Variational Bias Correction of TAMDAR Temperature Observations in the WRF Data Assimilation System

Feng Gao,[a,b] Zhiquan Liu,[c] Juhui Ma,[d] Neil A. Jacobs,[a,e]
Peter P. Childs,[a,f] and Hongli Wang[a,g,h]

[a] *Panasonic Avionics, Morrisville, North Carolina*
[b] *WeatherFlow, Fort Collins, Colorado*
[c] *Mesoscale and Microscale Meteorology Laboratory, National Center for Atmospheric Research, Boulder, Colorado*
[d] *Department of Marine, Earth and Atmospheric Sciences, North Carolina State University, Raleigh, North Carolina*
[e] *U.S. Department of Commerce/National Oceanic and Atmospheric Administration, Washington, D.C.*
[f] *Priogen Energy, Durham, North Carolina*
[g] *Cooperative Institute for Research in Environmental Sciences, Boulder, Colorado*
[h] *National Oceanic and Atmospheric Administration/Earth System Research Laboratory, Boulder, Colorado*

ABSTRACT

A variational bias correction (VarBC) scheme is developed and tested using regional Weather Research and Forecasting Model Data Assimilation (WRFDA) to correct systematic errors in aircraft-based measurements of temperature produced by the Tropospheric Airborne Meteorological Data Reporting (TAMDAR) system. Various bias models were investigated, using one or all of aircraft height tendency, Mach number, temperature tendency, and the observed temperature as predictors. These variables were expected to account for the representation of some well-known error sources contributing to uncertainties in TAMDAR temperature measurements. The parameters corresponding to these predictors were evolved in the model for a two-week period to generate initial estimates according to each unique aircraft tail number. Sensitivity experiments were then conducted for another one-month period. Finally, a case study using VarBC of a cold front precipitation event is examined. The implementation of VarBC reduces biases in TAMDAR temperature innovations. Even when using a bias model containing a single predictor, such as height tendency or Mach number, the VarBC produces positive impacts on analyses and short-range forecasts of temperature with smaller standard deviations and biases than the control run. Additionally, by employing a multiple-predictor bias model, which describes the statistical relations between innovations and predictors, and uses coefficients to control the evolution of components in the bias model with respect to their reference values, VarBC further reduces the average error of analyses and short-range forecasts with respect to observations. The potential impacts of VarBC on precipitation forecasts were evaluated, and the VarBC is able to indirectly improve the prediction of precipitation location by reducing the forecast error for wind-related synoptic circulation leading to precipitation.

1. Introduction

Most, if not all, data assimilation systems assume observations are unbiased (Le Dimet and Talagrand 1986; Parrish and Derber 1992; Barker et al. 2004; Huang et al. 2009; Shao et al. 2016). In these systems, the biases in observations are linearly transferred into the analysis, regardless of the specification of a gain operator (Dee 2005). Therefore, the systematic errors in observations and the representative errors must be corrected to maximize the potential benefit of observations during data assimilation.

To supplement existing in situ observing systems, aircraft-based observing systems have been expanded and a significantly increasing volume of such measurements have been collected over the past three decades (Drüe et al. 2008; Moninger et al. 2010). Tropospheric Airborne Meteorological Data Reporting (TAMDAR) sensors were developed in 2004 to improve mesoscale numerical weather prediction. These sensors offer high spatiotemporal resolution within the troposphere, where the majority of convective activity originates. However, aircraft-based measurements are subject to systematic errors due to imperfect measurement techniques and data processing procedures (World Meteorological Organization 2003; Jacobs et al. 2014).

Corresponding author: Feng Gao, gao.feng.nuist@gmail.com

Biases in aircraft temperature data generally come from the following primary sources: 1) incomplete correction of heating from adiabatic compression in the sensor chamber, which is a function of Mach number derived from total pressure and static pressure; 2) errors in the Mach number itself;[1] 3) time lag due to slightly wrong time stamp being associated with a particular data point; 4) thermal lag or hysteresis in the probe material and environment; and 5) aircraft maneuvering, which is likely to make measurements become unstable and affect the estimation accuracies of air heating caused by viscous friction and partial blocking of the thermistor. Although some aircraft data are excluded when the roll angle of aircraft is above a certain threshold, data observed during the period of maneuvering is possibly used when the maneuver is light.

Several investigations have addressed the biases of aircraft temperature measurement. When comparing with reference observations and background forecasts, aircraft temperature data generally have a warm bias, and the bias varies with the ascent and descent phases of the aircraft, the aircraft type, and even slightly with season (Drüe et al. 2008; Ballish and Kumar 2008). The presence of a warm bias in aircraft temperature observations means that these data are not used optimally in assimilation and may warm the analysis too much, although some studies have reported forecast improvements from using commercial aircraft data in regional and global models (Cardinali et al. 2003; Moninger et al. 2010; Gao et al. 2012; Petersen 2016).

Various approaches to estimating observational biases have been reported. Several studies suggest comparing the observations with uncorrelated comparator data, and then determining the bias by examining the discrepancies. In general, fine-mesh short-range model forecasts and observations with well-known accuracy are typically used as available comparator data. Benjamin et al. (1999) estimated the observation errors for Aircraft Communications Addressing and Reporting System (ACARS) by comparing with ACARS collected by aircraft with different tail numbers in a collocation study. Drüe et al. (2008) described the characteristics of observational errors of Aircraft Meteorological Data Relay (AMDAR) measurements against rawinsonde observations. In áreality, ACARS is a name of communication system through which aircraft-based reports are transmitted; AMDAR presents a complete observing system including aircraft onboard sensor, computers and communications systems where the communication systems can be based on satellite or ACARS. For simplicity, both ACARS and AMDAR are often used to refer directly to the meteorological observations in some studies (Benjamin et al. 1999). Compared with ACARS and AMDAR installed on airframes with long-range wide body, some TAMDAR units are installed on small unmanned aerial systems which provide more local weather information for high-resolution models. Moninger et al. (2010) evaluated TAMDAR observations by comparing them with the Rapid Update Cycle (RUC) 1-h forecast. These researches well addressed the pictures of aircraft observation errors. However, a noticeable disadvantage in these collocation studies is that multiple error sources involved in observations and references are combined in the error estimations.

Bias correction for aircraft observations should not aim to remove the biases present in the comparator data to avoid drifting the model uniformly toward the error in comparator. Additionally, the bias in aircraft observations is ever-changing, and at each analysis time, different error sources may dominate (World Meteorological Organization 2003). For daily bias correction of TAMDAR observations provided by several hundreds of aircraft, the temporal changes in bias can occur and should be accounted for.

Adaptive bias correction was first developed and implemented for use with satellite data and applies a variational framework (Derber and Wu 1998; Harris and Kelly 2001; Auligné et al. 2007; Zhu et al. 2014). In this approach, the observation operator in the cost function is updated by adding a bias model, which consists of a set of parameters and predictors. These predictors, often considered the key factors describing the systematic errors, are assumed to be capable of representing the bias in the data. Along with the model state variables, these parameters are updated during the cost function minimization step, which seeks the increments that minimize observational departures while fitting as many observations from all data sources as possible in the analysis. Thus, this approach is able to automatically track and integrate real-time changes of observations and forecast model variables. It should be noted that some other observation types are needed to anchor bias correction refraining from compensating for model errors. Eyre (2016) has addressed that the effect of model errors in analysis biases may increase when more observations are bias corrected and a smaller proportion are used as "anchor" observations, or more weight is given to the observations to be bias corrected than the other 'anchoring' observations.

---

[1] $\Delta T = 0.388 MT(1 + 0.194M^2)^{-1}\Delta M$, where T and M stand for temperature observation and Mach number, respectively, and Δ denotes the error of a variable. The equation can be seen in World Meteorological Organization (2003; p. 10).

Isaksen et al. (2012) extended the variational bias correction method (VarBC) to correct the bias of aircraft temperature observations by building a constant bias model into the European Centre for Medium-Range Weather Forecasts (ECMWF) global data assimilation system. Their results demonstrate that both biases and random error magnitude (standard deviation) of temperature analyses and short-term forecasts fit to aircraft and radiosonde temperature observations were reduced. The improvements were largest at 250–300 hPa, where the commercial aircraft typically cruised and thus the aircraft volume was large.

A similar scheme was implemented in the National Centers for Environmental Prediction (NCEP) Gridpoint Statistical Interpolation (GSI) analysis system by Zhu et al. (2015), who evaluated three extra bias models, such as ascent and descent rates, quadratic ascent and descent rates, and log pressure, as well as the one used by ECMWF. In their study, the bias correction scheme cooled the global analysis around 200 hPa in the Northern Hemisphere, and the forecasts were fitted closer to radiosonde temperature observations without degrading the fit at lower vertical levels. These contrasting experiments using different bias models demonstrate that the predictors of ascent and descent rates performed better than using predictors of constant and log pressure, and the addition of the predictors of quadratic ascent and descent rates produced higher skill scores on average. The results suggest that the predictors of ascent and descent rates had the ability to describe the bias of aircraft temperature observations, while an extra nonlinear constraint term (e.g., quadratic ascent and descent rates), is also needed to further improve the analysis.

Multiple error sources contribute to the uncertainty of aircraft temperature observations. In Isaksen et al. (2012) and Zhu et al. (2015), apart from sources that can be explicitly represented by the predictors of ascent and descent rates, other sources are represented by the constant-predictor term (first term) in the bias model. To supplement past studies, we investigate the TAMDAR temperature VarBC in a regional model, the Weather Research and Forecasting Data Assimilation system (WRFDA). By building a bias model, which describes the regression relations between systematic errors and predictors, we implement similar testing as previously carried out by NCEP and ECMWF to examine the performance of various bias models.

In section 2, we provide an overview of the variational bias correction (VarBC) methodology in WRFDA. Section 3 describes the model configuration, database, domain, and experimental setup. The implementation of bias correction procedure and the impacts of VarBC

on TAMDAR temperature assimilation are presented in section 4. In section 5, we illustrate a cold front precipitation case study. Finally, conclusions and discussion are included in section 6.

