A Machine Learning Approach to the Observation Operator for Satellite Radiance Data Assimilation

[1]Jianyu Liang

[1, 2] Koji Terasaki

[1]Takemasa Miyoshi

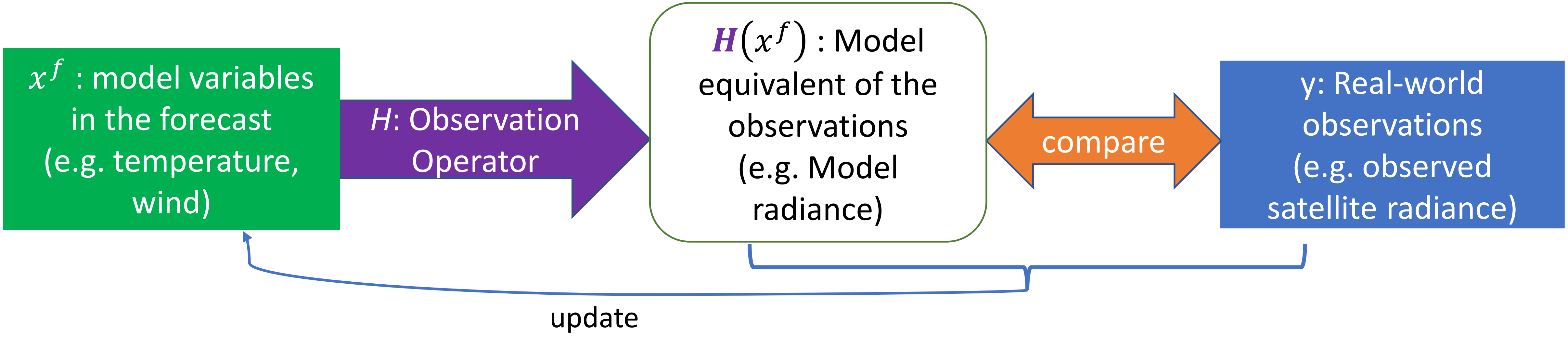
[1]Data Assimilation Research Team, RIKEN Center for Computational Science

[2] Meteorological Research Institute, Japan Meteorological Agency

ISDA Bologna 2023

Introduction

$$x^a = x^f + K[y - H(x^f)]$$



Two types of observation operators

(1) Physically-based OO (P-OO)

understand physics

DA: x^f : model variables in the forecast (e.g. temperature, wind) → *H*: Physical model physical relationship → $H(x^f)$: Model equivalent of the observations (e.g. Model radiance) ↔ compare ↔ y: Real-world observations (e.g. observed satellite radiance)

Issues: (1) P-OO may be time-consuming to develop; (2) bias correction may be needed.

Our goal: To accelerate the usage of new observations, build OO without physically-based model.

(2) Machine learning (ML) observation operator (ML-OO)

$$y = H(x^{true}) + \varepsilon$$

approximate

x^a and x^f

model variables in the Analyses or in the short-term forecasts after the analyses

y: Observations

Training

DA: x^f : model variables in the forecast (e.g. temperature, wind) → *H*: ML model statistical relationship → $H(x^f)$: Model equivalent of the observations (e.g. Model radiance) ↔ compare ↔ y: Real-world observations (e.g. observed satellite radiance)

Observations

(1) Conventional observations: NCEP PREPBUFR

(2) Satellite observations : Brightness temperatures (BT) from the Advanced Microwave Sounding Unit (AMSU-A)

standard deviation of the observation error of AMSU-A (unit: K)

Channel	NOAA-15	NOAA-18	NOAA-19	METOP-A	METO-B
6		0.3	0.3	0.3	0.3
7	0.3	0.3	0.3	0.3	
8	0.3	0.3		0.3	0.3

Pre-process AMSU-A observations

- Data thinning: 250 km. reduce spatial correlations
- Quality control: reduce the impact from cloud and rain. For example, channels 6 and 7 over the land were completely filtered out.
- Gross error checks: filtered out large departure data

