

## Research Article

# Impact of Multivariate Background Error Covariance on the WRF-3DVAR Assimilation for the Yellow Sea Fog Modeling

Xiaoyu Gao <sup>1,2</sup> and Shanhong Gao <sup>1</sup>

<sup>1</sup>Key Laboratory of Physical Oceanography, College of Oceanic and Atmospheric Sciences, Ocean University of China, Qingdao 266100, China

<sup>2</sup>Ministry of Education Key Laboratory for Earth System Modeling, Department of Earth System Science and Joint Center for Global Change Studies, Tsinghua University, Beijing 10084, China

Correspondence should be addressed to Shanhong Gao; gaosh@ouc.edu.cn

Received 9 July 2020; Revised 30 September 2020; Accepted 23 October 2020; Published 10 November 2020

Academic Editor: Haydee Salmun

Copyright © 2020 Xiaoyu Gao and Shanhong Gao. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Numerical modeling of sea fog is highly sensitive to initial conditions, especially to moisture in the marine atmospheric boundary layer (MABL). Data assimilation plays a vital role in the improvement of initial MABL moisture for sea fog modeling over the Yellow Sea. In this study, the weather research and forecasting (WRF) model and its three-dimensional variational (3DVAR) data assimilation module are employed for sea fog simulations. Two kinds of background error (BE) covariances with different control variables (CV) used in WRF-3DVAR, that is, CV5 and multivariate BE (CV6), are compared in detail by explorative case studies and a series of application experiments. Statistical verification metrics including probability of detection (POD) and equitable threat scores (ETS) of forecasted sea fog area are computed and compared for simulations with the implementations of CV5 and CV6 in the WRF-3DVAR system. The following is found: (1) there exists a dominant negative correlation between temperature and moisture in CV6 near the sea surface, which makes it possible to improve the initial moisture condition in the MABL by assimilation of observed temperature; (2) in general, the performance of the WRF-3DVAR assimilation with CV6 is distinctly better, and the results of 10 additional sea fog cases clearly suggest that CV6 is more suitable than CV5 for sea fog modeling. Compared to those with CV5, the average POD and ETS of forecasted sea fog area using 3DVAR with CV6 can be improved by 27.6% and 21.0%, respectively.

## 1. Introduction

Sea fog occurs within the marine atmospheric boundary layer (MABL). It is a type of fog that forms over an ocean surface and spreads over coastal areas. It can seriously affect many maritime activities (e.g., transport, navigation, fishery, and rescue) due to its low horizontal visibility [1, 2]. The Yellow Sea has the largest number of foggy days among the four seas of China [3–5], with a long fog season starting from March and ending in September. The Yellow Sea is located to the north of the Kuroshio Current and under influences of warm-moist air mass that is transported northward as the East Asia summer monsoon gradually intensifies. Due to the cold sea surface temperature (SST), an atmospheric inversion layer can easily develop in the MABL over the Yellow

Sea, leading to the formation of advection sea fog in the inversion layer [3]. Therefore, the sensitivity of sea fog modeling to the initial condition is actually largely related to temperature and moisture in the MABL [6–12].

After the textbook *Sea Fog* published in 1985 by Wang [4], Koraćin and Dorman [2] published a book in 2017, in which they summarize the challenges and advancements in observations, modeling, and forecasting of marine fog over the world. Modeling and forecasting have become an important research content of sea fog. In order to mitigate socioeconomic impacts of sea fog over the Yellow Sea and meet increasing practical demands, it is urgent to improve the forecast skill of sea fog. Undoubtedly, a high-quality initial condition is very beneficial to the numerical forecasting for sea fog. A vital way to improve the initial

condition in numerical modeling is data assimilation [13–19]. For the assimilation of given observations, different data assimilation methods yield different assimilation effects. In general, better effect can be obtained when using more advanced data assimilation method. For instance, in their modeling study of sea fog using the weather research and forecasting (WRF) model, Gao et al. [20] demonstrated that ensemble Kalman filter (EnKF) performs better than three-dimensional variational (3DVAR) assimilation, which is attributed to the flow-dependent background error (BE) covariance used in the EnKF. However, the EnKF requires much more computing resources compared to the 3DVAR, which is an issue that has to be considered for real-time forecast of sea fog. The WRF-3DVAR [18] has been widely used for sea fog modeling over the Yellow Sea due to its low computational cost [21–24].

3DVAR minimizes the analysis error variance by minimizing the cost function, which is defined by the weighted sum of analysis increment and observation innovation [25], respectively, with BE and observational error (OE) covariance. A better estimate of BE can definitely lead to a better assimilation analysis. Thus, how to conduct reasonable statistics of BE is an important issue, which has spawned a lot of related research works [26–31]. Currently, the National Meteorology Center (NMC; now called the National Centers for Environmental Prediction, NCEP) method [32] is the most commonly used to generate BE for the 3DVAR. It first conducts transformation of model variables to a set of control variables and then estimates the relationships between different variables by linear regression. The NMC method is implemented in the WRF-3DVAR [18]. There are two types of BE with different control variables (CV): one defines the control variables in the physical space (CV3), and the other defines the control variables in eigenvector space (CV5, CV6, and CV7). CV3 uses a vertical recursive filter to model the vertical covariance, while the others use an empirical orthogonal function (EOF) to represent the vertical covariance. CV3 is a global BE and can be used for any regional domain, while CV5, CV6, and CV7 are domain-dependent and therefore must be generated based on forecast or ensemble data over the same domain. Due to the poor performance of CV3, it is not suitable for sea fog modeling. In CV5, the moisture control variable is not related to any other variables. In CV6, the moisture control variable is the unbalanced portion of the pseudo-relative humidity, and six additional correlation coefficients in the definition of the balanced part of analysis control variables are introduced. Namely, CV6 is a multivariate background error covariance, especially for the moisture. CV7 is similar to CV5, but it uses a different set of control variables and there exist no correlations between the moisture and other control variables.

The issue on the strengths and shortcomings of CV5 and CV6 has been discussed. Based on the work to use multiple linear regressions to estimate multivariate analysis of humidity in a limited-area model, Berre [33] pointed that the relationships between forecast errors of humidity and those of mass and wind fields cannot be negligible. Dhanya and Chandrasekar [34, 35] found that the forecasts of rainfall

caused by monsoon depressions and tropical cyclones can be improved by applying data assimilation with CV6, which is better than CV5. However, among existing research work on sea fog over the Yellow Sea involving numerical modeling [24, 36–40], quite a few studies, especially those works with focus on the mechanism of sea fog evolution, have little or no mention of data assimilation. Even if data assimilation is employed in the numerical simulations for sea fog, CV5 is generally used [22, 23, 41–43]. The reason might be that CV5 is the default option except CV3 in the WRF-3DVAR, which results in the fact that little attention has been paid to CV6 in the application of data assimilation for sea fog modeling.

As stated previously, a successful sea fog modeling definitely requires the assimilation of observations over the sea [22]. Unfortunately, in spite of a large amount of satellite-retrieved data available over the sea, the accuracy of the retrieved water vapor profiles is unsatisfactory. Divakarla et al. [44] performed an evaluation of satellite-retrieved data using global radiosonde observations and found that the root mean square difference is larger than 15% for relative humidity at 1000 hPa. Such big observational errors lead to a malfunction of the retrieved humidity in data assimilation. The approach to correct errors in the observed humidity field and make it more realistic by increments of other control variables is probably an efficient way to modify the MBL moisture structure. Hence, the purpose of the present study is to investigate the impact of multivariate background error covariance on the WRF-3DVAR assimilation for fog modeling over the Yellow Sea. The assimilation effects with CV5 and CV6 are examined.

The remaining content is arranged as follows. Section 2 briefly describes the multivariate BE of the WRF-3DVAR and the data assimilation scheme, and Section 3 outlines numerical experiments, including data, sea fog cases, model configurations, and numerical experiments. Results, analysis, and evaluation are presented in detail in Section 4. Finally, summary and conclusions are provided in Section 5.

## 2. Methodology

*2.1. Multivariate BE.* The cost function for 3DVAR is defined by

$$J(\mathbf{x}) = J_b + J_o = \frac{1}{2} \left[ (\mathbf{x} - \mathbf{x}^b)^T \mathbf{B}^{-1} (\mathbf{x} - \mathbf{x}^b) + (\mathbf{y} - \mathbf{H}\mathbf{x})^T \mathbf{R}^{-1} (\mathbf{y} - \mathbf{H}\mathbf{x}) \right], \quad (1)$$

where  $\mathbf{x}$  and  $\mathbf{x}^b$  are, respectively, the model state and the background field,  $\mathbf{y}$  is the observation, and  $\mathbf{H}$  is the observation operator matrix that transforms data from model space to observation space. In equation (1), analysis increment  $(\mathbf{x} - \mathbf{x}^b)$  and observation innovation  $(\mathbf{y} - \mathbf{H}\mathbf{x})$  are, respectively, weighed by BE ( $\mathbf{B}$ ) and OE ( $\mathbf{R}$ ). OE is determined by instrument errors and representation errors caused by the transformation of data from model space to observation space; BE is determined by model forecast errors. The analysis  $\mathbf{x}^a$  represents a posterior maximum likelihood estimate of the true state of the atmosphere, and it can be obtained if the cost function  $J(\mathbf{x})$  is minimized. The WRF-

3DVAR system uses Quasi-Newton method [45] to realize the minimization process.

Assuming that the model has  $m$  degrees of freedom and the assimilated observation number is  $p$ ,  $\mathbf{x}$  and  $\mathbf{x}^b$  are  $m$ -dimensional vectors,  $\mathbf{y}$  is an  $p$ -dimensional vector,  $\mathbf{H}$  is a  $p \times m$ -dimensional matrix,  $\mathbf{B}$  is a  $m \times m$ -dimensional matrix, and  $\mathbf{R}$  is a  $p \times p$ -dimensional matrix that is diagonal because observational errors are usually supposed to be uncorrelated. In practice, the amount of calculation of  $J_b$  is much larger than that of  $J_o$ , since  $\mathbf{R}$  is diagonal and  $p$  is far less than  $m$ .

Usually, a numerical model has  $m$  degrees of freedom with a typical value of  $10^7$ ; thus it is almost impossible to explicitly solve  $J_b$ , which requires  $\sim O(m^2)$  calculations. However, in the model space, the correlations between different variables or between the same variables at different grid points are available only within a limited range, and thus BE is a sparse matrix. As a result, by decomposing BE, a lot of computational cost can be greatly reduced. This process is known as control variable transform; that is, the model variables  $\mathbf{x}$  are transformed to control variables  $\mathbf{v}$ , and  $J_b$  can then be calculated by  $\mathbf{v}$  instead of by  $\mathbf{x}$ . Let  $\mathbf{x}' = \mathbf{x} - \mathbf{x}^b$ ,  $\mathbf{x}' = \mathbf{U}\mathbf{v}$ ,  $\mathbf{y}' = \mathbf{y} - \mathbf{H}\mathbf{x}^b$ , so  $\mathbf{y} - \mathbf{H}\mathbf{x} = (\mathbf{y} - \mathbf{H}\mathbf{x}^b) - \mathbf{H}(\mathbf{x} - \mathbf{x}^b) = \mathbf{y}' - \mathbf{H}\mathbf{U}\mathbf{v}$ . Note that  $\mathbf{U}$  is the transform matrix defined in a way that satisfies  $\mathbf{B} = \mathbf{U}\mathbf{U}^T$ . Using the increment formulation [46], equation (1) can be rewritten as

$$J(\mathbf{v}) = J_b + J_o = \frac{1}{2} \left[ \mathbf{v}^T \mathbf{v} + (\mathbf{y}' - \mathbf{H}\mathbf{U}\mathbf{v})^T \mathbf{R}^{-1} (\mathbf{y}' - \mathbf{H}\mathbf{U}\mathbf{v}) \right]. \quad (2)$$

The WRF-3DVAR system takes three steps to complete the control variable transform: the horizontal transform  $\mathbf{U}_h$ , the vertical transform  $\mathbf{U}_v$ , and the physical transform  $\mathbf{U}_p$ . Therefore,  $\mathbf{U} = \mathbf{U}_p \mathbf{U}_v \mathbf{U}_h$ .  $\mathbf{U}_h$  uses recursive filter [47, 48] to calculate the correlations between data on different grid points at the same level,  $\mathbf{U}_v$  implements empirical orthogonal function (EOF) decomposition to obtain the correlations between data at different model levels, and  $\mathbf{U}_p$  does linear regression to get the correlations between different variables.

In the physical transform  $\mathbf{x}' = \mathbf{U}_p \mathbf{v}_p$ , the increments of model variables, that is, zonal wind  $u$ , meridional wind  $v$ , temperature  $T$ , pressure  $P$ , and mixing ratio  $Q$ , are first converted to a new set of variables, that is, stream function  $\psi$ , velocity potential  $\chi$ , temperature  $T$ , surface pressure  $P_s$ , and pseudo-relative humidity  $\text{RH}_s$ . These new variables are then divided into balanced and unbalanced components:  $\varphi = \varphi_b + \varphi_u$ , where  $\varphi$  can be any variable. The balanced part  $\varphi_b$  is related to other control variables  $\mu$  by linear regression  $\alpha_{\mu\varphi}$ ; that is,  $\varphi_b = \alpha_{\mu\varphi} \mu$ . The unbalanced part  $\varphi_u$  is independent and serves as input to the vertical transform. By the above procedures, the correlated model variables are transformed into uncorrelated variables.

In the WRF-3DVAR, the correlations are available only within a subset of the variables. CV5 has three basic coefficients:  $\alpha_{\psi\chi}$ ,  $\alpha_{\psi T}$ , and  $\alpha_{\psi P_s}$ , and these coefficients link  $\chi$ ,  $T$ , and  $P_s$  to  $\psi$ :

$$\chi(i, j, k) = \chi_u(i, j, k) + \alpha_{\psi\chi}(k) \psi(i, j, k), \quad (3a)$$

$$T(i, j, k) = T_u(i, j, k) + \sum_{l=1}^n \alpha_{\psi T}(k, l) \psi(i, j, l), \quad (3b)$$

$$P_s(i, j) = P_{su}(i, j) + \sum_{l=1}^n \alpha_{\psi P_s}(l) \psi(i, j, l), \quad (3c)$$

where  $i$  and  $j$  denote the horizontal dimension index,  $k$  and  $l$  indicate the vertical levels,  $n$  is the total number of vertical levels, and the regression coefficients are usually supposed to be horizontally homogeneous.

