An Adjoint-Free Alternating Direction Method for Four-Dimensional Variational Data Assimilation With Multiple Parameter Tikhonov Regularization

Xiangjun Tian1,2,3, Rui Han1,2,6, and Hongqin Zhang1,2,6

1International Center for Climate and Environment Sciences, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, China, 2University of Chinese Academy of Sciences, Beijing, China, 3Collaborative Innovation Center on Forecast and Evaluation of Meteorological Disasters, Nanjing University of Information Science and Technology, Nanjing, China

Abstract Tikhonov regularization is critical for accurately specifying both the background (B) and observational (R) error covariances in four-dimensional variational data assimilation (4DVar). The ratio of the background and observation error variances (referred to as the B-R ratio) is the key to ensuring that 4DVar maximizes the information extracted from the observations. However, it is difficult to specify the regularization parameters in a high-dimensional variational data assimilation (VDA) system for both the single- and multiple-parameter regularization schemes. In this study, we used a maximum likelihood estimation (MLE)-based inflation scheme that originated from the ensemble Kalman filter (EnKF) community and proposed an alternating direction method (ADM) to minimize the 4DVar cost function with the iterative application of multiple regularization parameters to simultaneously optimize the regularization parameters and model states under the framework of the nonlinear least-squares 4-D ensemble variational data assimilation method (NLS-4DVar). The big-data-driven version of NLS-4DVar (BD-NLS4DVar) with multiple-parameter Tikhonov regularization was able to adjust the B-R ratios more accurately. Several groups of observing system simulation experiments (OSSEs) based on 2-D shallow-water equations demonstrated that BD-NLS4DVar with multiple-parameter Tikhonov regularization produced a substantial performance improvement over the standard BD-NLS4DVar method with no regularization.

1. Introduction

Data assimilation can effectively combine observations and the background in an optimal way to yield the optimal estimation of model initial fields and thus improve the accuracy of numerical weather prediction (NWP) (e.g., Rabier et al., 2000). The effective fusion of the background states and observations is heavily dependent on the a priori specification of both the background (B) and observational (R) error covariances, as their accuracy will eventually influence the final assimilation performance (e.g., Li et al., 2009; Zhong et al., 2014). As the two major methods for data assimilation, both four-dimensional variational data assimilation (4DVar; Navon et al., 2005; Rabier et al., 2000) and the ensemble Kalman filter (EnKF; Evensen, 2009; Houtekamer et al., 2014) must estimate both the background and observational error covariances accurately to improve their assimilation accuracy.

In 4DVar, the ratio of the background and observation error variances (referred to as the B-R ratio) is believed to be critical for maximizing the information that can be extracted from observations (Johnson et al., 2000). However, an accurate estimate of the variance ratios is not generally available; 4DVar with multiple-parameter Tikhonov regularization is utilized to simultaneously optimize the model states and variance ratios by incorporating the Tikhonov regularization theory into 4DVar. Furthermore, the B-R ratio can be interpreted as a regularization parameter (Johnson et al., 2005). The accurate specification of the Tikhonov regularization parameter is the first and most important step in the implementation of the 4DVar with Tikhonov regularization. However, this specification process is extremely difficult in a high-dimensional variational data assimilation (VDA) system, and previous studies were largely implemented with low-dimensional toy models rather than actual high-dimensional systems. Using the L-curve rule...
regularization parameter selection, Zhao et al. (2011) first tested the Tikhonov regularization method to conduct radar observation data assimilation in a real high-dimensional VDA system based on the Weather Research and Forecasting (WRF) three-dimensional variational data assimilation (3DVar) system. Zhong et al. (2014) proposed a multiple-regularization-parameter selection method with acceptable computational complexity based on the posterior information of the 4DVar system. The proposed 4DVar with multiple regularization parameters as a weak constraint had superior performance compared with the traditional 4DVar for bogus data assimilation (BDA) in the initialization and simulation of a typhoon. In the EnKF, many studies have inflated or multiplied forecast ensemble perturbations by an appropriate inflation factor to adjust ensemble variance (Evensen, 2009). In the early stage, the inflation factor is estimated empirically. Subsequently, adaptive inflation approaches have been proposed to account for inhomogeneous covariance underestimation (Anderson, 2007; Li et al., 2009; Miyoshi, 2011; Wang & Bishop, 2003). Wang and Bishop (2003) developed an approach to optimize the inflation factor. Their scheme was further improved by Li et al. (2009). All of these schemes are based on a first moment estimation of the squared observation-minus-forecast residual, which was first introduced to data assimilation by Dee (1995). Zheng et al. (2013) proposed an alternative approach to optimize the inflation factor using maximum likelihood estimation (MLE), which was introduced by Dee and colleagues (Dee et al., 1999; Dee & da Silva, 1999). Algorithm analyses showed that the inflation factor in EnKF was expressed as a reciprocal relationship to its corresponding regularization parameter in 4DVar. Generally, these approaches belong to the parameter estimation algorithms based on EnKF (e.g., ELSheikh et al., 2013).

The inclusion of the EnKF-estimated (Evensen, 2009; Houtekamer et al., 2014) and flow-dependent error covariances in 4DVar (Rabier et al., 2000) has been a subject of recent intense study (Bannister, 2017; Gustafsson, 2007; Kalnay et al., 2007; Lorenc, 2017). The achievement of the background and observation error variance regularization in their (EnKF and 4DVar) fusion is a logical necessity. Existing composite estimation algorithms based on EnKF (e.g., ELSheikh et al., 2013). Bishop (2003) developed an approach to optimize the inflation factor in its 4DVar minimization, assisted by the tangent/adjoint models (Clayton et al., 2013; Zhang et al., 2009). Conversely, by approximating the tangent models with the linear assumption between the model state perturbations (MPs) and their corresponding simulated observation perturbations (OPs), the tangent/adjoint models are no longer needed in 4DEnVar (Qiu et al., 2007; Tian et al., 2008, 2018; Tian & Feng, 2015; Wang et al., 2010) methods. En/hybrid-4DVar completely or partially utilizes the ensemble background error covariance $B$, in its 4DVar minimization, assisted by the tangent/adjoint models (Clayton et al., 2013; Zhang et al., 2009). Among the many 4DVar methods, the nonlinear least-squares 4-D variational data assimilation method (NLS-4DVar; Tian et al., 2018; Tian & Feng, 2015) is distinctive as an advanced 4DEnVar algorithm that utilizes a Gauss-Newton iterative scheme (Dennis & Schnabel, 1996) to handle the nonlinearity of the 4DVar cost function. It therefore enhances assimilation accuracy while maintaining an overall structure that still resembles traditional 4DEnVar algorithms (i.e., with no dependence on the adjoint model) and provides a flow-dependent $B$. Very recently, NLS-4DVar was further upgraded to a big-data-driven version (BD-NLS4DVar; Tian & Zhang, 2019a, 2019b), which has better performance than the standard NLS-4DVar without increased computational costs.

