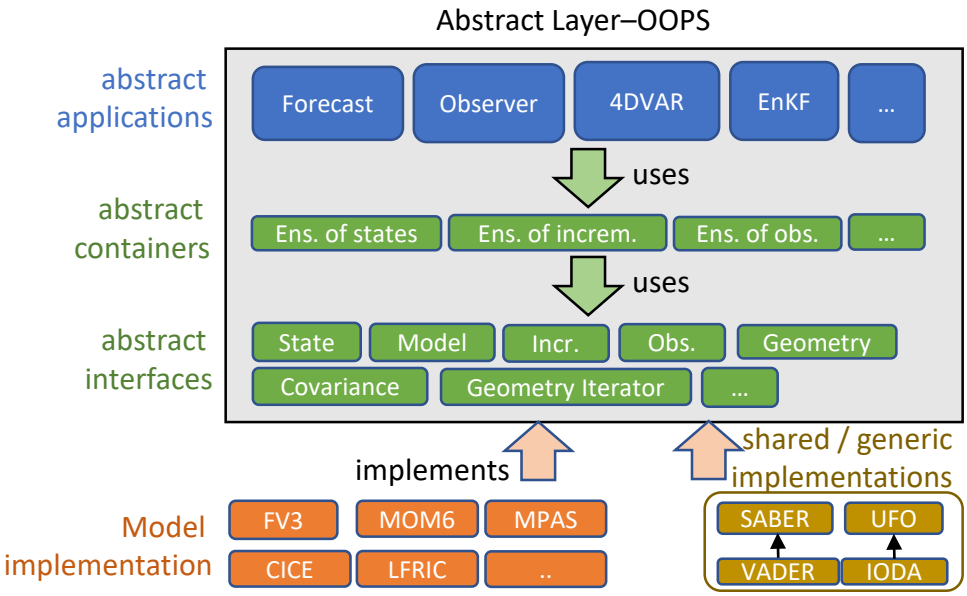

Figure 1: Separation of concerns in JEDI. Image adopted from JEDI academy slides (http://academy.jcsda.org/2021-10/slides/2021-06-21-WhyOOPSJEDI.pptx). Model implementation acronyms correspond to different model implementations and are defined in the main text of the paper. Other acronyms include: OOPS—Object Oriented Prediction System; SABER—System Agnostic Background Error Representation; UFO—Unified Forward Operator; VADER—Variable Derivation Repository; and IODA--Interface for Observational Data Access.

