A Tropical Cyclone Initialization in Multi-Scale Localization with Hybrid Four Dimensional Ensemble-Variational System: Preliminary Results

Sujeong Lim\textsuperscript{1,2}, Hyo-Jong Song\textsuperscript{1,3}, and In-Hyuk Kwon\textsuperscript{1}

\textsuperscript{1}Korea Institute of Atmospheric Prediction System (KIAPS), Seoul, Korea
\textsuperscript{2}Department of Climate and Energy System Engineering, Ewha Womans University, Seoul, Korea
\textsuperscript{3}Department of Environmental Engineering, Myongji University, Seoul, Korea

Abstract

One way of the tropical cyclone (TC) initialization is assimilating the official advisory sea-level pressure observation to specify the initial structures of a TC with the aid of a background error covariance (BEC). In the hybrid four dimensional ensemble-variational data assimilation system, a static BEC explains the geostrophic and cyclostrophic wind-mass balance, and an ensemble BEC expresses the flow-dependent feature. Assimilation of the minimum sea-level pressure using a larger localization length-scale with limited ensemble members yields the closest to the observations at the initial state, but an imbalance in the broad analysis increment distorts geopotential and wind fields. Moreover, the reduced central pressure of TC is rapidly returned to an intensity that a model resolution can represent during the prediction. We introduce the application of final-scale localization (FSL) at the last outer loop with the shortest one to improve the TC initialization. With the aid of FSL, we may conduct the shorter localization length-scale, especially adopted for the TC initialization. As preliminary results, both analysis and prediction become more stable and the large-scale environments are preserved better than in the control experiment.

(Citation: Lim, S., H.-J. Song, and I.-H. Kwon, 2020: A tropical cyclone initialization in multi-scale localization with hybrid four dimensional ensemble-variational system: Preliminary results. SOLA, 16, 145–150, doi:10.2151/sola.2020-025.)

2. Methods

2.1 Numerical weather prediction model

The Korea Institute of Atmospheric Prediction System (KIAPS) has been developing the global NWP model called the Korean Integrated Model (KIM, Hong et al. 2018). It is composed of non-hydrostatic governing equations that use the spectral element method on a cubed sphere implemented with state-of-the-art physics parameterization packages. The horizontal resolutions of deterministic and ensemble models in this study are 25 km and 50 km, respectively, with 91 vertical levels up to 80 km model top; the ensemble size is 50.

2.2 Hybrid four dimensional ensemble variational data assimilation

H4DEV is one of the advanced DA schemes, which minimizes a cost function with a hybridized BEC between a static BEC and the ensemble BEC from a local ensemble transform Kalman filter (LETKF) (Hunt et al. 2007; Shin et al. 2016). While the static BEC explains the relationship of the geostrophic and cyclostrophic wind-mass balances, the ensemble BEC expresses the flow-dependent feature. The cost function in H4DEV (Song et al. 2018) is:

\begin{equation}
J(\xi,\alpha) = \frac{1}{2} \varepsilon^T \varepsilon + \frac{1}{2} \alpha^T \alpha + \frac{1}{2} \left[ y^o - h(x^o + U(\xi,\alpha)) \right]^T \Omega^{-1} \left[ y^o - h(x^o + U(\xi,\alpha)) \right]
\end{equation}

\begin{equation}
U(\xi,\alpha) = \beta U' \cdot U, \xi + \beta \sum_{k=1}^{K} [U''_{k+1} U'_{k+1} \alpha] \xi^o
\end{equation}

where $\varepsilon$ is the analysis increment, $x^o$ and $y^o$ are the background state and the observation, respectively, $\Omega$ is the background error covariance matrix, $U$ is the analysis update, and $\beta$ is the analysis error variance. For more details, please see Song et al. (2018).
where \( \xi \) is the control variable that consists of stream function, unbalanced velocity potential, unbalanced temperature and surface pressure and pseudo relative humidity; \( \alpha \) is an ensemble control variable determining localization of ensemble BEC; \( x^a \) is the observation; \( h(x) \) is the observation operator; \( x^b \) is a guess field for the analysis; \( U \) is an operator transforming the control variable into the model variable, which composed of the vertical \( (U_v) \), horizontal \( (U_h) \), and parameter transforms \( (U_p) \); \( O \) is the observation error covariance. For H4DEV, the localization function is generated as a gaussian type in wavenumber space and the localization is conducted in the same space (Song et al. 2018). Currently, the weights \( \beta_1 \) \((= 0.7) \), \( \beta_2 \) \((= 0.3) \) are used since the global performance was the best with them for the current system (Song et al. 2017), although a smaller weight on the ensemble BEC could be disadvantageous to represent the TC analysis. Except TC initialization is not conducted in LETKF, the same real observations are assimilated in both variational and ensemble systems (Kang et al. 2018).

