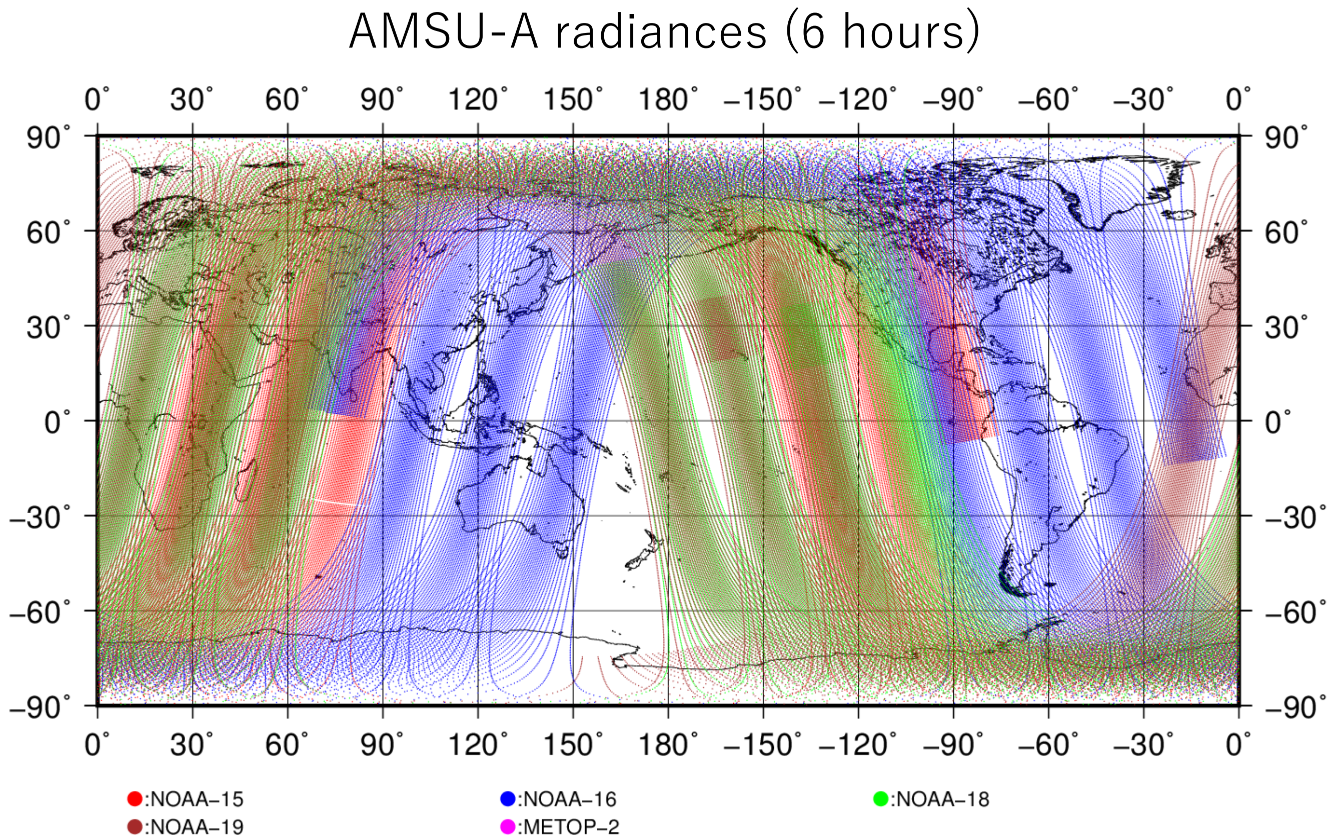
Including the horizontal observation error correlation in the assimilation of AMSU-A data

Koji Terasaki and Takemasa Miyoshi
RIKEN R-CCS Data Assimilation Research Team

ECMWF/EUMETSAT NWP SAF Workshop on the treatment of random and systematic errors in satellite data assimilation for NWP

Observation error correlations

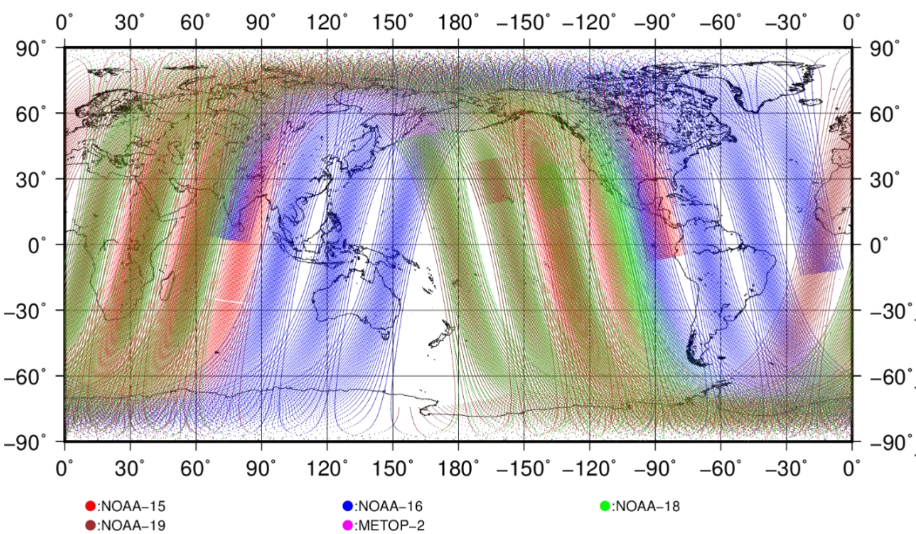
- Observations measured with the same instrument are known to have **error correlations**.
  - e.g., Satellite radiances, Atmospheric motion vector, Doppler radar



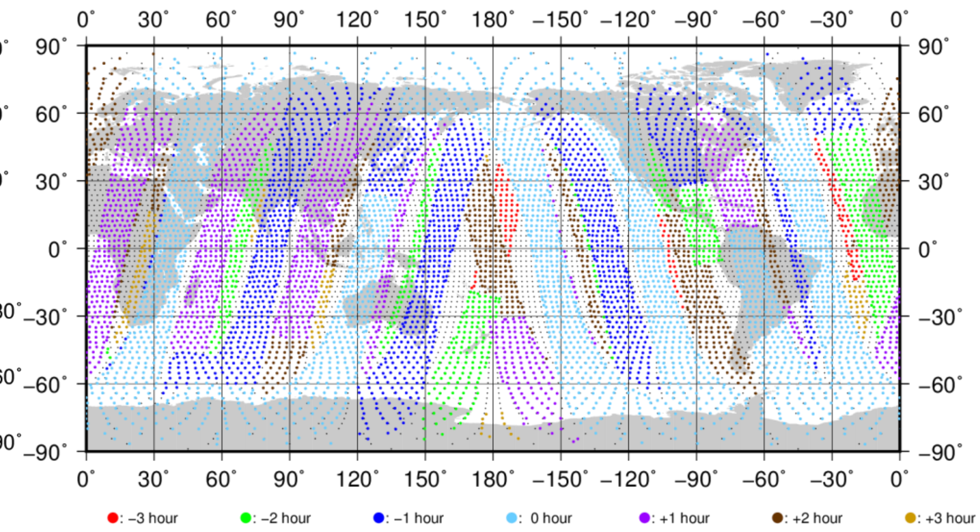
Observation error correlations

- There are some studies to estimate the **<u>horizontal observation error correlations</u>**, but not used in data assimilation (DA). We usually **<u>thin</u>** the horizontally dense observations and assume no-error correlations in DA.

Before thinning ($\approx$ **400,000 points**)



After thinning (dx=250km, $\approx$ **7,000 points**)

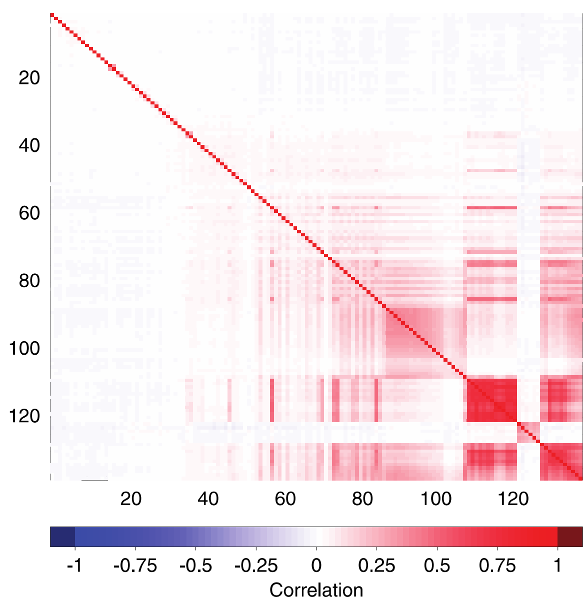


Observation error correlations

- There are some studies to estimate the **<u>horizontal observation error correlations</u>**, but not used in data assimilation (DA). We usually **<u>thin</u>** the horizontally dense observations and assume no-error correlations in DA.