2. WRFDA and new formulation

WRFDA is a multifunctional data assimilation model that uses a variational framework, and either three/four-dimensional variational or hybrid data assimilation (Barker et al. 2004; Huang et al. 2009; Wang et al. 2008a,b; Schwartz et al. 2015) can be used. In this study, the basic 3D-Var is employed.

a. Observation operator and bias model

Aiming to obtain a statistically optimal estimation of the atmospheric state by combining observations $\mathbf{y}$ with a background forecast $\mathbf{x}_b$, the cost function in (1) is minimized iteratively in WRFDA with respect to the model state $\mathbf{x}$:

$$J(\mathbf{x}) = \frac{1}{2}(\mathbf{x} - \mathbf{x}_b)^{\mathrm{T}}\mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}_b)$$
$$+ \frac{1}{2}[\mathbf{y} - h(\mathbf{x})]^{\mathrm{T}}\mathbf{R}^{-1}[\mathbf{y} - h(\mathbf{x})], \qquad (1)$$

where the covariance matrices $\mathbf{B}$ and $\mathbf{R}$ are the perceived error statistics of the background and observations, respectively; and h is an observation operator (i.e., a transformation from model state to the observation space) including the grid interpolation. The superscripts -1 and T denote the inverse and adjoint, respectively, of a matrix.

To solve (1) using an approximation to the best linear unbiased estimate, the observation vector is assumed to be bias-free. Since aircraft temperature observations contain nonnegligible systematic errors, the observation operator h must be updated by adding a bias model b in (2) in such a way that (3) is satisfied:

$$\tilde{h}(\mathbf{x}, \boldsymbol{\beta}, \mathbf{y}) = h(\mathbf{x}) - b(\boldsymbol{\beta}, \mathbf{y}) \qquad (2)$$

$$\langle \mathbf{e} \rangle = \langle \mathbf{y} - \tilde{h}(\mathbf{x}, \boldsymbol{\beta}, \mathbf{y}) \rangle = 0, \qquad (3)$$

where $\tilde{h}$ is the modified observation operator, $\mathbf{e}$ denotes the departure, and $\langle \ \rangle$ represents averaging over the departure. The bias model b consists of a set of parameters $\boldsymbol{\beta}$ and predictors. These predictors are supposed to account for the representation of the systematic error sources in TAMDAR temperature observation vector $\mathbf{y}$. If N is the number of predictors, b can be written as a linear combination of parameters β_n and predictors p_n, $n = 1, 2, \ldots, N$, as in (4):

$$b(\boldsymbol{\beta}, \mathbf{y}) = \sum_{n=1}^{N} \beta_n p_n(\mathbf{y}). \tag{4}$$

In the conventional practice of VarBC, β_1 is assigned to be the global offset, and the predictor p_1 is a constant of 1.0 (Dee and Uppala 2009). In the previously described NCEP global data assimilation systems, the aircraft ascent/descent rate (ADR) is used as the second predictor (Zhu et al. 2015). However, in our initial tests, the bias directly estimated by the bias model are occasionally far beyond the reasonable error range, especially during the phase of ascent when frequent maneuvers are conducted. In variational bias correction, the bias model is actually calculated by using the priori parameters and the real time computed predictors at analysis time. Therefore, the large ascent rates at one time may overestimate the correction to bias when the priori parameters are well estimated based on the past normal ascent rates. This often happens in TAMDAR measurements rather than ACARS and AMDAR because, compared with the large-size commercial aircraft flying in preset route, the vertical motions and airways of these small-size noncommercial aircraft and unmanned aerial vehicle generating TAMDAR are often adjusted according to local weather and environment whenever necessary. Since it is known that the height tendency of aircraft (i.e., ascent/descent rates) has a typical large range, a scaling method is needed in order to scale its magnitude to match with other predictor(s) in bias model. To conduct the bias correction within the typical error range of aircraft temperature measurements (Benjamin et al. 1999; Gao et al. 2012), two methods can be used. Either there can be an adjustment to the update of parameters via a scaling term or a coefficient is provided to adjust predictors. It is noted that the normalization method, regardless of the absolute magnitude, is inappropriate in this case because the absolute magnitudes of height tendency at different analysis time do matter to the variational bias correction. To keep the scaling terms consistent in all sensitivity experiments, a coefficient a_{ADR} of 0.1 is added into the ascent/descent rate predictor as the benchmark of ADR in this study. Thus, the bias model (4) is extended to (5) and used for these aircraft phases of ascent, descent and cruise, respectively:

$$b_{\mathrm{ADR}}(\boldsymbol{\beta}, \mathbf{y}) = \beta_1 + \beta_2 w a_{\mathrm{ADR}}, \tag{5}$$

where w is the ascent/descent rate of aircraft (with units of $\mathrm{m\,s^{-1}}$), estimated by the observed height tendency. Some error sources related to w (e.g., actual aircraft maneuvers and time lag of data recording as the aircraft moves in the vertical direction) can be explicitly expressed in (5). Apparently, by setting a_{ADR} to 0.1, the weight of the ascent/descent rate term in the bias model is artificially reduced. To examine the sensitivity of bias model to a_{ADR}, one can adjust the constant a_{ADR}. This question will be addressed later.

To investigate another key error source, we replace the predictor of the ascent/descent rate with Mach number M, as expressed in (6):

$$b_{\mathrm{MCN}}(\boldsymbol{\beta}, \mathbf{y}) = \beta_1 + \beta_2 M, \tag{6}$$

where the error sources related to Mach number are explicitly represented (e.g., the incomplete correction of air heating in the sensor chamber and the error in the Mach number itself). In the eastern half of contiguous United States (CONUS), all aircraft installing TAMDAR sensors are Dash-8 turboprops, which cruise at altitudes generally below 500 hPa, and have lower Mach values than jets collecting measurements for ACARS and AMDAR. The average Mach number for turboprops is below 0.4, so it is not necessary to adjust the predictor based on our initial tests.

Given the fact that no single predictor can fully represent error information in the data, a bias model describing multiple error sources [weighted multiple predictors (WMP)] is designed in this study based on the regression relations between OMB and various error sources. This will be discussed in the following section.

b. Augmentation of model state

In WRFDA, a routine approach to estimating parameters in connection with uncertainties in components of the cost function is the state augmentation method, which, by augmenting the dimension of the model state, adds uncertainty parameters to the other evolving model state variables to be estimated during the cost minimization step. The method is used widely, for instance, in the estimation of control variables α that determine the weights of ensemble perturbations in hybrid variational-ensemble data assimilation (Wang et al. 2008a,b). Similar to our study, the estimation of parameters $\boldsymbol{\beta}$ in satellite variational bias correction has been employed (Auligné et al. 2007). Following this procedure, we compute the parameters $\boldsymbol{\beta}$ jointly with the model state $\mathbf{x}$. Thus, the augmented model state $\mathbf{Z}$ can be expressed as

$$\mathbf{Z}^{\mathrm{T}} = \begin{bmatrix} \mathbf{x}^{\mathrm{T}} & \boldsymbol{\beta}^{\mathrm{T}} \end{bmatrix} \tag{7}$$

and the cost function (1) becomes

$$
\begin{aligned}
J(\mathbf{z}) = & (\mathbf{z} - \mathbf{z}_b)^{\mathrm{T}} \mathbf{B}_{\mathbf{z}}^{-1} (\mathbf{z} - \mathbf{z}_b) \\
& + [\mathbf{y} - \tilde{h}(\mathbf{z}, \mathbf{y})]^{\mathrm{T}} \mathbf{R}^{-1} [\mathbf{y} - \tilde{h}(\mathbf{z}, \mathbf{y})].
\end{aligned}
\tag{8}
$$

The assumption is made that the errors of parameters $\boldsymbol{\beta}$ are uncorrelated with the errors of the model state **x**. The background error covariance $\mathbf{B_z}$, therefore, can be expressed as a block-diagonal matrix:

$$\mathbf{B_z} = \begin{bmatrix} \mathbf{B_x} & 0 \\ 0 & \mathbf{B_\beta} \end{bmatrix}, \tag{9}$$

with $\mathbf{B_x}$ and $\mathbf{B_\beta}$ representing the background error covariance of model state **x** and parameters $\boldsymbol{\beta}$, respectively.

Two methods can be used to estimate $\mathbf{B_\beta}$. The analysis error of parameters $\boldsymbol{\beta}$ at the previous cycle can be regarded as the background error of parameters $\boldsymbol{\beta}$, which is estimable by the inverse Hessian of the cost function (8) (Zhu et al. 2015). The other feasible method is to use the averaged observational errors of observations involved, which indicates that the background of parameters $\boldsymbol{\beta}$ are assigned the same weights as these observations (Dee 2004; Sato 2007). In TAMDAR, each aircraft with a unique tail number is not guaranteed to take measurements at every analysis time; therefore, the latter method is employed in this study to avoid the impacts of flow-dependent parameter errors on extrapolating parameters.

c. Preconditioning

WRFDA employs the incremental formulation following Courtier et al. (1994). The background error covariance of model state is transformed by the preconditioning via the control variable in the incremental cost function (Barker et al. 2004). For the quadratic cost function (8), its condition number[2] determines the convergence properties of the minimization algorithm. To make the cost function rapidly converge to a solution, the condition number should be reduced. This can be achieved by introducing a preconditioning operator, which transforms the model state to the control variables, and meanwhile, forces the shape of the cost function in region of the control variables closer to spherical, so that the gradient of the cost function points toward the minimum as much as possible (Bouttier and Courtier 2002). The shape of the cost function (8) is described by its Hessian with respect to parameters $\boldsymbol{\beta}$:

$$\frac{\partial^2 J}{\partial \boldsymbol{\beta}^2} = \mathbf{B_\beta}^{-1} + \mathbf{H_\beta}^\mathrm{T} \mathbf{R}^{-1} \mathbf{H_\beta}, \tag{10}$$

where $\mathbf{H_\beta}$ is the derivative of the observation operator $\tilde{h}$ with respect to parameters $\boldsymbol{\beta}$ (i.e., the values of the predictors). Based on (10), when assuming that a bias model contains n predictors, and N observations from aircraft j with the observational error σ_o are involved in the variational bias correction procedure, the transformation from control variables to model variables can be expressed as

$$\mathbf{T}^j = \left[(\mathbf{B_\beta}^j)^{-1} + \frac{N}{\sigma_o^2} \mathbf{P} \right]^{-1/2}, \tag{11}$$

where **P** is an $n \times n$ matrix denoting the averaged covariance of predictors. This transformation $\mathbf{T}^j$ should be computed prior to minimization, and then multiplied by the control variables to complete the transformation process.

