We greatly appreciate the time and effort you invested in reviewing our work and providing your valuable comments on our manuscript. we have revised the manuscript accordingly. Below, we provide a point-by-point response to the reviewers' comments.

How does the choice of the Local Singular Evolutive Interpolated Kalman (LSEIK) filter impact the accuracy and stability of reanalysis, and what justifications were made for using a constant inflation factor rather than a spatially varying one?

Yang et al. (2015) demonstrated the effectiveness of the LSEIK filter in reducing model bias, improving the accuracy of model predictions and leading to more reliable results. Liu and Fu (2018) further showed that LSEIK enhances the performance of the NEMO-Nordic model by assimilating high-resolution SST data, providing more accurate representations of the ocean state. LSEIK has also been recognized as a strong candidate for stable and accurate reanalysis applications. This is evident in the CMEMS reanalysis used in the 3rd and 4th editions of the Copernicus Marine Service Ocean State Reports, where a constant inflation factor was applied.

The "inflation factor" refers to a method used to adjust the model error covariance matrix in numerical simulations. There are two main types of inflation factors:

- Constant inflation factor: This approach is simple and computationally efficient, assuming that uncertainty is the same across all locations. However, it may not account for regional variations in uncertainty.
- Spatially varying inflation factor: This method can better represent regional differences in uncertainty, potentially improving state estimates in data assimilation. However, it is more complex and requires more computational resources, making it less practical in some cases, especially when the distribution of model errors is unclear. This often leads to a preference for a constant inflation factor in many DA applications.

Given the limited understanding of how model errors vary spatially in the NEMO-Nordic domain, we treat uncertainty as constant and globally adjust the model error covariance matrix. Using a constant inflation factor helps maintain the overall spread of the ensemble, which is essential for accurate predictions.

Research by Scheffler et al. (2022) and Liu et al. (2021) has demonstrated the effectiveness of using a constant inflation factor in data assimilation. Based on its proven success and simplicity, we chose to apply this method in our offline setup to generate sample ensembles for data assimilation.

Qinghua Yang, Svetlana N. Losa, Martin Losch, Jiping Liu, Zhanhai Zhang, Lars Nerger and Hu Yang. 2015. Assimilating summer sea-ice concentration into a coupled ice–ocean model using a LSEIK filter. Annals of Glaciology, 56 , 38 – 44.

Ye Liu and Weiwei Fu. 2018. Assimilating high-resolution sea surface temperature data improves the ocean forecast potential in the Baltic Sea, Ocean Sci., 14, 525–541, https://doi.org/10.5194/os-14-525-2018.

Guillermo Scheffler, Alberto Carrassi Juan Ruiz, Manuel Pulido. 2022. Dynamical effects of inflation in ensemble-based data assimilation under the presence of model error, Quarterly Journal of the Royal Meteorological Society, 148, 2368-2383.

Liu, Y., Xie, J., Liu, Z., Gan, J., & Zhu, J. (2021). The assimilation of temperature and salinity profile observations for forecasting the river–estuary–shelf waters. Journal of Geophysical Research: Oceans, 126, e2020JC017043. https://doi.org/10.1029/2020JC017043.

Given that the reanalysis shows significant variability in simulation quality across different sub-basins, can the authors provide further details on how biases (especially for temperature and salinity) were corrected and what limitations persist in coastal and deepwater regions?

The biases in the temperature and salinity data are closely linked to the number of observations collected from different sub-basins. Figure 2 presents the variation in observation counts across various Baltic sub-basins. As a result, the quality of the data simulations differs significantly among these sub-basins. This variability is further validated using both in-situ and satellite observations. It is important to note that no bias corrections have been applied before the data assimilation. Due to lack of sufficient observations, a 3D bias correction before data assimilation may lead to extra errors. We only assimilated observations where four surrounding grids were all wet model grids, which leads to significant error differences between coastal and deeper water regions due to the complex topographic conditions of the model. We haven't other limitations for coastal and deep-water regions.

The study utilizes multiple validation datasets. How do differences between assimilated and validation datasets (e.g., OSISAF vs. CMEMS SST products) impact the interpretation of errors and the robustness of the results?