In this study, to minimize the 4DVar cost function with multiple regularization parameters, we utilized an alternating direction method (ADM; the convergence of ADM for two separable variables has been established for years; for more details, see Yang et al., 2017) iteratively to simultaneously optimize the regularization parameters and the model states under an ensemble forecasting framework: (1) The MLE-based approach (Zheng et al., 2013) was adopted to estimate the inflation factor (and thus the regularization parameter) by minimizing the 2log-likelihood of the observation-minus-forecast residual; and (2) the NLS-4DVar was implemented to seek the solution of the 4DVar problem with the optimized regularization parameters. We present the formulation of an adjoint-free ADM for 4DVar with multiple-parameter Tikhonov regularization in section 2. The performance evaluation of NLS-4DVar with Tikhonov regularization is presented in section 3, along with a comparison against the standard NLS-4DVar without Tikhonov regularization using the shallow-water equation model. Finally, we conclude with a summary and concluding remarks in section 4.
2. Methodology

With the growing number of observations available from remote sensing techniques, more asynchronous observations with different error properties and from a variety of sources are increasingly assimilated in real 4DVar data assimilations. In response, a multiple regularization parameters method was introduced by setting different regularization parameters for different observations and then transforming the traditional 4DVar cost function into the following function, with multiple-parameter Tikhonov regularization:

\[
J(x, \alpha, y_1, \cdots, y_{n_{\text{obs}}}) = \frac{1}{2} \alpha \|x\|_{R^{-1}}^2 + \frac{1}{2} \sum_{i=1}^{n_{\text{obs}}} y_i \|L_i(x) - y_{\text{obs}, i}\|_{R_i^{-1}},
\]

where

\[
\|x\|_{R^{-1}}^2 = x^T R^{-1} x,
\]

\[
L_k^{(i)} = H_k^{(i)} M_{b \rightarrow tk},
\]

\[
L_k^{(i)}(x') = L_k^{(i)}(x_b + x') - L_k^{(i)}(x_b),
\]

\[
y_{\text{obs}, k}^{(i)} = y_{\text{obs}, k} - L_k^{(i)}(x_b),
\]

\[
L_i(x) = \begin{pmatrix}
L_0^{(i)}(x) \\
L_1^{(i)}(x) \\
\vdots \\
L_{S}^{(i)}(x)
\end{pmatrix},
\]

\[
y'_{\text{obs}, i} = \begin{pmatrix}
y_{\text{obs}, 0}^{(i)} \\
y_{\text{obs}, 1}^{(i)} \\
\vdots \\
y_{\text{obs}, S}^{(i)}
\end{pmatrix},
\]

\[
R_i = \begin{pmatrix}
R_0^{(i)} \\
R_1^{(i)} \\
\vdots \\
R_S^{(i)}
\end{pmatrix},
\]

\[
\|L_i(x') - y'_{\text{obs}, i}\|_{R_i^{-1}} = \left[ L_i(x') - y'_{\text{obs}, i} \right]^T (R_i)^{-1} \left[ L_i(x') - y'_{\text{obs}, i} \right].
\]

Here, \(x' \in \mathbb{R}^{m_x} (m_x is the dimension of the state vector) is the perturbation of the background field \(x_b \in \mathbb{R}^{m_x} at \ t_b\), the superscript T stands for the matrix transpose, \(y_{\text{obs}, k}^{(i)} \in \mathbb{R}^{m_y^{(i)}}\) is the \(i\)th observation at \(t_b\), \(m_y^{(i)} \) (\(\sum_{k=0}^{S} m_{y, k} = m_y^{(i)}\) is the dimension of the \(i\)th observation) observation at \(t_b\), \(b\) is the background value, \(S+1\) is the total number of observation time points in the assimilation window, \(H_k^{(i)}\) is the \(i\)th observation operator at \(t_k\), \(M_{b \rightarrow tk}\) is the forecast model from \(t_0\) to \(t_k\), \(Y_i\) is the impact of the regularization parameter \(\alpha\) on each observation term \(\|L_i(x') - y'_{\text{obs}, i}\|_{R_i^{-1}}\), and \(B\) and \(R_k^{(i)}\) are the background and observation error covariance matrices, respectively.
The B-R ratio is therefore the key factor affecting the performance of 4DVar (Johnson et al., 2000). We further transformed Equation 1 into the following format:

\[
J\left(\mathbf{x}, \mu_1, \ldots, \mu_n\right) = \frac{1}{2} \|\mathbf{x}\|^2_{\mathbf{B}^{-1}} + \frac{1}{2} \sum_{i=1}^{n_{\text{obs}}} \left\|L_i\left(\mathbf{x}\right) - y_{\text{obs},i}\right\|_{\left(\mu_i \mathbf{R}_i\right)^{-1}},
\]

(10)

where \(\mu_i = \frac{\alpha}{\gamma_i}\). Thus, Equation 10 indicates that the ratio of the inflation parameters (Zheng et al., 2013) for the background error variance \(\mathbf{B}\) and \(i\)th observation error variance \(\mathbf{R}_i\) is 1 : \(\mu_i\) (\(i = 1, \ldots, n_{\text{obs}}\)).

In this study, an adjoint-free ADM is proposed to solve the 4DVar problem with the iterative application of multiple regularization parameters under a framework of ensemble forecasting as follows. For the \(j\)th \((j = 1, \ldots, j_{\text{max}})\) iterative step, (1) the MLE-based approach proposed by Zheng et al., 2013, with some modifications, was applied to estimate the inflation factor \(\lambda_i^j\) for the \(i\)th observation by minimizing the following 2log-likelihood of the observation-minus-forecast residual:

\[
L\left(\lambda_i^j\right) = \ln\left[\det\left(H\lambda_i^j\mathbf{B}(H_i)^T + \mathbf{R}_i^{(j-1)}\right)\right] + \left(\mathbf{y}_{\text{obs},i} - (H\lambda_i^j\mathbf{B}(H_i)^T + \mathbf{R}_i^{(j-1)})^{-1}\mathbf{y}_{\text{obs},i}\right)^T \left(H\lambda_i^j\mathbf{B}(H_i)^T + \mathbf{R}_i^{(j-1)}\right)^{-1} \left(H\lambda_i^j\mathbf{B}(H_i)^T + \mathbf{R}_i^{(j-1)}\right) - 1
\]

(11)

where \(H_i = \begin{pmatrix} H_1^i & 0 & \cdots & 0 \\
0 & H_2^i & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & \cdots & 0 & H_n^i \end{pmatrix}\). Thus, the ratio of the inflation parameters (Zheng et al., 2013) for the background error variance \(\mathbf{B}\) and the \(i\)th observation error variance \(\mathbf{R}_i^{(j)}\) is \(\lambda_i^j; 1\) (i.e., \(1 : \lambda_i^j\)), from which we deduce \(\mu_i^j = \frac{1}{\lambda_i^j}\).

