2020/02/10 2nd IMT Atlantique & RIKEN Joint WS

Toward hybrid NWP-AI system for precipitation nowcasting

Shigenori Otsuka, Yasumitsu Maejima, and Takemasa Miyoshi
(RIKEN Center for Computational Science, Japan)

Acknowledgment:

M. Ohhigashi, S. Kotsuki, K. Kondo, G. Tuerhong, Y. Ishikawa,
J. Ruiz, S. Satoh, T. Ushio, R. Kikuchi, Y. Kitano, Y. Taniguchi, MTI Ltd.
P. Tandeo, V. P. Huynh

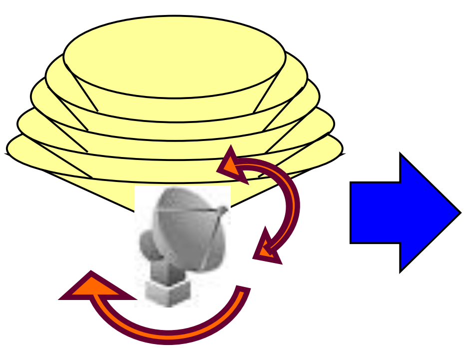
RIKEN
R-CCS
RIKEN Center for Computational Science
Data Assimilation Team
BDA BIG DATA ASSIMILATION

Background

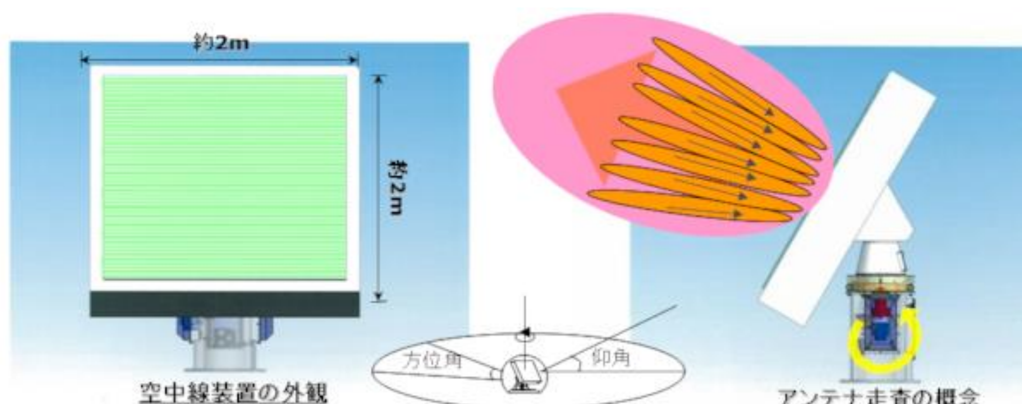
Flash flood in Kobe on July 28, 2008

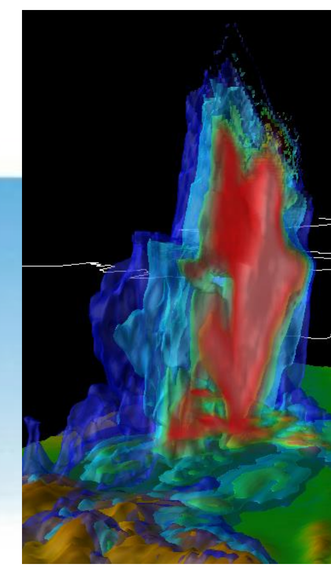
- Localized convective rain
  - Rapid development in 10 minutes
  - Needs for quick prediction
- Phased Array Weather Radar
  - Dense and frequent observations
    - 3D volume scans every 10-30 seconds
    - Range resolution of 100 m, 100 elevation angles
    - Operation started in 2012 at Osaka University

→ We developed ultra-rapid-update nowcast



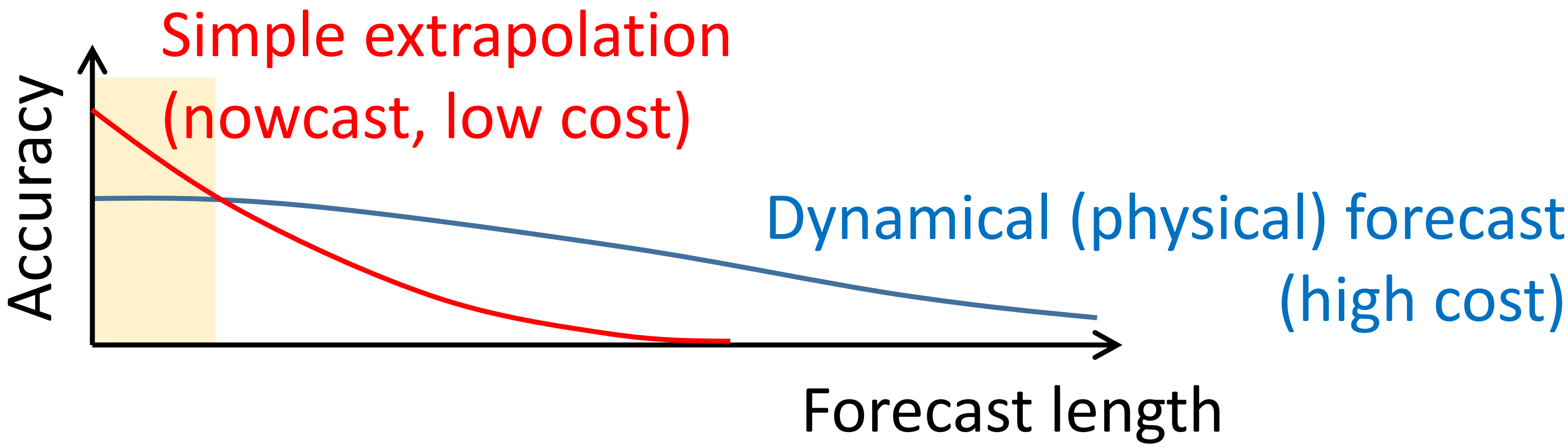
(NICT)





Background

- Nowcast outperforms dynamically-based forecast in short range precipitation forecast

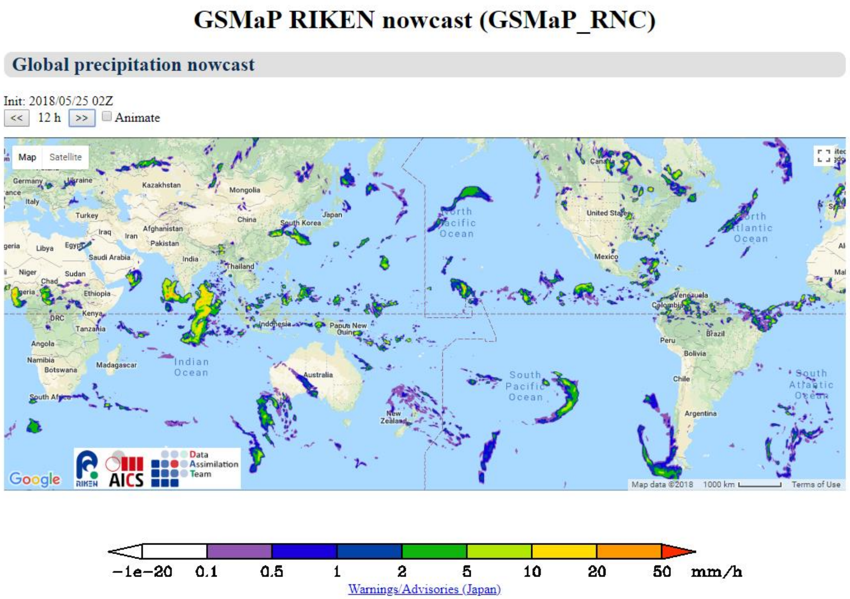


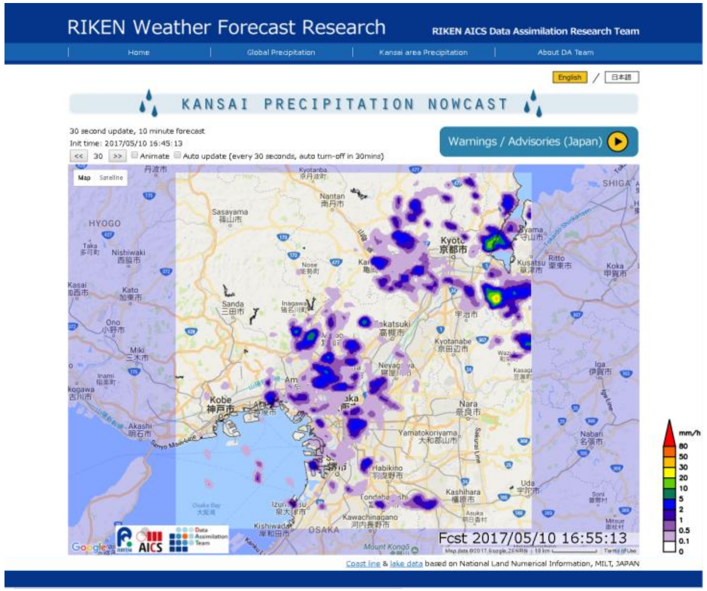
Background

- Nowcasting = Now + forecasting
- Four types
  - Numerical weather prediction
  - Kinematic prediction (e.g., motion vector)
  - Statistical prediction (e.g., neural network)
  - Conceptual modelling (e.g., 1D convection model)

Two precipitation nowcasting systems at RIKEN (https://weather.riken.jp)

Global precipitation nowcast (hourly updated)



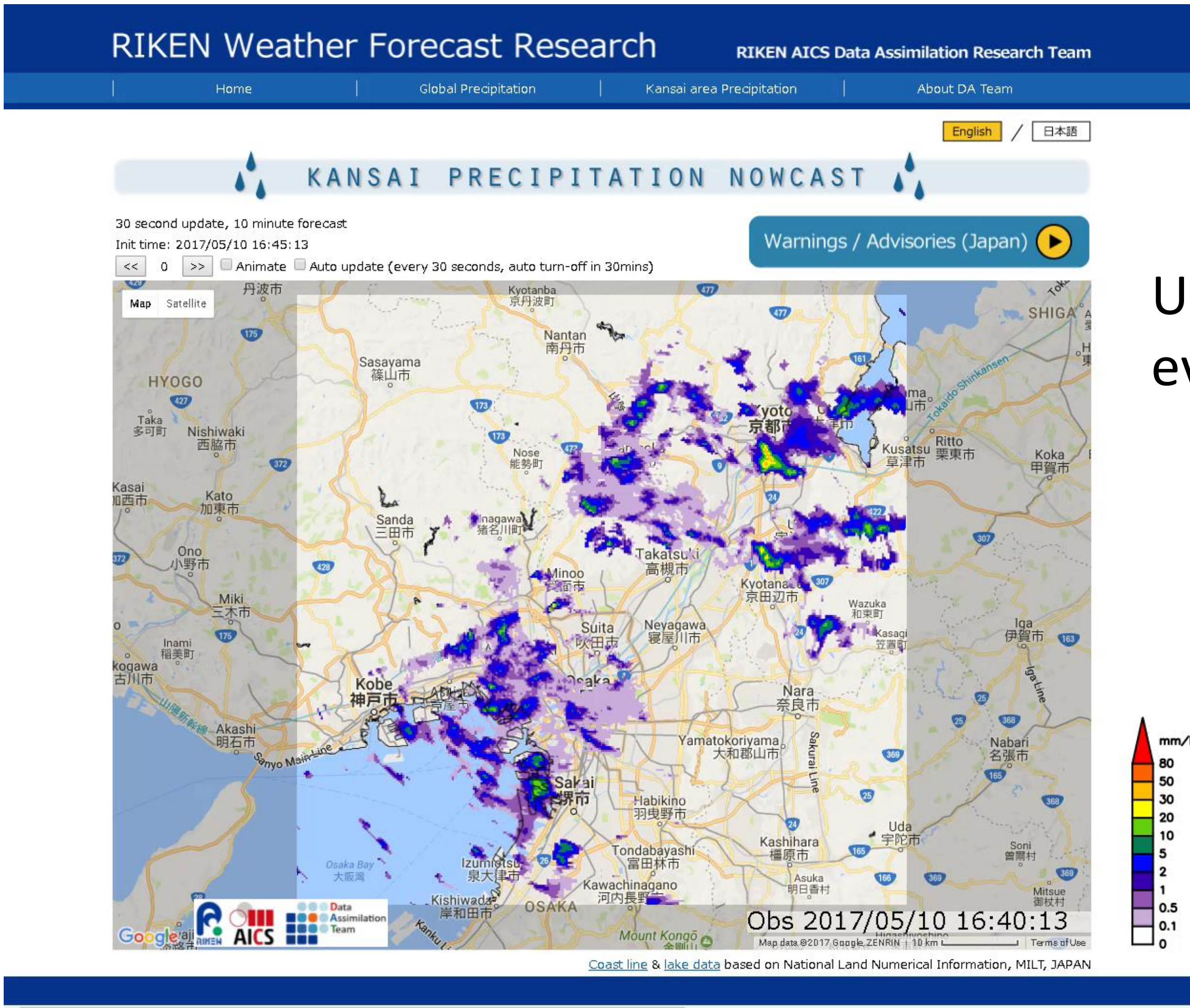


These two share the main part of the program codes

Phased-Array Weather Radar nowcast (updated every 30 seconds)

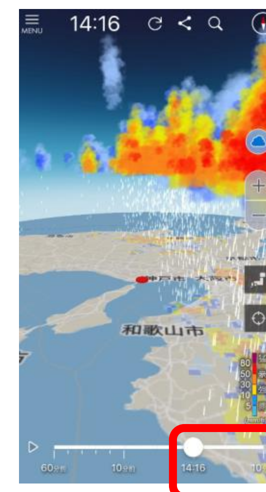
Phased-Array Weather Radar 3D nowcast

https://weather.riken.jp/



(NICT)

Updated every 30 s



App by MTI Ltd.

