1. INTRODUCTION

Rapid progress in remote sensing technologies including radar meteorology has brought us increasingly high-resolution and high-frequency observational data. The huge amount of detailed information of the atmospheric state contained in these data is expected to have a great potential to improve weather predictions. It is particularly important to effectively utilize the dense observations in numerical weather prediction in order to enhance accuracy in forecasting high-impact severe weather events, often involved with small scale phenomena in space and time.

However, data assimilation schemes to incorporate observational information into NWP are not necessarily sophisticated enough to appropriately handle these data. In particular, the spatial and temporal correlations of observation error become strong in dense observational data. Conventional data assimilation schemes are based on an assumption that the observation error correlation is negligible, allowing simple implementation and reduced computational cost. In order to meet this assumption, a severe thinning is usually applied to observations, discarding considerable part of the information contained in the data. Thus, it is an important issue to develop a methodology to appropriately handle correlated observational information in data assimilation.

Weather radar provides one of the important remote sensing observations, serving as a source of detailed information on the atmospheric situation. The Japan Meteorological Agency (JMA) operates the weather radar network consisting of 20 C-band weather Doppler radars and the aviation weather observation system including 9 C-band airport weather Doppler radars. Radial wind data from these radars, along with reflectivity and precipitation data, are assimilated in initializing the 5km-operational limited-area model, called the Meso-scale Model (MSM, JMA (2019)). In the data assimilation system, called the Meso-scale Analysis (MA), the four-dimensional variational (4D-Var) scheme based on JNoVA (JMA Nonhydrostatic model-based Variational Data Assimilation; Honda et al. 2005) is adopted. The diagnosis uses the methodology proposed by Desroziers (2005), estimating the error correlation from statistics of observation-minus-background and observation-minus-analysis residuals. This methodology has previously been applied to estimate observation error correlation of radial winds by Watterlot et al. (2012) and Waller et al. (2016b) on their operational three-dimensional variational data assimilation system. As discussed in Watterlot et al. (2012), Waller et al. (2016a, 2016b) etc., this methodology is only valid when the analysis update, which is determined from observation- and background-error covariances used in the assimilation, is consistent with the true error covariances. However, under this limitation, it is widely used to obtain results indicative of the properties of the observation error correlation.

The 3-hourly 4D-Var data assimilation cycles were performed over a period from 1 to 8 July 2018 to generate statistical samples. The optimization of the 4D-Var runs at a horizontal grid spacing of 15 km, generating the analysis increments to be added to the first guess at a resolution of 5 km. The super observation scheme and thinning are applied to the radial velocity data as in the operational configuration. Only observations at the hour are assimilated. On the other hand, the diagnosis uses all the super observations available at every cell (5 km by 5.625 deg.) and every 10 minutes to probe the detailed structure of the error correlation.

The figure 1 (a) displays the diagnosed observation error correlations along the beam for the weather Doppler radar at Sapporo (43.14N, 141.01E) at the elevation angle of 1.1 deg., averaged over all the azimuth angles. The diagnosed correlation has a half width of approximately 10 - 20 km around the diagonal. Although the result is noisy because of a limited number of statistical samples, the correlation distance looks to increase with range. This increase is consistent with the result by Waller et al. (2016b), shown to be likely to result from larger measurement volumes at far ranges. Consideration of the time correlation also is important in 4D-Var data assimilation. The figure 1 (b) displays the diagnosed time correlation of the observation error at the range of 75 km and the elevation angle of 1.1 deg. It is noted that the correlation from the statistics itself is not symmetric.
3. SIMPLE VARIATIONAL DATA ASSIMILATION WITH OBSERVATION ERROR CORRELATION

Effects of observation error correlations on data assimilation is investigated using a simple variational scheme. The two-dimensional variational (2D-Var) data assimilation runs on a space spanned by beam-range and azimuth-angle directions with 30 x 64 grid points, corresponding to a resolution of 5 km x 5.625 deg. Innovations of radial winds at 3 Jul. 2018 06UTC from the Sapporo radar at the elevation angle of 1.1 deg. are assimilated. For simplicity, observation operator is taken to be 1.

