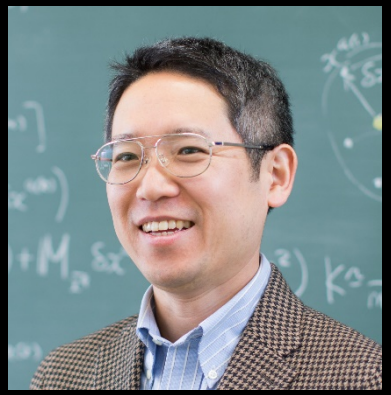
Big Data Assimilation

A New Science for Weather Prediction and Beyond



Takemasa Miyoshi

Ph.D. (Meteorology)
Data Assimilation Scientist

Data Assimilation Research Team

RIKEN

Who am I?

http://data-assimilation.riken.jp/~miyoshi/

B.S. from Kyoto U
↓
JMA administration (2y)
↓
JMA NWP (1.25y)
↓
UMD (2y, M.S. and Ph.D.)
↓
JMA NWP (3.5y)
↓
UMD (4y)
↓
RIKEN (6y+)

http://tedxsannomiya.com/en/speakers/takemasa-miyoshi/

Takemasa Miyoshi, Ph.D.
Team Leader
Data Assimilation Research Team
RIKEN Center for Computational Science

Deputy Director
RIKEN interdisciplinary Theoretical and Mathematical Sciences (iTHEMS) Program

Chief Scientist
Prediction Science Laboratory
RIKEN Cluster for Pioneering Research

Visiting Professor
University of Maryland, College Park

Affiliate Professor
Graduate School of Science, Kyoto University

Visiting Principal Scientist
Application Laboratory, JAMSTEC

Research Counselor
Servicio Meteorológico Nacional (National Meteorological Service), Argentina

Education

- **2005** Ph.D. in Meteorology, University of Maryland, College Park, Maryland, USA (Dissertation PDF)
- **2004** M.S. in Meteorology, University of Maryland, College Park, Maryland, USA (Scholarly Paper PDF)
- **2000** B.S. in Physics, Faculty of Science, Kyoto University, Kyoto, Japan

RIKEN Center for Computational Science (R-CCS)

Japan's flagship institute for computational science

Missions:

1) Development & operation of the Japanese flagship supercomputer

2) Center of Excellence for research on computational science

RIKEN Center for Computational Science (R-CCS)

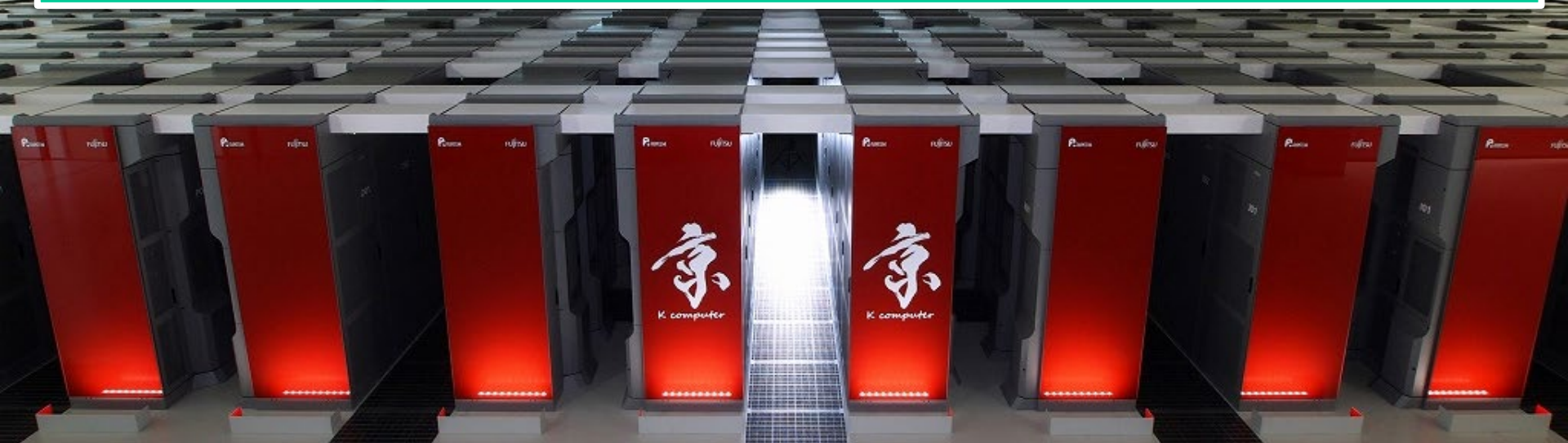
Japan's flagship institute for computational science

Missions:

1) Development & operation of the Japanese flagship supercomputer

2) Center of Excellence for research on computational science

New "Fugaku" or "富岳" is being developed



RIKEN R-CCS

Data Assimilation Team

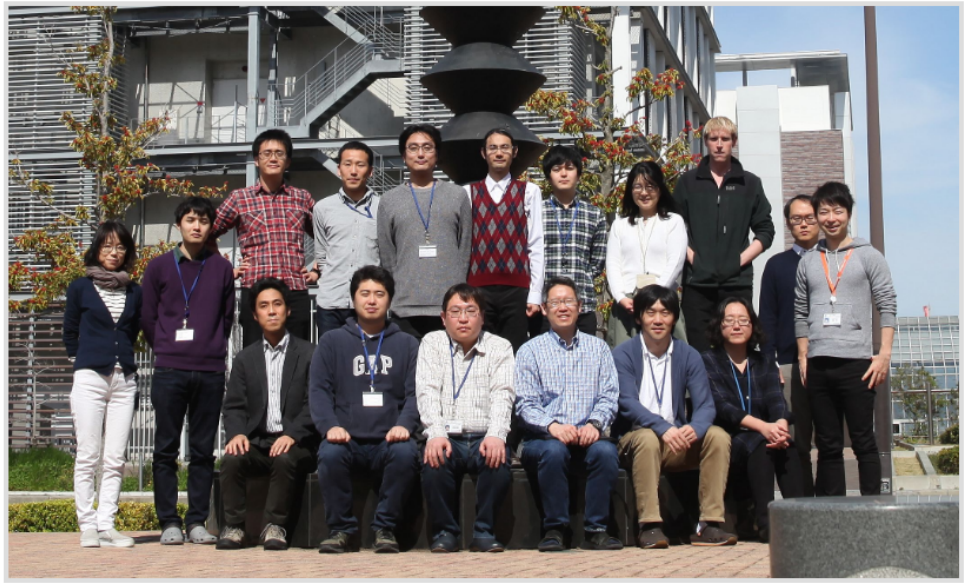
http://www.data-assimilation.riken.jp/

Data Assimilation Research Team

Data Assimilation Research Team was launched in October, 2012, in RIKEN Advanced Institute for Computational Science (AICS), conveniently located in the beautiful and historic city of Kobe. RIKEN is known as the flagship research institution in Japan. On April 1, 2018, RIKEN AICS was renamed RIKEN Center for Computational Science (R-CCS). R-CCS is operating the world's leading K computer, and also has a strong Research Division. R-CCS takes the lead in advancing the computational science and aims to be an international center of excellence for computational science in collaboration with a wide range of research organizations. R-CCS integrates the computer science and computational science to conduct most advanced research and development of a wide range of applied scientific computation, as well as of high performance computing technologies.

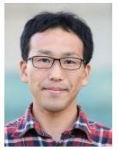
Data assimilation is a cross-disciplinary science to synergize numerical simulations and observational data, using statistical methods and applied mathematics. As computers become more powerful and enable more precise simulations, it will become more important to compare the simulation with actual observations.

Data Assimilation Research Team ("DA team") performs cutting-edge research and development on advanced data assimilation methodsand their wide applications, aiming at integrating computer simulations and observational data in the wisest way. Particularly, the DA team will tackle challenging problems of developing efficient and accurate data assimilation systems for high-dimensional simulations with large amount of data. The specific areas include 1) research on parallel-efficient algorithms for data assimilation with the super-parallel K computer, 2) research on data assimilation methods and applications by taking advantage of the world-leading K computer, and 3) development of most advanced data assimilation software optimized for the K computer.



April 5, 2019 at R-CCS

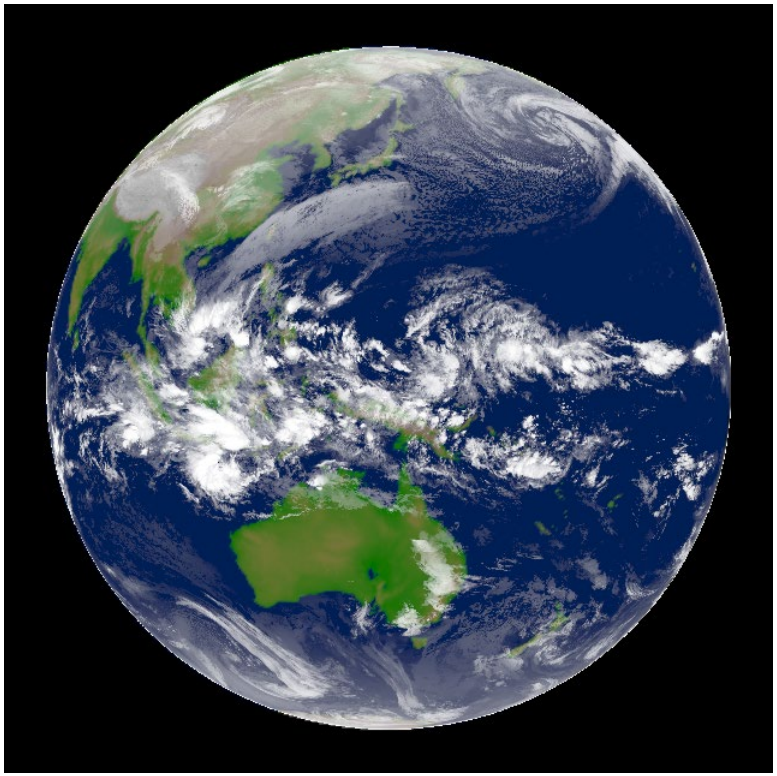
Team Leader	Research Scientist	Research Scientist	Research Scientist (Excellent Young Researcher)
			
Takemasa Miyoshi	Koji Terasaki	Shigenori Otsuka	Shunji Kotsuki

Postdoctoral Researcher	Postdoctoral Researcher	Postdoctoral Researcher	Postdoctoral Researcher	Postdoctoral Researcher
				
Takumi Honda	Kohei Takatama	James Taylor	Arata Amemiya	Hironori Arai

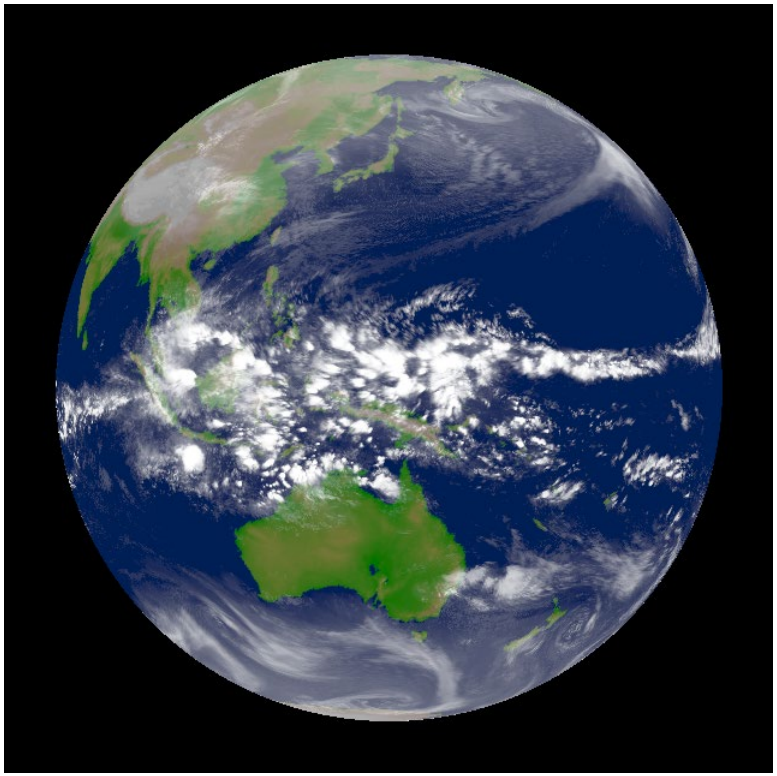
Research Associate	Technical Staff	Technical Staff	Technical Staff	Technical Staff
				
Yasumitsu Maejima	Hazuki Arakida	Hideyuki Sakamoto	Marimo Ohhigashi	Kenta Kurosawa

Which is real, or a simulation?