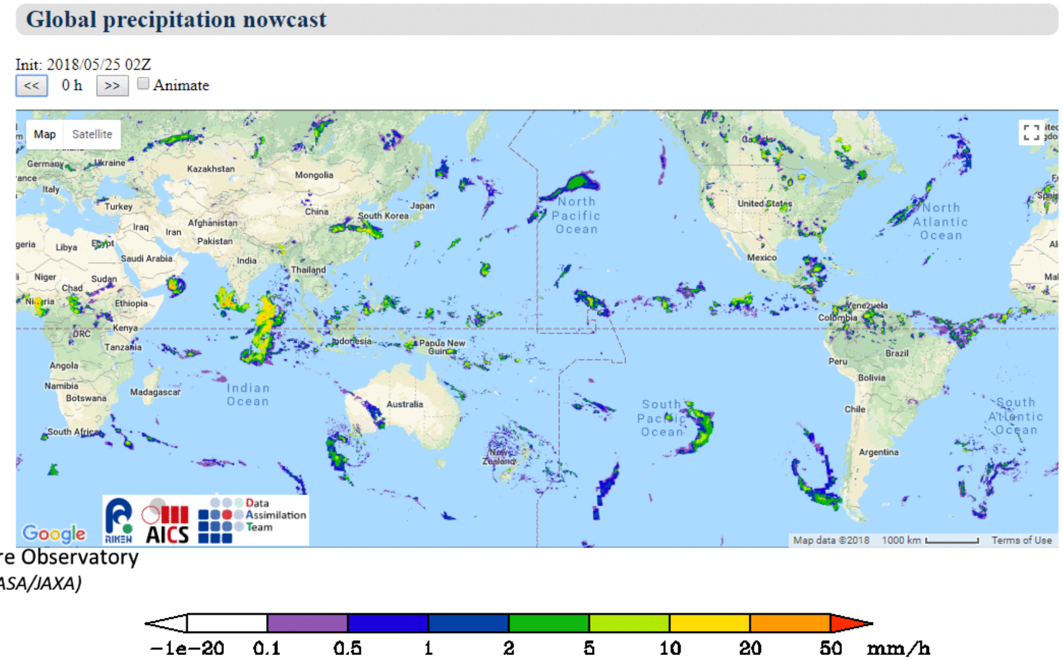
247,000+ downloaded

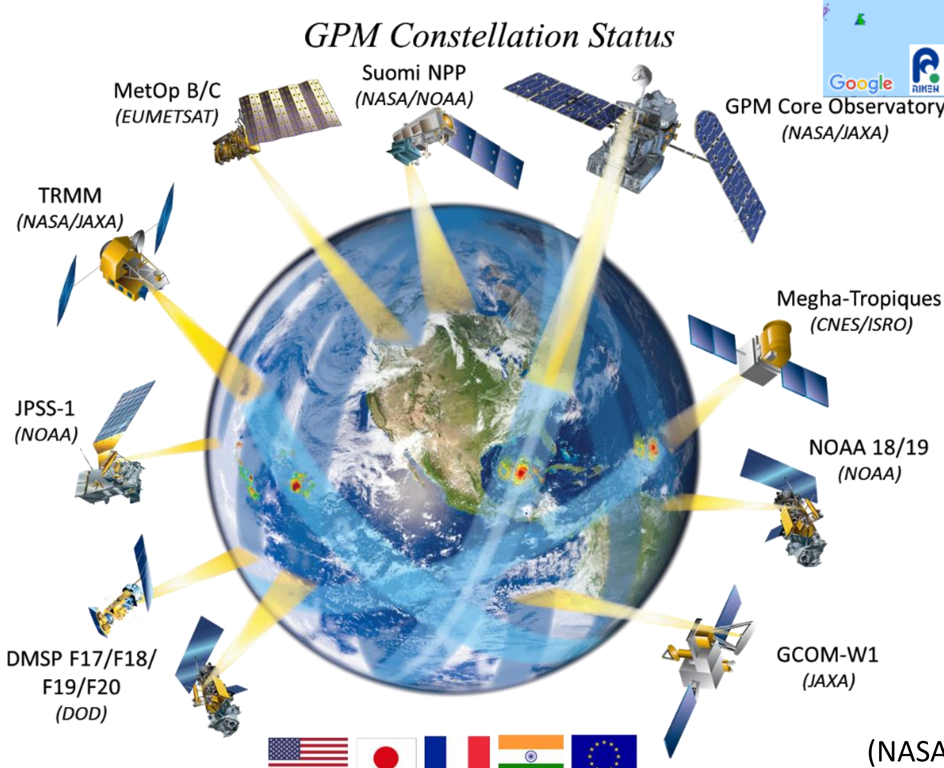
GSMaP RIKEN Nowcast

Global Satellite Mapping of Precipitation (GSMaP)

by Japan aerospace exploration agency (JAXA)

GSMaP RIKEN nowcast (GSMaP_RNC)





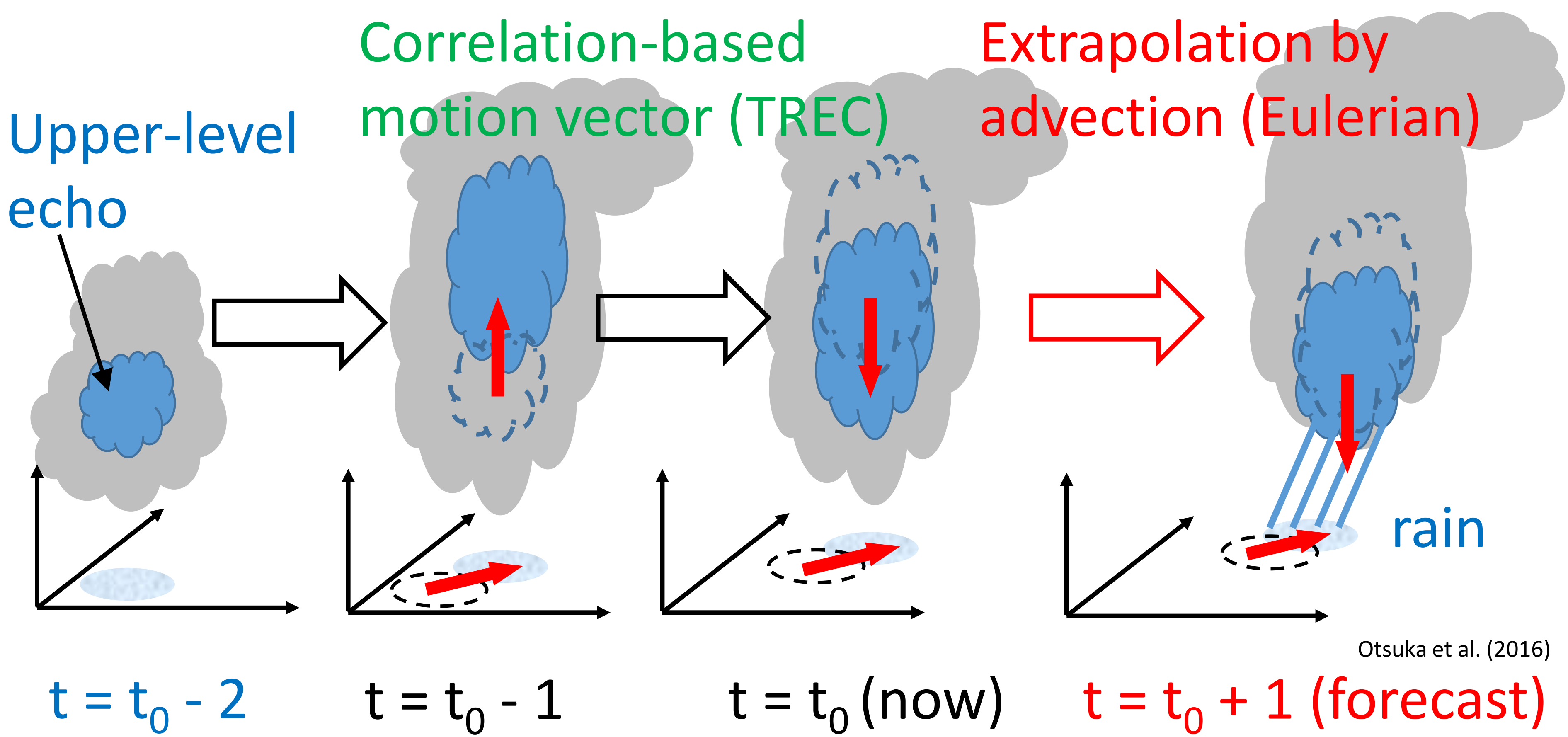
(NASA)