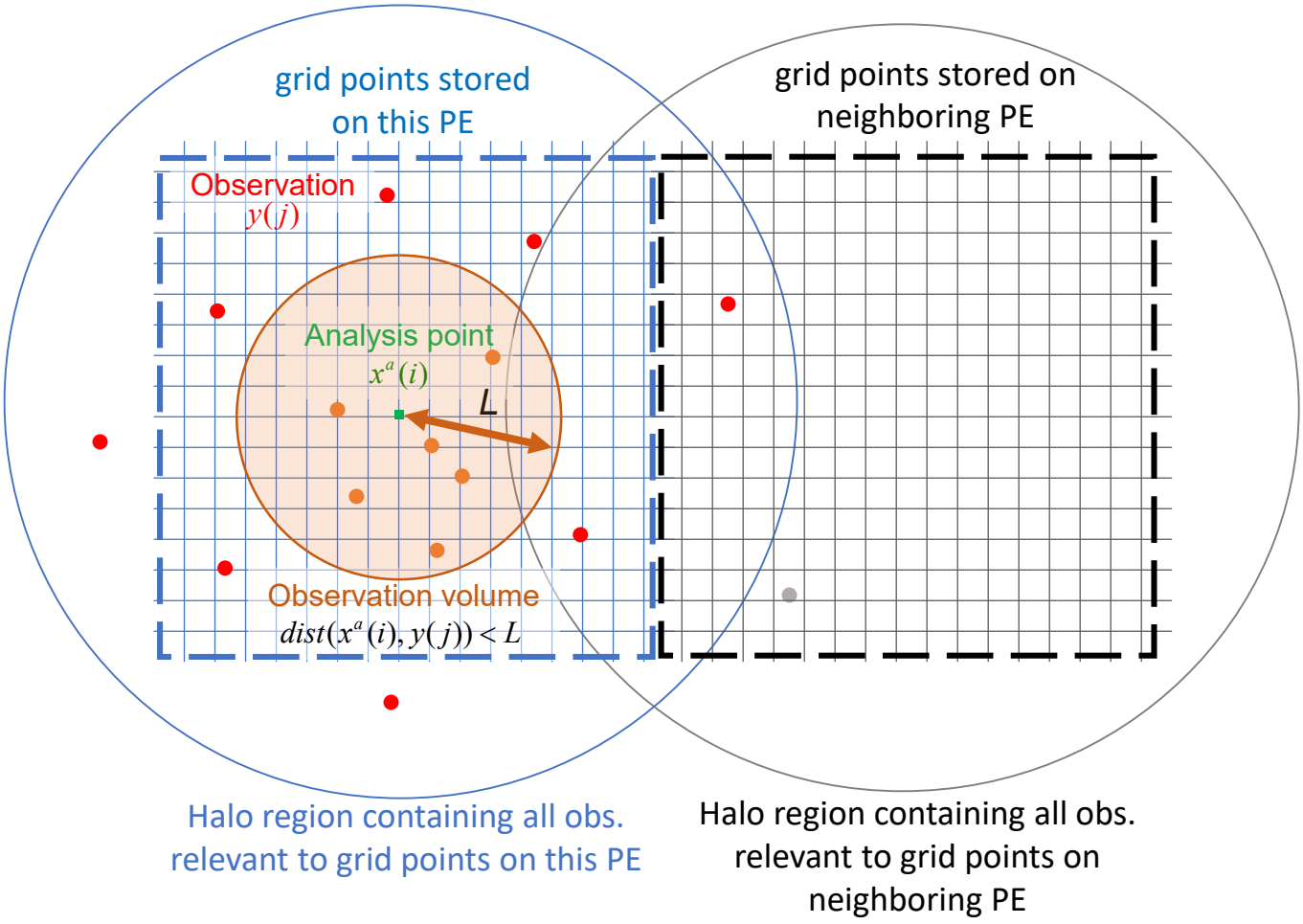

Figure 2: Local volume solver domain decomposition. Blue grid: grid points stored on the current (this) processing element (PE). Black grid: grid points on the neighboring PE. Green square indicates the updated, analysis point x(i). Dots indicate observations: (orange) observations used in the analysis; (red) observations that are needed to update all points on this PE; (gray) other observations that are not needed to update grid points on this PE. Circles indicate: (orange) the area used to collect local observations for the current local analysis, (blue) halo region that contains all possible observation locations needed to update grid points on this PE, (gray) observations halo region that contains all possible observation locations needed to update grid points on neighboring PE.