A

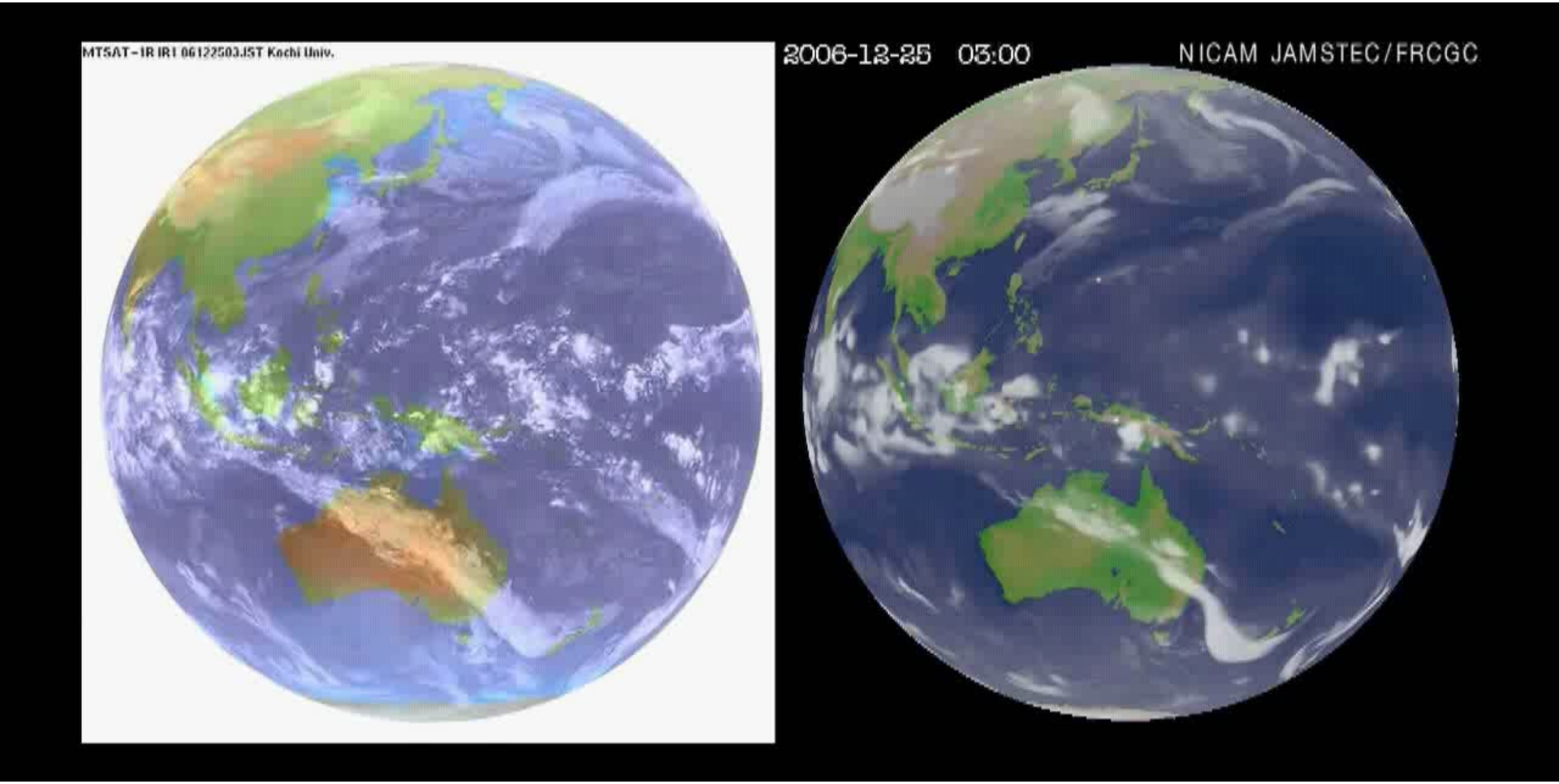


B



Courtesy of H. Miura

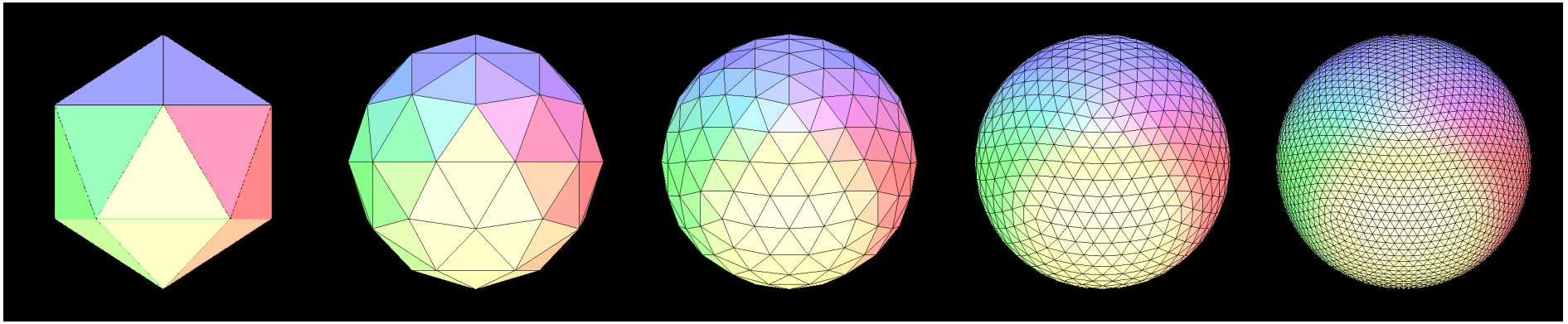
Simulation of global weather

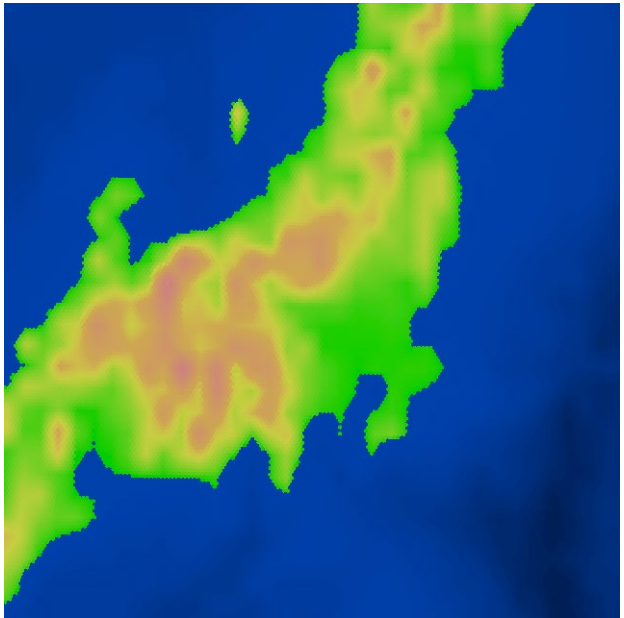


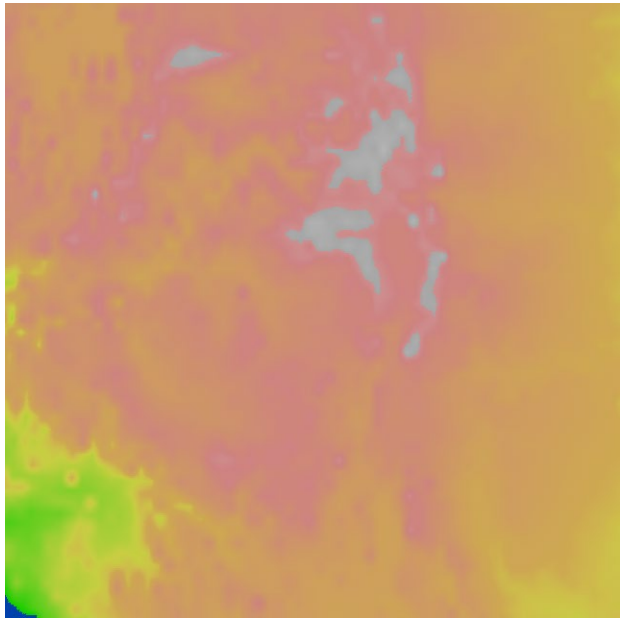
Method for weather simulation

1. Discretize the atmosphere

(courtesy of H. Miura)



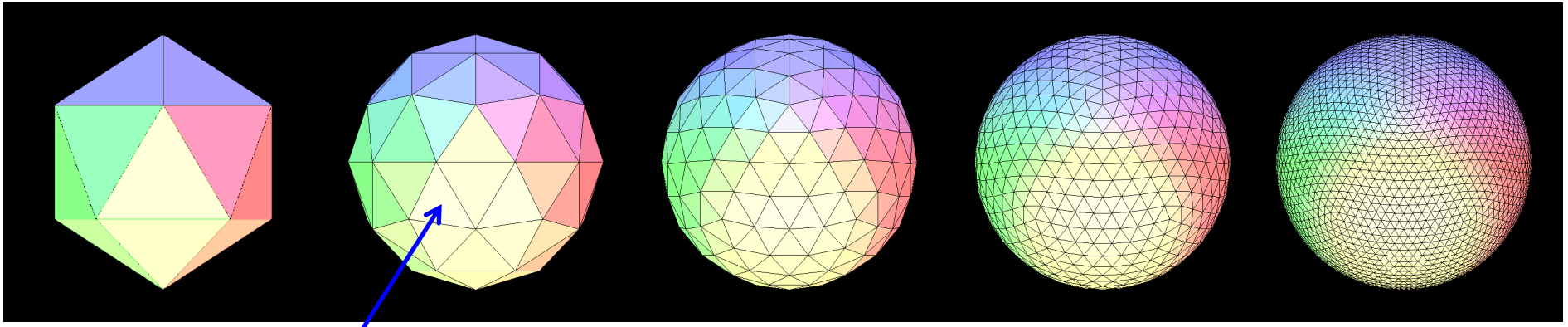




dx~480 km
dx~240 km
dx~120 km
dx~60 km
dx~30 km
dx~15 km
dx~7.5 km
dx~3.75 km

Method for weather simulation

1. Discretize the atmosphere

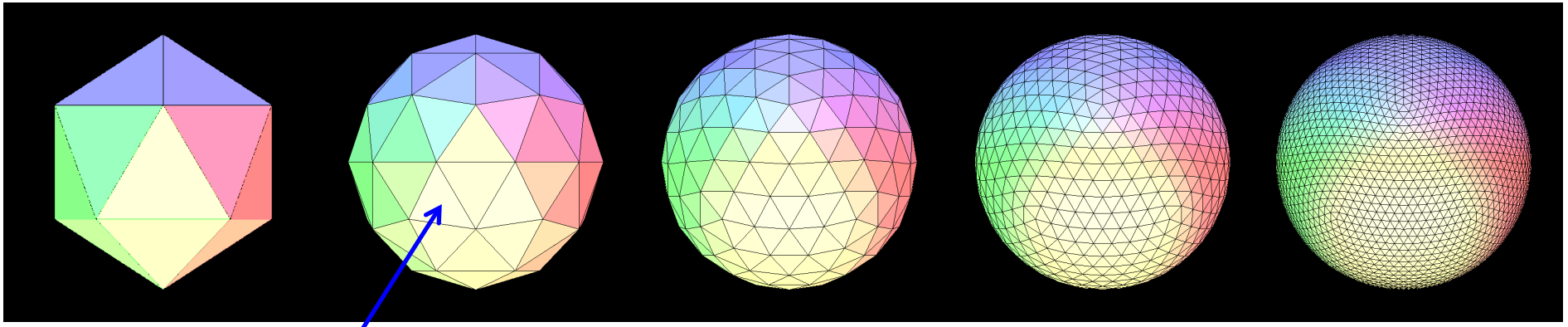


Each grid box has the meteorological variables:

winds, temperature, humidity, clouds, pressure

Method for weather simulation

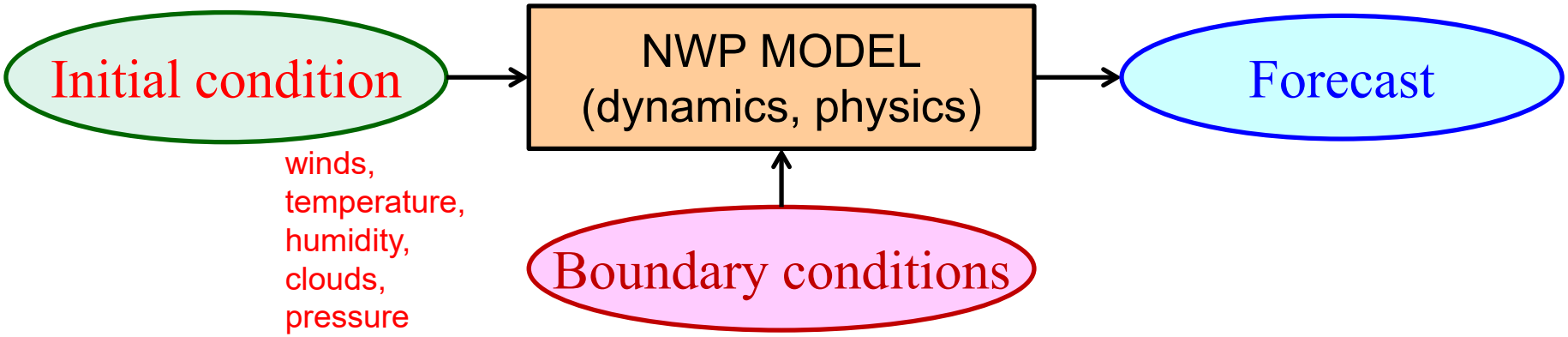
1. Discretize the atmosphere



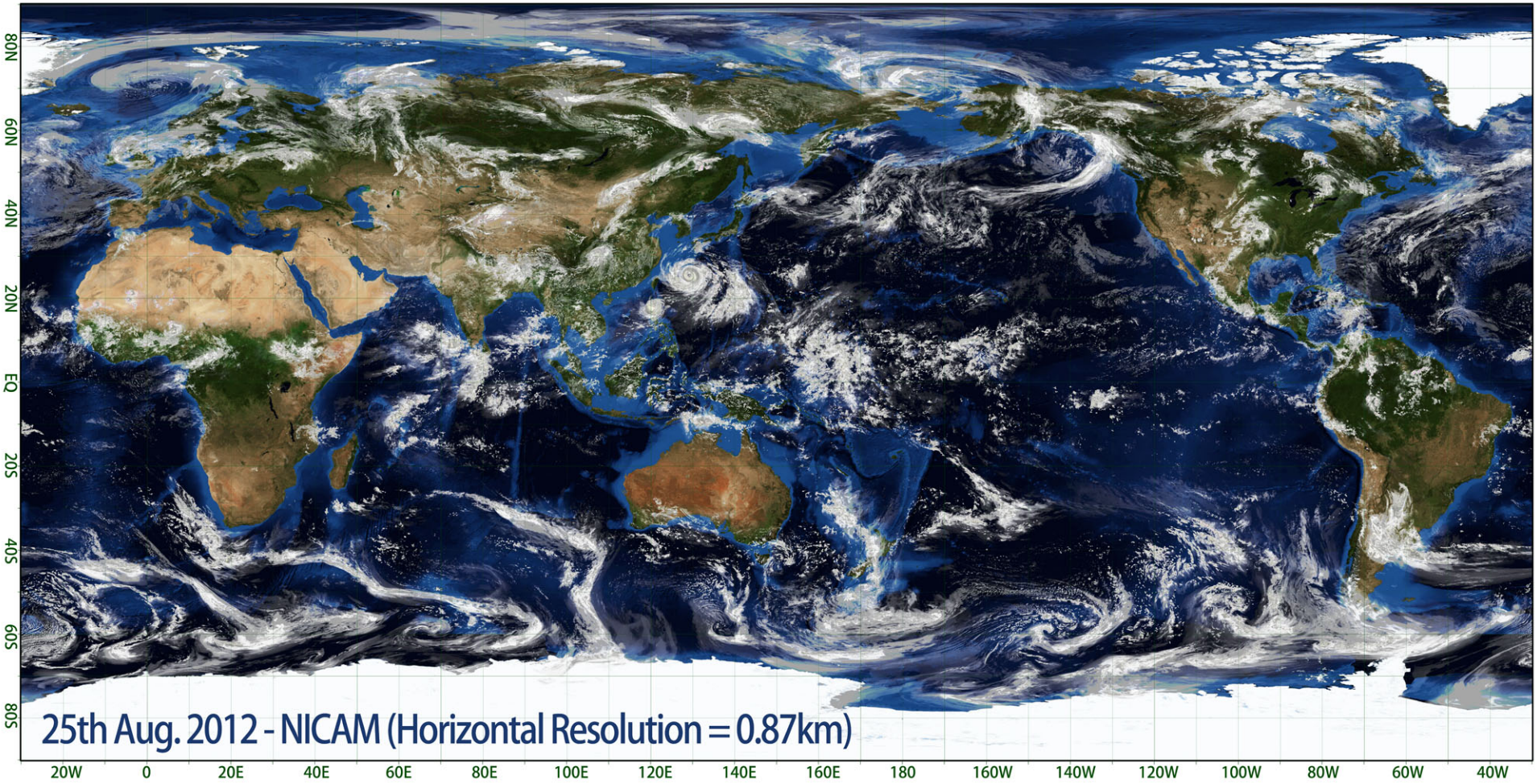
Each grid box has the meteorological variables:

winds, temperature, humidity, clouds, pressure

2. Solve the equations of atmospheric dynamics and physics

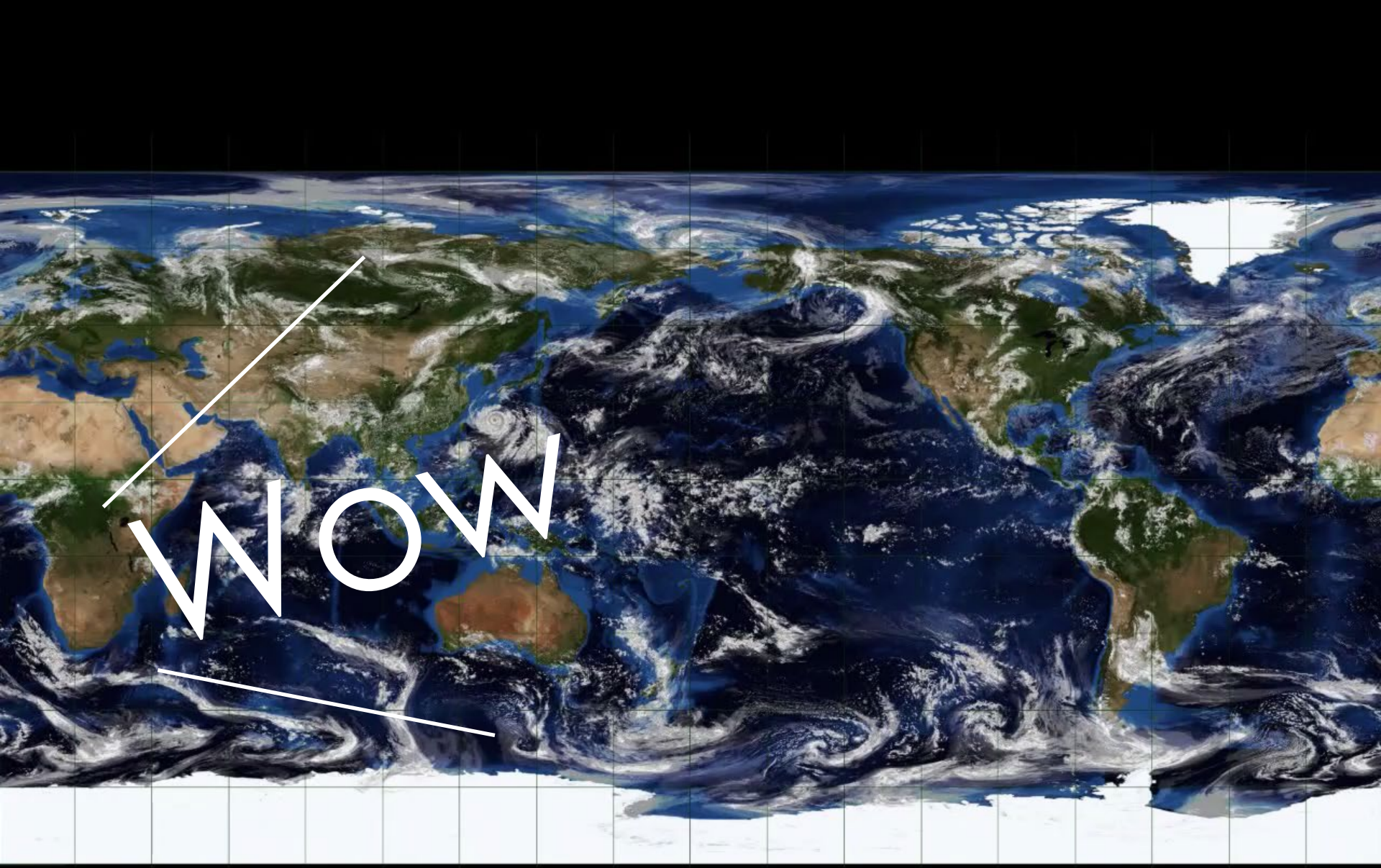


Global 870-m simulation (*Miyamoto et al. 2013*)



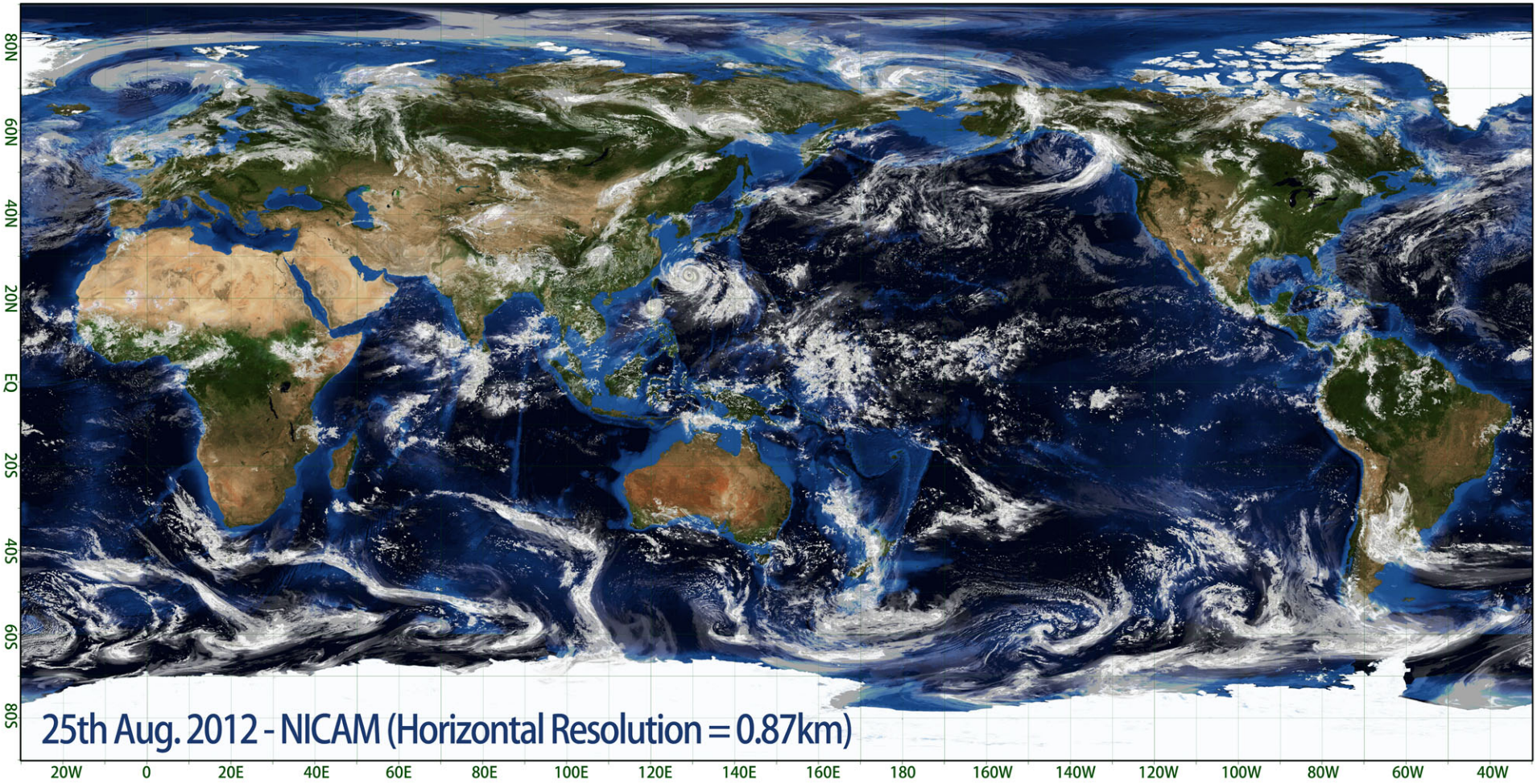
Visualized by Ryuji Yoshida


京
K computer
京
K computer
©RIKEN



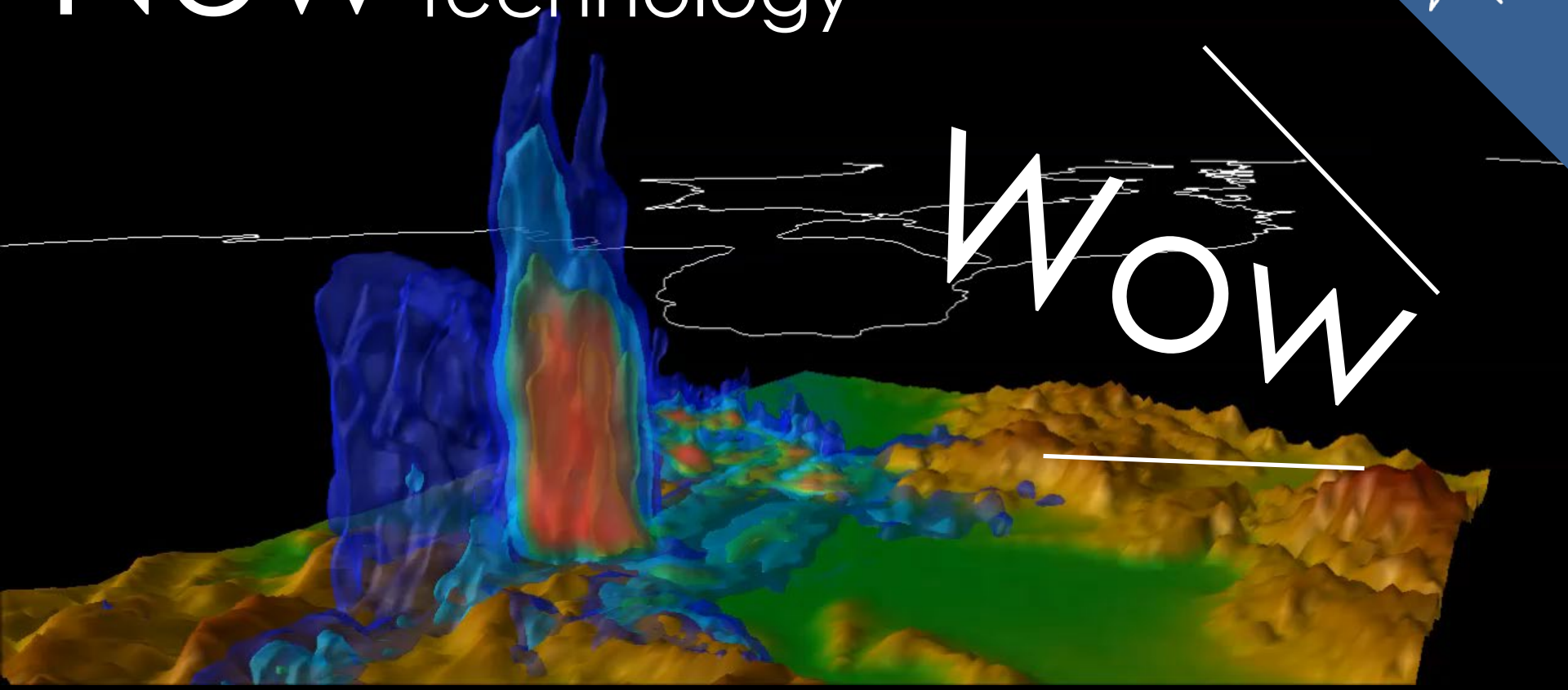
cf. TEDxSannomiya
http://tedxsannomiya.com/speakers/takemasa-miyoshi/

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

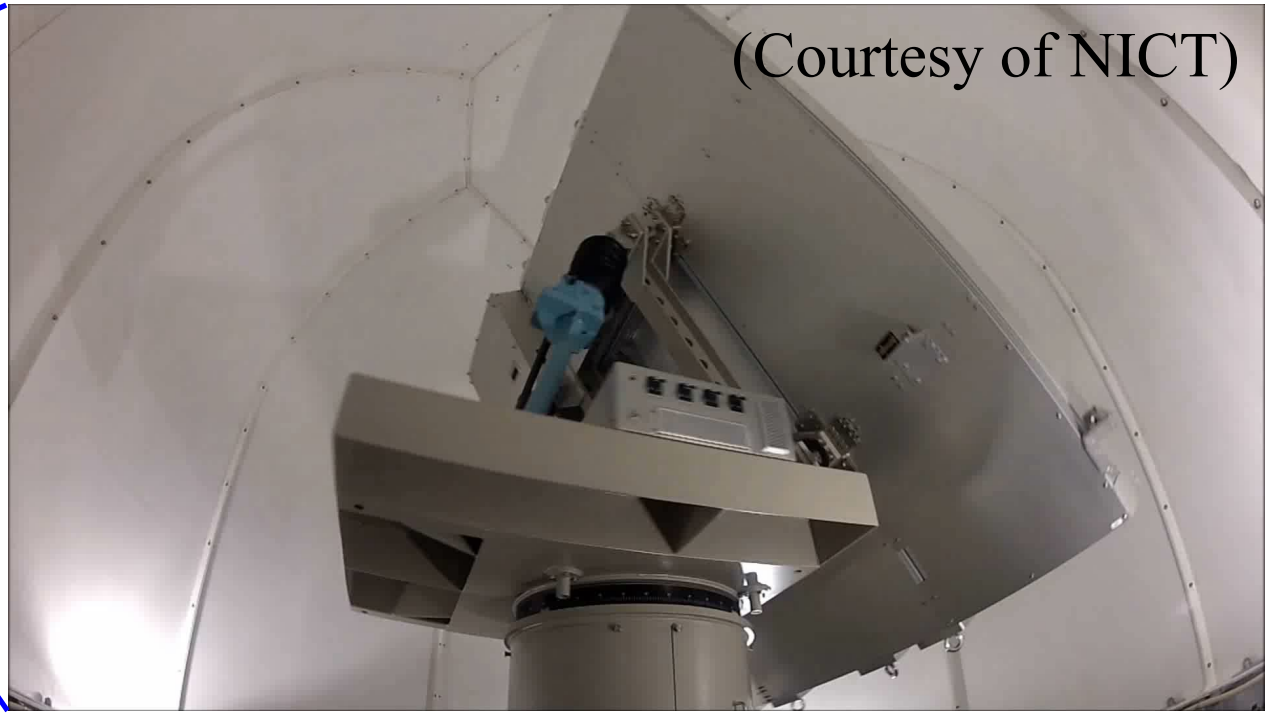


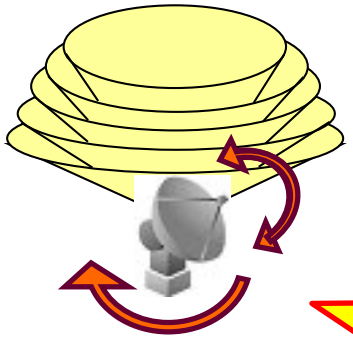
Visualized by Ryuji Yoshida

New radar technology



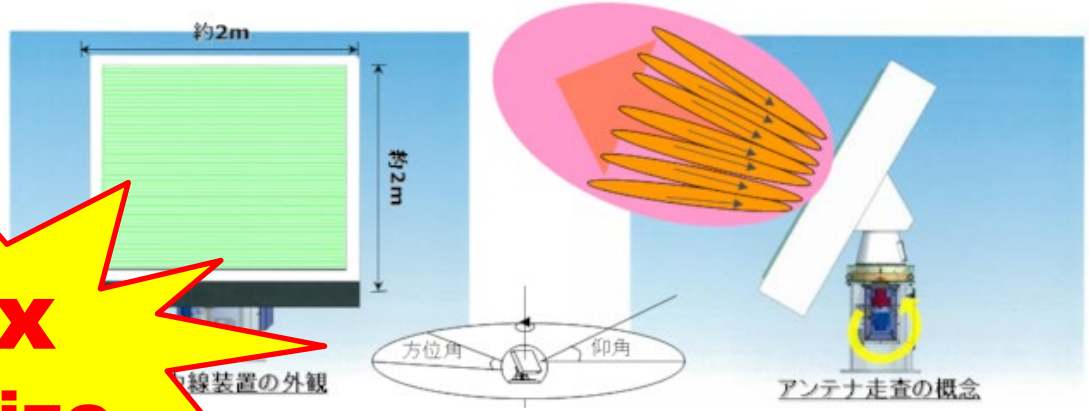
Phased Array Weather Radar (PAWR)