(2) NLS-4DVar was then implemented to obtain the solution of the 4DVar problem with multiple optimized regularization parameters \(\mu_i^j\):

\[
J\left(\mathbf{x}, \mu_1, \ldots, \mu_n\right) = \frac{1}{2} \|\mathbf{x}\|^2_{\mathbf{B}^{-1}} + \frac{1}{2} \sum_{i=1}^{n_{\text{obs}}} \left\|L_i\left(\mathbf{x}\right) - y_{\text{obs},i}\right\|_{\left(\mu_i \mathbf{R}_i\right)^{-1}},
\]

(12)

(3) The observation error variances \(\mathbf{R}_i^{(j)} = \mu_i^j(\mathbf{R}_i^{(j-1)} - 1)\) were updated, particularly, \(\mathbf{R}_i^{(0)} = \mathbf{R}_i\).

Ignoring the different types of observations, Equation 10 can then be degraded into the 4DVar with single-parameter Tikhonov regularization:

\[
J\left(\mathbf{x}, \mu\right) = \frac{1}{2} \|\mathbf{x}\|^2_{\mathbf{B}^{-1}} + \frac{1}{2} \sum_{i=1}^{n_{\text{obs}}} \left\|L_i\left(\mathbf{x}\right) - y_{\text{obs},i}\right\|_{\left(\mu \mathbf{R}_i\right)^{-1}},
\]

(13)

Correspondingly, Equation 11 is also simplified as

\[
L\left(\lambda^j\right) = \ln\left[\det\left(H\lambda^j\mathbf{B}(H)^T + \mathbf{R}^{(j-1)}\right)\right] + \left(\mathbf{y}_{\text{obs}} - (H\lambda^j\mathbf{B}(H)^T + \mathbf{R}^{(j-1)})^{-1}\mathbf{y}_{\text{obs}}\right)^T \left(H\lambda^j\mathbf{B}(H)^T + \mathbf{R}^{(j-1)}\right)^{-1} \left(H\lambda^j\mathbf{B}(H)^T + \mathbf{R}^{(j-1)}\right) - 1
\]

(14)

where

\[
H = \begin{pmatrix} H_1 & H_2 & \cdots & H_{n_{\text{obs}}} \\
\end{pmatrix}
\]

(15)
\[
\mathbf{R} = \begin{pmatrix}
\mathbf{R}_1 \\
\mathbf{R}_2 \\
\vdots \\
\mathbf{R}_{nobs}
\end{pmatrix},
\]  
(16)

and

\[
\mathbf{y}_\text{obs} = \begin{pmatrix}
\mathbf{y}_\text{obs, 1}^T \\
\mathbf{y}_\text{obs, 2}^T \\
\vdots \\
\mathbf{y}_\text{obs, nobs}^T
\end{pmatrix},
\]  
(17)

Steps 1 and 2 were conducted under a framework of ensemble forecasting. The initial MPs \(x_{0,j} \in \mathbb{R}^m\) (\(j = 1, \cdots, N\)) were prepared in advance. Subsequently, ensemble forecasts \(x_{tk,j} = M_{t_k \rightarrow t_k}(x_0 + x_j')\) initiated from \(x_0 + x_j'\) could be conducted over the assimilation window \([t_0, t_S]\).

Applying the \(i\)th observation operator \(H_i(\cdot)\) to the simulated model states, the corresponding simulated observation perturbations (OPs) were then yielded as follows:

\[
y_{ijk} = H_i(\mathbf{x}_{tk,j}) - L_i(\mathbf{x}_b(\cdot)),
\]  
(18)

Furthermore,

\[
\mathbf{P}_x = \begin{pmatrix}
x_1' \\
\vdots \\
x_N'
\end{pmatrix},
\]  
(19)

\[
\mathbf{P}_y^{(ij)} = \begin{pmatrix}
y_{ij1,k}' \\
\vdots \\
y_{ijN,k}'
\end{pmatrix},
\]  
(20)

and

\[
\mathbf{P}_y^j = \begin{pmatrix}
\mathbf{P}_{y,0}^j \\
\mathbf{P}_{y,1}^j \\
\vdots \\
\mathbf{P}_{y,S}^j
\end{pmatrix},
\]  
(21)

According to the following approximation (Tian & Zhang, 2019b),

\[
H_i\mathbf{x}_j^*\mathbf{B}(H_i)^T \approx H_i\mathbf{x}_j^*(\mathbf{P}_x(\mathbf{P}_x)^T - (H_i)^T = \frac{1}{N-1} (\mathbf{P}_y^j)(\mathbf{P}_y^j)^T,
\]  
(22)

Equation 14 can be rewritten as

\[
L_i(\mathbf{\lambda}_j) = \ln \left( \det \left( \mathbf{\lambda}_j - \frac{1}{N-1} \left( \mathbf{P}_y^j \mathbf{P}_y^j \right)^T + \mathbf{R}_i^{(j-1)} \right) \right) + \left( \mathbf{y}_\text{obs, i} \right)^T \left( \mathbf{y}_\text{obs, i} \right)^* + \left( \mathbf{P}_y^j \mathbf{P}_y^j \mathbf{P}_y^j \right)^T + \left( \mathbf{R}_i^{(j-1)} \right)^{-1} y_{\text{obs, i}} \right),
\]  
(23)