Input for nowcasting:
GSMaP Near-Real-Time product

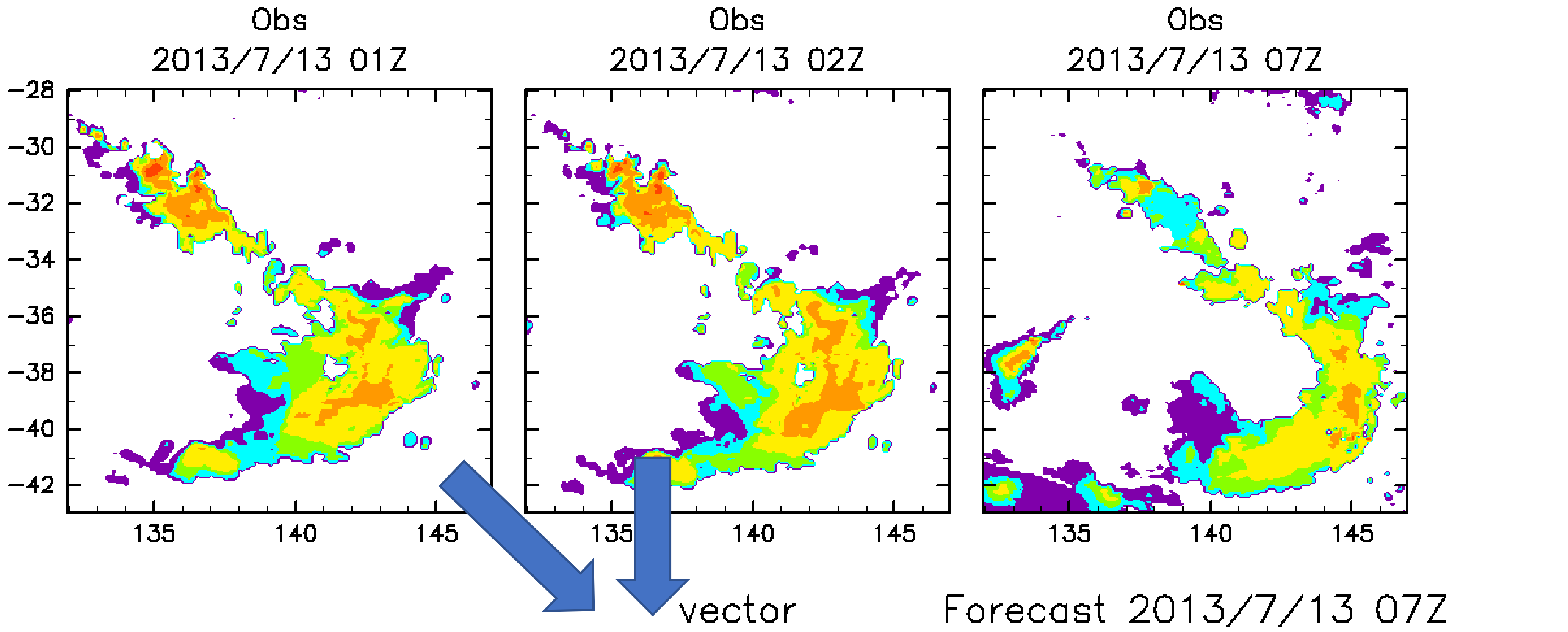
Methodology of kinematic nowcasting

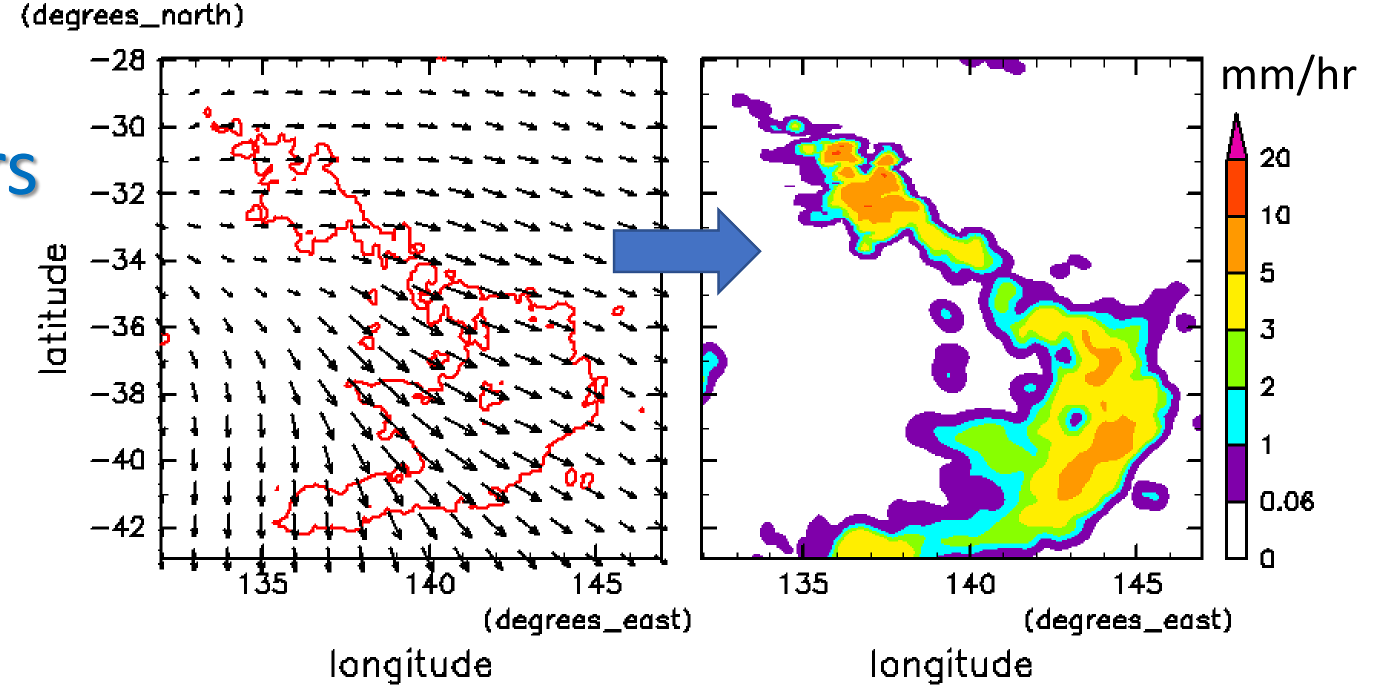
Method of kinematic nowcasting

- 3D space-time extrapolation (advection)
- Pure image processing (no physics considered)

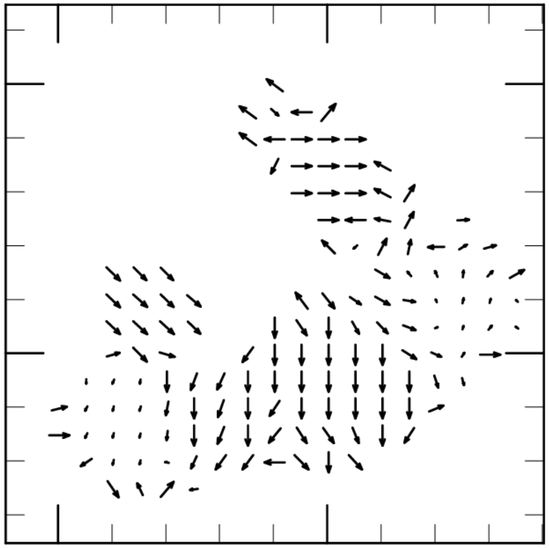


Computing motion vectors by pattern matching (TREC-based)





Estimating motion vectors by Data Assimilation



Noisy motion vectors

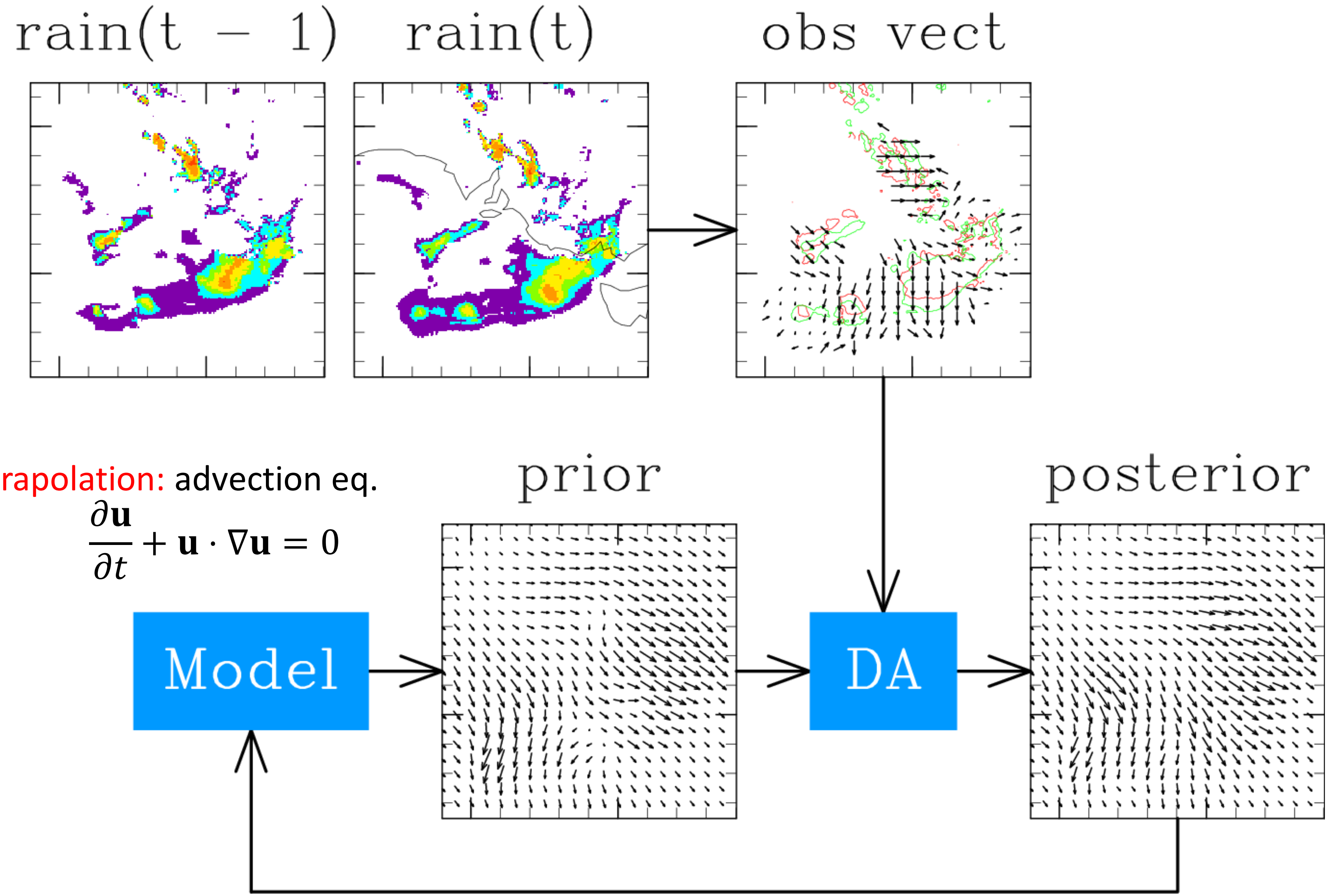
Motivation

- Motion vectors tend to be noisy due to observation errors
- Data assimilation may help reduce the noise

Methods

- Local Ensemble Transform Kalman Filter (LETKF)
  - Used for GSMaP RIKEN Nowcast
- Single-point temporal filtering
  - Used for PAWR 3D Nowcast

DA for space-time extrapolation



Extrapolation: advection eq.

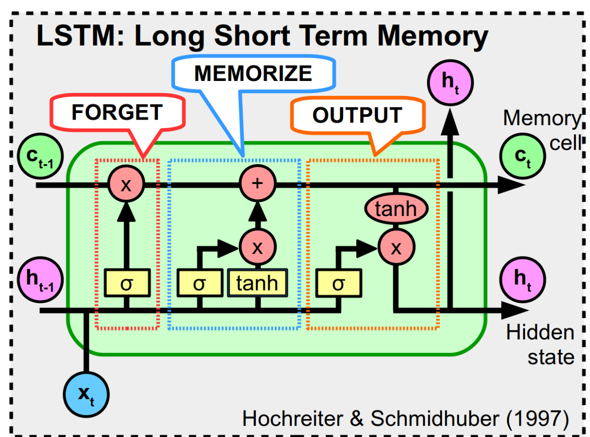
$$\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{u} = 0$$

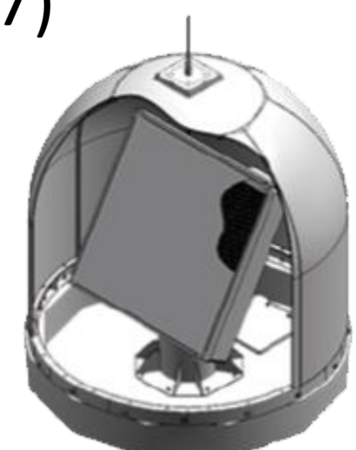
(Otsuka, Kotsuki, and Miyoshi, 2016)

Methodology of nowcasting with machine learning

Machine learning with the Convolutional LSTM

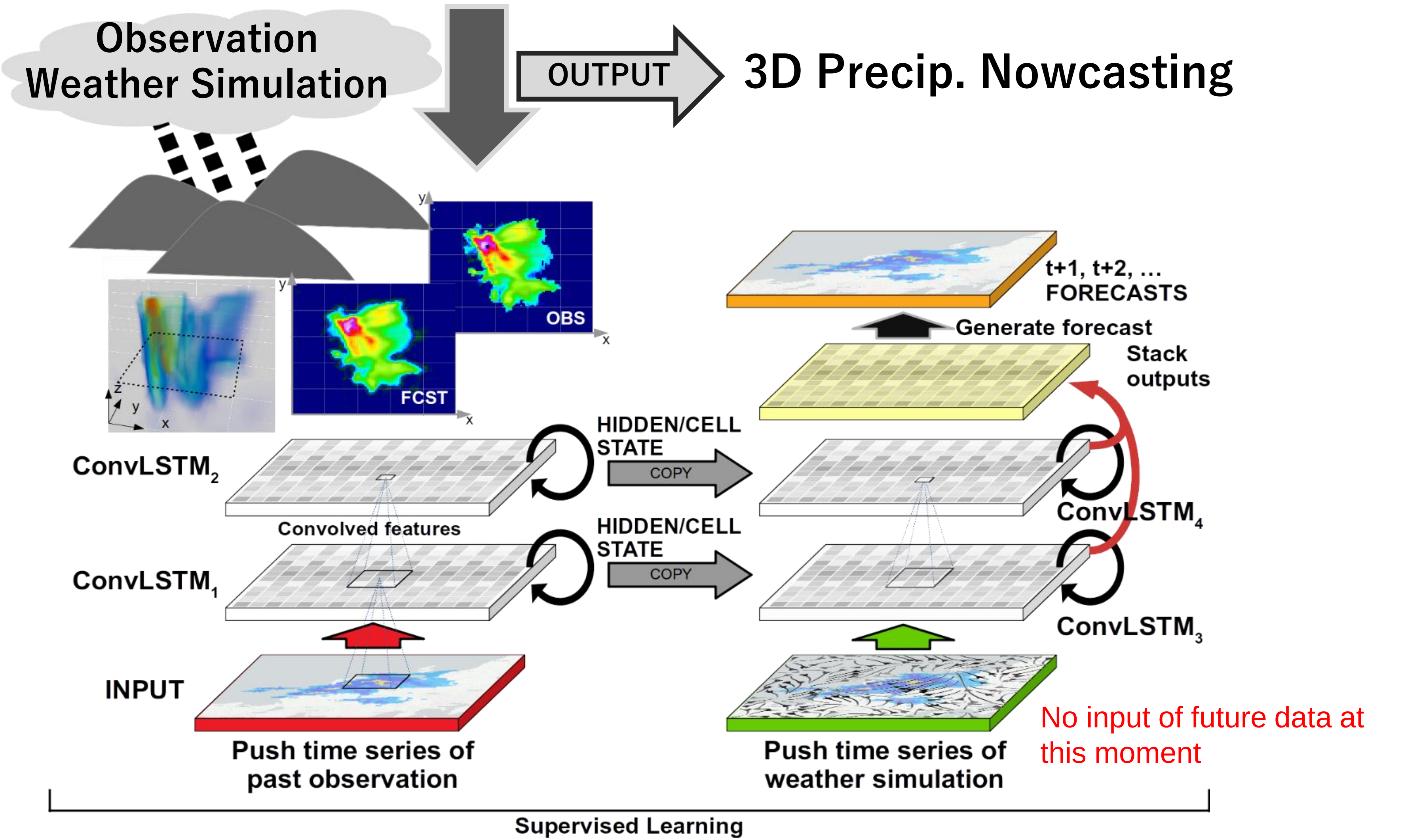
- Phased Array started its operation in 2012
- High-frequency, high-resolution data accumulated
  - 30-sec, 100-m-resolution, 100-elevations, 300-azimuth
- → Data driven prediction
- Long Short-Term Memory is suitable for sequence data (LSTM, Hochreiter & Schmidhuber 1997)
- Convolution is used to utilize spatial information (Shi et al. 2015)



