A DRP–4DVar Data Assimilation Scheme for Typhoon Initialization Using Sea Level Pressure Data

YING ZHAO
State Key Laboratory of Severe Weather, Chinese Academy of Meteorological Sciences, Beijing, China, and Institute of Sciences, University of Science and Technology of the Chinese People’s Liberation Army, Nanjing, China

BIN WANG AND JUANJUAN LIU
State Key Laboratory of Numerical Modeling for Atmospheric Sciences and Geophysical Fluid Dynamics, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, China

(Manuscript received 29 October 2010, in final form 16 November 2011)

ABSTRACT

In this study, a new data assimilation system based on a dimension-reduced projection (DRP) technique was developed for the fifth-generation Pennsylvania State University–NCAR Mesoscale Model (MM5) modeling system. As an initial step to test the newly developed system, observing system simulation experiments (OSSEs) were conducted using a simulated sea level pressure (SLP) field as “observations” and assimilation experiments using a specified SLP field to evaluate the effects of the new DRP–four-dimensional variational data assimilation (4DVar) method, initialization, and simulation of a tropical storm—Typhoon Bilis (2006) over the western North Pacific. In the OSSEs, the “nature” run, which was assumed to represent the “true” atmosphere, was simulated by the MM5 model, which was initialized with the 1.0° × 1.0° NCEP final global tropospheric analyses and integrated for 120 h. The simulated SLP field was then used as the observations in the data assimilation. It is shown that the MM5 DRP–4DVar system can successfully assimilate the (simulated) model output (used as observations) because the OSSEs resulted in improved storm-track forecasts. In addition, compared with an experiment that assimilated the SLP data fixed at the end of a 6-h assimilation window, the experiment that assimilated the SLP data every 3 min in a 30-min assimilation window further improved the typhoon-track forecasts, especially in terms of the initial vortex location and landfall location. Finally, the assimilation experiments with a specified SLP field have demonstrated the effectiveness of the new method.

1. Introduction

Considerable progress has been made over the past decade or so in the prediction of tropical cyclones (TCs) using numerical models and variational initialization techniques. The four-dimensional variational data assimilation (4DVar) with adjoint technique can produce dynamically and thermodynamically consistent initial conditions (ICs) and allow an optimal comparison between the model and observations at different times in the assimilation window through model trajectory (e.g., Guo et al. 2000; Wang et al. 2000; Xiao et al. 2000; Zou and Xiao, 2000; Zou et al. 2001; Pu et al. 2002; Xiao et al. 2002; Zhang et al. 2003; Zhao et al. 2005). However, establishing the adjoint model is difficult, and the cost of gradient computation is very large. As a result, although some numerical weather prediction centers are using 4DVar, many such centers and research institutions are still using three-dimensional variational data assimilation (3DVar) for model initialization because it is much less expensive than its 4DVar counterpart. Though some simple dynamical balance constraints have been utilized in 3DVar procedures without model constraints, the three-dimensional fields of the ICs from the 3DVar assimilation cannot be globally adjusted; thus, the ICs are not necessarily consistent with the full physics of the prediction model. Therefore, we need a more efficient data assimilation method, which is a crucial step toward improving the ICs for a forecast model.
Wang and Zhao (2006) proposed a new data assimilation approach, called three-dimensional variational data assimilation of mapped observations (3DVM). This approach operates under the supposition that all the model’s variables are measured. The 3DVM can produce an optimal IC at the end of the assimilation window without the need for an adjoint technique and, thus, greatly reduce the required computational cost to a level equivalent to a 3DVar approach. We developed the 3DVM assimilation system for the fifth-generation Pennsylvania State University–National Center for Atmospheric Research (Penn State–NCAR) Mesoscale Model (MM5), and have demonstrated its capability to assimilate the bogus sea level pressure (SLP) data and/or temperature and wind data retrieved from the Advanced Microwave Sounding Unit (AMSU-A) for initializing and simulating the landfall of Typhoon Dan (1999) (Zhao et al. 2007; Zhao and Wang 2008; Zhao et al. 2008). However, since the irregularly distributed observations usually have far fewer degrees of freedom than the model, deciding how to map these observations onto regular model grids is crucial to the 3DVM technique.

Recently, so-called hybrid (or ensemble-based variational data assimilation) techniques have become an attractive research topic. A hybrid ensemble transform Kalman filter 3DVar (ETKF–3DVar) method was recently introduced by Wang et al. (2008a). Wang et al. (2008b) demonstrated that the ETKF–3DVar system can provide analyses that are 15%–20% more accurate than the 3DVar but is only about twice as computationally expensive as the traditional 3DVar. Combining the necessary components from ensemble Kalman filter (EnKF) and 4DVar, Liu et al. (2008) proposed the En4DVar scheme. A great advantage of the En4DVar approach is that the tangent linear and adjoint models are not needed for its formulation and implementation.

These studies are encouraging, which suggests that the ensemble information may be used in the variational data assimilation to improve the analyses. However, to keep the classical EnKF algorithm, the background error covariance of an ensemble-based variational algorithm, which is estimated from the ensemble forecasts, must be constructed under the assumption that the error correlation is Gaussian and is a degeneration matrix.