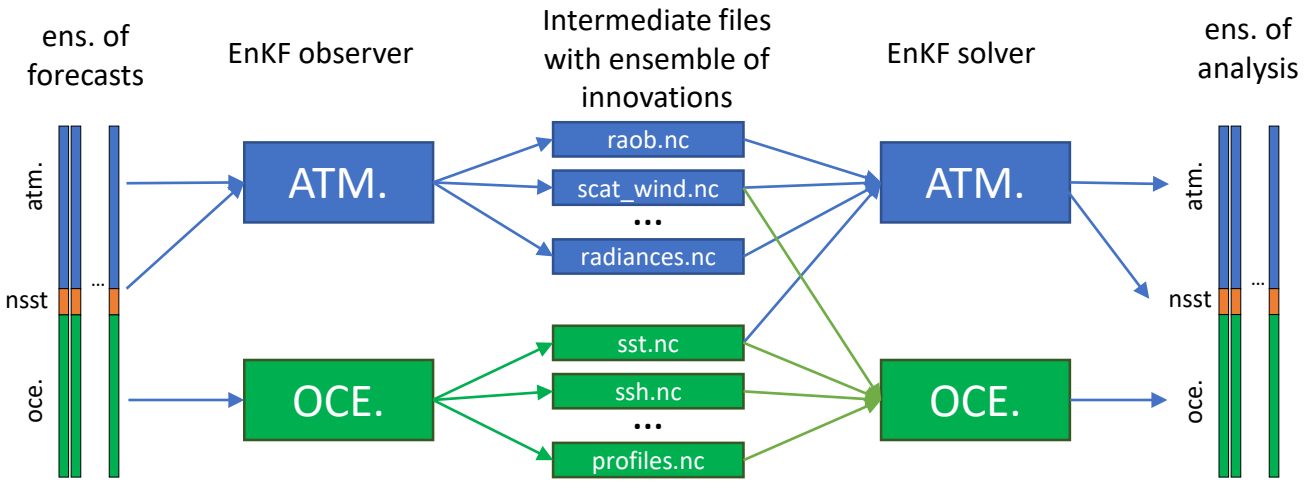

Figure 3: Example of interface solver implementation in JEDI framework. A pair of LETKF observers are run first obtaining ensemble observations of all atmospheric and oceanic variables. In this illustration, the near surface SST forecast (nsst) from the diurnal layer model is used as a first guess by the atmospheric observer for observations of the surface-sensitive radiances. Ocean state is used by the ocean observer to estimate the forecast of SST retrievals. Then a pair of LETKF solvers is executed one producing analysis on the atmospheric and the second on the oceanic grids. Each solver can use a combination of atmospheric and oceanic observations. In this example atmospheric solver is tasked to produce the analysis of the nsst variable as it has more complete picture of atmospheric observations that can potentially influence the nsst estimate.

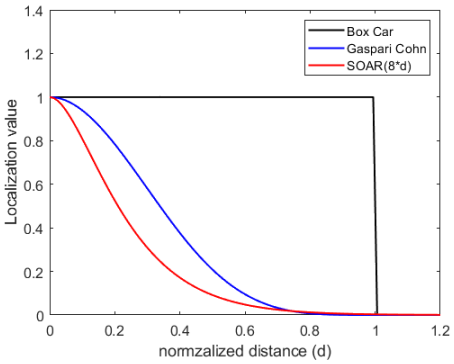

Figure 4: Localization as a function of normalized distance and localization type. For illustration purposes, distance was scaled by a factor of 8 before the SOAR function was calculated to achieve a near zero value for normalized distance of 1 (e.g. the SOAR function value is 0.003 for the distance of 8).

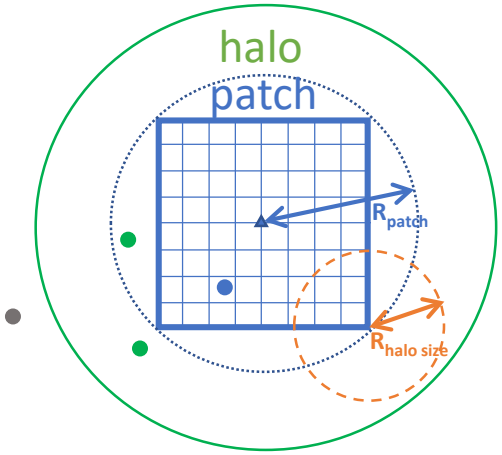

Figure 5: Halo distribution of observations. Shown are (blue grid) the grid points stored on this PE. (Blue triangle) halo center computed as a center of the grid points. (Blue circle) patch radius R_{patch} that encloses all the grid points on this PE; (Orange circle) halo size radius $R_{halo\ size}$ that is added to the patch circle. In practice, $R_{halo\ size}$ is related to the localization radius. (Green circle) the total halo circle that is computed as a summation of $R_{patch} + R_{halo\ size}$. Observations can then be divided in to patch observations that exclusively belong to this PE based on the shortest distance to the patch center (the blue dot). Halo observations (green dots) that are within the green total halo radius. And observations that are not stored on this PE because they are outside of the total halo radius (gray dots).

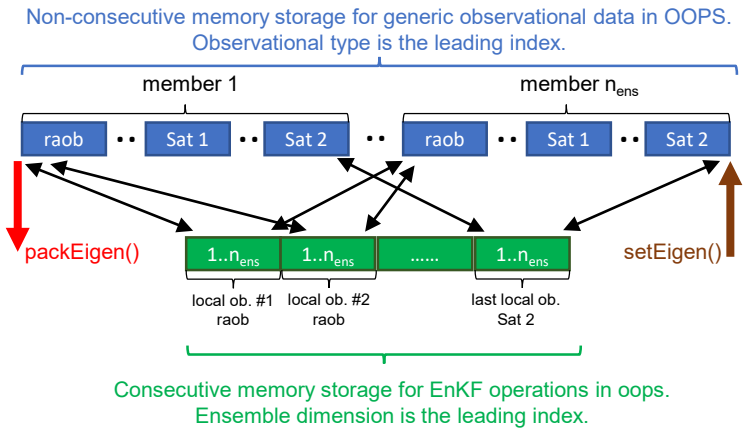

Figure 6: packEigen and setEigen methods operating on observational data in OOPS. Top row present data stored in generic containers for observational data in OOPS. Observational data stored as vectors of vectors, with each blue rectangle stored as contiguous data but with potential extraneous memory between blocks. Bottom row shows contiguous storage of ensembles of local observational data. The entire record is contiguous with the ensemble dimension as a leading index.