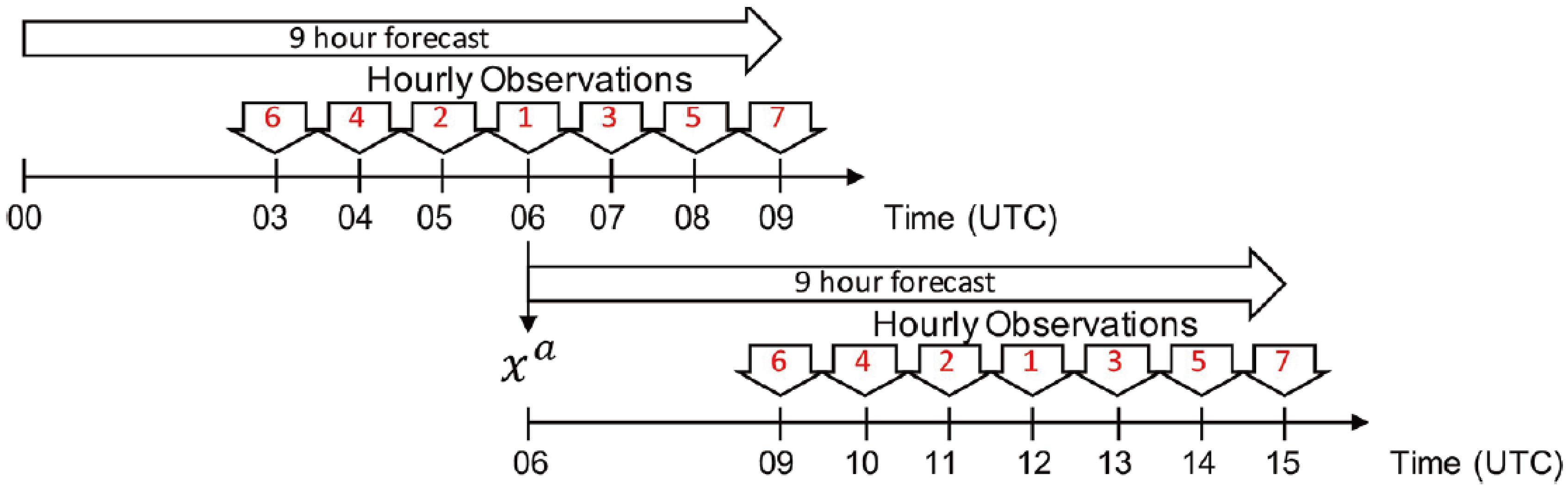
Data assimilation system

Model: Non-hydrostatic icosahedral atmospheric model (NICAM)

- 112 km horizontal resolution and 78 vertical levels up to 40 hPa

DA method: Local ensemble transform Kalman filter (LETKF)

- Assimilate conventional observations and satellite radiance every 6 hours
- 64 members
- The relaxation to prior spread (RTPS) is used for covariance inflation
- **No vertical localization** for AMSU-A brightness temperature