- Accounting for the **<u>inter-channel (vertical)</u>** observation error correlation will improve the analysis and forecast. (e.g., Weston et al. 2014, Bormann et al. 2016, Campbell et al. 2017)

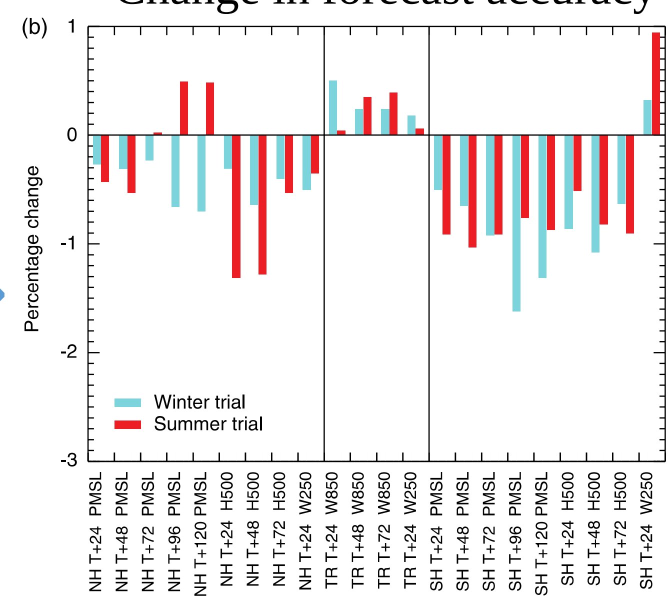
Estimated error correlation matrix (IASI)



(Fig. 10 of Weston et al. 2014)

Improved

Change in forecast accuracy



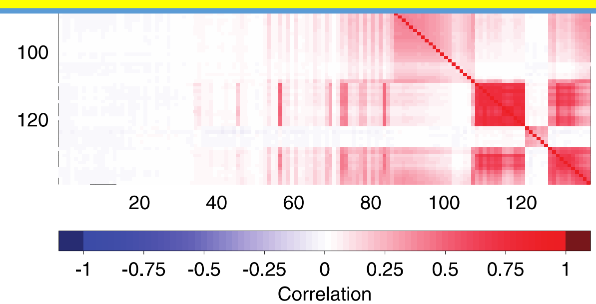
(Fig. 11 of Weston et al. 2014)

Observation error correlations

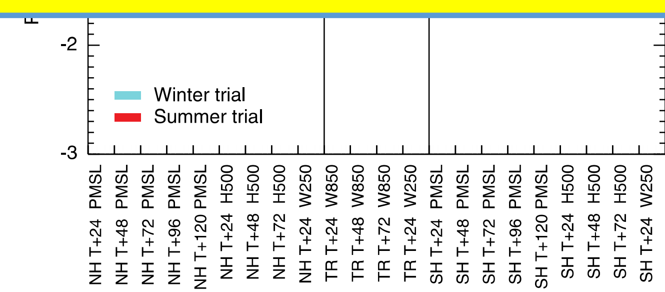
- There are some studies to estimate the **<u>horizontal observation error correlations</u>**, but not used in data assimilation (DA). We usually **<u>thin</u>** the horizontally dense observations and assume no-error correlations in DA.

- Accounting for the **<u>inter-channel (vertical)</u>** observation error correlation will improve the analysis and forecast. (e.g., Weston et al. 2014, Bormann et al. 2016, Campbell et al. 2017)

<u>Goal</u> is to investigate how to effectively utilize dense observations in horizontal by including the **<u>horizontal observation error correlations</u>** in DA and improve the weather forecast.



(Fig. 10 of Weston et al. 2014)



(Fig. 11 of Weston et al. 2014)

Local ensemble transform Kalman filter

- **Analysis Equation for LETKF**

Observation error covariance matrix

$$\boxed{\mathbf{x}_a} = \boxed{\bar{\mathbf{x}}_f} + \boxed{d\mathbf{x}_f\left\{\mathbf{U}\mathbf{D}^{-1}\mathbf{U}^{\mathrm{T}}\left(\mathbf{H}d\mathbf{x}_f\right)^{\mathrm{T}}\boxed{\mathbf{R}^{-1}}\left(\mathbf{y}-\overline{\mathbf{H}\mathbf{x}_f}\right)+\sqrt{m-1}\,\mathbf{U}\mathbf{D}^{-1/2}\mathbf{U}^{\mathrm{T}}\right\}}$$

Analysis　Ens. mean（FG）　Analysis Increment

y: observation, **x:** state variable,
H: observation operator,
a: analysis, f: forecast

Eigenvalue decomposition

$$\mathbf{U}\mathbf{D}\mathbf{U}^{\mathrm{T}} = (m-1)\mathbf{I} + \left(\mathbf{H}d\mathbf{x}_f\right)^{T}\mathbf{R}^{-1}\left(\mathbf{H}d\mathbf{x}_f\right)$$

- Assuming diagonal **R**

＜Merit＞
- Low computational cost for inverting **R**

＜Demerit＞
- Need to thin the observations （in spatial and between channels）

$$\mathbf{R} = \begin{pmatrix} 2 & 0.8 & 0.6 & 0.1 & 0.05 & 0.4 \\ 0.8 & 1.5 & 0.4 & 0.2 & 0.1 & 0.7 \\ 0.6 & 0.4 & 3 & 0.5 & 0.3 & 0.4 \\ 0.1 & 0.2 & 0.5 & 2 & 0.2 & 0.1 \\ 0.05 & 0.1 & 0.3 & 0.2 & 1.2 & 0.7 \\ 0.4 & 0.7 & 0.4 & 0.1 & 0.7 & 4 \end{pmatrix}$$

Observation error

Error correlations

Accounting for the OCEs in LETKF
Require $\mathbf{R}^{-1}$ → High computational cost
Unstable due to a **high condition number** → Stabilized by **reconditioning**
（Condition number: ratio between the largest and smallest eigenvalues）

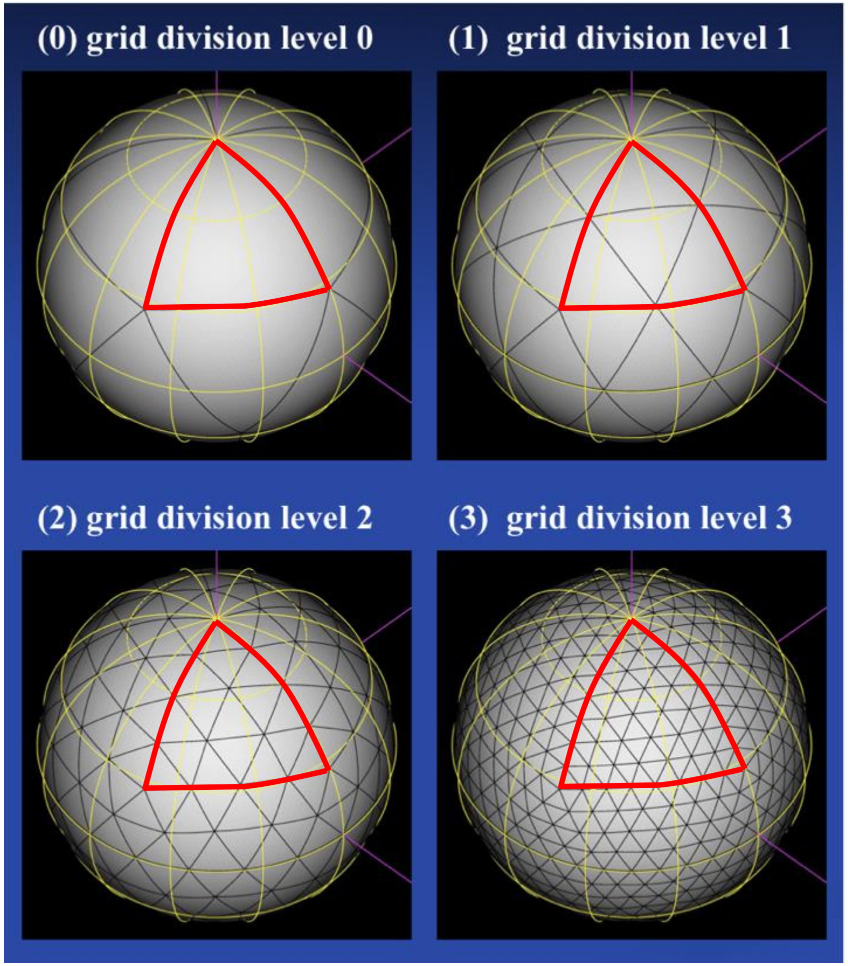
Idealized experiment

NICAM: Nonhydrostatic Icosahedral Atmospheric Model

Computer simulations create the future

Grid division level 0 is the original Icosahedron.
The horizontal resolution can be increased by splitting one triangle into four triangles.

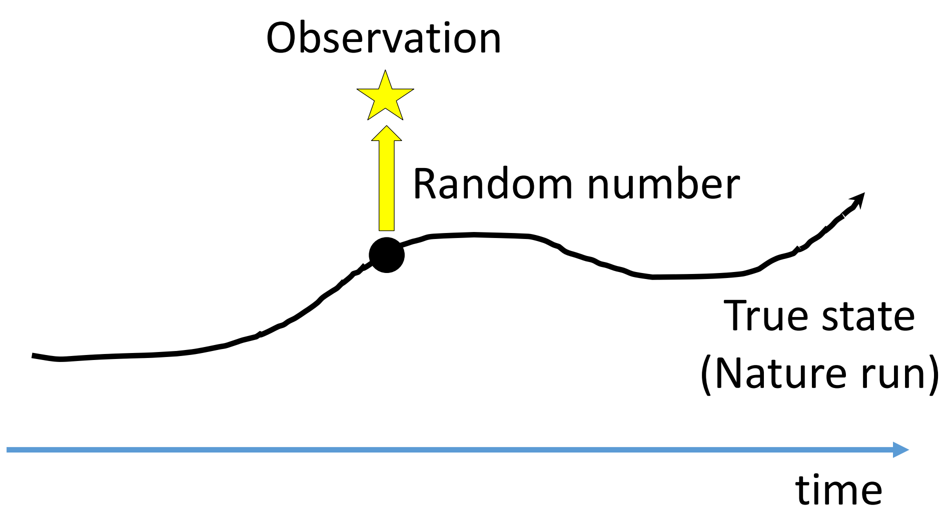
Grid division level	Horizontal resolution
6	112 km
7	56 km
8	28 km
9	14 km
10	7 km
11	3.5 km
12	1.7 km
13	0.87 km