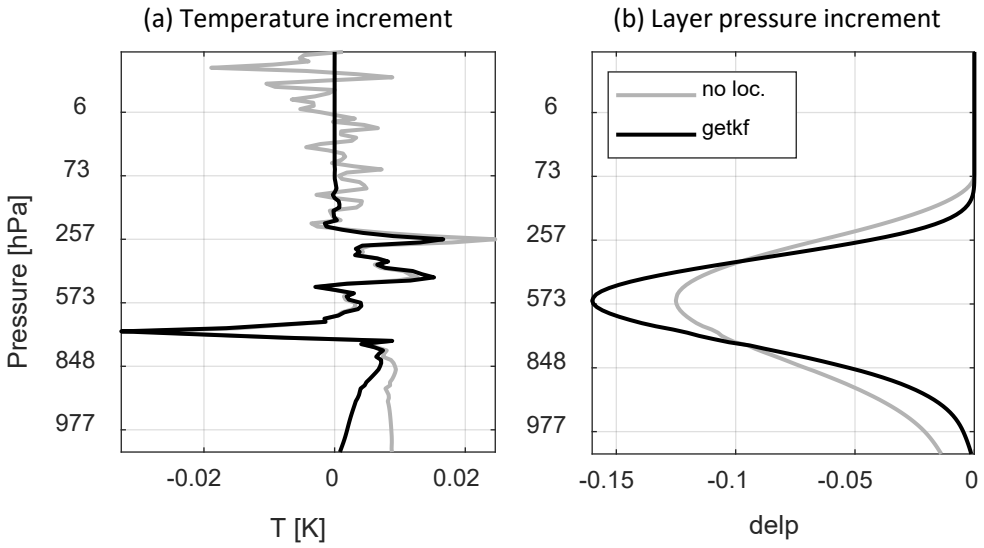

Figure 7: Increment to the vertical column of atmospheric temperature (a) and layer pressure (b) from assimilation of single surface pressure observation at 40.5°N, 160.5°E and innovation of 6 hPa. Background ensemble valid for 2015-12-05 18:00Z. GETKF vertical localization was used with localization scale of 30 levels and 10 eigen vectors capturing 96% of the variance.

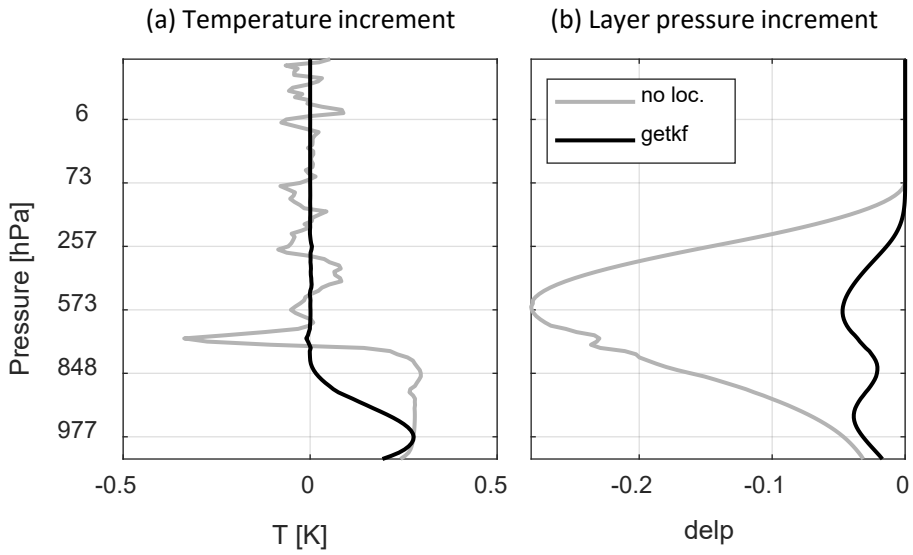

Figure 8: Increment to the vertical column of atmospheric temperature (a) and layer pressure (b) from assimilation of single temperature observations at 40.5°N, 160.5°E, 950 hPa, and innovation of 1 K. Background ensemble valid for 2015-12-05 18:00Z. GETKF vertical localization was used with localization scale of 30 levels and 10 eigen vectors capturing 96% of the variance.