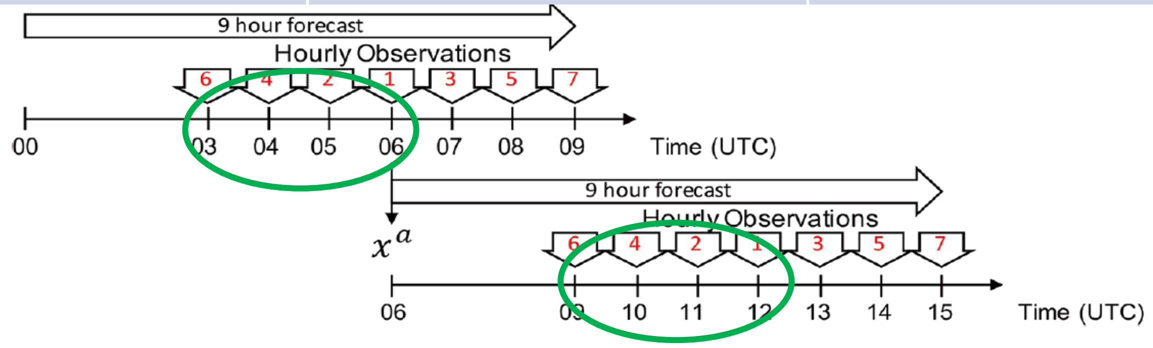
Experiments

	Training experiments		Testing experiments	
	Jan 2015 (DA spin-up)	Feb 2015 (DA cycle)	Jan 2016 (DA spin-up)	Feb 2016 (DA cycle)
Previous work Liang et al. 2023 JMSJ)	CONV+AMSUA (RTTOV)		CONV+AMSUA (RTTOV)	CONV+AMSUA (RTTOV) + bias correction CONV+AMSUA (ML) CONV
Current work	CONV		CONV	CONV + AMSUA(RTTOV) + bias correction
		(1) Interpolate the model to the AMSU-A locations (2) train the ML model		CONV + AMSUA (ML)
				CONV

