Impact of the Window Length of Four-Dimensional Local Ensemble Transform Kalman Filter: A Case of Convective Rain Event

Yasumitsu Maejima$^1$ and Takemasa Miyoshi$^{1,2,3,4,5}$

$^1$RIKEN Center for Computational Science, Kobe, Japan
$^2$RIKEN Center for Interdisciplinary Theoretical and Mathematical Sciences Program, Kobe, Japan
$^3$RIKEN Cluster for Pioneering Research, Kobe, Japan
$^4$Department of Atmospheric and Oceanic Science, University of Maryland, College Park, Maryland, U.S.A.
$^5$Application Laboratory, Japan Agency for Marine-Earth Science and Technology, Yokohama, Japan

Abstract

This study aims to investigate the tradeoff between the computational time and forecast accuracy with different data assimilation (DA) windows of four-dimensional local ensemble transform Kalman filter (4D-LETKF) for a single-case severe rainfall event. We perform a series of Observing System Simulation Experiments (OSSEs) with 1-, 3-, 5- and 15-minute DA window in a severe rainstorm event in Kobe, Japan, on July 28, 2008, following the prior OSSEs by Maejima et al. (2019). Running 1-minute DA cycles showed the best forecast accuracy but with the highest computational cost. The computational cost could be reduced by taking a long DA window, but the forecast became less accurate even though the same number of observations were used. A significant gap was found between the 3-minute window and 5-minute window. With the 1- and 3-minute windows, the forecasts captured the intense rainfall, while with the 5-minute window or longer, the rainfall intensity was drastically underestimated. This single-case study suggests that 3-minute or shorter DA window be a promising method for a severe rainfall forecast, although more case studies are necessary to draw general conclusion.

(Citation: Maejima, Y., and T. Miyoshi, 2020: Impact of the window length of four-dimensional local ensemble transform Kalman filter: A case of convective rain event. SOLA, 16, 37–42, doi:10.2151/sola.2020-007.)

1. Introduction

Miyoshi et al. (2016a, 2016b) have developed innovative “Big Data Assimilation” (BDA) technology for innovating numerical weather prediction (NWP) focusing on local severe weather. The BDA system implemented 30-second-update, 100-m-resolution NWP, which is 120 times higher data assimilation (DA) frequency with 400 times more horizontal grid points than the hourly updated Japan Meteorological Agency (JMA) operational Local Forecast Model (LFM) at 2-km resolution. As a part of the BDA project, Maejima et al. (2019) performed a series of Observing System Simulation Experiments (OSSEs) using the Local Ensemble Transform Kalman Filter (LETKF, Hunt et al. 2004; Miyoshi and Aranami 2006) is useful. The three-dimensional LETKF (3D-LETKF) uses the observational information at a single time. In contrast, the 4D-LETKF enables to use observation data at multiple times simultaneously in a time window (a.k.a. DA window). For example, if we run 10 consecutive 30-second-update 3D-LETKF cycles for a 5-minute period, a single 4D-LETKF cycle with the 5-minute window can take the same amount of data. This will save the Input/Output (I/O) costs by less frequent initialization of the model runs. In general, I/O costs are significant in the LETKF cycles, and a longer DA window saves more of the I/O costs.

However, we have a drawback by having a longer DA window. Namely, a longer DA window would degrade the forecast accuracy due to the chaotic nature of the weather system dynamics. Generally, the weather system becomes more chaotic and unpredictable with a longer time period. The limit to predictability suggests the limit to information transfer within the DA window. If we take a longer DA window beyond the predictability limit, the observational information cannot transfer effectively within the time window, and the resulting forecast will be degraded. Namely, the longer the DA window, the less accurate the forecast. The time scale depends on the scales of the phenomena of interest. In this study, we focus on convective scales, so that the time scale will be limited to several minutes or even shorter.

In this study, we aim to examine the tradeoff between the computational time and forecast accuracy with different DA windows of 4D-LETKF. Following Maejima et al. (2019), we perform experiments with various lengths of the DA window in a specific rainstorm event, and evaluate computational costs and forecast accuracy. The results will suggest a reasonable length of the DA window for a severe weather forecast for this particular case, an important first step toward more general studies in the future.
and observational data for the OSSEs. Refer to Maejima et al. (2019) for more details of the series of experiments.

