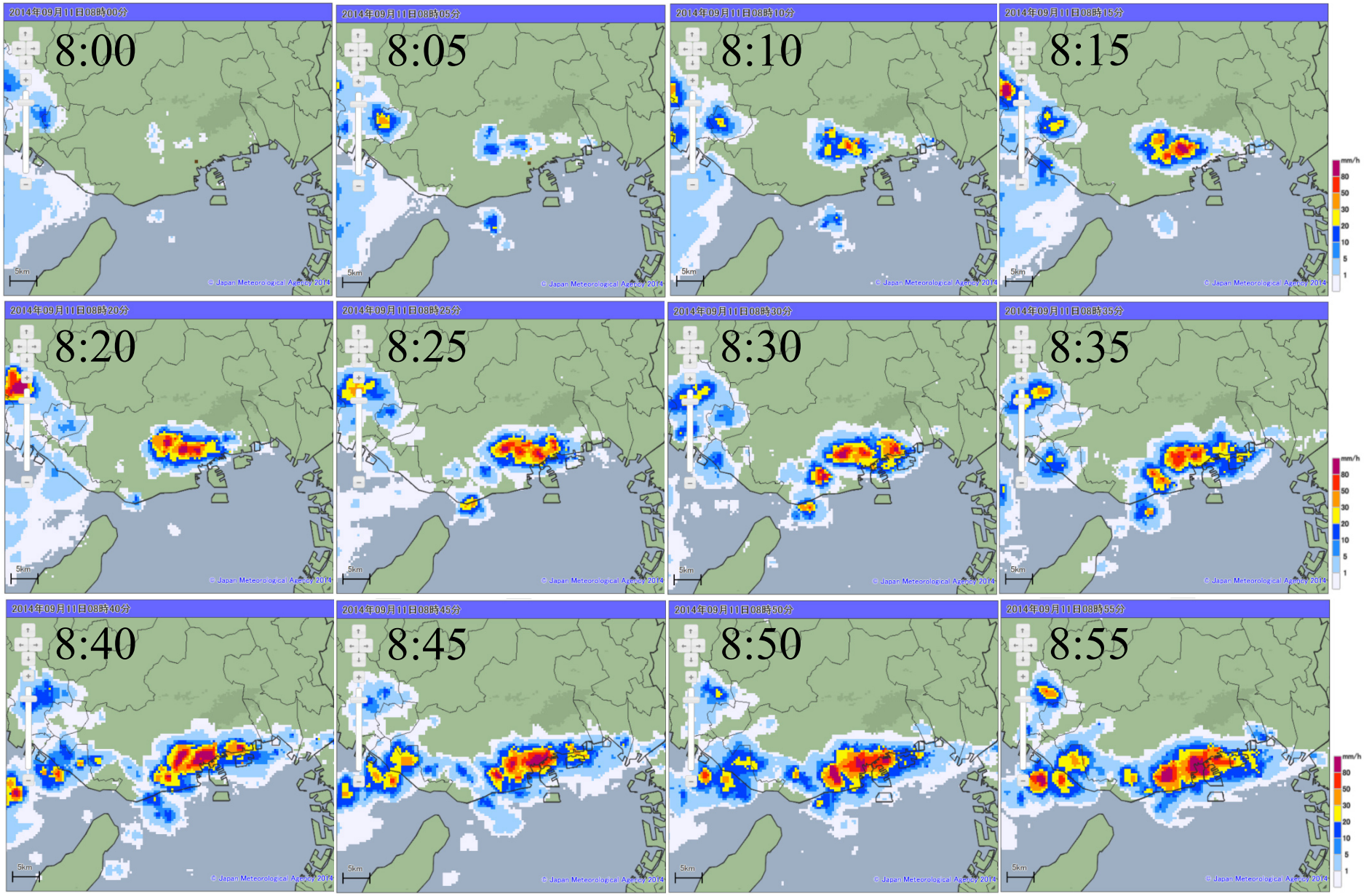
JMA observation



I looked at this obs
at 8 am.

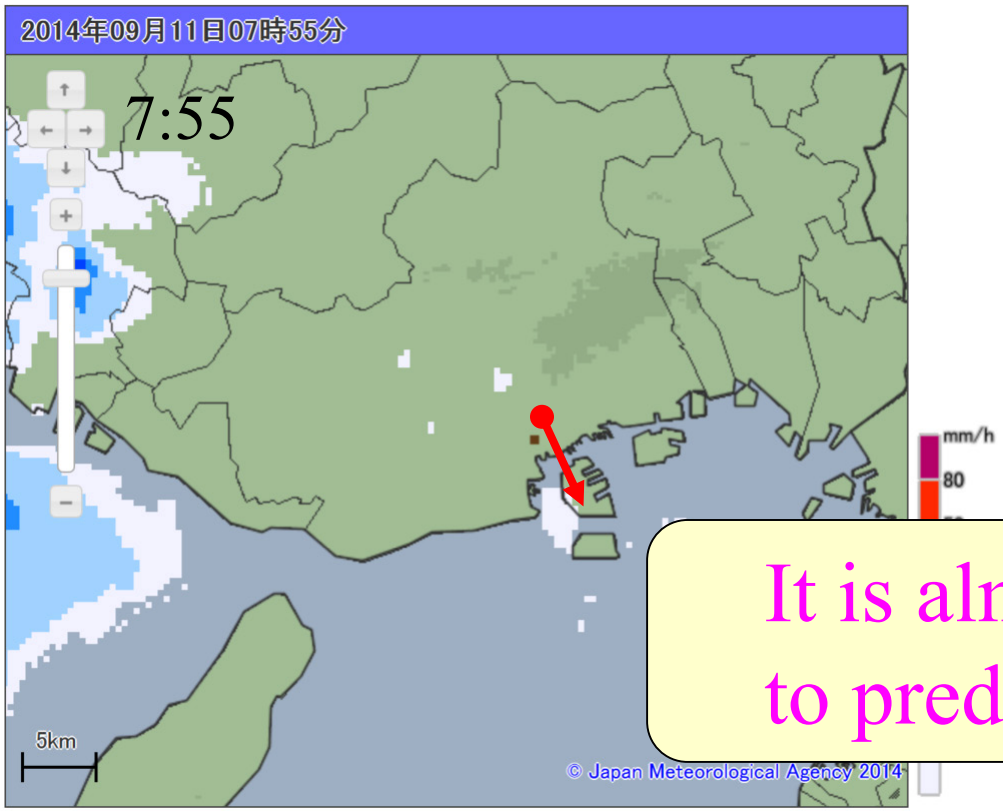
Decided to bike to office.
It takes about 30 min.

