# Developmenet of an EnKF-based ocean data assimilation system

● **Responsible Representative**

Riko Oki, Space Technology Directorate I, Earth Observation Research Center

● **Contact Information**

Misako Kachi(kachi.misako@jaxa.jp)

● **Members**

Misako Kachi, Shun Ohishi

● **Abstract**

With the recent development of satellites, sea surface temperatures at high spatiotemporal resolutions have been observed by infrared sensors, but the data is not available in the cloudy regions because clouds block infrared radiations. Data assimilation can estimate more accurate 3-dimensional ocean fields without missing values combining the model and observations. This study aims to develop an ensemble Kalman filter-based ocean data assimilation system to assimilate satellite observations from JAXA on the JAXA Supercomputer System Generation 2 and 3 (JSS2/3).


Ref. URL: https://www.eorc.jaxa.jp/en/research/topic/ocean/

● **Reasons and benefits of using JAXA Supercomputer System**

Ensemble integration of a high-resolution ocean model and data assimilation by the ensemble Kalman filter are large-scale computations. This computation becomes possible by using large-scale high-performance computer infrastructure such as the JSS2/3.

● **Achievements of the Year**

The recent developments of satellites and Argo profiling floats enable to observe at high spatiotemporal resolutions. Assimilating such observations in a short time interval could improve the analysis accuracy. However, since data assimilation might break the dynamical balance, initial shocks with gravity waves might occur. Consequently, data assimilation in a short time interval might result in the degradation of the analysis field. This study aims to construct an ensemble Kalman filter-based ocean data assimilation system with numerical stability even in a 1-day assimilation interval shorter than a general one of 1 week to utilize high-frequency satellite observations. Here, we apply an ensemble Kalman filter to explicitly deal with flow-dependent forecast errors. We also investigate optimal methods to inflate under-dispersive of the ensemble spread and reduce noise by gravity waves maintaining dynamical balance and analysis accuracy.

We have performed eight sensitivity experiments combining four covariance inflation methods to inflate under-dispersive ensemble spread [multiplicative inflation (MULT), no inflation (NONE), and relaxation-to-prior perturbation and spread (RTPP and RTPS, respectively; Zhang et al. 2004; Whitaker and Hamill 2012)] and with/without incremental analysis updates (IAU; Bloom et al. 1996) to gradually add the analysis increments during model integration. Besides, we conduct twelve sensitivity experiments with relaxation parameters in RTPP/RTPS (alfa*RTPP/RTPS = 0.5, 0.6, ... 1.1). Residual of a nonlinear balance equation (Zhang et al. 2001) and root mean square deviation (RMSD) relative to observations, respectively, are used to evaluate the dynamical imbalance and accuracy of the analysis.