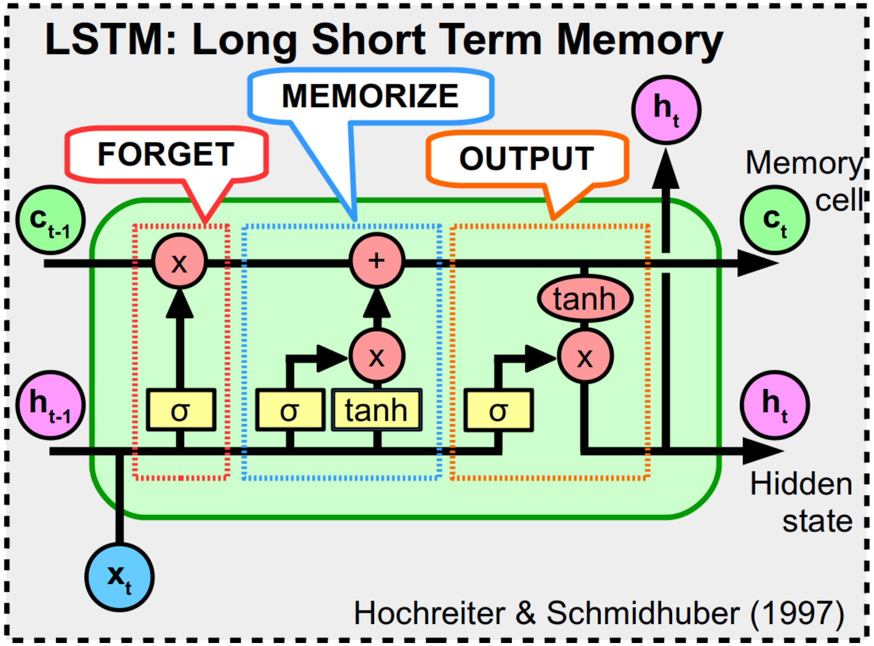


(NICT)

Conv-LSTM by Shi et al. (2015)
Extended to three-dimensional radar data



ConvLSTM



- Gate operations in LSTM become convolution
- Weights in LSTM become tensors of kernel size x kernel size x input channels x output channels

Input $i_t = \sigma(W_{xi} * \mathcal{X}_t + W_{hi} * \mathcal{H}_{t-1} + W_{ci} \circ \mathcal{C}_{t-1} + b_i)$

Forget $f_t = \sigma(W_{xf} * \mathcal{X}_t + W_{hf} * \mathcal{H}_{t-1} + W_{cf} \circ \mathcal{C}_{t-1} + b_f)$

Memorize $\mathcal{C}_t = f_t \circ \mathcal{C}_{t-1} + i_t \circ \tanh(W_{xc} * \mathcal{X}_t + W_{hc} * \mathcal{H}_{t-1} + b_c)$

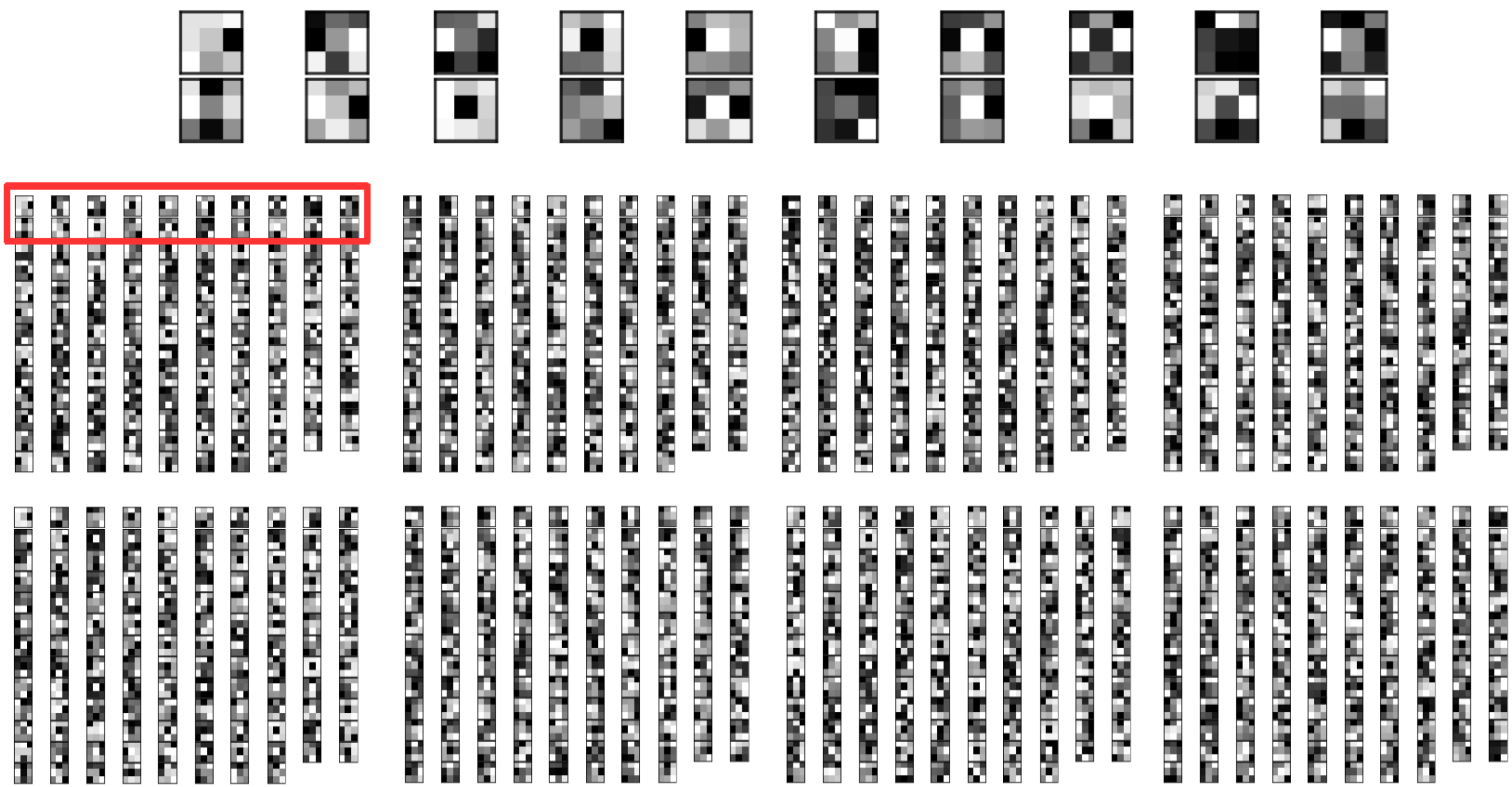
Output $o_t = \sigma(W_{xo} * \mathcal{X}_t + W_{ho} * \mathcal{H}_{t-1} + W_{co} \circ \mathcal{C}_t + b_o)$

Hidden state $\mathcal{H}_t = o_t \circ \tanh(\mathcal{C}_t)$

Shi et al. (2015)