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

100x data size

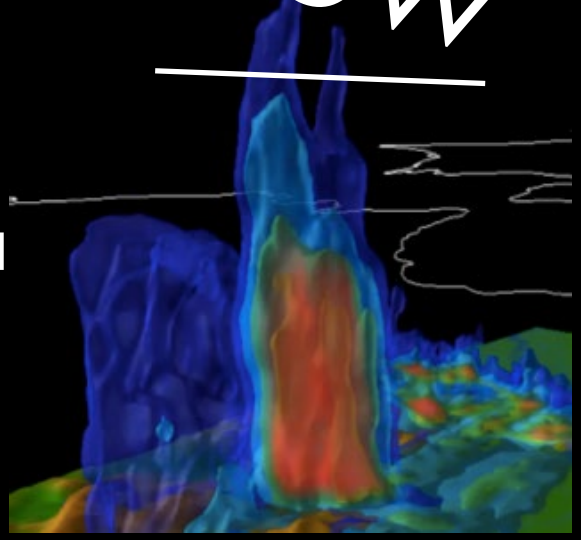


3-dim measurement using a phased array antenna
(100 m, 100 EL angles in 30 sec)

Wow

+

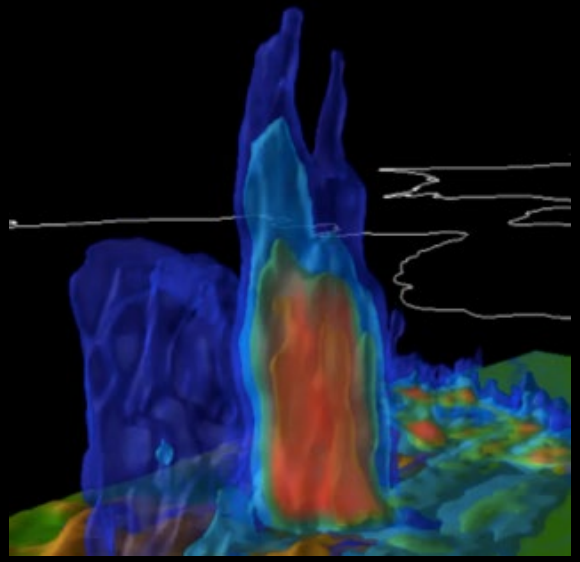
Wow



=

?

Data Assimilation



= ~~Sudden~~ heavy rain

Only in 10 minutes!!

Kobe newspaper (https://www.kobe-np.co.jp/news/kobe/201807/0011486822.shtml)

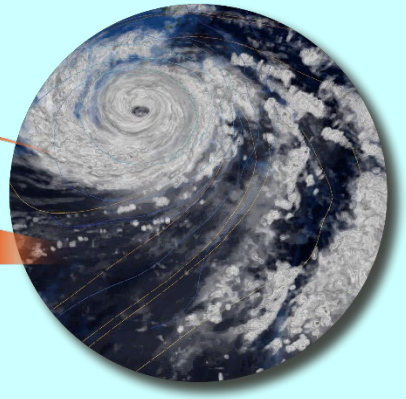
1.34 m in 10 minutes!!

5 people died in River Toga in Kobe on July 28, 2008

Big Data Assimilation

Observations

Simulations



Data Assimilation

Big Data

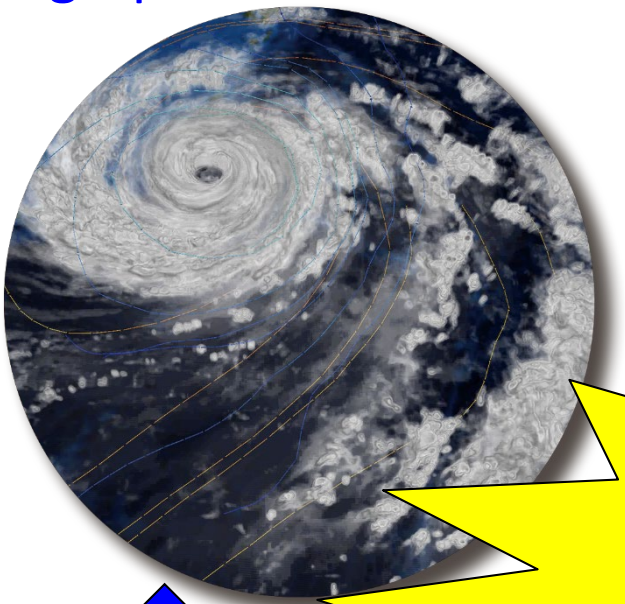
New sensors, IoT

Big Data

Powerful supercomputer

Pioneering “Big Data Assimilation” Era

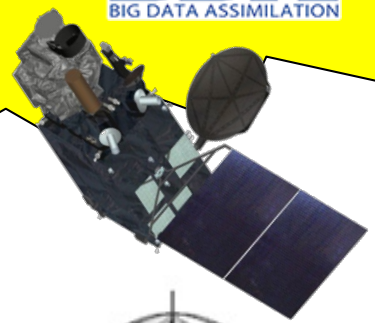
High-precision Simulations



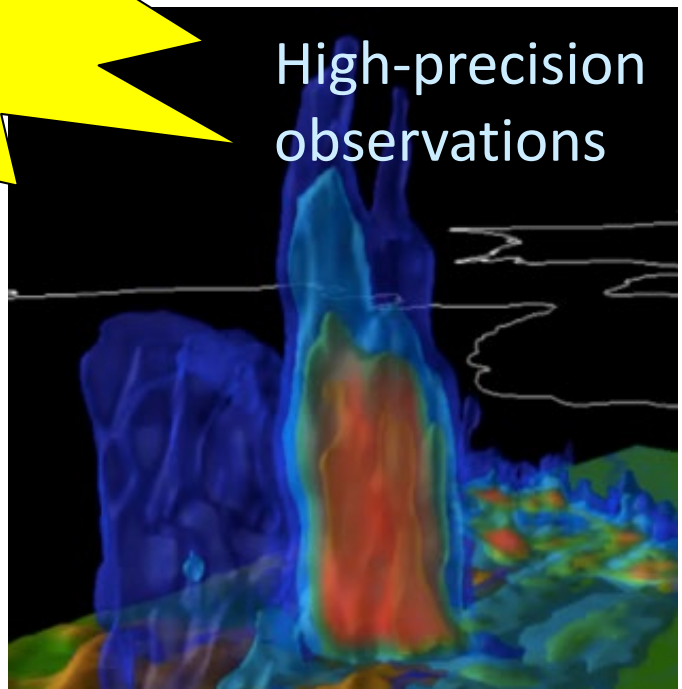
国立研究開発法人 科学技術振興機構 Japan Science and Technology Agency CREST