Compared to CV5, the multivariate BE (CV6) contains six extra coefficients, that is,  $\alpha_{\chi_u T}$ ,  $\alpha_{\chi_u P_s}$ ,  $\alpha_{\psi \text{RH}_s}$ ,  $\alpha_{\chi_u \text{RH}_s}$ ,  $\alpha_{T_u \text{RH}_s}$  and  $\alpha_{P_{su} \text{RH}_s}$ .  $\text{RH}_s$  also consists of balanced and unbalanced components. The control variable transform can be expressed by

$$\chi(i, j, k) = \chi_u(i, j, k) + \alpha_{\psi\chi}(k) \psi(i, j, k), \quad (4a)$$

$$T(i, j, k) = T_u(i, j, k) + \sum_{l=1}^n \alpha_{\psi T}(k, l) \psi(i, j, l) + \sum_{l=1}^n \alpha_{\chi_u T}(k, l) \chi_u(i, j, l), \quad (4b)$$

$$P_s(i, j) = P_{su}(i, j) + \sum_{l=1}^n \alpha_{\psi P_s}(i, j, l) \psi(i, j, l) + \sum_{l=1}^n \alpha_{\chi_u P_s}(l) \chi_u(i, j, l), \quad (4c)$$

$$\text{RH}_s(i, j, k) = \text{RH}_{su}(i, j, k) + \sum_{l=1}^n \alpha_{\psi \text{RH}_s}(k, l) \psi(i, j, l) + \sum_{l=1}^n \alpha_{\chi_u \text{RH}_s}(k, l) \chi_u(i, j, l) + \sum_{l=1}^n \alpha_{T_u \text{RH}_s}(k, l) T_u(i, j, l) + \alpha_{P_{su} \text{RH}_s}(k) P_{su}(i, j). \quad (4d)$$

Formula (4d) indicates that an increment of four other variables can lead to an increment of  $\text{RH}_s$  in CV6. As mentioned in Introduction, moisture information in the initial condition is vital for sea fog modeling. Note that the analysis of moisture is retrieved from the analysis of  $\text{RH}_s$ , background temperature, and pressure together. The moisture increment is directly proportional to  $\text{RH}_s$  increment because background data are recognized as constant.

**2.2. Data Assimilation Scheme.** A cycling-3DVAR scheme [21] is employed to do data assimilation (DA) in this study. The scheme is shown in Figure 1. CV5/6 presents CV5 or CV6, and ‘‘obs’’ stands for observations. Individual assimilation cycles are connected by the WRF integration (i.e., wrf.exe), and the final analysis  $x^a$  is taken as the initial condition for the WRF simulation. In this study, all WRF

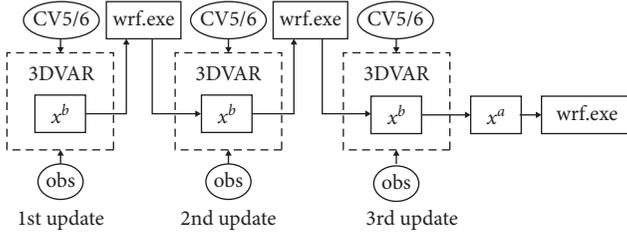


FIGURE 1: Flowchart of the 3DVAR data assimilation scheme. See details in the text.

simulations are designed with two domains in two-way nesting.

As shown in Figure 1, the cycling-3DVAR scheme includes three 3DVAR updates. Each update is only performed on the outer WRF model domain with its BE, and initial condition of the inner domain is obtained by interpolation of analysis field for the outer domain after its 3rd update. In the case study, the data assimilation period is 6 hours with a DA interval of 3 hours.

The NMC method [32] is used to generate CV5 and CV6, assuming that the BEs are well approximated by the averaged differences between 12-hour forecasts and 24-hour forecasts, which are expressed by

$$\mathbf{B} \approx \overline{(\mathbf{x}^{12} - \mathbf{x}^{24})(\mathbf{x}^{12} - \mathbf{x}^{24})^T}, \quad (5)$$

where  $\mathbf{x}^{12}$  and  $\mathbf{x}^{24}$  are 12-hour and 24-hour forecasts, respectively. These forecasts can be obtained from a period of time (e.g., 15–30 days) or can be got from an ensemble prediction system. The former way is adopted here, and thus equation (5) is rewritten as

$$\mathbf{B} \approx \frac{1}{T_n - 1} \sum_{t=1}^{T_n} (\mathbf{x}_t^{24h} - \mathbf{x}_t^{12h})(\mathbf{x}_t^{24h} - \mathbf{x}_t^{12h})^T, \quad (6)$$

where  $T_n$  is the time period for statistics. Here, the forecasts are collected within 15 days centered at the initial day of the case simulation.

### 3. Numerical Experiments

**3.1. Data.** Initial and lateral boundary conditions for numerical modeling were derived from the NCEP Final Analysis (FNL) Data ( $1^\circ \times 1^\circ$ ; 6 hourly). Bottom boundary conditions were extracted from the daily SST dataset ( $0.25^\circ \times 0.25^\circ$ ) of the North-East Asian Regional Global Ocean Observing System (NEAR-GOOS). Observational data for the 3DVAR data assimilation include conventional surface and upper air measurements from the Global Telecommunication System (GTS), buoy and island measurements, and AIRS (Atmospheric Infrared Sounder) retrieved temperature and humidity profiles.

Surface synoptic maps were downloaded from the Korea Meteorological Administration (KMA) to show weather situations of the sea fog cases in the present study. The observed fog areas were empirically retrieved following the approach of Wang et al. [22], which uses the multichannel data of Japanese Multifunctional Transport Satellite

(MTSAT), including albedo, infrared, and visible cloud imageries. The NCEP ship observations in PrepBUFR format, which can be downloaded from the NCAR Computational and Information Systems Laboratory (CISL) Research Data Archive, were used to evaluate model results.

**3.2. Sea Fog Cases.** Two typical sea fog cases over the Yellow Sea, one spreading narrowly over the eastern Yellow Sea along the Korean Peninsula coast (Figure 2(b)) and the other occupying a broad area over the southern Yellow Sea (Figure 3(b)), were selected for case studies. Additionally, extra 10 sea fog cases were chosen for quantitative evaluation on sea fog area forecast with the application of multivariate background error covariance.

The two typical sea fog cases occurred on 31 March 2011 and 10 April 2009 (hereafter referred to as Case-A and Case-B), respectively. Looking at the correspondence between spatial distributions of weather system and fog patches shown in Figures 2 and 3 (cf. Figure 2(a) with Figure 2(b); cf. Figure 3(a) with Figure 3(a)), both cases formed over the areas controlled by a high-pressure system. However, the SST conditions are different for the two cases. Case-A occurred along the west coast of the Korean Peninsula, where the SST was about  $5^\circ\text{C}$  (Figure 2(c)). The fog area was under the control of a weak high-pressure system with the center located at the Korean Peninsula (Figure 2(a)). Warm-moist air mass was transported from the central Yellow Sea to the cold SST area near the coast by the southwesterly wind (Figure 2(c)). Case-B occurred over an ocean area, where a warm SST tongue and a cold SST tongue (Figure 3(c)) coexisted with large SST gradients. Such an SST distribution is usually favorable for the formation of sea fog [49]. In addition, easterly winds developed over the southern Yellow Sea (Figure 3(c)) due to the influence of the high pressure (Figure 3(a)), transporting huge amounts of warm-moist air mass to the fog area.

**3.3. Model and Its Configuration.** The Advanced Research core of the WRF (ARW, version 3.8.1) and its corresponding WRF-3DVAR were employed in the present study for numerical simulations and data assimilations. Based on the previous work related to sea fog simulation over the Yellow Sea fog [22–24, 42], the Yonsei University (YSU) planetary boundary layer (PBL) scheme and the Lin microphysics scheme (LIN) were selected for the present study. Appropriate, vertical levels were specified as well. Detailed model configurations are listed in Table 1. Due to the distinctly different occurrence locations and areas between the two sea fog cases, the model domains and resolutions were separately designed for the two cases (Figure 4). Two-way nesting was implemented in the simulations.

**3.4. Experimental Design.** The experiments in this study were divided into 3 categories: single-observation experiments, case study experiments, and evaluation experiments. The single-observation experiments were conducted based on Case-A, with a focus on RHs in CV6. The case study

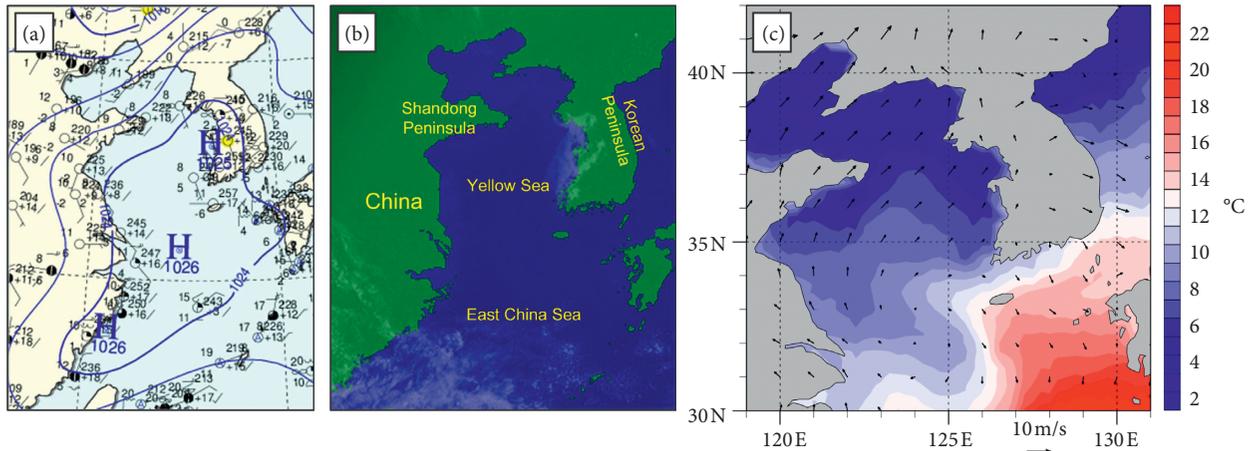


FIGURE 2: Korea Meteorological Administration (KMA) surface synoptic chart (a), Multifunctional Transport Satellite (MTSAT) visible image (b), and 10 m wind vectors over the sea and SST (c) at 00UTC, 31 March 2011.

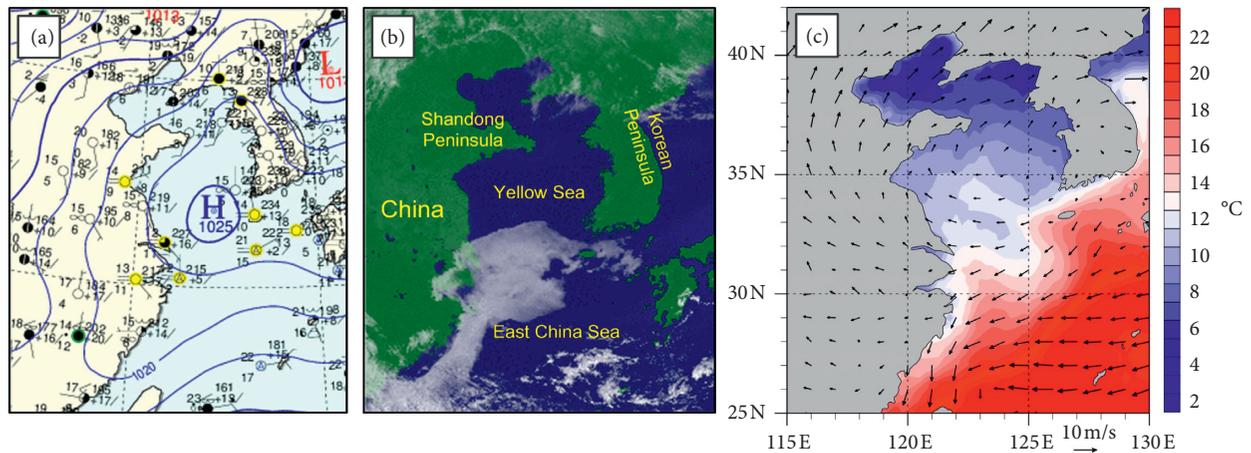


FIGURE 3: Korea Meteorological Administration (KMA) surface synoptic chart (a), Multifunctional Transport Satellite (MTSAT) visible image (b), and 10 m wind vectors over the sea and SST (c) at 00UTC, 10 April 2009.

experiments include six experiments, which were conducted to explore the impacts of CV6 on the 3DVAR assimilation. The evaluation experiments were conducted for the purpose of evaluating the effect of CV6 application.

Table 2 lists the case study experiments for Case-A and Case-B. The initial times of these simulations are presented in Table 1, and all the simulations ran out to 24 hours. The assimilation of observations collected at different platforms may have complicated superposition effects, which is not favorable for clear analysis on the assimilation impact. Therefore, only the GTS routine observations were used for assimilation. For the coastal sea fog case, only observations over the southern Korean Peninsula (Figure 4(a)) were assimilated and the data collected at other far away observational sites were excluded. For the widespread sea fog case, all the GTS observations in D1 were assimilated (Figure 4(b)).

A suite of experiments were conducted for 10 additional sea fog cases that occurred over the Yellow Sea during 2007–2014. In order to produce better initial conditions for the WRF model, the assimilation window was extended from 6 hours for the case study to 12 hours for these cases, and the

observations assimilated include the GTS routine observations, buoy and ship observations and island measurements, and AIRS retrieved data. As with the BEs of those experiments listed in Table 2, CV5 and CV6 were generated in advance separately for each case using the method described in Section 2.2. The model configurations are the same as those used for the widespread sea fog case (see the description for Case-B in Table 1 and the domains in (Figure 4(b))). The experiments consist of two groups: Group-CV5 running with CV5 and Group-CV6 running with CV6, respectively.