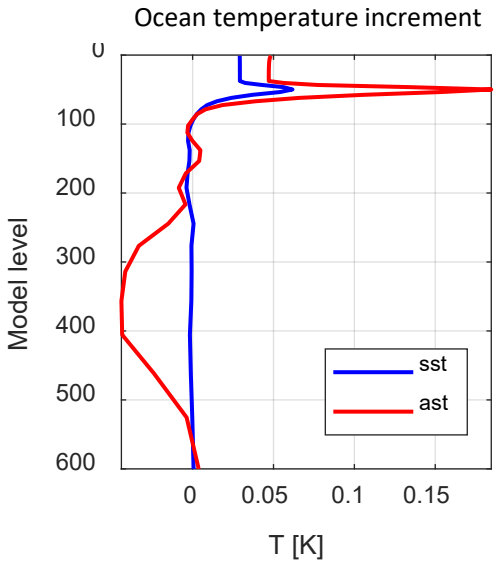

Figure 9: Increment to the vertical column of ocean temperature from assimilation of a single atmospheric temperature observations at 27.5°S, 154.5°W, 950 hPa (red) and sea surface temperature observation at the same horizontal location (blue). Both with the innovation of 1 K. Background ensemble valid for 2015-12-05 18:00Z. No vertical localization was used.

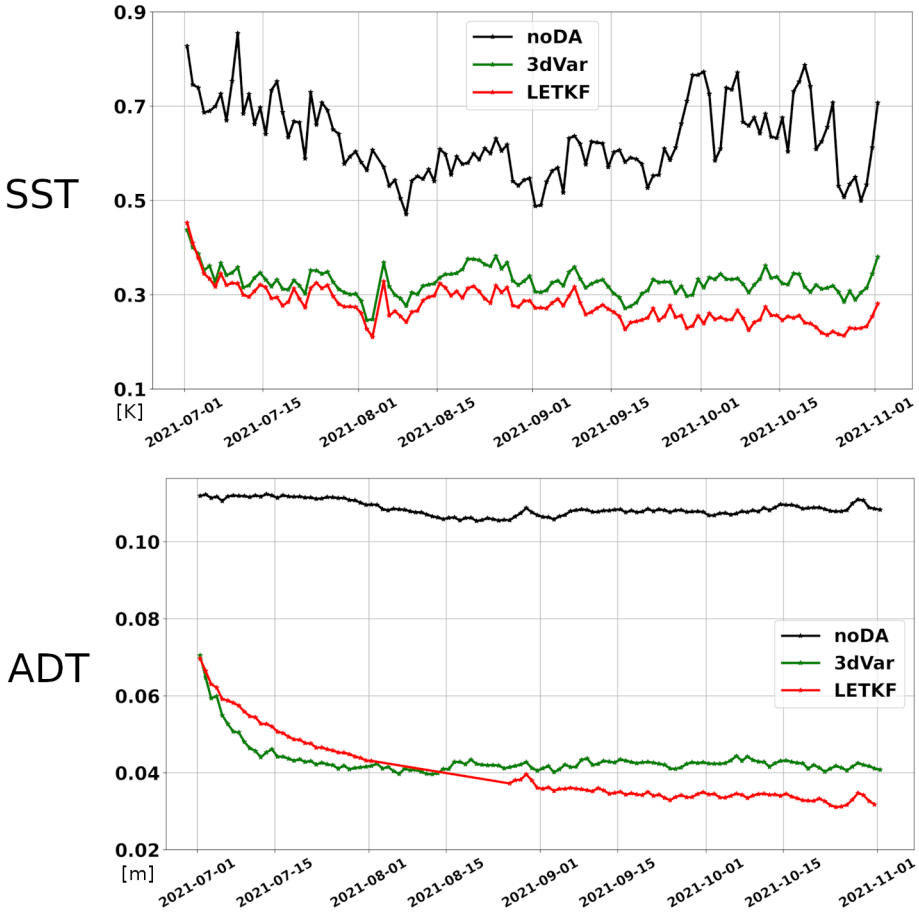

Figure 10: RMSE for the 24-hour forecast of (a) SST and (b) ADT observations. Colors indicate: (black) no data assimilation, (green) 3DVAR algorithm, and (red) LETKF.