Future-generation technologies available 10 years in advance

BDA BIG DATA ASSIMILATION

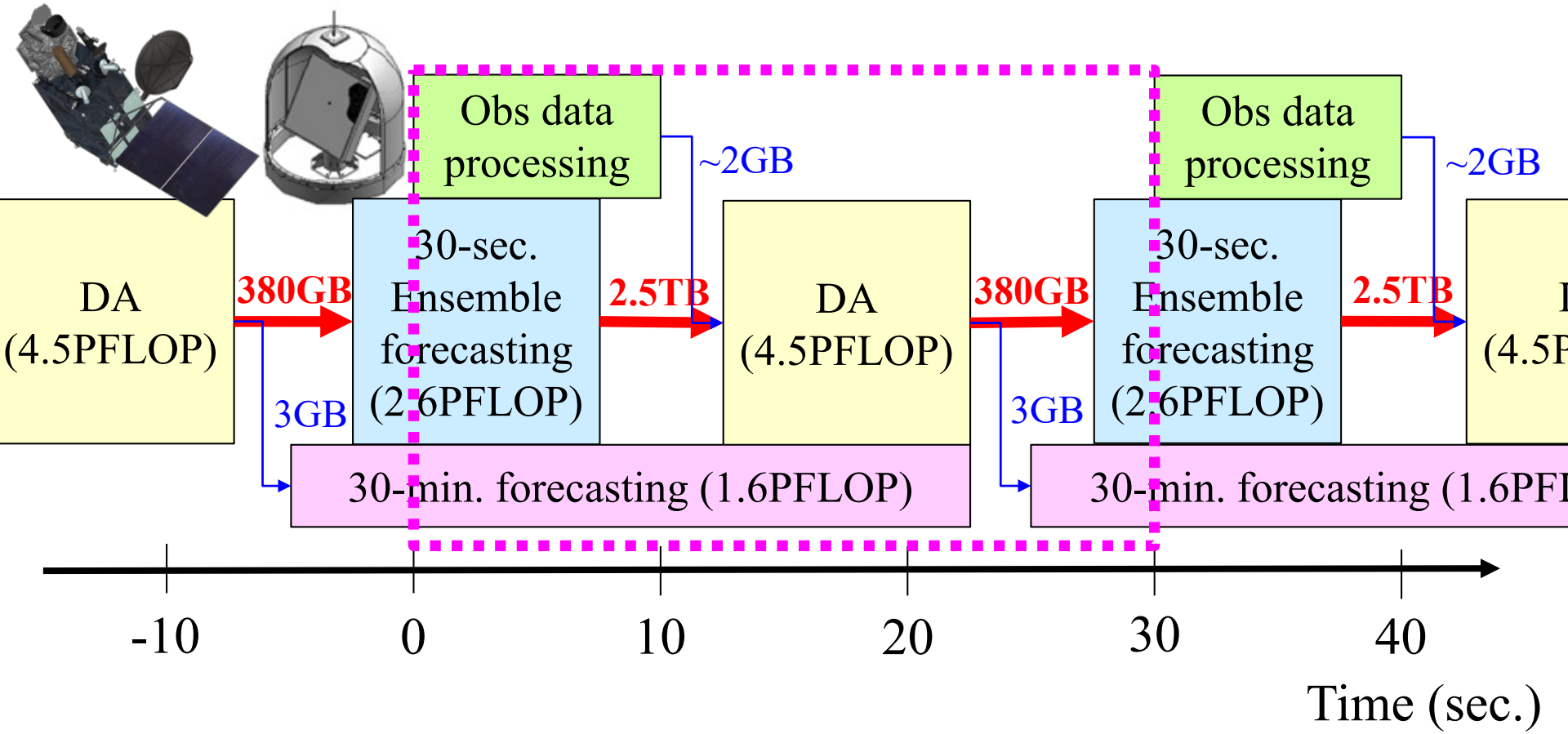


High-precision observations



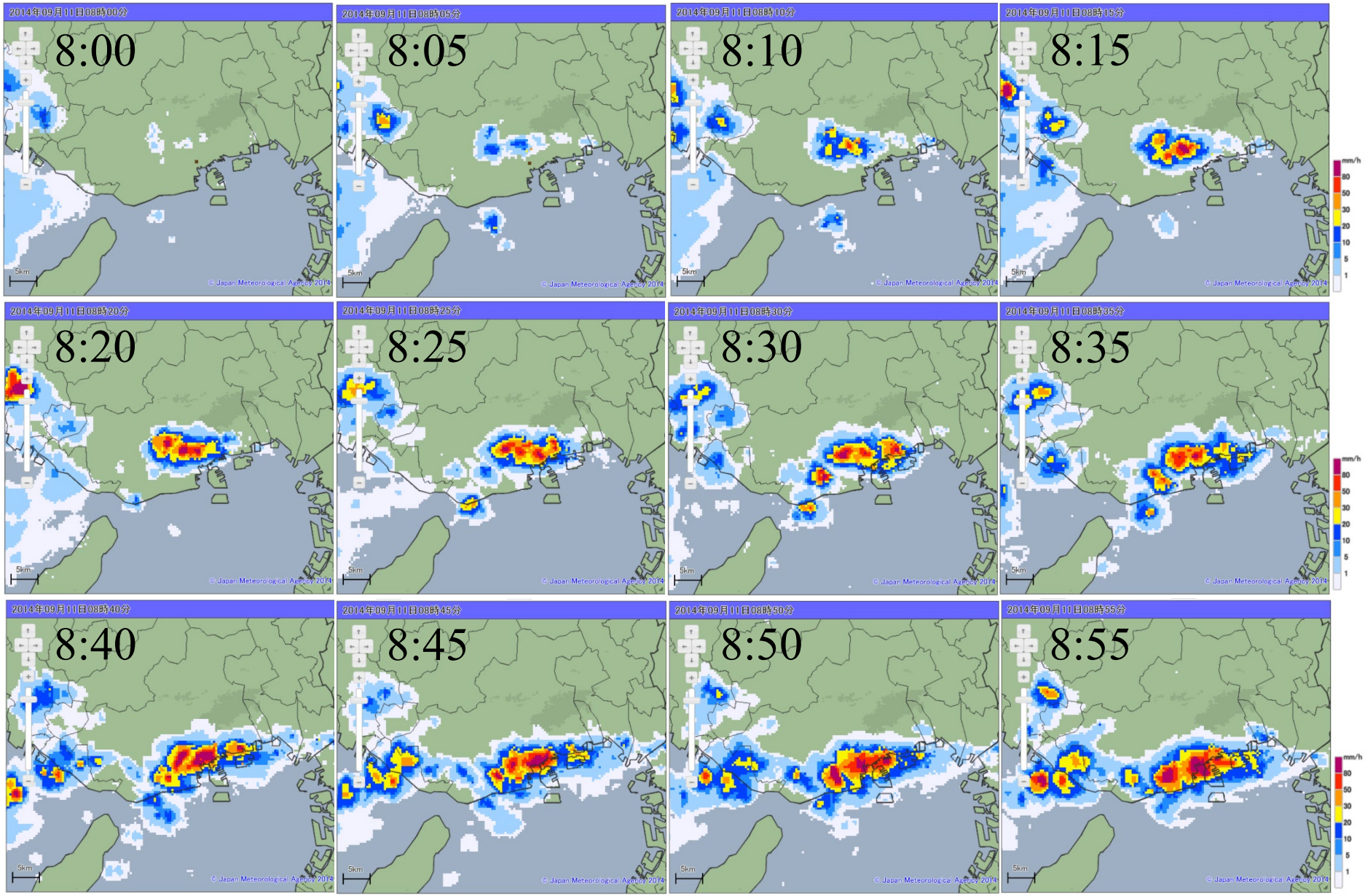
Mutual feedback

Revolutionary super-rapid 30-sec. cycle



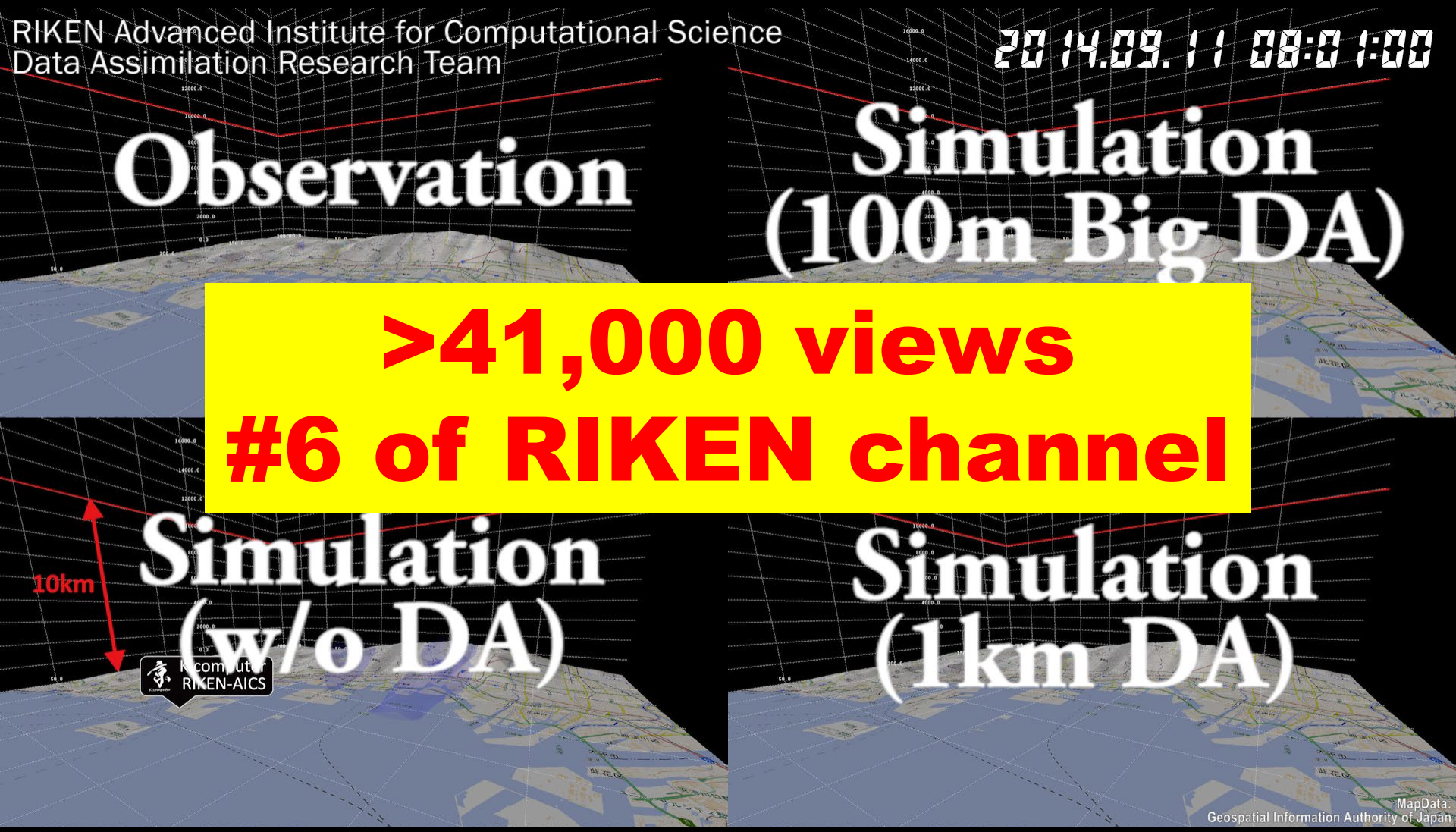
120 times more rapid than hourly update cycles

9/11/2014 morning, sudden rain

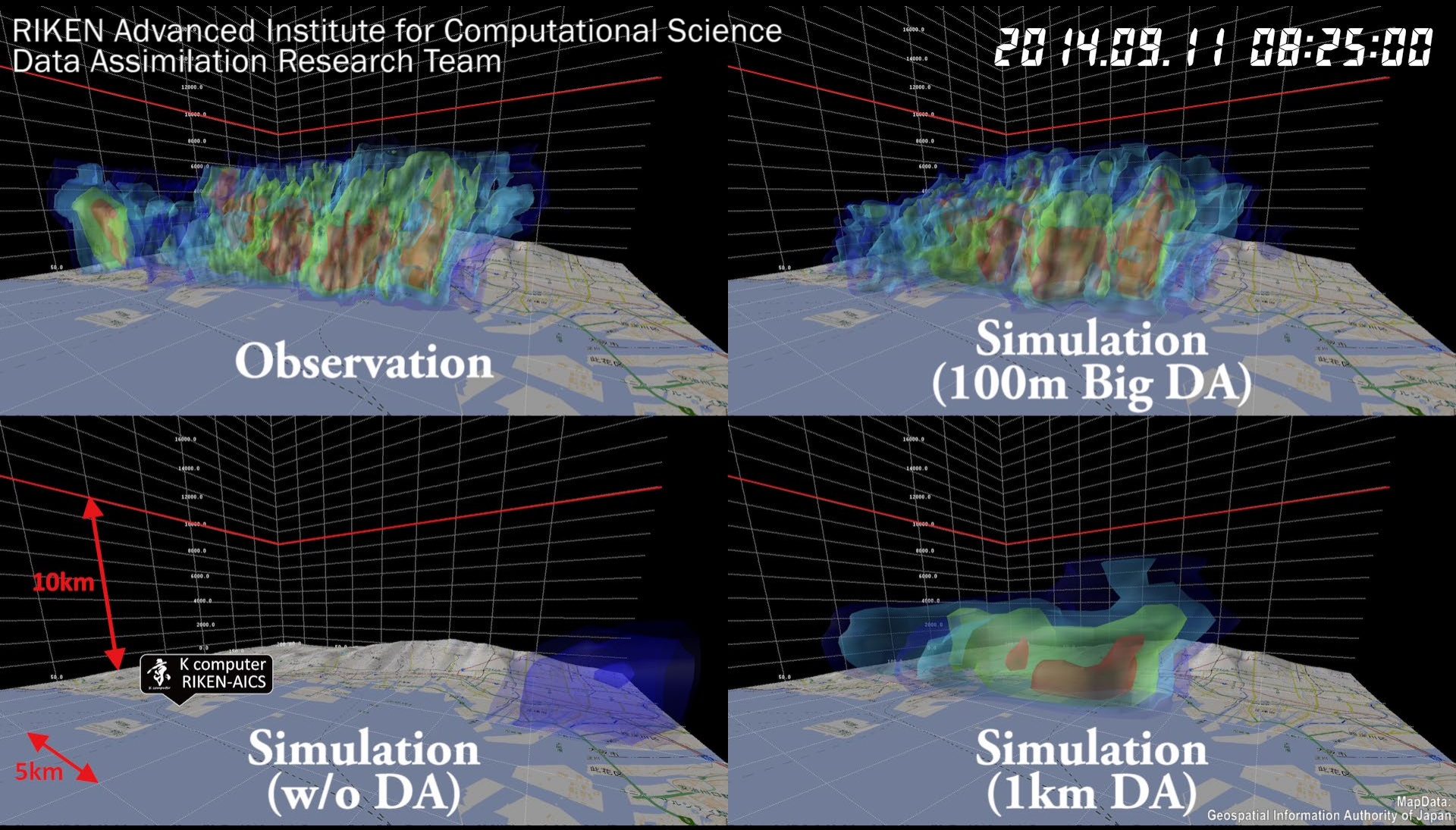


9/11/2014, sudden local rain

9/11/2014, sudden local rain



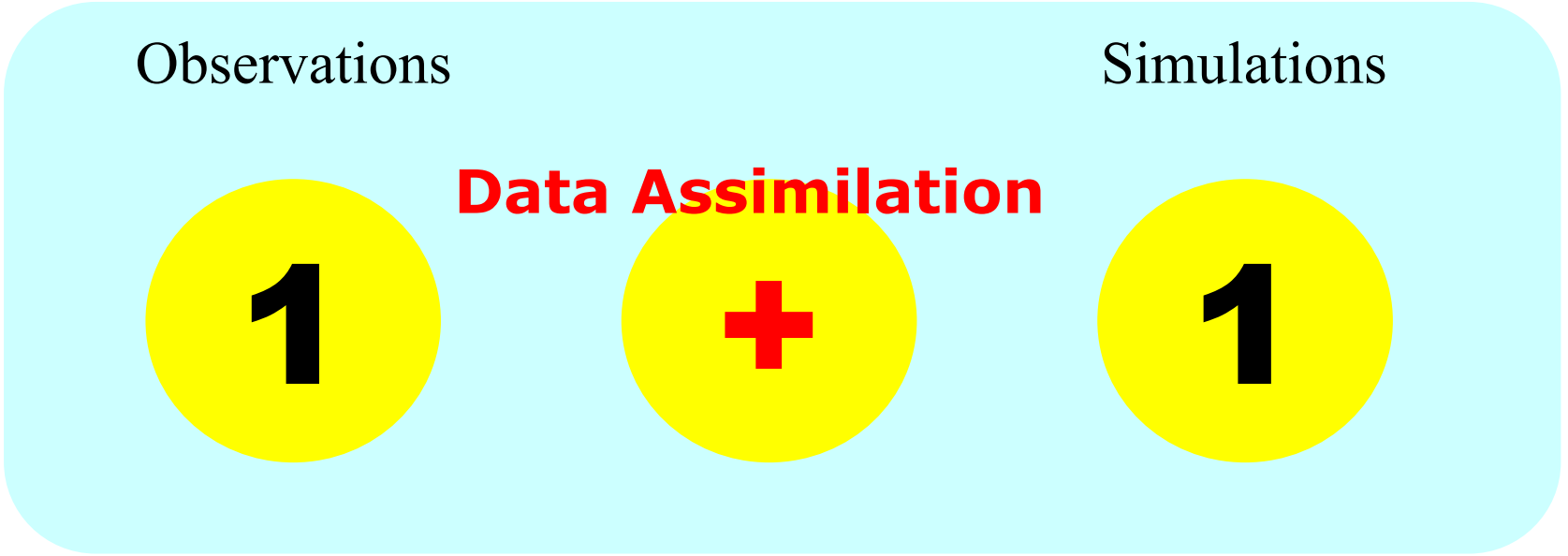
9/11/2014, sudden local rain



Data Assimilation (DA)

Data assimilation best combines observations and a model, and brings synergy.

Data Assimilation (DA)



>2

Data Assimilation (DA)

