Tests of observation error variance tuning with multiple regularization parameters method for typhoon initialization with microwave radiances

Jian Zhong¹, Min Jiang, Yuqin Wu and Jiong Zhou

China Satellite Maritime Tracking and Control Department, Jiangyin, Jiangsu, China

¹Email: zhj_zhongjian@163.com

Abstract. Based on the Tikhonov regularization inversion problem theory, a Tikh–4D–Var data assimilation method, which using multiple regularization parameters as a weak constraint for four dimensional variational data assimilation (4D-Var), has been developed by the authors recently. We have shown that the new method is effective in improving the typhoon prediction compared with traditional 4D–Var for the assimilation of bogus observations. In this study, the new method is first applied to the typhoon initialization with microwave radiance observations assimilation, and the results are compared with 4D–Var for typhoon Chaba (2010). Results show that compared with 4D–Var experiment, Tikh–4D–Var experiment can effectively reduce the iterations to accelerate the convergence, and obviously improve the initial structure, the simulated structure, and the intensity and track prediction when radiance observations assimilated. Meanwhile, due to the optimal use of the radiance observations, the warm-core structure and secondary circulation of Tikh–4D–Var experiments improve more significant than the experiments without radiance observations. Moreover, the regularization parameters for different observations are un-correlated, and increasing or decreasing observation type don’t influence the calculation of regularization parameters.

1. Introduction

Variational data assimilation (VDA) technique which includes three-dimensional variational data assimilation (3D-Var) and four-dimensional variational data assimilation (4D-Var), is one of the most efficient methods to improve the initial conditions (IC) for numerical weather prediction (NWP) model [1-3]. Compared with 3D-Var, 4D-Var is one of the most advanced in algorithm formulation and technique design, which can produce dynamically and thermodynamically consistent initial conditions by using the model equations as constraints and the observations at different time in the assimilation window. As the development of computer, 4D-Var has been implemented for operational numerical weather prediction at many operational centers, e.g. ECMWF, Metero France, UK Met Office, and JMA [1]. The background and observation error covariances should be set appropriately when 4D-Var is implemented. When the observation errors are assumed to uncorrelated and the observation error covariances are assumed to be diagonal in operational data assimilation, the observation errors determine the relative weight of each part in the cost function, and the contribution of each observation to the analysis. However, the observation errors are always set to constant values with experience or empirical statistics, especially for radiance observations, which makes the corresponding observation not sufficiently used in data assimilation system.
One of the approaches for optimal setting the relative weights of observations is the regularization method in mathematics. Johnson et al. [4] have suggested that the high-dimensional variation data assimilation problem is an ill-posed inverse problem, and the variance ratio of errors of observation and background can be improved and estimated with Tikhonov regularization theory, and the theory is tested with a simple low-dimensional model for estimating the ratio. From the point of singular value decomposition, Johnson et al. [5] have suggested that the 4D-Var is equivalent to Tikhonov regularization theory using certain transformation. Budd et al. [6] have shown that the background term in the variational data assimilation corresponds to the regularization term in the regularization theory. Because the regularization parameter selection is difficult for real high-dimensional system, the regularization method is rarely implemented in real high-dimensional system.

Zhao et al. [7] have first applied the regularization method into real high-dimensional system of the Weather and Forecasting (WRF) 3D-Var data assimilation system. The results have implied that the regularization method improves the rain prediction by extracting more information from the radar observations. In this paper, the L-cure rule regularization parameter selection method is used, and it suggested that the selection method is not suitable for 4D-Var due to the large amount of computation. Recently, a Tikh-4D-Var data assimilation method, which using multiple regularization parameters in mathematic as a weak constraint for 4D-Var has been recently proposed by Zhong et al. [8]. Compared with the traditional 4D-Var, regularization parameters are added for each observation type to optimally set the relative weight of each observation. Also, a multiple regularization parameters selection method which is used for error tuning [9], is introduced for implementing Tikh-4D-Var with WRF 4D-Var system. The computation needed for the Tikh-4D-Var method is about (less than) three times of traditional 4D-Var, which is remarkably smaller than the traditional single regularization parameter selection method. The new approach has been implemented into the WRF 4D-Var system for optimal setting variances of bogus and conventional observations.