As for the LETKF in H4DEV (Shin et al. 2018), the ICs for ensemble forecasts are generated by adding randomly selected perturbations from static BEC samples to the deterministic analysis (Kwon et al. 2018). To prevent the filter divergence, the adaptive multiplicative and additive inflations are implemented. It autonomously has the horizontal localization pieces given by a Gaussian-like piecewise fifth-order rational function (Gaspari and Cohn 1999; Miyoshi et al. 2011), varying \( 660 \) km to \( 1800 \) km depending on the level. The vertical localization functions are given by the Gaussian-like rational function with \( 0.1 \) \((\text{wind})\) or \( 0.2 \) \(\text{hPa (mass)}\) in logarithmic pressure for conventional data and a gradient of transmittance of the measured radiance for radiance data (Thépaut 2003).

### 2.3 Tropical cyclone initialization

The TC initialization (Kleist 2011) assimilates the estimates of MinSLP produced by the Regional Specialized Meteorological Center (RSMC) – Tokyo. 3-hourly generated TC advisory data covers the western North Pacific and the South China Sea (Muroi 2018). This single observation is treated as ordinary surface observation, but the quality controls (QC) are handled differently:

1. The conventional gross check is removed to prevent the observation from being filtered out due to a large observation innovation.
2. Only observations with good and fair QC indices can be used.
3. The observation over the land is not used because of the large uncertainty in surface pressure estimation.
4. The observation is assimilated when the background surface pressure of the TC center is higher than the TC advisory data. Since the gross check is in it is removed, it is hard to filter out the observation having the large observation innovation. This could produce excessive AI generating an unnecessary gravity wave and harming the large scale environment. It eventually occurs over the large uncertainty area to diagnose the TC in particularly in the ensemble system (Holt et al. 2015). To remedy excessive AIs, we devise an additional process:

1. The observation errors (OE) linearly vary with the TC intensity (Kleist 2011; Holt et al. 2015), then it diminishes with excessively fitted to observation (O). Based on the NMC statistics and previous papers, empirical values of \( 2-6 \) \(\text{hPa} \) were obtained when unwanted gravity waves do not appear:

   \[
   \text{OE} [\text{Pa}] = \frac{A}{1 + e^{\frac{1000 - 1000}{10}}} + B
   \]

   where \( A = 400 \) (\( A + B \), the maximum value) and \( B = 200 \) (the minimum value).

2. The observation innovation is empirically limited to \( 30 \) hPa at the maximum to suppress the initial shock. The incremental analysis updating (IAU, Bloom et al. 1996) is an alternative prescription, but currently, H4DEV does not have initialization methods.

3. The main idea in this study is the application of FSL to prevent excessive AI. The detailed method will be covered in the next subsection.

### 2.4 Multi-scale localization

Figure 1 illustrates the configuration of the multi-resolution OL composed of four iterations. During a cost function minimization, the OL contributes to updating the basic state with respect to which a nonlinear observation operator is linearized. At the first OL, static and ensemble AIs are generated with 42 and 10 total wave numbers, respectively, to fit the observation innovation. At the second (third) OL, they are generated with 85 (170) and 15 (20) total wavenumbers, and the last OL are made with 170 and 40 total wavenumbers.

In this study, we use LLSs of 5, 7, 10 and 20 total wavenumbers for ensemble BEC, corresponding to approximately \( 7200 \) km, \( 5400 \) km, \( 3600 \) km and \( 1800 \) km, in each OL. It determines how far the observation could be propagated by the background error correlation. A recent study suggests that we need to consider the observation scales, which are covered from large-scale to small-scale motions (Song et al. 2018), to make the best use of observation information. Since the TC advisory data, MinSLP, corresponds to the synoptic scale, we propose that the TC initialization should be assimilated at the last OL of the synoptic–scale localization (20 total wavenumber, the same as \( 1800 \) km). This approach enables the MinSLP observation to change the TC and its near-environment only, not large-scale features that are unwanted to be modified with the TC initialization.