OSISAF data were assimilated when making the reanalysis, which is Level 3 Sea-ice Surface Temperature. Then reanalysis was validated against CMEMS L4 SST, which was an optimal interpolation product based on multiple satellites including OSISAF satellites. OSISAF has a lower spatial resolution and lower coverage compared to CMEMS L4 SST. Therefore, it isn't surprise that the CMEMS-L4 shows different features relative to OSISAF. Different satellite products could show different nature of the datasets due to using different satellite platforms, algorithms, and processing methods. The differences between OSISAF and CMEMS L4SST may have some impacts on the interpretation of SST error features. We revised the SST validation figure and included a comparison between the CMEMS L4 SST and the assimilated satellite SST to assess those influences.

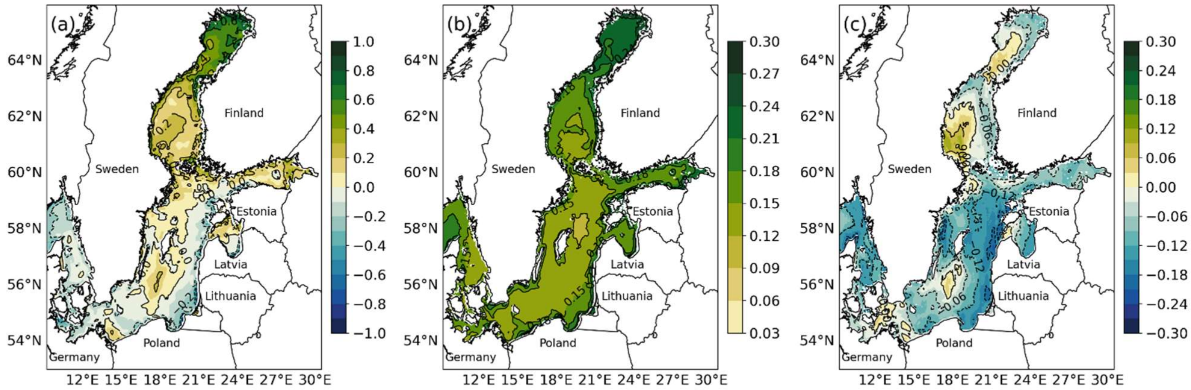

Figure 1. The overall averaged (a) bias, (b) NRMSD of reanalyzed sea surface temperature, and (c) difference in assimilated satellite sea surface temperature relative to the CMEMS L4 sea surface temperature product for the period 1990-2020.

We have revised the text with further elaborated texts:

“However, the differences between the DA and validation satellite SSTs could potentially impact the accuracy of the SST validation process. The discrepancy between assimilated combined satellite SST from IceMap and OSISAF and CMEMS-L4 SST varies in time and space (not shown). The CMEMS-L4 SST was warmer than both assimilated satellite SSTs and the reanalysis SST in the southwest of both the Bothnia Sea and the Bothnian Bay and the south of Gotland basin. Therefore, it is reasonable to believe that the differences between assimilated combined satellite SST and CMEMS-L4 SST contributed to the reanalyzed biases of SST in those regions relative to CMEMS-L4 SST. The opposite phenomenon occurred in the eastern Baltic Proper, the Bornholm Basin, Eastern Arkona Basin, southern Gulf of Riga, and the transition region between the North Sea and Baltic Sea (Fig. 7). These results highlight the importance of considering such discrepancies for a comprehensive assessment of the reanalysis results and their implications for understanding SST variations. The reprocessed product on multiple satellites is better to represent a general ocean state than a single satellite product. In this study, we used a reprocessed SST (CMEMS-L4) based on multiple satellites to validate the assimilation results to verify the robustness and generalizability of the reproduced SST”

Given the reported warming and salinity trends, how do these findings compare with observed trends from other datasets covering different periods? Are the differences statistically significant and, if so, what could be contributing to these discrepancies?

We have already intercompared our results with other observation-based trend analysis in the Baltic Sea, here we add two more recent researches for intercomparing:

Zalewska, T., Wilman, B., Lapeta, B., Marosz, M., Biernacik, D., Wochna, A., Saniewski, M., Grajewska, A., Iwaniak, M. 2023. Seawater temperature changes in the southern Baltic Sea (1959–2019) forced by climate change, Oceanologia, 66, 37-55. https://doi.org/10.1016/j.oceano.2023.08.001. Its main conclusions are:

- SST in the southern Baltic in 1959–2019 increased by 0.6˚C per decade.
- The change in temperature in the water column varies with the depth.
- The most intense water warming occurs in the spring–summer (0.8–1˚C/decade).
- The average increase in seawater temperature in the coastal zone was 0.2˚C/decade.

