18A.2 **Targeted Observation for Radar Data Assimilation**

Weiguang Chang* and IsztarZawadzki
McGill University, Montreal, Quebec, Canada

1. Introduction

Radar provides large amount of radial velocity and reflectivity in high spatial and temporal resolutions as the most important observation data at convective scale. However, traditional radar scans the atmosphere uniformly in space regardless of the spatial distribution of precipitation. Modern radar facilities, such as phased array radar (PAR), have much greater operational flexibility to scan the atmosphere adaptively according to different weather conditions. Therefore it is beneficial to develop methods of adaptive radar operation that takes the advantage of this new operational flexibility.

We will consider the problem from the perspective of data assimilation: targeted radar observation is the scanning strategy from which input information to an assimilation system leads analysis and forecast with the lowest possible uncertainty. By combining model with radar observations, information of radar data spreads to other unobserved atmosphere state variables, which helps us understand better the atmosphere phenomena. Moreover, this quantitative use of radar data produces analysis and forecast with less uncertainties due to the constrain of the dynamics and physics in the model. The goal of targeted radar observation is to reduce further the uncertainties by adaptively scanning the atmosphere.

The following experiments focus on the impact of different observation densities and different observation numbers on analysis uncertainties.

2. Estimation of analysis error

2.1 *Kalman filter equations*

Ensemble Kalman filter (EnKF) and variational approach are two approaches to data assimilation. One advantage of EnKF is that it can, at the same time as producing the analysis and forecast, estimate their error covariance matrices. Therefore, Kalman filter equations will be used in the following experiments to calculate the analysis error covariance matrix (ECM), as shown in the following equation.

* *Corresponding author address:* Weiguang Chang, Room 945, Burnside Hall, 805 Sherbrooke Street West, Montreal, Quebec, H3A 2K6.
Email: weiguang.chang@mail.mcgill.ca

$$\mathbf{P}^a = \mathbf{P}^f - \mathbf{P}^f\mathbf{H}^T(\mathbf{H}\mathbf{P}^f\mathbf{H}^T + \mathbf{R})^{-1}\mathbf{H}\mathbf{P}^f \tag{1}$$

where $\mathbf{P}^a$ is the analysis error covariance matrix; $\mathbf{P}^f$ is the forecast error covariance matrix; $\mathbf{H}$ is the observation operator; $\mathbf{R}$ is the observation error covariance matrix which describes the observation error structure.

2.2 *The ill estimated observation error structure.*

Equation 1 gives precise estimation of analysis error covariance matrix only when correct $\mathbf{P}^f$ and $\mathbf{R}$ are given, while $\mathbf{P}^f$ is always considered non-diagonal and can be calculated from ensemble members in EnKF, the observation ECM is poorly known (a diagonal $\mathbf{R}$ is often used ignoring correlations between observation errors). In the case of an estimated observation ECM, estimation of analysis error covariance matrix can be calculated by Eq. 2:

$$\begin{aligned}\mathbf{P}^a = \mathbf{P}^f &+ \mathbf{P}^f\mathbf{H}^T(\mathbf{H}\mathbf{P}^f\mathbf{H}^T + \mathbf{R}^e)^{-1} \\ &(\mathbf{H}\mathbf{P}^f\mathbf{H}^T + \mathbf{R})(\mathbf{H}\mathbf{P}^f\mathbf{H}^T + \mathbf{R}^e)^{-1}\mathbf{H}\mathbf{P}^f \\ &- 2\mathbf{P}^f\mathbf{H}^T(\mathbf{H}\mathbf{P}^f\mathbf{H}^T + \mathbf{R}^e)^{-1}\mathbf{H}\mathbf{P}^f\end{aligned} \tag{2}$$

where $\mathbf{R}^e$ and $\mathbf{R}$ are the estimated and true observation ECM respectively. Other symbols are the same as in Eq. 1.

3. Experiments on observation distribution

3.1 *Optimal observation distribution with fixed observation number*

The purpose of this experiment is to decide the observation location under the condition that only a fixed number of observations are allowed. This condition is reasonable because there is always a compromise between time and space sampling, that is, only a fixed number of observations can be obtained during the available period. Here a simple experiment is implemented to simulate this circumstance. The setting of this experiment is that there are 200 one-dimension state variables with same forecast error variance 1.0. Only two observations with error variance 1.0can be obtained. The error correlation functionis supposed to be Gaussian for both forecast and observation, as shown in Eq. 3.

$$\rho = \exp(-\frac{x^2}{r^2}) \quad (3)$$

where x is distance between model grids or observations, and r is the correlation distance.