Similar to the classical 4DVar, which can simultaneously and optimally combine the background and all observations during the assimilation window, an economical ensemble-based approach implementing 4DVar was recently proposed by Wang et al. (2010) using the dimension-reduced projection (DRP) technique. An optimal solution is directly obtained in a lower-dimension sample space by fitting observations with the historical time series generated by the model to form consistent forecast states. The assimilation system operates at a low cost when it is used by any numerical model that does not implement the adjoint of tangent linear approximation. Idealized experiments with the Lorenz-96 model have been conducted to show that the DRP–4DVar method works well and produces more accurate analyses than does the EnKF when assimilating observations in an 18-h assimilation window (Liu et al. 2011). The performance of the DRP–4DVar assimilation of 6-hourly accumulated rainfall and temperature observations has been assessed using observing system simulation experiments (OSSEs) (Wang et al. 2010).

In this study, we introduce a DRP–4DVar system for the assimilation of typhoon SLP information using the MM5. Our main objective is to test the efficiency and robustness of the proposed assimilation technique for typhoon initialization with SLP fields using Typhoon Bilis (2006) as an example. We will test the MM5 DRP–4DVar system using both OSSEs with simulated SLP fields and DRP–4DVar assimilation experiments with bogus SLP observations. The rest of the paper is organized as follows. In section 2, we introduce the MM5 DRP–4DVar system. Sections 3 and 4 describe our design of the OSSEs and their results. Section 5 demonstrates the DRP–4DVar performance using assimilation experiments with specified SLP data. Concluding remarks are given in the last section.

2. MM5 DRP–4DVar system

An adjoint-based 4DVar simulation produces an optimal increment $x^*_t$ of the ICs by minimizing the cost function $J(\cdot)$ in an assimilation window of $[t_0, t_N]$ ($t_N - t_0 \leq 6$ h):

$$ J(x^*_t) = \min_x J(x^*_t) $$

$$ J(x^*_t) = \frac{1}{2} (x^*_t)^T B^{-1} x^*_t + \frac{1}{2} (y'(x^*_t) - y_{obs})^T,$$

where

$$ J(x^*_t) = \frac{1}{2} (x^*_t)^T B^{-1} x^*_t + \frac{1}{2} (y'(x^*_t) - y_{obs})^T,$$

$$ y^T = [y_{obs,1}, y_{obs,2}, \ldots, y_{obs,N}]^T,$$

$$ y'(x^*_t) = [y'_1(x^*_t), y'_2(x^*_t), \ldots, y'_N(x^*_t)]^T,$$

$$ y_{obs} = [y_{obs,1}, y_{obs,2}, \ldots, y_{obs,N}]^T,$$

$$ y_b = y(x_b) = [y_1(x_b), y_2(x_b), \ldots, y_N(x_b)]^T,$$

$$ y_i(x) = H_i[M_{t_0 \rightarrow t}(x, \tau)],$$

$$ y_{obs,i} = y_{obs,i} - y_i(x_b), i = 1, \ldots, N,$$

and

$$ y_i(x) = H_i[M_{t_0 \rightarrow t}(x, \tau)].$$
Thus, we have

\[
\tilde{y} = \tilde{y}'(x') = R^{-1}Lx' = R^{-1}
\begin{bmatrix}
L_1 \\
L_2 \\
L_N
\end{bmatrix}x',
\]  

(8)

where \(L_i = H_i'M_{i-1}' \quad i = 1, \ldots, N \), \(H_i'\) is the tangent linear operator of \(H_i\), and \(M_{i-1}'\) is the tangent linear model of \(M_{i-1}\). Therefore, if \(\tilde{y}'\) is similar to \(\tilde{y}\) \(\tilde{y}'\) is a more closely available increment of the IC.

Let us assume that there are linear independent samples of \(\tilde{y}_1', \tilde{y}_2', \ldots, \tilde{y}_m'\) generated by the relation (8) with the IC perturbation samples \(x'_1, x'_2, \ldots, x'_m\). According to ensemble forecasting, if every sample \(\tilde{y}_i' (i = 1, \ldots, m)\) is correlated with \(\tilde{y}\) \(\tilde{y}\) the ensemble mean of these samples can provide a better forecast of \(\tilde{y}'\) as follows:

\[
\tilde{y}' = \alpha_1\tilde{y}_1' + \alpha_2\tilde{y}_2' + \cdots + \alpha_m\tilde{y}_m' = P_x\alpha \quad \text{and} \quad P_y = \begin{bmatrix} \tilde{y}_1', \tilde{y}_2', \ldots, \tilde{y}_m' \end{bmatrix} \quad \alpha = [\alpha_1, \alpha_2, \ldots, \alpha_m]^T,
\]  