## 4. Results

### 4.1. Single-Observation Experiments

4.1.1. Relationship between  $RH_s$  and  $T_u$ . As mentioned in Introduction, there is no relationship between moisture control variable and any other variable in CV5. However, such relationships exist in CV6. Equation (4a) shows that, in CV6,  $RH_s$  is related to four other variables, and the

TABLE 1: Configurations of the weather research and forecasting (WRF).

Model setting and option		Specification	
		Case-A	Case-B
Domain	Map projection	Lambert	
	Central point	(37.0°N, 125.0°E)	(34.2°N, 124.1°E)
	Grid number	80 × 80 for D1, 151 × 151 for D2	166 × 190 for D1, 120 × 120 for D2
	Horizontal resolution	30 km for D1, 6 km for D2	30 km for D1, 10 km for D2
	Vertical grid	44 $\eta^*$ with a pressure top at 50 hPa	
Initial time	00UTC, 31 March 10 2011		12UTC, 9 April 2009
Time step	Adaptive time step (60–120s for D1)		
PBL scheme	YSU scheme [50]		
Cumulus parameterization	Kain-Fritsch scheme [51, 52]		
Microphysics	Lin (Perdue) scheme [53]		
Long-shortwave radiation	RRTMG scheme [54]		
Land surface model	Noah [55]		

\* $\eta = 1.0000, 0.9975, 0.9925, 0.9850, 0.9775, 0.9700, 0.9540, 0.9340, 0.9090, 0.8800, 0.8506, 0.8212, 0.7918, 0.7625, 0.7084, 0.6573, 0.6090, 0.5634, 0.5204, 0.4798, 0.4415, 0.4055, 0.3716, 0.3397, 0.3097, 0.2815, 0.2551, 0.2303, 0.2071, 0.1854, 0.1651, 0.1461, 0.1284, 0.1118, 0.0965, 0.0822, 0.0689, 0.0566, 0.0452, 0.0346, 0.0249, 0.0159, 0.0076, \text{ and } 0.0000.$

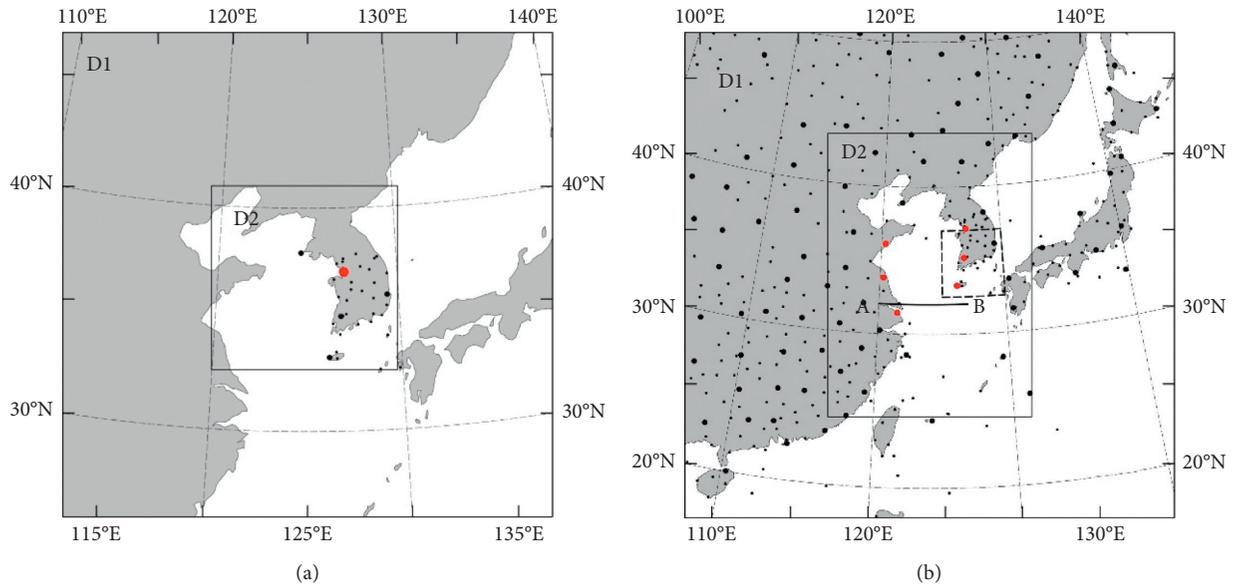


FIGURE 4: Geographic map of the two nesting domains of the weather research and forecasting (WRF) simulations for Case-A (a) and Case-B (b). Sites of surface (small dots) and upper-air (large dots) observations are indicated. The red dot in (a) is the observation location for single-observation experiments in Section 4.1, and the red dots in (b) are locations of upper-air observations for verification of Case-B.

TABLE 2: Design of experiments.

Experiment	Fog case	BE type	Assimilated observations
Exp-A_CV5	Case-A	CV5	Radiosonde and surface measurements over the southern Korean Peninsula
Exp-A_CV6		CV6	
Exp-B_CV5		CV5	
Exp-B_CV6	Case-B	CV6	Radiosonde and surface measurements
Exp-B_CV5e		CV5	Same as Exp-B_CV5/6 but excluding the observations over the southern Korean Peninsula
Exp-B_CV6e		CV6	

relationships between them are decided by the regression coefficients. In order to investigate how close these variables are related, contribution rates of other control variables to  $RH_s$  were calculated from CV6 generated in advance via the NMC method for Case-A. The results are shown in

Figure 5(a). At low levels of the model, the balanced component accounts for 30–40% of the total  $RH_s$ , among which the term  $T_u$  contributes up to 20–30% of the total. Other individual terms can only make a contribution of less than 5%. As a result, the term  $T_u$  accounts for more than

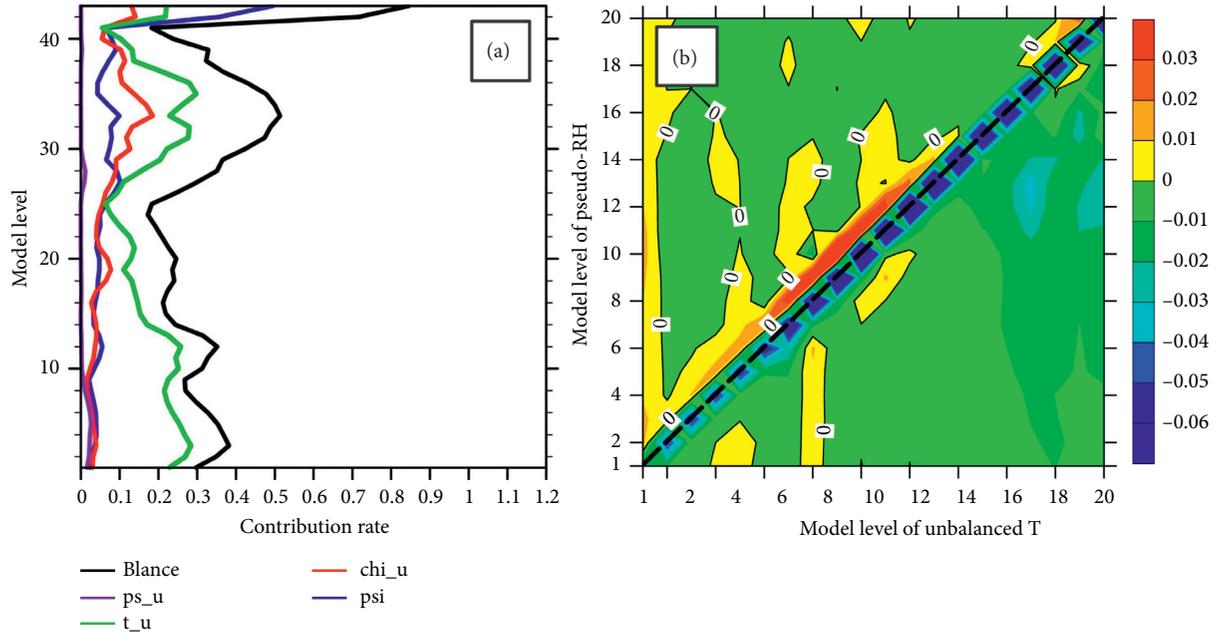


FIGURE 5: Contribution rates of individual terms to  $RH_s$  at different model levels (a) and regression coefficients of  $T_u$  (b) at vertical model level (horizontal axis) for  $RH_s$  calculation (vertical axis). Thick-dashed line in (b) indicates the regression coefficients at the same level.

~70% of the balanced component of  $RH_s$ , indicating that the temperature increment is much more important for the humidity increment than the wind and pressure increments. Similar results are also found for Case-B (figure not shown).

Figure 5(b) presents the distribution of regression coefficient  $\alpha_{T_u, RH_s}(k_1, k_2)$  of  $T_u$  at level  $k_2$  to calculate  $RH_s$  at level  $k_1$  (see equation (4a)). The horizontal/vertical axis stands for  $k_2/k_1$ . The coefficients along the diagonal of Figure 5(b) are all negative, and their absolute values are much larger than other data on the same row. As a result, the humidity increment at any specific point is basically decided by temperature increment at the same point, and a negative increment of  $T_u$  always causes a positive increment of  $RH_s$ .

**4.1.2. Contribution Rate of  $T_u$  to  $RH_s$ .** A series of single-observation experiments were designed based on Case-A with CV5/CV6, and the single observation is marked by the red dot shown in Figure 4(a). Each experiment was conducted at different vertical model level with a temperature observation of 1 K lower than the background. These experiments are denoted by Exp-Sk, where  $k$  (from 1 to 20) stands for the vertical level number.

Since  $\alpha_{T_u, RH_s}(1, 1)$  is negative, decrease of temperature with a maximum value of about  $-0.3$  K (Figure 6(a)) produces an increase of mixing ratio with a maximum value of about  $0.06$  g/kg (Figure 6(c)). The maximum values of temperature and mixing ratio increments at levels 1–20 from Exp-S1–20 are illustrated in Figures 6(b) and 6(d). Through vertical transform, the negative temperature increment spreads to several vertical levels and causes increases of mixing ratio there. Figure 6(d) has a very similar pattern to that of Figure 6(b), which again shows that the humidity change mainly results from the change of temperature at the

same level, though there exist some slight influences from neighboring levels. Compared with CV6, although similar patterns and a little bit weaker intensities of temperature patterns and a little bit weaker intensities of temperature assimilation result are obtained for CV5 (cf. Figure 6(e) with Figure 6(f) and Figure 6(a) with Figure 6(b), respectively), there is not obviously any gain for mixing ratio (Figures 6(g) and 6(h)) due to the fact that there is no relationship between moisture control variable and any other variable.

Overall, for CV6, regression coefficients  $\alpha_{T_u, RH_s}$  of the same level are all negative and their absolute values are usually much larger than those of neighboring levels, leading to an increase of mixing ratio when a lower temperature observation is assimilated. For 3DVAR with CV5, assimilation of temperature observations cannot change humidity analysis, since  $RH_s$  is an independent variable in CV5.

## 4.2. Case Study 1: The Coastal Sea Fog

**4.2.1. Verification.** Simulated sea fog is defined to those areas where the following criteria are met [22]: cloud liquid water (CLW) at the model's lowest level is  $\geq 0.016$  g/kg or cloud top is  $\leq 400$  m.  $CLW \geq 0.016$  g/kg is equivalent to a visibility  $\leq 1$  km. The cloud top is calculated based on the threshold of  $CLW \geq 0.016$  g/kg; and observations indicate that advection sea fogs are deeper than other types of fog but rarely exceed 400 m [56].

The fog occurred over a narrow ocean area along the western coast of the Korean Peninsula (Figures 7(a)–7(e)). As mentioned in Section 3.4, only observations over the southern Korean Peninsula Figure 4(a) were assimilated, because we want to explore the impact of assimilating observations that are close to the fog area. The fog areas forecasted by Exp-A\_CV5 and Exp-A\_CV6 are shown in

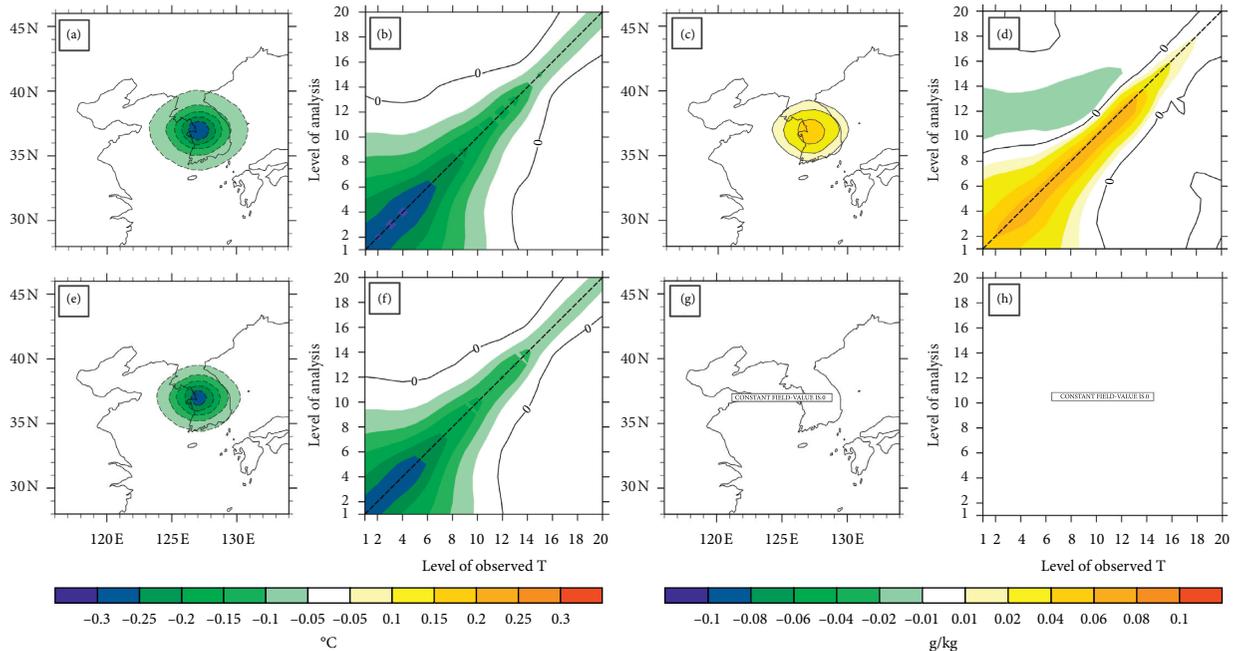


FIGURE 6: Analysis increments in the single-observation experiments for CV6 (upper row) and CV5 (lower row): horizontal distribution of analysis increments of temperature (a, e) and mixing ratio (c, g) at the bottom model level when the single observation is at the same level; analysis increments of temperature (b, f) and mixing ratio (d, h) at individual levels (vertical axis) when the single observation is specified at each level (horizontal axis).