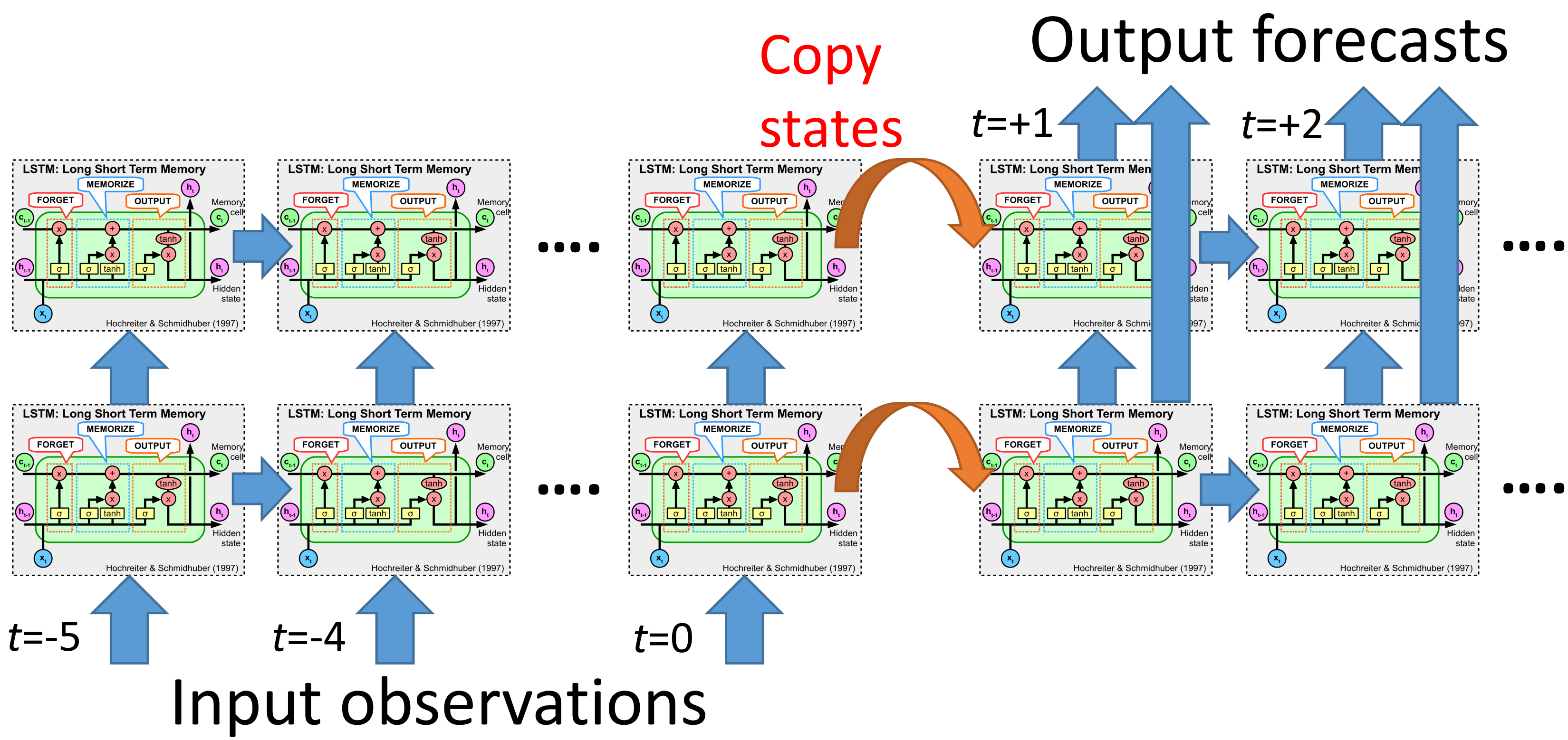
*: convolution, ∘: Hadamard product

An example of filters trained by ConvLSTM



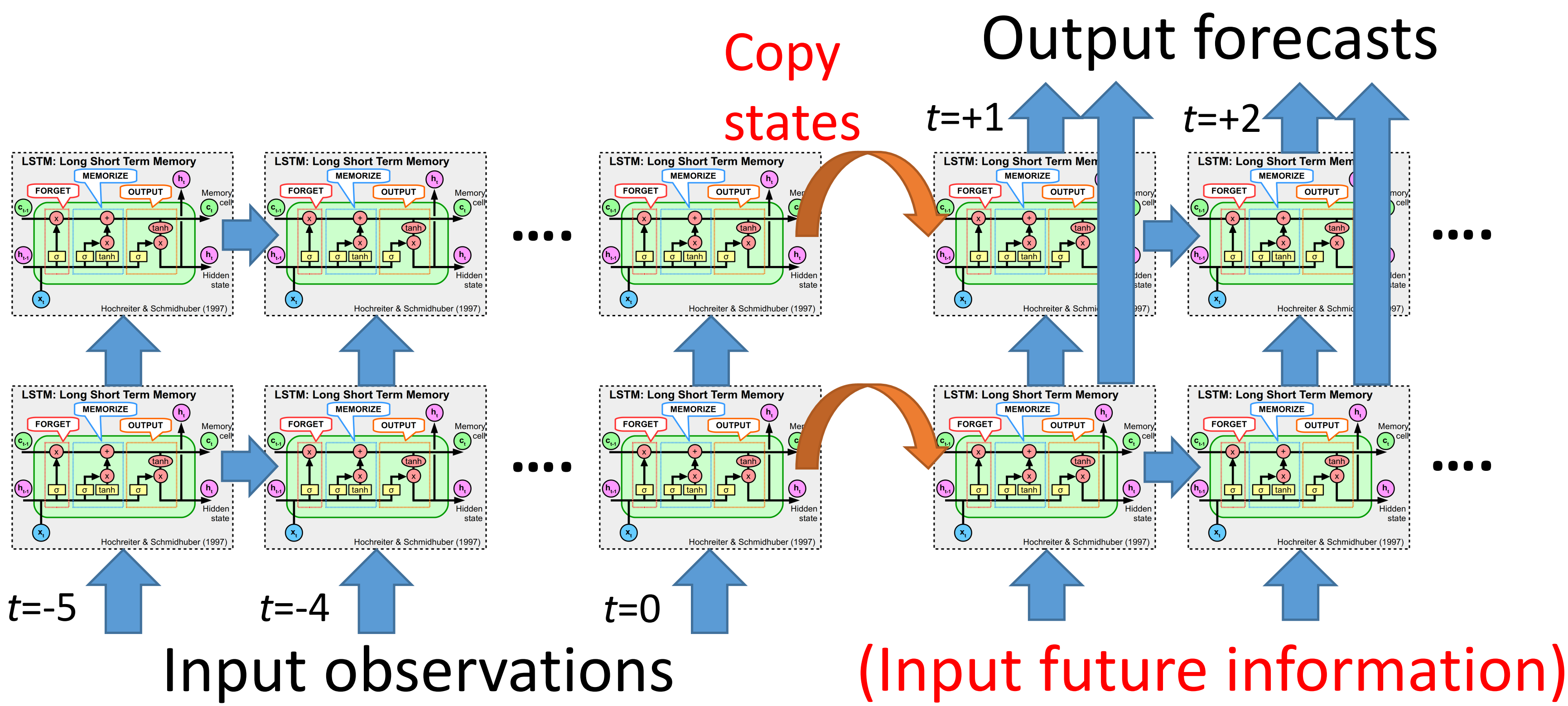
ConvLSTM

- Encoder-Decoder with 2 LSTM layers



ConvLSTM

- Encoder-Decoder with 2 LSTM layers

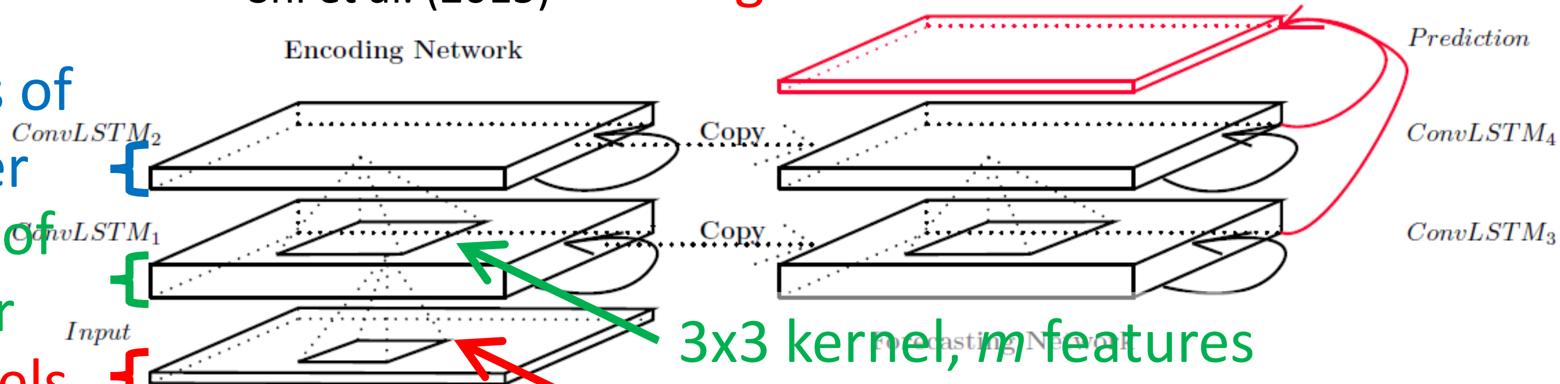


Extension to 3D radar data

- 9 vertical levels from 1.5 km to 3.5 km are used
  - Every 0.25 km
- Vertical levels are treated as "channels" of input images

9-channel output is generated from *n* + *m* channels

Shi et al. (2015)



m channels of the 2nd layer

n channels of the 1st layer

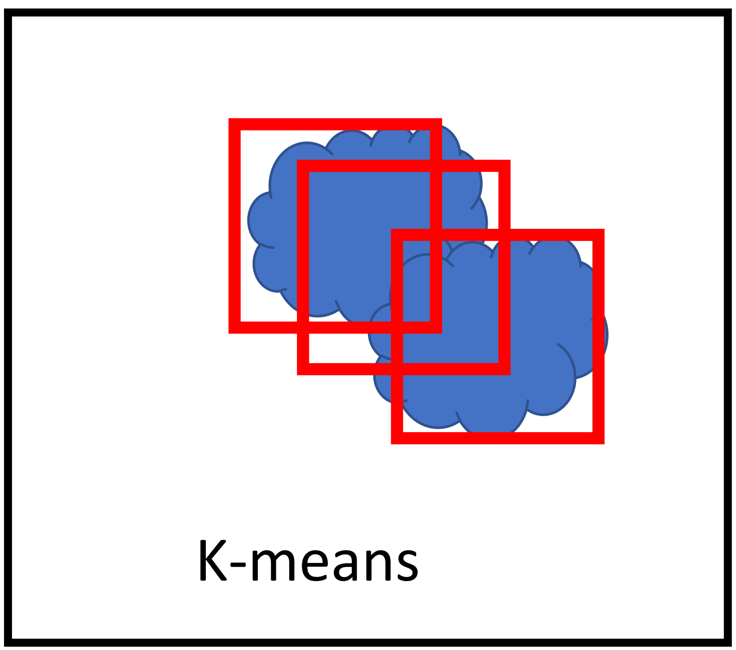
9 vertical levels
→ treated as 9 color channels of a 2D image

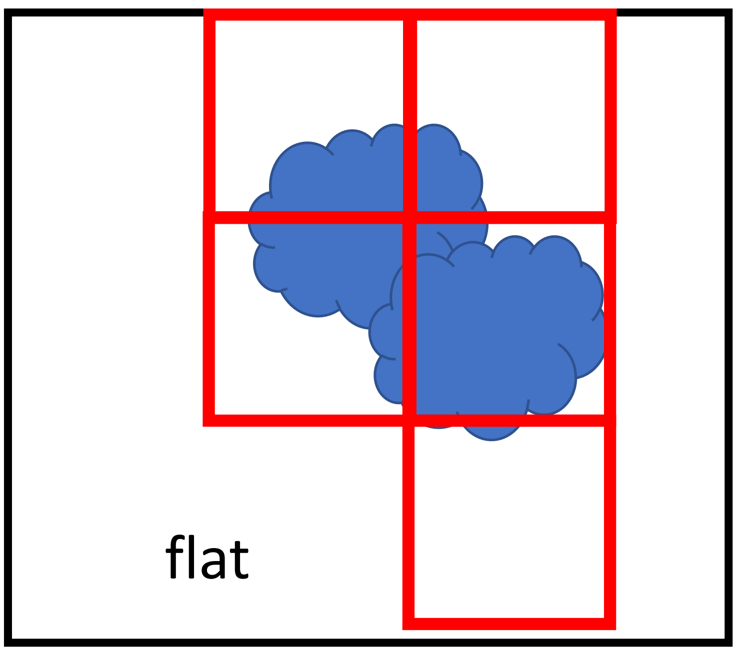
3x3 kernel, *m* features

3x3 kernel, *n* features

Region of Interest (ROI) extraction

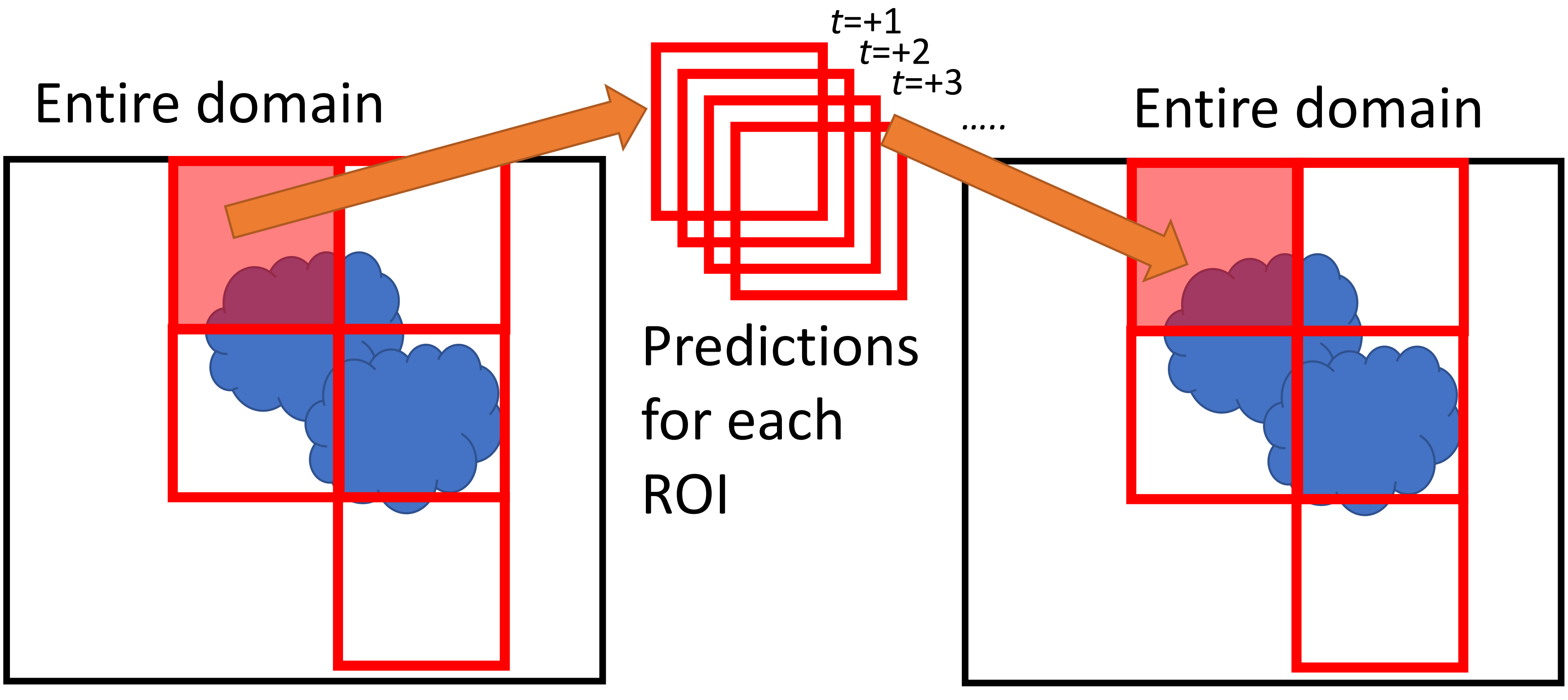
- Extracting areas including rainy pixels
- Increasing the effective sample size
- Two approaches: K-means / flat





ROI merge

- Predictions for each ROI are merged to reconstruct the entire domain



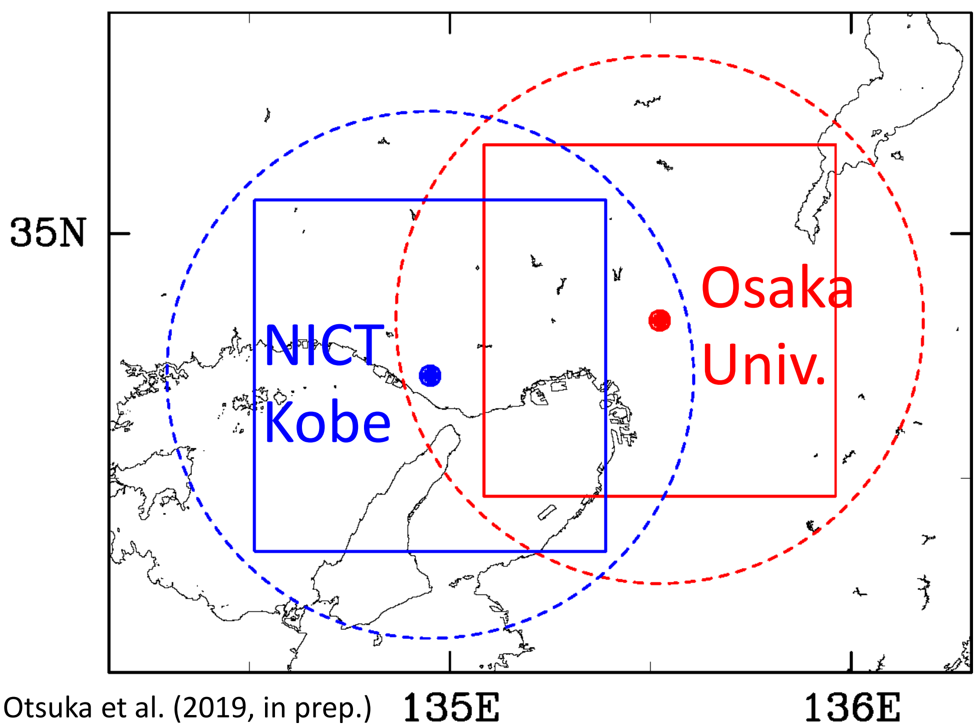
Changes from the original ConvLSTM

	Shi et.al. (2015)	Huynh & Tandeo	Ohhigashi & Otsuka
Data	2D radar images	3D radar data	3D radar data
Interval	6 minutes	30 seconds	30 seconds
ROI extraction	K-means	K-means	Flat ROI extraction
ROI size	100 x 100 transformed to 50 x 50 x 4	61 x 61 x 9	60 x 60 x 9
Filtering	Disk filter	No-filter	No-filter
Input	Past 5 steps	Past 6 steps	Past 6 steps Forecast 20 steps
Output	Future 15 steps	Future 5 steps	Future 20 steps
ROI merge	No	No	Flat ROI merge
Framework	Theano	Theano	Chainer
Optimizer		AdaDelta, learning rate: 10^{-4}	Adam, learning rate: optimized later
Convolution	3 x 3, 64 features	3 x 3, 32 features	3 x 3, optimized later