The uncertainties of the analysis results are investigated by Eq. 1 to identify the optimal observation strategy. By summing the diagonal elements of $\mathbf{P}^a$, we obtain the total variance of analysis error. Because of the correlations between state variables and between observations, the total variance of analysis erroris a function of the distance between two observations, as shown in Fig. 1. In this case, two observations should be at least 10 points away to obtain the smallest total error variance.

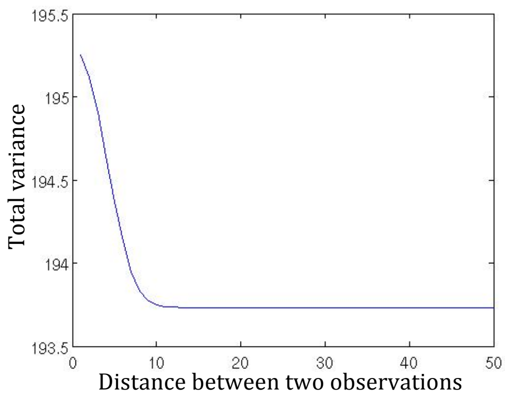

Fig.1 *Total variance of analysis error as a function of distance between two observations. Correlation distance is 5 and the variance is 1 for both observation error and forecast error. It shows that the smallest variance can be obtained when the distance between observations is larger than 10.*

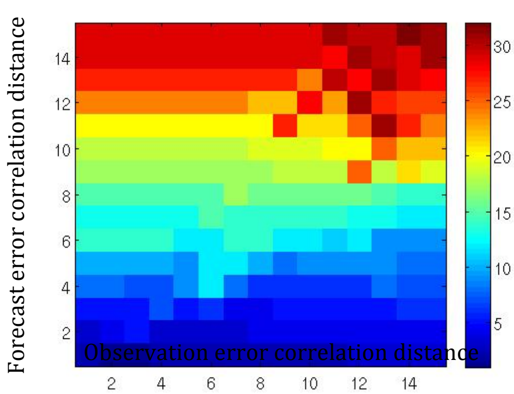

Fig.2 *Optimal observation distance as a function of correlation distances of forecast error and observation error. The variance is set to be 1 for both observation error and forecast error.*

By changing the correlation distance of forecast error and observation error, the optimal distance between observations is changed as well. Figure 2 shows the optimal distance as a function of the correlation distances of both forecast and observation. When observation error correlation distance changes, and forecast error correlation distance remains the same, the optimal distance varies little. But if forecast error correlation distance changes, the optimal distance between observations is significantly affected. Figure 2 indicates that the forecast error correlation distance is much more important than observation error correlation distance for deciding the optimal observation distance.

The optimal distance between observations is also related to the variances of observation and forecast errors. Fig. 3 shows how the optimal observation distance varies for different variances of observation and forecast errors. We can tell that both observation and forecast error variances are important for determining the optimal observation distance.

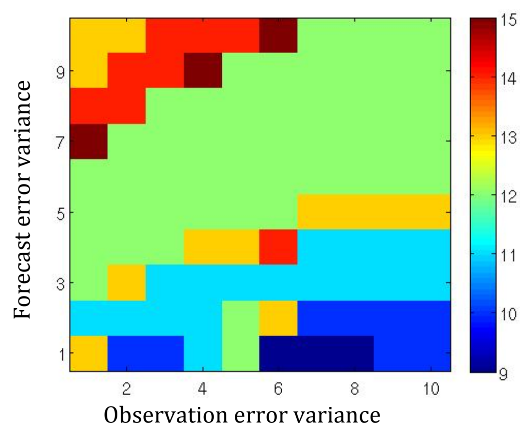

Fig. 3 *Optimal observation distance as a function of variances of forecast error and observation error. The error correlation distance is set to be 5 for both observation error and forecast error.*

The reason why observation error correlation distance plays a more important role than forecast error correlation distance is that while the observed variables are updated by both forecast and observation, the unobserved variables are updated according to the forecast error correlation only. Additionally, variances of both observation and forecast errors are important because both variances affect the variances of error of the updated observed variables, and then the variance information spreads to the unobserved variables.

For example, if only two radar beams are allowed, we want those two beams to be separate by a particular distance, which is determined by the error structures of the system, in order to have small analysis uncertainty.