Observation error correlations are approximated using Gaussian functions, \( e^{-0.5(\Delta/\sigma)^2} \), where \( \Delta \) indicates a separation between the points. Based on the statistics in the previous section, \( \sigma \) is taken to be 15 km in the beam direction and 14.7 km in the azimuth-angle direction (corresponding to a single grid spacing at a beam range of 150 km). The correlation length in azimuth angle direction is inverse proportion to the beam range in the grid space. The background error correlation lengths are set to be smaller values compared to the actual situation, 10 km x 14.7 km, in order to clearly see the effect from the observation error correlation. Observation and background error variances are taken to be the same, (1 m/s)².

In implementation of the correlated observation error in variational assimilation, calculation of inverse of the observation error covariance is required. Full calculation of the inverse of the gaussian covariance matrix resulted in severe spurious noise. In order to avoid this, modes with larger eigenvalues are used, up to the mode with accumulated sum of eigenvalues reaching 99% of the full trace. In this case with 562 observations, 122 modes are used in the inverse calculation. It is a future work to investigate sensitivity of the numbers of available modes to different modelling functions of covariance.

Figure 2 (a) displays analysis increments (analysis – the first guess) thus obtained from the 2D-Var assimilation taking into account the observation error correlations. The increments reflect detailed structure of the spatial distribution of innovations (observation – the first guess) evenly over whole the domain (Fig. 2 (b)). Neglecting the observation error correlations, detailed structure is still present in the increment (Fig. 2 (c)). However, larger increments dominate where innovations with identical sign are grouped, giving the analysis very close to the observations. Multiplying the observation error standard deviation by 3 (Fig. 2 (d)), peak values of the increment gets closer to the case with the full observation error covariance, but the increment pattern is smoothed out, and the detailed structure is lost. Figure 2 (e) displays increments from another experiment without observation error correlation, but with a thinning applied on the observations with an equal interval, taking every 3 grid points in the both directions (Fig. 2 (f)). The increments again have peak values closer to Fig. 2 (a), but with a smoothed structure. Figure 2 (g) displays increments from another assimilation without the observation error correlation, but thinning observations by sequentially searching those within 20 km intervals (Fig. 2 (h)). The increments are still smooth compared to the full case (Fig. 2 (a)), lacking some of the observational information, but recovers many of signals located at far ranges and those isolated.

Using a full observation error covariance offers a detailed reflection of innovation distributions. Careful design of thinning and inflation is required, taking into account density of observations, grouped distribution
of similar innovations, resolution of the assimilation, etc., in order to appropriately keep information from observations.

4. 4D-VAR AND HYBRID 4D-VAR ASSIMILATION WITH TEMPORAL AND SPATIAL CORRELATED OBSERVATION ERROR

In this section, the temporal and spatial correlation of radial velocity observation error is applied to the JNoVA 4D-Var. As an investigation using the operational data assimilation system, Simonin et al. (2019) reports on experiments with a correlated observation error using their 3D-Var system. In this study, the radial velocity observations from the Sapporo radar are assimilated using the 4D-Var with a 3 hour assimilation window to generate an analysis valid at 3 Jul, 2018 06UTC.

The observation error correlation scales, \( \sigma \), in the beam direction and the azimuth-angle direction are
taken to be 15 km and 14.7 km, respectively, as in the previous section. Errors on different elevation angles are assumed to be uncorrelated. As a 4D-Var system, the observation error also is assumed to be correlated in time with a scale of 45 minutes, based on the statistics in section 2. On the other hand, the observation error variance is taken to be about $(3 \text{ m/s})^2$ simply as in the operational MA with a diagonal error covariance, not based on or optimized to the present statistics.

The first guess of the experiment is generated by a 4D-Var analysis-forecast cycle at a horizontal grid spacing of 5 km initiated at 1 Jul. 2018 00UTC. The incremental approach is used, running the 4D-Var optimization at a lower resolution of 15 km. All the observation data assimilated in the operational MA, including conventional observations, ground-based and satellite remote sensing data etc., are assimilated in the cycle without observation error correlation.