Data-driven
Induction
Real world

Process-driven
Deduction
Cyber world

Observations

Simulations

Data Assimilation

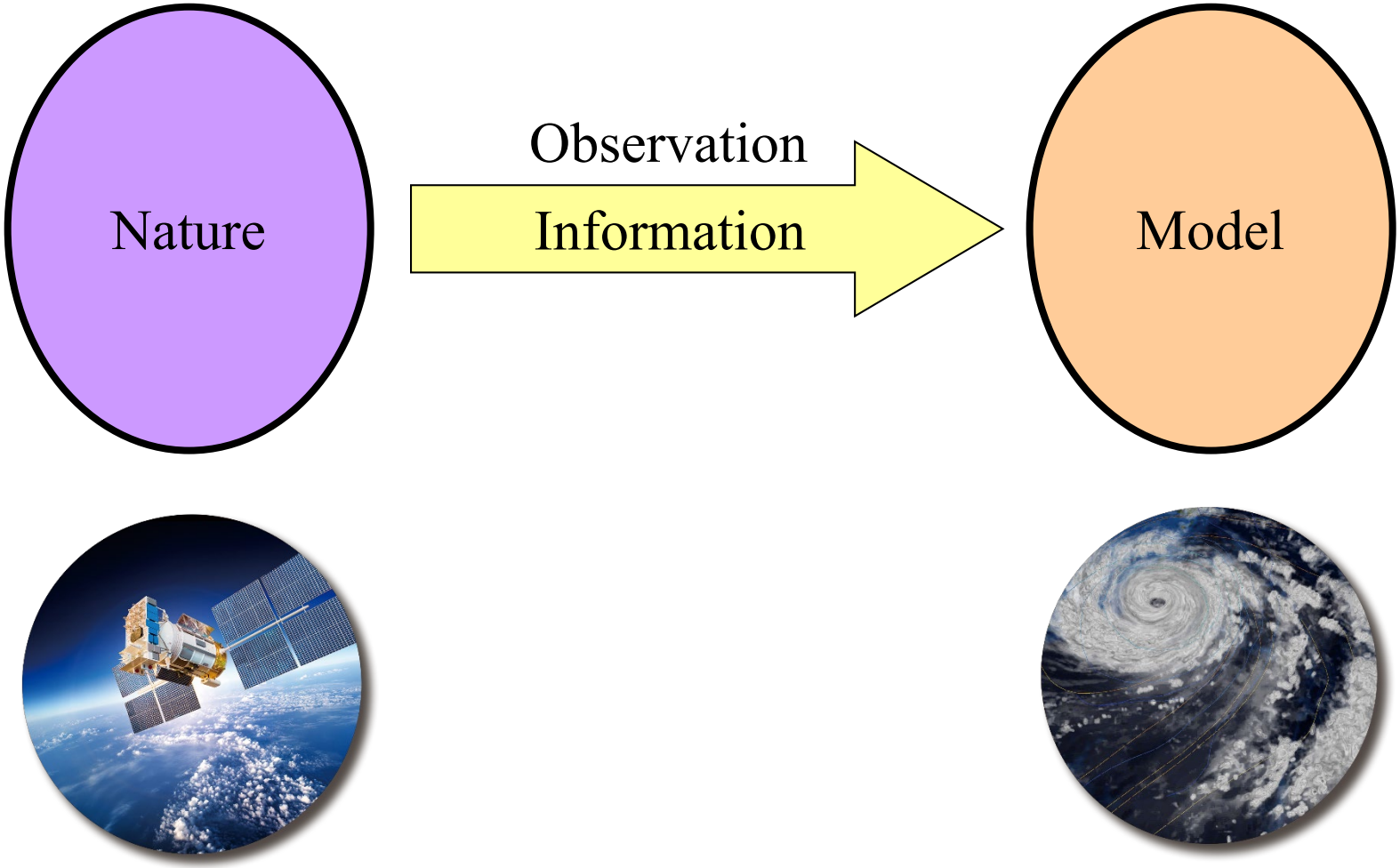
1 **+** **1**

>2

DA and Chaos Synchronization *(Yang et al. 2006)*

Master (drive) system

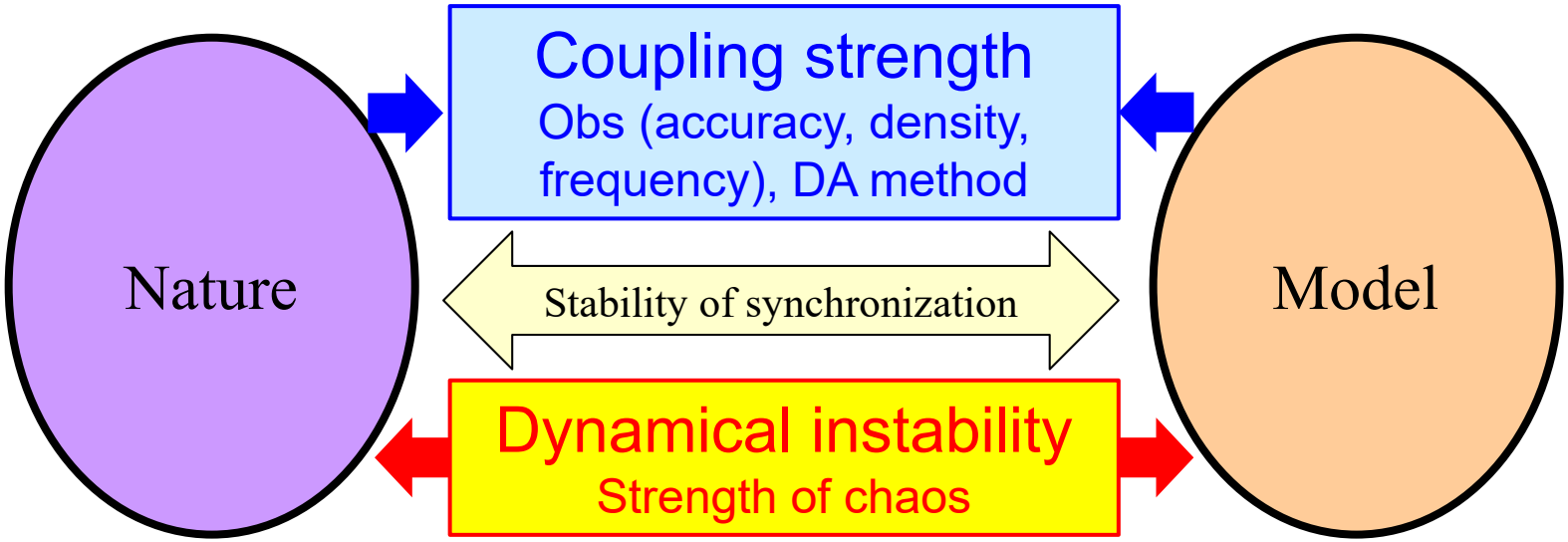
Slave (response) system

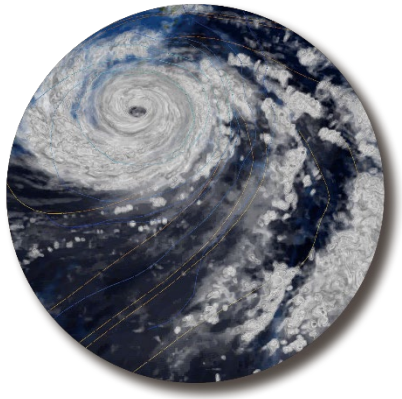


DA and Chaos Synchronization *(Yang et al. 2006)*

Master (drive) system

Slave (response) system





Logistic map

A simple system: $x_{n+1} = ax_n(1 - x_n)$

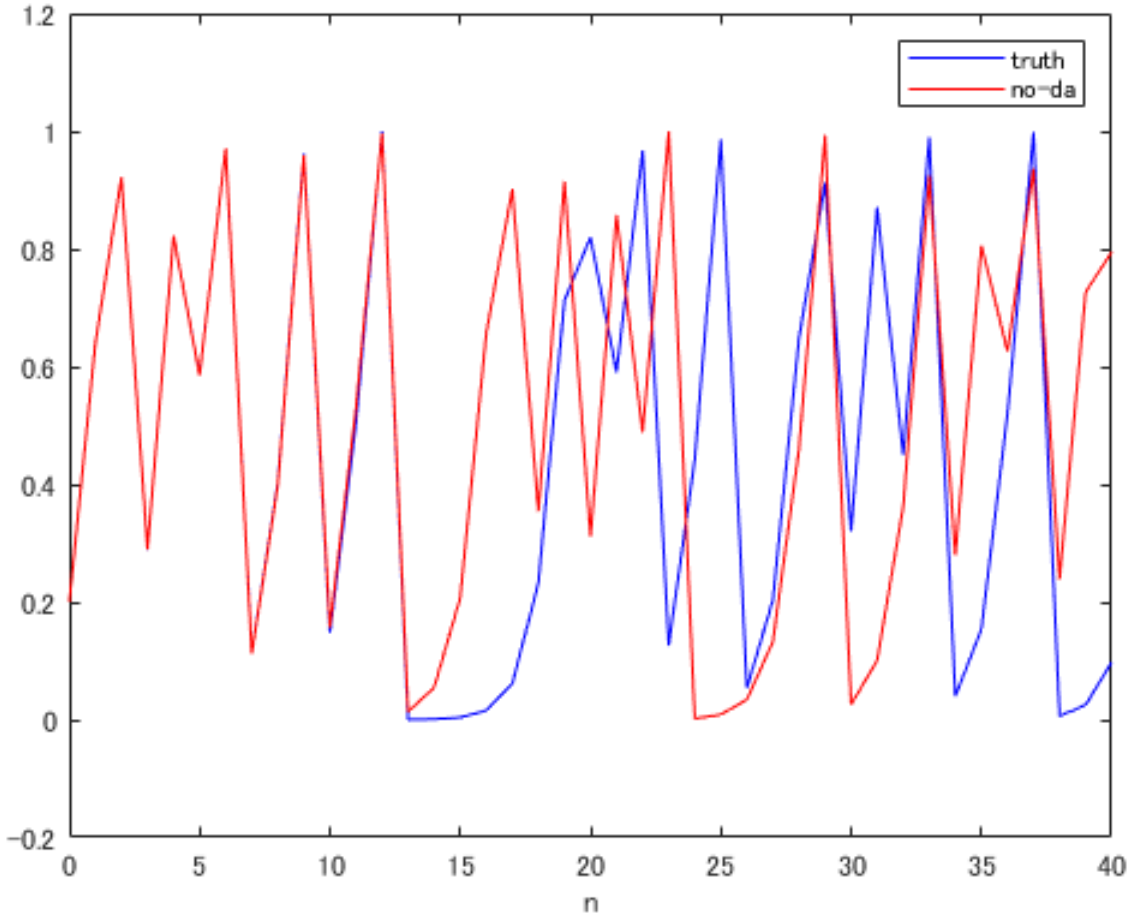
Assume $a = 4$

1	0.200000	0.200010
2	0.640000	0.640024
3	0.921600	0.921573
4	0.289014	0.289104
5	0.821939	0.822092
6	0.585421	0.585026
7	0.970813	0.971082
8	0.113339	0.112327
9	0.401974	0.398838
10	0.961563	0.959065
11	0.147837	0.157037
12	0.503924	0.529506
13	0.999938	0.996518
14	0.000246	0.013881
15	0.000985	0.054755
16	0.003936	0.207027
17	0.015682	0.656668
18	0.061745	0.901820
19	0.231730	0.354162
20	0.712124	0.914925

Logistic map

A simple system: $x_{n+1} = ax_n(1 - x_n)$

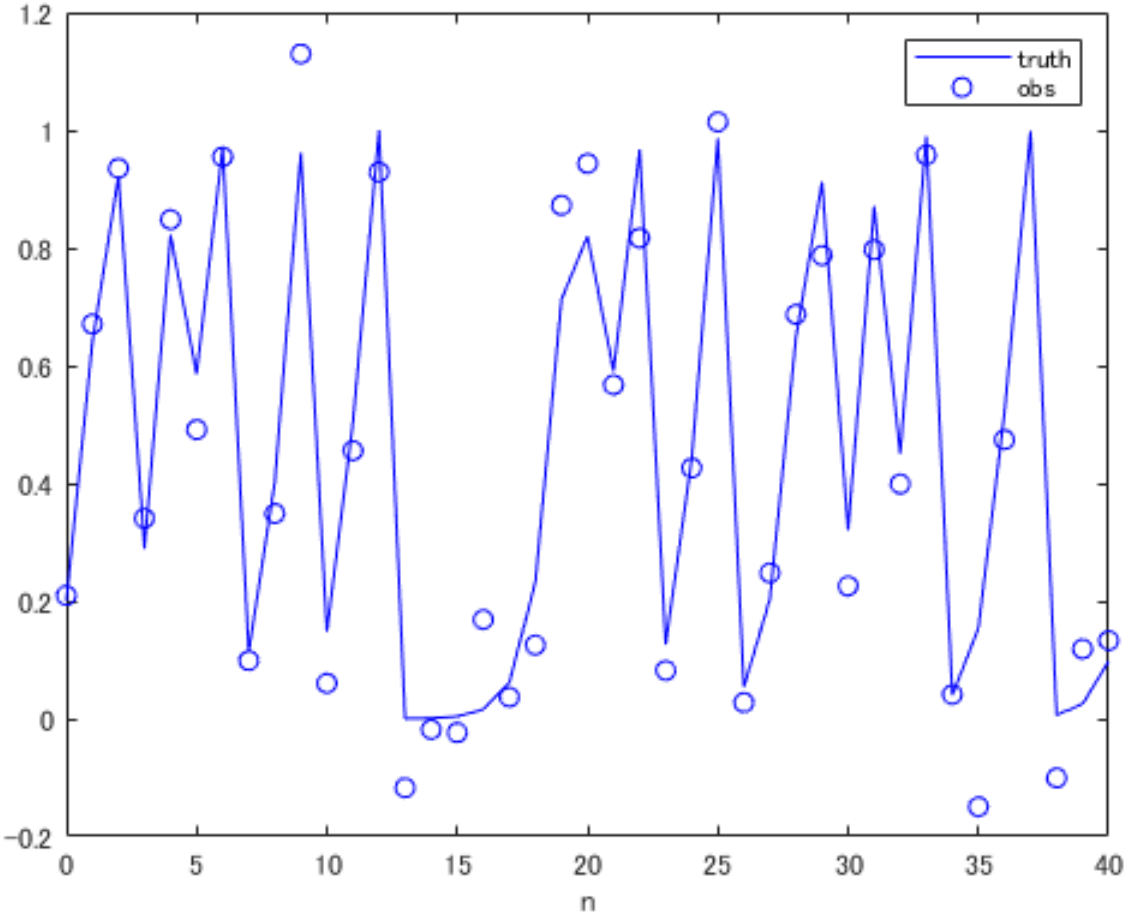
Assume $a = 4$



	truth	no-da
1	0.200000	0.200010
2	0.640000	0.640024
3	0.921600	0.921573
4	0.289014	0.289104
5	0.821939	0.822092
6	0.585421	0.585026
7	0.970813	0.971082
8	0.113339	0.112327
9	0.401974	0.398838
10	0.961563	0.959065
11	0.147837	0.157037
12	0.503924	0.529506
13	0.999938	0.996518
14	0.000246	0.013881
15	0.000985	0.054755
16	0.003936	0.207027
17	0.015682	0.656668
18	0.061745	0.901820
19	0.231730	0.354162
20	0.712124	0.914925

Logistic map

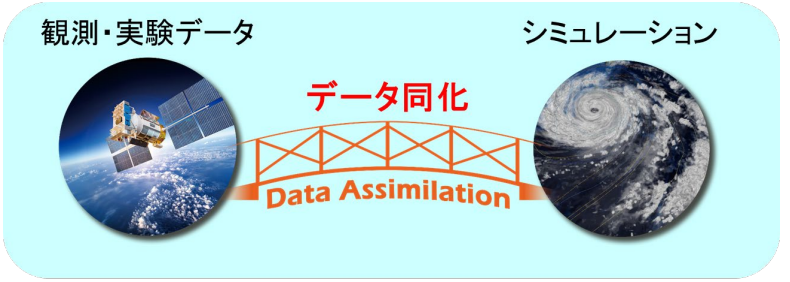
Noisy observations ○ are generated with random numbers.