(9)

where \(\alpha_1, \alpha_2, \ldots, \alpha_m\) are the weight coefficients. Given the approximately linear relationship between \(\tilde{y}'\) and \(x'\), as indicated in (8), the corresponding IC perturbation \(x'\) can be expressed as the weighted mean of the IC perturbations \(x'_1, x'_2, \ldots, x'_m\). That is,

\[
x' = \alpha_1x'_1 + \alpha_2x'_2 + \cdots + \alpha_mx'_m = P_x\alpha \quad \text{and} \quad P_x = [x'_1, x'_2, \ldots, x'_m].
\]  

(10)

A new concept called DRP–4DVar (Wang et al. 2010) is derived from these analyses by minimizing the cost function as follows:

\[
\begin{aligned}
x_a &= x_b + x'_a = x_b + P_x\alpha_a, \\
\tilde{J}(\alpha_a) &= \min \tilde{J}(\alpha) \\
\tilde{J}(\alpha) &= \frac{1}{2}\alpha^TB^{-1}\alpha + \frac{1}{2}(P_y\alpha - \tilde{y})^T(P_y\alpha - \tilde{y}),
\end{aligned}
\]  

(11)

(12)

The solution to the above minimization problem is derived from the reduced \(m\) dimension; namely,

\[
\alpha_a = (B^{-1} + P_y^TP_y)^{-1}P_y\tilde{y},
\]  

(13)

(14)

where the matrices \(B\) and \(P_y^TP_y\) are of full rank. Here, the background error covariance matrix is estimated by initial perturbation samples (Wang et al. 2010), as in EnKF (Houtekamer et al. 1996; Fisher 1999):

\[
\begin{aligned}
B &= b_m^Tb_m, \\
b_m &= \frac{1}{\sqrt{m-1}}(x_1' - \bar{x}', x_2' - \bar{x}', \ldots, x_m' - \bar{x}'), \\
x' &= \frac{1}{m}(x_1' + x_2' + \cdots + x_m').
\end{aligned}
\]  

(15)

Using the relation (11), \(B\) in the \(m\)-dimension reduced space can be estimated in the same way:

\[
(x')^TB^{-1}x' = \alpha^TP_x^TP_y^Tb_m^Tb_m^T P_x\alpha
= \alpha^TP_x^T(P_xb_m)(P_xb_m)^TP_x\alpha
= \alpha^Tb_m^Tb_m\alpha \propto \alpha^T(B_a)^{-1}\alpha
\]  

(16)
Table 1. Experimental design.

<table>
<thead>
<tr>
<th>Numerical exp</th>
<th>Observation data</th>
<th>Assimilation time window</th>
<th>Model initial conditions (initial time)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nature run (NR)</td>
<td>Simulated SLP data from NR at 1200 UTC 12 Jul 2006</td>
<td>0600 to 1200 UTC</td>
<td>NCEP FNL analysis (0000 UTC 10 Jul 2006)</td>
</tr>
<tr>
<td>CTRL1</td>
<td>Simulated SLP data from NR at 1200 UTC 12 Jul 2006</td>
<td>12 Jul 2006</td>
<td>ECMWF global reanalysis (0600 UTC 12 Jul 2006)</td>
</tr>
<tr>
<td>DRP_OSSE1</td>
<td>Simulated SLP data from NR at 0600 UTC 12 Jul 2006</td>
<td>0600 to 1200 UTC</td>
<td>DRP–4DVar analysis (0600 UTC 12 Jul 2006)</td>
</tr>
<tr>
<td>DRP_OSSE2</td>
<td>Simulated SLP data from NR at 0600 UTC 12 Jul 2006</td>
<td>12 Jul 2006</td>
<td>DRP–4DVar analysis (0600 UTC 12 Jul 2006)</td>
</tr>
<tr>
<td>CTRL2</td>
<td>Simulated SLP data from NR at 1200 UTC 12 Jul 2006</td>
<td>0600 to 1200 UTC</td>
<td>DRP–4DVar analysis (0600 UTC 12 Jul 2006)</td>
</tr>
<tr>
<td>DRP–4DV6H</td>
<td>Simulated SLP data at 1200 UTC 12 Jul 2006</td>
<td>0600 to 0630 UTC</td>
<td>DRP–4DVar analysis (0600 UTC 12 Jul 2006)</td>
</tr>
<tr>
<td>DRP–4DV30m</td>
<td>Simulated SLP data at 0600 UTC 12 Jul 2006</td>
<td>12 Jul 2006</td>
<td>DRP–4DVar analysis (0600 UTC 12 Jul 2006)</td>
</tr>
</tbody>
</table>

Given a set of observations \( y_{\text{obs}} \) and a set of model variables \( x \), the observation operator \( A \) is defined as:

\[
A = \begin{bmatrix} A_1 & \cdots & A_m \end{bmatrix}
\]

where \( A_i : \mathbb{R}^n \rightarrow \mathbb{R}^{m_i} \) is the observation operator for the \( i \)-th observation. The IC perturbation sample \( P \) is constructed as follows:

\[
\begin{align*}
B_\alpha & = b_\alpha^a (b_\alpha^a)^T \\
b_\alpha^a & = \frac{1}{\sqrt{m-1}} \left( \begin{array}{cccc}
1 & 0 & \cdots & 0 \\
0 & 1 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & 1 
\end{array} \right) \\
\alpha & = \frac{1}{m} (\alpha_1 + \alpha_2 + \cdots + \alpha_m) 
\end{align*}
\]