Figure 7. It is apparent that Exp-A\_CV6 performs better than Exp-A\_CV5, because the former can well capture the sea fog evolution, while the latter completely fails to reproduce the sea fog (cf. Figures 7(f)–7(j) and 7(k)–7(o) with 7(a)–7(e), respectively). Note that the difference of 925 hPa wind speeds between these two experiments is less than 1 m/s (figure not shown), which indicates that CV6 makes a trivial change for wind field.

The agreement between the initial conditions and the radiosonde observations (see the big dots in Figure 4(a) for their locations) over the Korean Peninsula was examined. Figure 8 shows the average vertical profiles of bias and root mean square error (RMSE) of temperature and mixing ratio for Exp-A\_CV5 and Exp-A\_CV6. Below the height of 600 m, the initial temperatures of both experiments are higher than observations (Figure 8(a)), while most of the initial mixing ratios are smaller than observations (Figure 8(b)). Nevertheless, the mixing ratio bias of Exp-A\_CV6 is much smaller than that of Exp-A\_CV5. The same is true for the RMSE. It means that the initial moisture condition of Exp-A\_CV6 is much wetter than that of Exp-A\_CV5, which results in the better forecasted sea fog area by Exp-A\_CV6.

In addition, ship measurements (not assimilated; see their sites in Figure 9(a)) were used to verify the moisture condition near the sea surface. Biases between model output and the measurements were calculated and results are shown in Figure 9(b). It is obvious that the initial moisture conditions of both Exp-A\_CV5 and Exp-A\_CV6 are drier than the observed ones. However, the dry condition is significantly improved in Exp-A\_CV6 compared to that in Exp-A\_CV5. Note that the sites S6, S7, and S8 are close to the

Korean Peninsula. Compared to Exp-A\_CV5, the improvements at these three sites in Exp-A\_CV6 are, respectively, 48.7%, 76.7%, and 100%, and the corresponding biases are reduced by 0.39, 0.23, and 0.47 g/kg, respectively.

**4.2.2. Analysis Increments.** Exp-A\_CV5 performs poorly. As shown in Figures 7(f)–7(j), it produces almost no sea fog during the whole forecast period. In contrast, Exp-A\_CV6 successfully simulates the sea fog event (Figures 7(k)–7(o)). From Figure 9, it is found that the initial condition of Exp-A\_CV5 is quite dry. For example, the moisture bias at the site of S6 is about  $-0.8$  g/kg. Note that this dry bias is corrected in Exp-A\_CV6 and the bias is reduced to 0.4 g/kg. The reason why Exp-A\_CV6 can gain such a significant moisture improvement is discussed below.

During the analysis update of the WRF-3DVAR, the minimization of the cost function (equation (1)) produces the analysis field  $\mathbf{x}^a$ .  $\mathbf{x}^a$  is determined by the balance of analysis increment ( $\mathbf{x}-\mathbf{x}^b$ ) and observation innovation ( $\mathbf{y}-\mathbf{H}\mathbf{x}$ ) under the joint constraint of OE and BE. Because OE is usually prescribed regardless of a specific sea fog case, BE plays the key role. There are three updates for  $\mathbf{x}^a$  in the designed data assimilation scheme (Figure 1), and the 3rd update yields the initial condition for the experiments. Figure 10 demonstrates analysis fields of temperature, wind, and mixing ratio at the lowest model level for three DA updates in Exp-A\_CV5 (upper row) and Exp-A\_CV6 (lower row). It is clearly seen that the analysis fields of both temperature and wind are almost the same between Exp-A\_CV5 and Exp-A\_CV6. The analysis fields of mixing ratio

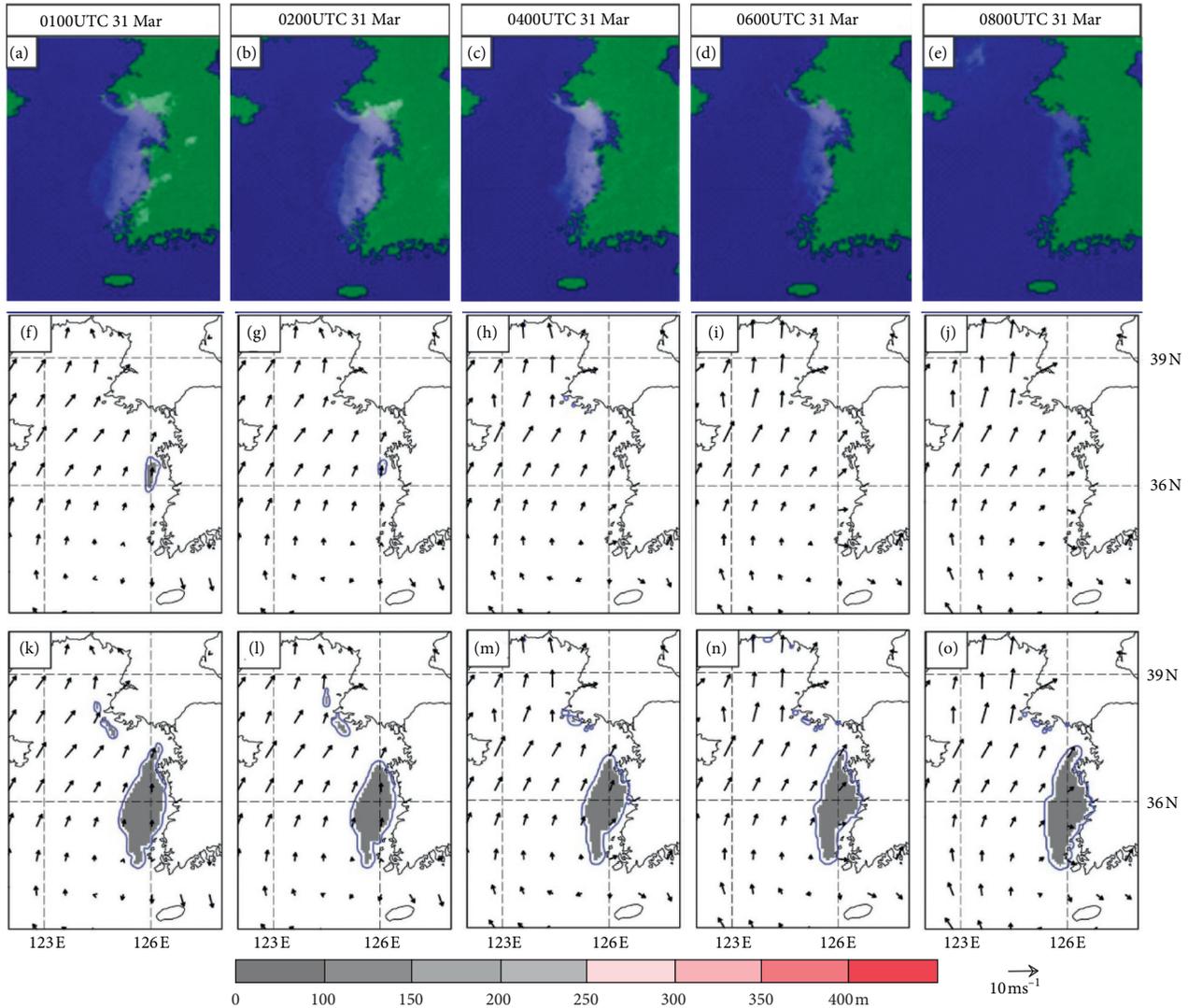


FIGURE 7: Comparison between the forecasted and observed sea fog areas (upper row; MTSAT visible imageries) for Case-A. The middle and lower rows show the results of Exp-A\_CV5 and Exp-A\_CV6, respectively. Fog depth is shown by shadings, and surface winds are displayed by vectors.

are obviously different, especially at the 3rd update (cf. Figure 10(c) with Figure 10(f); see bold lines with the value of 5.5 g/kg). Water vapor is much less near surface over sea area west of Korean Peninsula in Exp-A\_CV5 than that in Exp-A\_CV6, which leads to its failure. Analysis fields at upper levels are also checked and significant difference of mixing ratio between Exp-A\_CV5 and Exp-A\_CV6 exists below ~600 m (not shown). Next, we try to investigate in detail the difference between the observation innovations of Exp-A\_CV5 and Exp-A\_CV6 at the 3rd update, as well as the consequence of the difference on analysis increment.

Figure 11 illustrates the observation innovations and analysis increments at the lowest model level for Exp-A\_CV5 and Exp-A\_CV6. For the temperature innovation, there is nearly no difference between Exp-A\_CV5 and Exp-A\_CV6 (cf. Figure 11(a) with Figure 11(e)). Meanwhile, the difference in moisture innovation is very small, and the moisture innovation in the area denoted by a red-

dashed frame (Figure 11(f)) is a little bit larger in Exp-A\_CV6 than in Exp-A\_CV5 (Figure 11(b)). However, Exp-A\_CV6 obtains much more moisture increment than Exp-A\_CV5 (cf. Figure 11(c) with Figure 11(g)). Does this moisture increment still result from temperature innovation through the way revealed in the single-observation experiments of Section 4.1? To answer this question, two extra experiments were conducted, and the value of  $T_u$  (see equations (3a)–(3c) and (4a)–(4d)) is set to 0 in the 3rd cycle of the WRF-3DVAR; namely, the temperature information is not used in the assimilation process. Results of the two extra experiments show that the difference in moisture increment between Exp-A\_CV6 and Exp-A\_CV5 is very small (cf. Figure 11(d) with Figure 11(h)), which confirms that the larger moisture increment shown in Figure 11(g) is due to the negative temperature innovation shown in Figure 11(e) by the function of CV6 in Exp-A\_CV6. Note that the area of positive moisture

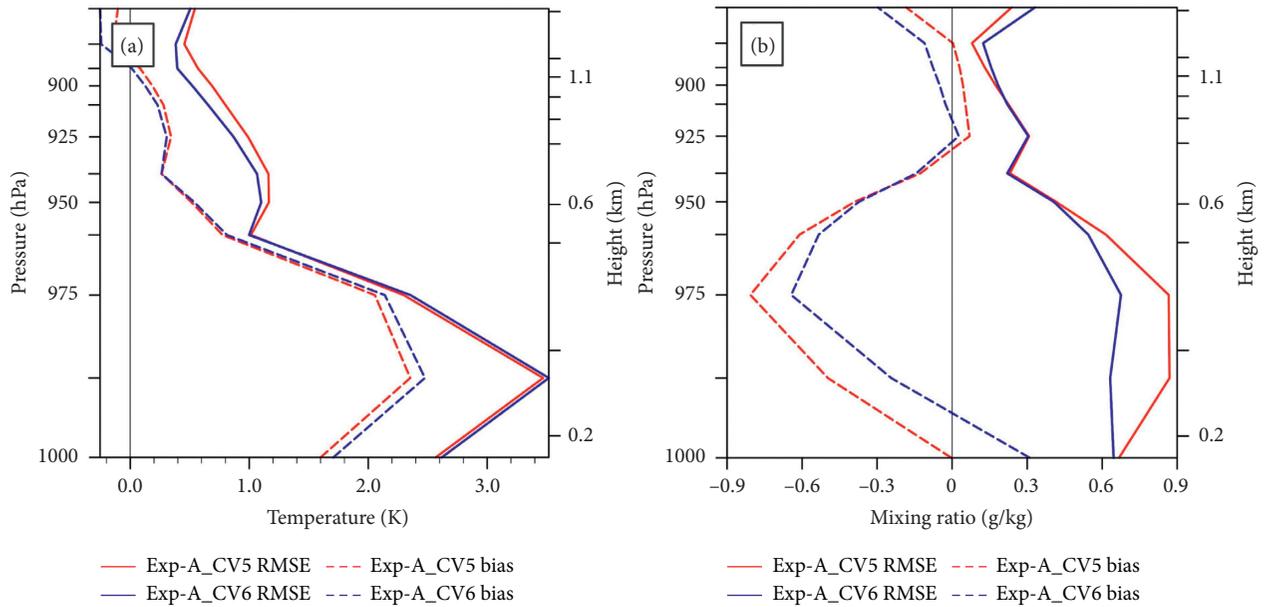


FIGURE 8: Vertical profiles of root mean square errors (RMSEs) (solid lines) and biases (dashed lines) between the initial analysis and the radiosonde observations of temperature (a) and mixing ratio (b) for Exp-A\_CV5 and Exp-A\_CV6.

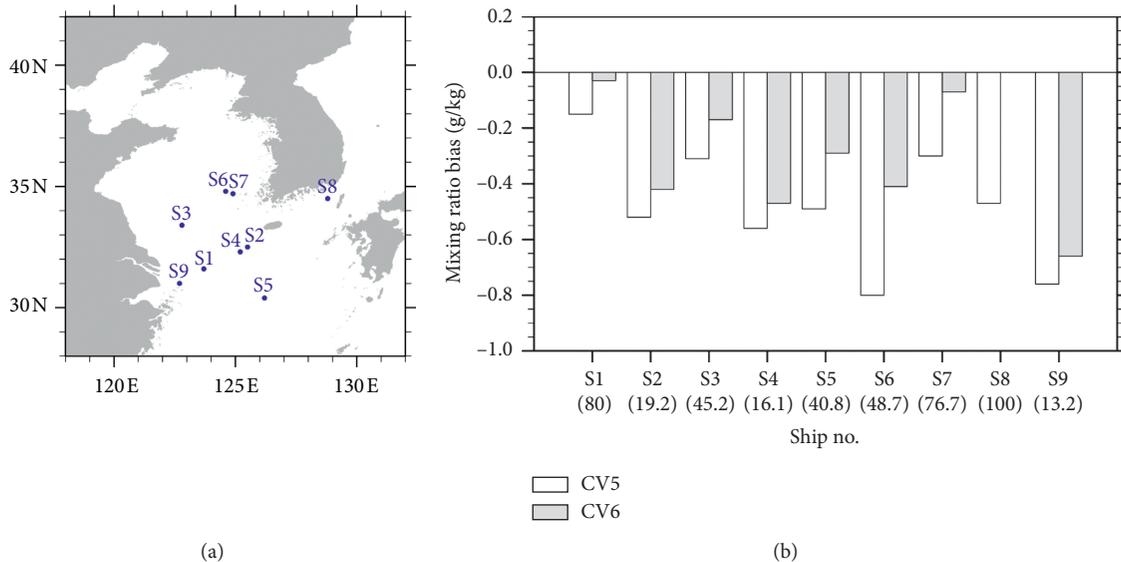


FIGURE 9: Locations of ship measurements (a) and biases of forecasted mixing ratio compared with observation (b) for Case-A. The improvements (%) of Exp-A\_CV6 compared to Exp-A\_CV5 are presented below the S# (identified mark of ship sites) along horizontal coordinates.

increments in Figure 11(g) only covers the fog area (see Figure 7). Therefore, the success of Exp-A\_CV6 is undoubtedly the consequence of the moisture improvement in its initial condition, and CV6 is vital to guarantee this improvement.