And the optimal observation strategy depends largely on the forecast error correlation distance, but not on the observation error correlation distance.

3.2 *Optimal number of observations*

More observations provide more information. Therefore, we may want as many observations as possible. For example, we could direct the radar to scan everywhere to get all the information that can be obtained. However, this scan strategy is optimal for radar assimilation system only when the error structure of forecast and observation are well estimated beforehand. Increasing information with unknown error structure may be detrimental to the forecast. When an ill estimated error covariance matrix is applied, the question that how many observations are needed is not easy to answer.

The next experiment set the observation number to be 80 or less, and other parameters are the same as experiment 3.1.

First, if the assumption of well-estimated error covariance matrix is fulfilled, Fig. 4 shows that when more and more observations are assimilated, the total variance of analysis error is reduced. The optimal number of observations is 80, which means we want as many observations as possible.

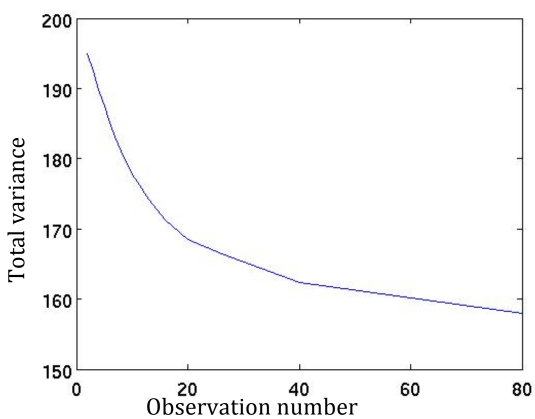

Fig.4 *Total variance of analysis error as a function of observation number. Observation error covariance matrix is well estimated. Correlation distance is 5 and the variance is 1 for both observation error and forecast error. Smallest total variance corresponds to 80 observations.*

Second, if a diagonal observation error covariance matrix is used, but the true observation errors have significant correlation, $\mathbf{R}$ is ill estimated. From Eq. 2, we can obtain the analysis error variances. Fig. 5 shows that in this case, 80 observations give larger analysis error variance than 40 observations. Therefore, reducing the number of observations is better due to the incomplete information about the observation error structure.

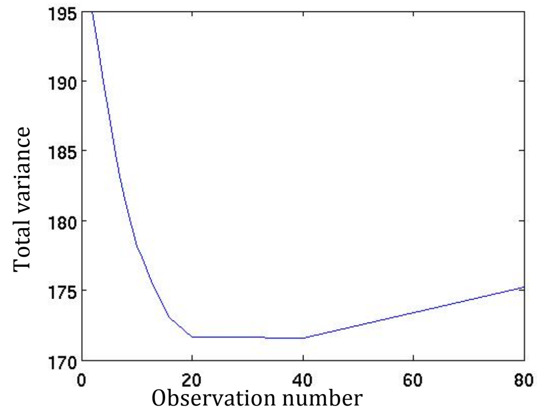

Fig.5 *Total variance of analysis error as a function of observation number. Observation error covariance matrix is ill estimated. Correlation distance is 5 and the variance is 1 for both observation error and forecast error. Smallest total variance corresponds to 40 observations.*

This experiment shows that when the well estimated observation ECM is used, no data thinning is needed because more observations with precise error structures are always better for data assimilation. However, when the observation error correlation is difficult to estimate, there is advantage to do some thinning to the observations. Dropping some observations whose error structure is unknown is better than keeping all these observations. The number of removed observations depends on the error structures of forecast and observation. Therefore, if a diagonal observation ECM is applied in data assimilation system, removing some radar data from the data assimilation system will be beneficial for the analysis.

3.3 *Non-uniform background error variance*

In reality, the forecast error structure cannot be the same for all the model grids. Therefore it is necessary to locate the observations according to different forecast error structures at different regions. In this experiment, 100 state variables are used. The forecast error variance is equal to 1.2between x=1 and 50, and is equal to 1between x=51 and 100, as shown in Fig. 6. And only 7 observations are allowed. To reduce the uncertainty of the analysis, more observations are needed where the forecast error variance is larger. We need to determine how many observations should be placed between x=1 and 50, and how many observations between x=51 and 100. The optimal placement gives the smallest total variance of analysis error. Fig. 7 indicates that the optimal placement in this experiment should be 5 observations between x=1 and 50, and 2 observations between x=51 and 100.