In order to have an accurate intensity, track and vortex structure of typhoon in model forecasts, the initial conditions should be improved using observations in the typhoon’s inner core region. Unfortunately, typhoons are always over oceans and there are usually no radar data and no sufficient convection data in the typhoon region. To define a typhoon structure in the model initial condition, many studies have shown that assimilating radiance observations e.g. microwave radiances with VDA algorithms can improve typhoon initialization and prediction. When the satellite radiance observations are used in 4D-Var, one of the important aspects is appropriately appoint the radiance observation errors, which influence the relative weight of radiance observations [10]. However, Current radiance observation errors in the VDA, which are always set to constant with empirical statistics, may lead to suboptimal assimilation of radiance observations. The inaccurate radiance observation errors setting will influence typhoon initialization and prediction. How to optimally set the observation error especially for radiance observations is very important.

In this paper, based on the WRF 4D-Var system, the Tikh-4D-Var method is first tested for assimilating of microwave radiance observations through typhoon initialization of Chaba (2010). The rest of this paper is organized as follows. The detail of the Tikh-4D-Var approach is introduced in the next section, including the multiple regularization parameters selection method. Details of experimental design are given in Section 3. The simulation results from both 4D-Var and Tikh-4D-Var experiments, including the convergence property, the initial structure, and the typhoon track, intensity and structure prediction are compared in Section 4. The major results are summarized in the last section.

2. Theory of Tikh-4D-Var data assimilation method

The Tikh-4D-Var data assimilation approach which is proposed by Zhong et al. [8], is based on the biased minimum norm solution. The cost function of Tikh-4D-Var is

\[
J(x^a, \alpha, \Gamma_1, ..., \Gamma_{nu}) = \frac{1}{2} \alpha \| x - x^f \|^2 + \frac{1}{2} \sum_{j=1}^{nu} \Gamma_j \| x_j^f - H_j(x_j) \|^2 = \min!
\]
Where $x^a$ and $x^b$ are the analysis and background at initial time separately, $H_i$ is the observation operator for observation type $i$, $y^{o,j}_i$ is the observation in time $j$ and for observation type $i$, $\alpha$ is always called as regularization parameter, $\Gamma_i$ is the regularization parameter for each observation type, $n_{obs}$ is the number of observation type, $N$ stands for the assimilation window time, $\|x\|$ stands for the bound norm, which is

$$
\|y^{o,j}_i - H_i(x_j)\|^2 = \left\langle y^{o,j}_i - H_i(x_j), R_i^{-1}(y^{o,j}_i - H_i(x_j)) \right\rangle
$$

$$
\|x - x^b\|^2 = \left\langle x - x^b, B^1(x - x^b) \right\rangle
$$

(2)

Where $R_i$ and $B$ stand for the observation type $i$ and background error covariance separately. $(\bullet, \bullet)$ and $\langle \bullet, \bullet \rangle$ are the inner product defined in the model state and observation Euclidean space separately.

The multiple regularization parameters are the key factors for successfully implementing the Tikhonov 4D-Var. How to select the multiple regularization parameters is very important. In this paper, the multiple regularization parameters are selected with the method as proposed in Zhong et al. [8], which is based on the posterior information of the assimilation system [11, 12]. The multiple regularization parameters are calculated as [8]

$$\alpha = \frac{2E\left\{ J^b(x^a(s)) \right\}}{Tr(K(s)H)}$$

$$\Gamma_k = \frac{2E\left\{ J^o_k(x^a(s)) \right\}}{Tr\left( \Pi_k^{o}\left( I - HH^T(s) \right) \Pi_k^{oT} \right)}, \quad k=1, \ldots, n_{obs},$$

(3)

Where $s = (\alpha, \Gamma)$ is the Tikhonov regularization parameters vector, $E\left\{ J^b(x^a(s)) \right\}$ and $E\left\{ J^o_k(x^a(s)) \right\}$ stand for the average value of the cost function of background and the $k$th observation separately, $Tr$ stands for the trace of matrix; $\Pi_k^{o}$ is the projection matrix which extract the subpart $z^o_k$ associated with the $k$th observation from the data vector $z^o$; $K(s)$ is the Kalman gain matrix constructed from the first guess regularization parameters vector, i.e. for any $s$:

$$K(s) = B(s)H^T(s)H^T + R(s)^{-1}$$

(4)

With $B(s) = \Sigma B$ and $R(s) = \Sigma R_i$, $B$ and $R_i$ are the background error covariance and $k$th observation error covariance for traditional 4D-Var.