Obviously, the minimal solution to Equation 23 can be computed iteratively by performing evaluations of the cost function of Equation 23 and its gradient using a suitable descent algorithm (e.g., the limited memory quasi-Newton method [L-BFGS]; Liu & Nocedal, 1989). For the convenience of calculation, we first simplified the computation of the cost function of Equation 23 and its gradient based on a singular value decomposition (SVD):
\[
\left( R_i^{(j-1)} \right)^{-1/2} \frac{P_j}{N-1} = U_j D_j V_j, \tag{24}
\]
where
\[
\left( R_i^{(j-1)} \right)^{1/2} = R_i^{(j-1)},
\]
yielding
\[
\det \left( \frac{1}{N-1} \frac{d_i^{(j)}}{P_i^{(j)}} \left( P_i^{(j)} \right)^T + R_i^{(j-1)} \right) = \det \left( R_i^{(j-1)} \right) \times \prod_{k=1}^{N} \left( \chi_i^{d_{j,k}} + 1 \right), \tag{25}
\]
and
\[
\left( \frac{y_{\text{obs}, i}}{N-1} \right)^T \left( \frac{1}{N-1} \frac{d_i^{(j)}}{P_i^{(j)}} \left( P_i^{(j)} \right)^T + R_i^{(j-1)} \right)^{-1} \frac{y_{\text{obs}, i}}{N-1} = 
\left( \left( R_i^{(j-1)} \right)^{-1/2} y_{\text{obs}, i} \right)^T U_j \text{diag}_{k=1, N} \left( \frac{1}{\chi_i^{d_{j,k}} + 1} \right) \left( U_j \right)^T \left( R_i^{(j-1)} \right)^{-1/2} \frac{y_{\text{obs}, i}}{N-1}. \tag{26}
\]
Here,
\[
\text{diag}_{k=1, N} \left( \frac{1}{\chi_i^{d_{j,k}} + 1} \right) = 
\begin{bmatrix}
1 & 0 & \cdots & 0 \\
\frac{1}{\chi_i^{d_{j,1}} + 1} & \frac{1}{\chi_i^{d_{j,2}} + 1} & \ddots & \\
\vdots & \ddots & \ddots & 0 \\
0 & \cdots & 0 & \frac{1}{\chi_i^{d_{j,N}} + 1}
\end{bmatrix}, \tag{27}
\]
and, \( d_{j,k} \) is the \( k \)th diagonal element of \( D_j \).

Simple mathematical calculations (see Zheng et al., 2013) yield
\[
L_i^{'}(\chi_i) = \ln \left[ \det \left( R_i^{(j-1)} \right) \right] + \sum_{k=1}^{N} \ln \left( \chi_i^{d_{j,k}} + 1 \right)
+ \left( \left( R_i^{(j-1)} \right)^{-1/2} y_{\text{obs}, i} \right)^T U_j \text{diag}_{k=1, N} \left( \frac{1}{\chi_i^{d_{j,k}} + 1} \right) \left( U_j \right)^T \left( R_i^{(j-1)} \right)^{-1/2} \frac{y_{\text{obs}, i}}{N-1}, \tag{28}
\]
and
\[
\frac{\partial L_i^{'}(\chi_i)}{\partial \chi_i} = \sum_{k=1}^{N} \frac{d_{j,k}^2}{\chi_i^{d_{j,k}} + 1} + 
\left( \left( R_i^{(j-1)} \right)^{-1/2} y_{\text{obs}, i} \right)^T U_j \text{diag}_{k=1, N} \left( -\frac{d_{j,k}^2}{\left( \chi_i^{d_{j,k}} + 1 \right)^2} \right) \left( U_j \right)^T \left( R_i^{(j-1)} \right)^{-1/2} \frac{y_{\text{obs}, i}}{N-1}, \tag{29}
\]
where
\[
\text{diag}_{k=1, N} \left( -\frac{d_{j,k}^2}{\chi_i^{d_{j,k}} + 1} \right) = 
\begin{bmatrix}
-\frac{d_{j,1}^2}{\left( \chi_i^{d_{j,1}} + 1 \right)^2} & 0 & \cdots & 0 \\
0 & -\frac{d_{j,2}^2}{\left( \chi_i^{d_{j,2}} + 1 \right)^2} & \ddots & \\
\vdots & \ddots & \ddots & 0 \\
0 & \cdots & 0 & -\frac{d_{j,N}^2}{\left( \chi_i^{d_{j,N}} + 1 \right)^2}
\end{bmatrix}. \tag{30}
\]
We used L-BFGS software to minimize Equation 23 iteratively by performing evaluations of Equations 28 and 29 to obtain $\lambda^l_m$; thus, $\mu^l_m = \frac{1}{\lambda^l_m}$. We accomplished the first step to estimate the inflation factor $\lambda^l_m$ for all of the ith observations by minimizing Equation 11 in the jth outer iteration.

In Step 2, NLS-4DVar was applied to express the analysis increment $x^l_i \in \mathbb{R}^{m_i}$ by the linear combinations of the MPs, $P_x \in \mathbb{R}^{m_x \times N}$, as

$$x^l_i = P_x \beta^l_i.$$  

Substituting Equation 31 and the ensemble background error covariance $B = \frac{(P_x)(P_x)^T}{N - 1}$ (Evensen, 2009) into Equation 12 with the optimized $\mu^l_i$ and expressing the cost function in terms of the new control yields $\hat{\beta}$ (for more details, see Tian et al., 2018):

$$J(\beta) = \frac{1}{2}(N - 1)\beta^T \beta + \frac{1}{2} \sum_i n_{\text{obs}} \left| L^i_i(P_x \beta) - y_{\text{obs},i} \right|^2 (\mu^l_i R^i_i)^{-1}.$$  

To circumvent the use of the adjoint models, Equation 32 is further rewritten as a nonlinear least-squares problem, which, by making some necessary approximations, is solved by the Gauss-Newton iteration scheme as follows:

$$\beta^l_i = \beta^l_{i-1} - \left( (N - 1)I_N + \frac{\sum_i n_{\text{obs}}}{2} \left( \mu^l_i R^i_i \right)^{-1} \left( P_x^i \right)^T \left( P_x^i \right) \right)^{-1} \times \left\{ \sum_i n_{\text{obs}} \left( \mu^l_i R^i_i \right)^{-1} L^i_i (P_x \beta_i^{l-1}) - y_{\text{obs},i} \right\} + (N - 1)\beta^l_{i-1},$$

for $l = 1, \cdots, l_{\text{max}}$, where $l_{\text{max}}$ is the maximum iteration number, and $I_N$ denotes the $N \times N$ identity matrix (see Tian et al., 2018 for more details).