### 2.5 Experimental design

Three experiments were performed to identify the response

![Fig. 1. A schematic diagram for the multiple-resolution outer-loop (OL) frame with different localization length-scales (LLS). The width of each box with the wavenumber in parentheses implies each LLS in physical space. As a result, it has a combination of the analysis increments localized with 7200, 5400, 3600, and 1800 km LLS, which correspond to truncation wavenumbers 5, 7, 10, and 20 in wavenumber space, respectively.](image)
of TC initialization in the frame of MSL (Table 1). Basically, the first experiment (NOTC) without TC initialization is performed to demonstrate why the TC initialization is needed in the H4DEV. The second experiment (MUL1) assimilates the single MinSLP over the TC center during the whole OL processes with approximately 7200, 5400, 3600, and 1800 km LLSs as for the ordinary observations. Otherwise, the last experiment (MUL4; i.e., FSL) assimilates the MinSLP at the last OL, whose minimization is conducted with the highest resolution (T40 and T170, for ensemble and static control variables, respectively) and shortest LLS (approximately 1800 km) for ensemble BEC. Here, the other observations are minimized starting from the first OL as before. Through the comparison between the two experiments (MUL1 and MUL4), it is identified that the TC initialization needs to be conducted in the last OL, which has the shortest LLS (about 1800 km) with the highest resolution (about 900 km, T40, for the ensemble, and 212 km, T170, for the static control variables).

H4DEV system was performed in a warm cycle run from 1800 UTC on 22 June 2018 to 1200 UTC on 11 July 2018 during which TC Maria (2018) existed. Figure 2 shows the best track of TC Maria (2018) that reached a super typhoon, which is classified into category 5 on the Saffir-Simpson scale. Since the discernible difference between the MUL1 and MUL4 is emerging from 18 UTC on 5 July 2018 in the intensified TC with few cycles, a detailed analysis was conducted at 0000 UTC on July 7 2018.

3. Results

3.1 TC analysis results

To demonstrate the advantage of the TC initialization at the last OL with the shortest LLS, we investigated which strategy is more appropriate for representing the TC structure. Figure 3 shows

![Table 1. Description of experiments: Yes (No) indicates that TC advisory data is used (not used) in the corresponding outer-loop (OL). The localization length-scale (LLS) in physical space is represented in parentheses. As a result, the MUL1 has a combination of the analysis increments (AIs) of TC advisory data localized with 7200, 5400, 3600, and 1800 km LLSs while the AI of TC advisory in MUL4 is only obtained from 1800 km LLS, namely final-scale localization.

|            | NOTC | MUL1       | MUL4       |
|------------|------|------------|------------|
| TC advisory data in 1st OL | No   | Yes (7200 km) | No         |
| TC advisory data in 2nd OL  | No   | Yes (5400 km) | No         |
| TC advisory data in 3rd OL  | No   | Yes (3600 km) | No         |
| TC advisory data in 4th OL  | No   | Yes (1800 km) | Yes (1800 km) |

![Fig. 2. The best track of TC Maria (2018) from http://www.typhoon.or.kr.](https://example.com/track)

![Fig. 3. AIs of MSLP (shading, hPa) and 850 hPa horizontal wind (vector, m s⁻¹) valid at 0000 UTC on 7 July 2018: (a) NOTC, (b) MUL1, and (c) MUL4.](https://example.com/ai)
the AIs of a mean sea-level pressure (MSLP) and a horizontal wind vector at 850 hPa. In the TC Maria (2018) area, we anticipate that the DA makes the TC stronger over the ocean where the observation is deficient. For the NOTC that does not conduct TC initialization, it generates an increased MSLP and an anti-cyclonic circulation increment (Fig. 3a). It weakens, rather than strengthens, the TC, because there was no observation to express the TC over the ocean. On the contrary, the TC initialization produces a decreased MSLP and a cyclonic circulation to strengthen the TC structure in cases of the MUL1 and MUL4. However, the magnitude and width of AIs depend on which OL the TC initialization is involved from. MUL1 decreases the MinSLP by 26 hPa and generates unnecessary wind vectors around the TC (Fig. 3b) since MUL1 has AIs obtained from longer LLS (Table 1). While the TC having a radius between 222 and 888 km is meso-α or synoptic scale, it is localized with longer length-scales in the first (7200 km), second (5400 km), and third (3600 km) OL in the frame of MSL. As a result, the single observation related to synoptic or less-scale has been over-propagated up to longer effective radii. On the other hand, the MUL4 experiment assimilates the MinSLP with at the last OL of the LLS of T20 or 1800 km; then it lowers the MinSLP by 13 hPa. Furthermore, a cyclonic wind vector is made while an unnecessary large-scale circulation is hard to be derived. The shorter LLS makes it reduced, which is the spurious error correlation due to the limited ensemble size. Figure 4 explains how these AIs are related to other atmospheric fields. To identify the results reasonable, we used the Integrated Forecast System (IFS) analysis data of the European Centre for Medium-Range Weather Forecasts (ECWMF), which is well known as the best quality of the forecast skill (https://apps.ecmwf.int/wmolcdnv/scores/mean/500_z). It has the 0.25-degree latitude and longitude grids with 25 pressure levels from 1000 hPa to 1 hPa. At 0000 UTC on 7 July 2018, the IFS analysis simulates the central pressure of TC Maria (2018) as 948.5 hPa. Although it is higher than the RSMC best track information, 935 hPa, it well simulates, for a global NWP model result, lower-level TC structure at the analysis time. Compared to the IFS analysis, the MUL1 shows the best result of the central pressure of TC, 953.7 hPa, which the MUL4 follows as 957.1 hPa; NOTC is 984 hPa. While the geopotential height (GPH) at 500 hPa forms a closed contour in the Figs. 4a and 4d, the NOTC as well as MUL1 (Figs. 4b and 4c) similarly shows a wiggle in GPH in the southwest region of the TC. In particular, MUL1 has shown a distorted shape of TC due to an excessive AI along with the left side of the TC (Fig. 4c). These features were also identified as similar in the analysis of another experiment, which is with larger $\beta_e (= 0.7)$, at 0000 UTC on 7 July 2018 (not shown).