In this paper, the trace of large matrices is computed with the Girard [13] method. Considering $\xi \in N(0, I)$ (Gaussian with mean 0 and covariance matrix the identity I), the trace for matrix $A$ can be written as

$$Tr(A) = E(\xi^T A \xi)$$

(5)
\( E(\xi^T A_\xi) \) stands for the average value of \( \xi^T A_\xi \). By setting \( A \) as \( HK \), the \( Tr(HK) \) can be expressed as

\[
Tr(HK) = (R^{1/2} \xi)^T \left( H \left( \delta x^e_{\nu+\sigma^p} \right) - H \left( \delta x^e_\nu \right) \right)
\]

(6)

With \( \delta x^e_{\nu+\sigma^p} \) and \( \delta y^0 = R^{1/2} \xi \). The regularization parameters are calculated with only one iteration, and the accuracy for the trace calculation is impacted by the observation number. The main steps for implementing the Tikh-4D-Var are shown in Zhong et al. [8].

3. Experimental design

In this paper, initialization and simulation of typhoon Chaba (2010) is implemented with Tikh-4D-Var for assimilating microwave radiance observations, bogus observations and convectional observations. The initial time for the case is 0600 UTC 25 October 2010, and the assimilation window is 6h. The parameter schemes used in this paper are the same as in Zhong et al [8], which are all provided in the version 3.4.1 [14] of WRF and WRF Data assimilation (WRFDA). They are the large scale condensation, surface drag, and a cumulus scheme for assimilation experiments, and the Kain-Fritsch cumulus parameterization scheme, YSU planetary boundary layer parameterization, and the WSM-6 microphysics scheme for forecast experiments. The model has 28 vertical levels and the model domain is doubly nested with horizontal resolutions of 45-km and 15-km for the coarse mesh and fine mesh separately, and the horizontal grid numbers are \( 70 \times 90 \) and \( 160 \times 214 \) separately. The NCEP global reanalysis data are used to define the initial and lateral boundary conditions. The typhoon initialization with different data assimilation methods is carried out on the coarse mesh domain, and is interpolated to the fine mesh domain.

Two data assimilation experiments which are traditional 4D-Var and Tikh-4D-Var, are carried out on the coarse mesh domain. The initial conditions obtained from the data assimilation experiments, combined with the initial conditions from NCEP reanalysis data without any VDA, are conducted for the 72-h simulation of typhoon Chaba (2010) using a double nested-mesh WRF. The initial conditions from NCEP reanalysis data without any VDA are denoted as CTRL experiment, and serves as a benchmark to demonstrate how the 4D-Var and Tikh-4D-Var experiments improve the typhoon forecast. The observations used in the assimilation include amsua and amsub radiance observations, and the bogus and conventional observations. Details of the configuration of bogus data can be found in Huang et al. [15].

4. Property analysis

Figure 1 shows multiple regularization parameters for bogus and conventional observations. The regularization parameters calculated from different observations data assimilation experiments are shown to analyze the property of the regularization parameters. Black bar stands for the Tikh-4D-Var experiment with only conventional and bogus observations assimilated [16], and gray bar stands for the Tikh-4D-Var experiment with radiance observations assimilated. As shown in Zhong et al [8], the regularization parameters for different observations are different. Moreover, the differences of regularization parameters for different observations assimilation experiments are very small for all observation types except for the airep observation. The results indicate that the regularization parameters for different observations are un-correlated, increasing or removing observations from the data assimilation system doesn’t influence the regularization parameter calculation for the other observations. The reason for large difference of airep observation is that regularization parameters calculation is mainly influenced by the observation amount, and the number of airep observation is very small (only 21) for data assimilation experiments.
The regularization parameters of amsua and amsub observations are shown in figure 2. The regularization parameters for different channels and platforms are different, which is an indication of the necessity of introducing multiple regularization parameters for each radiance observation type. Most of the regularization parameters of radiance observations are larger than 1, which indicates that the observation errors of amsua and amsub in assimilation system is small for traditional 4D-Var experiment, and can be correctly tuned closer to the true error by introducing regularization parameters.