Idealized experiment with NICAM-LETKF

Horizontal resolution:	Glevel-6 (112km)
Vertical resolution:	38 layers (model top = 40km)
Ensemble size:	40
Period:	2 months (From 2012/1/1/00Z - 2012/2/29/18Z)

Observing System Simulation Experiment (OSSE)



- Error-correlated random numbers

$$\mathbf{R} = \mathbf{C}\mathbf{C}^T$$

$$\boldsymbol{\varepsilon} = \mathbf{C}\boldsymbol{\mu}$$

R: Observation error covariance matrix
C: Cholesky decomposition of **R**
μ: Independent random numbers
ε: Error-correlated random numbers

Idealized experiment with NICAM-LETKF

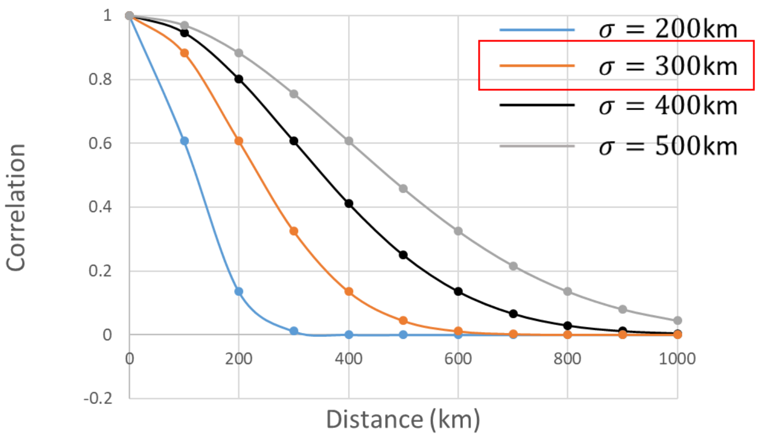
Computer simulations create the future

- **Simulated observations with dx=150km**
  - Error standard deviations
    - T = 2 (K), U & V = 4 (m/s)
  - Error correlations
    - 15 pressure levels
    - No error-correlation in different levels
    - **Condition number > 10^{10}**