(17)

However, \( b_\alpha^a \) is a singular matrix with a rank of \((m - 1)\). For example, if \( \alpha_1, \alpha_2, \ldots, \alpha_m \) is ideally obtained as \((1, 0, 0, \ldots, 0)^T, (0, 1, 0, \ldots, 0)^T, \ldots, (0, 0, 0, \ldots, 1)^T\), then

\[
b_\alpha^a = \frac{1}{\sqrt{m-1}} \left( \begin{array}{cccc}
1 & -1 & \cdots & -1 \\
-1 & 1 & \cdots & -1 \\
\vdots & \vdots & \ddots & \vdots \\
-1 & -1 & \cdots & 1 
\end{array} \right)_{m \times m} 
\]

(18)

To avoid the dimension reduction of the projection space, a zero IC perturbation sample, \( x_{m+1} = x_b - x_b = 0 \), is introduced to build a \((m + 1)\)-dimension Euclidean space, in which \( b_\alpha^a \) could be estimated with its rank of \( m \) (Wang et al. 2010):

\[
b_\alpha^a = \frac{1}{\sqrt{m}} \left( \begin{array}{cccc}
1 & -1 & \cdots & -1 \\
-1 & 1 & \cdots & -1 \\
\vdots & \vdots & \ddots & \vdots \\
-1 & -1 & \cdots & 1 
\end{array} \right)_{m \times m} 
\]

(19)

Though \( B_\alpha \) is relatively simple, it has good flow-dependent features that have been demonstrated in single-point experiments through comparisons with adjoint-based 4DVar and 3DVar assimilations (Liu et al. 2010).

To evaluate the performance of the DRP–4DVar method for typhoon initialization, we constructed a DRP–4DVar system based on the MM5 model, which is based on the classical MM5 4DVar system developed by Zou et al. (1995). There are two keys to the evaluation of the practicability of the MM5 DRP–4DVar system. One is how to choose the IC perturbation sample \( P \), which requires that

1) the column vectors of \( P \) be linearly independent and
2) all of the column vectors of \( P \) be significantly correlated with \( y_{\text{obs}} \).

Here, the historical time series of consistent forecasts generated by the MM5 are used to generate the IC perturbation samples, which can be obtained without any additional computational cost at an operational numerical weather forecast center (Wang et al. 2010; Zhao and Wang 2010). Another key is how to address the spurious correlations between the locations of the observations and the model grids. These spurious correlations result from the reduced assimilation space, which is composed of far fewer members than both the quantity of observational data and the degrees of freedom of the model variables. Here, we realize the localization using the new approach with the following formula, which is also used in EnKF:

\[
x_a' = \rho^* P_X \alpha_a \\
= \rho^* [P_X (B_\alpha^{-1} + P_Y P_Y^{-1} P_T) y'_{\text{obs}}] \\
= \rho^* [P_X (P_Y B_\alpha^{-1} + P_Y P_Y^{-1} P_T) y'_{\text{obs}}] 
\]

(20)
3. Design of OSSEs

Previous studies using the MM5 indicated that the variational data assimilation techniques using a synthetic SLP field to initialize the tropical cyclone structure are able to improve the forecasts of both the track and intensity of a TC (Xiao et al. 2000; Zou and Xiao, 2000; Pu and Braun 2001; Park and Zou 2004; Zhao et al. 2005; Xiao et al. 2006; Zhao et al. 2007; Zhao et al. 2008). As an initial test of the newly developed MM5 DRP–4DVar system within the context of typhoon SLP information assimilation, OSSEs (Arnold and Dey 1986) are first designed to evaluate the performance of this new assimilation method in the initialization and simulation of typhoons with simulated SLP data. The case studied here is Typhoon Bilis (2006) during the period from 0600 UTC on 12 July to 0000 UTC on 15 July 2006. Bilis began to ravage the Philippines and Taiwan on 12 July and hit China 2 days later. The storm made landfall in Fujian Province on 14 July, and caused prolonged and intense precipitation for 120 h.

In this study, all experiments are performed using the MM5 with a 36-km grid spacing at $124 \times 155$ horizontal dimensions and 15 vertical levels with the top at 100 hPa. The physics packages employed include the Burk–Thompson planetary boundary layer parameterization (Burk and Thompson 1989), Dudhia’s simple ice cloud microphysics scheme (Dudhia 1989) and dry convective adjustment, the Grell cumulus parameterization scheme (Grell 1993), and a cloud radiation cooling scheme for atmospheric radiation (Dudhia 1989). More details on the MM5 can be found in Grell et al. (1994).

The “true” state for our OSSEs is generated by the MM5 model simulation initialized at 0000 UTC on 10 July 2006. It is initialized with an National Centers for Environmental Prediction (NCEP) final (FNL) global tropospheric analyses ($1.0^\circ \times 1.0^\circ$ resolution) and is integrated for 5 days from 0000 UTC on 10 July to 0000 UTC on 15 July. During the period of this case study, the simulation from the high-resolution ICs, known as the nature run (NR hereinafter; Table 1), simulated the 66-h evolution of the central SLP (Fig. 1) and the landfall (Fig. 2) of Typhoon Bilis reasonably well, when compared with the best-track central SLP and the track of Typhoon Bilis from the China Meteorological Administration (CMA; http://www.typhoon.gov.cn/data).

