A Global Ocean Assimilation and Forecast System using the localized weighted ensemble Kalman filter

Yan Chen, Xiaohui Wang, Senliang Bao

Changsha, China

National University of Defense Technology

国防科技大学气象海洋学院
College of Meteorology and Oceanography of NUDT

Motivation

The global ocean forecasting system has played a significant role in the fields of marine resource development, disaster early warning, and climate change research. Data assimilation is a core technical means to improve the accuracy of ocean forecasts. Ocean forecasting relies on the accuracy of the initial ocean state, such as temperature, salinity, and current velocity. Data assimilation enhances forecast accuracy by combining observational data (e.g., from satellites, buoys, and ships) with numerical models to optimize the initial fields and reduce model initial errors. Currently, the mainstream global ocean assimilation and forecasting systems internationally include GLO12v4(Mercator Ocean), GIOPS 3.5.0(Environment Canada), FOAM GOSI9(UK Met Office), RTOFS v2(NOAA/NCEP), GOFS 3.5(U.S. Navy), and NMEFC-NEMO(NMEFC), etc. These systems predominantly employ traditional assimilation methods, such as three-dimensional variational (3DVar) assimilation, Kalman filter (KF) and its variants, ensemble Kalman filter (EnKF) and its variants, etc. However, their theoretical frameworks lack the capability to simulate non-Gaussian background errors, making it difficult to meet the requirements of the development of new observational techniques and the improvement of model resolution.

Method

1. Data Assimilation Method

The Localized Weighted Ensemble Kalman Filter (LWEnKF) is an innovative data assimilation technique that integrates the strengths of Particle Filter (PF) and Ensemble Kalman Filter. It introduces a localization scheme in the weight calculation, thereby overcoming the traditional assumptions of linearity or weak nonlinearity of the model and observation operators. Additionally, it does not rely on the Gaussian distribution assumption of errors. As a fully nonlinear and non-Gaussian assimilation method, LWEnKF demonstrates significant advantages in addressing nonlinear problems and non-Gaussian error distributions in real complex ocean systems.

2. Ocean Model

The Mass Conservation Ocean Model (MaCOM) is an ocean circulation numerical model for operational applications by the National Marine Environmental Forecasting Center(NMEFC), Ministry of Natural Resources. The model achieves mass conservation of seawater in the dynamical framework, overcoming the inherent shortcomings of traditional volume-conserving ocean models in accurately simulating sea surface height and salinity. The model primarily focuses on ocean circulation numerical forecasting capabilities, while also accommodating additional applications such as sea ice, tracer trajectory, and coupling.

3. Fast Sparse Method for Unstructured Grids

Horizontal thinning averages observations within each model grid to form super-observations, thereby eliminating redundant information. However, the longitudes and latitudes of unstructured grids are not stored in ascending or descending order, and non-parallel brute-force search algorithms have extremely long computational times. Since KD-trees have efficient search capabilities for unordered data, this method is adopted to obtain super-observations without parallelization. For instance, thinning 1,036,800 sea surface height observations takes only 25 seconds.

Experiment design

We conducted a quasi-operational assimilation and forecasting experiment. The global ocean model is based on MaCOM, utilizing a hexahedral grid with a horizontal resolution of $1/12^{\circ}$ and 75 vertical levels. The assimilation method employed is the Localized Weighted Ensemble Kalman Filter, with localization coefficients determined through experimental tuning. The LWEnKF involves 60 ensemble perturbations, which are sampled from the model's climatological state.

Considering the substantial computational cost associated with particles in the global ocean forecasting process, this study referenced an efficient approximation method similar to EnOI, conducting numerical forecasts only for the ensemble mean. The assimilated observations include satellite sea surface temperature, satellite sea surface height anomalies, Argo temperature and salinity profiles, and reconstructed temperature and salinity profiles. The control variables are sea surface height, bottom pressure, temperature and salinity. The vertical coordinate uses pressure coordinates, which is also consistent with the pressure coordinates of Argo observations. The assimilation window is 1 day, with forecasts extending to 10 days. The experimental period spans from August 15, 2024, to September 15, 2024.