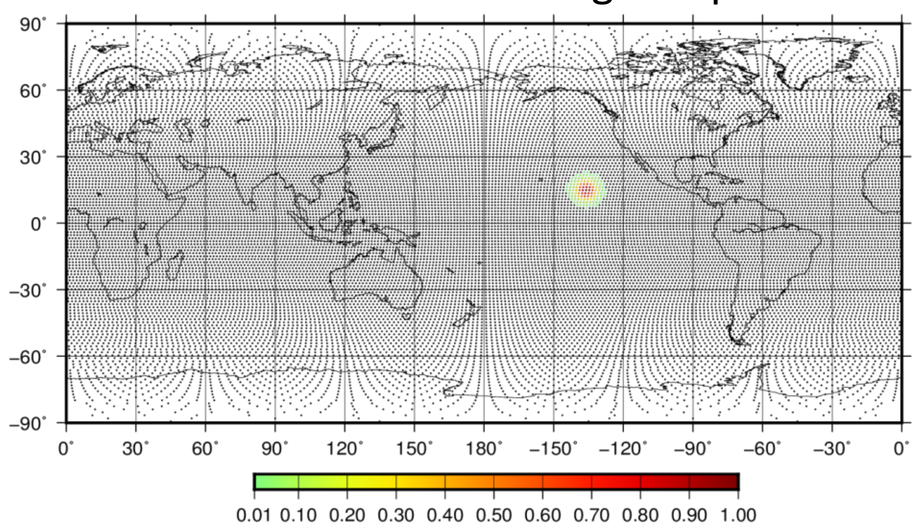
Error correlations

✓Gaussian Function

$$(R_{ij}) = \exp(-\frac{1}{2}(\frac{d(i,j)}{\sigma})^2)$$

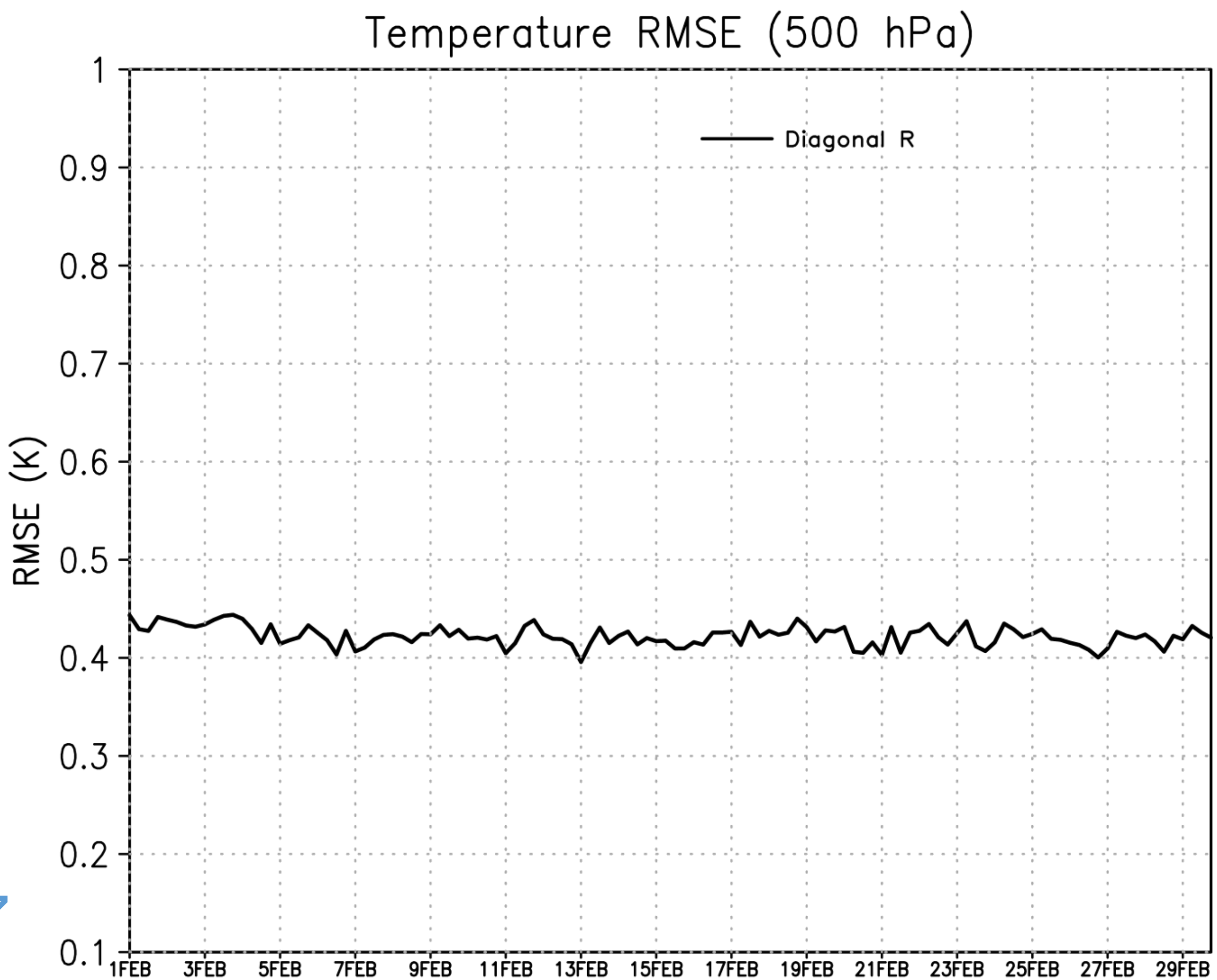


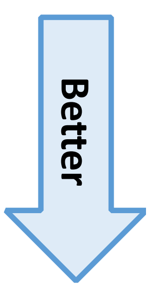
Observation coverage map



Error correlation for the observation located at 136.047°W and 14.887°N

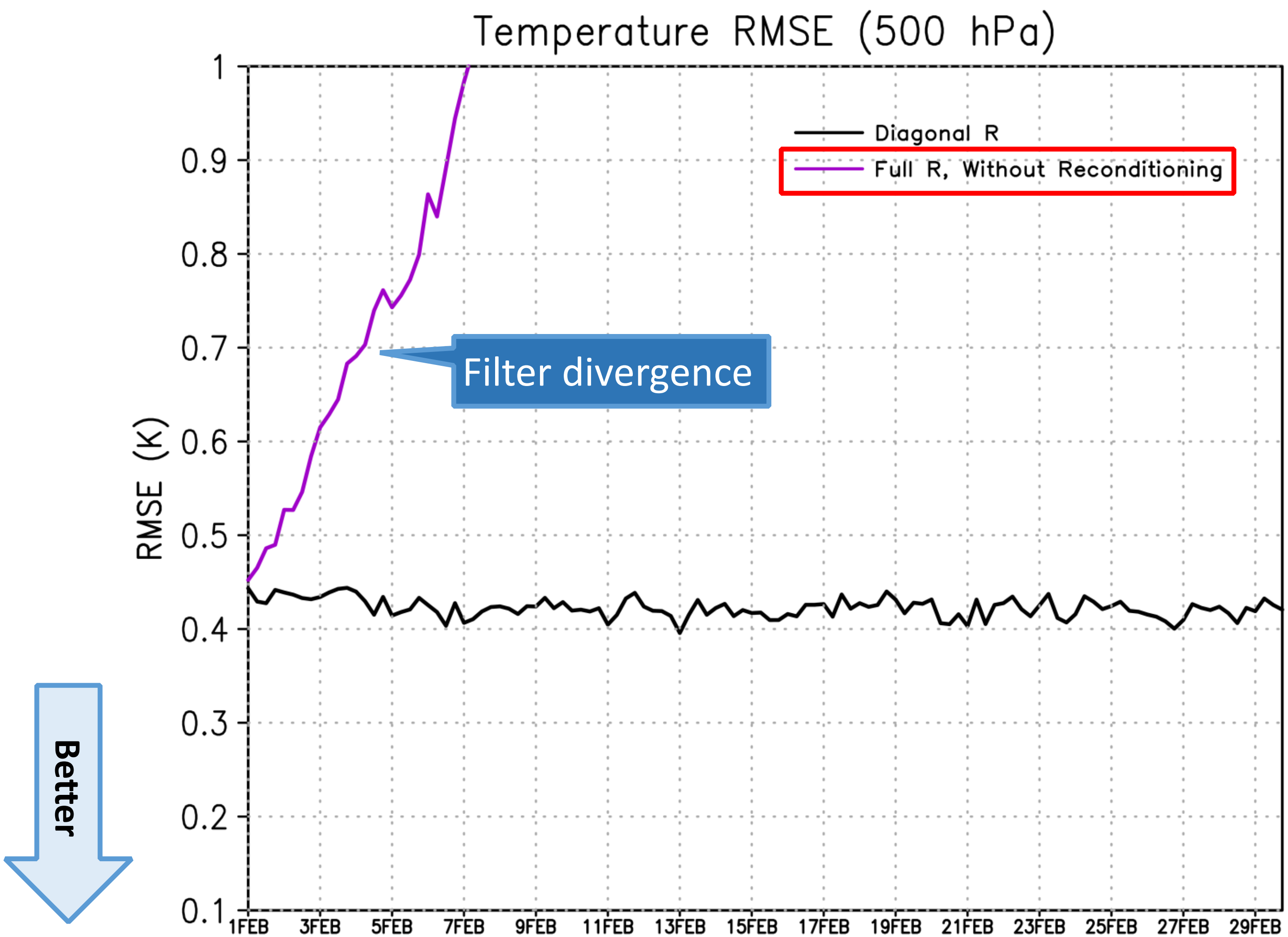
Analysis RMSE (Temperature)





Analysis RMSE (Temperature)

Computer simulations create the future



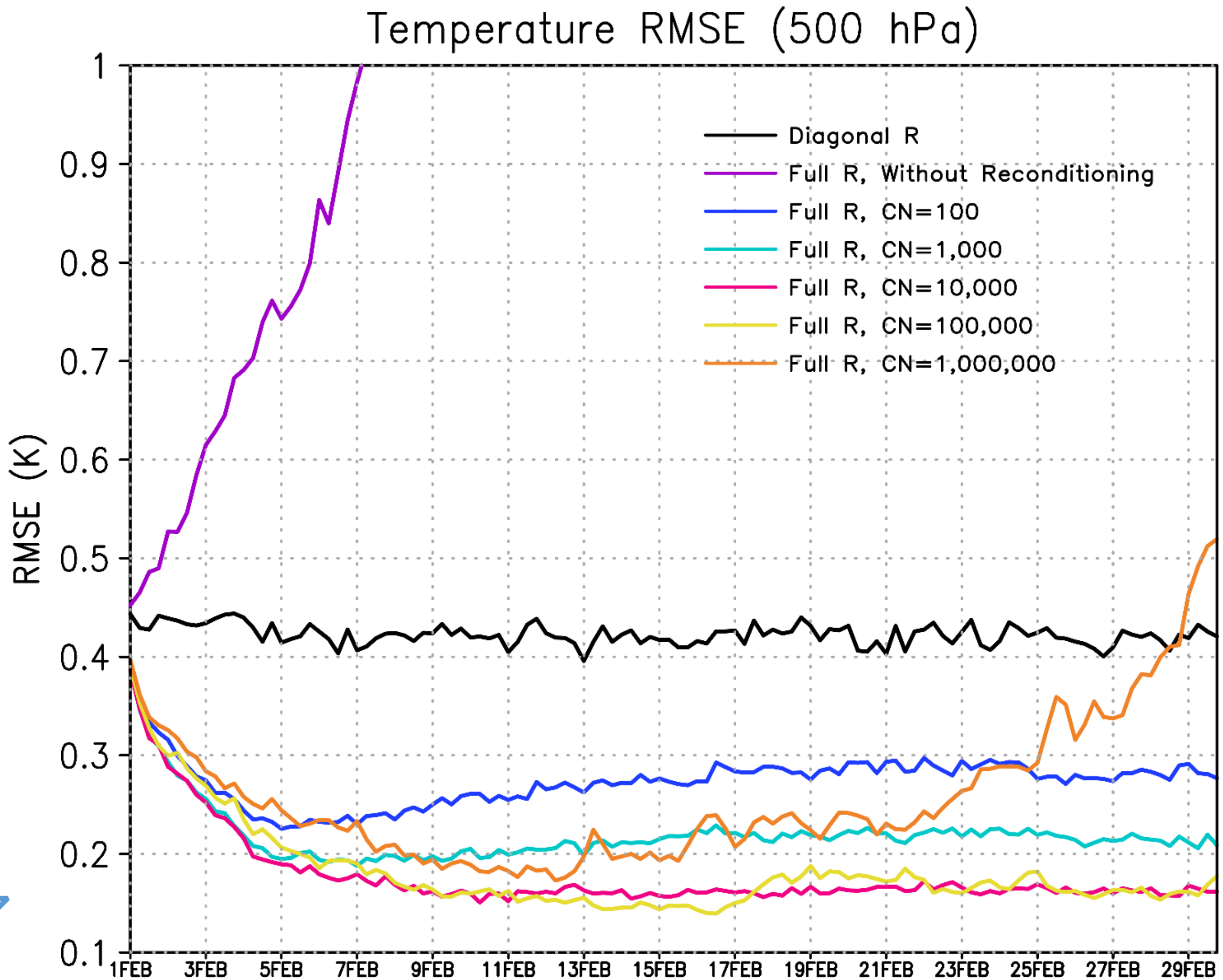
Including full R without any reconditioning → Filter divergence

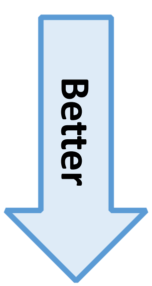
Reducing the condition number by reconditioning

- **The purpose of reconditioning is to stabilize the LETKF by reducing the condition number of the R matrix (cf. Weston et al. 2014 for 4D-VAR)**
  - $\lambda_{inc} = \frac{\lambda_{max} - \lambda_{min}\kappa_{req}}{\kappa_{req} - 1} \approx \frac{\boldsymbol{\lambda_{max}}}{\boldsymbol{\kappa_{req}}}$
    - Friedman et al. (1981)…Estimate the largest eigenvalue of correlation matrix
      - $\lambda_{max} \geq 1 + (n-1)\bar{r}$, $\bar{r}$: average of the non-diagonal components

Analysis RMSE (Temperature)

Computer simulations create the future





The analysis is improved and the best with condition number 10,000 or 100,000.

Experiment with real observations (AMSU-A)

Experimental setting

Horizontal resolution: Glevel-6 (112km)

Vertical resolution: 38 layers (model top = 40km)

Ensemble size: 32

Period: From 2018/6/10/00UTC to 2018/9/1/00UTC

Observations: Conventional observations, **AMSU-A (chs. 6, 7, 8)**

	Observation error correlation	Thinning distance of AMSU-A
DIAG250 (Control experiment)		**250 km**
DIAG125		**125 km**
FULL125	✓	**125 km**

More observation

Full **R**

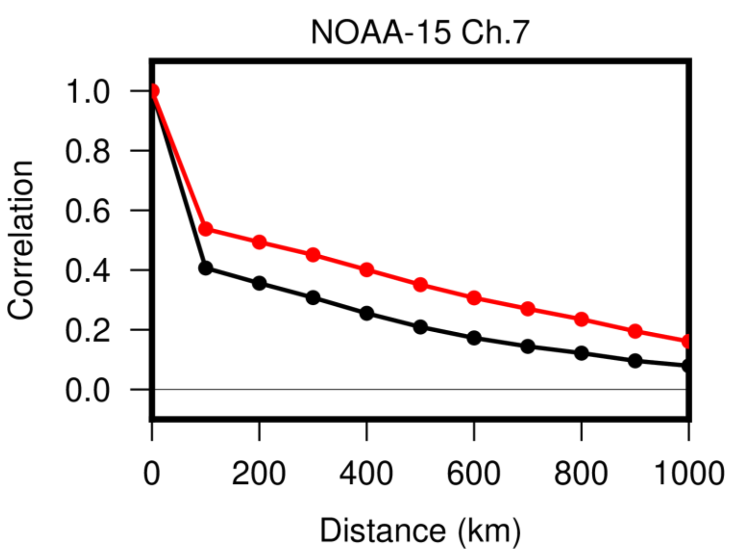
Estimation of R matrix

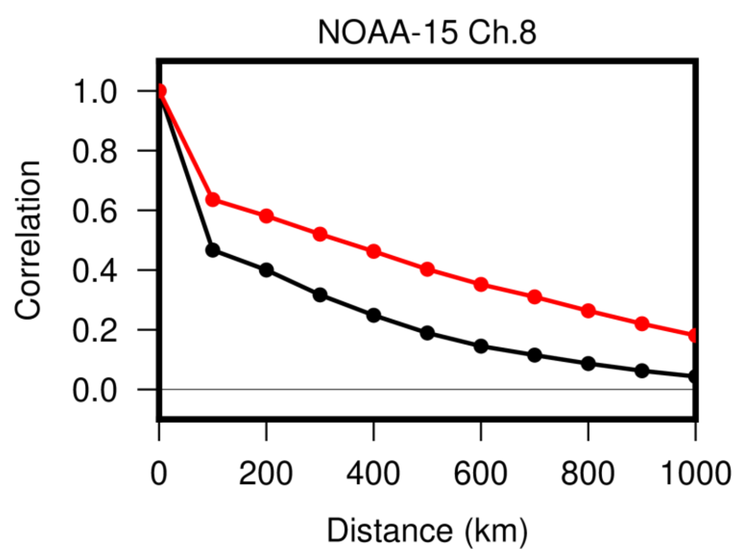
- **R** is estimated using innovation statistics (Desroziers et al. 2005)
  - $< \mathbf{d}_a \mathbf{d}_b^{\mathrm{T}} > = < (\mathbf{y} - \mathbf{H}\mathbf{x}_a)(\mathbf{y} - \mathbf{H}\mathbf{x}_b)^{\mathrm{T}} > = \mathbf{R}$

a: analysis, b: forecast, < >: statistical expectation

- This estimation assumes that appropriate **R** is used in DA.