Introducing a flow-dependent background error can help better extract effective information from the observations. The control variables of JNoVA are extended to include flow dependency from ensemble perturbations (Buehner 2005). The ensemble perturbations are generated using an EDA (Ensemble of Data Assimilations, Isaksen et al. 2010). The EDA runs 6 analysis-forecast 4D-Var cycles assimilating all the MA observations at the same resolution as the deterministic 4D-Var. All the observations except for precipitation rate data are randomly perturbed according to their diagonal observation error covariance. The 6 cycles are initiated at 1 Jul. 2018

Figure 3: Increments of U wind component at 850 hPa. Hourly propagation of increments are displayed from the beginning (FT (forecast time) = -3) to the end (FT = 0) of the 3-hour assimilation window. The 4-hour forecasts are also displayed on the right. (a): 4D-Var with observations assimilated every 10 minutes. (b): 4D-Var with observations assimilated every 60 minutes. (c): hybrid 4D-Var with observations assimilated every 10 minutes. (d): hybrid 4D-Var with observations assimilated every 60 minutes.
00 UTC, and forecasts from the latest 9 initial runs valid at 3 Jul. 2018 03UTC form a 54 member ensemble to be used in the 3 Jul. 06UTC analysis. Experiments are performed with 4 different configurations. All the experiments assimilate only radial wind data from the Sapporo radar, averaged over 5 km by 5.625 deg. cells, taking into account the temporal and spatial correlations of observation error described above.

a) 4D-Var (without flow-dependent background error). Radar observations are assimilated every 10 minutes from 03:10 to 06:00 UTC 3 July 2018.

b) As in a), but radar observations are assimilated every 60 minutes from 04:00 to 06:00 UTC 3 July 2018.

c) Hybrid 4D-Var with the flow-dependent background error. The EDA and climatological background errors are equally weighted. A Gaussian localization function $\epsilon [-0.5(\Delta/\sigma)^2]$ with $\sigma = 75$ km is applied to the EDA background error covariance. Radar observations are assimilated every 10 minutes.

d) As in c), but radar observations are assimilated every 60 minutes.

Figure 3 displays the increments of U wind component at 850 hPa. Increments from the hybrid 4D-Var (Fig. 3 (c) and (d)) have a detailed structure from the beginning of the assimilation window. The hourly evolution of the increments also shows patterns continuous in time, indicating that the system is likely to have a potential to reflect balanced scenarios in detail through the assimilation. On the other hand, the 4D-Var gives less structured increment (Fig. 3 (a) and (b)), especially at the beginning of the assimilation window (FT = -3). Flow dependent patterns gradually grow through the 3 hour assimilation window, but they have still crude...
patterns even at the end of the assimilation window (FT=0) as compared to the hybrid 4D-Var cases.

Comparing the hybrid 4D-Var runs with observations assimilated every 10 minutes (Fig. 3 (c)) and every 60 minutes (Fig. 3 (d)), changes in the pattern of the increments are not as drastic as the changes between 4D-Var and hybrid 4D-Var. However, southeastern part of the increments is different throughout the assimilation window (shown by the red arrows). The difference propagates to the east, and reaches to the southeast of Hokkaido at 4 hour forecast. Also, the assimilation in finer time interval (c) is likely to give more small scale patterns from the beginning of the assimilation window (FT = 3).

In the 4D-Var cases (Fig. 3 (a) and (b)), difference mainly lies in the northeastern part of the increments. At 4 hour forecast, increments from the run with observations assimilated every 10 minutes (Fig. 3 (a)) overall have larger amplitude than the case with 60 minutes (Fig. 3 (b)). This looks more consistent with hybrid 4D-Var runs (Fig. 3 (c) and (d)), including line-shaped positive increment in northern Hokkaido (shown with black arrows).

Figure 5 displays RMSE (root mean square error) of the forecast verified against radial velocities from the Sapporo and the Kushiro radars (Fig. 4 (a)). The Kushiro radar is located at (42.96N, 144.52E), about 280 km east of the Sapporo radar. In this case, a front extends stationary over the Hokkaido, bringing rainfall in wide area including the Sapporo and the Kushiro radar sites (Fig.4 (b)). Thus, the impact from assimilating wind information around the Sapporo site is expected to be detected around the Kushiro site. The verification was performed every 10 minutes from the assimilation window through the 12 hour forecast. Data from elevation angles 0.1, 1.1, 2.6, and 4.3 deg. for the Sapporo radar and 0.3, 1.8, 3.4 and 5.2 deg. for the Kushiro radar are included in the verification.

The 4D-Var and the hybrid 4D-Var assimilation both reduce RMSE against the Sapporo radar (Fig. 5 (a)) from the first guess over most of the assimilation window. Main impact from assimilating radial velocity lasts up to about 6 hour forecast in both 4D-Var and hybrid 4D-Var. The hybrid 4D-Var tends to give smaller RMSE compared to the 4D-Var.