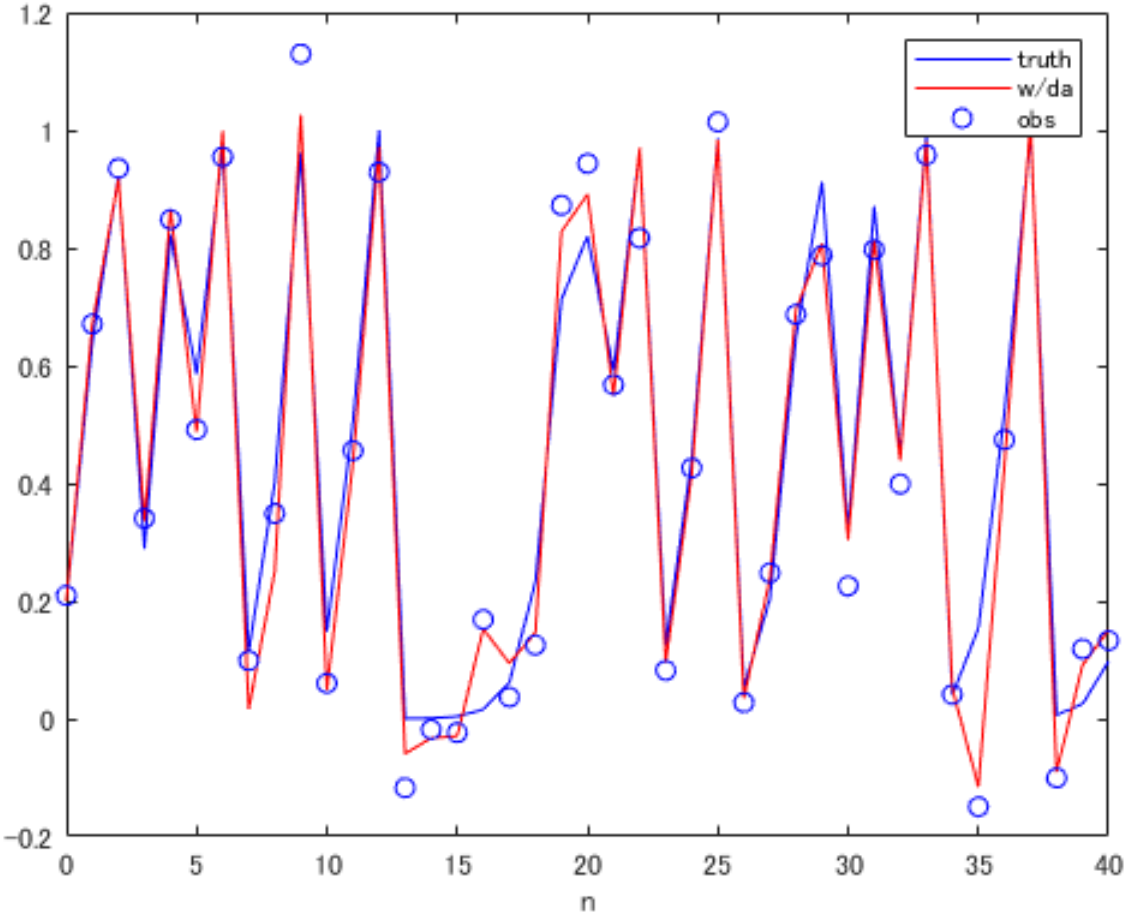


	truth	no-da
1	0.200000	0.200010
2	0.640000	0.640024
3	0.921600	0.921573
4	0.289014	0.289104
5	0.821939	0.822092
6	0.585421	0.585026
7	0.970813	0.971082
8	0.113339	0.112327
9	0.401974	0.398838
10	0.961563	0.959065
11	0.147837	0.157037
12	0.503924	0.529506
13	0.999938	0.996518
14	0.000246	0.013881
15	0.000985	0.054755
16	0.003936	0.207027
17	0.015682	0.656668
18	0.061745	0.901820
19	0.231730	0.354162
20	0.712124	0.914925

Logistic map: DA experiment

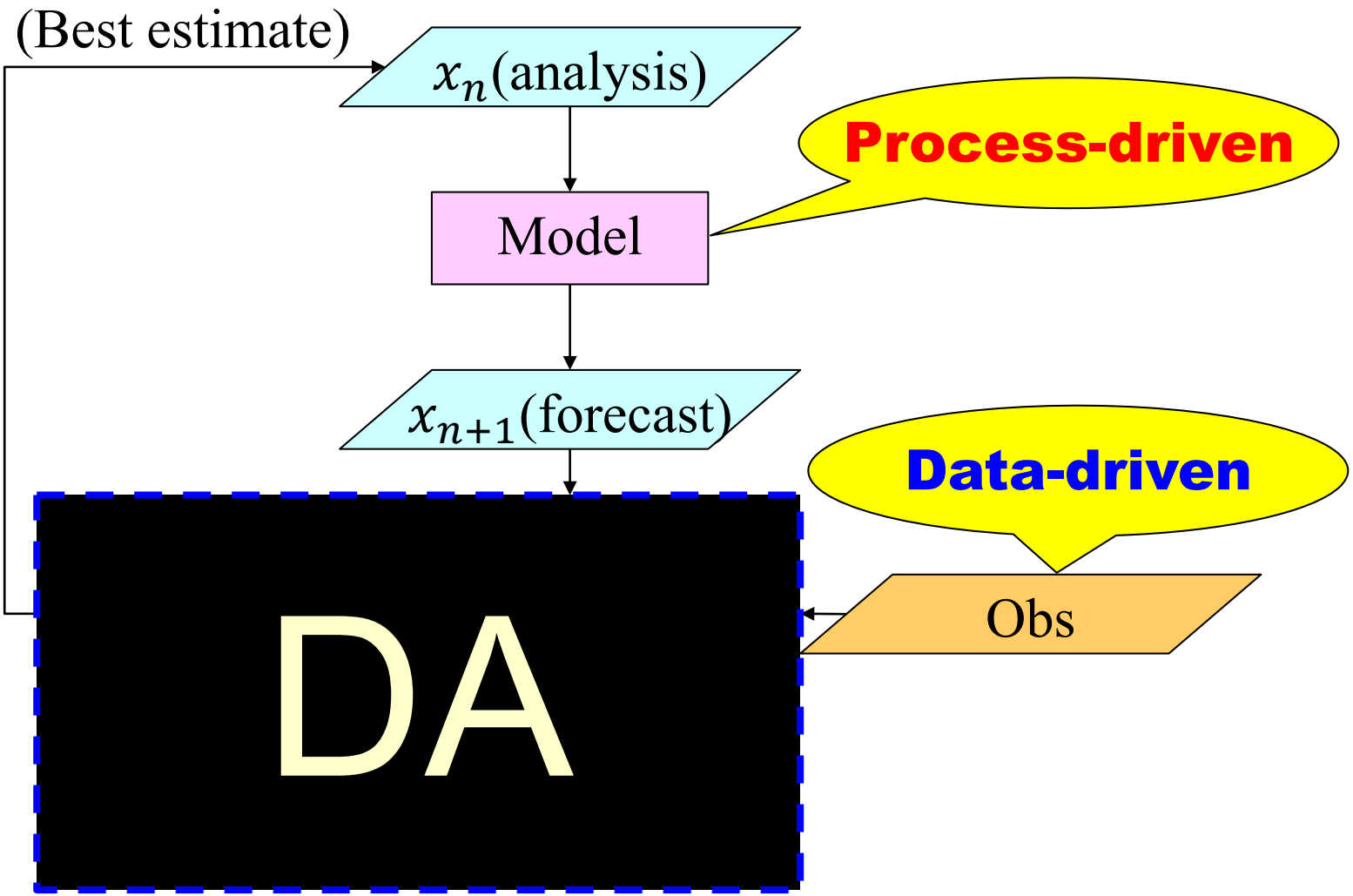
Noisy observations ○ are assimilated.



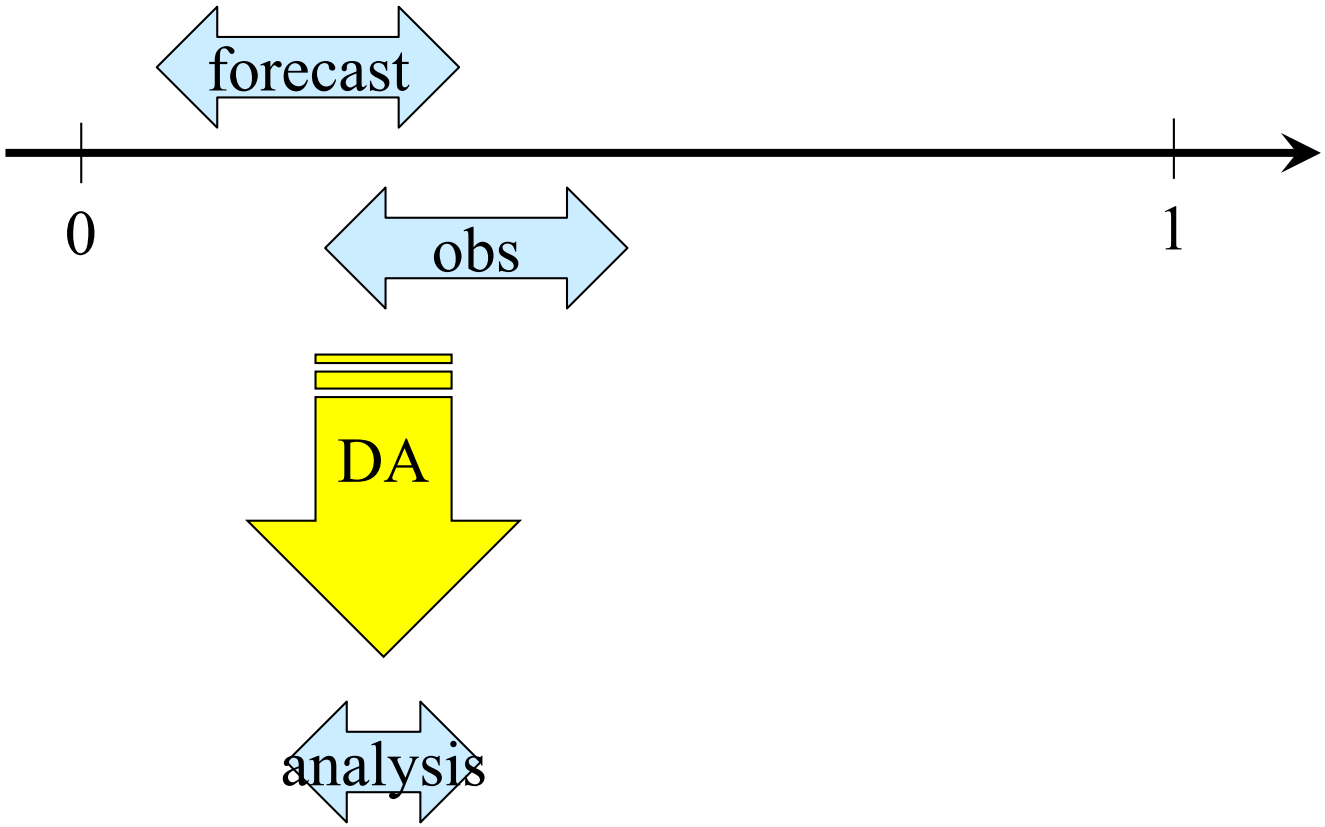


➔ Synchronized chaos!

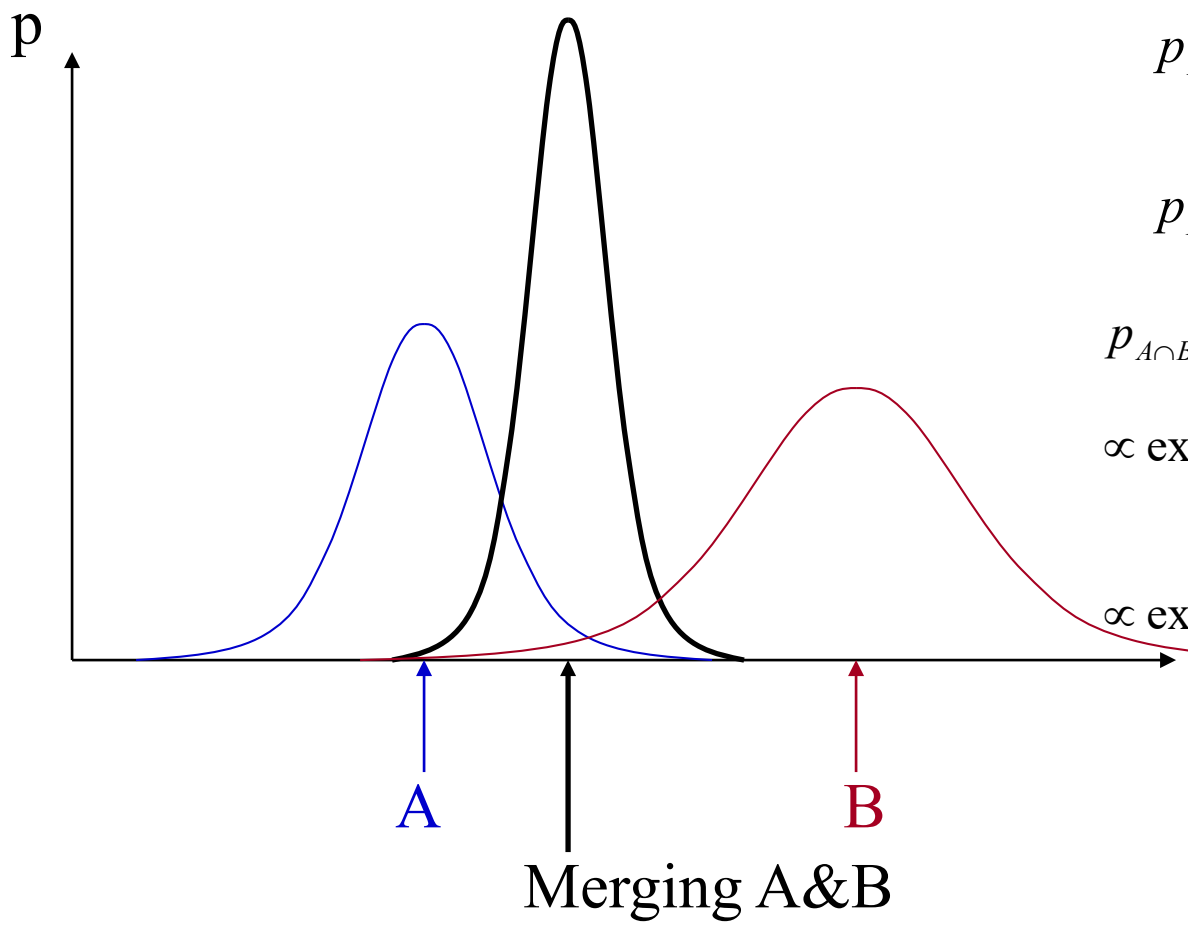
DA workflow


(Best estimate)
x_n(analysis)
Process-driven
Model
x_{n+1}(forecast)
Data-driven
DA
Obs