In the series of six OSSEs, 4D-LETKF technique is adopted. Hunt et al. (2004) described how to input observation data at multiple times in a single DA cycle. Figure 2 illustrates an example of having every-minute observations with (b) every-minute 3D-LETKF cycles and (c) every-3-minute 4D-LETKF cycles. Taking a longer DA window reduces the number of LETKF cycles while taking the same number of observations. An LETKF cycle contains I/O and MPI (Message Passing Interface) initialization which are required only once per LETKF cycle. Therefore, we can reduce the number of these processes by taking a longer DA window and can save the computational time. Namely, a longer DA window will be more efficient computationally while taking the same number of observations.

However, as mentioned in introduction, taking a long DA window would degrade the forecast accuracy if the DA window is long enough that the limit to predictability plays a role. Therefore, it is necessary to evaluate how effectively the LETKF works with longer DA windows.

To investigate the tradeoff between the computational time and forecast performances, we considered six scenarios for 15 minutes from 0245 UTC to 0300 UTC after the 1-minute-update DA cycles from 0230 UTC to 0245 UTC (see Fig. 2a): (A) 1-minute-update 3D-LETKF cycles as if there were no suspension, (B) five 4D-LETKF cycles with 3-minute windows, (C) three 4D-LETKF cycles with 5-minute windows, (D) a single 4D-LETKF with a 15-minute window, (E) a single step of 3D-LETKF at 0300 UTC rejecting all data during the suspension from 0246 UTC to 0259 UTC, and (F) no DA during the suspension. Experiments (A)−(E) use all available observations. After 0300 UTC, 30-minute extended forecasts initialized by the analysis ensemble means are performed.

The experiments were performed on the Fujitsu FX10 supercomputer of the University of Tokyo. We used 480 nodes (7680 cores) for the forecasts and 160 nodes (2560 cores) for the LETKF. Although the computational resources are not sufficient to run the DA computations in real time, we can address the main purpose of this study, that is, the tradeoff between the computational time and forecast performances.

3. Results

We first investigate the forecast accuracy and computational time for each experiment (Fig. 3). In the nature run, intense rainfall areas over 25 mm per 30 minutes spread from east to west. Experiment (A) used all of the PAWR and surface station data as if there were no suspension but shows generally weaker rainfall and different shapes mainly because of the difference of the model resolution. Namely, the nature run was performed at 100-m resolution. To generate 40 initial ensemble members, a 5-km-mesh downscaling simulation from the JMA Global Spectral Model (JMA-GSM) forecasts initialized at 0000 UTC July 27 was performed, and we chose initial ensemble members from the simulation at different times up to 13 hours away (see Maejima et al. 2019 for details). Every minute synthetic observations simulate the Phased Array Weather Radar (PAWR) at Osaka University and 167 surface weather stations at RIKEN Center for Computational Science and Kobe city elementary schools (Fig. 1b). The synthetic observations were generated by adding Gaussian noise to the observed quantities computed from the nature run. The assimilated observations and the magnitude of the observation errors are shown in Table 1. The error values are somewhat larger than the instrumental errors. Six OSSEs using NHM-LETKF at 1-km resolution with 40 ensemble members were initialized at 0230 UTC July 28, 2008 (Fig. 2a). Here, the model resolution was degraded by a factor of 10 compared with the nature run, and accordingly, the turbulence scheme is changed from Deardorff (1973) to improved Mellor and Yamada Level 3 (Nakanishi and Niino 2006, 2009). This is the only difference of the model settings. The entire domain of Fig. 1a shows the 300-km-by-300-km model domain for the OSSEs, and Table 1 provides the model settings

and observational data for the OSSEs. Refer to Maejima et al. (2019) for more details of the series of experiments.

In the series of six OSSEs, 4D-LETKF technique is adopted. Hunt et al. (2004) described how to input observation data at multiple times in a single DA cycle. Figure 2 illustrates an example of having every-minute observations with (b) every-minute 3D-LETKF cycles and (c) every-3-minute 4D-LETKF cycles. Taking a longer DA window reduces the number of LETKF cycles while taking the same number of observations. An LETKF cycle contains I/O and MPI (Message Passing Interface) initialization which are required only once per LETKF cycle. Therefore, we can reduce the number of these processes by taking a longer DA window and can save the computational time. Namely, a longer DA window will be more efficient computationally while taking the same number of observations.