Comparison of kinematic nowcasting and machine learning

Computational domain and specification of kinematic nowcast

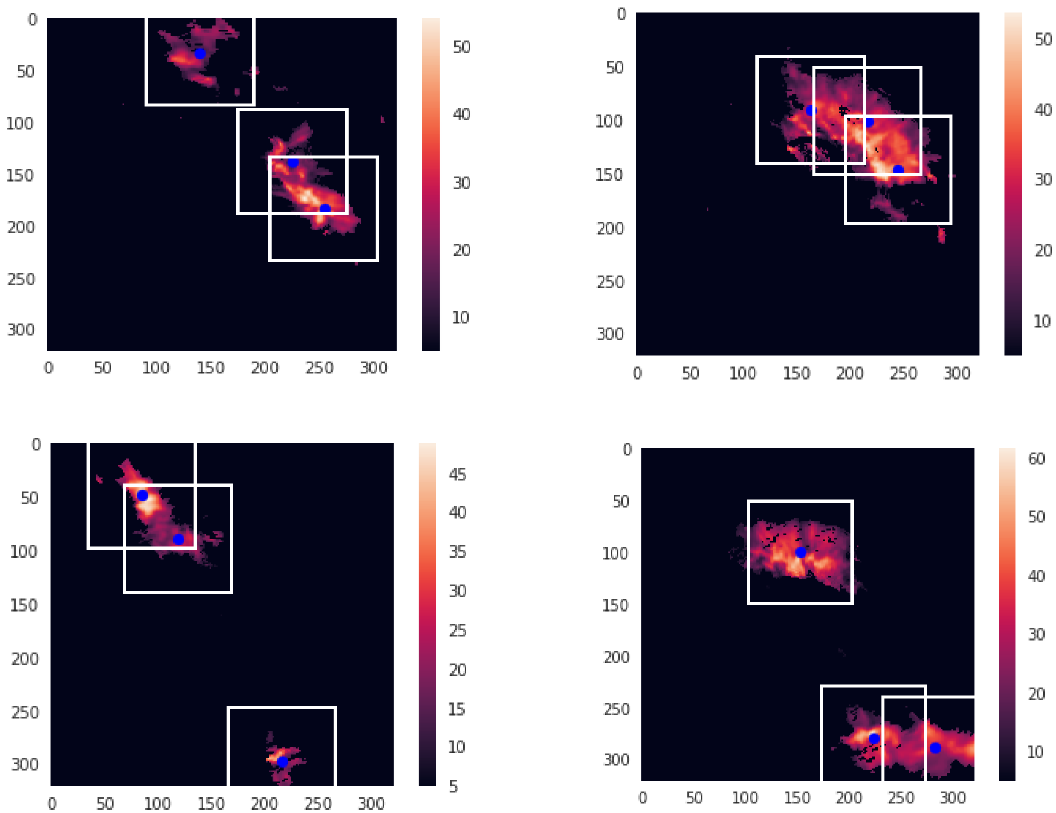
- Three-dimensional nowcast since May 2017
- NICT Kobe PAWR is used



Otsuka et al. (2019, in prep.)

Resolution	250 m (Hor. / Ver.)
Grid points	321 x 321 x 57
Update Freq.	30 seconds
Motion vector	TREC (fractional)
Quality control	Provided by NICT
Quality control of motion vector	Correlation *r*, range of *r*, sample size for *r*, solid body rot.
Temporal filter for motion vector	Obs.:Fcst = 1:10 weighted average
Forecast length	3D: 5 minutes + 2D: 10 minutes

Test case



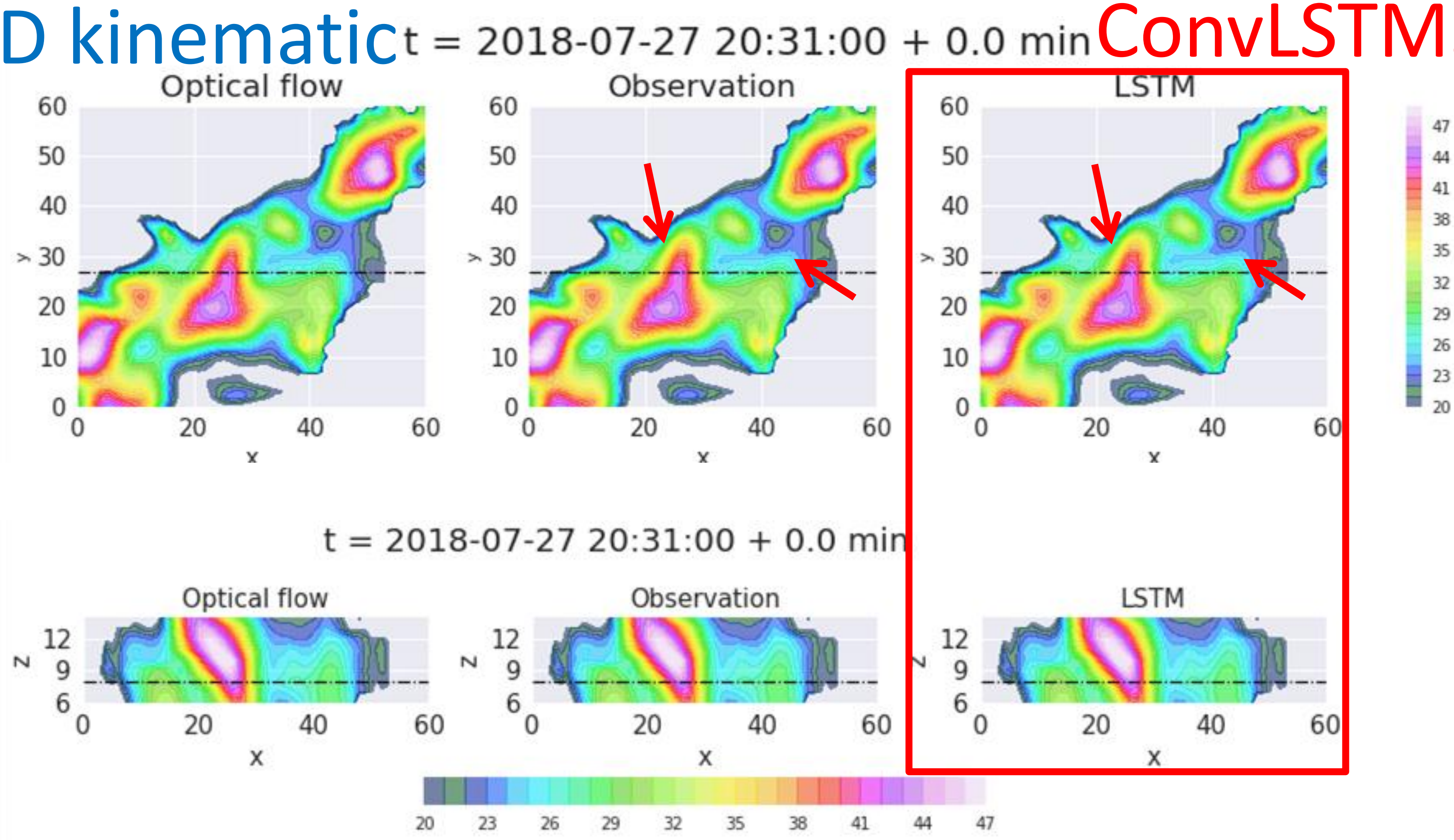
- Data:
  31 May &
  26-27 July, 2018
- Box size:
  61 x 61 x 9 pixels
- 6 steps for input,
  5 steps for forecast
- 3,500 samples for training
  - K-means to extract patches
- 300 samples for test

(Work with Mr. Viet Phi Huynh and Prof. Pierre Tandeo)

Prediction by ConvLSTM3D

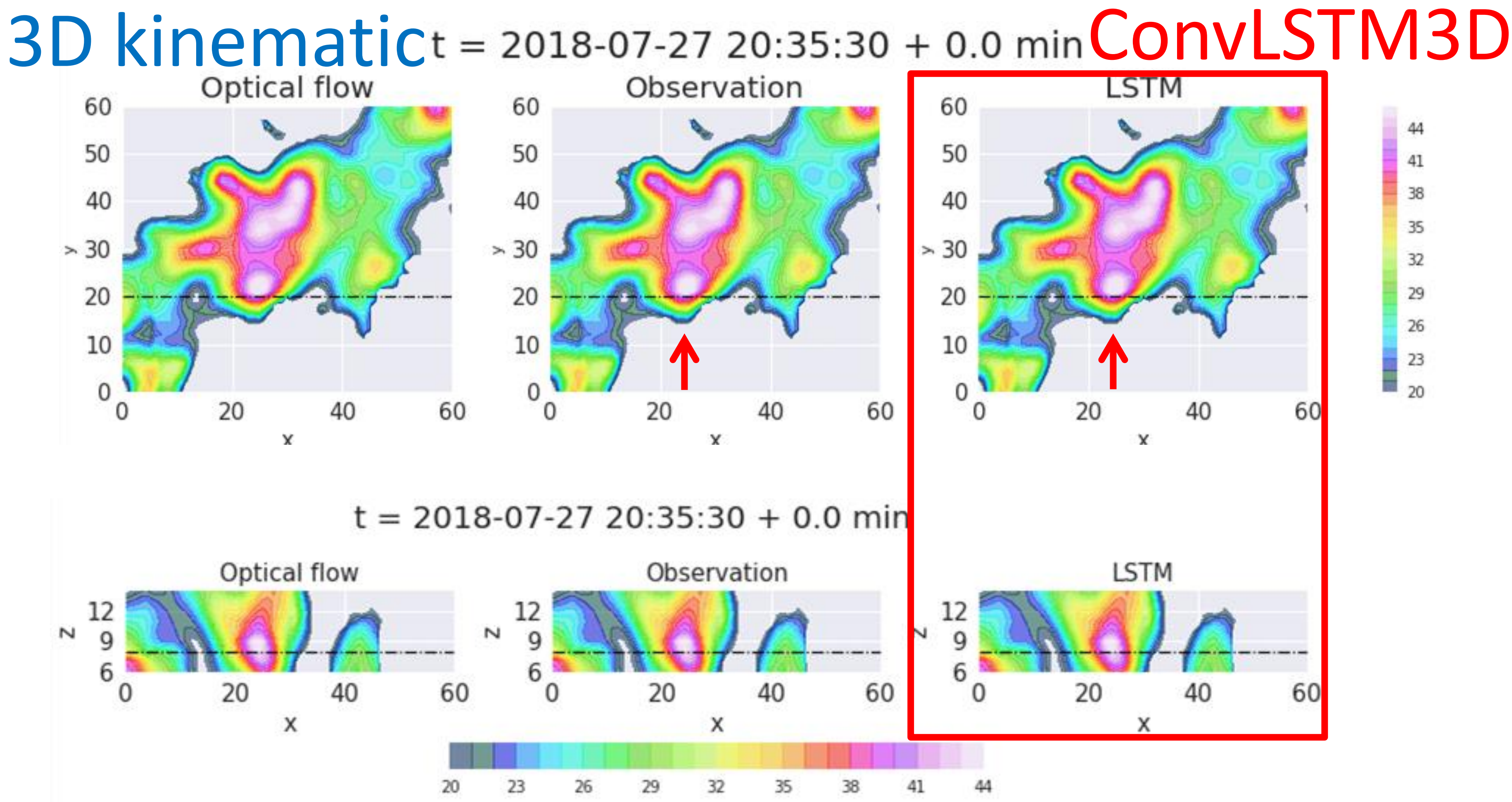
3D kinematic

ConvLSTM3D



(Work with Mr. Viet Phi Huynh and Prof. Pierre Tandeo)