### 4.3. Case Study 2: The Widespread Sea Fog

4.3.1. Verification. The results of forecasted sea fog area for Exp-B\_CV5 and Exp-B\_CV6 are shown in Figure 12. Similar to the results of Case-A, the sea surface wind

fields of the two experiments are almost the same (cf. Figures 12(f)–12(j) with 12(k)–12(o), respectively). Although the two experiments both underestimate the sea fog areas, it is clear that Exp-B\_CV6 still performs better than Exp-B\_CV5. Compared with that produced by Exp-B\_CV5, the fog area produced by Exp-B\_CV6 is much closer to the observed one (cf. Figures 12(f)–12(j) and 12(k)–12(o) with 12(a)–12(e), respectively). Similar to the coastal sea fog, the difference in winds between the two experiments in the MABL is quite small (figure not shown).

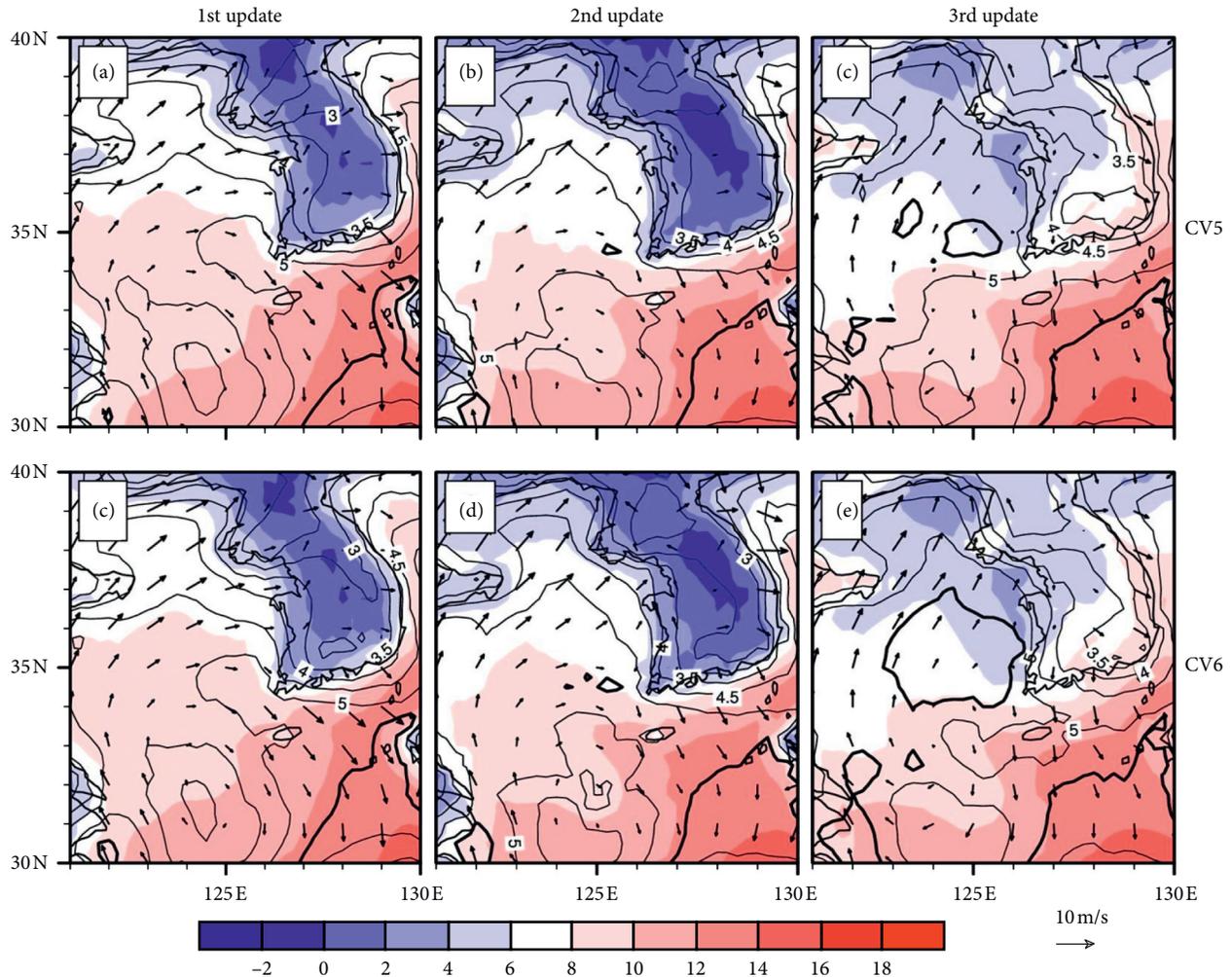


FIGURE 10: Analysis fields of temperature (color shades), wind (vectors), and mixing ratio (contours) at the lowest model level after three DA updates: (a)–(c) for CV5 and (d)–(f) for CV6. The bold lines indicate mixing ratio of 5.5 g/kg.

The temperature and moisture structures of the MABL in the initial conditions of Exp-B\_CV5 and Exp-B\_CV6 were compared with six radiosonde soundings (red dots in Figure 4(b)). Vertical profiles of RMSE and bias of temperature and mixing ratio for Exp-A\_CV5 and Exp-A\_CV6 are presented in Figure 13. Similar to Figure 8, the initial temperatures of both experiments are higher than observations (Figure 13(a)). However, the initial mixing ratios are smaller than observations below 300 m see the biases in (Figure 13(b)). Compared to Exp-B\_CV5, the drier condition is alleviated by 0.2 g/kg in Exp-B\_CV6, which results in better performance. Ship measurements (not assimilated; see their sites in (Figure 14(a))) were also employed to check the moisture condition near the sea surface. The biases between model outputs and the measurements are displayed in (Figure 14(b)), which shows clearly that Exp-B\_CV6 outperforms Exp-B\_CV5 at six out of eight sites and the maximum improvement is up to 36.7%.

**4.3.2. Differences in Initial Condition.** Since Exp-B\_CV5 and Exp-B\_CV6 share the same model configuration, their different performances are attributed to differences between

their initial conditions. Figure 15 shows the differences in initial condition near the surface between the two experiments (Exp-B\_CV6 minus Exp-B\_CV5). The differences in pressure and wind are not presented because they are very small. The difference in moisture is relatively larger than that in temperature. The differences in moisture and temperature over the sea fog area (see the sea fog in Figures 12(a), 12(f), and 12(k)) are quite small. For example, the difference in temperature is within the range of about  $-0.2$ – $+0.2$ °C, and the differences in mixing ratio and relative humidity are within the ranges of about  $-0.2$ – $+0.2$  g/kg and 0–5%, respectively. Note that there are two areas with large positive and negative differences in mixing ratio, respectively (Figure 15(b)). The area of positive difference is located over the northwestern Yellow Sea (see the red shading around Shandong Peninsula marked in Figure 2(b)), and the area of negative difference is located over the southern Korea Peninsula. The area of positive mixing ratio difference stretches to the southeast, while positive difference in relative humidity occurs over almost the entire Yellow Sea (Figure 15(c)). Different from the positive analysis increment of mixing ratio over the Yellow Sea in Exp-B\_CV6,

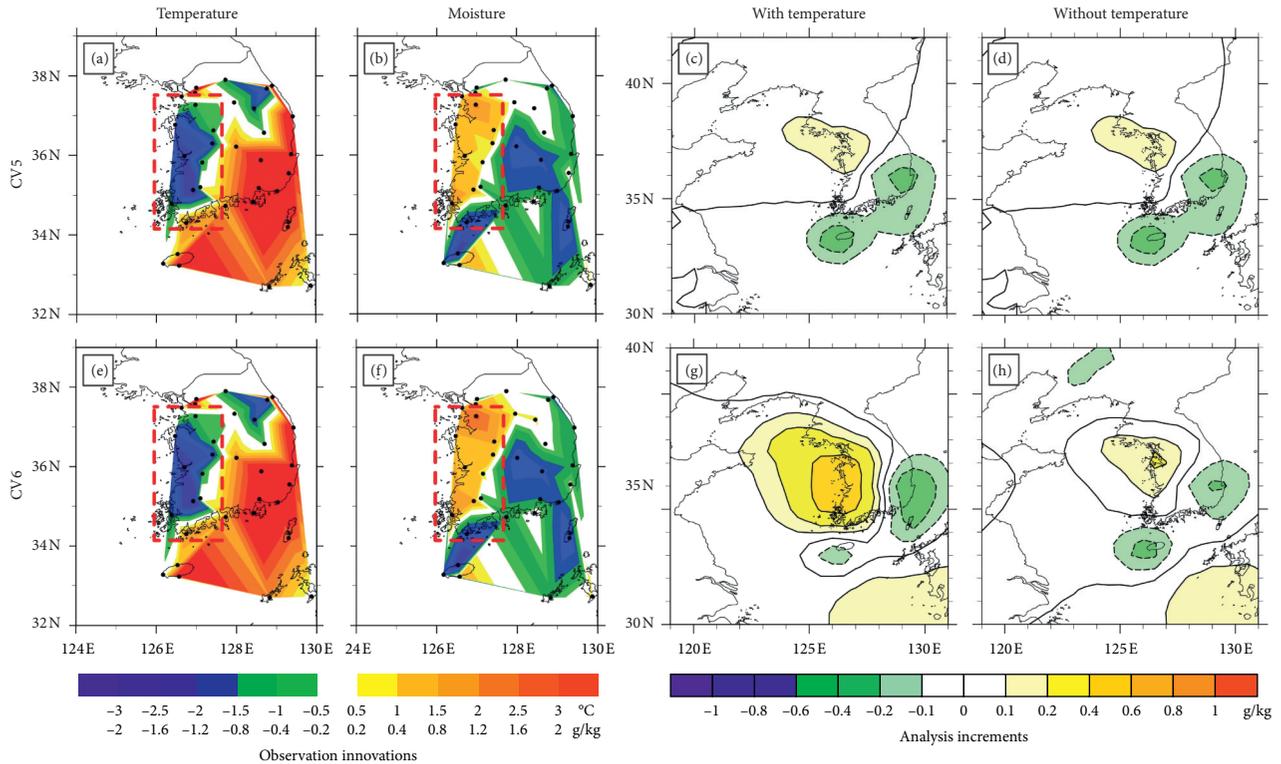


FIGURE 11: Observation innovations and analysis increments of the 3rd update at the lowest model level for CV5 (upper row) and CV6 (lower row): innovations of temperature (a, e) and mixing ratio (b, f); increments of mixing ratio with (c, g) and without (d, h) assimilation of temperature. The dots indicate the observational sites.

there is almost no positive gain of mixing ratio in Exp-B\_CV5 (cf. Figures 16(a) and 16(b)). The above results demonstrate that the initial condition of Exp-B\_CV6 is wetter than that of Exp-B\_CV5 in the MABL, which leads to better performance.

For Case-A described in Section 4.2, the assimilation of the observations over the Korea Peninsula (see their sites in Figure 4(a)) can noticeably improve the modeling result of Exp-A\_CV6, because positive analysis increment of moisture is produced by negative temperature observation innovation (see Figure 11). On the contrary, for Case-B, both Exp-B\_CV5 and Exp-B\_CV6 get negative analysis increments of moisture over Area-K (shown in Figure 16), and the increment for Exp-B\_CV6 is larger than that for Exp-B\_CV5. This can be explained by the observation innovations of moisture and temperature over Area-K. The moisture innovation is negative and the temperature innovation is positive over Area-K for both Exp-B\_CV5 and Exp-B\_CV6 (figures not shown). For Exp-B\_CV5, the negative increment of moisture is produced because  $RH_s$  is an independent variable in CV5. For Exp-B\_CV6, however, since there exist correlations between  $RH_s$  and other variables, especially temperature (equation (4a)), negative moisture increment and positive temperature increment jointly produce a large negative increment of moisture over Area-K. In contrast, Exp-B\_CV6 gets positive moisture increment (see the yellow shading in Figure 16(b)) over other areas of the

Yellow Sea due to the assimilation of other observations beyond Area-K.

Forced by the sea surface wind field (see the vectors in Figure 12), the southwesterly advection pulls drier air mass from Area-K to the southern Yellow Sea, which is not conducive to the formation and development of sea fog there. Despite this unfavorable situation, Exp-B\_CV6 still performs better than Exp-B\_CV5, which is attributed to the fact that its initial condition contains positive moisture increment over the Yellow Sea caused by the correlation between moisture and other control variables in CV6. To further explore the impact of CV6, two sensitivity experiments—Exp-B\_CV5e and Exp-B\_CV6e—were, respectively, conducted based on Exp-B\_CV5 and Exp-B\_CV6 but without assimilation of the observations over the southern Korea Peninsula (the domain framed by dashed line in Figure 4(b)). Figure 17 compares the analysis increments of mixing ratio near the surface between Exp-B\_CV5 and Exp-B\_CV5e and initial difference in mixing ratio between them. It can be seen clearly that the assimilation of the observations over the southern Korea Peninsula results in an obvious decrease of mixing ratio in the MABL in the initial condition. Comparative results of the forecasted sea fog area between Exp-B\_CV5/6 and Exp-B\_CV5e/6e are shown in Figure 18. Compared to Exp-B\_CV5, the fog area simulated by Exp-B\_CV5e is significantly improved by 29.0–45.4%. However, the difference between results of Exp-B\_CV6e and Exp-B\_CV6 is no more than ~5%. The above results indicate that the assimilation effect of CV6 is more reliable than CV5.