To filter out the spurious long-range correlations resulting from the finite ensemble number $N$, Tian et al. (2018) localized Equation 33 as follows:

$$\beta^l_{i,\rho} = \beta^l_{i-1} + \sum_{l=1}^{n_{\text{max}}} \left( \rho^{i,\rho}_{0} < \rho < \rho^{i,\rho}_{e} \right) L^i_i (P_x \beta_i^{l-1}) + \sum_{l=1}^{n_{\text{max}}} \left( \rho^{i,\rho}_{0} < \rho < \rho^{i,\rho}_{e} \right) \left( \mu^l_i R^i_i \right)^{-1} \left( y_{\text{obs},i} - L^i_i (P_x \beta_i^{l-1}) \right),$$

and

$$x^{l-1} = (\rho_m < \rho < P_x) \beta^l_{i,\rho},$$

where

$$P_x^{i,\rho} = - (N - 1) \left( P_x^i \right)^T \left( P_x^i \right)^{-1} \left( (N - 1)I_N + \left( P_x^i \right)^T \left( \mu^l_i R^i_i \right)^{-1} \left( P_x^i \right) \right)^{-1},$$

and

$$P_x^{i,\rho} = \left( P_x^i \right)^T \left( (N - 1)I_N + \left( P_x^i \right)^T \left( \mu^l_i R^i_i \right)^{-1} \left( P_x^i \right) \right)^{-1}.$$  

Here, $\beta^l_{i,\rho} \in \mathbb{R}^{N \times r}$, $\rho_m \in \mathbb{R}^{m_x \times r}$, $\rho_m(\rho_m)^T = C_{mm} \in \mathbb{R}^{m_x \times m_x}$, $C_{mm}$ is the spatial correlation matrix, $\rho^{i,\rho}_{0} \in \mathbb{R}^{m_x \times r}$ is computed together with $\rho_m$, and $r$ is the selected truncation mode number (Zhang & Tian, 2018). The definition of the notation "$(\cdot < \cdot < \cdot)$" is provided in Zhang and Tian (2018).
Additionally, NLS-4DVar uses the square root analysis scheme (Evensen, 2009) for the online ensemble update in the assimilation cycles, as follows:

\[
P_a^x = P_x V_1 \sqrt{I - \Sigma_2^T \Sigma_2} \Phi^T, \tag{38}
\]

where

\[
U_2 \Sigma_2 V_2^T = X_2, \tag{39}
\]

\[
X_2 = \Lambda^{-1/2} Z^T P_y, \tag{40}
\]

\[
Z \Lambda^{-1} Z^T = \left[ P_y P_y^T + (N - 1) R \right]^{-1}, \tag{41}
\]

\[
P_y = \begin{pmatrix}
P^1_y \\
P^2_y \\
\vdots \\
P^n_{obs}^y
\end{pmatrix}, \tag{42}
\]

and \( \Phi \) is a random orthogonal matrix (for more details, see Evensen, 2009). Noticeably, the diagonal elements of \( \sqrt{I - \Sigma_2^T \Sigma_2} \) are less than 1.0, which would result in a gradual ensemble collapse. Since \( R \) is symmetric, it has the Cholesky factorization \( R = R_{+}^{1/2} \left( R_{+}^{1/2} \right)^T \), and the eigenvalue decomposition

\[
Z \Lambda^{-1} Z^T = \left[ P_y P_y^T + (N - 1) R \right]^{-1}
\]

\[
= R_{+}^{-1/2} \left[ \left( R_{+}^{-1/2} P_y \right) \left( R_{+}^{-1/2} P_y \right)^T + (N - 1) I \right]^{-1} \left( R_{+}^{-1/2} \right)^T
\]

\[
= R_{+}^{-1/2} \left[ \left( U \Sigma V^T \right) \left( U \Sigma V^T \right)^T + (N - 1) I \right]^{-1} \left( R_{+}^{-1/2} \right)^T
\]

\[
= R_{+}^{-1/2} U \left[ \Sigma^2 + (N - 1) I \right]^{-1} U^T \left( R_{+}^{-1/2} \right)^T,
\]

\[
Z = R_{+}^{1/2} U, \tag{44}
\]

and

\[
\Lambda = \left[ \Sigma^2 + (N - 1) I \right]. \tag{45}
\]

To maintain the ensemble spread, we further transformed Equation 38 into the following format:

\[
P_a^x = P_x V_2 \Phi^T. \tag{46}
\]

It can be seen that Equation 46 is actually a special inflation of Equation 38, which is used for the online ensemble update in the assimilation cycles. However, we still used Equation 38 to construct the analysis error covariance \( B^a \).
3. Observing System Simulation Experiments Using the Shallow-Water Equation Model

Several groups of observing system simulation experiments (OSSEs) based on the shallow-water equation model were designed to demonstrate the potential merits and advantages of the NLS-4DVar with single/multiple parameter Tikhonov regularization, especially compared to the standard NLS-4DVar method with no regularization.

3.1. Shallow-Water Equation Model

The following 2-D shallow-water equations, with the f-plane formulation, were utilized as the forecast model for the OSSEs:

\[
\frac{\partial u}{\partial t} = -u \frac{\partial u}{\partial x} - v \frac{\partial u}{\partial y} + f v - g \frac{\partial h}{\partial x}, \tag{47}
\]

\[
\frac{\partial v}{\partial t} = -u \frac{\partial v}{\partial x} - v \frac{\partial v}{\partial y} - f u - g \frac{\partial h}{\partial y}, \tag{48}
\]
\[ \frac{\partial h}{\partial t} = -u \frac{\partial (h - h_s)}{\partial x} - v \frac{\partial (h - h_s)}{\partial y} - (H + h - h_s) \left( \frac{\partial u}{\partial x} + \frac{\partial u}{\partial y} \right), \quad (49) \]

where \( f = 7.272 \times 10^{-5} \text{s}^{-1} \) is the Coriolis parameter, \( H = 3,000 \text{ m} \) is the basic state depth, \( h_s = h_0 \sin(4\pi x/L_x)[\sin(4\pi y/L_y)]^2 \) is the terrain height, \( h_0 = 250 \text{ m} \), and the lengths of the two sides of the computational domain are \( D = L_x = L_y = 44d \text{ km} \), respectively, where \( d = \Delta x = \Delta y = 300 \text{ km} \) is the uniform grid size. The domain was partitioned into a square mesh with 45 grid points in each coordinate direction, and periodic boundary conditions were imposed at \( x = (0, L_x) \) and \( y = (0, L_y) \). The second-order central finite difference and two-step backward difference schemes of Matsuno (1996) were selected to discretize the spatial and local time derivatives, respectively, and thereby ensure computational stability (Qiu et al., 2007). The time step was 360 s (6 min). It should be noted that the damping effect deserves special attention in the Matsuno scheme (Pfeffer et al., 1992). The model state vector was represented by height \( h \), and the horizontal velocity components were represented by \( u \) and \( v \) at the grid points.