1. **Estimate R using DIAG125 experiment (Black line)**
2. **Run FULL125_pre experiment using full R (Black line)**
3. **Estimate R using FULL125_pre experiment (Red line) → Run FULL125**





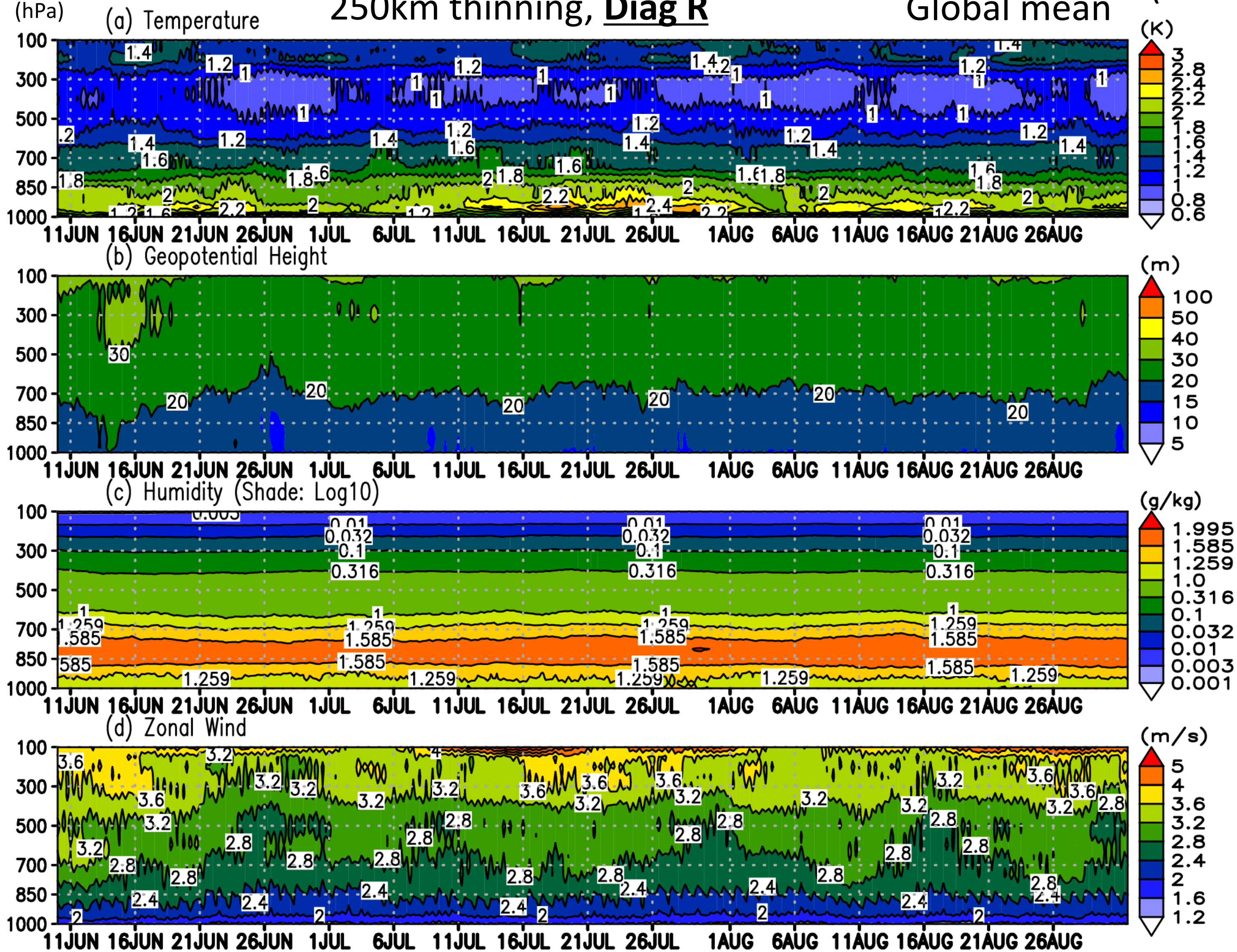
The condition number of **R** is not so large.

Analysis RMSE (DIAG250 : Control experiment)

(vs. ERA-interim)

250km thinning, **Diag R**

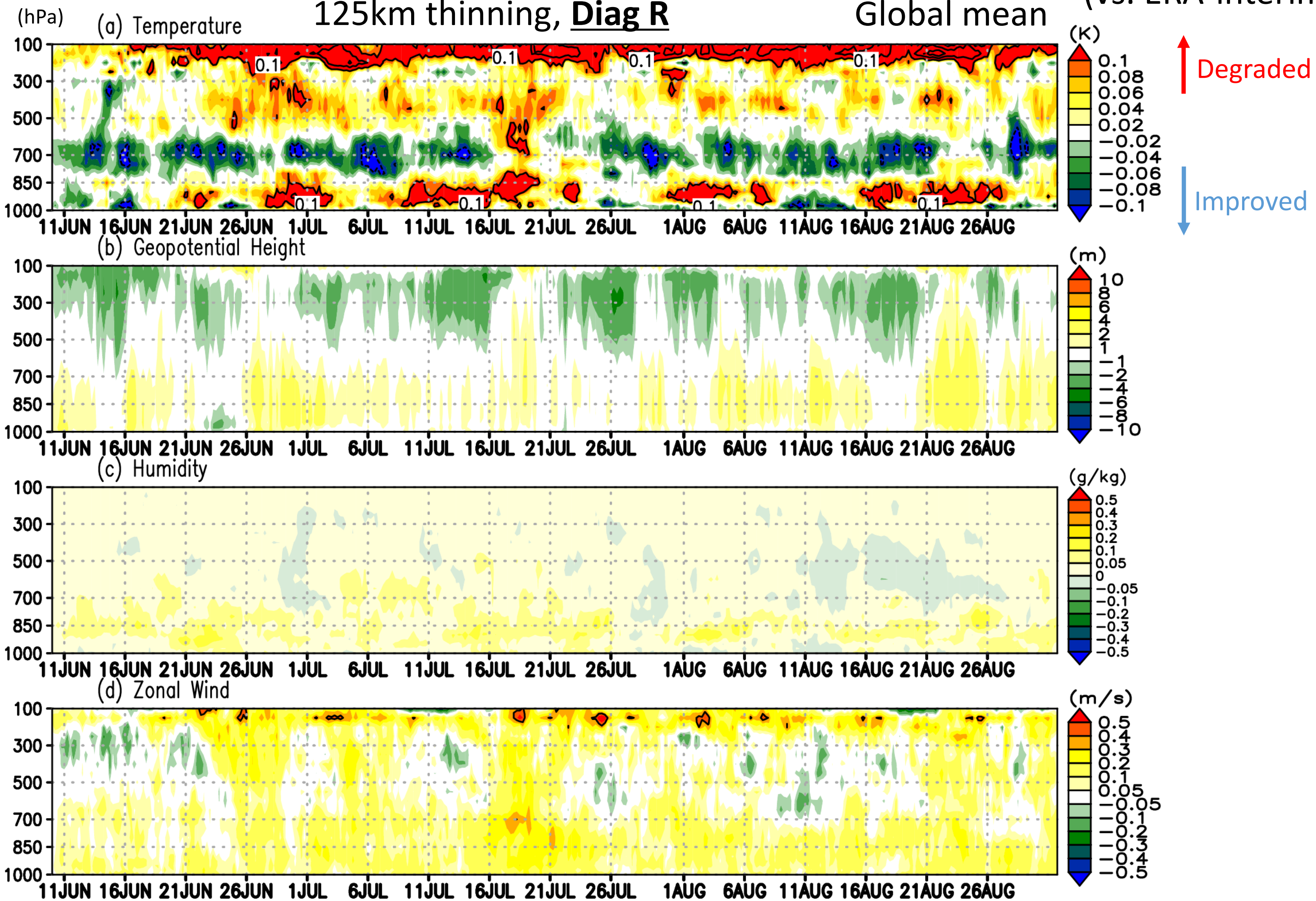
Global mean



Analysis RMSE change (DIAG125 vs DIAG250)

(vs. ERA-interim)