However, as mentioned in introduction, taking a long DA window would degrade the forecast accuracy if the DA window is long enough that the limit to predictability plays a role. Therefore, it is necessary to evaluate how effectively the LETKF works with longer DA windows.

To investigate the tradeoff between the computational time and forecast performances, we considered six scenarios for 15 minutes from 0245 UTC to 0300 UTC after the 1-minute-update DA cycles from 0230 UTC to 0245 UTC (see Fig. 2a): (A) 1-minute-update 3D-LETKF cycles as if there were no suspension, (B) five 4D-LETKF cycles with 3-minute windows, (C) three 4D-LETKF cycles with 5-minute windows, (D) a single 4D-LETKF with a 15-minute window, (E) a single step of 3D-LETKF at 0300 UTC rejecting all data during the suspension from 0246 UTC to 0259 UTC, and (F) no DA during the suspension. Experiments (A)−(E) use all available observations. After 0300 UTC, 30-minute extended forecasts initialized by the analysis ensemble means are performed.

The experiments were performed on the Fujitsu FX10 supercomputer of the University of Tokyo. We used 480 nodes (7680 cores) for the forecasts and 160 nodes (2560 cores) for the LETKF. Although the computational resources are not sufficient to run the DA computations in real time, we can address the main purpose of this study, that is, the tradeoff between the computational time and forecast performances.

3. Results

We first investigate the forecast accuracy and computational time for each experiment (Fig. 3). In the nature run, intense rainfall areas over 25 mm per 30 minutes spread from east to west. Experiment (A) used all of the PAWR and surface station data as if there were no suspension but shows generally weaker rainfall and different shapes mainly because of the difference of the model resolution. Namely, the nature run was performed at 100-m resolution. To generate 40 initial ensemble members, a 5-km-mesh downscaling simulation from the JMA Global Spectral Model (JMA-GSM) forecasts initialized at 0000 UTC July 27 was performed, and we chose initial ensemble members from the simulation at different times up to 13 hours away (see Maejima et al. 2019 for details). Every minute synthetic observations simulate the Phased Array Weather Radar (PAWR) at Osaka University and 167 surface weather stations at RIKEN Center for Computational Science and Kobe city elementary schools (Fig. 1b). The synthetic observations were generated by adding Gaussian noise to the observed quantities computed from the nature run. The assimilated observations and the magnitude of the observation errors are shown in Table 1. The error values are somewhat larger than the instrumental errors. Six OSSEs using NHM-LETKF at 1-km resolution with 40 ensemble members were initialized at 0230 UTC July 28, 2008 (Fig. 2a). Here, the model resolution was degraded by a factor of 10 compared with the nature run, and accordingly, the turbulence scheme is changed from Deardorff (1973) to improved Mellor and Yamada Level 3 (Nakanishi and Niino 2006, 2009). This is the only difference of the model settings. The entire domain of Fig. 1a shows the 300-km-by-300-km model domain for the OSSEs, and Table 1 provides the model settings
Fig. 2. An example of two series of DA cycles with different DA windows: (a) Six scenarios during the 15-minute suspension between 0245 to 0300 UTC. Black arrows denote the observational inputs. (b) 1-minute-update 3D-LETKF cycles, (c) 3-minute-update 4D-LETKF cycles.

Fig. 3. (Top) Computational time of all experiments [seconds] from 0245 to 0300 UTC. Red and blue bars correspond to the timings for LETKF and forecasts, respectively. (Bottom) Accumulated surface rainfall amount [mm] in 30-minutes forecasts from 0300 to 0330 UTC. The nature run is shown in the left-most panel (red box).
mm in Experiment (A). This is considered reasonably well since we would expect systematic underestimation of rainfall intensity due to the resolution degradation. The 3D-LETKF with 1-minute window rapidly updates the analyses before significant nonlinear dynamics appear, and shows the best result. Comparing the different OSSEs, we find a significant gap of peak rainfall intensity between Experiments (B) and (C). Experiment (B) maintained the rainfall amount of 25 mm. However, when the DA window is extended to 5 minutes or more, the peak intensity was drastically decreased to less than 30% of Experiment (B). Although over 25 mm precipitation was found in Experiments (A) and (B), Experiments (C), (D) and (E) showed only less than 8 mm precipitation. Even though the number of input observations was the same, DA window clearly affects the forecast accuracy of this local severe rainfall event.