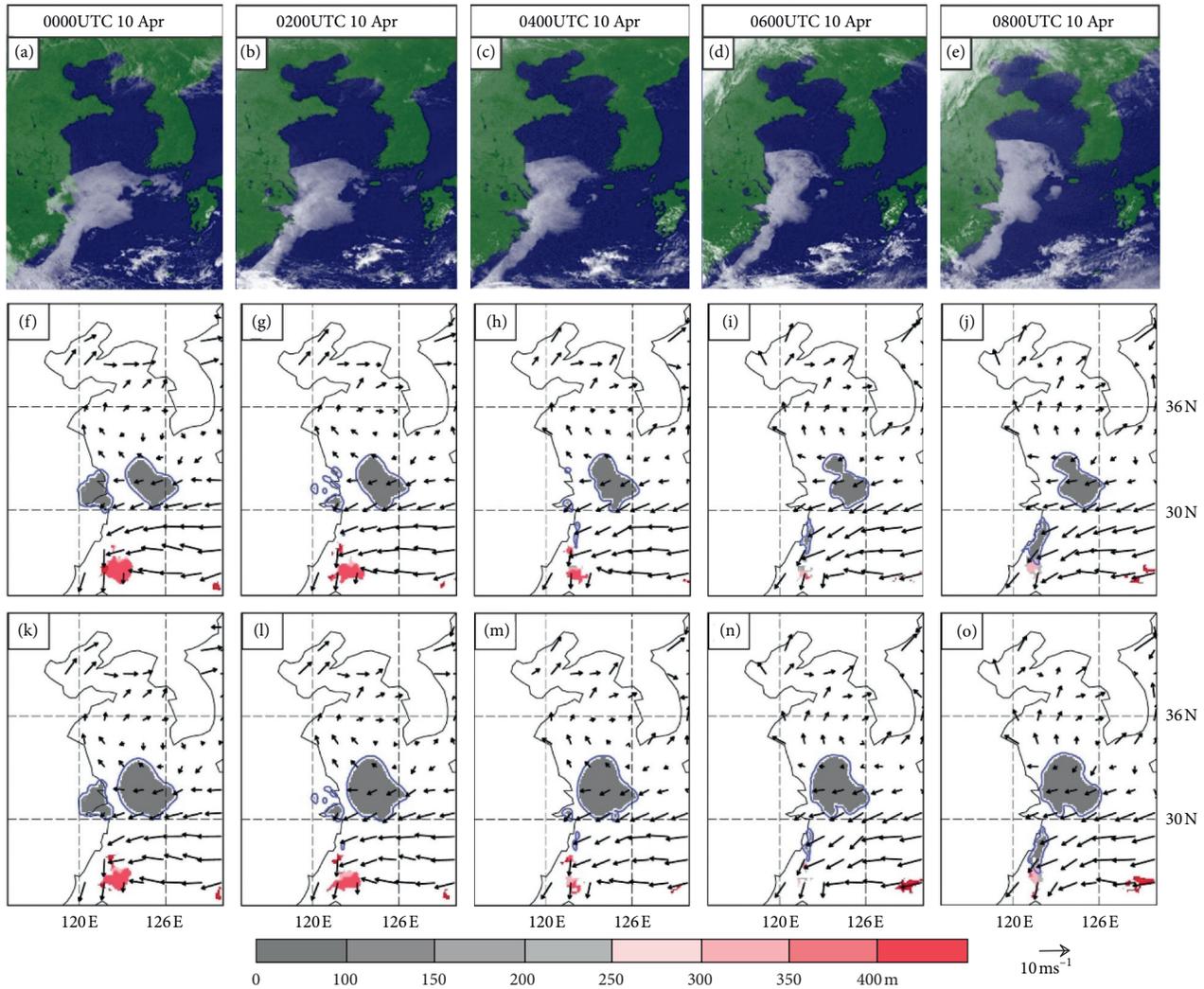


FIGURE 12: Comparison between the forecasted and observed sea fog areas (upper row; MTSAT visible imageries) for Case-B. The middle and lower rows show the results of Exp-B\_CV5 and Exp-B\_CV6, respectively. Fog depth is shown by shadings, and surface winds are displayed by vectors.

4.4. Discussion on the Impact of CV6. The results of the above case studies show that CV6 outstrips CV5 in the both cases. It seems that the performance of CV6 is much better in Case-A than in Case-B (see Figures 7 and 12). In Case-A, the sea fog patch is small and very near the western coast of the Korean Peninsula. The assimilation of temperature observations over the Korean Peninsula can affect the area where the sea fog occurred (Figure 11(g)), which leads to relatively great improvement in Case-A. In contrast, the sea fog patch in Case-B is much larger than that in Case-A, and most of the sea fog patch in Case-B is far away from the coast; the assimilation information of observations over land is difficult to spread over the area where sea fog occurred. This results in a relatively poor performance in Case-B.

Due to the negative correlation between temperature and moisture in CV6, assimilation of negative observational innovations of temperature can definitely get positive moisture increments, which leads to improvement of sea fog simulation.

However, degree of improvement (i.e., development of sea fog area) is primarily dominated by the localization scales in BE, which limits the spatial spread of observational information. The case studies indicate that the impact of CV6 depends on weather and real-time condition. Actually, this is directly due to the individual localization scales in BEs for Case-A and Case-B.

The localization scales are calculated when BE is generated by the NMC method. As mentioned in Section 2.2, forecast differences during 15 days are used to produce BE in the NMC method. Thus, BE comes from the temporal average of forecast differences, which means that BE does not vary with time and sometimes the localization scales are not suitable for a specific real-time condition, especially that the weather system changes rapidly. If ensemble forecast members are available from an ensemble forecast system, ensemble-based perturbations can be used instead of using forecast differences to generate CV6 in the NMC method. Perhaps this kind of CV6 becomes more effective, because some flow-dependent information could be introduced into the BE.

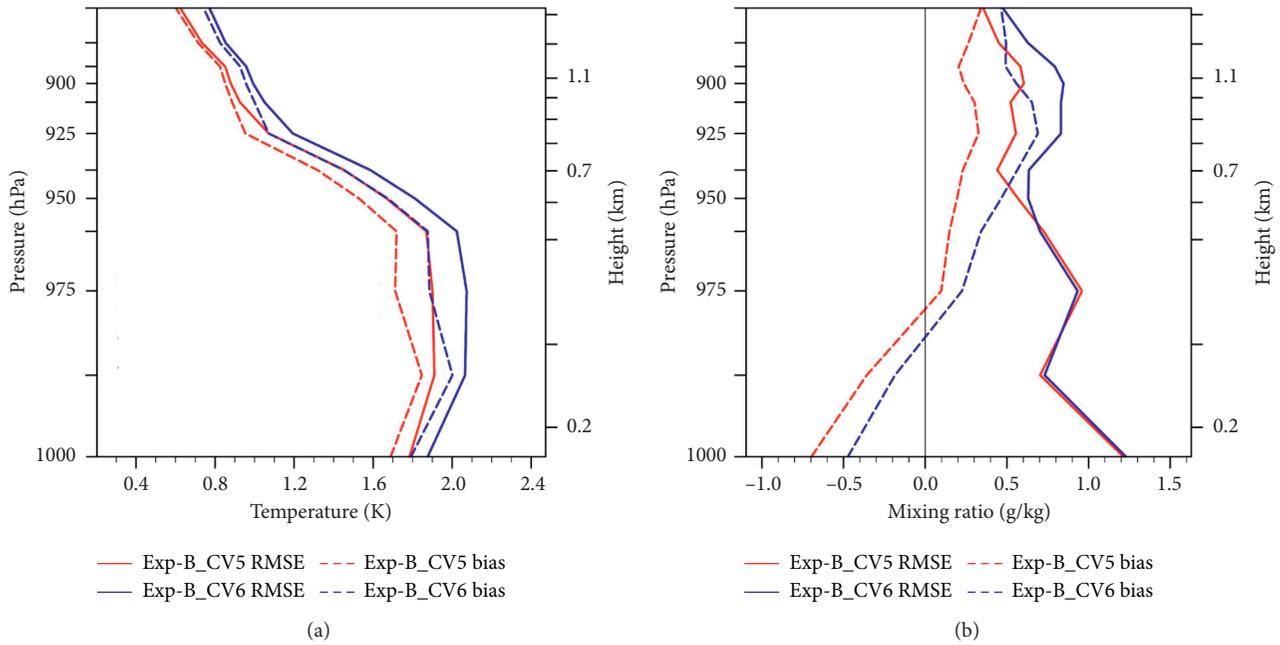


FIGURE 13: Vertical profiles of root mean square errors (RMSEs) (solid lines) and biases (dashed lines) between the initial analysis and the radiosonde observations of temperature (a) and mixing ratio (b) for Exp-B\_CV5 and Exp-B\_CV6.

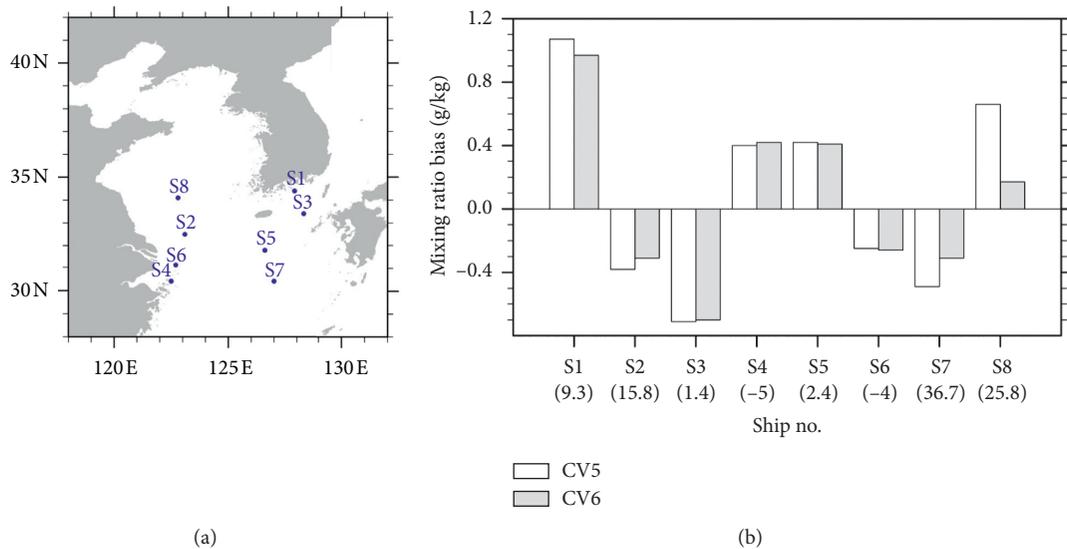


FIGURE 14: Locations of ship measurements (a) and biases of forecasted mixing ratio compared with observation (b) for Case-B. The improvements (%) of Exp-B\_CV6 compared to Exp-B\_CV5 are presented below the S# (identified mark of ship sites) along horizontal coordinates.

4.5. *Quantitative Evaluation.* The evaluation method of the simulated sea fog area in case study is a so-called eyeball method, which visually compares forecasted fog and satellite observed fog. This is a qualitative and subjective method, which cannot provide a quantitative statistical result for multiple cases. According to Zhou and Du [56], a more objective and comprehensive method is to remap both the observed and forecasted fog areas onto the same grids, in which point-to-point comparisons can be conducted. Fog area prediction can

be regarded as a binary event (yes or no; 1 or 0). Statistical scores, including the probability of detection (POD), false alarm ratio (FAR), bias, and equitable threat score (ETS), are used for evaluation. They are defined as follows:

$$POD = \frac{H}{O}, \tag{7a}$$

$$FAR = \frac{F-H}{F}, \tag{7b}$$

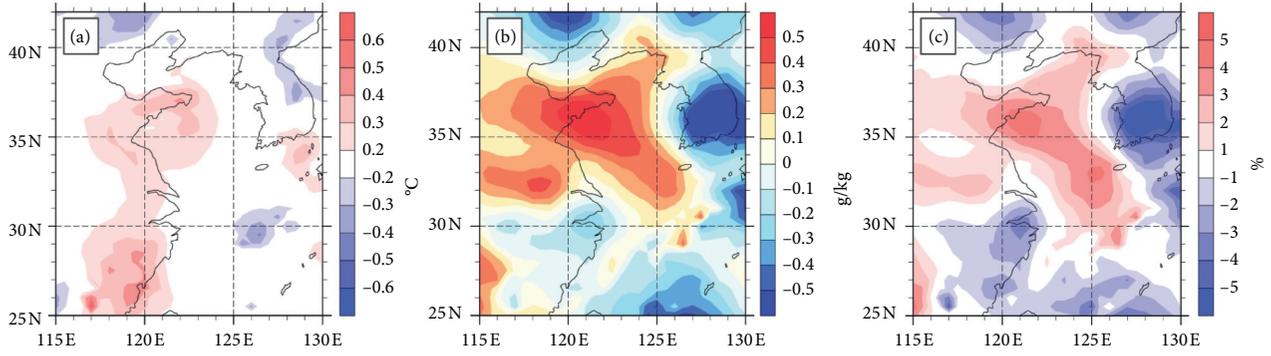


FIGURE 15: Initial differences of temperature (a), mixing ratio (b), and relative humidity (c) near the surface between Exp-B\_CV6 and Exp-B\_CV5 (the former minus the latter).