3. Experiment design

The Weather Research and Forecasting (WRF) Model (Skamarock et al. 2008) was run with 9-km grid spacing covering the eastern half of contiguous United States (Fig. 1) with a model top of 50 hPa. TAMDAR,[3] along with the datasets collected from the Global Telecommunication System (GTS),[4] which included surface land and marine reports, rawinsondes, atmospheric motion vectors, and the aircraft-based data, such as ACARS and AMDAR, were used in the data assimilation process. The rawinsonde, ACARS, and AMDAR temperature reports, well processed by internal quality control and bias correction procedures[5] in the National Centers for Environmental Prediction (NCEP), are also considered the anchor dataset during variational bias correction. The satellite-based data are not employed in this study in order to avoid introducing the local systematic errors in this small model domain. The experimental domain and examples of spatial and temporal distributions of TAMDAR and rawinsonde (raob) observations are shown in Fig. 1. Typically, TAMDAR provides two soundings, one during the ascent portion of the flight and one during the descent. This helps fill the spatial and temporal coverage gaps of upper-air temperature observations around the Charlotte, North Carolina, airport. TAMDAR from these regional aircraft complements RAOBs normally from day 1 at 1200 UTC to day 2 at 0000 UTC daily (0700–1900 eastern standard time).

---

[2] The condition number is defined as the ratio between the maximum and minimum eigenvalues of Hessian of cost function, which measures the ellipticity of the isosurfaces of cost function.

---

[3] Provided by Panasonic Avionics Corporation.

[4] Available in PrepBUFR format from NCEP.

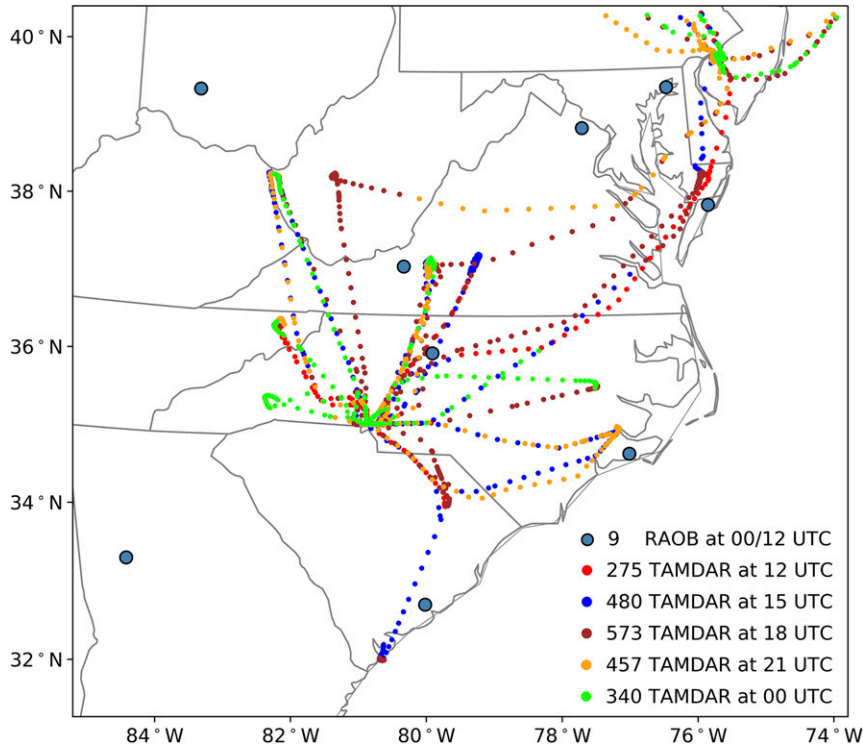[5] www.emc.ncep.noaa.gov/mmb/data_processing/prepbufr.doc/document.htm.

FIG. 1. The experimental domain and the spatial and temporal distributions of temperature observations of TAMDAR and rawinsonde (raob) at 1200 UTC 1 Feb–0000 UTC 2 Feb 2017 with a time window of ± 1.5 h. The UTC times are represented by colors. The observation count is shown by UTC time in the legend.

To evaluate the performance of TAMDAR temperature VarBC, parallel experiments were conducted. CTL is the control run, which assimilated observations from TAMDAR and GTS. ADR, MCN, and WMP are identical to CTL in every way except that they run the VarBC scheme by using the corresponding bias models. Additionally, we ran a version titled ADR05 and ADR10 by adjusting the coefficient a_{ADR} from 0.1 in (5) to 0.5 and 1.0 for the purpose of inspecting the impacts of predictor weights on bias model.

These experiments were started on 1 February 2017 and lasted six weeks ending on 15 March 2017. This assimilation employs the partial cycling data assimilation scheme as depicted in Fig. 2, which initializes at 1200 UTC each day and run five times per day (day 1 at 1200 UTC to day 2 at 0000 UTC) with 3-h cycling. In detail, the first data assimilation procedure in WRFDA starts at 1200 UTC in each day, and its background is a 3-h forecast produced from regional model WRF, which is initialized at 0900 UTC by the interpolation of Global Forecast System (GFS) short-term forecast. At other analysis time, the previous 3-h WRF model forecast is used as a background to produce a new analysis in WRFDA. The time window of the assimilation
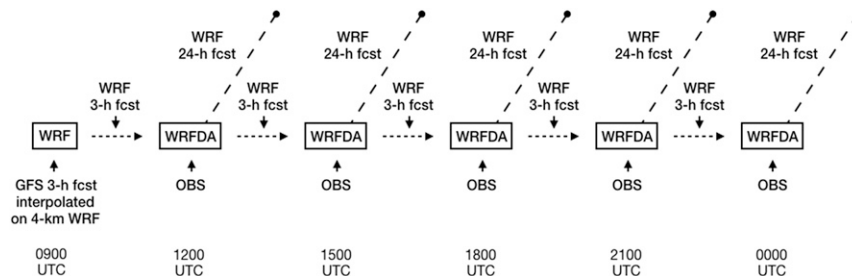
FIG. 2. Schematic of 3-h partial cycling data assimilation and forecast configuration.
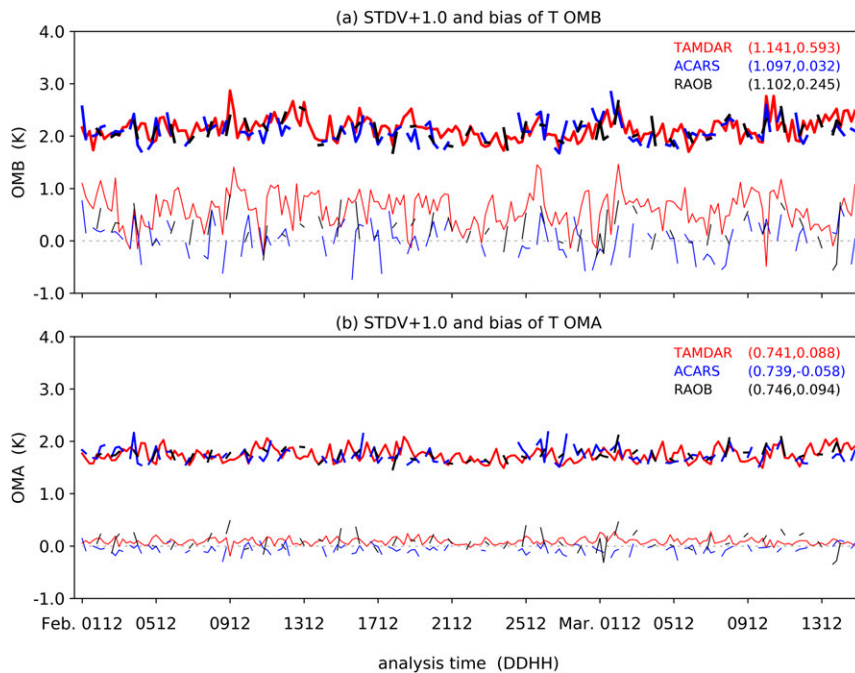
FIG. 3. The time series of standard deviation (STDV, thick lines) and bias (thin lines) of temperature (a) observation-minus-background (OMB) and (b) observation-minus-analysis (OMA) of TAMDAR (red lines), ACARS (blue lines), and raob (black lines) below 500 hPa in the control experiment (CTL). All STDV statistics are shown by initially adding 1.0 for better visualization. The average values of STDV and bias are summarized in the first and second columns of the legends, respectively. The discontinuous lines indicate that observations are not available at these analysis times.

is −1.5/+1.5 h on either side of the analysis time. All experiments generate 24-h forecasts for which the lateral boundary conditions are provided by NCEP 0.5° × 0.5° global forecasts. The partial cycling in daylight is employed partly because TAMDAR takes measurements principally in local daylight in order to monitor convective activity, and moreover, restarting regional cycling daily by using GFS short-term forecast may limit the accumulation of model errors derived from imperfect physics. It should be noted that the parameters kept updating consecutively for all six weeks of experiments, even while the partial cycling data assimilation scheme was being used.