References

- Chen, Yan; Zhang, Weimin; Zhu, Mengbin. A localized weighted ensemble Kalman filter for high-dimensional systems, Quarterly Journal of the Royal Meteorological Society, 2020, 146(726):438-453
- Chen, Yan; Zhang, Weimin; Wang, Pinqiang. An application of the localized weighted ensemble Kalman filter for ocean data assimilation, Quarterly Journal of the Royal Meteorological Society, 2020, 146(732): 3029-3047
- Yan Chen, Weimin Zhang, Huizan Wang, and Yu Cao. Data assimilation system for the finite-volume community ocean model based on a localized weighted ensemble Kalman filter. Journal of Applied Remote Sensing, 2023, 17(02)

Contact Information

cy_msc@126.com

Results

The following figure presents a comparative analysis of the Root Mean Square Error (RMSE) for sea surface height anomaly (SLA) and sea surface temperature (SST) forecasts over a 10-day lead period, evaluating two distinct systems: LWEnKF and GLO12V4. The results indicate that both systems perform admirably in generating forecasts with a reasonable degree of accuracy. However, the LWEnKF system demonstrates a marginal improvement in forecast precision, particularly as the lead time extends. This enhanced performance may be attributed to the sophisticated algorithms and data assimilation techniques employed by LWEnKF, which appear to more effectively capture the intricate dynamics of the ocean surface. The observed increase in RMSE with longer lead times is consistent with expectations, as the cumulative effects of initial condition uncertainties and model inaccuracies are likely to amplify the forecast errors over time.

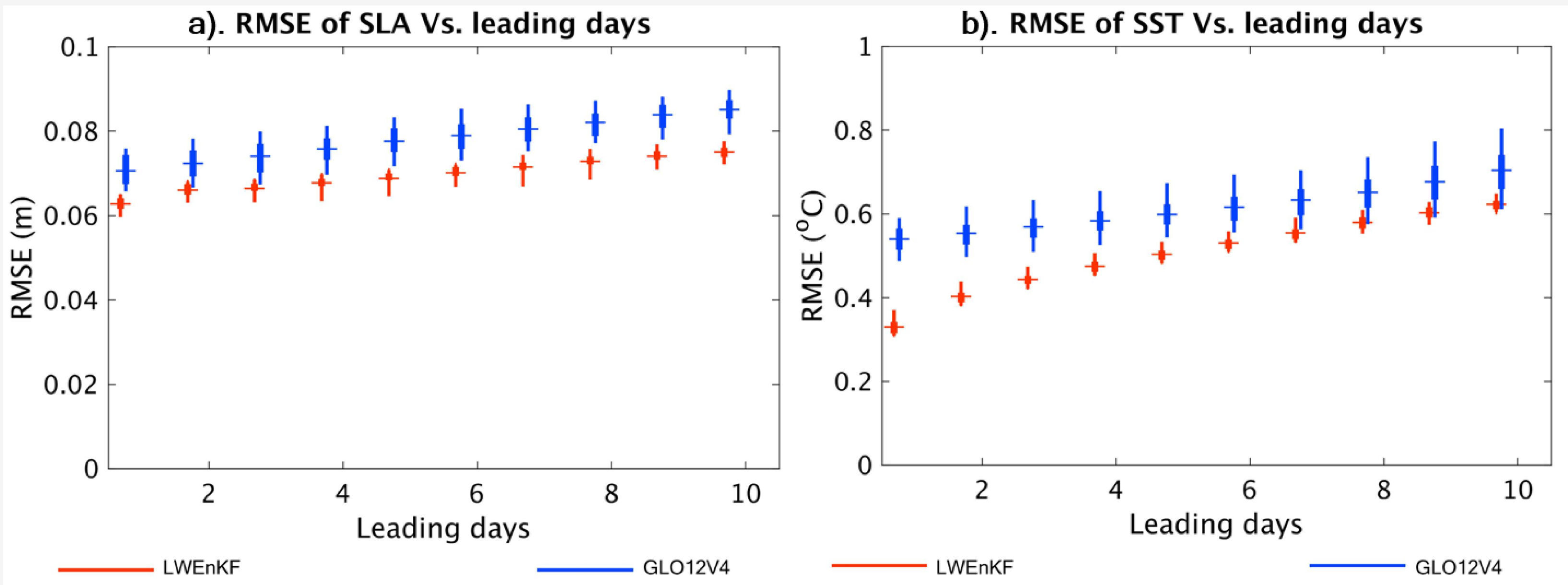

Figure: RMSE for a).SLA and b).SST

The following figure compares the performance of the two systems in forecasting temperature and salinity profiles at various depths over a one-day lead time. Figure a) presents the RMSE of temperature profiles, while Figure b) displays the RMSE of salinity profiles. Regarding temperature forecasts (Figure a), both systems exhibit similar performance, with RMSE values being closely aligned across most depths. However, the LWEnKF system has slightly lower RMSE values at the sea surface, while GLO12V4 has lower RMSE in the thermocline. For salinity forecasts (Figure b), the GLO12V4 system shows a slight advantage, especially in the shallower regions up to 300 meters. The RMSE values for LWEnKF are consistently higher than those for GLO12V4, indicating that the LWEnKF system may not represent salinity profiles as accurately.