$$y = H(x^{true}) + \varepsilon$$

approximate

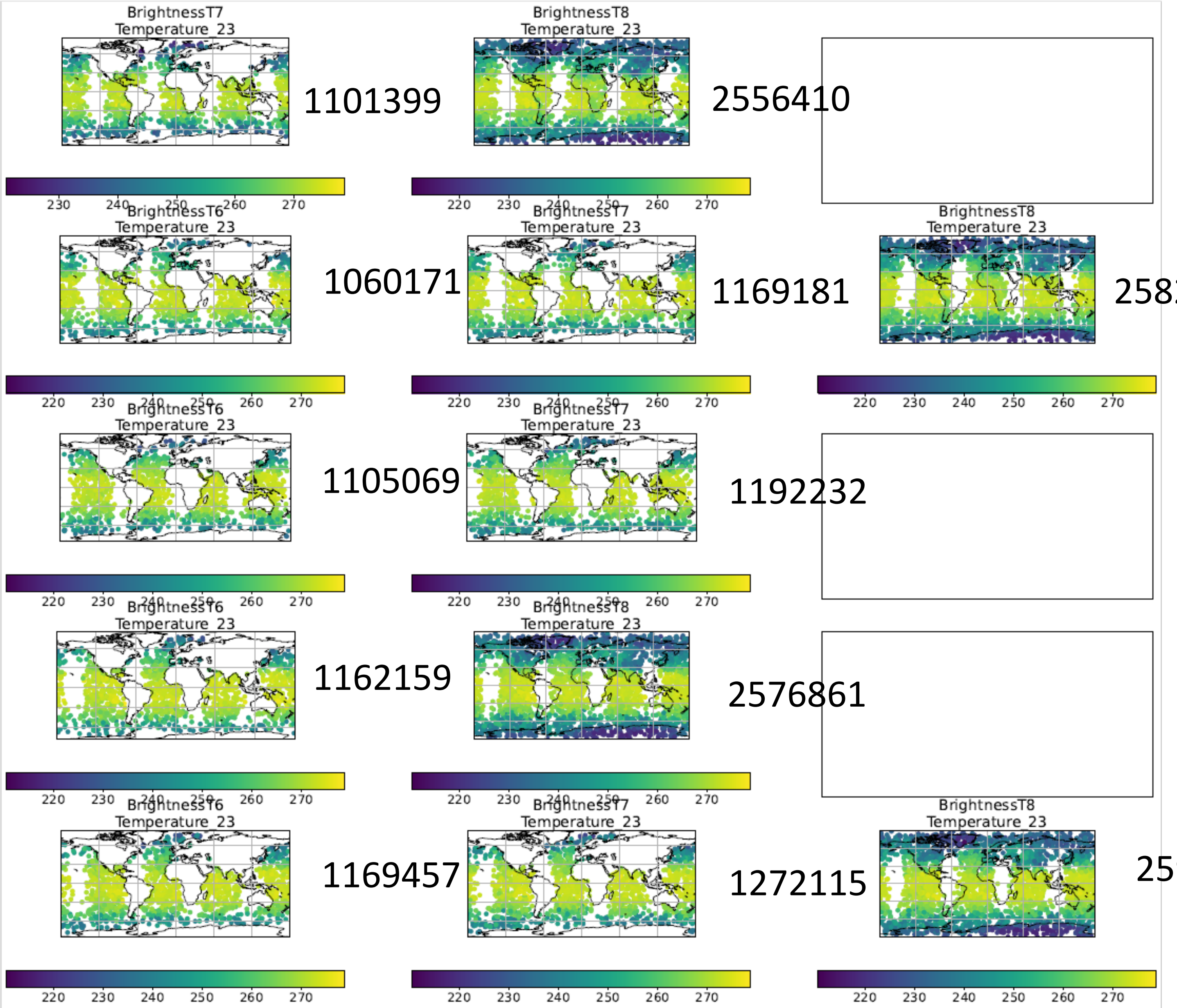
x^a and x^f



Observation coverage in the training data

Interpolate the model to the AMSU-A locations

temperature at 500 hPa



NOAA15 1101399 2556410

NOAA18 1060171 1169181 2582306

NOAA19 1105069 1192232

METOP-B 1162159 2576861

METOP-A 1169457 1272115 2592228

Design of the machine learning models

$$y = h_{ml}(x_b, \theta, \varnothing, p) + e_{ml}$$

output input

input

3D variables:

pressure (78 levels)

temperature (78 levels)