4. Performance of VarBC

In this section, the results are presented in the following order: To begin, we will show the potential value of TAMDAR temperature observations. Second, the statistical relations between predictors and innovations are illustrated. Additionally, the evolution characteristics of terms in the bias models in these sensitivity experiments are contrasted and analyzed. A total of 13 aircraft are used in this study and the aircraft with tail

number 0751 is taken as a key example due to its good data continuity. Finally, the impacts of VarBC on analyses and ensuing forecasts are verified.

a. TAMDAR temperature OMB and OMA

The statistics of TAMDAR temperature observation-minus-background (OMB) and observation-minus-analysis (OMA) in CTL are presented in Fig. 3. The continuous data stream demonstrates that TAMDAR are the only available upper-air temperature observations during some of the analysis times in this region. The average standard deviation of TAMDAR is comparable to that of ACARS and raob for both OMB and OMA, which suggests that TAMDAR temperature observations are able to provide useful information for analysis. However, as evidenced by the bias changing from 0.593 K in OMB to 0.088 K in OMA, the analysis is contaminated because the biases in TAMDAR are almost fully transferred into the analysis during the data assimilation process.

b. Predictors in WMP bias model

To correct the bias, the key error sources dominating the systematic errors should be included and addressed
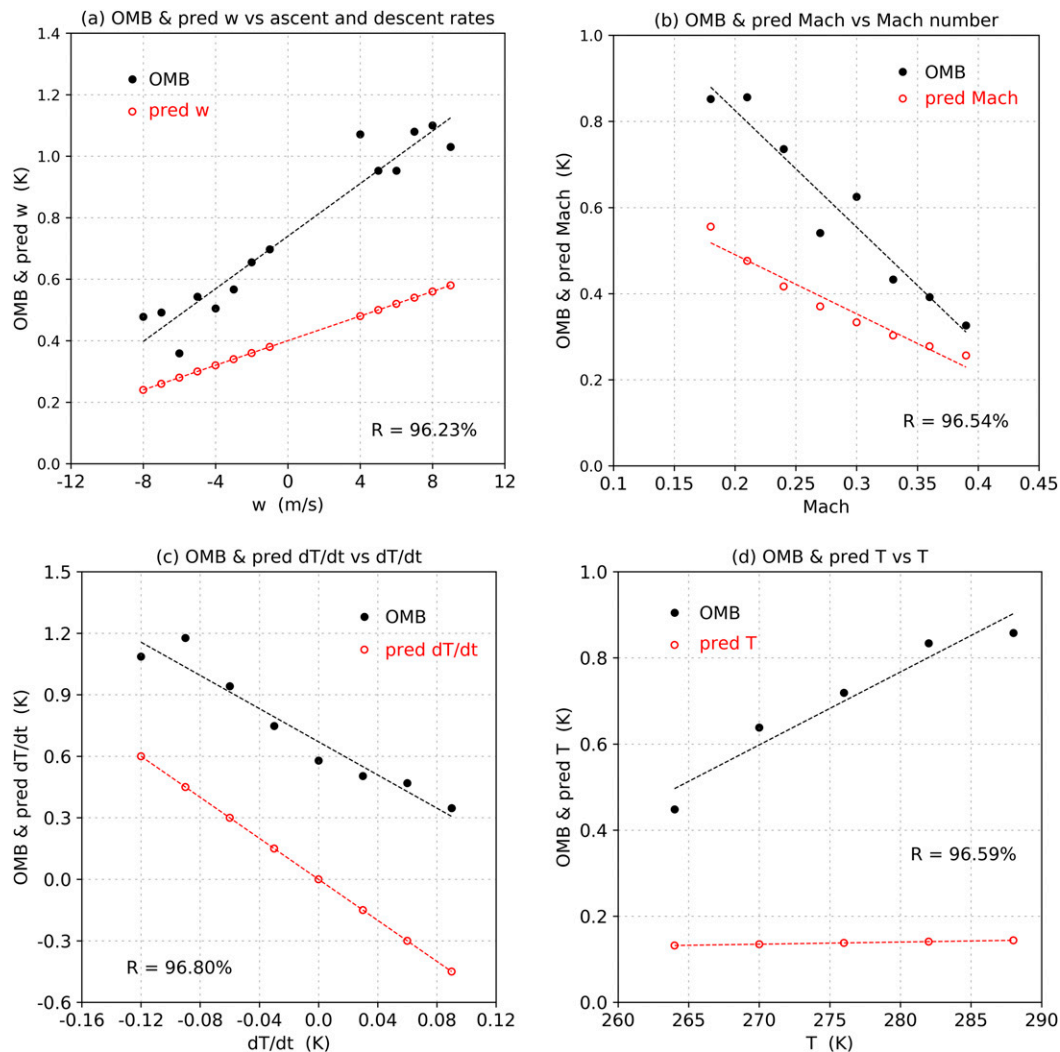
FIG. 4. Scatterplots of TAMDAR temperature observation-minus-background (OMB) of aircraft 0751 in control experiment (CTL) and predictors (pred) in the WMP bias model (12) vs (a) aircraft ascent and descent rates (w), (b) Mach number (Mach), (c) temperature tendency (dT/dt), and (d) the observed temperature (T). The black dots represent the OMB statistics and the red circles represent the predictors in bias model (12). The black and red dashed lines are the linear regression curves for OMB and predictors, respectively.

via a bias model in order to truly reflect the systematic errors in TAMDAR temperature measurements. In other words, the predictors in bias model need to contain variables which OMB statistics are rapidly responsive to. To investigate these predictors, the regression relations between OMB and some potential error sources are presented. Figure 4 shows scatterplots of TAMDAR temperature OMB versus aircraft height tendency (w), Mach number (Mach), temperature tendency (dT/dt), and the observed temperature (T), respectively. The black dot represents the average of a set of data which fall within a box binned according to the variable of the x axis. Any dot is marked only when the number of data falling in a binned box exceeds 100.

Generally, OMB is approximatively proportional to height tendency (Fig. 4a) and inversely proportional to temperature tendency (Fig. 4c). It is expected because the reading time lag produces a warm bias during ascent and a cold bias during descent. It should be noted that positive w and negative dT/dt (except temperature inversion) correspond with ascent, and vice versa. Moreover, aircraft maneuvering and thermal hysteresis are more frequent and sharper during ascent than descent, especially given that ascents typically involve both steep vertical angles and greater acceleration. Thus, larger warm bias is generated during ascent than descent.

These two linear regressions demonstrate that bias in TAMDAR temperature observation is rapidly responsive

to the tendencies of height and temperature. In reality, the temperature tendency is not independent from the height tendency due to the adiabatic lapse rate of air. It seems straightforward enough to consider either height tendency or temperature tendency as a predictor. However, this becomes unwieldy during cruise of aircraft. The tendencies of height and temperature are approximately zero in average during cruise when light ascent and descent maneuvers are conducted (i.e., $\langle w \rangle \sim 0$, $\langle dT/dt \rangle \sim 0$). It is understandable that the contribution of reading time lag (denoted by dT/dt) to bias during cruise is negligible. However, as seen in Figs. 4a and 4c, when the tendencies of height and temperature are equal to zero, OMB still reach up to 0.740 and 0.579 K, respectively. These values far exceed the typical errors which are estimable by the equation related to Mach number effect (World Meteorological Organization 2003). Apparently, besides the reading time lag and air heating caused by adiabatic compression (represented by Mach number), aircraft maneuvering and thermal hysteresis also do contribute to the bias. Following that, a reasonable interpretation is that the extra warm bias in addition to the effect of reading time lag and Mach number is mainly derived from the thermal lag of the metal probe and unstable measurements during aircraft maneuvers which can alter the airflow, even during cruise. Thus, the contribution of aircraft maneuver (represented by w) to bias must be considered effect accumulated, that is, the contribution should not be offset during cruise when light ascent (positive w) and descent (negative w) are averagely conducted. To achieve this, we define an initial predictor containing height tendency as $(0.1w + 2)$. In addition, the linear regression function of OMB statistics versus height tendency cannot be directly employed for predictors because the OMB statistics contains biases in both observations and model. Here we assume the magnitude of each predictor contributing to 60% of OMB statistics in average. In the case of height tendency, a coefficient of 0.2 applies for the initial predictor, and the final predictor becomes $(0.1w + 2) \times 0.2$. The change of predictor with height tendency is shown in red circles in Fig. 4a. In fact, the simulation by predictor is not strictly 60% of OMB statistics because the simple coefficient is employed in predictors. Depending on locations and times of measuring, the percent might be lower or higher than truth, and thus, not optimal.

In Fig. 4b, the linear regression between OMB and Mach number results from several aspects, such as Mach number errors, incomplete correction of the air heating from adiabatic compression, thermal lag, and other possible factors. It is complicated to describe these possibilities and their interaction in a bias model.

In general, Mach number increases with the flight altitude of aircraft. More observations with large temperature OMB tend to appear at lower Mach numbers, which often correspond to the start of ascent and end of descent with large height tendency (not shown). Following that, the predictor related to Mach number simply employs the inverse proportion relation to OMB and is scaled as $(10 \text{ Mach})^{-1}$, denoted by the red circles in Fig. 4b. The magnitude of predictor is 64.84% of OMB on average.

Conversely, the predictor of the observed temperature (T) needs more consideration. As we discussed above, the larger OMB typically corresponds with lower Mach number, which is often located at lower height with higher temperature. Thus, it is somewhat expected to see that OMB increases with the observed temperature (Fig. 4d). Because of the adiabatic lapse rate of air, temperature is a function of height and thus approximately equivalent in the predictor sense to pressure, which was used as predictor in Zhu et al. (2015). However, this temperature predictor doesn't perform well in correcting bias during ascent and descent because temperature observations are averaged vertically in order to calculate the predictor. The temperature average is not responsive to OMB in a season. Therefore, we consider the temperature predictor only as a constraint in this study, which changes interseasonally, by scaling the observed temperature by a large number of 2000.