### 3.2. Experimental Setup

The true initial fields for all OSSEs were first produced by integrating the perfect (i.e., \( h_0 = 250 \text{ m} \)) shallow-water equation model from the following initial conditions for 60 hr:

![Figure 1. Time series of spatially and temporally averaged root mean square errors (RMSEs) of height and wind for BD-NLS4DVar with multiple-/single-parameter (NLS-TK_m) and without (NLS) Tikhonov regularization for the (a, c) perfect model (\( h_0 = 250 \text{ m} \)) and the (b, d) imperfect model (\( h_0 = 0 \text{ m} \)).](image-url)
\[ h = 360 \left( \sin \left( \frac{\pi y}{D} \right) \right)^2 + 120 \sin \left( \frac{2\pi x}{D} \right) \sin \left( \frac{2\pi y}{D} \right), \]  
\[ u = -f^{-1} \frac{\partial h}{\partial y} \text{ and } v = -f^{-1} \frac{\partial h}{\partial x} \text{ at } t = -60. \]  

Correspondingly, the background state \( x_0 \) was produced in a similar manner but using the imperfect model (i.e., \( h_0 = 0 \) m). The different model configurations resulted in significant differences between the background and true states (not shown) due to the two 60-hr model integrations under the perfect- and
imperfect-model scenarios, respectively. Specifically, the spatially averaged root mean square errors (RMSEs) were 23.4 m, 1.53 m s\(^{-1}\), and 2.58 m s\(^{-1}\) for \(h\), \(u\), and \(v\), respectively.

There were six assimilation cycles (windows) in each evaluation experiment. The length of each window was 12 hr, and observations were available every 3 hr (i.e., at 3, 6, 9, and 12 hr during each assimilation window). Each model grid had one randomly distributed observation site, resulting in a total of 44 × 44 observational sites at each observation time point. The observations were generated by adding random noise to the true values at the observational locations, which were obtained using a simple bilinear interpolation method. Therefore, the observation operator \(H_k\) was simply the bilinear interpolation operator.

The big-data-driven version of NLS-4DVar was adopted to minimize the 4DVar cost function with/without Tikhonov regularization in the OSSEs. The total initial MPs \(P_x = P_x, h, o\) ∈ \(\mathbb{R}^{m_x × N}\) of BD-NLS4DVar was composed of two ensembles, a preprepared historical big-data ensemble \((P_x, h) ∈ \mathbb{R}^{m_x × N_h}\) and a small online ensemble \((P_x, o) ∈ \mathbb{R}^{m_x × N_o}\), where \(N_h+N_o = N(N_h > N_o)\). See Tian and Zhang (2019a, 2019b) for more details about the preparation and update of the ensemble for the BD-NLS4DVar. The default parameter values were \(N_h = 10, N_o = 100, \alpha = 11, \beta = 2, j_{max} = 1, \) and \(j_{max} = 4\), and the covariance localization Schür radius \(r_x = r_y = 15 \times 300\) km for the \(x\) and \(y\) coordinates, respectively (Gaspari & Cohn, 1999). Also see Tian and Zhang (2019a) for the definition of \(\alpha\) and \(\beta\). The 4-D moving strategy (Tian & Feng, 2015) was adopted to produce initial MPs for BD-NLS4DVar with/without Tikhonov regularization, respectively. Both the perfect \((h_0 = 250\) m) and imperfect \((h_0 = 0\) m) models were used as forecast models in all
subsequent OSSEs to investigate the BD-NLS4DVar with Tikhonov regularization under the perfect and imperfect model scenarios, respectively.

### 3.3. Experimental Results

We first evaluated the performance of BD-NLS4DVar with multiple- (NLS-TK_m) and single-parameter (NLS-TK_s) Tikhonov regularization and without regularization (referred to as NLS). The spatially and temporally averaged RMSEs of height (\(r_{h_{mse}}\)) and wind (\(r_{v_{mse}}\))(see Tian & Zhang, 2019b for definitions of \(r_{h_{mse}}\) and \(r_{v_{mse}}\)) for NLS-TK_m/s and NLS were compared for each assimilation window under the perfect-model assumption (\(h_0 = 250\) m) for all true, forecast, and assimilation runs (Figures 1a and 1c). Figures 1a and 1c show that both NLS-TK_m and NLS-TK_s performed significantly better than NLS. We suggest this is due to the proposed (multiple/single) regularization parameter ADM selection approach, which could produce an accurate ratio of background and observation error covariances iteratively, resulting in superior performance compared to the NLS (i.e., the BD-NLS4DVar without regularization). In addition, the accuracy of the observation error variances could be effectively improved by the multiple regularization parameters, which contributed significantly to the substantially better performance of NLS-TK_m compared with NLS-TK_s.

Another group of experiments using severe model error was conducted by changing the maximum terrain height (\(h_0 = 0\) m) from that used in the truth simulations (\(h_0 = 250\) m). Under this model error, BD-NLS4DVar with and without Tikhonov regularization performed slightly worse than under the perfect model scenario, with increased RMSEs for both the height and wind components. Nevertheless, the NLS-

![Figure 4](https://example.com/figure4.png)
TK_m and NLS-TK_s methods still outperformed the original BD-NLS4DVar (i.e., without Tikhonov regularization), leading to substantially lower RMSEs of NLS-TK_m and NLS-TK_s in terms of both the height and wind fields than were observed in NLS (Figures 1b–1d). Not surprisingly, NLS-TK_m still performed much better than NLS-TK_s under this severe model error due to its better derived ratios of the background and observation error variances. These results clearly demonstrate that the implementation of the Tikhonov regularization scheme (especially the multiple schemes) improved the data assimilation performance of BD-NLS4DVar with and without model error.

To further investigate the impacts of Tikhonov regularization on the (NLS-)4DVar method, the variations in the multiple ($\mu_i, i = 1,2,3$) and single ($\mu$) regulation parameters with increasing iteration steps in the first assimilation window are shown in Figure 2. It was clear that, along with the increase in iteration time, both the multiple and single regularization parameters would gradually approach 1.0 for both the perfect and imperfect model assumptions. This indicated that the $B$-$R$ ratio was adjusted to an accurate level through the proposed ADM. Consequently, the improvement of the $B$-$R$ ratio helped to increase NLS-4DVar accuracy. Additionally, the differences in the three regulation parameters (i.e., $\mu_i, i = 1,2,3$) in NLS-TK_m with much lower RMSEs compared with NLS-TK_s further demonstrated the necessity of setting different regularization parameters for different observations.