**Figure 1.** Multiple regularization parameters for the bogus and conventional observations and background. Black (Bogus) stands for only bogus and conventional observations are assimilated as in Zhong [16]. Grey (Bogus+Rad) stands for radiance observations are assimilated.

**Figure 2.** Multiple regularization parameters for the amsua and amsub radiance observations.
Meanwhile, Tikh-4D-Var experiment shows littler iterations than 4D-Var experiment, which is the same as the results in Zhong et al [8]. The iterations needed for Tikh-4D-Var and 4D-Var experiments are 35 and 47 separately. The results suggest that assimilating more observations of radiance observations, Tikh-4D-Var experiment still can reduce the iterations for accelerating the convergence.

5. Track and intensity
In this section, the influence of different data assimilation experiments and the NCEP experiment on the prediction of track and intensity of typhoon Chaba from the fine mesh with horizontal resolution of 15-km is discussed.

The mean errors of 72-h prediction for typhoon track, maximum wind speed and center surface level pressure are given in table 1. Compared with NCEP experiment, the 4D-Var and Tikh-4D-Var experiments reduce the average track prediction error slightly with mean track error of about 58.5 km and 52.1 km separately, and the Tikh-4D-Var improves more. However, the track forecast from Tikh-4D-Var experiment shows no-obvious improvement compared with 4D-Var experiment from the 72-h prediction error as shown in figure 3. The track forecast error of Tikh-4D-Var experiment is smaller than 4D-Var experiment for the whole 72-h prediction except for 12 – 24 hours and 42 hour.

Figure 4 depict the time evolution of the center surface level pressure error and the maximum wind speed error at 6-h intervals separately. With a cold start without any observed data being assimilation into the initial condition, NCEP experiment fails to get the initial typhoon structure, which results in a mean center surface pressure error of about 18.5 hPa and maximum wind speed error of about 11.8 m/s in 72-h forecast (table 1). The data assimilation experiments with addition of bogus and radiance observations information improve the prediction of intensity to various extents compared with the NCEP experiment. Moreover, Tikh-4D-Var experiment improves the intensity simulation more significantly than the 4D-Var experiment. The mean center surface level pressure and maximum wind speed error of 4D-Var experiment are 8.6 hPa and 5.0 m/s, which are greatly reduced to 3.6 hPa and 3.7 m/s separately for Tikh-4D-Var experiment for 72-h integration. The time evolution of the intensity at 6-h interval for 72-h prediction (figure 4) show that the intensity are improved for almost all the time in 72-h prediction for Tikh-4D-Var experiment compared with 4D-Var experiment, except for the maximum wind speed prediction of 12 hour and 24 – 30 hours.

As show in section 4, different regularization parameters are set for different observations. For examples, the regularization parameters for bogus observations are smaller than 1, and for almost radiance observations are larger than 1, which means that the observation errors setting for bogus observations in 4D-Var are too large, and for almost radiance observations are too small. By applying the multiple regularization parameters in Tikh-4D-Var, the relative weight of observations are tuned closer to their corresponding real weight, and the information of different observations especially for bogus and radiance observations are more effectively and fully explored, which results in more improvement in typhoon intensity prediction. These results indicate that compared with the 4D-Var experiment, the Tikh-4D-Var experiment can improve the typhoon prediction, especially for the intensity prediction, by setting the observations more proper relative weight to fully explore the information of observations.

Table 1. Mean error of typhoon mean position error of the simulated track and intensity error.

|                          | NCEP | 4D-Var | Tikh-4D-Var |
|--------------------------|------|--------|-------------|
| Mean track-error/km      | 61.8 | 58.5   | 52.1        |
| Mean surface pressure-error/hPa | 18.5 | 8.6    | 3.6         |
| Mean maximum wind speed error/(m/s) | 11.8 | 5.0    | 3.7         |
Figure 3. 72-h variations of typhoon track (km) at 6-h interval for all experiments.