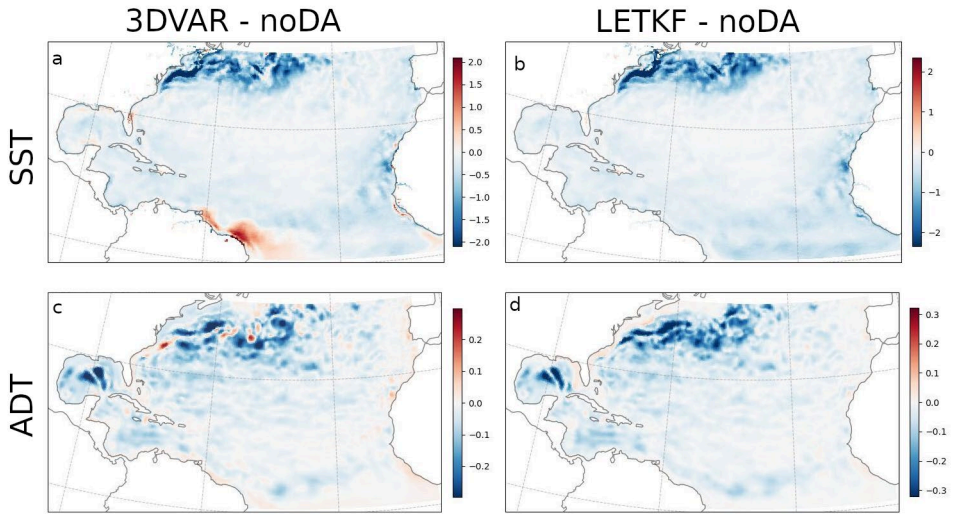

Figure 11: RMSE for the difference between run with no data assimilation and 3DVAR (left column, panels a and c) and LETKF (right column, panels b and d). Top row shows differences for SST (panels a and b) and the bottom row shows differences for ADT (panels c, d).

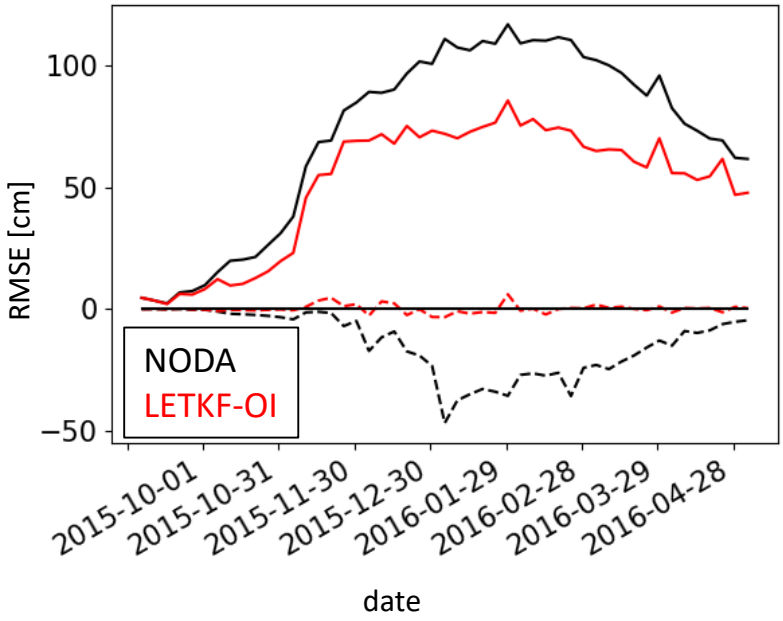

Figure 12: RMSE for the 6-hour forecast of the snow depth (GHCN network). Colors indicate: (black) no data assimilation, and (red) LETKF-OI assimilation of snow depth measurements. Line style indicates: (solid) standard deviation of error (dashed) mean error.