Kankaanpää, H. T., Alenius, P., Kotilainen, P., Roiha, P., 2023. Decreased surface and bottom salinity and elevated bottom temperature in the Northern Baltic Sea over the past six decades, Science of The Total Environment, 859, https://doi.org/10.1016/j.scitotenv.2022.160241. Its main conclusions are:

- Changes in northern Baltic Sea (North of 57N) temperature and salinity in 1957–2021 were investigated
- Temperatures in the near-bottom layer increased by 0.75–2.9 °C (0.013–0.115 °C a^{-1}) since the 1960's
- Surface layer salinities declined by 0.31–1.14 units (0.005–0.019 a^{-1}) since the 1960's
- Varying trends occurred in bottom salinity; declines by 0.35–1.45 units (0.007–0.025 a^{-1}) occurred at 19 out of 33 locations

To the entire Baltic Sea, this study reported a warming rate of 0.037 °C/year in the Baltic surface waters for the period 1990 to 2020, which is very similar to previous estimates from in-situ observations: 0.03–0.06 °C/year for the period from 1982 to 2016 (Liblik and Lips, 2019) and 0.028 to 0.039 oC/year in the southeastern Baltic Sea for the period 1950 to 2020 (Stockmayer and Lehmann, 2023). However, SST trends from satellite measurements in the Baltic Sea showed a little higher estimate: 0.059 °C/year for the period 1990−2018 (Siegel and Gerth, 2019) and 0.048 °C/year for the period 1982-2021 (Jamali et al., 2023). Further, spatial pattern of the SST trend shown in the present study is very similar to earlier studies from both model or observations. Specifically, a smaller warming trend is observed in the southern Baltic Sea, while a larger warming trend appears in the northern Baltic Sea, consistent with the results of Siegel and Gerth (2019), Liblik and Lips (2019), and Jamali et al. (2023). Additionally, a significant warming trend is evident in the Gulf of Finland, which has been noted by Liblik and Lips (2019) and Stockmayer and Lehmann (2023). In the northern Baltic Sea, the warming trend in SBT reported in this study (0.031°C per year) aligns well with the warming rate (0.013–0.115°C/year) from in-situ observations for the period 1957–2021(Kankaanpää et al., 2023).

For the salinity of the Baltic Sea, the present study reported a freshening trend of -0.004±0.002 PSU/year at the surface for the period 1990−2020, which is very similar to the previous estimations: -0.005 to -0.014 PSU/year in the upper layers for the period 1982−2016 (Liblik and Lips 2019) and -0.005 to -0.019 PSU/year in the surface waters in the northern Baltic Sea for the period 1957–2021 (Kankaanpää et al., 2023). Furthermore, this study reported a salinization trend of 0.015 to 0.049 PSU/year in the deeper waters, which is consistent with earlier estimates as well: 0.02–0.04 PSU/year in the deeper waters for the period 1982−2016 (Liblik and Lips 2019). Besides, the salinity are declined by 0.009 PSU/year in northern Baltic bottom in this study, which is consistent to the findings of and 0.007–0.025 PSU/year in Kankaanpää et al. (2023) for the period 1957–2021.

The differences in SST trends observed between the satellite-based studies and this study can be attributed to several factors, such as the time periods analyzed, the sources of data used, and the spatial coverage of the observations. While satellite data provide broad spatial coverage but may be influenced by retrieval uncertainties, particularly in coastal areas. The lower SST trend reported in this study, compared to those in Siegel and Gerth (2019) and Jamali et al. (2023), may result from differences in spatial resolution and observational datasets. For instance, the satellite SSTs used in this study have relatively coarse horizontal resolutions (5 km). In contrast, the satellite SSTs used in their trend derivation have a higher resolution (4 km), which may be better at capturing fine-scale spatial variations. The NEMO-Nordic model used in this study also has a limited horizontal resolution (3.8 km) and vertical resolution (3-22 m). Accordingly, the simulation of vertical mixing processes is restricted, which could be a reason for the underestimated depth of the thermocline/halocline in the eastern Gotland Basin and for differing structures in the trend profiles. It may also struggle to exact spread water masses.

The results indicate significant seasonal and depth-dependent variability in temperature and salinity. Can the authors elaborate on how these variations align with physical oceanographic processes and potential anthropogenic impacts on the Baltic Sea?

Seasonality of T/S in different depths have been displayed in the paper. We’ve added a few lines in the revision.