The MPs have a very important role in the proposed ADM because they approximate the background error variance $B$ and the composite tangent models $H_0M_0$ (Tian et al., 2018) in our NLS-4DVar method. To explore
the sensitivity of the proposed ADM to the initial MPs, we compared NLS-TK_m performance with that of three different initial online MPs (i.e., $5P_{x,h}$, $P_{x,o}$, and $1/5P_{x,o}$). Their impacts on the NLS-4DVar for the imperfect model scenario (Figures 3b and 3d) were far less significant than those for the perfect model scenario (Figures 3a and 3c). However, as the NLS-TK_m proceeded, the adverse impacts of the initial MPs gradually decreased due to the Tikhonov regularization.

Obviously, the predetermined observation error variances $R$ would be of much more importance to the final NLS-TK_m assimilation performance. To investigate the impact of the predetermined $R$ on the NLS-TK_m results, three further experiments were conducted using NLS-TK_m with three different predetermined $R$s (i.e., $10^2 R$, $R$, and $10^{-2} R$) to contrast with each other. Similarly, the NLS-TK_m RMSEs derived from the imperfect model scenario (Figures 4b and 4d) were actually more influenced by the various predetermined $R$s than those from the perfect model scenario (Figures 4a and 4c). Compared with the initial MPs, the effect of contamination of and inaccurate $R$ value on the NLS-TK_m results was difficult to correct, even though the impact could be gradually reduced, especially for the imperfect model scenario (Figures 4b and 4d). Fortunately, compared with the background error variance $B$, the observation error variances $R$ were usually confirmed easily and accurately.

Our final experiment examined the sensitivity of the impact of ensemble sizes $N_h$ and $N_o$ on the NLS-TK_m results by varying $N_h (=5, 10, and 20)$/$N_o (=80, 100, and 120$) and comparing with fixed values $N_h = 100/ N_o = 10$ for the imperfect model scenario. As with the original BD-NLS4DVar with no regularization, the increase in the $N_h = 100$ usually resulted in improved NLS-TK_m assimilation performance, with much lower RMSEs (Figures 5b and 5d). Conversely, larger sizes ($N_o$) of the online MPs did not result in better assimilation results (Figures 5a and 5c). These results were combined with those from Tian and Zhang (2019a) to further verify the results, and the following conclusions were reached. The BD-NLS4DVar method (with/without Tikhonov regularization) was similar to the hybrid-4DVar method, where the historical ensemble $P_{x,h}$ (size $N_h$) was analogous to the climatological ensemble (to construct the climatological covariance, $B$), and the online ensemble represented the instantaneous and flow-dependent perturbations $P_{x,o}$ (size $N_o$), as with its counterpart in the hybrid-4DVar.

### 4. Summary and Concluding Remarks

Accurate specification of both the background ($B$) and observational ($R$) error covariances is essential to the success of data assimilation. For 4DVar, the $B$-$R$ ratio is particularly critical to maximizing the information that can be extracted from the observations. To provide an accurate $B$-$R$ ratio, the Tikhonov regularization theory was thus simultaneously incorporated into 4DVar to optimize the model states and the variance ratios. Furthermore, the $B$-$R$ ratios can be interpreted as the (Tikhonov) regularization parameters (Johnson et al., 2005).

However, it is difficult to specify the regularization parameters in a high-dimensional VDA system for both the single- and especially the multiple-parameter regularization schemes. In the multiple-parameter regularization scheme, it became increasingly important to deal with different error properties from a variety of observations. Correspondingly, the EnKF usually adopts the inflation strategy to adjust the ensemble variances, which is relatively simple to implement, especially in terms of computational complexity.

In this study, we incorporated an MLE-based inflation approach into the 4DVar Tikhonov regularization. An ADM was proposed to minimize the 4DVar cost function with multiple regularization parameters iteratively to simultaneously optimize the regularization parameters and the model states under the framework of the NLS-4DVar method. The MLE-based approach was first adopted to estimate the regularization parameters (according to the reciprocal relationship between the regularization and inflation parameters) by minimizing the log-likelihood of the observation-minus-forecast residual. The NLS-4DVar was then implemented to obtain the solution of the 4DVar problem with optimized regularization parameters.

We conducted a comprehensive performance evaluation of the BD-NLS4DVar method with the proposed single-/multiple-parameter Tikhonov regularization using several groups of OSSEs based on 2-D shallow-water equations and compared the results with those obtained by the standard BD-NLS4DVar method. These experiments showed that the BD-NLS4DVar method with single- or multiple-parameter Tikhonov regularization performed substantially better than the BD-NLS4DVar method alone when all had the
same online/historical ensemble sizes. Additionally, the BD-NLS4DVar method with multiple-parameter regularization (NLS-TK_m) performed much better than the corresponding single-parameter regularization method because NLS-TK_m provided better B-R ratios.

The proposed NLS-TK_m method has many advantages and great potential due to its superior performance, which was achieved without noticeably increasing computational costs, as well as its ease of implementation. The potential applications of NLS-TK_m in satellite radiance data assimilation are particularly promising due to the possible large discrepancies between the measured and simulated radiance, especially within regional models (Zhu et al., 2019). It is generally known that satellite data assimilation is responsible for much of the improvement in forecasting skill in global NWP over recent decades, especially in the Southern Hemisphere (Le Marshall et al., 2007). Therefore, our future work will focus on applying this proposed method in more complex real-world models to further investigate its potential application in real satellite data assimilation. Moreover, how to improve the B-R ratios and model perturbations simultaneously is indeed quite important in ensemble-based assimilation methods and still needs to be addressed.

Data Availability Statement

The data used in this paper were obtained by the authors through numerical experiments using the shallow-water equation model and are all freely available (https://pan.baidu.com/s/1mLpbGQHxgTsC4wNW2RjCw. Password: kgjl).