Figure 4. As in figure 3, but for surface pressure error (left) and maximum wind speed error (right).

6. Structure

6.1. Initial structure

Table 2 shows the mean absolute analysis increments of the data assimilation experiments. Compared with 4D-Var experiment, the zonal and meridional wind and the temperature variables exhibit larger analysis increments for Tikh-4D-Var experiment. This is because Tikh-4D-Var experiment will make the analysis more closer to the observations through using more proper relative weights for different observations. The reason for larger relative humidity analysis increments existing in 4D-Var is that improper large analysis increment exists in the model top layer as shown in figure 5.

Table 2. Mean absolute analysis increments of the data assimilation experiments.

|       | U(m/s) | V(m/s) | T(K)  | Q(g/kg) |
|-------|--------|--------|-------|---------|
| 4D-Var| 2.390  | 2.356  | 0.501 | 2.546   |
| Tikh-4D-Var| 2.520  | 2.381  | 0.502 | 1.038   |

The cross sections for the relative humidity and temperature analysis increments around typhoon center are given in figure 5. Compared with 4D-Var experiment, larger analysis increments for both relative humidity and temperature in the typhoon center are exhibited in Tikh-4D-Var experiment. The largest temperature analysis increment is 6 K for Tikh-4D-Var experiment, larger than 4 K for 4D-Var experiment. Meanwhile, the relative humidity analysis increment is 5 g/kg for Tikh-4D-Var experiment, which is larger than 4 g/kg for 4D-Var experiment. The results indicate that when the
radiance observations assimilated, the Tikh-4D-Var experiment still can improve the initial relative humidity and warm-core structure.

Figure 5. Cross-section of ((a) and (c)) temperature and ((b) and (d)) relative humidity analysis increments. (a) and (b) 4D-Var, (c) and (d) Tikh-4D-Var. Contour interval for (a) and (c) is 1 K. Contour interval for (b) and (d) is 1 g/kg.

6.2. Simulated structure

The azimuthally and radially averaged temperature anomaly of the data assimilation experiments are given in figure 6. The values are computed with a radius of 50 km from the TC center. The 4D-Var and Tikh-4D-Var experiments both exhibit obvious warm-core structures during the 72-h simulation. The value of the temperature anomaly approximately coincides with the intensity, and the warm core intensifies until the storm reaches maturity. The largest temperature anomaly occurs in 4D-Var experiment up to 7 K smaller than the Tikh-4D-Var experiment which exhibits a larger largest temperature anomaly up to 9 K. The Tikh-4D-Var experiment shows larger temperature anomaly than 4D-Var experiment for almost all the 72-h prediction, especially for the middle vertical layer.

The temporally and azimuthally averaged radial wind and vertical motion for the data assimilation experiments are displayed in figure 7. The values are calculated within the period of 60 - 66 h prediction. The 4D-Var and Tikh-4D-Var experiments both exhibit a typical secondary circulation: the bottom layer (about 0 - 2 km in altitude) with a strong radial inflow, the upper layer (about 13 - 17 km in altitude) with an intense radial outflow, and the middle layer (about 3 km in altitude) with a weak outflow. In addition, compared with 4D-Var, Tikh-4D-Var experiment shows a larger radial wind of 2 m s⁻¹ for the upper layer and 4 m s⁻¹ for the bottom layer. Meanwhile, the vertical velocity is larger for Tikh-4D-Var experiment in the radius of 75 - 100 km from the typhoon center. The results indicate that compared with 4D-Var experiment, Tikh-4D-Var experiment still exhibits a stronger secondary circulation for radiance observations assimilation.

Compared with the bogus and conventional observations data assimilation experiments in Zhong [16] which is not shown in this paper, the Tikh-4D-Var experiment improves the warm-core structure
and secondary circulation more significant when the radiance observations assimilated. This is because the internal structure, such as the temperature and relative humidity prediction, which are impacted by radiance observations, are improved more significantly by setting proper relative weights with regularization parameters to fully explore the information of radiance observations.