- The seasonality in the surface salinity has a minimum in April-May, which is correspondent to the highest river runoff during the months
- The seasonality of water temperature in intermediate waters (as shown at 60m depth) is directly linked to the seasonal shoaling and deepening of thermocline and associated temperature variations, which are caused by vertical transport of seasonal variability of surface heat flux. The phase shift of the seasonality from the surface temperature is also a result of vertical heat transport.
- The trend of seasonality may be related to human-induced climate change. The strengthening of SST seasonality due to anthropogenic impacts have been documented by several papers while the weakening of seasonality in water temperature below thermocline in extratropical oceans was also found by Liu etc. (2024). There are also climate change signals in seasonality in our results, e.g. at 60m, the minimum value of 10-year mean monthly temperature (in April)

has increased by 1.2^{o}C from 1990-1999 to 2010-2020 while the maximum value (in December) has only increased by 0.7C. This means that the seasonality of water temperature at 60m has been weakened by 41.6%**.**

We also defined a seasonality index (Tmax-Tmin and Smax-Smin) to perform a detailed seasonality climate change analysis on Fig.12.

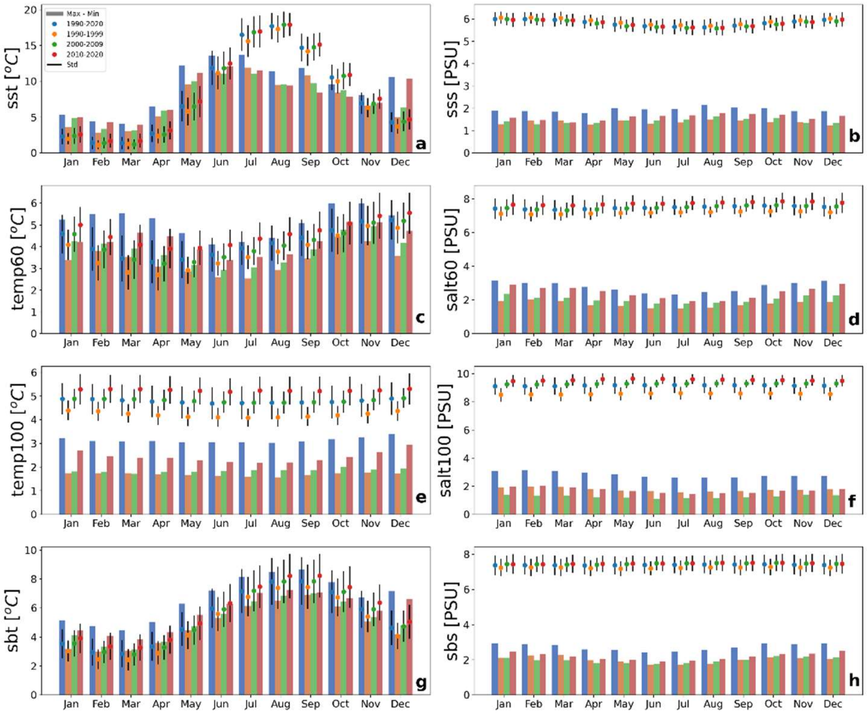

Figure 12. The monthly temperature and salinity of the Baltic Sea at the surface, 60 m, 100 m, and the bottom for different decades, derived from reanalysis. The dots, lines, and bars show the mean value and standard deviations, variability range of monthly temperature and salinity.

How do the authors address limitations of a 2-nautical mile horizontal resolution and the potential impacts of parameterization choices on model accuracy, particularly in areas with complex bathymetry?

Insufficient horizontal resolution, particularly in narrow Danish straits, will lead to certain errors in simulating Baltic-North Sea water transport. However, NEMO-Nordic1.0 with a 2-nm horizontal resolution used in this study has been calibrated and validated with observations by Hordoir et al. (2019), confirming that the parameterization of the model can perform reasonably in the field, even in areas with complex bathymetry. In this study, the biases of model have been reduced by data assimilation and data assimilation is to remedy the model’s disadvantages in some extent. here is another example to verify our model’s accuracy in the areas with complex bathymetry.