Based on these discussions above, a bias model containing the weighted multiple predictors is designed in (12):

$$b_{\text{WMP}}(\boldsymbol{\beta}, \mathbf{y}) = \beta_1 + \beta_2 (0.1w + 2) \times 0.2 + \beta_3 (10M)^{-1}$$
$$+ \beta_4 \frac{dT}{dt} \times 5 + \beta_5 \frac{T}{2000}, \qquad (12)$$

where T is the observed static temperature, and dT/dt is the temperature tendency between two consecutive measurement locations. The fourth term expresses the contribution of data reading time lag to the overall bias, and the fifth term relates the temperature being measured to the sensitivity of the sensor to the atmospheric environment or seasonal factors (Ballish and Kumar 2008). Within a given season, the last predictor should become stable, and thus may be approximately equivalent to certain variables that scale nearly linearly with height (e.g., pressure). Indeed, pressure was explicitly used as a predictor in Zhu et al. (2015). Additionally, other existing uncertainties, such as calibration and correction for exposure, de-icing, and wet evaporation of probe, are unlikely to be estimated during minimization. As a result, these uncertainties are expected to be reflected by the constant-predictor term. These coefficients
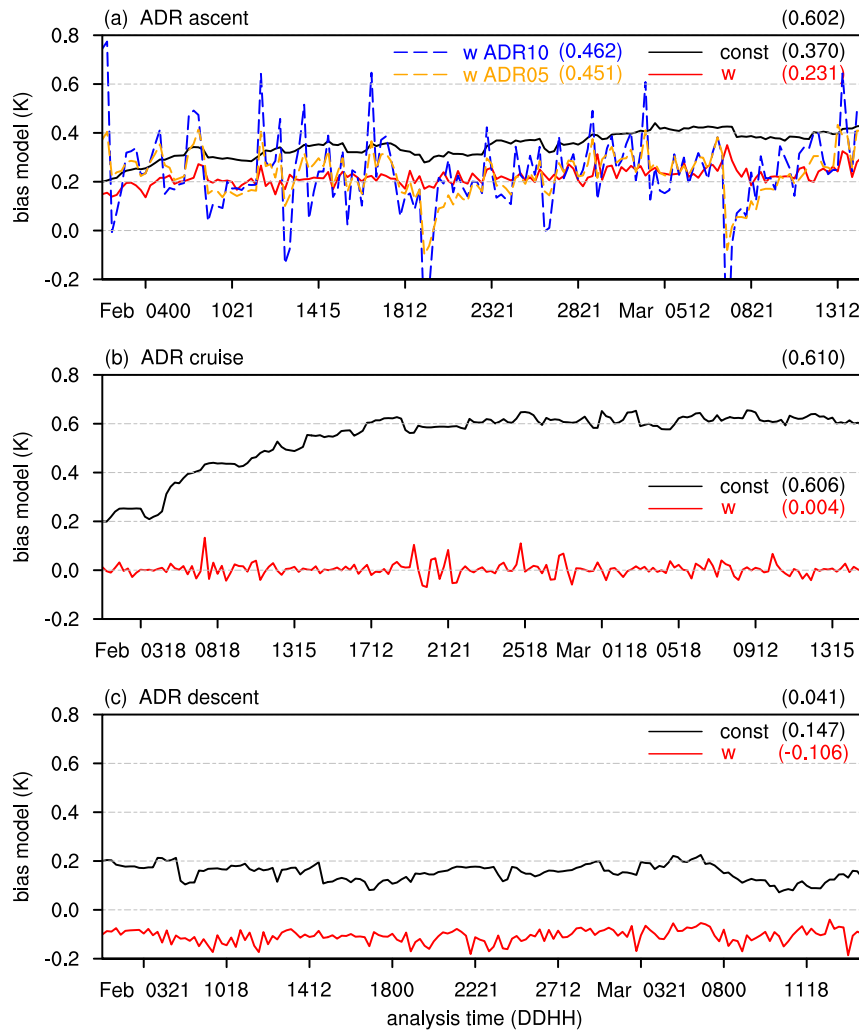
FIG. 5. The evolution of terms in the ADR bias model (5) during (a) ascent, (b) cruise, and (c) descent for aircraft 0751. The const and *w* series correspond to terms in bias model (5), for predictors of constant and vertical motion rates (represented by height tendency), respectively. The average value of each term from 1200 UTC 14 Feb to 0000 UTC 15 Mar 2017 is shown in legend. The average bias model is shown on top. The blue and orange dashed lines in (a) represent the *w* terms in bias model (5) in ADR10 and ADR05, scaled by 1/2, respectively.

in (12) are given based on the simple regression relations between OMB and predictors in direct or inverse proportion, therefore, are not optimal. In fact, one can correct the bias by vertical levels based on the regression relation, when/where a large sample of aircraft are available to take measurements at each target height level and each aircraft is equipped with the unified sensor with same measurement accuracy.

c. Initialization and cycling of bias models

In the bias models, the predictors are functions of aircraft measurements while the parameters need to be obtained by minimizing the cost function (8). In the

absence of any prior information, all parameters are initially set to 0.2, which, during the first data assimilation window, will guarantee a bias correction of approximately 0.2 K depending on the phase of flight. The update of parameter is independent from its first guess in incremental VarBC, in other words, the obtained parameter is not affected by the initial estimation value. However, using a positive number as an initial estimation of parameters can decrease the time spent in cycling these parameters since the positive bias in aircraft temperature observations is well addressed.

Figure 5 shows the evolution of terms in the bias models in ADR for aircraft 0751 at the phases of ascent,

cruise, and descent. Overall, the parameters adjust rapidly in the first few days of the run and then settle down gradually. The statistics of OMB in Fig. 4a demonstrate that the bias during ascent is larger than during descent, and the cruise bias is intermediate between the ascent and descent biases. However, due to averaged height tendency (w) being approximately equal to zero during cruise, the bias correction represented by w in ADR, which is only 0.004 K, is evidently underestimated (Fig. 5b). As a result, almost all error sources have to be reflected by the constant-predictor term, leading to a model that cannot discern the physical significance of particular contributors to the bias. A potential weakness is that bias correction could compensate for model errors under the weak constraints of the predictor. Such evidence can be found by the larger bias correction at cruise (0.610 K) than ascent (0.602 K), which conflicts with the OMB statistics in Fig. 4a.

To inspect the impacts of scaling the predictor of w on bias model, we run ADR05 and ADR10 by adjusting a_{ADR} from 0.1 in ADR to 0.5 and 1.0. The results using aircraft 0751 at ascent demonstrate that compared with ADR, the uncertainty of the w term increases significantly in ADR05 and ADR10 (Fig. 5a). The unsteady bias correction in ADR05 and ADR10 makes it less likely to obtain statistically meaningful estimates of parameters by long-term cycling data assimilation (often overestimating), and thus, it fails to use these estimates to extrapolate bias corrections accurately into the next cycle.

It must be mentioned that one reviewer considered that the results in Fig. 5 suggested that the w predictor was very noisy and that smoothing w, as in Zhu et al. (2015), should be tried. This algorithm did achieve the goal in Zhu et al. (2015). However, the predictor of w is vertically averaged for one sounding (ascent or descent) prior to estimating the bias model in this study. The vertically averaged ascent/descent rates are not sensitive to the smoothing algorithm. Therefore, we reserve our views about this issue for further investigation.

By comparing ADR with MCN in Fig. 6, the bias corrections are comparable, especially during the phases of cruise and descent. Although the second terms in the bias models demonstrate differences of 0.1 K at cruise and 0.145 K at descent between MCN and ADR, the constant-predictor terms take up most of the difference. VarBC does not appear to be sensitive to the second predictors used in ADR and MCN. Moreover, as discussed above, aircraft normally have larger Mach number during cruise than ascent and descent. However, the bias corrections represented by the Mach term are larger during cruise than ascent and descent, which is opposite to the linear regression relation of OMB

versus Mach number presented in Fig. 4b. The results indicate that VarBC with a plain bias model is prone to being dominated by other factors and has little capability to automatically correct the problematic statistical relationship between the systematic errors and the predictors described in the bias model.

In WMP, as shown in Fig. 7, the height tendency (w) plays the leading role among predictors except during descent. Compared with the negligible bias correction represented by w in ADR (Fig. 5b), the bias correction represented by w in WMP suggests that the bias at cruise (Fig. 7b) reflects the accumulated biases during light ascent and descent phase of flight. It also demonstrates that the contribution of the reading time lag term (dT/dt) to the bias correction is well fit to its typical value in each flight phase. However, concern was raised by one reviewer that the height tendency w and temperature tendency would be highly anticorrelated making it unsuitable to use them together as predictors. In reality, the correlation coefficient is -0.33 for ascents and -0.13 for descents: making their use together acceptable but again raising questions about noise in the predictors.

Unlike ADR and MCN, the bias correction is slightly smaller at cruise than that at ascent in WMP, which is expected because the contributions from Mach number and reading time lag predictors are slightly reduced at cruise. These results demonstrate that the bias model (12) captures the bias in TAMDAR temperature observations qualitatively and quantitatively better than (5) and (6).

Although the same Dash-8 turboprops aircraft are used in this study, the bias corrections are different for these aircraft according to the error sources at the measuring time. In WMP, the minimum correction is 0.139 K for aircraft 0862 while the maximum correction is 0.69 K for aircraft 1004. Nonetheless, the evolutions of bias models are qualitatively consistent among these aircraft.

d. Impacts of VarBC on analyses and forecasts

Figures 5–7 indicate that these bias models become stable after the first two weeks of cycling. From that model time, we then verify analyses and forecasts for another four weeks from 1200 UTC 14 February to 0000 UTC 15 March 2017.

The statistics of TAMDAR temperature OMB in CTL, ADR, MCN, and WMP and the TAMDAR observation count in CTL are presented in Fig. 8. Compared with CTL, the application of VarBC corrects the warm bias at all height levels where TAMDAR observations are present, and WMP further reduces the bias by 0.06° and 0.05°C compared with ADR and MCN, respectively. The random errors in observations
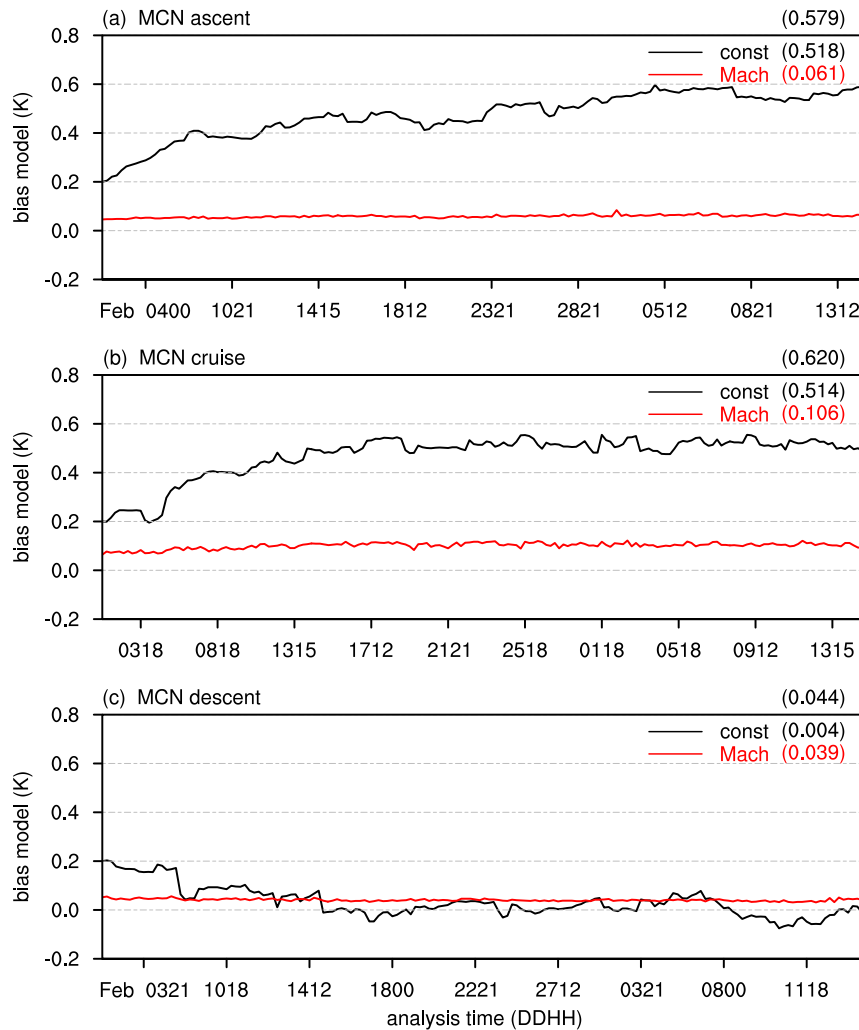
FIG. 6. As in Fig. 5, but for MCN.