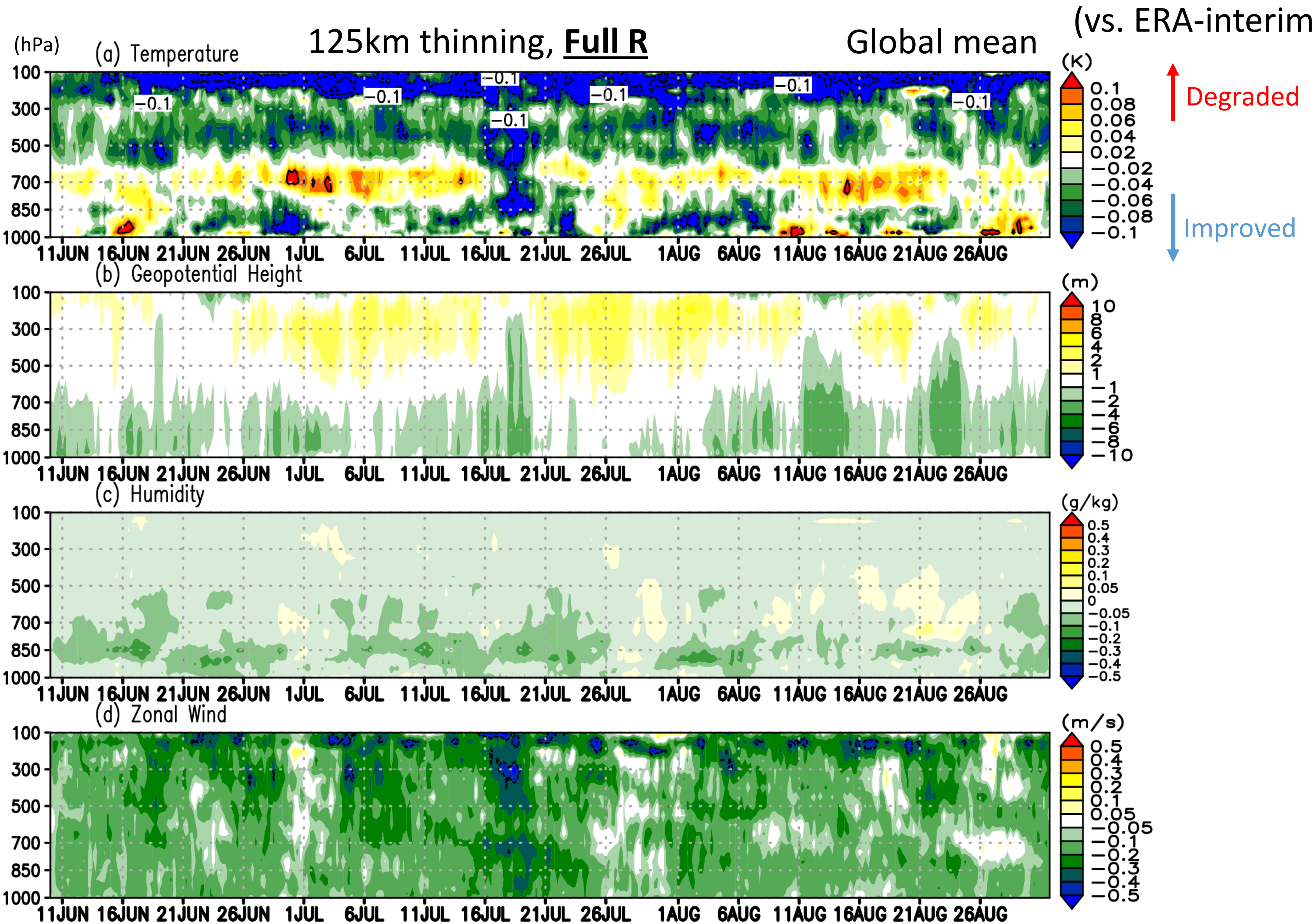
125km thinning, **Diag R**   Global mean



Assimilating dense observations with diagonal **R** makes the analysis **worse**.

Analysis RMSE change (FULL125 vs DIAG125)

Computer simulations create the future

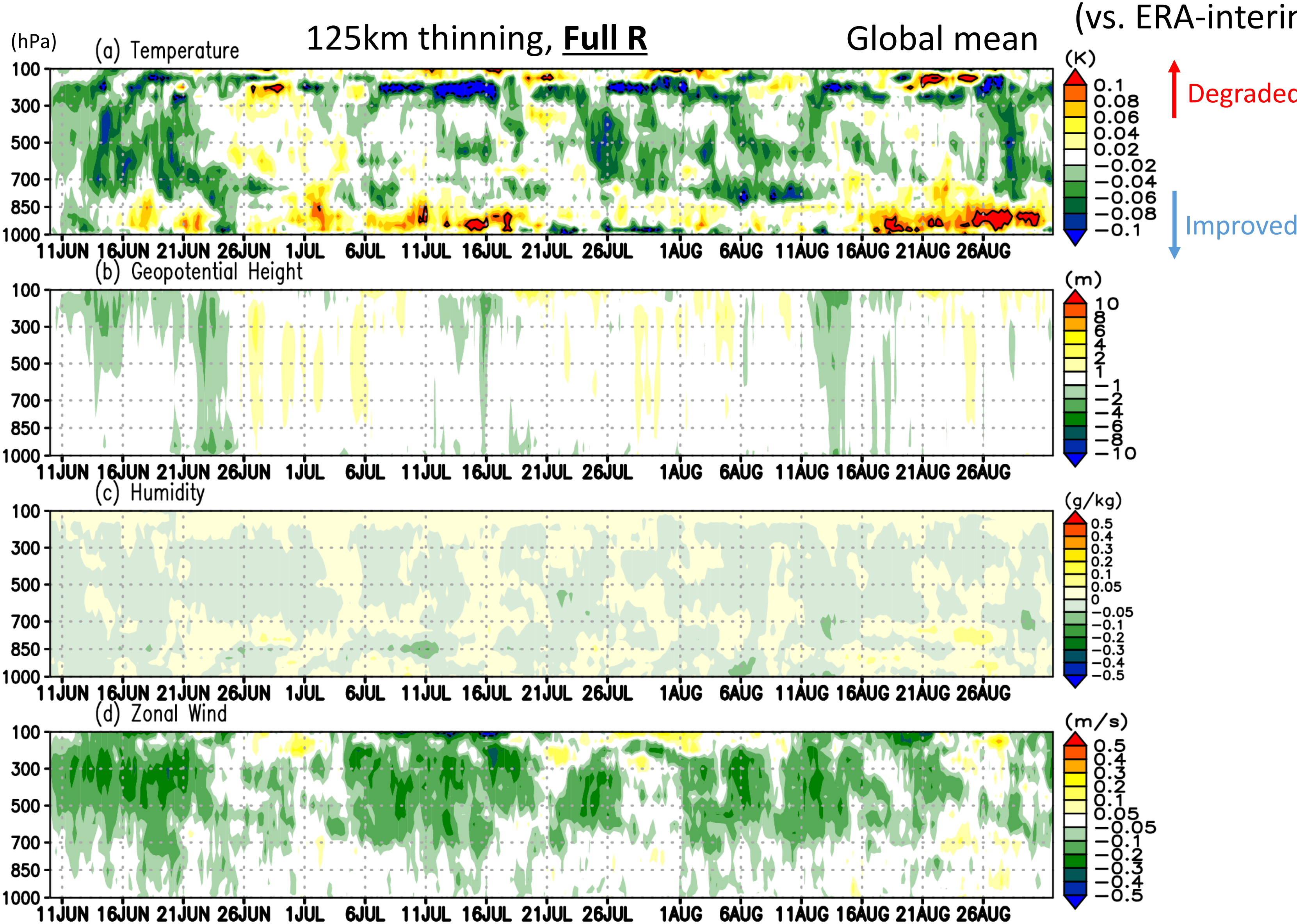


Including full **R** improves the analysis.

Analysis RMSE change (FULL125 vs DIAG250)

Computer simulations create the future

(vs. ERA-interim)



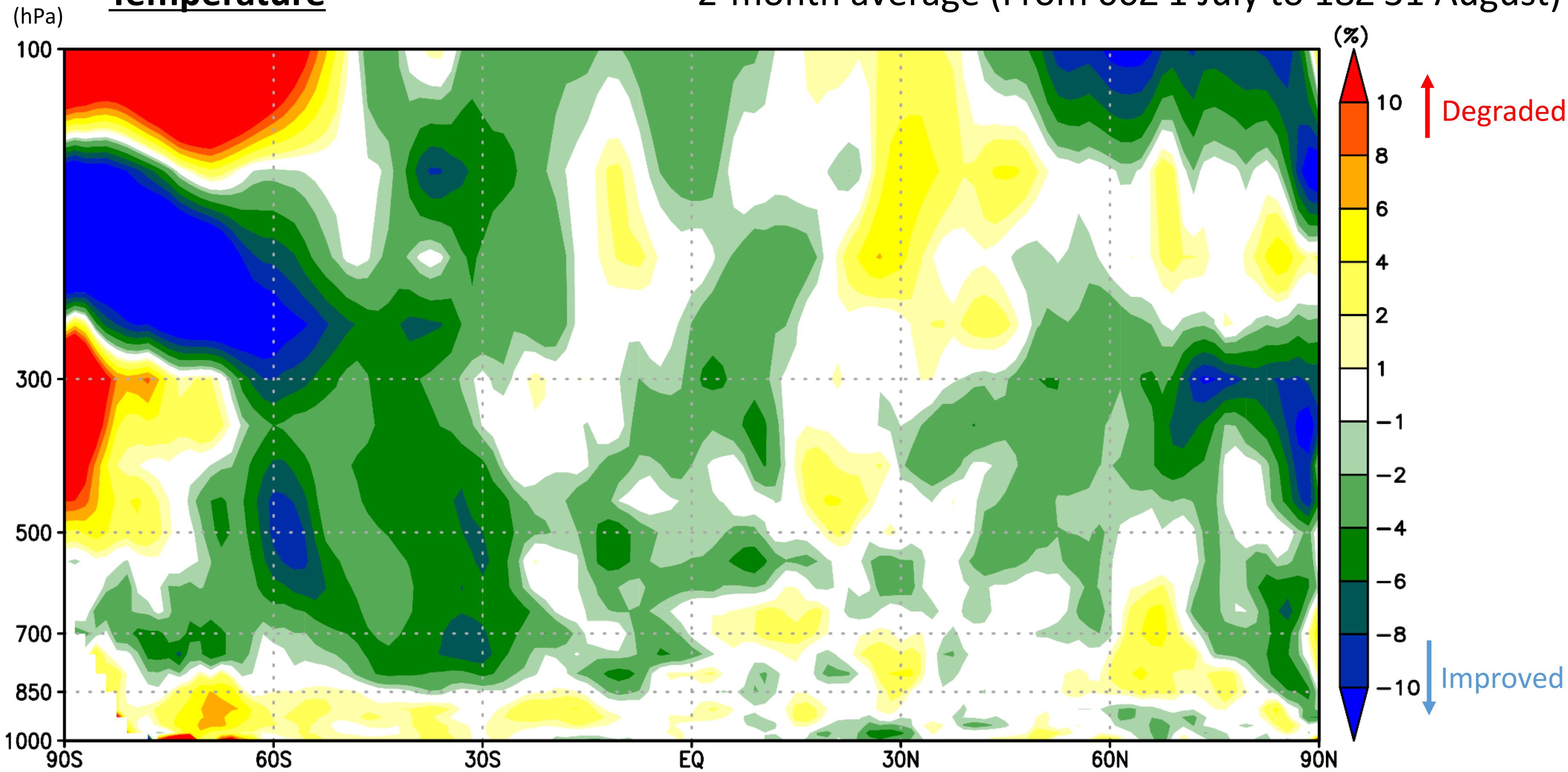
Temperature and zonal wind at mid- and upper troposphere improved