specific humidity (40 levels)

2D variables:

surface pressure

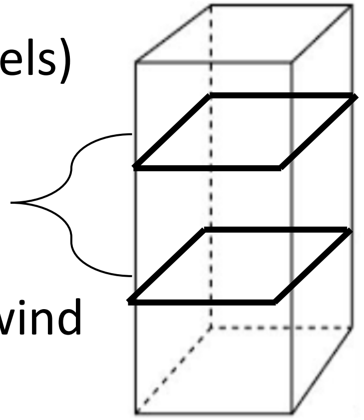
surface temperature

10-meter u-wind and v-wind

2-meter temperature

2-m specific humidity

other bias predictor: Satellite zenith angle, Scan angle, latitude

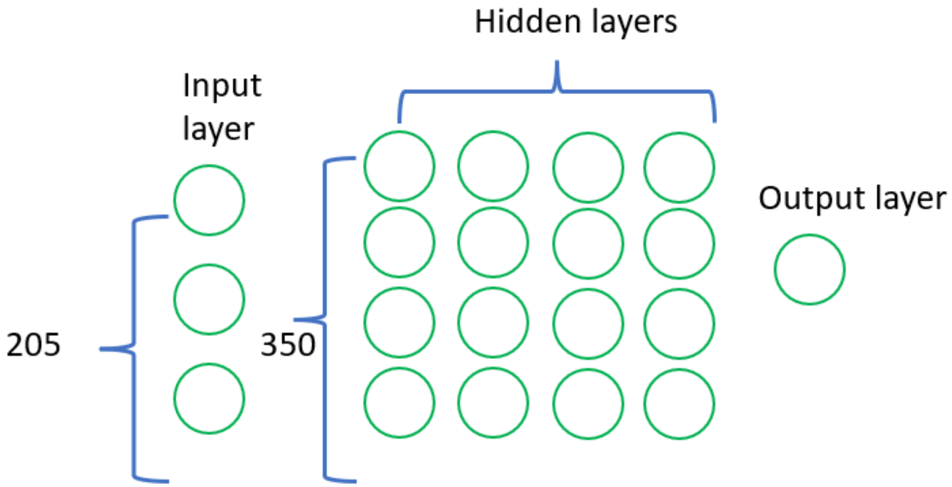


Output: y

satellite brightness temperature from channel 6, 7, 8

Deep neural networks (DNNs) for each channel and satellite

205 features, includes different vertical levels



Hyperparameters searching:

		This presentation
activation function	ReLu	ReLu
learning rate	1e-05, 1e-6	1e-6
Units	250~400	300
Layers	2, 3, 4	4

Can ML model handles bias ?

$$y = H(x^{true}) + \varepsilon$$

$$\underset{\text{obs}}{y} = h_{ml}(x_b, \theta, \varnothing, p) + e_{ml}$$

scan angle

latitude

p: state-dependent predictors

The ML algorithm minimizes the mean squared error (MSE).
The MSE can be decomposed into the variance of the error and the square of the bias

$$E\{[y - h_{ml}(x_b, \theta, \varnothing, p)]^2\}$$
$$= Var[y - h_{ml}(x_b, \theta, \varnothing, p)] + bias^2.$$

Evaluate the machine learning models

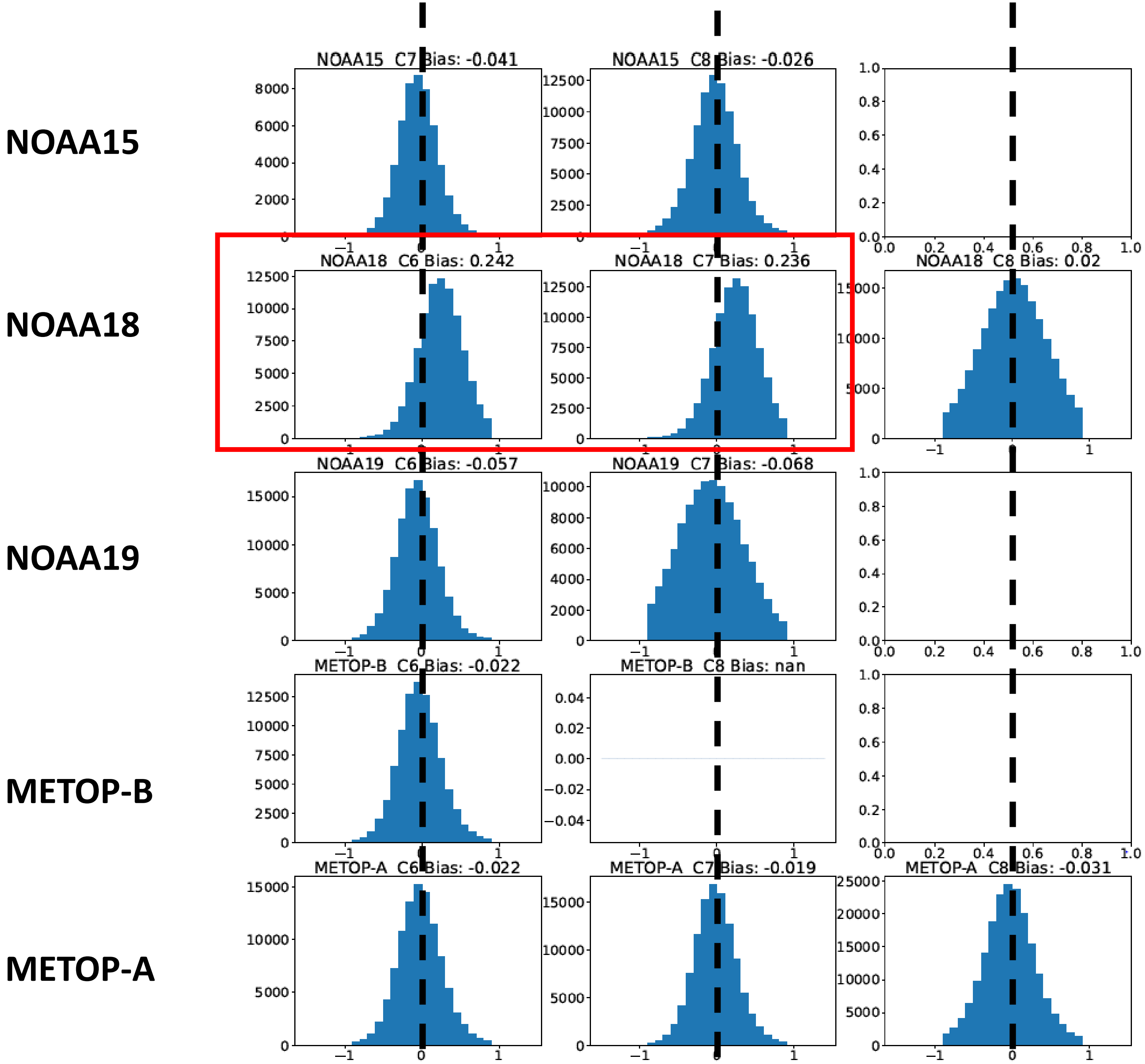
		rmse	bias	r2
NOAA-15	C7	0.205199	-0.00424	0.997418
	C8	0.236601	-0.00244	0.998823
NOAA-18	C6	0.205312	-0.00277	0.998464
	C7	0.22366	-0.00265	0.997381
	C8	0.301168	-0.00483	0.998129
NOAA-19	C6	0.187959	-0.00055	0.998679
	C7	0.371082	0.004732	0.992673
METOP-B	C6	0.195633	0.001825	0.9986
	C8	0.984203	-0.00397	0.980018
METOP-A	C6	0.189656	0.003823	0.998689
	C7	0.204728	0.001413	0.99776
	C8	0.261701	-0.00208	0.998571