References

Anderson, J. L. (2007). An adaptive covariance inflation error correction algorithm for ensemble filters. Tellus A, 59(2), 210–224. https://doi.org/10.1111/j.1600-0870.2006.00216.x

Bannister, R. N. (2017). A review of operational methods of variational and ensemble-variational data assimilation. Quarterly Journal of the Royal Meteorological Society, 143(703), 607–633. https://doi.org/10.1002/qj.2882

Clayton, A. M., Lorenz, A. C., & Barker, D. M. (2013). Operational implementation of a hybrid ensemble/4D-Var global data assimilation system at the Met Office. Quarterly Journal of the Royal Meteorological Society, 139(675), 1445–1461. https://doi.org/10.1002/qj.2054

Dec, D. P. (1995). On-line estimation of error covariance parameters for atmospheric data assimilation. Monthly Weather Review, 123(4), 1128–1145. https://doi.org/10.1175/1520-0493(1995)123<1128:OEOLIF>2.0.CO;2

Dee, D. P., & da Silva, A. M. (1999). Maximum-likelihood estimation of forecast and observation error covariance parameters Part I: Methodology. Monthly Weather Review, 127, 1822–1834. https://doi.org/10.1175/1520-0493(1999)127<1822:MLEOCP>2.0.CO;2

Dee, D. P., Gaspari, G., Redder, C., Rukhovets, L., & da Silva, A. M. (1999). Maximum-likelihood estimation of forecast and observation error covariance parameters Part II: Applications. Monthly Weather Review, 127, 1835–1849. https://doi.org/10.1175/1520-0493(1999)127<1835:MLEOCP>2.0.CO;2

Dennis, J. E., & Schnabel, R. B. Jr. (1996). Numerical methods for unconstrained optimization and nonlinear equations. In Classics in applied mathematics (pp. 221–233). Philadelphia: SIAM. https://doi.org/10.1137/1.9781611971970

ELSheikh, A. H., Pain, C. C., Fang, F., Gomes, J. L. M. A., & Navon, I. M. (2013). Navon parameter estimation of subsurface flow models using iterative regularized ensemble Kalman filter. Stochastic Environmental Research and Risk Assessment, 27(4), 877–897. https://doi.org/10.1007/s00477-012-0613-x

Evensen, G. (2009). Ensemble Kalman filter. In Data assimilation: The ensemble Kalman filter. Springer Dordrecht Heidelberg London New York: Springer Science & Business Media. https://doi.org/10.1007/978-1-4419-5515-8

Gaspari, G., & Cohn, S. E. (1999). Construction of correlation functions in two and three dimensions. Quarterly Journal of the Royal Meteorological Society, 125(554), 723–757. https://doi.org/10.1002/qj.49712555417

Gustafsson, N. (2007). Discussion on ‘4D-Var or EnKF?’. Tellus A, 59(5), 774–777. https://doi.org/10.1111/j.1600-0870.2007.00262.x

Houtekamer, P. L., He, B., & Mitchell, H. L. (2014). Parallel implementation of an ensemble Kalman filter. Monthly Weather Review, 142(3), 1163–1182. https://doi.org/10.1175/MWR-D-13-00011.1

Johnson, C., Hoskins, B. J., & Nichols, N. K. (2005). A singular vector perspective of 4D-Var: Filtering and interpolation. Quarterly Journal of the Royal Meteorological Society, 131(605), 1–19. https://doi.org/10.1256/qj.03.231

Johnson, C., Nichols, N. K., & Hoskins, B. J. (2000). Very large inverse problems in atmosphere and ocean modeling. International Journal for Numerical Methods in Fluids, 47, 759–771. https://doi.org/10.1002/fld.869

Kalnay, E., Li, H., Miyoshi, T., Yang, S. C., & Ballabres-Poy, J. (2007). 4D-Var or ensemble Kalman filter? Tellus A: Dynamic Meteorology and Oceanography, 59A, 758–773. https://doi.org/10.1111/j.1600-0870.2007.00261.x

Le Marshall, J., Uccellini, L., Einaudi, F., Colton, M., Chang, S., Weng, F., et al. (2007). The joint center for satellite data assimilation. Bulletin of the American Meteorological Society, 88(3), 329–340. https://doi.org/10.1175/BAMS-88-3-329

Li, H., Kalnay, E., & Miyoshi, T. (2009). Simultaneous estimation of covariance inflation and observation errors within an ensemble Kalman filter. Quarterly Journal of the Royal Meteorological Society, 135(639), 523–533. https://doi.org/10.1002/qj.371

Liu, D. C., & Nocedal, J. (1989). On the limited memory BFGS method for large scale optimization. Mathematical Programming, 45(1–3), 503–528. https://doi.org/10.1007/BF01589116

Lorenc, A. C. (2003). Recommended nomenclature for EnVar data assimilation methods. In Research Activities in Atmospheric and Oceanic Modelling (p. 2). Retrieved from http://www.wcrp-climate.org/WGNE/BlueBook/2013/individual-articles/03_Lorenc_Arnel_EnVar_nomenclature.pdf

Lorenc, A. C. (2017). Improving ensemble covariances in hybrid variational data assimilation without increasing ensemble size. Quarterly Journal of the Royal Meteorological Society, 143(703), 1062–1072. https://doi.org/10.1002/qj.2990

Acknowledgments

The work was partially supported by the National Key Research and Development Program of China (2016YFA0600203), the National Natural Science Foundation of China (41575100), and the Key Research Program of Frontier Sciences, Chinese Academy of Sciences (CAS) (QYZDYSSW-DQC012). The authors acknowledge two anonymous reviewers for their helpful suggestions.
Tian, X. J., Zhang, H. Q., Feng, X., & Xie, Y. (2018). Nonlinear least squares En4DVar to 4DEnVar methods for data assimilation: Theoretical formulation and conceptual evaluation. *Earth and Space Science*, 11(2), 12. https://doi.org/10.1002/2019EA000963

Tian, X. J., & Feng, X. B. (2015). A nonlinear least squares enhanced POD-4DVar algorithm. *Earth and Space Science*, 6. https://doi.org/10.1002/2019EA000735

Zhong, J., Fei, J. F., Cheng, X. P., & Huang, X. G. (2014). The four dimensional variational data assimilation with multiple regularization parameters as a weak constraint (Tikh-4D-Var) and its preliminary application on typhoon initialization. *Advances in Atmospheric Sciences*, 31(3), 2690–2701. https://doi.org/10.1007/s00376-014-4873-9

Zhu, K. F., Xue, M., Pan, Y. J., Hu, M., Benjamin, S. G., Wegrandt, S. S., & Lin, H. D. (2019). The impact of satellite radiance data assimilation within a frequently updated regional forecast system using a GSI-based ensemble Kalman filter. *Advances in Atmospheric Sciences*, 36(12), 1308–1326. https://doi.org/10.1007/s00376-019-9011-3