Next, we focus on the computational time (top panel bars of Fig. 3). We found tradeoffs between the computational cost and forecast accuracy. Experiment (A) achieved the best forecast accuracy in the price of the highest computational cost. Experiment (B) saved about a half of the computational time of Experiment (A), with similar rainfall intensity to Experiment (A). The computational costs of Experiments (C), (D) and (E) were consistently cheaper as expected, but with significantly lower forecast skills. Experiment (E) did not take observation data at all from 0246 UTC to 0259 UTC as described in Section 2, so that Experiment (E) was much cheaper than Experiment (D). However, Experiments (C), (D) and (E) did not predict heavy rainfall well as mentioned above.

To evaluate the forecast accuracy statistically, the root mean square errors (RMSE) for water vapor mixing ratio [g kg\(^{-1}\)] at the 2-km level of the terrain following vertical coordinate (z* = 2 km, Saito et al. 2006) was computed. This measure is chosen because water vapor in the lower troposphere is strongly related to precipitation (Maejima et al. 2019). The RMSE was obtained by the difference between the ensemble mean fields of the analyses and the nature run, and was computed in the same domain of the nature run (Fig. 2 of Maejima et al. 2019). Figure 4 shows the time series of the RMSE of Experiment (A). The location of the rich LWP areas X and Y in experiments (A) and (B) was similar to the nature run. The RMSE was computed in the same domain of the nature run (Maejima et al. 2019). The RMSE was obtained by the difference between the ensemble mean fields of the analyses and the nature run, and was computed in the same domain of the nature run (Fig. 2 of Maejima et al. 2019).

To summarize, a 3-minute or shorter DA window would be a good choice to maintain the forecast accuracy while reducing the computational cost based on the results from this particular single case.

We investigated six scenarios to evaluate the forecast accuracy and computational costs. Running full 1-minute cycles showed the best forecast accuracy but with the highest computational cost. The computational time could be shortened by taking a long DA window, but the forecast accuracy became worse even though the same number of observations were used. The 3-minute DA window saved about a half of the computational costs of the 1-minute DA window, but predicted the peak rainfall amount well. Therefore, 3-minute DA window is more cost-effective. Although 5-minute window or longer were cheaper, the heavy rainfall was not predicted well.

We found a significant gap between the 3-minute window and 5-minute window. The 3-minute window maintained similar rainfall amount to full 1-minute cycles. However, with the 5-minute window or longer, rainfall amount was drastically underestimated. The time series of the RMSE showed a gap between the 3-minute window and 5-minute window up to about 25-minute forecasts.

To summarize, a 3-minute or shorter DA window would be a good choice to maintain the forecast accuracy while reducing the computational cost based on the results from this particular single case.

We generalize the implication, further experiments with dif-
Fig. 5. (a) The analyses of Liquid water path [kg m\(^{-2}\)] and (b) mean-sea-level temperature [°C] at 0300UTC. The X and Y denote the active convections.
ferent cases would be necessary. Although this study investigates only a single case of heavy convective rainfall, we need to investigate other cases including different types of events. The BDA technology opened the door to very precise, 100-m-mesh, 30-second-update severe weather prediction, and this study provides the first step toward a failsafe workflow for real time application of the BDA-based NWP system.

Acknowledgements

The authors thank the members of Data Assimilation Research Team at R-CCS, Mr. Masaru Kunii and Dr. Hiromu Seko for insightful discussions. This work was supported by CREST, JST projects ‘EBD: Extreme Big Data – Convergence of Big Data and HPC for Yottabyte Processing’ (grant number: JPMJCR1303), ‘Innovating “Big Data Assimilation” technology for revolutionizing very-short-range severe weather prediction’ (grant number: JPMJCR1312) and JST AIP acceleration research ‘Big Data Assimilation and AI Creating New Development in Real-time Weather Prediction’ (grand number: JPMJCR19U2). This research was supported by CREST, JST Team at R-CCS, Mr. Masaru Kunii and Dr. Hiromu Seko for insightful discussions. This work was supported by CREST, JST Team at R-CCS, Mr. Masaru Kunii and Dr. Hiromu Seko for insightful discussions.

References

Deardorff, J. W., 1973: The use of subgrid transport equations in a three-dimensional model of atmospheric turbulence. J. Fluids Eng., 95, 429–438.