(taken from the standard deviation of OMB) are reduced as well because the temperature observations deviating substantially from the background are made considerably more accurate after employing the VarBC. The black dots in left panel suggest that the difference of OMB between WMP and CTL is significant at or above the 95% confidence level by applying the t test.

The temperature increments produced by VarBC not only generate increments on other model variables but also can be extended beyond the original observation locations by using the information in the background error covariance matrix. Figure 9 shows the differences of the temperature analysis in WMP from CTL at 1200 UTC 1 March 2017. Although the observations are distributed along narrow flight tracks, the increments produced by VarBC spread out smoothly and extend to a wider area in both horizontal and vertical

dimensions. This information transmission is important for enlarging the influence range of these observations. Figure 10 presents the statistics of ACARS and TAMDAR temperature OMA in CTL, ADR, MCN and WMP. The analyses become fit closer to both ACARS and TAMDAR observations by employing VarBC, and the application of WMP bias model further reduce the mean analysis differences as expected. The differences produced by VarBC are visible in 9-h forecasts (Fig. 11), however, these improvements are wiped out in this small domain after 9 h when the airflow from lateral boundary starts to move in the verification domain.

5. A cold frontal rainfall event

In this section, TAMDAR temperature VarBC is investigated using a cold frontal rainfall event that occurred
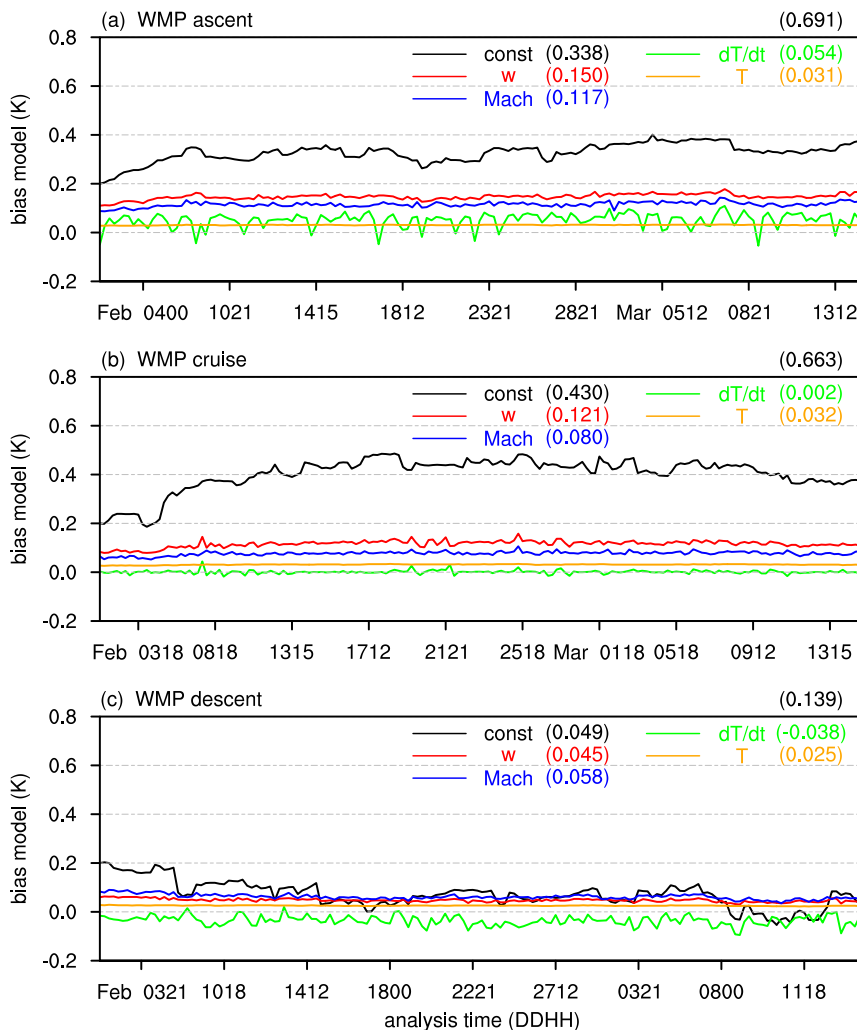
FIG. 7. As in Fig. 5, but for WMP.

during the period of the six-week experiments. We explore how VarBC affects the prediction of precipitation location and intensity. The precipitation forecast is verified against the NCEP stage IV analysis with 4-km resolution, which is a mosaic from the regional hourly/6-hourly multisensor (radar + gauges) precipitation analyses produced by the 12 River Forecast Centers over the continental United States (Lin and Mitchell 2005). The synoptic circulation situation is described, and then the results are presented.

a. Synoptic circulation

A strong stream of cold and dry air from the west flowed nearly perpendicular to isotherms toward the east-central United States, joining with the southwest warm and wet jet stream by 0000 UTC 2 March 2017 (Fig. 12a). The dewpoint deficit was less than 2.8 K

within the wet core area of $10 \, \mathrm{g \, kg^{-1}}$ indicated by blue lines. At that time, a cold front formed along the North Carolina (NC)/Tennessee (TN) border, extending through southwestern Virginia (VA). The simulated maximum reflectivity (not shown) demonstrates that some mesoscale convective cells were generated ahead of the cold frontal surface along the southwest air mass. The cold front and the transverse trough were intensified by 0600 UTC 2 March 2017 (Fig. 12b) which sharply strengthened the temperature decrease along the front near the western border of NC and VA. The stage IV analysis shows that the precipitation was distributed mainly downstream of TAMDAR coverage along a band which is oriented northeast–southwest in central NC and northern SC. Based on this weather pattern and location, the TAMDAR temperature VarBC is expected to have impacts on precipitation forecasts.
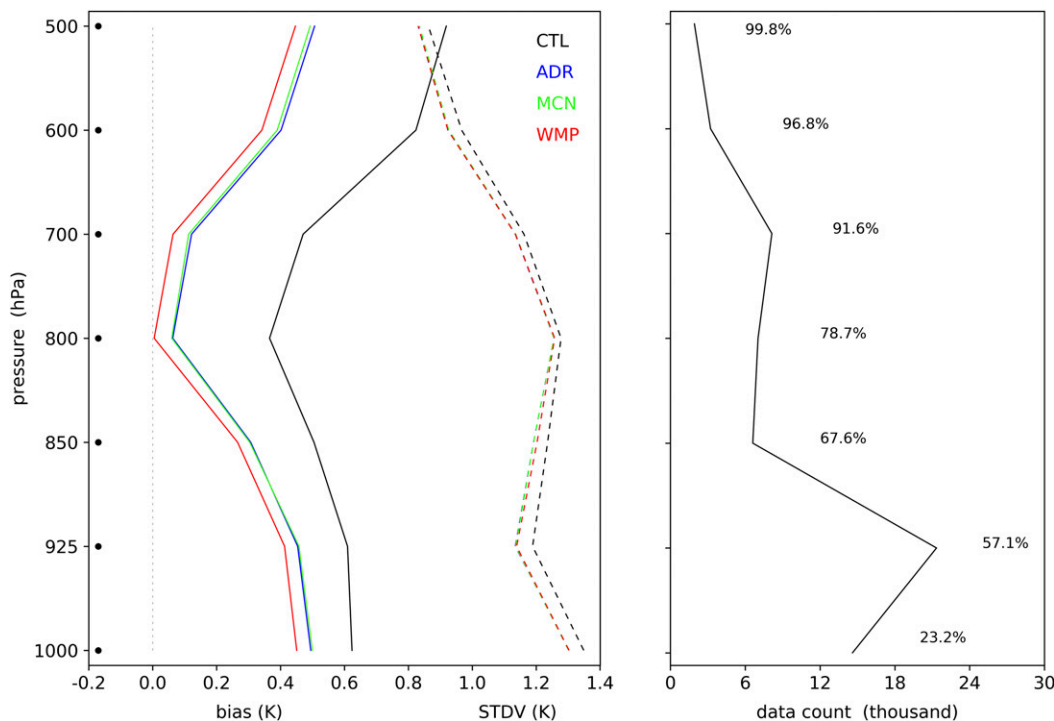
FIG. 8. (left) The vertical profiles of standard deviation (STDV, dotted lines) and bias (solid lines) of TAMDAR temperature observation-minus-background (OMB) in CTL, ADR, MCN, and WMP, and (right) the count of TAMDAR temperature observations assimilated in CTL from 1200 UTC 14 Feb to 0000 UTC 15 Mar 2017. The accumulated count percentage from the surface level in CTL is presented in the right panel. The black dots in the left panel suggest that the difference of OMB between WMP and CTL is significant at or above the 95% confidence level by applying the t test.

b. Impacts of VarBC on analyses and forecasts of wind and humidity

The typical analysis increments of wind and relative humidity (RH) produced by TAMDAR temperature VarBC are presented in Fig. 13. As shown in Fig. 2, the same background forecast is used as a first guess in data assimilation procedure in different experiments at each 1200 UTC instance, therefore, the differences of analyses in Fig. 13 are completely generated by the application of VarBC.