Prediction by ConvLSTM3D



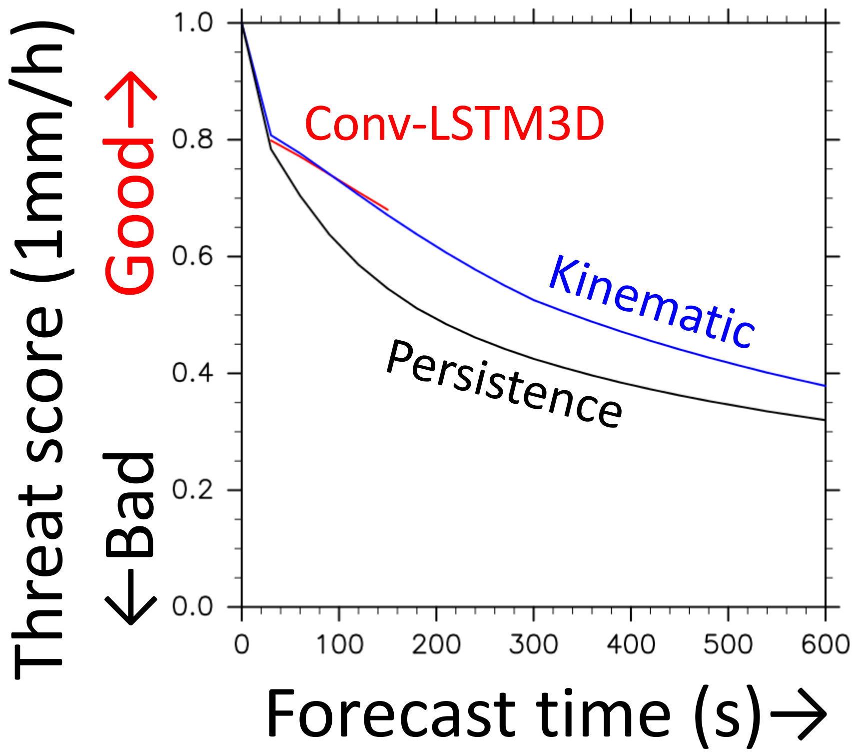
(Work with Mr. Viet Phi Huynh and Prof. Pierre Tandeo)

ConvLSTM3D real-time run

- Training with 97,935 rainy samples during April-May 2018
  - Training: 91,308
  - Test: 6,627

- Real-time run
  - Since June 2019
  - 2.5-min forecast
  - Inference over small areas, finally merged

Average in July 2019



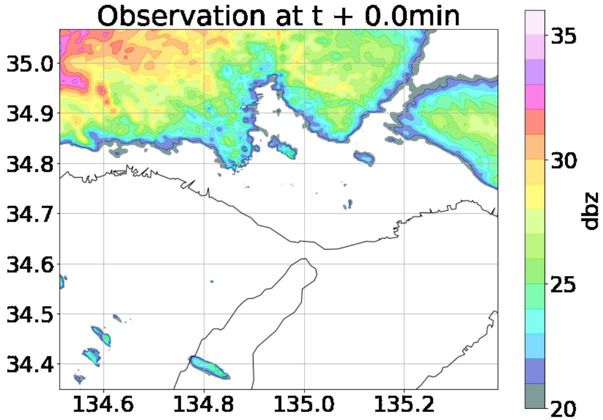
Example: 2019/4/10 (full domain)

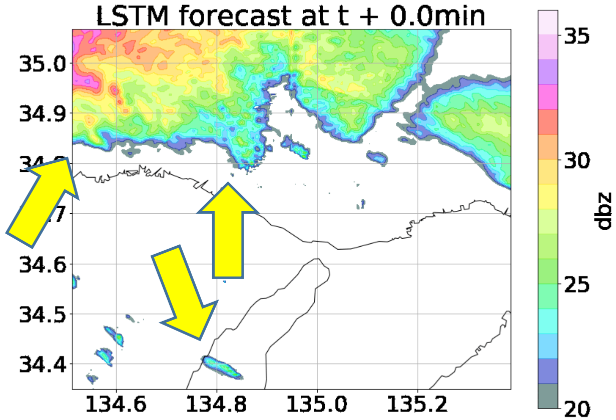
Observation

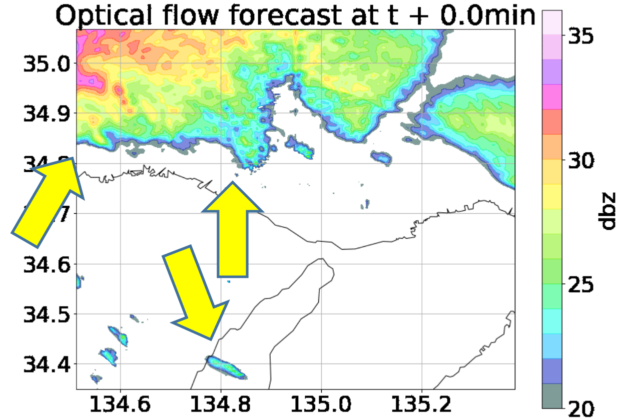
Conv-LSTM3D

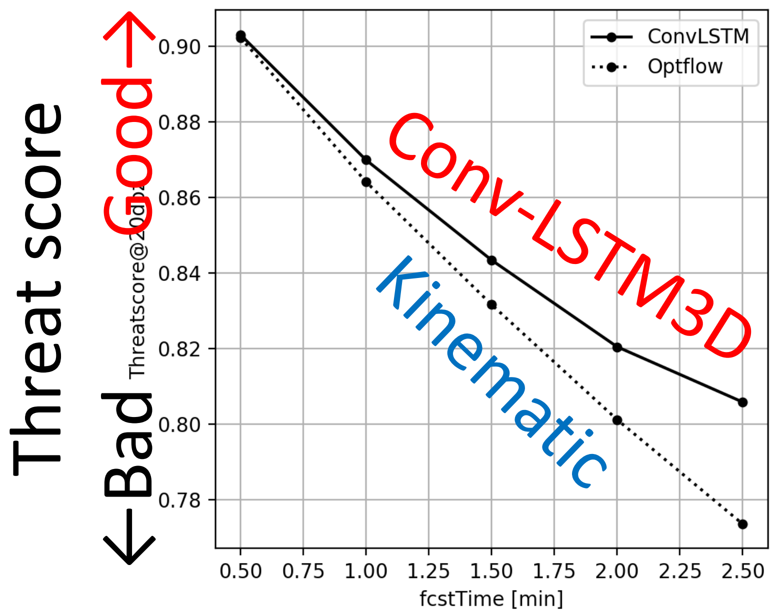
2019-04-10 06:05:30

Kinematic

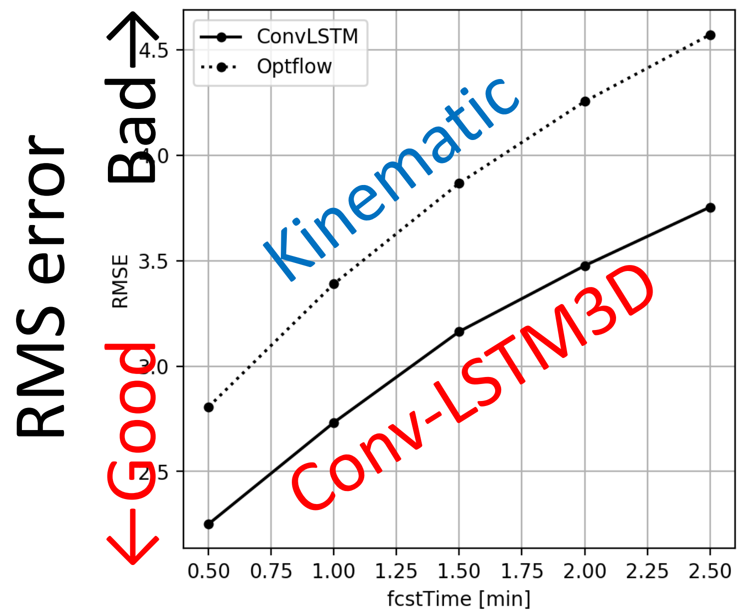






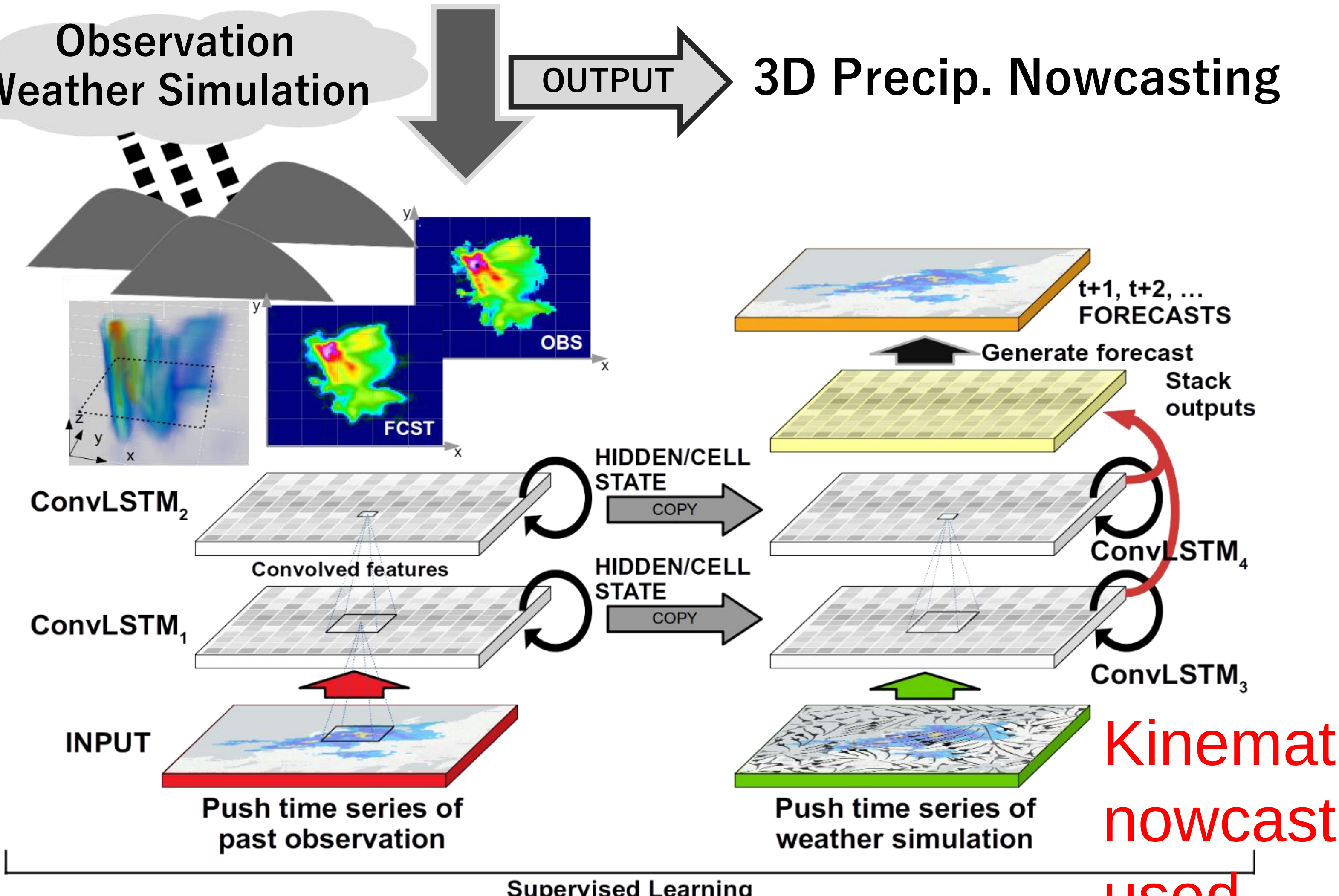


Forecast time (s) →



Forecast time (s) →

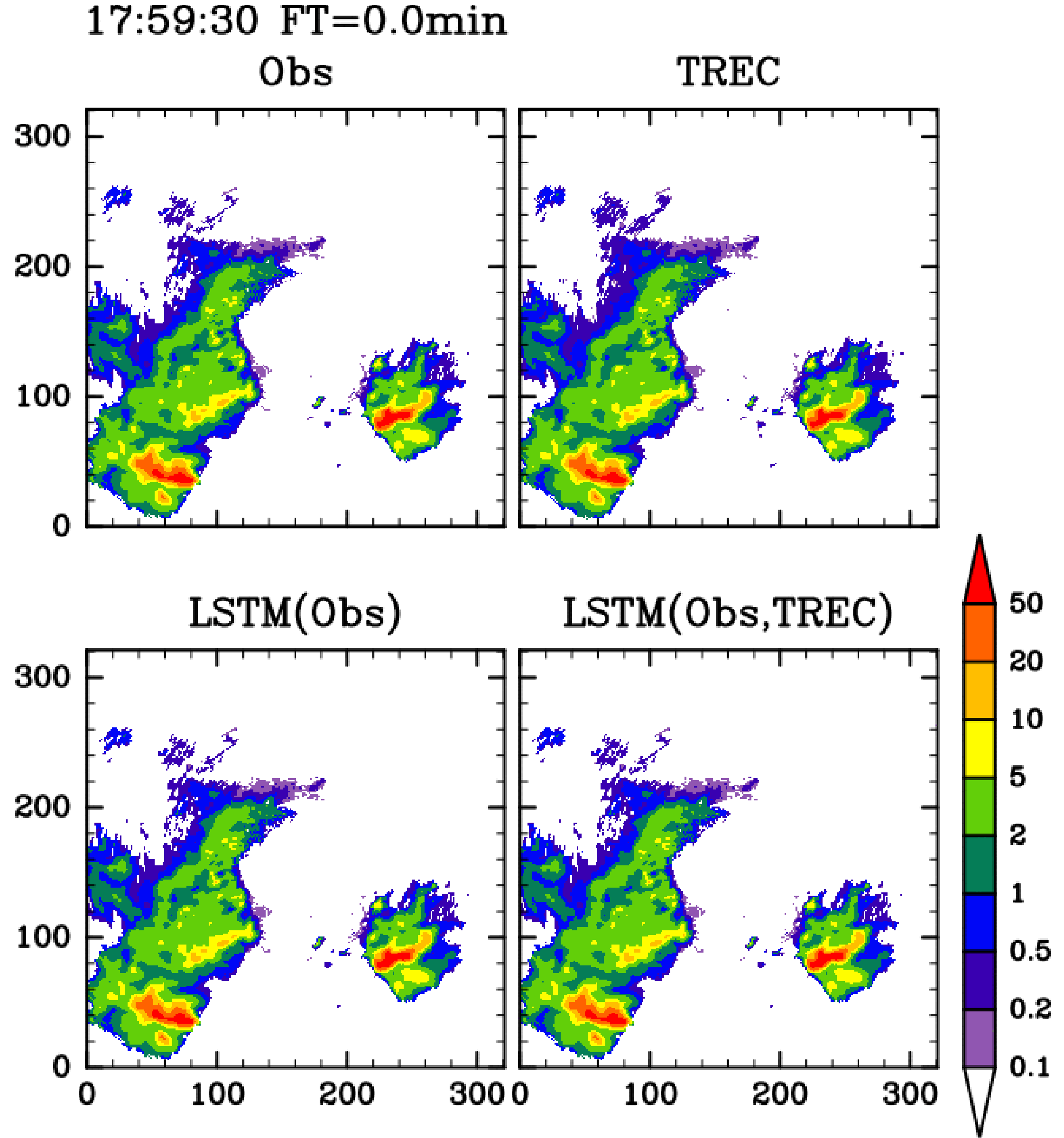
Conv-LSTM by Shi et al. (2015)
Extended to three-dimensional radar data