- use the test set (20% of all data) to evaluate the performances.
- RMSE and bias **are small;** coefficient of determination (R^2) is **high**.
- ML for one channel is not good

Experiments

	Training experiments		Testing experiments	
	Jan 2015 (DA spin-up)	Feb 2015 (DA cycle)	Jan 2016 (DA spin-up)	Feb 2016 (DA cycle)
Previous work Liang et al. 2023 JMSJ)	CONV+AMSUA (RTTOV)		CONV+AMSUA (RTTOV)	CONV+AMSUA (RTTOV) + bias correction CONV+AMSUA (ML) CONV
Current work	CONV		CONV	CONV + AMSUA(RTTOV) + bias correction
		(1) Interpolate the model to the AMSU-A locations (2) train the ML model		CONV + AMSUA (ML)
				CONV

Bias



O-B

$y - H(x^f)$

ML-OO was based on data from February 2015 and could not treat the bias well in February 2016 **with a significant change in satellite characteristics**.

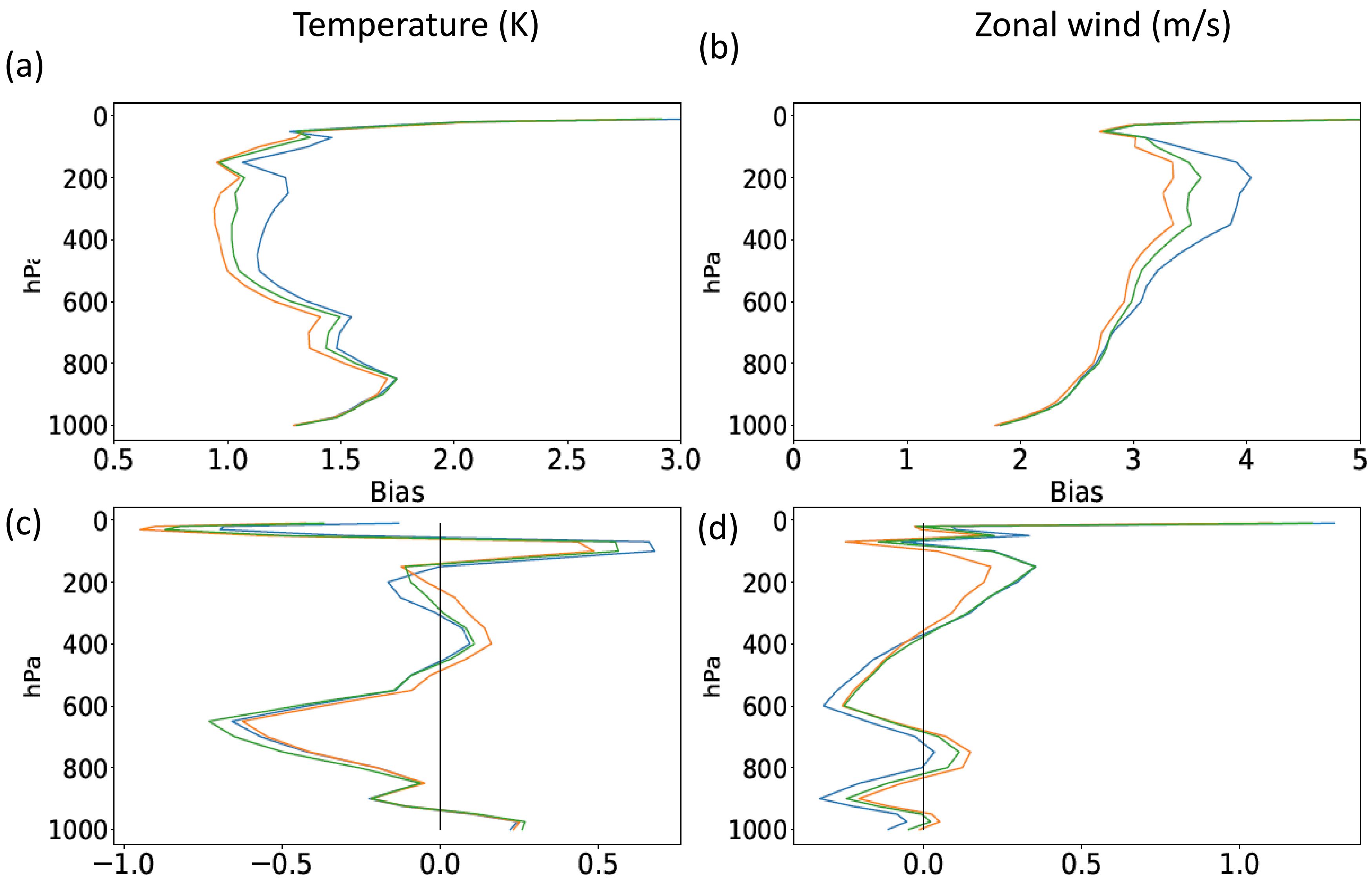
Evaluate ML-OO
compare to ERA Interim

1 month average

CONV
CONV+AMSUA (RTTOV)
CONV+AMSUA (ML)

RMSD

Bias



- ML-OO (green line) is slightly worse than RTTOV-OO (orange line) but better than only assimilating conventional observations (blue line).

Summary

1. Model forecasts and AMSU-A observations were used to train neural networks for generating ML-OO
2. Assimilating additional AMSU-A observations using ML-OO shows better performance compared to assimilating only conventional observations, although its performance is slightly inferior than that of RTTOV.
3. A **separate bias correction procedure may not needed** if no significant change in satellite characteristics.
4. This data-driven approach **enables earlier assimilation** of satellite data following the launch of a new satellite

Future works

1. Evaluate the performance of ML-OO during different time periods.
2. Evaluate the feature importance.
3. Examine the level of accuracy necessary in the model background to effectively train a robust ML-OO.

Reference

Liang, J., T. Koji, and M. Takemasa, 2023: A Machine Learning Approach to the Observation Operator for Satellite Radiance Data Assimilation. JMSJ, 101, 79–95, https://doi.org/10.2151/jmsj.2023-005.

Poster: Understanding the Dynamics of Venus Atmosphere using Bred Vectors
Jianyu Liang (RIKEN R-CCS), Norihiko Sugimoto (Keio University), Takemasa Miyoshi (RIKEN R-CCS)

email: jianyu.liang@riken.jp