The background error covariance matrix employed does not include the covariances between specific humidity and other model variables. Therefore, the temperature increments derived from temperature VarBC cannot impact the specific humidity analysis in WMP. As a result, as shown in Fig. 13, RH increases in the environment, which becomes cooler after warm bias is corrected (Fig. 9), when specific humidity in the analysis is unchanged. In general, the increased RH may intensify the prediction of precipitation intensity in WMP because the threshold for RH to reach a condition for precipitation will be easier to reach in WMP than CTL. It is also seen that anticyclonic wind increments are

produced, which reflect the geostrophic adjustment of wind to the temperature correction. Although the magnitude of these wind increments ia small in the analysis, the information is able to produce a notable influence on the wind forecasts in a nonlinear forecast model.

c. Verification of precipitation prediction

The stage IV analysis and 15 mm isolines of 6-h accumulated precipitation forecasts in CTL and WMP, valid 0600 UTC 2 March 2017, are presented in Fig. 14. Overall, CTL and WMP both capture the main observed characteristics of precipitation distribution. From the isolines of 15-mm precipitation, the precipitation location in central North Carolina predicted in CTL has a southward displacement error, which is noticeably reduced in WMP. Based on a point-to-point threat score (TS; Schaefer 1990) and a neighborhood-based fractions skill score (FSS; Roberts and Lean 2008) (not shown here), WMP outperforms CTL at thresholds of 25 mm and less while there is little or no skill for precipitation amounts of 30 mm and more because both CTL and WMP have eastward displacement errors for the
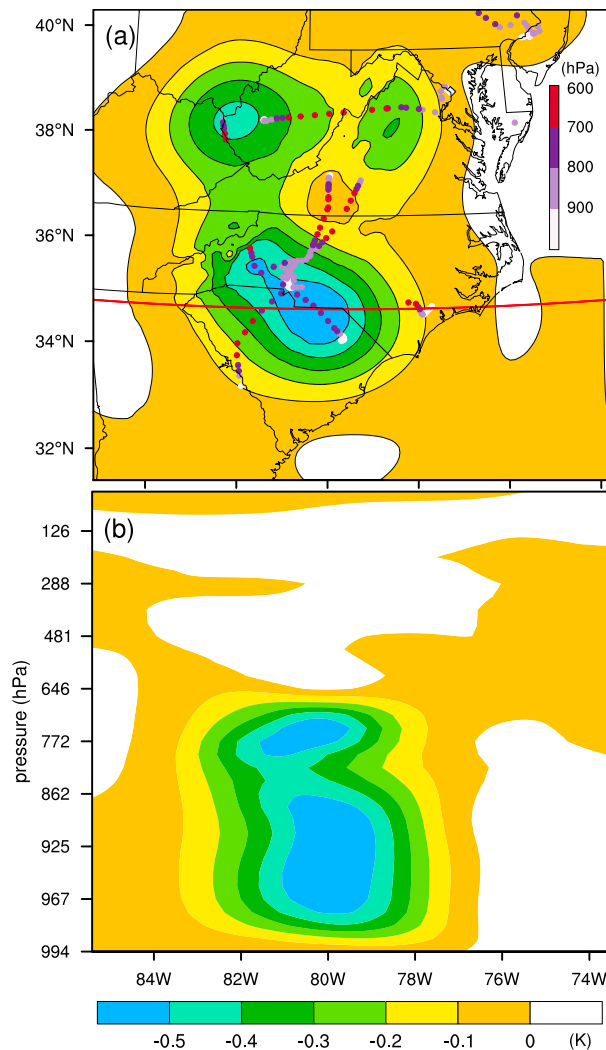
FIG. 9. (a) The departure of the temperature analysis in WMP from CTL (WMP minus CTL) at 1200 UTC 1 Mar 2017 at 750 hPa; (b) the pressure–longitude cross section along 34.8°N illustrated by a red solid line in (a). The TAMDAR observation locations are indicated by dots and the observational pressures are denoted by the vertical color bar in (a).

observed maximum precipitation area located in northern South Carolina (SC). Although the improved wind by WMP intensify the southwesterly jet which, combined with the RH increments generated by VarBC, contributes to the RH differences of up to 30%, precipitation is still underestimated in this region by WMP.

The results demonstrate that in this case, the implementation of temperature VarBC improves the prediction of precipitation location and intensity mainly by improving the forecasts of wind-related synoptic circulation and humidity. However, the positive RH increments are always produced in a colder analysis after using VarBC as discussed above. The potential

weakness is that impacts of VarBC on analysis and forecast may change by cases. To optimize the application of VarBC, the background error covariance matrix containing the covariance between specific humidity and other model variables should be provided in data assimilation procedure. The hybrid ensemble-variational data assimilation (Wang et al. 2008a,b; Schwartz et al. 2015) and the application of multivariate background error statistics (Chen et al. 2013) may overcome the persistent increase of RH in the analysis due to the correction of the warm bias in TAMDAR temperature observations.

6. Discussion

This paper investigates the performance of variational bias correction for TAMDAR temperature observations in WRFDA model. The potential benefit of this method is evaluated over the eastern half of the CONUS for a six-week period by using TAMDAR observations collected around Charlotte airport in North Carolina. The indirect impacts of the VarBC on precipitation prediction are explored using a cold front precipitation event.

Besides using the predictor of aircraft height tendency, which is similar to methods used elsewhere in the NCEP and ECMWF global data assimilation systems, we try extra predictors related to the well-known error sources, such as Mach number, the temperature tendency between two consecutive measuring locations, and the static temperature measured by aircraft. The four predictors together are expected to represent the contributions of aircraft maneuvers, incomplete correction of air heating caused by adiabatic compression in sensor chamber, data reading time lag, and thermal hysteresis of the metal. Although these ''bad'' data with roll angles exceeding a certain threshold during frequent maneuvers have been eliminated from the dataset, the remaining ''bad'' data with light maneuvers and roll angles are still used which contribute to the bias.

Results demonstrate that the VarBC approach works well by automatically correcting TAMDAR temperature biases, and produces better analyses and short-range forecasts, even when using a relatively simple bias model. However, some weaknesses are noticed in the basic bias model formulations. For example, when using height tendency as a predictor, the error sources related to height tendency cannot be explicitly expressed during cruise due to the averaged predictor value being approximately equal to zero. Thus, the ability of VarBC to identify the source of bias in innovations becomes weak and could make the bias correction compensate
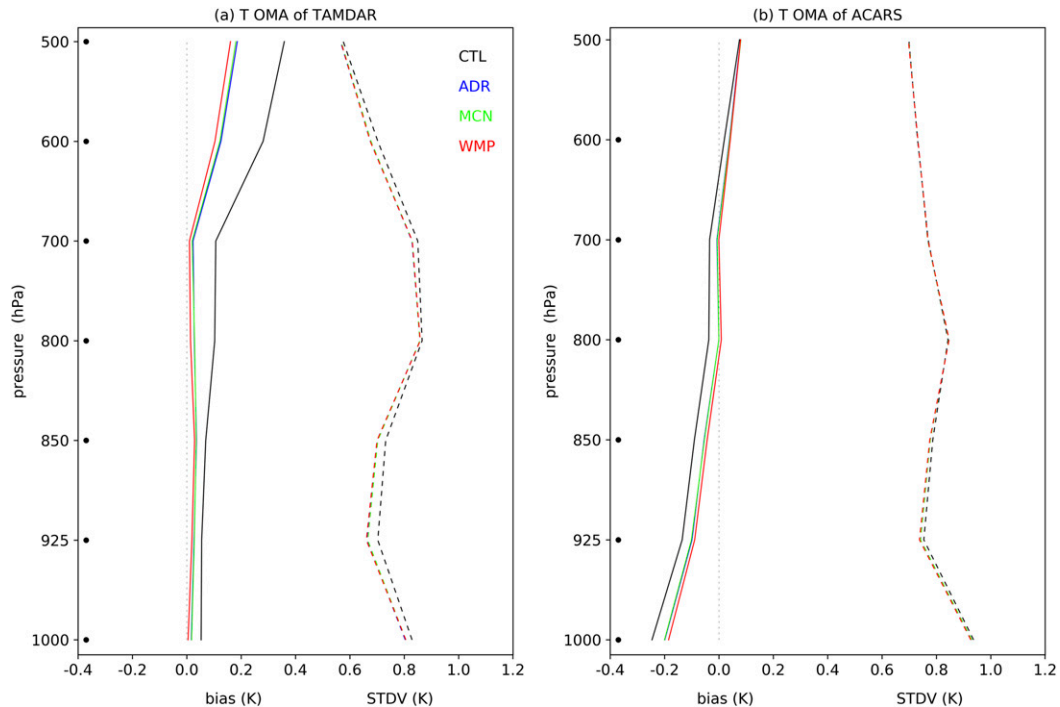
FIG. 10. The vertical profiles of standard deviations (STDV, dotted lines) and biases (solid lines) of temperature observation-minus-analysis (OMA) in CTL, ADR, MCN, and WMP for (a) TAMDAR and (b) ACARS from 1200 UTC 14 Feb to 0000 UTC 15 Mar 2017. The black dots suggest that the differences of OMA of ACAR and TAMDAR between WMP and CTL are significant at or above the 95% confidence level by applying the *t* test, respectively.

for model errors instead. Further, when using Mach number as a predictor, the bias corrections represented by the Mach number term at different phases of aircraft flight fail to match the linear regression relation of OMB versus Mach number. This suggests that VarBC has little capability to correct inaccurate statistical relations between variables using such a simple bias model.

To remove these weaknesses, the bias model (12) is designed, which contains the weighted multiple predictors and describes the linear regression relations between innovations and predictors. Results demonstrate that this bias model removes bias in TAMDAR temperature observations qualitatively and quantitatively better than the simpler bias models. As expected, TAMDAR temperature innovations are further reduced, and the analyses and short-range forecasts are improved.