The 90-h evolution of the central sea level pressure (CSLP) (hPa) at 6-h intervals for Typhoon Bilis from 0600 UTC 11 Jul to 0000 UTC 15 Jul 2006. The solid line with asterisks is the observed central SLP (downloaded from http://www.typhoon.gov.cn/data/), and the solid line with dots is the forecast from the NR experiment.

The 66-h track forecast of Typhoon Bilis from 0600 UTC 12 Jul to 0000 UTC 15 Jul 2006. The solid line with typhoon symbols is the best track (observations; downloaded from http://www.typhoon.gov.cn/data/), and the solid line with plus signs is the forecast from the NR experiment. The positions of the storm center are given every 6 h.

Simulated SLP with contour interval of 3 hPa at 1200 UTC 12 Jul 2006 from the NR experiment.
For the OSSEs, our overall goal is not to provide an accurate forecast for the selected typhoon case; the bias in this simulation is acceptable because the NR is used only to provide a four-dimensional dynamically consistent dataset to serve as a “reference dataset.”

In the control run (CTRL1 hereinafter; Table 1), the model was initialized with the 1.5° × 1.5° European Centre for Medium-Range Weather Forecasts (ECMWF) global reanalysis without a variational initialization procedure. This experiment serves as a benchmark for evaluating how the variational initialization technique improves typhoon forecasts. Two assimilation experiments, DRP_OSE1 and DRP_OSE2 (Table 1), were initialized at 0600 UTC on 12 July 2006 using the DRP–4DVar scheme over 6- and 30-min assimilation windows, respectively. The background information is based on the ECMWF global reanalysis, and the assimilated observations of SLP are obtained from the NR simulation.

In DRP_OSE1, the simulated SLP field in the region (11.4°–34.1°N, 115.8°–135.3°E) from the NR simulation at 1200 UTC on 12 July 2006 (Fig. 3) was incorporated at the end of a 6-h window. The IC perturbation sample $P_x$ was chosen from the historical forecast of MM5. First, we ran the MM5 model for a 108-h forecast from 1200 UTC on 10 July 2006 and another 96-h forecast from 0000 UTC on 11 July 2006. The model output was saved every hour. Next, 194 pairs of sampled $x^i$ and $y^i$ were prepared from these two forecasts with a 6-h assimilation window. Quality control and empirical orthogonal function (EOF) decomposition are used to evaluate $y^i$ (Wang et al. 2010). As a result, 22 members of $P_y$ were chosen and 22 corresponding members of the IC perturbation samples $P_x$ could then be generated.

In DRP_OSE2, the simulated SLP field in the region 17.9°–26.6°N, 121.6°–130.4°E from the NR simulation at 0600 UTC on 12 July 2006 was incorporated at the first
of a 30-min window and carried out over a 30-min window every 3 min with the same values, as in Zou and Xiao (2000). The 25 pairs of perturbation samples \( P_X \) and \( P_Y \) were generated using the same method as in DRP_OSSE1.

### 4. Results from OSSEs

The root-mean-square error (RMSE) among the analysis, forecast from CTRL1, and the DRP–4DVar experiments is used to quantitatively assess the quality of the analysis and to assess the efficiency of DRP–4DVar for typhoon initialization. The analyses and forecasts were verified against the true atmospheric state from the NR.

In DRP_OSSE1, the observational cost function shows a reasonable decrease from 1396 to 452, and its RSME decreased from 2.11 hPa in CTRL1 to 1.28 hPa. This result demonstrates that the true SLP information was assimilated by the DRP–4DVar technique. With the incorporation of SLP data at the end of the 6-h assimilation window, the horizontal distribution of the initial SLP showed a slight improvement in the initial location in the DRP_OSSE1 analysis (Fig. 4b1). Figure 5 shows the cross section of the initial temperature increment in OSSE1. Figure 6 shows the vertical profiles of the RMSEs of (a) perturbation pressure \( pp \), (b) zonal wind \( u \), (c) meridional wind \( v \), (d) vertical velocity \( w \), (e) temperature \( T \), and (f) the water vapor mixing ratio \( q \) from CTRL1 (solid line with circles) and DRP_OSSE1 (solid line with dots) at the first assimilation window. These were calculated at all model grid points on each level. The smaller the RMSE of a model variable is, the more similar this variable is to the simulated true state. As can be seen in Fig. 5, the warm core of the modeled typhoon (Fig. 5b) was closer to the true one (Fig. 5a) than that in CTRL1 (Fig. 5c). However, the RMSE of temperature (Fig. 6e) in DRP_4DVar is slightly larger than that in CTRL1, because the locations of the warm core are different. As can be seen in Fig. 6, in the analysis given the IC of DRP_OSSE1, the RMSEs in \( pp \), \( u \), and \( v \) on each model level were almost all reduced. Although the reduction of error in \( pp \) on each model level was less than 1 hPa, the reduction of error in \( v \) was less than 1.5 m s\(^{-1}\), and the reduction of error in \( u \) was less than 0.5 m s\(^{-1}\), slightly better improvement in the horizontal wind speed near the core was presented in DRP–4DVar (Fig. 4b2). Since only the simulated SLP data are incorporated at the end of the 6-h assimilation window, some structural features can only show slight improvement, while other structural features of the initial typhoon vortex (Figs. 6d–f) are worse than those in CTRL1. These results indicate that with SLP information assimilated alone, the MM5 DRP–4DVar system has some ability to recover the model typhoon structure.