![Figure 6](image1.png)

**Figure 6.** Temporal variation of the azimuthally and radially averaged temperature anomaly (K) for (a) 4D-Var (b) Tikh-4D-Var. The ordinate is height (km), and the abscissa is simulation time (h).

![Figure 7](image2.png)

**Figure 7.** As in figure 6, but for the vertical motion (m/s, shaded) and radial winds (m/s, contoured).

7. **Summary and Conclusions**

One of the challenging tasks in the traditional 4D-Var experiment is how to effectively setting the relative weight for different observations. In this study, the Tikh-4D-Var data assimilation method, which using multiple regularization parameters as a weak constraint for 4D-Var is firstly applied to data assimilation with microwave radiance observations. Initialization and simulation of typhoon Chaba (2010) are conducted for testing the Tikh-4D-Var method with radiance observations assimilation.

Results suggest that the regularization parameters for different observations are un-correlated, increasing or removing observations from the data assimilation system doesn’t influence the regularization parameter calculation of the other observations. The precision of the regularization parameters is mainly determined by the number of observations. Meanwhile, Tikh-4D-Var experiment still can accelerate the convergence with less iterations for radiance observations assimilation. Moreover, compared with 4D-Var experiment, when the radiance observations assimilation, the
typhoon track and intensity prediction are still effectively improved for Tikh-4D-Var experiment, and the same for typhoon initial structure, the simulated warm-core structure and secondary circulation. Compared with the Tikh-4D-Var experiment only assimilating the bogus and conventional observations, Tikh-4D-Var experiment improves the warm-core structure and secondary circulation more significant when the radiance observations assimilated through optimal exploring the information of radiance observations.

In this paper, only one case study is performed for typhoon initialization to evaluate the new method for radiance observations assimilation. With the encouraging results herein and in Zhong et al. [8], we plan to apply this method to a large number of typhoon cases over the western North Pacific to allow a systematic evaluation of the performance of the new method.

Acknowledgments
This work was supported by the National Natural Science Foundation of China (Grants No. 41605075).

References
[1] Huang X Y, Xiao Q N, Barker D M, Zhang X, Michalakes J, Huang W, Henderson T, Bray J, Chen Y S, Ma Z Z, Dudhia J, Guo Y R, Zhang X Y, Won D J, Lin H C and Kuo Y H 2009 Mon Wea Rev 137 299
[2] Zhong J, Fei J F, Cheng X P and Huang X G 2014 Acta Phys Sin 63 149201
[3] Zhong J, Fei J F, Huang S X and Du H D 2012 Acta Phys Sin 61 149203
[4] Johnson C, Nichols N K and Hoskins B J 2000 Int J Numer Methods Fluids 00 1
[5] Johnson C, Hoskins B J and Nichols N K 2005 Q J R Meteorol Soc 131 1
[6] Budd C J, Freitag M A and Nichols N K 2011 Computer Fluids 46 168
[7] Zhao Y L, Huang S X, Du H D and Zhong J Q 2011 Acta Phys Sin 60 079202
[8] Zhong J, Fei J F, Cheng X P and Huang X G 2014 Science China: Earth Sciences 57 2690
[9] Desroziers G and Ivanov S 2001 Quart J R Meteorol Soc 127 1433
[10] Schwartz C S, Liu Z Q, Cheng Y S and Huang X Y 2012 Weather Forecasting 27 424
[11] Chapnik B, Desroziers G, Rabier F and Talagrand O 2004 Quart J R Meteorol Soc 130 2253
[12] Chapnik B, Desroziers G, Rabier F and Talagrand O 2006 Quart J R Meteorol Soc 132 543
[13] Girard D 1987 Technical Report. IMAG 1
[14] Wang W, Cindy B, Michael D, Jimy D, Dave G, Michael K, Kelly K, Lin H C, John M, Syed R and Zhang X 2012 ARW Version 3 Modeling system User’s Guide (New York: National Center for Atmospheric Research) 22
[15] Huang X G, Fei J F, Zhang G S and Lu H C 2004 Journal of Tropical Meteorology 20 129
[16] Zhong J 2013 Theoretical study and numerical experiments on weak constraint in four-dimensional variational data assimilation (Nanjing, University of Science and Technology)