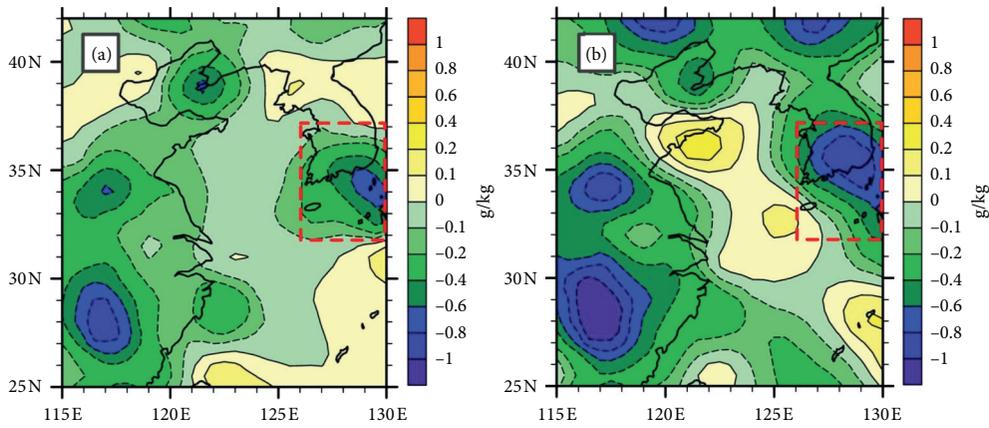


FIGURE 16: Analysis increments of mixing ratio in Exp-B\_CV5 (a) and Exp-B\_CV6 (b) near the surface. The area framed by a red-dashed line box is named Area-K.

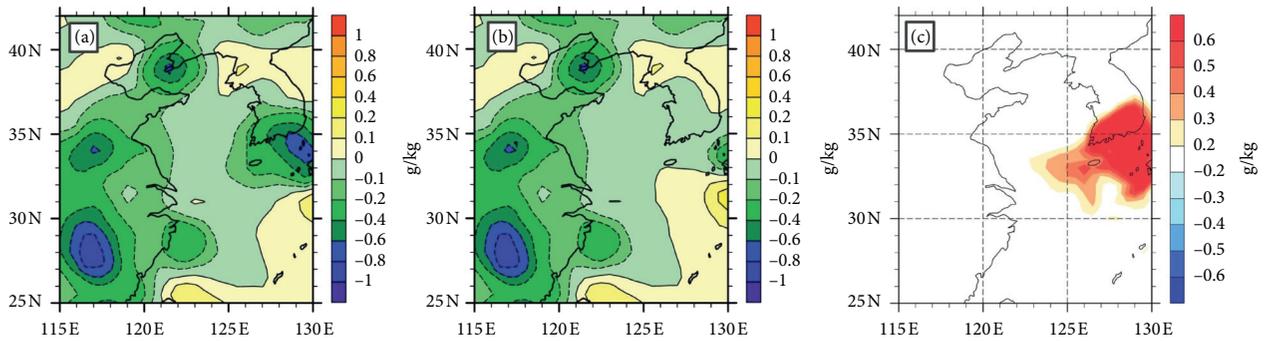


FIGURE 17: (a) Analysis increments of mixing ratio near the surface for Exp-B\_CV5 (a) and Exp-B\_CV5e (b) and differences in the initial mixing ratio between them (Exp-B\_CV5e minus Exp-B\_CV5) (c).

$$\text{bias} = \frac{F}{O} - 1, \quad (7c)$$

$$\text{ETS} = \frac{H-R}{F+O-H-R}. \quad (7d)$$

Here,  $H$ ,  $F$ , and  $O$  refer to the numbers of correctly forecast points (hits), forecast points, and observed points;  $R = F(O/N)$  is a random hit penalty, and  $N$  is the total number of grid points. The observed fog areas are retrieved

from the MTSAT using the method designed by [22], and D2 is taken as the verification domain with a mesh resolution of  $0.1^\circ$ . Grids covered by high-altitude clouds are excluded over both the land and the sea areas. Verification was carried out for all model outputs at 1-hour intervals. Quantitative evaluation was conducted using the statistical scores defined above for each case, and the scores were then temporally averaged. The evaluation results are presented in Table 3, in which the cases are arranged chronologically according to initial time.

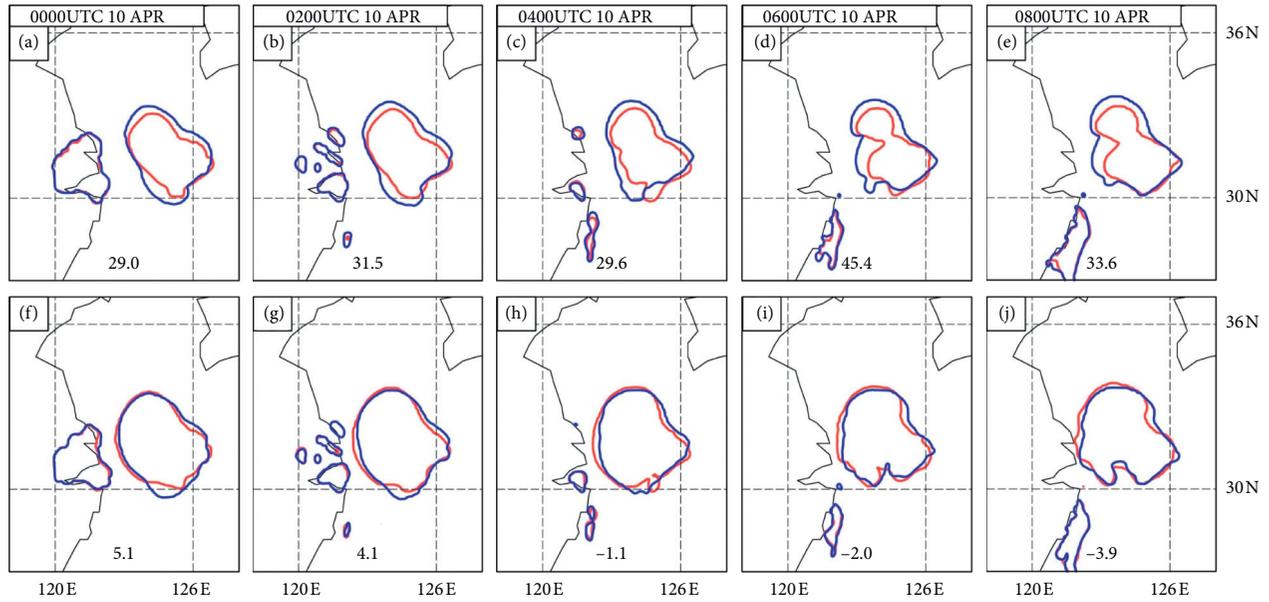


FIGURE 18: Boundaries of the forecasted sea fog area for Exp-B\_CV5 (red lines) and Exp-B\_CV5e (blue lines) (upper row) and for Exp-B\_CV6 (red lines) and Exp-B\_CV6e (blue lines) (lower row). The number in each panel is the enlargement rate (%) of sea fog area produced by Exp-B\_CV#e compared to that by Exp-B\_CV#.

TABLE 3: Statistical verification result of the experiments of the 10 sea fog cases. The improvements (%) in Group-CV6 relative to Group-CV5 are in parentheses.

Case and initial time no., yyyy-mm-dd_hh	Group-CV5				Group-CV6			
	POD	FAR	Bias	ETS	POD	FAR	Bias	ETS
1 2007-02-05_12	0.833	0.506	0.702	0.329	0.840 ( <b>0.8</b> )	0.491 ( <b>3.0</b> )	0.665 (3.7)	0.343 ( <b>4.3</b> )
2 2008-04-28_12	0.510	0.281	-0.292	0.289	0.781 ( <b>53.1</b> )	0.302 ( <b>-2.9</b> )	0.128 ( <b>16.4</b> )	0.432 ( <b>49.5</b> )
3 2009-05-02_12	0.935	0.290	0.317	0.528	0.933 ( <b>-0.2</b> )	0.287 ( <b>0.4</b> )	0.309 ( <b>0.8</b> )	0.532 ( <b>0.8</b> )
4 2010-02-22_12	0.733	0.307	0.057	0.340	0.827 ( <b>12.8</b> )	0.328 ( <b>-3.0</b> )	0.232 ( <b>-17.5</b> )	0.362 ( <b>6.5</b> )
5 2011-03-12_12	0.674	0.481	0.299	0.212	0.681 ( <b>1.0</b> )	0.523 ( <b>-8.1</b> )	0.429 ( <b>-13.0</b> )	0.172 ( <b>-18.9</b> )
6 2011-05-17_00	0.439	0.094	-0.516	0.344	0.823 ( <b>87.5</b> )	0.243 ( <b>-16.4</b> )	0.087 ( <b>42.9</b> )	0.556 ( <b>61.6</b> )
7 2012-03-27_12	0.578	0.487	0.126	0.302	0.892 ( <b>54.3</b> )	0.546 ( <b>-11.5</b> )	0.968 ( <b>-84.2</b> )	0.343 ( <b>13.6</b> )
8 2012-04-13_18	0.408	0.412	-0.305	0.113	0.489 ( <b>19.9</b> )	0.402 ( <b>1.7</b> )	-0.180 ( <b>12.5</b> )	0.142 ( <b>25.7</b> )
9 2012-05-09_12	0.245	0.384	-0.603	0.114	0.528 ( <b>115.5</b> )	0.391 ( <b>-1.1</b> )	-0.134 ( <b>46.9</b> )	0.236 ( <b>107.0</b> )
10 2014-05-22_02	0.471	0.368	-0.199	0.242	0.644 ( <b>36.7</b> )	0.446 ( <b>-12.3</b> )	0.197 ( <b>0.2</b> )	0.284 ( <b>17.4</b> )
Average	0.583	0.361	-0.041	0.281	0.744 ( <b>27.6</b> )	0.396 ( <b>-5.5</b> )	0.270 ( <b>-22.9</b> )	0.340 ( <b>21.0</b> )

Because ETS is a comprehensive scoring indicator, the result of ETS is analyzed first. It is clearly seen from Table 3 that the average ETS of Group-CV6 is much higher than that of Group-CV5 with an improvement of 21%, and only one case (Case-5 in Table 3) gets a worse ETS among the ten cases. It demonstrates that implementing CV-6 BE in the WRF-3DVAR can obviously improve sea fog modeling. Further case-by-case analyses of POD, FAR, bias, and ETS indicate that the improvement of ETS is mainly attributed to the increase of POD. Compared to that for Group-CV5, POD for Group-CV6 improved by 27.6% on average. Meanwhile, the average decreases in FAR and bias were -5.5% and -22.9%, respectively. Similar to the cases studied in the present study, CV6 results in more water vapor in the initial condition compared to CV5 for almost every case here

and hence promotes sea fog development and larger fog area (see POD improvements in Table 3). Although the results of FAR and bias show that the sea fog areas are a bit over-forecasted in Group-CV6, this is acceptable because the average deterioration of FAR is only 5.5%.

## 5. Summary

Data assimilation is absolutely necessary for numerical modeling of sea fog, because modeling result is highly sensitive to initial condition, especially to initial moisture status. The WRF-ARW model and WRF-3DVAR module have been widely employed for sea fog modeling, and the WRF-3DVAR has already been proven to be able to effectively improve the initial conditions for sea fog simulations

over the Yellow Sea. However, the default domain-dependent BE of the WRF-3DVAR is CV5, which does not consider the relationship between moisture and other control variables. As a result, the improvement of moisture in the initial condition has to be only dependent on the direct assimilation of moisture information. This approach causes some disadvantageous issues; for example, other control variables like temperature can hardly contribute to the improvement of moisture status.

By theoretically comparing CV5 with CV6 (multivariate BE), it is found that CV6 includes the correlation between moisture and temperature, which is absent in CV5. In order to explore the impact of multivariate BE on the 3DVAR assimilation effect over the Yellow Sea fog modeling, two sea fog cases that differ greatly were selected for case study. The advantage of CV6 compared to CV5 has been primarily explained by single-observation experiments and further revealed in detail based on analysis of the modeling results from the cases study. In addition, the impact of CV6 application in the WRF-3DVAR has been evaluated by a series of experiments of extra ten sea fog events over the Yellow Sea. The major conclusions are as follows:

- (1) The performance of the WRF-3DVAR assimilation with CV6 is obviously better than that with CV5. The assimilation with CV6 can significantly improve the forecasted sea fog area, which is clearly demonstrated in the study of the selected cases. Particularly, for the case of spreading narrowly along the coast, the sea fog evolution is successfully reproduced when the assimilation with CV6 is implemented. In contrast, the model completely fails to capture this fog event when using the assimilation with CV5.
- (2) There exists a dominant negative correlation between temperature and moisture in CV6 near the sea surface, which makes it possible to improve initial moisture condition in the marine atmospheric boundary layer by assimilating temperature observations near the sea surface. However, the primary way to improve moisture in the WRF-3DVAR with CV5 is to assimilate moisture observations.
- (3) Experimental results of the 10 additional sea fog cases clearly indicate that CV6 is more suitable than CV5 in the WRF-3DVAR assimilation for sea fog modeling. Compared to that with CV5, the average POD and ETS of forecasted sea fog area using the WRF-3DVAR with CV6 can be improved by 27.6% and 21.0%, respectively, and positive impacts are found in 9 out of the 10 cases in the present study.

The present study has demonstrated the importance of improving moisture condition in the initial condition of sea fog modeling and the necessity of using CV6 in the WRF-3DVAR to achieve this improvement. This is of reference value for the real-time operational forecast of sea fog when using the WRF-3DVAR. Despite encouraging results that have been achieved in this study, further in-depth studies are still needed. CV6 is essentially a type of climatological background error covariance, which usually varies on

monthly basis and may also have a feature of rapid variation in the low-level atmosphere due to transient weather systems. Thus, when ensemble forecast members are available, it is better to use ensemble-based perturbations instead of using forecast differences in the NMC method. This is because some flow-dependent information might be introduced into the BE. In addition, the localizations of horizontal and vertical correlation in CV6 represent an issue that needs to be addressed.

## Data Availability

FNL data can be downloaded at <https://rda.ucar.edu/datasets/ds083.2>. NEAR-GOOS SST data are available at <http://ds.data.jma.go.jp/gmd/goos/data>. AIRS-retrieved temperature and humidity profiles can be accessed at <https://rda.ucar.edu/datasets/ds735.0>. MTSAT data are available at <http://weather.is.kochi-u.ac.jp/sat/GAME>, including albedo, infrared, and visible cloud imageries. The NCEP ship observations are from <https://rda.ucar.edu/datasets/ds337.0>. Other data for this study are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

## Acknowledgments

This research was funded by the National Key Research and Development Program of China (2017YFC1404200 and 2017YFC1404100), the Key Research and Development Program of Shandong Province (2019GSF111066), and the National Natural Science Foundation of China (42075069). The authors acknowledge UCAR/NCAR/CISL/TDD for providing the powerful graphic tool NCL (<https://doi.org/10.5065/D6WD3XH5>).