The sensitivity experiments demonstrate that RTPP/RTPS and IAU (multiplicative inflation) significantly improve (degrade) the dynamical balance and accuracy (Figs. 1, 2). The combination of RTPP/RTPS and IAU provides the best dynamical balance and accuracy maintaining numerical stability. The best relaxation parameters in RTPP and RTPS are about 0.8-0.9 (Figs. 3, 4).
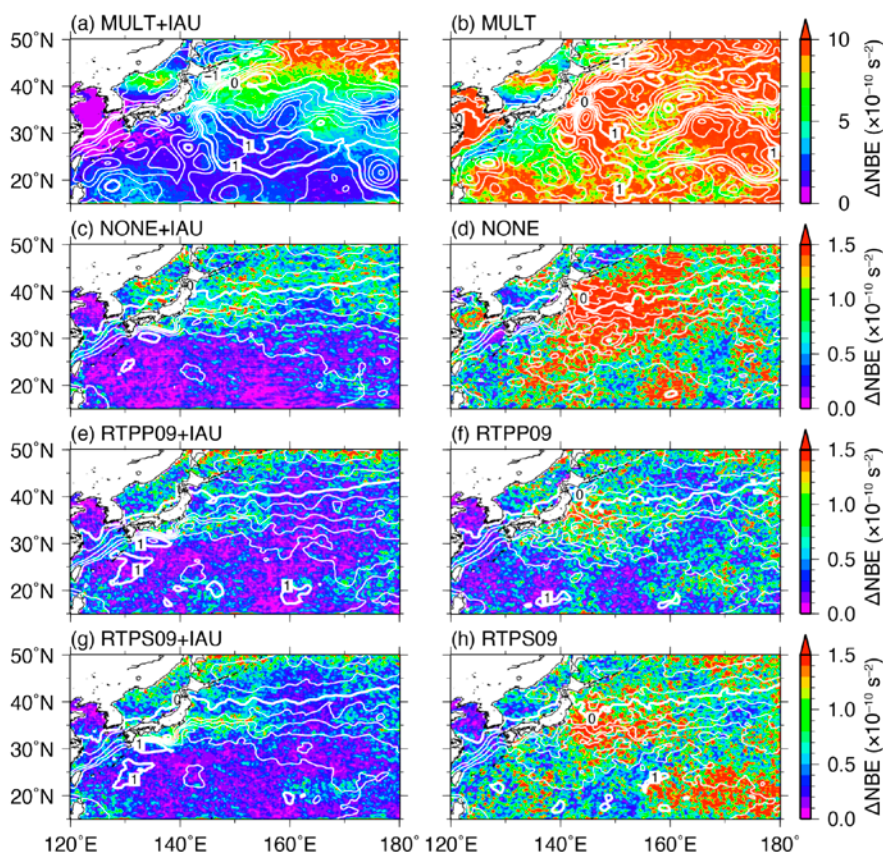