The effective performance of DRP–4DVar can also be seen from the RMSE of the simulated typhoon. During the 54-h integration, the RMSE in SLP was all reduced slightly in DRP_OSSE1 (Fig. 7a) compared to that in CTRL1. Furthermore, DRP_OSSE1 had a better integral performance in the SLP forecasts (Fig. 7b). The predicted \( v \) (Fig. 8b) and \( q \) (Fig. 8c) at the lowest model level \( (\sigma = 0.975) \) by DRP_OSSE1 were also closer to the corresponding true values, while the others, like \( pp \)
FIG. 6. Vertical profiles of RMSEs of (a) perturbation pressure (pp, hPa), (b) zonal wind ($u$, m s$^{-1}$), (c) meridional wind ($v$, m s$^{-1}$), (d) vertical velocity ($w$, cm s$^{-1}$), (e) temperature ($T$, °C), and (f) water vapor mixing ratio ($q$, g kg$^{-1}$) of CTRL1 (solid line with circles) and DPR_OSSE1 (solid line with dots) at the first of the assimilation windows ($t = 0$).
Fig. 7. (a) The 54-h variation of RMSEs of central SLP (hPa) predicted by CTRL1 (solid line with circles) and DPR_OSSE1 (solid line with dots). (b) The 60-h variation of CSLP (hPa) of Typhoon Bilis from 0600 UTC 12 Jul to 1800 UTC 14 Jul 2006. The solid line with dots is the forecast from the NR experiment, the solid line with open rectangles is the forecast from the CTRL1 experiment, and the solid line with plus signs is the forecast from the DRP_OSSE1 experiment.

Fig. 8. The 54-h variation of the RMSEs of (a) pressure perturbation (pp, Pa), (b) meridional wind (υ, m s⁻¹), and (c) water vapor mixing ratio (q, g kg⁻¹) at T = 0.975 predicted by CTRL1 (solid line with circles) and DPR_OSSE1 (solid line with dots).

For example, cannot completely perform well during the simulation period. The results imply that the assimilation of SLP information alone has a considerable positive effect on the mass of the wind field. This indicates that with the DRP-4DVar technique, the MM5 is able to simulate a more realistic “model” Typhoon Bilis. However, it is necessary to provide a more accurate simulation for a typhoon using a better assimilation of more observation data.

Figure 9 shows the impact of SLP information at the end of the assimilation window on the typhoon track forecast. The track forecast in DRP_OSSE1 was closer to the true track than that in CTRL1, but it failed to capture the initial location of the typhoon (Fig. 9a). Thus, we constructed the DRP_OSSE2 experiment, which assimilated the SLP field at the first assimilation window, repeatedly incorporating it every 3 min over a 30-min window. The total cost function of these 11 observations decreased from 17 819 to 1052, and the corresponding RSME decreased from 3.03 to 2.67 hPa. This assimilation scheme improved the prediction of the typhoon track (Table 2) and captured the initial location and landfall of the typhoon (Fig. 9b) better than did the DRP_OSSE1 experiment. This demonstrates that DRP-4DVar could be a useful tool for improving the initialization and prediction of typhoons.
5. DRP–4DVar experiments with bogus SLP field

Given the encouraging results seen above, we further tested the newly developed MM5 DRP–4DVar data assimilation system with specified SLP observations; our results are presented in this section. We ran MM5 with the same model configuration and generated IC perturbation samples using the same method as for the OSSEs in section 3. The control run (CTRL2) was integrated for 66 h and initialized using the NCEP FNL analysis at 0600 UTC on 12 July 2006. Experiments DRP–4DV30m and DRP–4DV6H were started from the ICs obtained by the DRP–4DVar scheme over a 30-min or 6-h assimilation window, respectively. The background information was based on the NCEP FNL analysis at 0600 UTC on 12 July 2006, and the assimilated observations of SLP were specified based on Fujita’s (1952) formula, which is expressed as a function of $r$ (radial distance from the cyclone center) as follows:

$$P_{\text{bogus}}(r) = P_c + \Delta P \left(1 - \left[1 + \frac{1}{2} \left(\frac{r}{R} \right)^2\right]^{-1/2}\right), \quad r \leq R_{\text{out}},$$