Kinematic nowcast is used

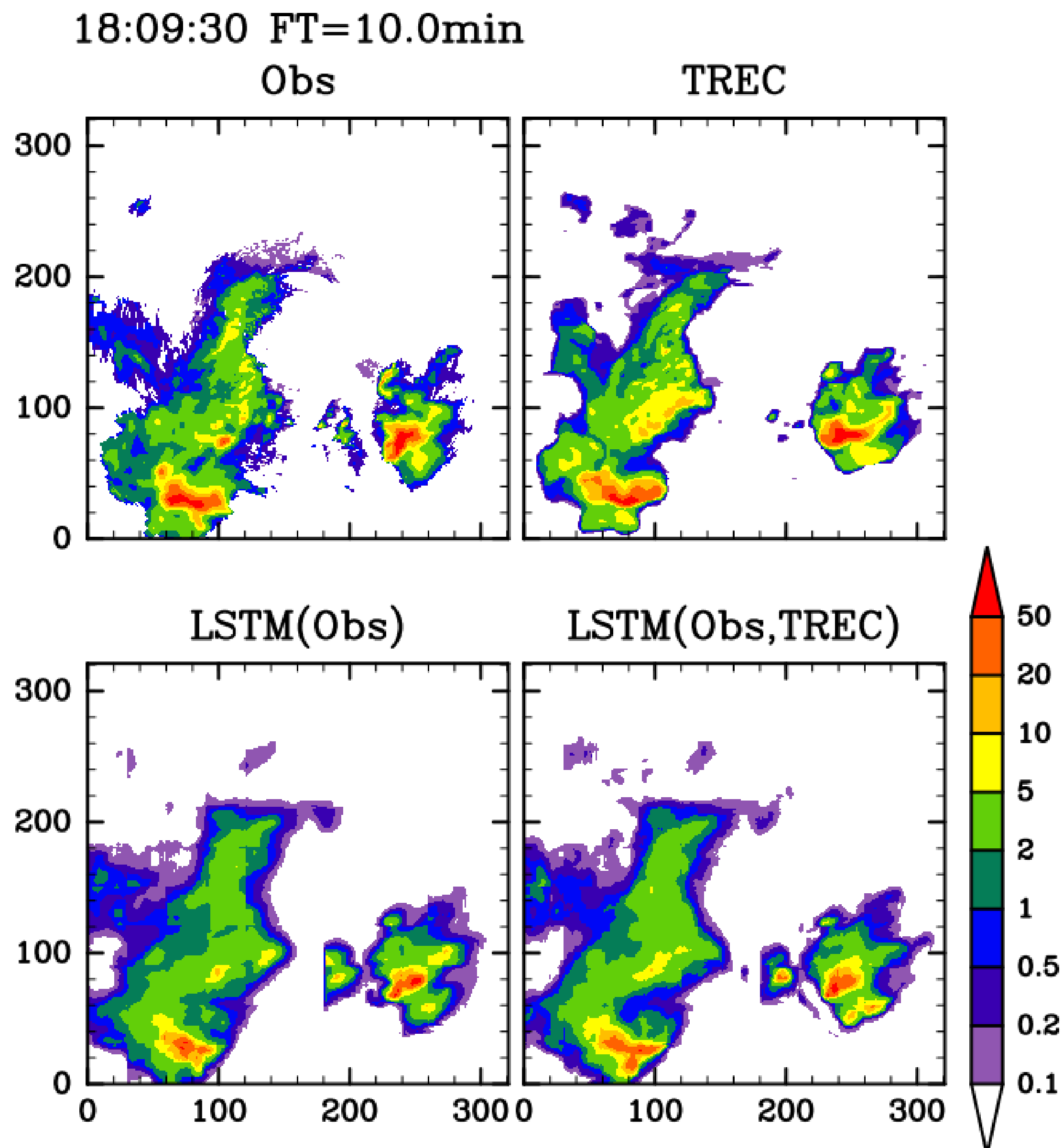
Preliminary results

- Training:
  5,852 samples
  - 2019/6/10 16:30:00 - 17:49:30
- Test:
  1,101 samples
  - 2019/6/10 17:50:00 - 17:59:30



Preliminary results

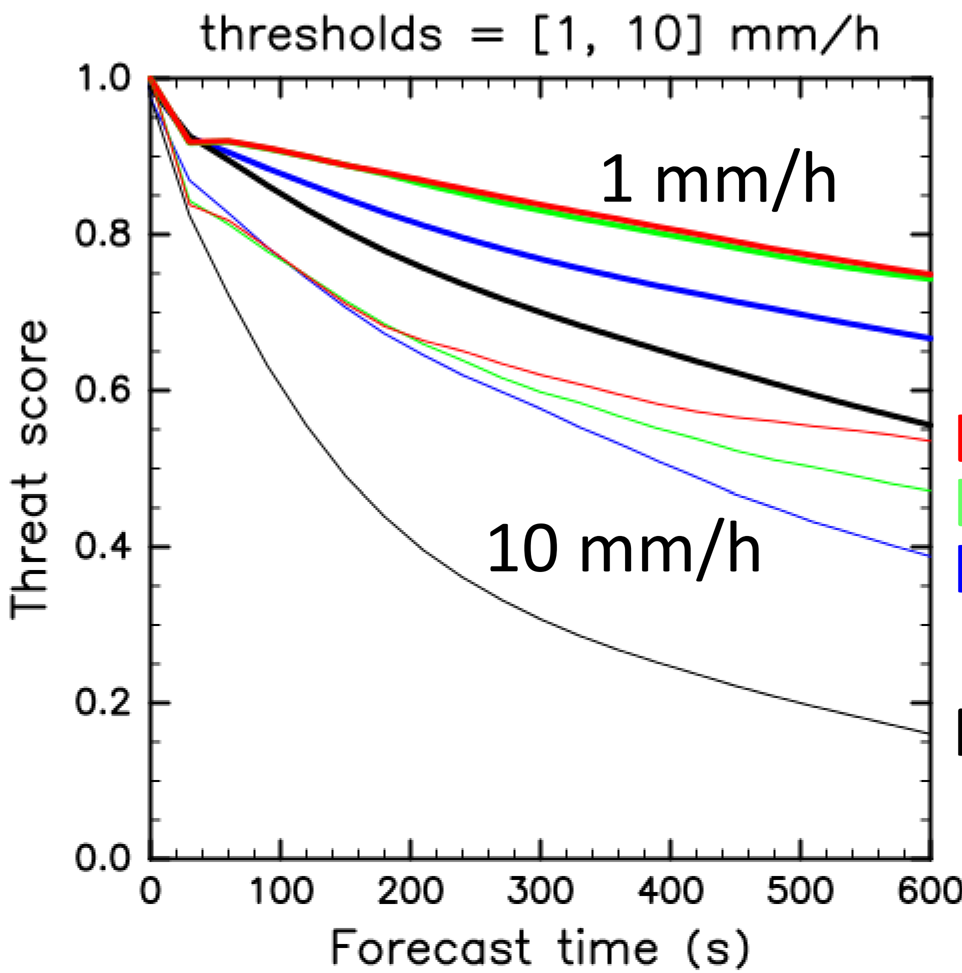
- Training:
  5,852 samples
  - 2019/6/10 16:30:00 - 17:49:30
- Test:
  1,101 samples
  - 2019/6/10 17:50:00 - 17:59:30



Preliminary results

Average of 20 forecasts

2019/6/10 17:50:00 - 17:59:30



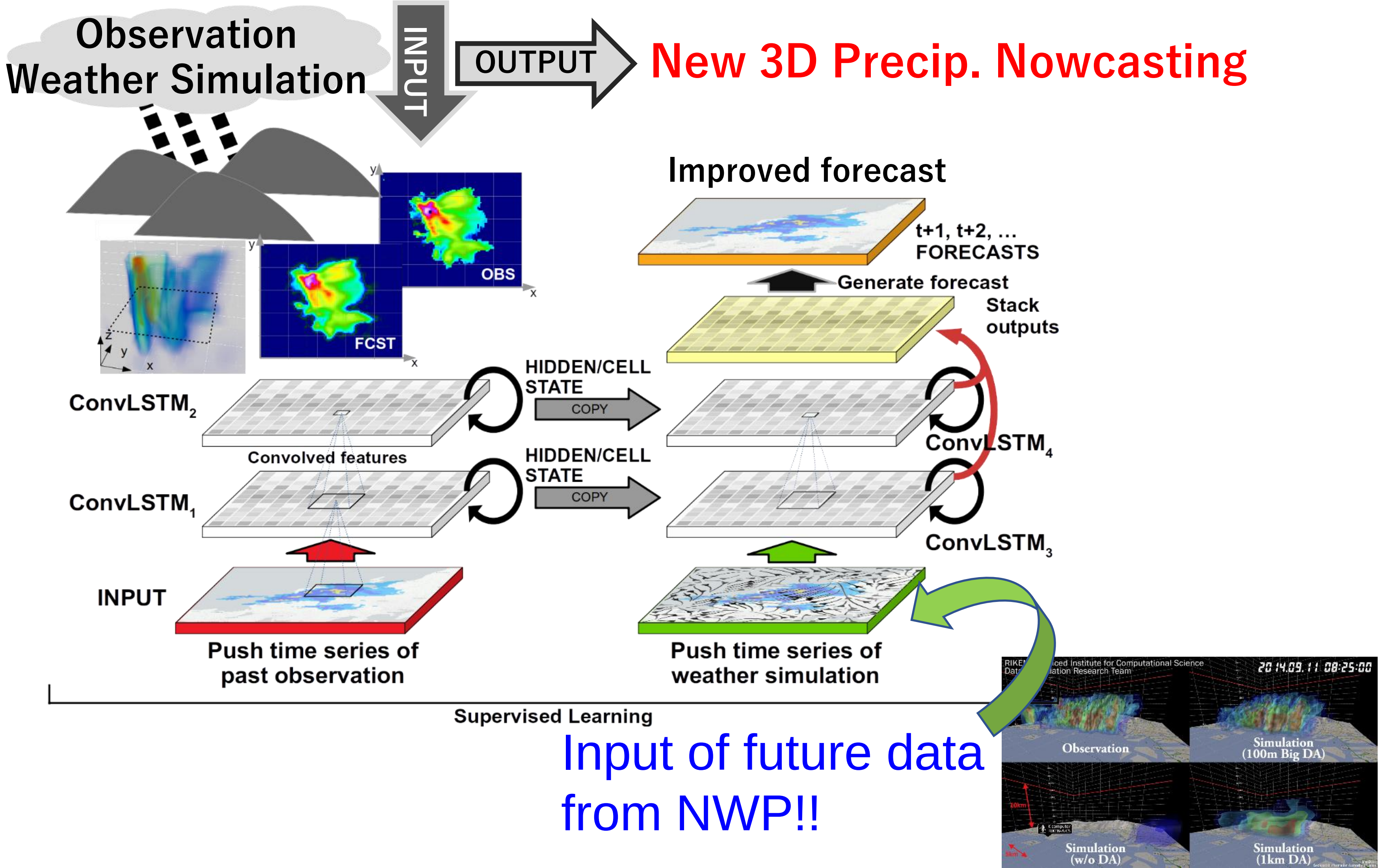
LSTM(Obs. + Kinematic)

LSTM(Obs. only)

Kinematic

Persistence

Future direction: **Fusing ML+DA+Simulation**



Summary

- Phased-Array Weather Radar 3D nowcasting
  - TREC-based algorithm + Data assimilation
  - Operated since July 2017

- Phased-Array Weather Radar + Machine Learning
  - ConvLSTM3D captured rapid change of radar reflectivity
    - Difficult for kinematic nowcasting
  - Ingesting kinematic nowcast / numerical weather prediction into ConvLSTM was tested
  - Future
    - More samples, training for different seasons, online training