Hunt, B. R., E. Kalnay, E. J. Kostelich, E. Ott, D. J. Patil, T. Sauer, I. Szunyogh, J. A. Yorke, and A. V. Zimin, 2004: Four-dimensional ensemble Kalman filtering. Tellus A, 56, 273–277.

Hunt, B. R., E. J. Kostelich, and I. Szunyogh, 2007: Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter. Physica D, 230, 112–126.

Ikawa, M., and K. Saito 1991: Description of a non-hydrostatic model developed at the Forecast Research Department of the MRI. Tec. Rep. MRI, 28, 238 pp.

Kalman, R. E., 1960: A new approach to linear filtering and prediction problems. Trans. ASME J. Basic. Eng., 82, 35–45.

Kunii, M., 2014: Mesoscale data assimilation for a local severe rainfall event with the NHM-LETKF system. Wea. Forecasting, 29, 1093–1105.

Kusabiraki, H., Y. Koyama, T. Kanamori, and H. Seko, 2011: Line-shaped precipitation system observed in the Kinki area and heavy rainfall in the Toga River on 28 July 2008. Tenki, 58, 395–412 (in Japanese).

Lahoz, W. A., R. Brugge, D. R. Jackson, S. Migliorini, R. Swinbank, D. Lary, and A. Lee, 2005: An Observing System Simulation Experiment to evaluate the scientific merit of wind and ozone measurements from the future SWIFT instrument. Quart. J. Roy. Meteor., 131, 503–524.

Lorenz, E. N., 1963: Deterministic nonperiodic flow. J. Atmos. Sci., 20, 130–141.

Maejima, Y., T. Miyoshi, M. Kunii, H. Seko, and K. Sato, 2019: Impact of dense and frequent surface observations on 1-minute-update severe rainstorm prediction: A simulation study. J. Meteor. Soc. Japan, 97, 253–273.

Miyoshi, T., and K. Aranami, 2006: Applying a four-dimensional local ensemble transform Kalman filter (4D-LETKF) to the JMA nonhydrostatic model (NHM). SOLA, 2, 128–131.

Miyoshi, T., and S. Yamane, 2007: Local ensemble transform Kalman filtering with an AGCM at a T159/L48 resolution. Mon. Wea. Rev., 135, 3841–3861.

Miyoshi, T., M. Kunii, J. Ruiz, G.-Y. Lien, S. Sato, T. Ushio, K. Bessho, H. Seko, H. Tomita, and Y. Ishikawa, 2016: “Big Data Assimilation” revolutionizing severe weather prediction. Bull. Amer. Meteor. Soc., 97, 1347–1354.

Miyoshi, T., G.-Y. Lien, S. Satoh, T. Ushio, K. Bessho, H. Tomita, S. Nishizawa, R. Yoshida, S. A. Adachi, J. Liao, B. Gerofi, Y. Ishikawa, M. Kunii, J. Ruiz, Y. Maejima, S. Otsuka, M. Otsuka, K. Okamoto, and H. Seko, 2016: “Big Data Assimilation” toward post-petascale severe weather prediction: An overview and progress. Proc. IEEE, 104, 2155–2179.

Nakanishi, M., and H. Niino, 2006: An improved Mellor-Yamada Level-3 model: Its numerical stability and application to a regional prediction of advection fog. Bound.-Layer Meteor., 119, 397–407.

Nakanishi, M., and H. Niino, 2009: Development of an improved turbulence closure model for the atmospheric boundary layer. J. Meteor. Soc. Japan, 87, 895–912.

Saito, K., T. Fujita, Y. Yamada, J. Ishida, Y. Kumagai, K. Aranami, S. Ohmori, R. Nagasawa, S. Kumagai, C. Muroi, T. Kato, H. Eito, and Y. Yamazaki, 2006: The operational JMA nonhydrostatic mesoscale model. Mon. Wea. Rev., 134, 1266–1298.

Saito, K., J. Ishida, K. Aranami, T. Hara, T. Segawa, M. Narita, and Y. Honda, 2007: Nonhydrostatic atmospheric models and operational development at JMA. J. Meteor. Soc. Japan, 85B, 271–304.

Saito, K., 2012: The Japan Meteorological Agency nonhydrostatic model and its application to operation and research. In Tech. Atmos. Model Appl., 85–110.

Manuscript received 2 October 2019, accepted 17 January 2020 SOLA: https://www.jstage.jst.go.jp/browse/sola/