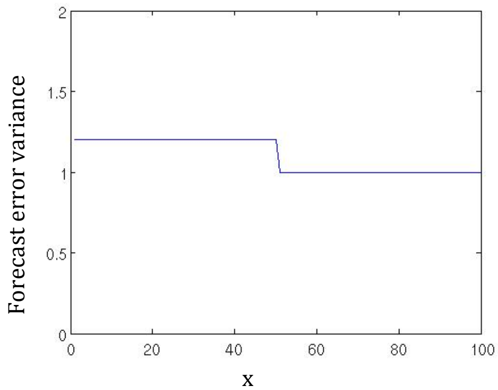

Fig.6 Forecast error variance at different locations.

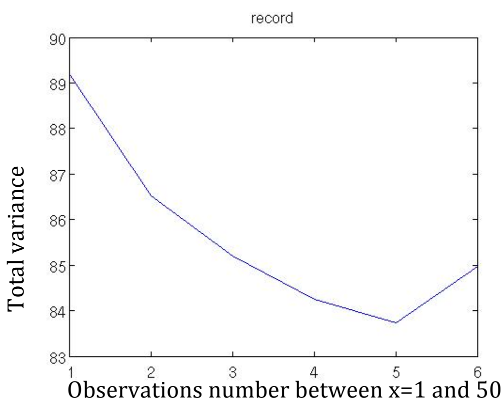

Fig.7 *Total variance of analysis error as a function of the location of observations. Smallest total variance corresponds to 5 observations between x=1 and 50.*

Moreover, there exists a mathematical approach to identify the important regions where the observations will reduce the total variance of analysis error most significantly. For example, ensemble transform Kalman filter (ETKF) introduced by Bishop (2001) allows to identify the most important observation region. We will develop a comprehensive methodology for optimal scanning of the atmosphere to take advantage of future possibilities of scanning agility.

4. Summary

The above experiments try to find out the optimal quantity of observations and the optimal distance between adjacent observations. The experiments prove that if the observation error correlation is known exactly, more observations produce more precise data assimilation result for sure. But if the observation error correlation is not known exactly, and is assumed to be zero, as in most realistic cases, there is an optimal quantity of observation that generates the analysis with minimum total error variance. If the amount of observations is fixed, there is an optimal distance between adjacent observations, which depends more on the correlation distance of forecast errors than that of observation errors. When the forecast error variance varies at different locations, there exists an optimal observation distribution, which depends on the error structures of the forecast and observation.

The idea of these simple experiments will be applied on the radar data assimilation system in the future. It is expected that the locations and the amount of observations will significantly affect the analysis uncertainty. However, the smallest total variance of analysis error, or an analysis with smallest uncertainty, cannot guaranty the forecast to be precise. In the future research, wewill consider the effect of targeted observation on the forecast.

References:

Bishop, C. H., B. J. Etherton, et al., 2001: "Adaptive Sampling with the Ensemble Transform Kalman Filter. Part I: Theoretical Aspects." Monthly Weather Review 129(3): 420-436.

Buizza, R., C. Cardinali, et al., 2007: "The value of observations. II: The value of observations located in singular-vector-based target areas." Quarterly Journal of the Royal Meteorological Society 133(628): 1817-1832.

Buizza, R. and A. Montani, 1999: "Targeting Observations Using Singular Vectors." Journal of the Atmospheric Sciences 56(17): 2965-2985.

Cardinali, C., R. Buizza, et al., 2007: "The value of observations. III: Influence of weather regimes on targeting." Quarterly Journal of the Royal Meteorological Society 133(628): 1833-1842.

Kelly, G., J. N. Thépaut, et al., 2007: "The value of observations. I: Data denial experiments for the Atlantic and the Pacific." Quarterly Journal of the Royal Meteorological Society 133(628): 1803-1815.

Langland, R. H., 2005: "Issues in targeted observing." Quarterly Journal of the Royal Meteorological Society 131(613): 3409-3425.

Majumdar, S. J., C. H. Bishop, et al., 2002: "Adaptive Sampling with the Ensemble Transform Kalman Filter. Part II: Field Program Implementation." Monthly Weather Review 130(5): 1356-1369.

Palmer, T. N., R. Gelaro, et al., 1998: "Singular Vectors, Metrics, and Adaptive Observations." Journal of the Atmospheric Sciences 55(4): 633-653.

Rabier, F., E. Klinker, et al., 1996: "Sensitivity of forecast errors to initial conditions." Quarterly Journal of the Royal Meteorological Society 122(529): 121-150.