Fig. 1: Geostrophic balance indicator delta_NBE (color) and sea surface height (contour) averaged in 2016–17 in MULT+IAU, (b) MULT, (c) NONE+IAU, (d) NONE, (e) RTPP09+IAU, (f) RTPP09, (g) RTPS09+IAU, and (g) RTPS09 experiments. Here, for example, experiments with the MULT and IAU method and RTPP with the relaxation parameter of alfa*RTPP=0.9 are described as MULT+IAU and RTPP09 experiments, respectively. Smaller and larger delta_NBE means more and less geostrophic balance, respectively. We note that MULT+IAU and MULT experiments use a different color scale from other experiments.
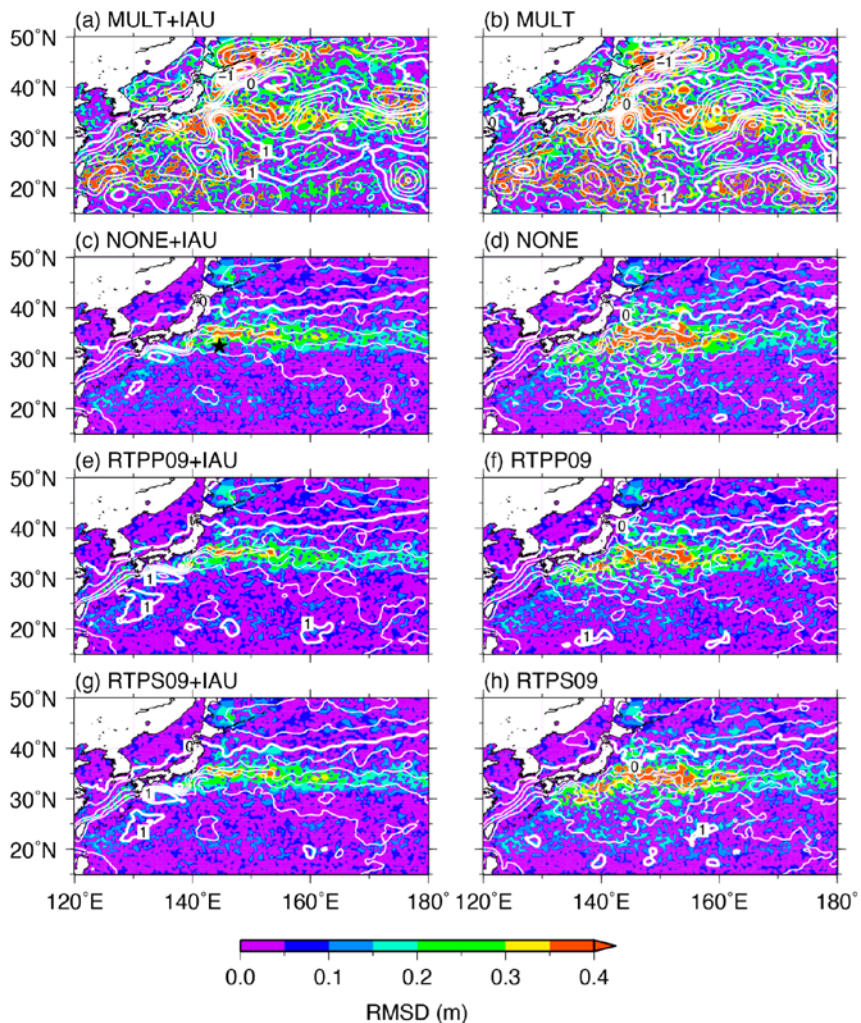
Fig. 2: As in Fig. 1, but for RMSD (color) relative to observational sea surface height anomaly dataset (Ducet et al. 2000). Smaller and larger RMSD indicates more and less analysis accuracy, respectively.
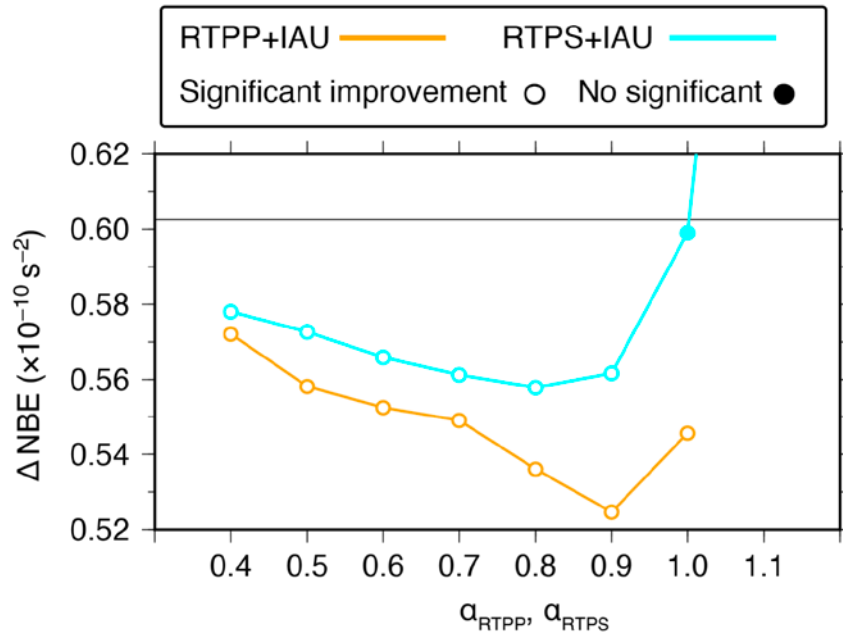
Fig. 3: delta_NBE averaged within the system domain and in 2016-17 in RTPP04-09+IAU (orange) and RTPS04-09+IAU experiments (cyan). Open and closed circles indicate significant improvement and no significant improvement/degradation, respectively, relative to the NONE+IAU experiment. A black line denotes the NONE+IAU experiment.
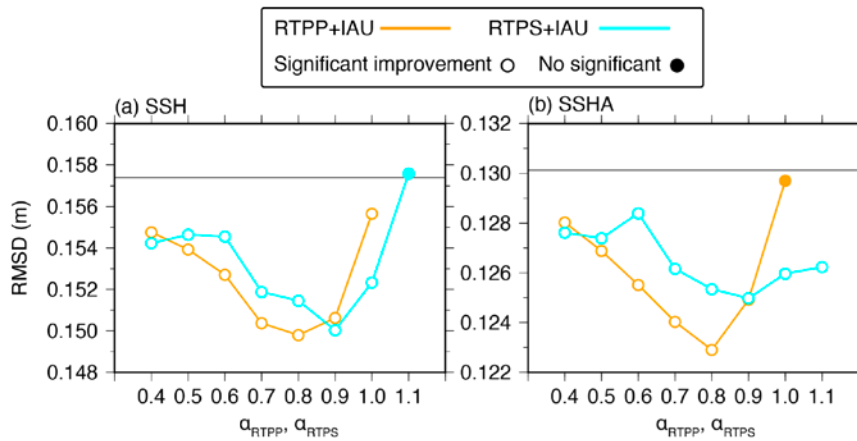