Figure 5 shows the box plots of the RMSD of GPH at 500 hPa.
for NOTC, MUL1, and MUL4. In the vicinity of TC, excessive AI of MUL1 harms large-scale field and worsens the RMSD of GPH. However, the MUL4 may solve the imbalance problem, thereby tending to preserve the original GPH, resulting in the RMSD at a level similar to that of NOTC.

3.2 TC forecast results

We identified the 120-hour predicted tracks of TC Maria (2018) from 0000 UTC on 7 July 2018. The application of FSL (MUL4) at the last OL with the shortest LLS to TC initialization could not, however, reach to deriving improvements of prediction of the TC track (not shown). Next, we verified the time series of MinSLP for experiments, beginning at the same time (Fig. 6). At the initial time, the TC initialization (MUL1 and MUL4) leads to a decrease of about 30 hPa compared to NOTC; as a result, MUL1 produces MinSLP that was most similar to the RSMC best track; the MUL4 also generates MinSLP with similar performance. However, the reduced MinSLP could not be maintained during the forecast time especially in MUL1. In the case of MUL1, the strongest MinSLP among the experiments rapidly weakens during the initial 6-hours; however, its intensity is even weaker than NOTC after that. The excessive AI of 26 hPa is probably associated with an increase in acoustic and gravity wave activity, and the MinSLP eventually returns to a reduced intensity that the grid resolution can resolve (Kleist 2011). Otherwise, while the MUL4 shows the improved MinSLP at the analysis time, the reduced MinSLP is maintained until the 120-hour forecast lead time. The reduced central pressure of the assimilation of TC MinSLP is sustained during the prediction of the TC. Although the reduced MinSLP is increased from the analysis time, this amount is much smaller compared to MUL1. The changes in the track and intensity forecasts due to the multi-localization strategy for the TC initialization were modest but positive.

4. Conclusion

This study proposes the application of FSL at the last OL with the shortest LLS to improve the TC initialization. It is based on the idea that TC initialization should be assimilated with a LLS between meso-α and synoptic. Since the localization can reduce the propagation of unnecessary observation information (Anderson 2007), the shorter length-scale localization at the last OL with the minimum (1800 km) LLS is conducted. In TC Maria (2018) case, the mid-level GPH in the analysis shows a more stable TC structure and the reduced central pressure intensity is sustained during 120-hour forecast, compared to the experiment including the TC initialization with larger-scale localizations (3600, 5400, and 7200 km).

As a further study, experiments with increased ensemble ratio and resolution in H4DEV, including the TC initialization in LETKF as well, are planned. Besides, the additional impact of the IAU on suppression of initial shock can be investigated. With more TC cases, we anticipate the FSL in MSL will be more effective for TC initialization.

Acknowledgements

This work has been carried out as part of the R&D project on the development of global numerical weather prediction systems of the Korea Institute of Atmospheric Prediction System, funded by the Korea Meteorological Administration, and supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. 2020R1F1A1072315). The authors declare that there is no conflict of interest regarding the publication of this article. We acknowledge the comments of the anonymous reviewers and the editor.

Edited by: S.-H. Chen

References

American Meteorological Society (AMS), 2018: Glossary of Meteorology (2018). The version of 26 June 2018. (Available online at http://glossary.ametsoc.org/wiki/Main_Page, accessed 20 April 2020)

Anderson, J. L., 2007: Exploring the need for localization in ensemble data assimilation using a hierarchical ensemble filter. *Physica D: Nonlinear Phenomena*, 239, 99–111.

