RESPONSE TO REVIEWER 2:

Dear Dr. Benjamin Ménétrier,

We thank you for carefully reading our manuscript and giving useful comments. We revised the manuscript based on your comments. Our responses to your comments are described in the following, where your comments are italicized.

Major comments 1

As mentioned by referee #1 in his/her major comment, this paper is a restrictive application of the full MGBF method. With only one filtering level left, equation 10 seems similar to the NICAS method I developed independently in 2020 (https://doi.org/10.5281/zenodo.4058620, and mentioned in equation 10 of https://doi.org/10.5194/gmd-15-7859-2022). However, the NICAS method can use adaptive unstructured subgrids, handle complex boundaries, and produce inhomogeneous and anisotropic localization functions.

Response:

As you pointed out, the idea to apply the filter with the coarse resolution is the same as the NICAS. While the NICAS applies a localization matrix on the unstructured coarse filter grid and interpolates it to the analysis grid directly, the MGBF-based localization applies a compact-support filter on the structured coarse filter grid and interpolates it from the coarsest filter grid g_T to the analysis grid g_0 step by step. One advantage of the MGBF-based localization is high parallelization efficiency with the compact-support filter and the step-by-step interpolation. We will explain it in the revised manuscript. It is technically possible to produce inhomogeneous and anisotropic localization functions in MGBF (Purser et al. 2022) although to show the impact is the future task.

Major comments 2

This kind of explicit convolution method on coarse subgrids is computationally efficient when the localization length-scale is large compared to the analysis grid cell size, since the subgrid can be coarse. However, I agree with referee #1 that it can become very expensive for smaller localization length-scales, because in this case a fine subgrid must be kept to maintain the localization function sharpness.

Response:

As you pointed out, the computational cost of the analysis with small localization length in MGBF is not necessarily smaller than that in RF since the interval of the filter grid should be smaller than the localization length. We will explain this disadvantage in the revised manuscript.

Major comments 3

Another issue properly handled in the NICAS method and missing here is the localization normalization (i.e. diagonal coefficients of the localization matrix should all be equal to one). Figure 4 suggests that the MGBF method is perfectly normalized with all curves going to 1 at zero separation. However, I believe this is true only if the observation is located on a coarse grid node. Indeed, even if the continuous function B_p(x) is normalized (as mentioned after equation 11), the discrete low-resolution filters F_{BF} might not be, and even if they were, the final interpolation to then analysis grid would break this normalization. Only an outer diagonal scaling matrix taking all the operators (filters and interpolations) into account can ensure a proper normalization.

Response:

As you point out, the interpolation to the analysis grid slightly breaks the normalization, but the impact is negligible. The impact of the interpolation is shown as difference between MGBF00 (the filter grid is the same as the analysis grid) and MGBF04 (the filter grid is coarser than the analysis grid) in Fig. 4. The difference of the peak value is very small. We will revise the description of the normalization to make it more accurate.

Minor comments 1

In section 2.1, equation (2) is already an approximation of the general 3DEnVar formulation. Indeed, the authors are using the same 3D localization matrix for all the auto- and cross-localization blocks between different analysis variables. This method is sometimes referred to as "Mark Buehner's trick" (used in https://doi.org/10.1175/2009MWR3157.1 and clearly described in section 3.4.2. of https://doi.org/10.1002/qj.2325). It assumes that all the analysis variables have roughly the same error correlation length-scale. Whether this assumption holds here or not, I think it should be mentioned.

Response:

As you pointed out, this formulation uses the same localization matrix in all analysis variables. We will explain it in the revised manuscript.

Minor comments 2

In equation (10) of section 2.3, the rightmost interpolation operator (D from g1 to gt) is actually not required if only one grid and one scale are used, as DD^T = I. If several grids are needed (e.g. g2 and g4 as in experiment MGBF03SDL), this interpolation operator is required to combine the scales with operator E, but the destination grid should be the finest grid used (here g2), not necessarily g1.

Response:

As you pointed out, $\mathbf{D}_{g_T \leftarrow g_1}$ is not required in the single-scale localization. Actually, it was not calculated in the sensitivity experiments with single-scale localization in this study. We will explain it in the revised manuscript. Even in SDL, it can be changed to $\mathbf{D}_{g_T \leftarrow g_2}$ if the finest filter grid is g_2. However, this change to

$\mathbf{D}_{g_T \leftarrow g_2}$ hardly shortens the calculation time because of the load imbalance; g_2 is calculated in the limited number of processors (see section 4d in Purser et al. 2022). Therefore, we adopted $\mathbf{D}_{g_T \leftarrow g_1}$ in SDL.

Minor comments 3

Finally, I think that the experiments with slightly reduced length-scales (with a sigma suffix) are not really necessary. As shown in https://doi.org/10.1175/MWR-D-22-0255.1, the analysis quality is not very sensitive to the localization length-scale, as long as this length-scale is good enough. Given all the other uncertainties about the localization function shape and the fact that it should actually be anisotropic and inhomogeneous, the optimization of the localization length-scale does not seem really relevant here. Removing it (or better keeping it and removing the non-sigma case) would make the article a bit lighter and easier to read.

Response:

As you pointed out, the analysis quality in MGBF04σ was very similar to that in MGBF04. However, the difference of analysis quality was not negligible. Namely, MGBF04σ showed a smaller deviation from RF than MGBF04 (Figs. 6-9). Since this is one of the important conclusions of this study, we showed both MGBF04 and MGBF04σ and compared them. This result possibly implies that the analysis quality is occasionally sensitive to the localization length in the operational DA system unlike the idealized cases shown in Morzheld and Hodyss (2023).

References:

Purser, R. J., Rancic, M., and De Pondeca, M. S. F. V.: The multigrid beta function approach for modeling of background error covariance in the Real-Time Mesoscale Analysis (RTMA). Monthly Weather Review, 150(4), 715–732, https://doi.org/10.1175/MWR-D-20-0405.1, 2022.

Morzfeld, M., and Hodyss, D.: A theory for why even simple covariance localization is so useful in ensemble data assimilation. Monthly Weather Review, 151(3), 717–736, https://doi.org/10.1175/MWR-D-22-0255.1.