Diagnosis of Data Assimilation Systems: Observation Impact Estimation, Error Covariance Matrix Optimization, and Analysis Error Estimation

Toshiyuki Ishibashi
Meteorological Research Institute, Japan

In this study, we develop diagnostic schemes of data assimilation systems (DASs), observation impact estimation, error covariance matrix optimization, and analysis error estimation. These diagnostics are essential to improve data assimilation systems and to design future observation networks.

1. Observation impact estimation

In data assimilation systems (DASs), the effect of each assimilated observation dataset on an analysis field is one of the main factors in determining analysis and succeeding forecast accuracy. We call the effect the linear observation impact, which is fully determined by the Kalman gain. However, the Kalman gain is never constructed explicitly in variational DASs. Therefore, the estimation of the linear observation impact is a difficult problem. In this section, we analyze linear observation impacts using two methods, the adjoint-based method and the tangent linear (TL) approximation based method.

Firstly, the adjoint-based method is implemented to the JMA global 4D-Var DAS. One-month long experiments of the observation impact estimation (using a dry total energy norm, and 15 or 27 hour forecasts) show that almost all observation data contribute forecast error reduction in average, and this result is consistent with those of past OSEs in JMA. However, the experiments show that the impact of a total satellite radiance data is about the same magnitude as that of radio sonde data. This result implies that there is room for improvement of the forecast accuracy by improving usage of the radiance data, since, in previous studies in other numerical weather prediction (NWP) centers shows larger impacts from radiances. We also find impacts of water vapor channel radiances increase in case of using a norm including humidity energy, and the method can detect wrong observation data which are given artificially deflated (too small) observation error settings.

Secondly, we construct the tangent linear (TL) approximation based method. The method estimates the linear observation impacts as a partial analysis increment vector (PIV) that is generated by each observation dataset. The method enables us to see how the Kalman gain transforms information of observations into analysis increments. We can also estimate the observation impacts on a forecast field by the method, while the tangent linear approximation of a forecast model is valid. We perform a preliminary experiment of one analysis cycle on the global variational DAS of the JMA to demonstrate the validity and performance of the method. The experimental results are as follows:

a. The method can identify the observation impacts of each observation dataset on an analysis and a forecast field in the operational variational DAS.
b. There are negative correlations between an integrated background error vector and dominant PIVs.
c. The nonlinear interactions between observation impacts in the known methods can be estimated quantitatively.

The method is useful for observational data impact estimation in variational DASs.

Here, we consider a general relationship between the adjoint-based and the TL-based estimation. As mentioned previously, linear observation impacts are perfectly determined by the Kalman gain. Therefore, all estimation methods of them are proxies of the direct estimation using the Kalman gain matrix. Since, the Kalman gain has two indices, a row index for an analysis field space and a column index for an observation space, there are two ways to approximate the direct method. First, we can use all of our computation ability to estimate the action of the Kalman gain in the observation space. This approximation estimates the action in the observation (analysis) space with high (low) resolution. Second, we can also use all of our computation ability to estimate the action of the Kalman gain in the analysis space. This approximation estimates the action in the analysis (observation) space with high (low) resolution. The former corresponds to the adjoint-based estimation and the latter corresponds to the TL-based estimation. Therefore, these two methods are in complementary relationship.

2. Error covariance matrix optimization

Error covariance matrices are elements of the Kalman gain. Here, we describe the optimization of error covariance matrices using sensitivities of forecast errors to the matrices. These sensitivities can be calculated by extending the adjoint-based observation impact estimation method. This adjoint-based covariance sensitivity estimation method is implemented to the JMA global 4D-Var, and month long sensitivity calculations show that observation error variances of radiance data are too large for almost all these data types. These results are consistent with past results based on degrees of freedom for signal (DFS). We find that root mean square errors of forecasts are reduced by optimizing observation error variances of these data according to the results of the sensitivity estimation.

.

3. Analysis error estimation

There have been many studies on an analysis error estimation using forecast error information, in other words, the construction of more accurate analyses than current DASs. We construct new analysis error estimation methods. If analysis errors can be known, we may be able to design future observation systems to detect them, and construct a proxy of the truth instead of, or complementary to, a nature run in OSSEs (observing simulation system experiments). Furthermore, although, these analyses cannot be used for initial fields of NWPs because of the time constraint of NWPs, while, “reanalysis” is free from this time constraint.

Corresponding Author:
Name: Toshiyuki Ishibashi
Organization: Meteorological Research Institute
Address: 1-1 Nagamine, Tsukuba, Ibaraki, Japan
Email Address: ishibasi@mri-jma.go.jp