(21)

$$\Delta P = P_\infty - P_c,$$

and

$$P_\infty = \frac{P_{\text{out}}(R_{\text{out}}) \left[1 + \frac{1}{2} \left(\frac{R_{\text{out}}}{R} \right)^2\right]^{1/2}}{\left[1 + \frac{1}{2} \left(\frac{R_{\text{out}}}{R} \right)^2\right]^{1/2}} - 1,$$

(22)

where $P_c$ is the typhoon’s central pressure, $R$ is the estimated radius of the maximum SLP gradient, and $R_{\text{out}}$ is the radius of the outermost closed isobar $P_{\text{out}}$. The parameters $P_c$, $P_{\text{out}}$, and $R_{\text{out}}$ are specified according to the annual TC observational report. Here, $R$ can be determined from the radius of the 34-kt wind speed ($\sim 17.5 \text{ m s}^{-1}$) (Park and Zou 2004) or can be chosen such that the initial vortex approximates these characteristics of the actual typhoon.

---

![Fig. 9](image_url)

**Fig. 9.** The (a),(b) 60-h track forecast and (c) track error of Typhoon Bilis from 0600 UTC 12 Jul to 1800 UTC 14 Jul 2006. The solid line with dots is the forecast from the NR experiment, the solid line with open squares is the forecast from the CTRL1 experiment, the solid line with plus signs is the forecast from the DRP_OSSE1 experiment in (a), and the solid line with closed squares is the forecast from the DRP_OSSE2 experiment in (b). The positions of the storm center are given every 6 h.
At 0600 UTC on 12 July 2006, Typhoon Bilis was located at 21.3°N, 125.3°E with a central SLP of 985 hPa and a maximum wind speed of 28 m s\(^{-1}\). The values of \(P_{\text{out}}(R_{\text{out}}) = 1002 \text{ hPa and } R_{\text{out}} = 900 \text{ km are estimated based on the NCEP FNL analysis.}\) In addition, \(R\) is chosen to be 200 km. These bogus SLP data for DRP–4DV30m are incorporated at the first assimilation window and repeatedly carried out every 3 min with the same values over the 30-min window from 0600 to 0630 UTC on 12 July.

At 1200 UTC on 12 July 2006, Typhoon Bilis was located at 21.8°N, 124.7°E with a central SLP of 980 hPa and a maximum wind speed of 30 m s\(^{-1}\). The values of \(P_{\text{out}}(R_{\text{out}}) = 1002 \text{ hPa and } R_{\text{out}} = 900 \text{ km are estimated based on the NCEP FNL analysis.}\) Here, \(R\) is chosen to be 200 km. These bogus SLP data for DRP–4DV6H are incorporated at the end of the 6-h window from 0600 to 1200 UTC on 12 July.

Figure 10 shows the model track and its error at 6-h interval for the entire forecast period and is compared with the best track (http://www.typhoon.gov.cn/data/). With a cold start without any observed data being assimilated into the IC, CTRL2 has a mean track error of about 223 km in the 66-h forecast and a delay in landfall with the position error of about 614 km (Fig. 10c). With the addition of bogus SLP information by using DRP–4DVar initialization method, DRP–4DV6H and DRP–4DV30m improve the prediction of the track to various extents compared with the CTRL2 experiment (Table 2).

<table>
<thead>
<tr>
<th></th>
<th>0 h</th>
<th>6 h</th>
<th>12 h</th>
<th>18 h</th>
<th>24 h</th>
<th>30 h</th>
<th>36 h</th>
<th>42 h</th>
<th>48 h</th>
<th>54 h</th>
<th>60 h</th>
</tr>
</thead>
<tbody>
<tr>
<td>CTRL1</td>
<td>165.1</td>
<td>243.9</td>
<td>176.2</td>
<td>305.4</td>
<td>235.6</td>
<td>315.4</td>
<td>212.1</td>
<td>230.0</td>
<td>219.0</td>
<td>188.2</td>
<td>254</td>
</tr>
<tr>
<td>DRP OSSE1</td>
<td>72.7</td>
<td>66.0</td>
<td>108.0</td>
<td>317.1</td>
<td>332.3</td>
<td>213.2</td>
<td>204.1</td>
<td>186.8</td>
<td>260.6</td>
<td>205.2</td>
<td>136.7</td>
</tr>
<tr>
<td>DRP OSSE2</td>
<td>55.6</td>
<td>11.1</td>
<td>121.4</td>
<td>255.7</td>
<td>161.2</td>
<td>111.7</td>
<td>190.5</td>
<td>186.8</td>
<td>157.5</td>
<td>140.5</td>
<td>122.4</td>
</tr>
<tr>
<td>CTRL2</td>
<td>15.6</td>
<td>22.55</td>
<td>74.2</td>
<td>196.4</td>
<td>118.4</td>
<td>86.6</td>
<td>214.5</td>
<td>264.2</td>
<td>324.0</td>
<td>326.6</td>
<td>420.4</td>
</tr>
<tr>
<td>DRP–4DV6H</td>
<td>36.2</td>
<td>61.9</td>
<td>111.8</td>
<td>48.9</td>
<td>165.7</td>
<td>115.5</td>
<td>6.6</td>
<td>73.2</td>
<td>135.1</td>
<td>150.7</td>
<td>181.8</td>
</tr>
<tr>
<td>DRP–4DV30m</td>
<td>41.5</td>
<td>112.7</td>
<td>131.2</td>
<td>124.1</td>
<td>277.1</td>
<td>353.2</td>
<td>318.2</td>
<td>99.1</td>
<td>94.5</td>
<td>70.0</td>
<td>68.8</td>
</tr>
</tbody>
</table>