Analysis RMSE change（FULL125 vs DIAG250)

(vs. ERA-interim)

125km thinning, **Full R**

Temperature 2-month average (From 00Z 1 July to 18Z 31 August)

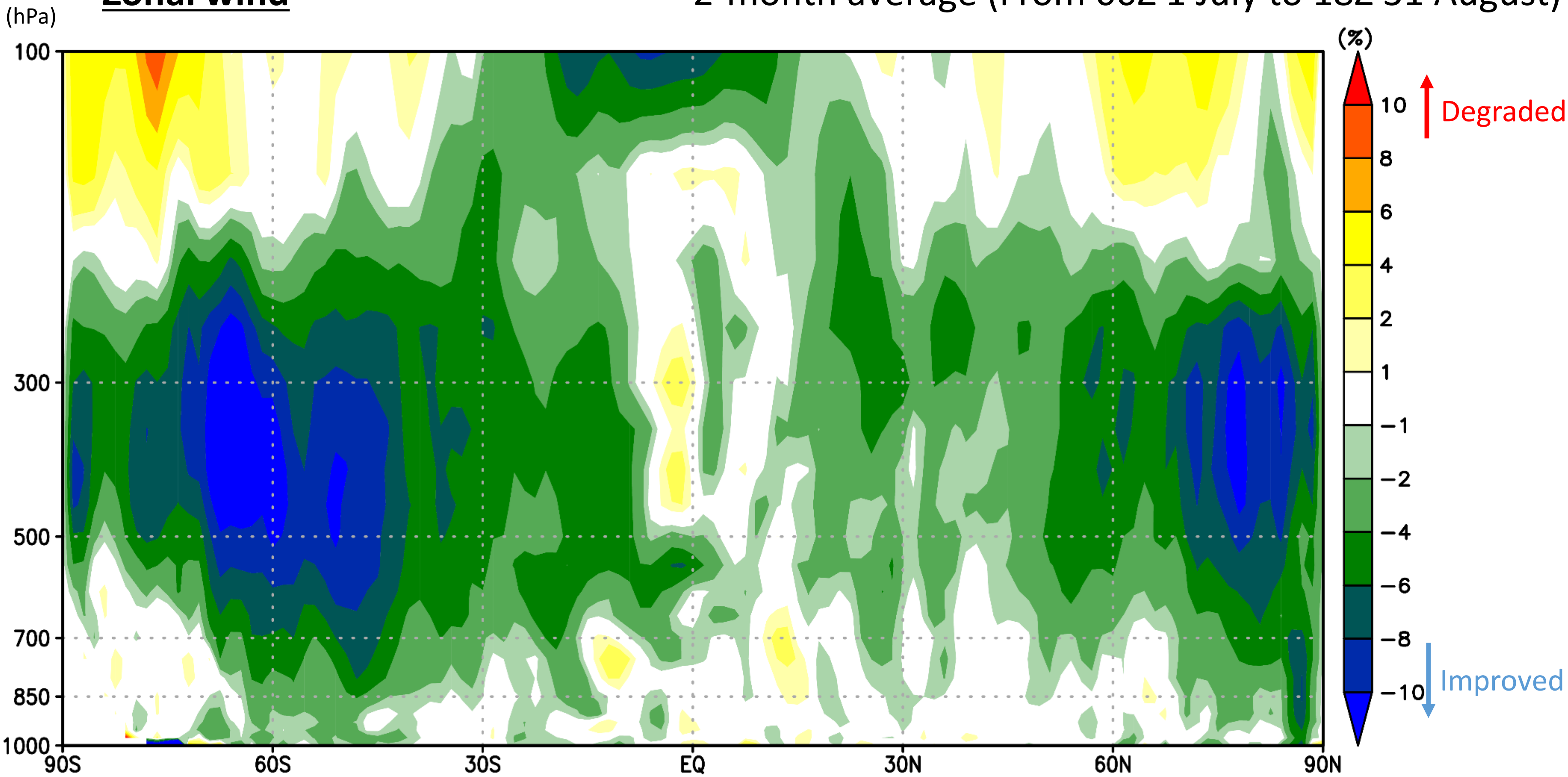


Analysis RMSE change (FULL125 vs DIAG250)

(vs. ERA-interim)

125km thinning, **Full R**

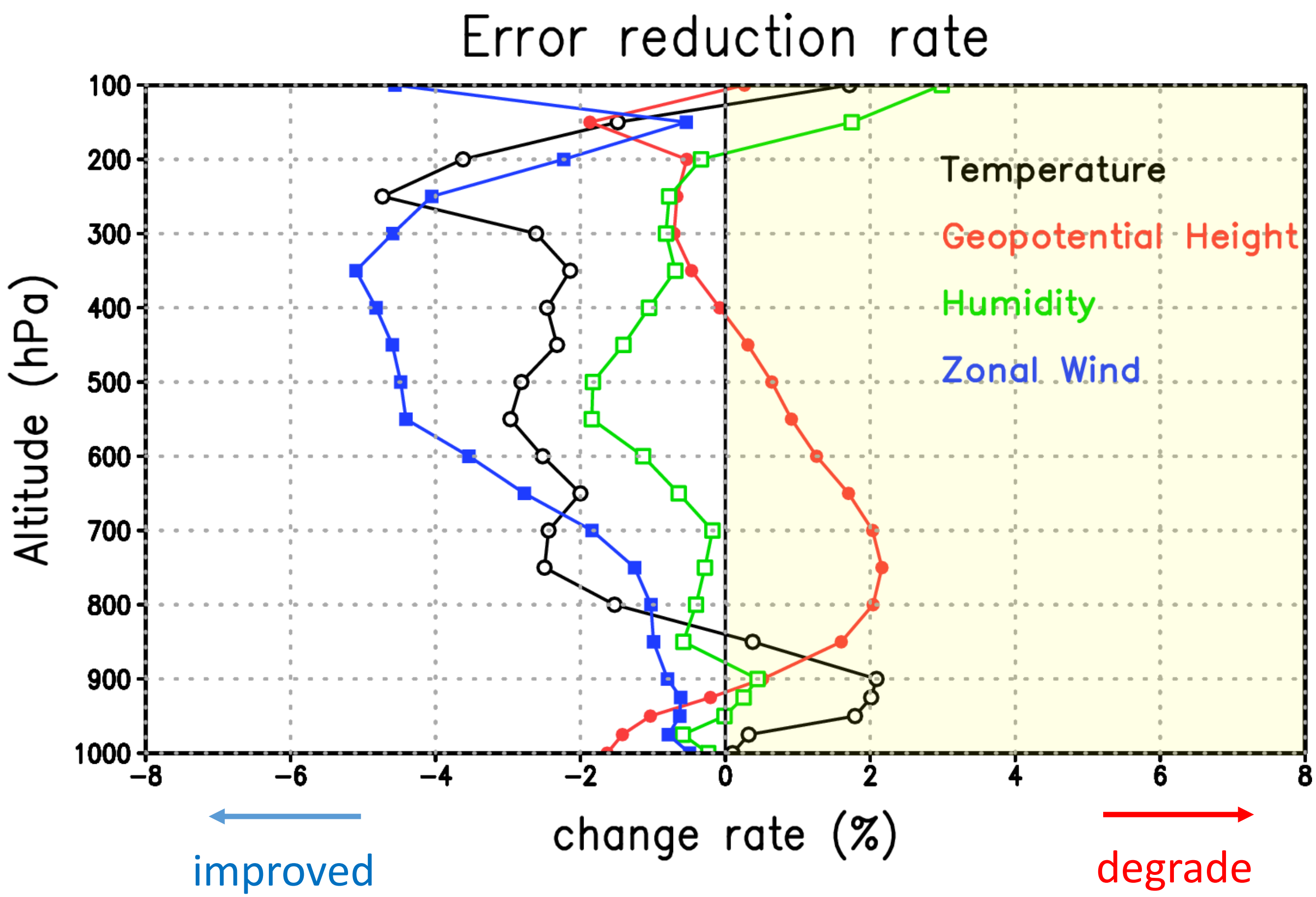
Zonal wind 2-month average (From 00Z 1 July to 18Z 31 August)