Barker, D. M., W. Huang, Y. R. Guo, A. J. Bourgeois, and Q. N. Xiao, 2004: A three-dimensional variational data assimilation system for MM5: Implementation and initial results. *Mon. Wea. Rev.*, 132, 897–914.

Bloom, S. C., L. L. Takacs, A. M. da Silva, and D. Ledvina, 1996: Data assimilation using incremental analysis updates. *Mon. Wea. Rev.*, 124, 1256–1271.

Gaspari, G., and S. E. Cohn, 1999: Construction of correlation functions in two and three dimensions. *Quart. J. Roy. Meteor. Soc.*, 125, 723–757.

Kang, J.-H, H.-W. Chun, S. Lee, H.-J. Song, J.-H. Ha, I.-H. Kwon, H.-J. Han, H. Jeong, and H.-N Kwon, 2018: Development of an observation processing package for data assimilation in KIAPS. *Asia-Pac. J. Atmos. Sci.*, 54, 303–318.

Kwon, I.-H., H.-J. Song, J.-H. Ha, H.-W. Chun, J.-H. Kang, S.
Lee, S. Lim, Y. Jo, H.-J. Han, H. Jeong, H.-N. Kwon, S. Shin, and T.-H. Kim, 2018: Development of operational hybrid data assimilation system at KIAPS. *Asia-Pac. J. Atmos. Sci.*, 54, 319–335.

Kleist, D. T., 2011: Assimilation of tropical cyclone advisory minimum sea level pressure in the NCEP global data assimilation system. *Wea. Forecasting*, 26, 1085–1091.

Hong, S.-Y., Y. C. Kwon, T.-H. Kim, J.-E. Kim, S.-J. Choi, I.-H. Kwon, J. Kim, E.-H. Lee, R.-S. Park, and D.-I. Kim, 2018: The Korean Integrated Model (KIM) system for global weather forecasting. *Asia-Pac. J. Atmos. Sci.*, 54, 267–292.

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

Miyoshi, T., 2011: The Gaussian approach to adaptive covariance inflation and its implementation with the local ensemble transform Kalman filter. *Mon. Wea. Rev.*, 139, 1519–1535.

Muroi, C., 2018: Brief history and recent activities of RSMC Tokyo-Typhoon Center. *Tropical Cyclone Res. Rev.*, 7, 57–64.

Orlanski, I., 1975: A rational subdivision of scales for atmospheric processes. *Bull. Amer. Meteor. Soc.*, 56, 527–530.

Shin, S., J.-S. Kang, and Y. Jo, 2016: The local ensemble transform Kalman filter (LETKF) with a global NWP model on the cubed sphere. *Pure Appl. Geophys.*, 173, 2555–2570.

Shin, S., J.-H. Kang, H.-W. Chun, S. Lee, K. Sung, K. Cho, Y. Jo, J.-E. KIM, I.-H. Kwon, S. Lim, and J.-S. Kang, 2018: Real data assimilation using the local ensemble transform Kalman filter (LETKF) system for a global non-hydrostatic NWP model on the cubed-sphere. *Asia-Pacific J. Atmos. Sci.*, 54, 351–360.

Song, H. J., S. Shin, J.-H. Ha, and S. Lim, 2017: The advantages of hybrid 4D-EnVar in the context of the forecast sensitivity to initial conditions. *J. Geophys. Res.*, 122, 12226–12244.

Song, H.-J., J.-H. Ha, I.-H. Kwon, J. Kim, and J. Kwun, 2018: Multi-resolution hybrid data assimilation core on a cubed-sphere grid (HybDA). *Asia-Pacific J. Atmos. Sci.*, 54, 334–350.

Song, H.-J., 2019: Existence of multiple scales in uncertainty of numerical weather prediction. *Sci. Rep.*, 9, 1–10.

Thépaut, J.-N., 2003: Satellite data assimilation in numerical weather prediction: An overview. *Proceedings of the Annual Seminar on Recent Development in Data Assimilation for Atmosphere and Ocean, 8–12 September 2003*, Reading, United Kingdom: ECMWF, 75–95.

Torn, R. D., 2010: Performance of a mesoscale ensemble Kalman filter (EnKF) during the NOAA high-resolution hurricane test. *Mon. Wea. Rev.*, 138, 4375–4392.

*Manuscript received 20 April 2020, accepted 22 June 2020*

SOLA: https://www.jstage.jst.go.jp/browse/sola/