### 6. Summary and conclusions

Wang et al. (2010) introduced a time-saving approach to 4DVar based on the concept of dimension reduction projection (DRP–4DVar). In this study, we developed a new data assimilation system for MM5 based on the DRP–4DVar method. The MM5 DRP–4DVar data assimilation system, we conducted OSSEs to evaluate the effectiveness of DRP–4DVar scheme for typhoon initialization with a case study, and demonstrated it with specified SLP data assimilations for Typhoon Bilis (2006). Our main results are summarized here.

1) The MM5 DRP–4DVar system assimilated the simulated SLP data into the ICs and maintained the observational information during the assimilation window. After minimizing the cost function with the model constraint, the comparison of the RMSEs in the ICs revealed that the model variables at the lowest model level could be adjusted more. Furthermore, the horizontal wind near the core was improved.

2) The RMSEs of some basic model variables were significantly reduced during the prediction, demonstrating that the new approach can effectively improve the model atmosphere in both space and time.

3) When the typhoon ICs are initialized with SLP data alone using the MM5 DRP–4DVar system, the RMSEs of some model variables cannot perform better at the initial time and during the model integration. This suggests that if both SLP and other information are included in the typhoon initialization schemes, the structure of the typhoon could be further improved.

4) The numerical results indicate that the optimal ICs after incorporating SLP data alone can greatly improve typhoon track forecasts. Compared with the typhoon initialization scheme with the SLP data included at the end of the 6-h window, the scheme with the SLP data included at the first assimilation window can further improve the typhoon track forecasts, especially in terms of the initial location and landfall for the case used in this study.

<table>
<thead>
<tr>
<th></th>
<th>0 h</th>
<th>6 h</th>
<th>12 h</th>
<th>18 h</th>
<th>24 h</th>
<th>30 h</th>
<th>36 h</th>
<th>42 h</th>
<th>48 h</th>
<th>54 h</th>
<th>60 h</th>
</tr>
</thead>
<tbody>
<tr>
<td>CTRL1</td>
<td>165.1</td>
<td>243.9</td>
<td>176.2</td>
<td>305.4</td>
<td>235.6</td>
<td>315.4</td>
<td>212.1</td>
<td>230.0</td>
<td>219.0</td>
<td>188.2</td>
<td>254</td>
</tr>
<tr>
<td>DRP OSSE1</td>
<td>72.7</td>
<td>66.0</td>
<td>108.0</td>
<td>317.1</td>
<td>332.3</td>
<td>213.2</td>
<td>204.1</td>
<td>186.8</td>
<td>260.6</td>
<td>205.2</td>
<td>136.7</td>
</tr>
<tr>
<td>DRP OSSE2</td>
<td>55.6</td>
<td>11.1</td>
<td>121.4</td>
<td>255.7</td>
<td>161.2</td>
<td>111.7</td>
<td>190.5</td>
<td>186.8</td>
<td>157.5</td>
<td>140.5</td>
<td>122.4</td>
</tr>
<tr>
<td>CTRL2</td>
<td>15.6</td>
<td>22.55</td>
<td>74.2</td>
<td>196.4</td>
<td>118.4</td>
<td>86.6</td>
<td>214.5</td>
<td>264.2</td>
<td>324.0</td>
<td>326.6</td>
<td>420.4</td>
</tr>
<tr>
<td>DRP–4DV6H</td>
<td>36.2</td>
<td>61.9</td>
<td>111.8</td>
<td>48.9</td>
<td>165.7</td>
<td>115.5</td>
<td>6.6</td>
<td>73.2</td>
<td>135.1</td>
<td>150.7</td>
<td>181.8</td>
</tr>
<tr>
<td>DRP–4DV30m</td>
<td>41.5</td>
<td>112.7</td>
<td>131.2</td>
<td>124.1</td>
<td>277.1</td>
<td>353.2</td>
<td>318.2</td>
<td>99.1</td>
<td>94.5</td>
<td>70.0</td>
<td>68.8</td>
</tr>
</tbody>
</table>
In summary, we have shown that the DRP–4DVar method is effective for typhoon initialization and track forecast. With the encouraging results presented here, we plan to further investigate the DRP–4DVar assimilation method with satellite data to improve typhoon track and intensity predictions.

Acknowledgments. This study has been supported in part by the National High-tech R&D Program of China for funding the 863 Program (Grant 2010AA012304) and by the Ministry of Finance of China and the China Meteorological Administration for the Special Project of the Meteorological Sector [Grant GYHY (QX) 2009416011].

REFERENCES
Fisher, M., 1999: Background error statistics derived from an ensemble of analyses. ECMWF Research Department Tech. Memo. 79, ECMWF, 1–12.