DA = math of errors



Merging 2 information (Bayesian estimation)



$$p_A(T) \propto \exp\left[-\frac{(T-T_A)^2}{2\sigma_A^2}\right]$$

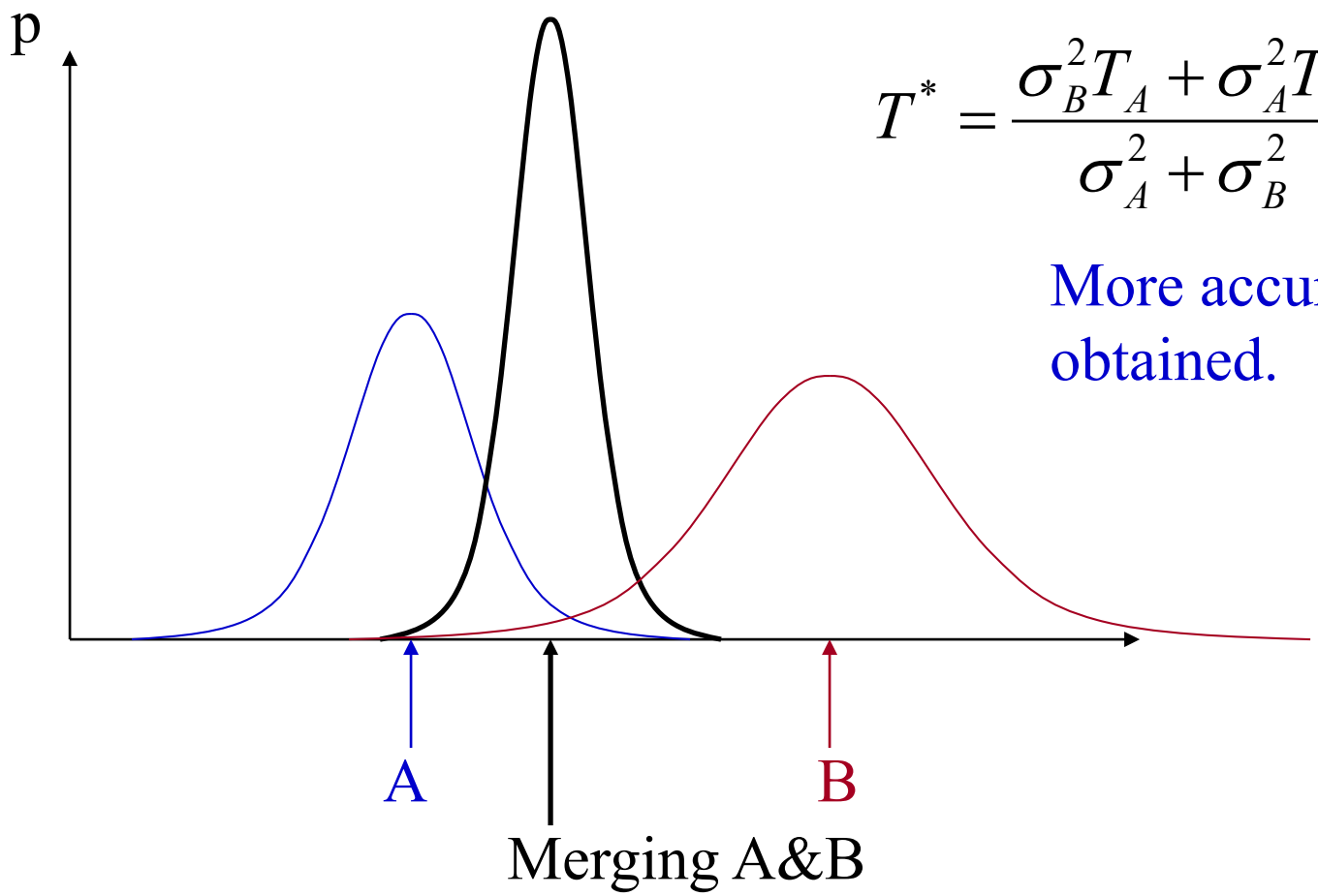
$$p_B(T) \propto \exp\left[-\frac{(T-T_B)^2}{2\sigma_B^2}\right]$$

$$p_{A\cap B}(T) = p_A(T) \bullet p_B(T)$$

$$\propto \exp\left[-\frac{(T-T_A)^2}{2\sigma_A^2} - \frac{(T-T_B)^2}{2\sigma_B^2}\right]$$

$$\propto \exp\left[-\frac{\sigma_A^2+\sigma_B^2}{2\sigma_A^2\sigma_B^2}\left(T - \frac{\sigma_B^2 T_A + \sigma_A^2 T_B}{\sigma_A^2+\sigma_B^2}\right)^2\right]$$

Merging 2 information (Bayesian estimation)

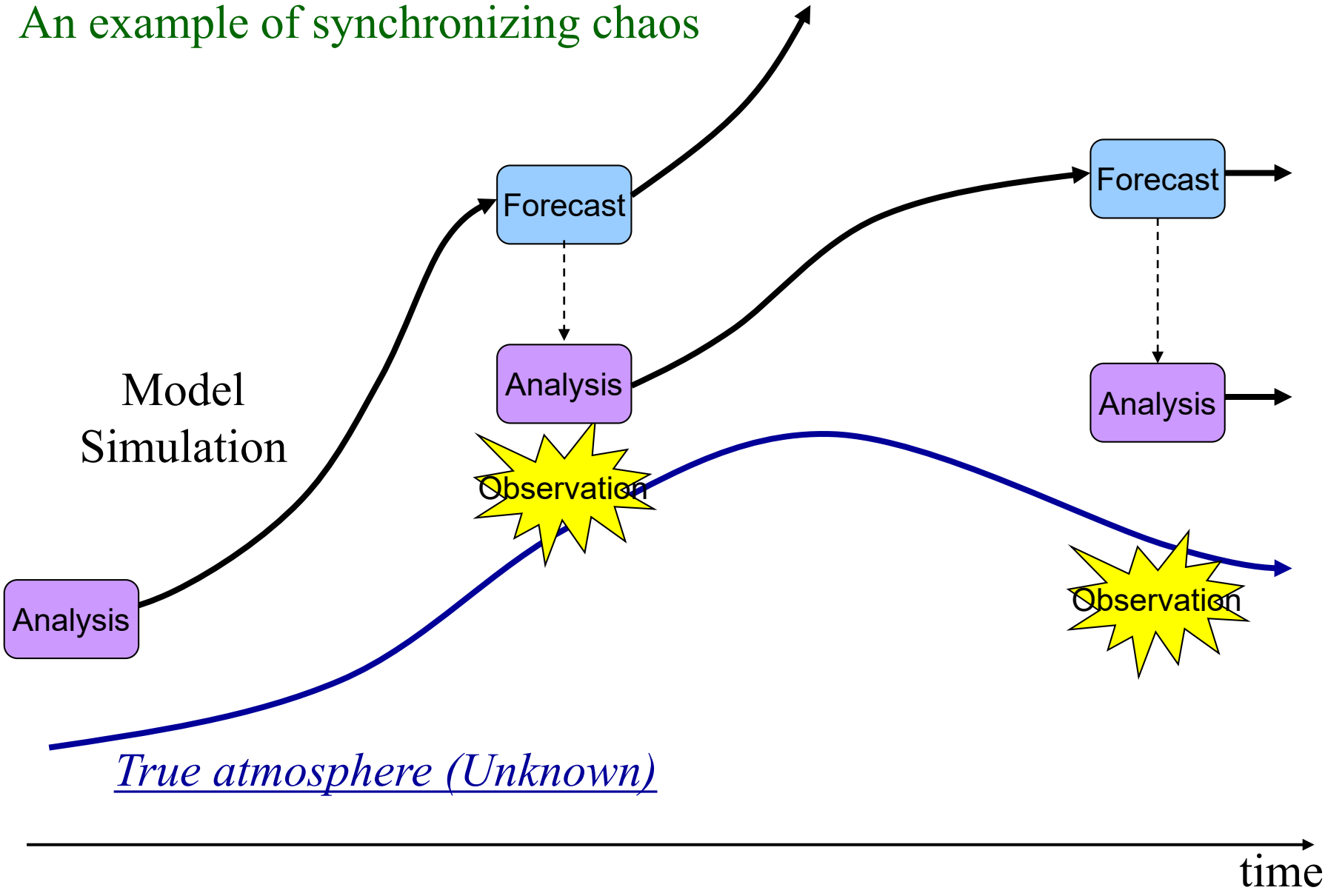


$$T^* = \frac{\sigma_B^2 T_A + \sigma_A^2 T_B}{\sigma_A^2 + \sigma_B^2}, \sigma^* = \frac{\sigma_A^2 \sigma_B^2}{\sigma_A^2 + \sigma_B^2}$$

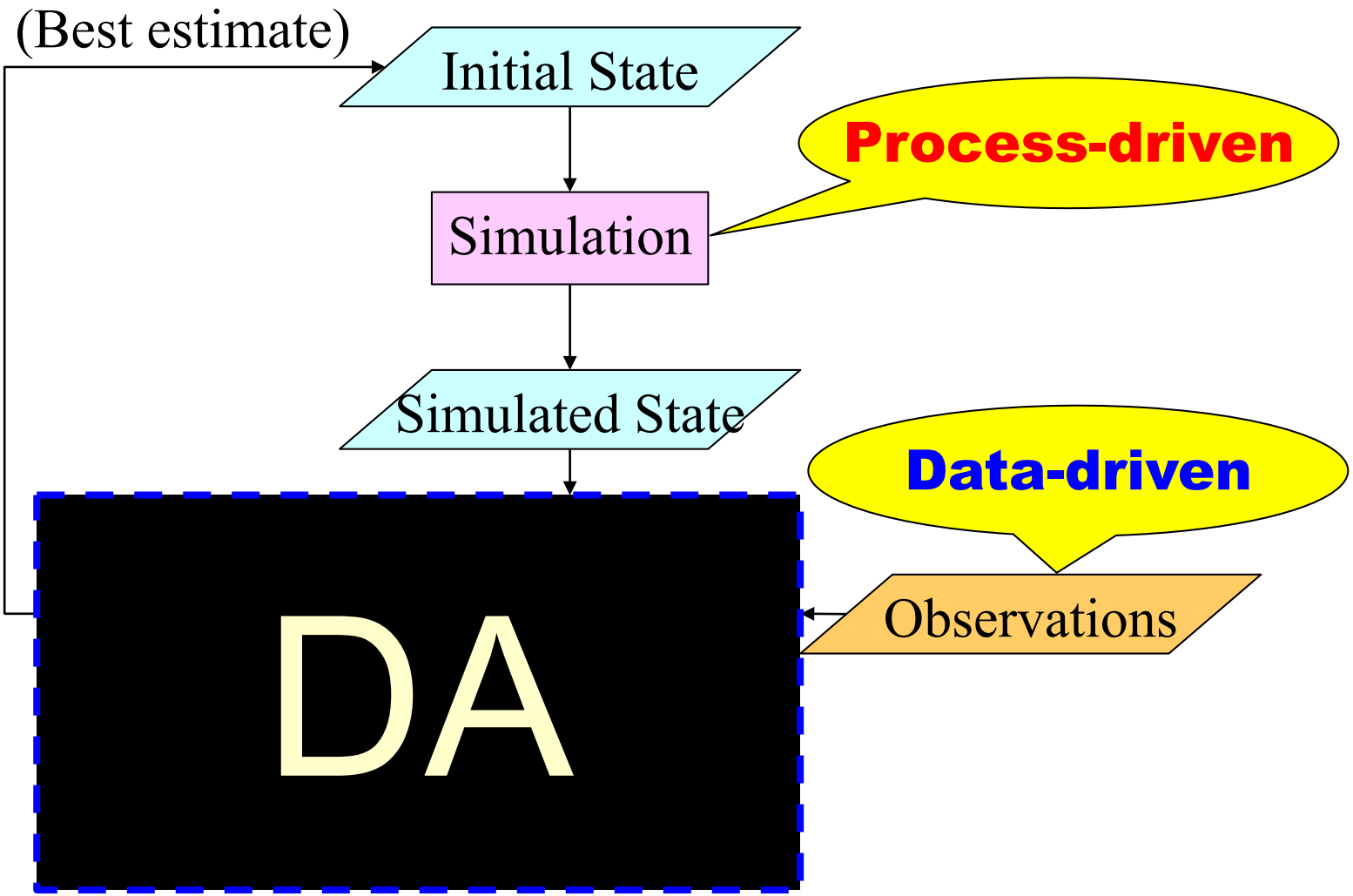
More accurate analysis is obtained.

Numerical Weather Prediction

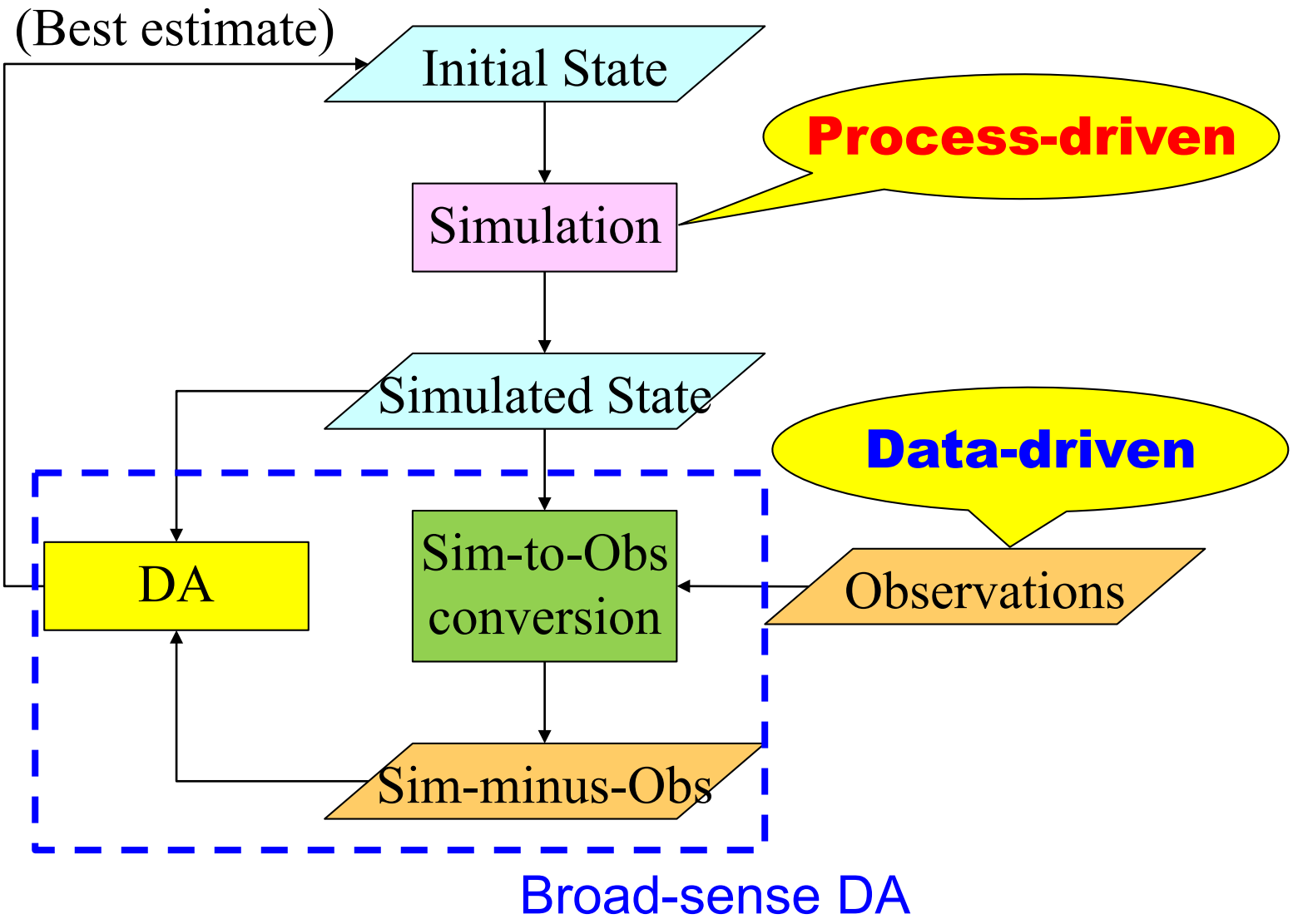
An example of synchronizing chaos



DA workflow


(Best estimate)
Initial State
Process-driven
Simulation
Simulated State
Data-driven
DA
Observations

DA workflow


(Best estimate)
Initial State
Process-driven
Simulation
Simulated State
Data-driven
DA
Sim-to-Obs conversion
Observations
Sim-minus-Obs
Broad-sense DA

What DA can do