## References

- [1] I. Gulpepe, R. Tardif, S. C. Michaelides et al., "Fog research: a review of past achievements and future perspectives," *Fog and Boundary Layer Clouds: Fog Visibility and Forecasting*, vol. 164, no. 6-7, pp. 1121-1159, 2007.
- [2] D. Koraćin and C. E. Dorman, *Marine Fog: Challenges and Advancements in Observations, Modeling, and Forecasting*, Springer International Publishing, San Diego, CA, USA, 1st edition, 2017.
- [3] S. Gao, H. Lin, B. Shen, and G. Fu, "A heavy sea fog event over the Yellow Sea in March 2005: analysis and numerical modeling," *Advances in Atmospheric Sciences*, vol. 24, no. 1, pp. 65-81, 2007.
- [4] B. Wang, *Sea Fog*, China Ocean Press, Beijing, China, 1st edition, 1985.
- [5] S. Zhang and W. Bao, "The main advances in sea fog research in China," *Journal of Ocean University of China*, vol. 38, no. 3, pp. 359-366, 2008, in Chinese.
- [6] S. Nicholls, "The dynamics of stratocumulus: aircraft observations and comparisons with a mixed layer model,"

- Quarterly Journal of the Royal Meteorological Society*, vol. 110, no. 466, pp. 783–820, 1984.
- [7] J. Findlater, W. T. Roach, and B. C. McHugh, “The haar of north-east Scotland,” *Quarterly Journal of the Royal Meteorological Society*, vol. 115, no. 487, pp. 581–608, 1989.
  - [8] S. P. Ballard, B. W. Golding, and R. N. B. Smith, “Mesoscale model experimental forecasts of the haar of northeast Scotland,” *Monthly Weather Review*, vol. 119, no. 9, pp. 2107–2123, 1991.
  - [9] J. M. Lewis, D. Koraćin, R. Rabin, and J. Businger, “Sea fog off the California coast: viewed in the context of transient weather systems,” *Journal of Geophysical Research*, vol. 108, no. D15, 2003.
  - [10] D. Koraćin, J. Lewis, W. T. Thompson, C. E. Dorman, and J. A. Businger, “Transition of stratus into fog along the California coast: observation and modeling,” *Journal of the Atmospheric Sciences*, vol. 58, no. 13, pp. 1714–1731, 2001.
  - [11] D. Koraćin, J. A. Businger, C. E. Dorman, and J. M. Lewis, “Formation, evolution, and dissipation of coastal sea fog,” *Boundary-Layer Meteorology*, vol. 117, no. 3, pp. 447–478, 2005.
  - [12] D. Koraćin, D. F. Leipper, and J. M. Lewis, “Modeling sea fog on the U.S. California coast during a hot spell event,” *Geofizika*, vol. 22, no. 1, pp. 59–82, 2005.
  - [13] P. Courtier, E. Andersson, W. Heckley et al., “The ECMWF implementation of three-dimensional variational assimilation (3D-Var). I: formulation,” *Quarterly Journal of the Royal Meteorological Society*, vol. 124, no. 550, pp. 1783–1807, 1998.
  - [14] F. Rabier, A. McNally, E. Andersson et al., “The ECMWF implementation of three-dimensional variational assimilation (3D-Var). II: structure functions,” *Quarterly Journal of the Royal Meteorological Society*, vol. 124, no. 550, pp. 1809–1829, 1998.
  - [15] E. Andersson, J. Haseler, P. Undén et al., “The ECMWF implementation of three-dimensional variational assimilation (3D-Var). III: experimental results,” *Quarterly Journal of the Royal Meteorological Society*, vol. 124, no. 550, pp. 1831–1860, 1998.
  - [16] H. Zhang, J. Xue, S. Zhuang, G. Zhu, and Z. Zhu, “Idea experiments of GRAPES three-dimensional variational data assimilation system,” *Acta Meteorologica Sinica*, vol. 62, no. 1, pp. 31–41, 2004, in Chinese.
  - [17] J. Xue, S. Zhuang, G. Zhu et al., “The new global/regional variational assimilation system of GRAPES,” *Chinese Science Bulletin*, vol. 20, pp. 18–27, 2008, in Chinese.
  - [18] D. M. Barker, W. Huang, Y.-R. Guo, A. J. Bourgeois, and Q. N. Xiao, “A three-dimensional variational data assimilation system for MM5: implementation and initial results,” *Monthly Weather Review*, vol. 132, no. 4, pp. 897–914, 2004.
  - [19] D. Barker, X.-Y. Huang, Z. Liu et al., “The weather research and forecasting model’s community variational/ensemble data assimilation system: WRFDA,” *Bulletin of the American Meteorological Society*, vol. 93, no. 6, pp. 831–843, 2012.
  - [20] X. Gao, S. Gao, and Y. Yang, “A comparison between 3DVAR and EnKF for data assimilation effects on the Yellow Sea fog forecast,” *Atmosphere*, vol. 9346 pages, 2018.
  - [21] S. Gao, Y. Qi, S. Zhang, and G. Fu, “Initial conditions improvement of sea fog numerical modeling over the Yellow Sea by using cycling 3DVAR. Part I: WRF numerical experiments,” *Journal of Ocean University of China*, vol. 40, no. 10, pp. 1–9, 2010, in Chinese.
  - [22] Y. Wang, S. Gao, G. Fu, J. Sun, and S. Zhang, “Assimilating MTSAT-derived humidity in nowcasting sea fog over the Yellow Sea,” *Weather and Forecasting*, vol. 29, no. 2, pp. 205–225, 2014.
  - [23] Y. Yang, X. M. Hu, S. Gao, and Y. Wang, “Sensitivity of WRF simulations with the YSU PBL scheme to the lowest model level height for a sea fog event over the Yellow Sea,” *Atmospheric Research*, vol. 215, no. 1, pp. 253–267, 2018.
  - [24] Y. Yang and S. Gao, “The impact of turbulent diffusion driven by fog-top cooling on sea fog development,” *Journal of Geophysical Research: Atmospheres*, vol. 125, no. 4, 2020.
  - [25] K. Ide, P. Courtier, M. Ghil, and A. C. Lorenc, “Unified notation for data assimilation: operational, sequential and variational (gtSpecial Issue) Data assimilation in meteorology and oceanography: theory and practice,” *Journal of the Meteorological Society of Japan. Ser. II*, vol. 75, no. 1B, pp. 181–189, 1997.
  - [26] R. N. Bannister, “A review of forecast error covariance statistics in atmospheric variational data assimilation. I: characteristics and measurements of forecast error covariances,” *Quarterly Journal of the Royal Meteorological Society*, vol. 134, no. 637, pp. 1951–1970, 2008.
  - [27] R. N. Bannister, “A review of forecast error covariance statistics in atmospheric variational data assimilation. II: modelling the forecast error covariance statistics,” *Quarterly Journal of the Royal Meteorological Society*, vol. 134, no. 637, pp. 1971–1996, 2008.
  - [28] M. Buehner, “Ensemble-derived stationary and flow-dependent background-error covariances: evaluation in a quasi-operational NWP setting,” *Quarterly Journal of the Royal Meteorological Society*, vol. 131, no. 607, pp. 1013–1043, 2005.
  - [29] J. Derber and F. Bouttier, “A reformulation of the background error covariance in the ECMWF global data assimilation system,” *Tellus A: Dynamic Meteorology and Oceanography*, vol. 51, no. 2, pp. 195–221, 1999.
  - [30] D. Katz, A. S. Lawless, N. K. Nichols, M. J. P. Cullen, and R. N. Bannister, “Correlations of control variables in variational data assimilation,” *Quarterly Journal of the Royal Meteorological Society*, vol. 137, no. 656, pp. 620–630, 2011.
  - [31] V. Rakesh and P. Goswami, “Impact of background error statistics on 3D-Var assimilation: case study over the Indian region,” *Meteorology and Atmospheric Physics*, vol. 112, no. 1–2, pp. 63–79, 2011.
  - [32] D. F. Parrish and J. C. Derber, “The national meteorological center’s spectral statistical-interpolation analysis system,” *Monthly Weather Review*, vol. 120, no. 8, pp. 1747–1763, 1992.
  - [33] L. Berre, “Estimation of synoptic and mesoscale forecast error covariances in a limited-area model,” *Monthly Weather Review*, vol. 128, pp. 644–667, 1998.
  - [34] M. Dhanya and A. Chandrasekar, “Impact of variational assimilation using multivariate background error covariances on the simulation of monsoon depressions over India,” *Annales Geophysicae*, vol. 34, no. 2, pp. 187–201, 2016.
  - [35] M. Dhanya and A. Chandrasekar, “Multivariate background error covariances in the assimilation of SAPHIR radiances in the simulation of three tropical cyclones over the Bay of Bengal using the WRF model,” *International Journal of Remote Sensing*, vol. 39, no. 1, pp. 191–209, 2018.
  - [36] G. Fu, J. Guo, S.-P. Xie, Y. Duan, and M. Zhang, “Analysis and high-resolution modeling of a dense sea fog event over the Yellow Sea,” *Atmospheric Research*, vol. 81, no. 4, pp. 293–303, 2006.
  - [37] K.-Y. Heo and K.-J. Ha, “A coupled model study on the formation and dissipation of sea fogs,” *Monthly Weather Review*, vol. 138, no. 4, pp. 1186–1205, 2010.

- [38] P. Li, G. Fu, C. Lu, D. Fu, and S. Wang, "The formation mechanism of a spring sea fog event over the Yellow Sea associated with a low-level jet," *Weather and Forecasting*, vol. 27, no. 6, pp. 1538–1553, 2012.
- [39] C. K. Kim and S. S. Yum, "A numerical study of sea-fog formation over cold sea surface using a one-dimensional turbulence model coupled with the weather research and forecasting model," *Boundary-Layer Meteorology*, vol. 143, no. 3, pp. 481–505, 2012.
- [40] H.-Y. Lee and E.-C. Chang, "Impact of land-sea thermal contrast on the inland penetration of sea fog over the coastal area around the Korean Peninsula," *Journal of Geophysical Research: Atmospheres*, vol. 123, no. 12, pp. 6487–6504, 2018.
- [41] R. Li, S. Gao, and Y. Wang, "Numerical study on direct assimilation of satellite radiances for sea fog over the Yellow Sea," *Journal of Ocean University of China*, vol. 42, no. 3, pp. 10–20, 2012, in Chinese.
- [42] X. Lu, S. Gao, L. Rao, and Y. Wang, "Sensitivity study of WRF parameterization schemes for the spring sea fog in the Yellow Sea," *Journal of Applied Meteorological Science*, vol. 25, no. 3, pp. 312–320, 2014, in Chinese.
- [43] Y. Wang and S. Gao, "Assimilation of Doppler Radar radial velocity in Yellow Sea fog numerical modeling," *Journal of Ocean University of China*, vol. 46, no. 1–12, 2016, in Chinese.
- [44] M. G. Divakarla, C. D. Barnet, M. D. Goldberg et al., "Validation of Atmospheric Infrared Sounder temperature and water vapor retrievals with matched radiosonde measurements and forecasts," *Journal of Geophysical Research*, vol. 111, no. D9, Article ID D09S15, 2006.
- [45] D. C. Liu and J. Nocedal, "On the limited memory BFGS method for large scale optimization," *Mathematical Programming*, vol. 45, no. 1–3, pp. 503–528, 1989.
- [46] P. Courtier, J. N. Thépaut, and A. Hollingsworth, "A strategy for operational implementation of 4D-Var, using an incremental approach," *Quarterly Journal of the Royal Meteorological Society*, vol. 120, no. 1, pp. 1367–1387, 1994.
- [47] C. M. Hayden and R. J. Purser, "Recursive filter objective analysis of meteorological fields: applications to NESDIS operational processing," *Journal of Applied Meteorology*, vol. 34, no. 1, pp. 3–15, 1995.
- [48] R. J. Purser, W.-S. Wu, D. F. Parrish, and N. M. Roberts, "Numerical aspects of the application of recursive filters to variational statistical analysis. Part I: spatially homogeneous and isotropic Gaussian covariances," *Monthly Weather Review*, vol. 131, no. 8, pp. 1524–1535, 2003.
- [49] M. Li and S. Zhang, "Impact of sea surface temperature front on stratus-sea fog over the Yellow and East China Seas—a case study with implications for climatology," *Journal of Ocean University of China*, vol. 12, no. 2, pp. 301–311, 2013.
- [50] S. Y. Hong, Y. Noh, and J. Dudhia, "A new vertical diffusion package with an explicit treatment of entrainment processes," *Monthly Weather Review*, vol. 134, no. 4, pp. 2318–2341, 2006.
- [51] J. S. Kain, "The Kain-Fritsch convective parameterization: an update," *Journal of Applied Meteorology*, vol. 43, no. 1, pp. 170–181, 2004.
- [52] J. S. Kain and J. M. Fritsch, "A one-dimensional entraining/detraining plume model and its application in convective parameterization," *Journal of the Atmospheric Sciences*, vol. 47, no. 23, pp. 2784–2802, 1990.
- [53] Y.-L. Lin, R. D. Farley, and H. D. Orville, "Bulk parameterization of the snow field in a cloud model," *Journal of Climate and Applied Meteorology*, vol. 22, no. 6, pp. 1065–1092, 1983.
- [54] M. J. Iacono, J. S. Delamere, E. J. Mlawer, M. W. Shephard, S. A. Clough, and W. D. Collins, "Radiative forcing by long-lived greenhouse gases: calculations with the AER radiative transfer models," *Journal of Geophysical Research*, vol. 113, no. D13103, 2008.
- [55] F. Chen and J. Dudhia, "Coupling an advanced land surface-hydrology model with the Penn State-NCAR MM5 modeling system. Part I: model implementation and sensitivity," *Monthly Weather Review*, vol. 129, no. 4, pp. 569–585, 2001.
- [56] B. Zhou and J. Du, "Fog prediction from a multimodel mesoscale ensemble prediction system," *Weather and Forecasting*, vol. 25, no. 1, pp. 303–322, 2010.