The Kushiro radar is located downstream of the flow from the Sapporo radar. Thus, the main impact on the RMSE ranges up to 8 hour forecasts. The 4D-Var and hybrid 4D-Var runs again give smaller RMSE than the first guess, with a larger reduction in the hybrid 4D-Var.

Time interval of the observations brings a larger difference in the hybrid 4D-Var than in the 4D-Var. The hybrid 4D-Var with a 10 minute observation interval gives smaller RMSE than the run with a 60 minute interval during the forecast range with the main impact. The flow-dependent background error may help appropriately reflect detailed scenarios in the assimilation, although more cases are needed to be investigated to get a conclusion. In 4D-Var, impact from the observation time interval is smaller. The RMSE from the run with a 10 minute interval is even slightly larger than the one from the run with a 60 minute interval in the Sapporo radar (Fig. 5 (a)).

On the other hand, the hybrid 4D-Var with a 10 minute observation interval gives the largest RMSE against the Kushiro radar over the assimilation window and 8-11 hour forecast (Fig. 5 (b)). Although more cases are needed to see whether this degradation is generally seen, it is possible that there is a contribution from the increments far from the Sapporo radar site, where influence from sampling errors can deteriorate the analysis. Investigation of a sophisticated configuration of the ensemble is one of the important issues.

5. Summary

Time and space correlation of observation error is investigated for radial velocity from weather Doppler radars.

Properties of the observation error correlation of radial velocity is diagnosed using statistics of observation-minus-background and observation-minus-analysis residuals. The correlation length along the beam is found to increase with distance from the radar site. Time correlation range also is found to increase with forecast time, implying possible contributions to the correlations from observational operator and forecast model.

A simplified 2D-Var is used to investigate the effects of observation error correlation on data assimilation. A data assimilation incorporating observation error correlation reflects detailed structure of innovations, keeping information from observations. Simply neglecting off-diagonal part of the correlation results in excessive increments over regions with grouped innovations with an identical sign. A simple Inflation of the error or a naive thinning helps mitigate the excessive increments, but can result in loss of detailed structures in the increment. Another thinning based on physical distance and keeping isolated observations recovers more signals. Careful investigation is needed to appropriately design thinning and inflation, considering resolution of the data assimilation and forecast system, density of observation, etc., to prevent loss of information from observations.

The observation error correlation is incorporated into the 4D-Var based on the JMA operational system and its extension to hybrid 4D-Var. In a case study, assimilating radial velocity data from a single radar site, impact from assimilation lasts up to 6-8 hour forecast in terms of RMSE against radial velocity. The hybrid 4D-Var outperforms the 4D-Var. Comparing the increment patterns suggest that the hybrid 4D-Var may have more potential to extract detailed scenario from the observational information.

Effect from utilizing observations with a small time-interval also is investigated using the same case. When interval is reduced from 60 to 10 minutes, the hybrid 4D-Var shows a larger sensitivity than the 4D-Var. In this case, the hybrid 4D-Var with a 10 minute observation interval tend to reduce RMSE against radial velocity.

Simplifications are applied in the present investigation.
The observation correlation in the present study is estimated from statistics of only 1 week using data from only a single radar site. Statistics is required to include various cases under different weather conditions from different radar sites to get general properties of the correlation.

The diagnosis of the correlation is based on the assumption that analysis update is consistent with the true background and observation error covariances. For possible enhancement of consistency, future works include iteration of analysis cycles with updates of the observation error covariance, using the estimation from the latest cycle.

The samples in the impact study using 4D-Var and hybrid 4D-Var is limited. Also, optimization is required for a configuration of the ensemble in hybrid 4D-Var (ensemble size, localization scale and weight of the ensemble and climatological background errors, etc.) to provide flow-dependency to extract as much observational information as possible.

Further investigations are required to evaluate robust characteristics of the observation error covariance and its appropriate handling in data assimilation. However, these results suggest a possibility that observation error correlation can become important in extensive use of radial wind observations.

Acknowledgements

This work was supported by JST AIP Grant Number JPMJCR19U2, Japan. This work is based on the operational NWP system developed by Numerical Prediction Division, Japan Meteorological Agency.

References