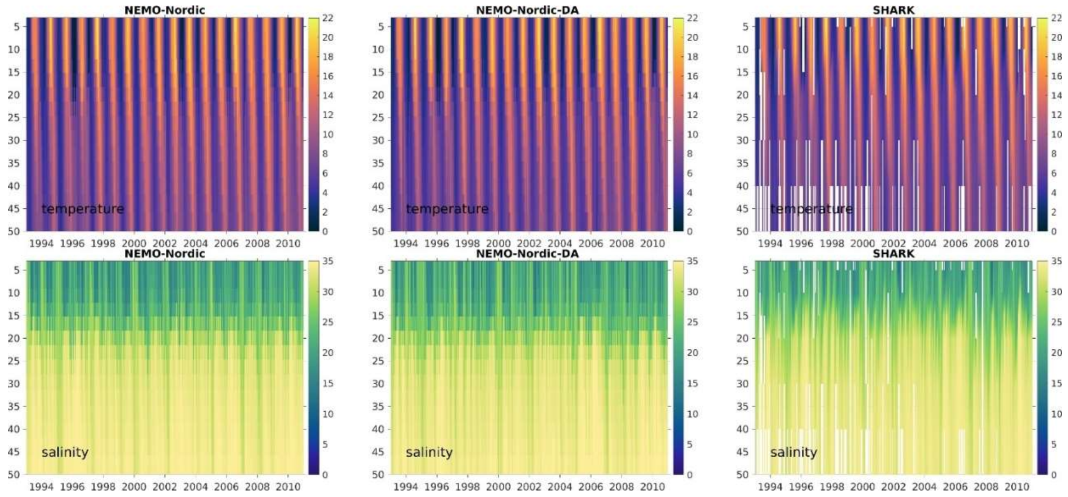

Figure 3. the temperature and salinity from model, DA simulation, and the SHARK observation for the period 1993 to 2010 in the ANHO (56.667°N, 12.117°E) in the Kattegat.

Can more details be provided about the data quality control measures applied to observations before assimilation? How do these controls influence data coverage and assimilation accuracy, especially in areas with sparse data?

All observed data underwent quality control prior to assimilation. To ensure high-quality assimilation results, an additional quality control was applied to the data before it was used for assimilation and validation. This included examining the differences between forecast and observed values, and excluding observations where the difference exceeded specified maximum deviations (3°C for temperature and 3 PSU for salinity). These thresholds were empirically determined based on past validation results of the model (Hordoir et al., 2019). Additionally, stations located on land, according to the NEMO-Nordic grid, were excluded, and duplicate records were removed.

When multiple observations were available for the same layer, the average value was used. This data control process helps correct the model fields reasonably and prevents instability in the simulation, which could arise from large adjustments in data assimilation when the difference between forecast and observation is too large. No specific controls were applied for areas with dense or sparse data. After applying quality control, the dataset was reduced by 1.7%.

The study emphasizes the climate relevance of Baltic Sea warming and salinity trends. What specific recommendations can be made for policy or management actions based on the observed trends?

Based on our findings on the Baltic Sea warming and salinity trends, there are several policy or management recommendations that could help address the impacts of these trends in the Baltic Sea, e.g. in the adaptation of fishery management, improved coastal and marine spatial planning, management of salinity changes and enhanced monitoring of ecosystem change related to the warming and salinity trends.

- Fishery management: the warming of sea temperatures can lead to shifts in fish populations, with some species thriving in warmer waters while others may decline. This is similar to salinity-sensitive species. Management should consider adapting fisheries policies to account for these changes. For instance, adjusting fishing quotas and seasons to align with the new distribution of species, while ensuring that vulnerable species are protected. Proactive measures such as the establishment of marine protected areas (MPAs) or fishing bans during critical breeding seasons could be considered.

- Integration of Climate Data: Incorporating long-term climate data into coastal and marine spatial planning is important for anticipating future changes and implementing preventive measures, such as adjusting infrastructure planning or developing resilient coastal communities.
- Sustainable Tourism and Development: As the Baltic Sea changes, there may be impacts on tourism, particularly for coastal areas that rely on biodiversity and specific environmental conditions. Sustainable tourism policies should be promoted to protect both the environment and local economies.

Were any sensitivity analyses performed on key model parameters or external forcing datasets to evaluate their influence on simulation accuracy and uncertainty in trend estimations?

The model configuration has been calibrated and validated against observations, as part of the NEMO-Nordic model development (Hordoir et al.,2019). However, no specific model optimization has been made to improve the trend analysis. Major trend uncertainty reduction in this paper is based on temperature and salinity assimilation.

How does this study's approach and findings compare to other similar long-term reanalysis efforts in the Baltic Sea? Are there specific methodological advantages that led to better accuracy or unique insights?