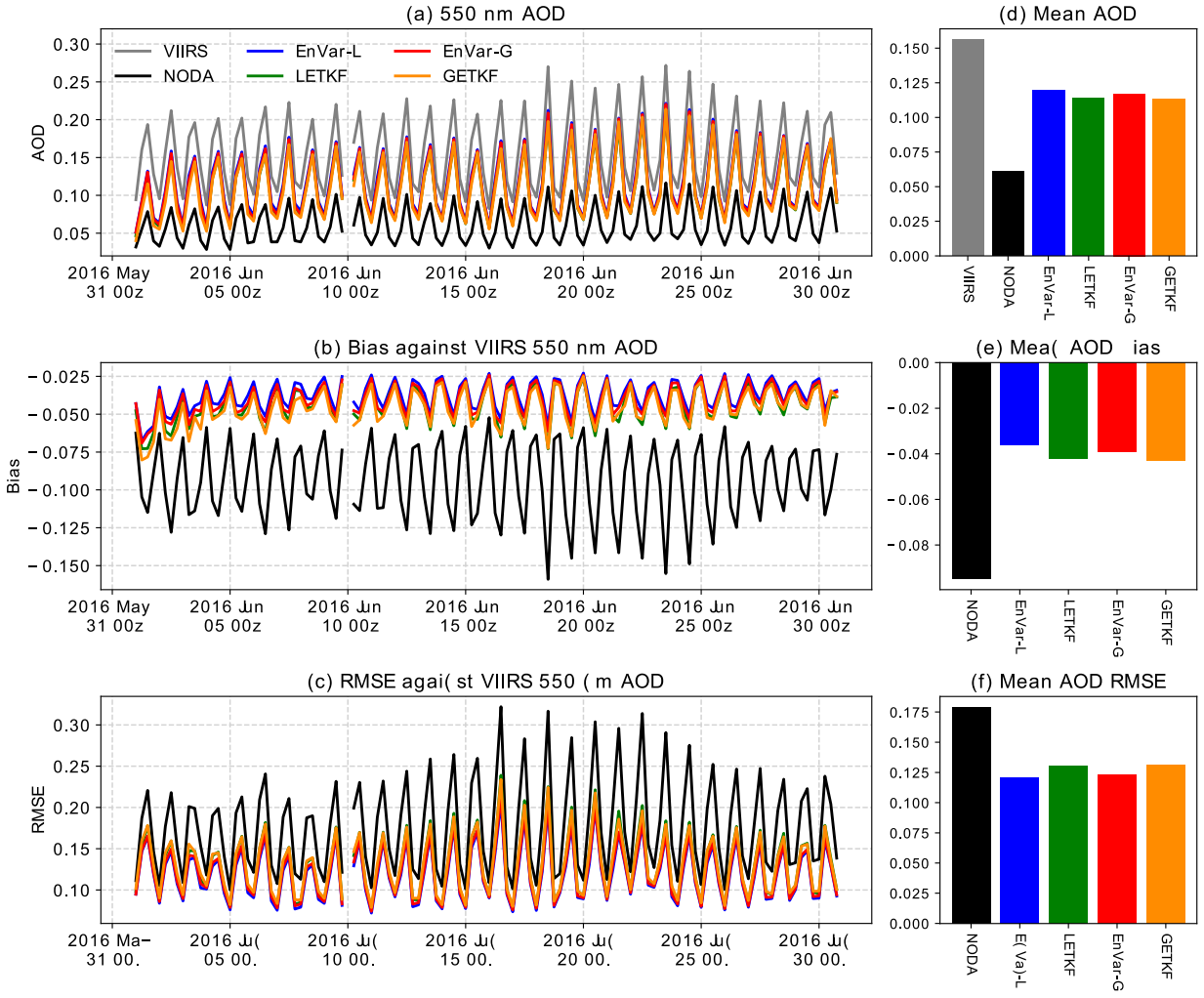

Figure 13: (a) Time-series of 550 nm AOD means from VIIRS (gray), no data assimilation six-hour forecast (black), the one-member control and ensemble mean analyses in the EnVar-L (blue and green, respectively) and EnVar-G (red and orange, respectively) experiments in a six-hourly interval in June 2016, and their corresponding (b) biases and (c) root-square-mean errors (RMSEs) against VIIRS AOD retrievals. (d)-(f) are the temporally means of the corresponding (a)-(c) over the last three weeks in June 2016.