Fig. 4: Figure 4: As Fig. 3, but for RMSD relative to observational (a) sea surface height and (b) its anomaly datasets.

● **Publications**

- Poster Presentations

1. Ohishi, Shun, Tsutomu Hihara, Hidenori Aiki, Joji Ishizaka, Yasumasa Miyazawa, Misako Kachi, and Takemasa Miyoshi, 'Development of an ensemble Kalman filter-based regional ocean data assimilation system', JpGU-AGU Joint Meeting 2020, On-line, Jul. 2020

2. Ohishi, Shun, Tsutomu Hihara, Hidenori Aiki, Joji Ishizaka, Yasumasa Miyazawa, Misako Kachi, and Takemasa Miyoshi 'Development of an ensemble Kalman filter-based regional ocean data assimilation system', The 3rd R-CCS International Symposium, On-line, Feb. 2021

● **Usage of JSS**

●      **Computational Information**

| Process Parallelization Methods | MPI |
|---|---|
| Thread Parallelization Methods | Automatic Parallelization |
| Number of Processes | 4 - 400 |
| Elapsed Time per Case | 200 Minute(s) |

●   **Resources Used(JSS2)**

Fraction of Usage in Total Resources[*1](%): 0.98

Details

| Computational Resources | | |
|---|---|---|
| System Name | Amount of Core Time (core x hours) | Fraction of Usage[*2](%) |
| SORA-MA | 5,697,212.59 | 1.08 |
| SORA-PP | 0.00 | 0.00 |
| SORA-LM | 0.00 | 0.00 |
| SORA-TPP | 0.00 | 0.00 |

| File System Resources | | |
|---|---|---|
| File System Name | Storage Assigned (GiB) | Fraction of Usage[*2](%) |
| /home | 42.92 | 0.04 |
| /data | 83,866.16 | 1.62 |
| /ltmp | 8,789.07 | 0.75 |

| Archiver Resources | | |
|---|---|---|
| Archiver Name | Storage Used (TiB) | Fraction of Usage[*2](%) |
| J-SPACE | 0.00 | 0.00 |

[*1]: Fraction of Usage in Total Resources: Weighted average of three resource types (Computing, File System, and Archiver).

[*2]: Fraction of Usage：Percentage of usage relative to each resource used in one year.

- **Resources Used(JSS3)**

Fraction of Usage in Total Resources[*1](%): 1.39

Details

| Computational Resources | | |
|---|---|---|
| System Name | Amount of Core Time (core x hours) | Fraction of Usage[*2](%) |
| TOKI-SORA | 7,154,288.48 | 1.54 |
| TOKI-RURI | 0.00 | 0.00 |
| TOKI-TRURI | 0.00 | 0.00 |

| File System Resources | | |
|---|---|---|
| File System Name | Storage Assigned (GiB) | Fraction of Usage[*2](%) |
| /home | 14.31 | 0.01 |
| /data | 146,627.49 | 2.46 |
| /ssd | 143.05 | 0.07 |

| Archiver Resources | | |
|---|---|---|
| Archiver Name | Storage Used (TiB) | Fraction of Usage[*2](%) |
| J-SPACE | 0.00 | 0.00 |

[*1]: Fraction of Usage in Total Resources: Weighted average of three resource types (Computing, File System, and Archiver).

[*2]: Fraction of Usage：Percentage of usage relative to each resource used in one year.