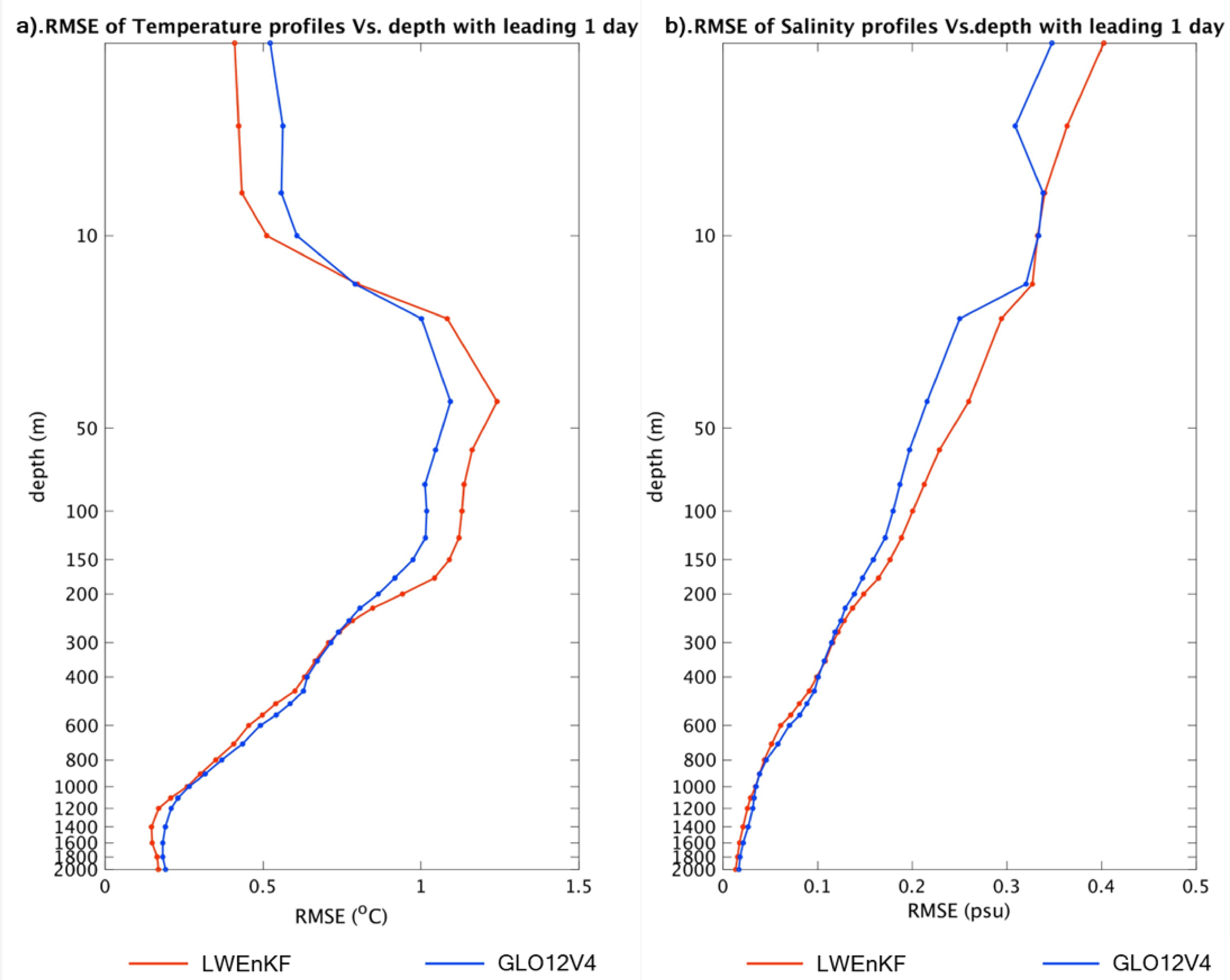

Figure: 1-day Forecast RMSE of a).Temperature and b).Salinity Profiles

In summary, both systems provide forecasts of acceptable accuracy. However, the LWEnKF system demonstrates a noticeable enhancement in the precision of sea surface height anomaly and sea surface temperature forecasts. Despite this, its performance in salinity profile prediction remains suboptimal. This shortfall may stem from the limited availability of salinity observations. Looking ahead, we intend to integrate a broader array of in-situ observational data, including measurements from CTD, XBT, gliders, and ship-based instruments, to enhance the salinity profile forecasting capabilities.