We have made a comparison with other reanalysis in the Baltic Sea in the discussion section 5.3. The dataset from both this study and other Baltic reanalysis provides valuable insights into the Baltic Sea’s physical conditions, but they differ significantly in terms of model setup, atmospheric forcing, data assimilation methods, and the choice of assimilated satellite products of SST.

- In the reanalysis, Fu et al., (2012) and Axell and Liu (2016) used a version of HIROMB which is old BALMFC Baltic forecast model. However, this study and CMEMS Baltic physical reanalysis use the state-of-the-art model of the Baltic region, NEMO-Nordic with different version.
- The study’s simulation is forced by atmospheric data from the UERRA reanalysis (11 km) and ECMWF ORAS4/5 reanalysis products along with a barotropic surge model for the northeast Atlantic. The resoluation of this atmospheric forcing is higher than other forcing for the Baltic reanalysis, For example, the CMEMS product is forced by the ERA5 dataset (31 km) and uses daily mean values from the CMEMS North West Shelf multi-year product at lateral boundaries.
- Different from other reanalyszed simulations, this study uses a multivariate EOF method to generate a sample ensemble of background error covariances, which effectively captures key large-scale spatial and temporal variability patterns while reducing dimensionality and noise. This method allows sampling different combinations of patterns from the ensemble space to increase the diversity of the ensemble and thus improve the prediction skill, leading to a more reliable simulation. Futhermore, a varying observation error with water depths in this study is benifitical for adjusting the models in waters deeper than the thermocline/ halocline.
- The study uses both OSISAF SST and IceMap satellite SST, whereas the CMEMS product uses Copernicus Marine Service L3S SST multi-year observations (SST_BAL_PHY_L3S_MY_010_040). Further, Fu et al. (2012) didn’t assimialted satellite SST and Axell and Liu (2016) only assimlated the IceMap SST. The OSISAF used in this study remedy the data coverage of IceMap in the Ice-free period.

These differences could lead to variations in the oceanographic predictions for the Baltic Sea, with potential impacts on temperature, salinity, and ice dynamics. As a result, this study showed a slightly better bottom simulation than those reported by Liu et al. (2013) for temperature and by Axell and Liu (2016) for salinity, reducing bias by 0.5 °C and 0.2 PSU in the central region of the eastern Gotland Basin, and 0.3 °C and 0.1 PSU in the Bornholm Basin. Additionally, this study outperforms previous ones that did not incorporate satellite observations. For instance, Fu et al. (2012) made an error in the baroclinic inflows of 2006, a mistake that was avoided in this research. When compared to the CMEMS Baltic Reanalysis Product (CBRP), the sea level anomaly in this study exhibits a stronger correlation with observations, except in Visby, Talinn, Grena, and Viken. A comparison of cost function revealed both datasets have comparable salinity quality at the same Baltic stations. However, there are notable discrepancies in bottom salinity in both the southern Bothnian Sea and the southwestern Bothnian Bay. In the southwestern Bothnian Bay, the quality of sea bottom salinty reported in this study was classified as good, whereas the CBRP was classified as poor. Conversely, in the southern Bothnian Sea, the CBRP data was of higher quality than the SBS in this study. Regarding temperature, the CBRP was of high quality across all stations, while this study reported slightly lower quality in the southern Bothnian Sea, though still within acceptable limits.

I appreciate the comprehensive approach taken in this study, particularly about the assimilation of temperature and salinity data. However, I believe that incorporating references to relevant works on streamflow assimilation could enrich the context and highlight parallel methodologies or complementary findings. For instance, streamflow assimilation studies often deal with similar challenges of data assimilation accuracy, bias correction, and spatiotemporal variability, which could offer valuable insights or comparisons. Including such references would provide a broader perspective on assimilation techniques and their applications, potentially benefiting both this study's conclusions and the wider research community. I strongly suggest citing below papers:

Mazrooei, A., Sankarasubramanian, A., Woodm A. W., 2021. Potential in improving monthly streamflow forecasting through variational assimilation of observed streamflow. Journal of Hydrology, 600. https://doi.org/10.1016/j.jhydrol.2021.126559

Zafarmomen, N., Alizadeh, H., Bayat, M., Ehtiat, M., Moradkha, H., 2024. Assimilation of sentinel-based leaf area index for modeling surface-ground water interactions in irrigation districts. Water Resources Research, 10, e2023WR036080. https://doi.org/10.1029/2023WR036080

Thanks, we add these two references in the revised manuscript.