9/11/2014 morning, biked to office



9/11/2014 morning, biked to office

JMA observation

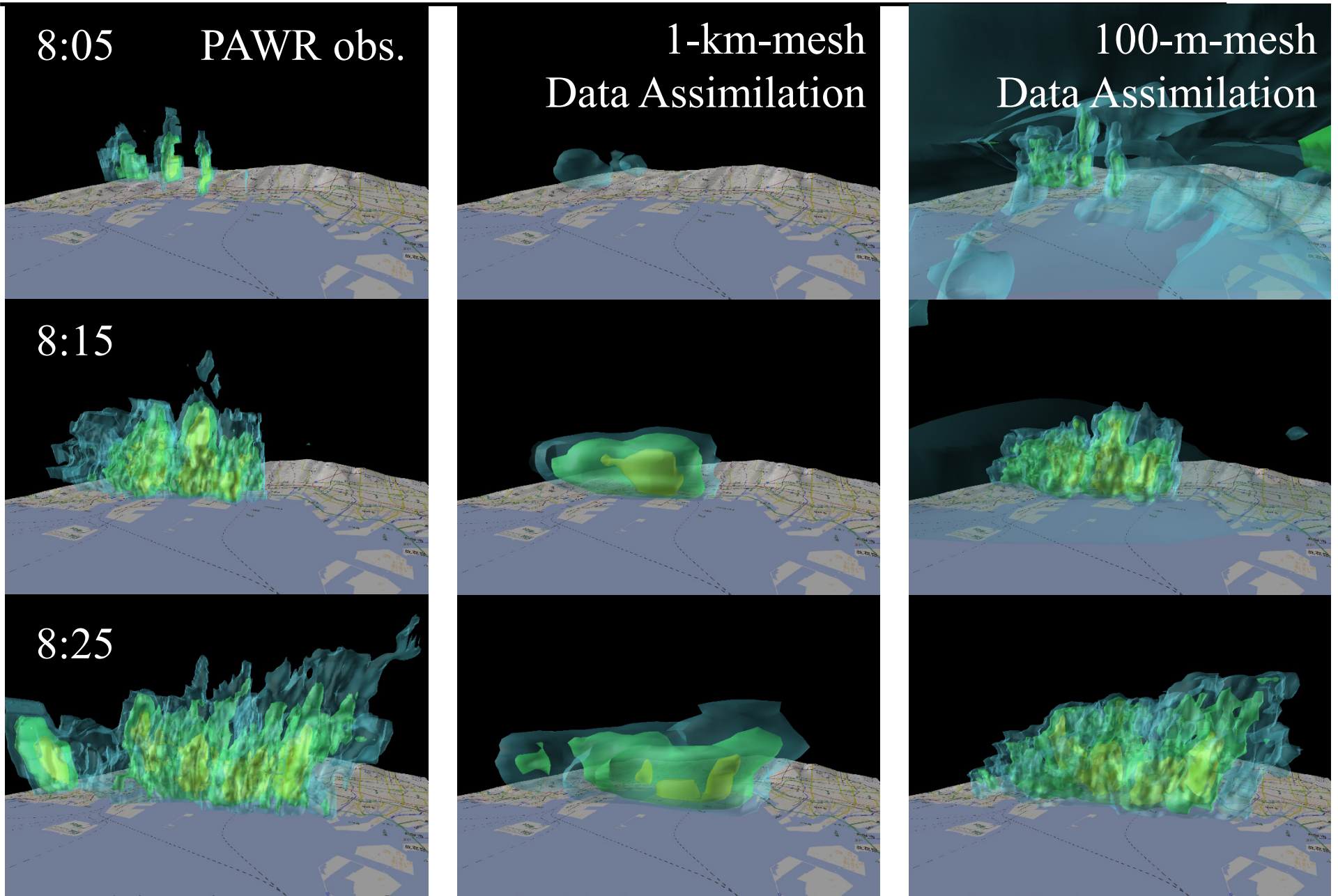


I looked at this obs
at 8 am.

Decided to bike to office.
It takes about 30 min.

It is almost impossible
to predict from this obs!

9/11/14 morning torrential rain



Summary and a perspective

"*Big Data*" ⇔ "*Big Simulation*"

Exa-scale computing in the next decade

Use dense, error-correlated data effectively

With the Post-K, we aim to run 1000-member global NICAM-LETKF at 3.5-km resolution

➔ Revolutionize Weather Prediction