Analysis RMSE change (FULL125 vs DIAG250)

Global mean

2-month average (From 00Z 1 July to 18Z 31 August)



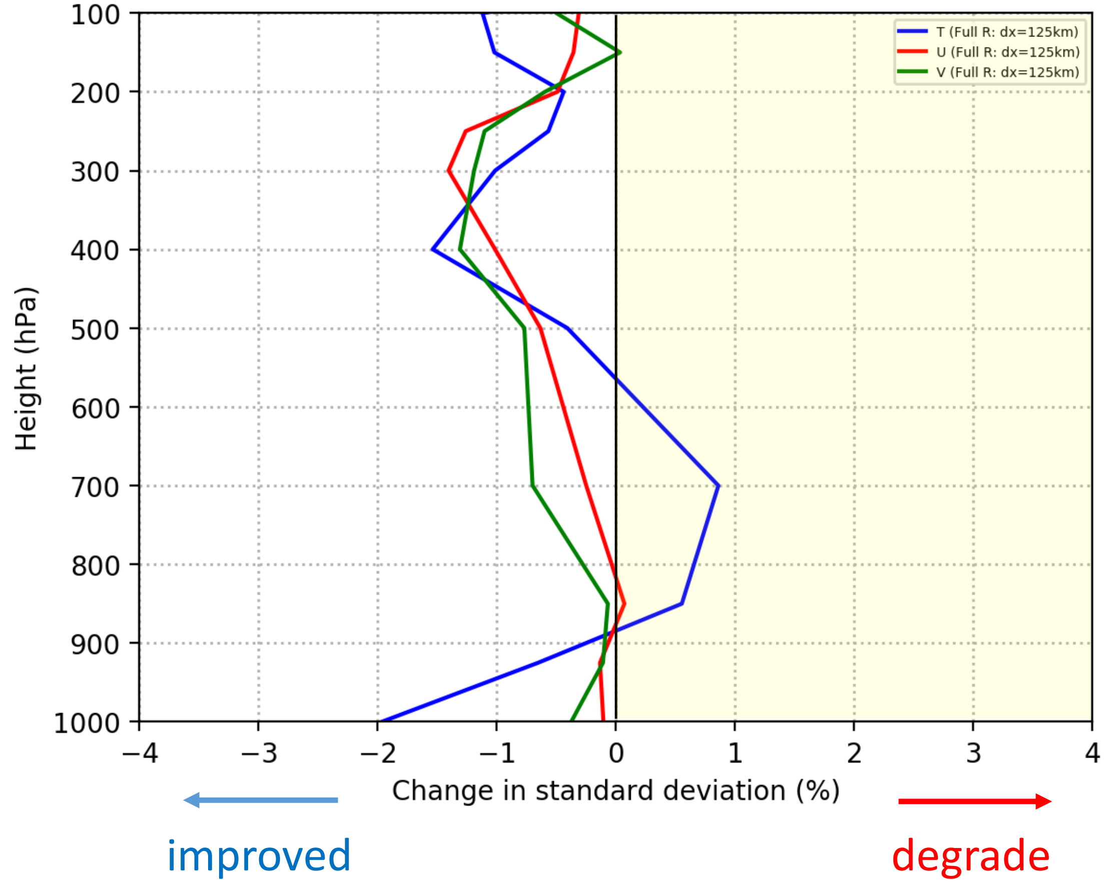
- Positive impact on zonal wind and temperature
- Slightly degrade geopotential height

Verification against observation

6-hour forecast

FULL125 (vs DIAG250)

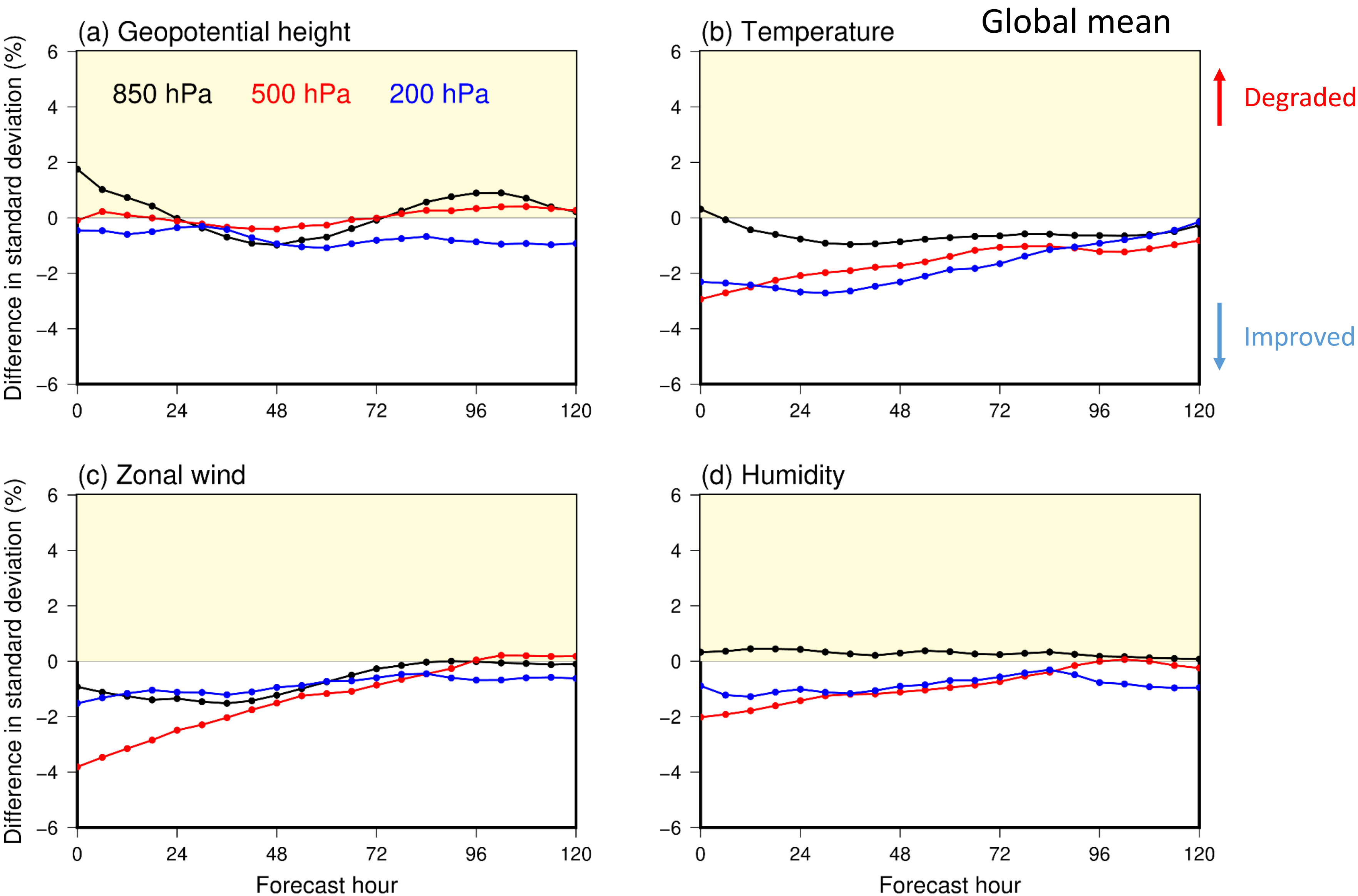
250 km (Diag **R**) → 125 km (Full **R**)



- Improved except for temperature of lower troposphere

Forecast improvements

1-month average (From 00Z 1 August to 18Z 31 August)



Computational Cost

- Inverting **R** will increase when the non-diagonal components are considered.

- **R** matrix becomes block diagonal because the error correlation between satellites and channels is not considered.

- Inverting the small block diagonal matrix reduces the increase in computational cost. (Up to 13 %)

$$\mathbf{R} = \begin{pmatrix} \mathbf{R}_1 & & & \mathbf{0} \\ & \mathbf{R}_2 & & \\ & & \mathbf{R}_3 & \\ \mathbf{0} & & & \mathbf{R}_4 \end{pmatrix} \Rightarrow \mathbf{R}^{-1} = \begin{pmatrix} \mathbf{R}_1^{-1} & & & \mathbf{0} \\ & \mathbf{R}_2^{-1} & & \\ & & \mathbf{R}_3^{-1} & \\ \mathbf{0} & & & \mathbf{R}_4^{-1} \end{pmatrix}$$

Computational Cost

Computer simulations create the future

Using 32 nodes of Fujitsu's supercomputer (FX100)

	Obs Cor	Thinning		
			Obs. Op.	**LETKF**
DIAG250		250 km	**24.54 (s)**	**70.89 (s)**
DIAG125		125 km	**32.36 (s)**	**75.18 (s)**
FULL125	✓	125 km	**32.11 (s)**	**84.89 (s)**

13% increase

Summary

- Including the horizontal observation error correlation in DA
  - Idealized case experiment
    - LETKF computation was unstable when the condition number of **R** was extremely large.
    - Reducing the condition number of **R** by reconditioning stabilized the data assimilation cycles.
    - The analysis was greatly improved by including the horizontal observation error correlation.
  - Experiment with real observations (AMSU-A)
    - **R** was estimated with innovation statistics.
    - The analysis was improved by up to 5% by including the horizontal observation error correlation.
    - The forecast was also improved especially for temperature and zonal wind.
- **Future directions**
  - Increasing the ensemble size.
  - Increasing the spatial resolution of NWP.