A tricky practice conducted for scaling ascent/descent rates is to use a constant α_{ADR} in bias models (5) and (12). To make these predictors match with other predictors in magnitudes, a scaling process is necessary. However, the absolute magnitudes of height tendency at different analysis time do matter to the variational bias correction. Thus, the normalization method is

inappropriate in this case. One reviewer suggested that a smooth algorithm should be employed to reduce the noises in the predictors of ascent/descent rates. Future work will focus on refining the ability
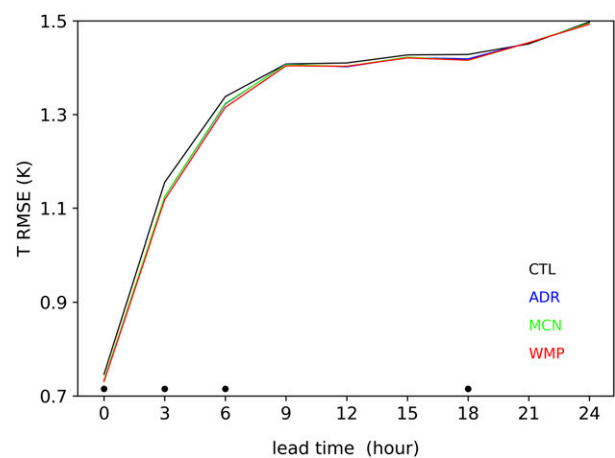

FIG. 11. The rmse of temperature forecasts in CTL, ADR, MCN, and WMP fit to ACARS temperature observations by forecast lead time from 1200 UTC 14 Feb to 0000 UTC 15 Mar 2017. The black dots suggest that the difference of temperature rmse between WMP and CTL is significant at or above the 95% confidence level by applying the *t* test.
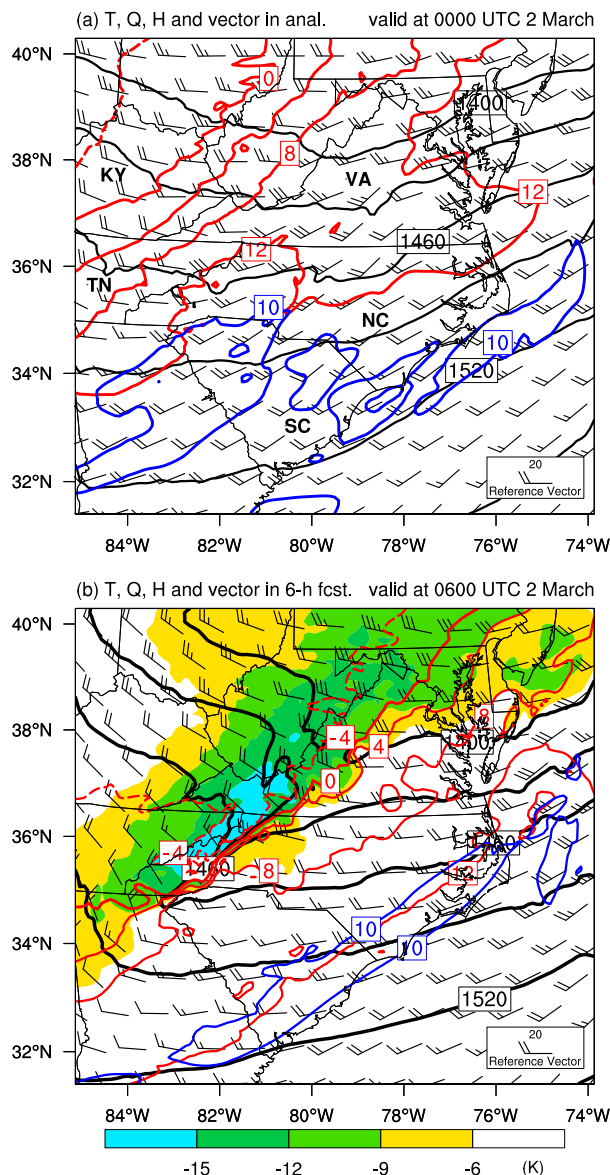
FIG. 12. Temperature (°C, red lines), specific humidity (g kg^{-1}, blue lines), height (m, black lines), and wind vectors at 850 hPa in (a) analysis and (b) 6-h forecast in WMP, initialized at 0000 UTC 2 Mar 2017. The predicted 6-h temperature changes are shaded in (b).

of a smooth algorithm to describe the predictors of ascent/descent rates.

The indirect impacts of VarBC on precipitation prediction are examined for a cold front precipitation event that occurred during the experimental period. VarBC is able to improve the forecasts of precipitation location and intensity by improving the prediction of wind-related synoptic circulation and humidity. However, due to the employment of the background error covariance, which does not include the covariance between specific humidity and other model variables, positive relative
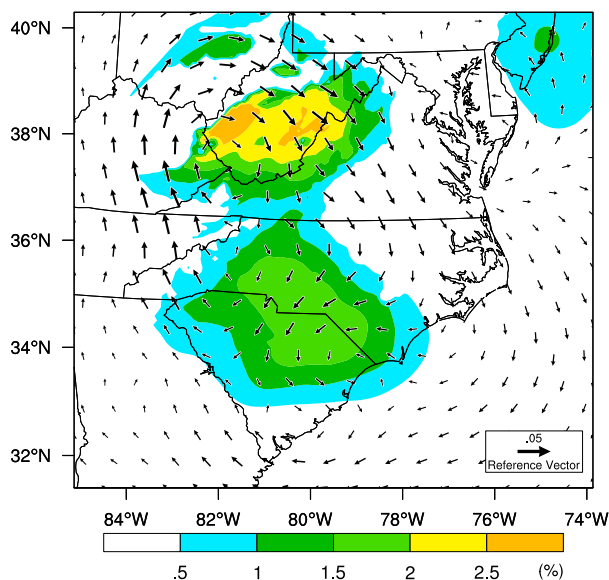
FIG. 13. The differences of analyses of relative humidity (%, shaded) and wind vectors at 850 hPa between WMP and CTL at 1200 UTC 1 Mar 2017.

humidity increments are always produced by the correction of warm bias in TAMDAR temperature observations. As a result, VarBC has capability to improve precipitation prediction for some cases but also may overestimate precipitation for other cases. To extract the maximum benefit from VarBC, a background error covariance matrix containing the covariance between
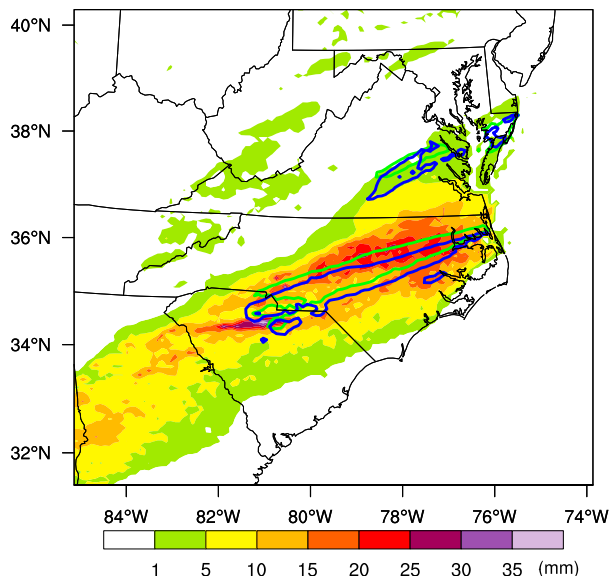

FIG. 14. The stage IV precipitation analyses (shaded) and 15 mm isolines of 6-h accumulated precipitation forecasts in CTL (blue lines) and WMP (green lines), valid at 0600 UTC 2 Mar 2017.

specific humidity and other model variables should be employed.

Overall, these results are encouraging; however, the bias model with the weighted multiple predictors in (12) is not optimal when it comes to the linear regression and the coefficients. It is practical to apply more sophisticated regression relations between innovations and predictors by a bias model or advanced methodology.

Another way to improve the representativeness of systematic error sources could be to use extra predictors, for example, the maneuvers, which usually involve a substantial roll angle, produce significant errors; or humidity, which could lead to a cold bias when any evaporation takes place during the period of air heating in the sensor chamber. Therefore, the roll angle of aircraft and humidity may further improve the modeling of bias. It is noted that the humidity changes rapidly with height, and therefore, it is not straightforward to define a coefficient for such a predictor.

Besides the Charlotte airport in North Carolina, TAMDAR also provides observations for some other major hubs over the CONUS (e.g., New York, Philadelphia, Seattle, etc.), as well as Alaska, Mexico, Europe, and Asia. Local climates, geographic environments, and aircraft types may cause additional diverse systematic biases, which would require considerable care in building bias models with more general applicability. Therefore, VarBC may be more useful when implemented on a regional basis. In addition, a study over a larger domain may improve the understanding of applicability of VarBC.

REFERENCES

Auligné, T., P. A. McNally, and D. P. Dee, 2007: Adaptive bias correction for satellite data in a numerical weather prediction system. *Quart. J. Roy. Meteor. Soc.*, **133**, 631–642, https://doi.org/10.1002/qj.56.

Ballish, B. A., and V. K. Kumar, 2008: Systematic differences in aircraft and radiosonde temperatures: Implications for NWP and climate studies. *Bull. Amer. Meteor. Soc.*, **89**, 1689–1708, https://doi.org/10.1175/2008BAMS2332.1.

Barker, D. M., W. Huang, Y.-R. Guo, A. J. Bourgeois, and Q. N. Xiao, 2004: A three-dimensional variational data assimilation system for MM5: Implementation and initial results. *Mon. Wea. Rev.*, **132**, 897–914, https://doi.org/10.1175/1520-0493(2004)132<0897:ATVDAS>2.0.CO;2.

Benjamin, S. G., B. E. Schwartz, and R. E. Cole, 1999: Accuracy of ACARS wind and temperature observations determined by collocation. *Wea. Forecasting*, **14**, 1032–1038, https://doi.org/10.1175/1520-0434(1999)014<1032:AOAWAT>2.0.CO;2.

Bouttier, F., and P. Courtier, 2002: Data assimilation concepts and methods. Meteorological training course lecture series, ECMWF, Reading, United Kingdom, 59 pp., https://www.ecmwf.int/en/elibrary/16928-data-assimilation-concepts-and-methods.

Cardinali, C., L. Isaksen, and E. Andersson, 2003: Use and impact of automated aircraft data in a global 4DVAR data assimilation system. *Mon. Wea. Rev.*, **131**, 1865–1877, https://doi.org/10.1175//2569.1.

Chen, Y., S. R. H. Rizvi, X.-Y. Huang, J. Min, and X. Zhang, 2013: Balance characteristics of multivariate background error covariances and their impact on analyses and forecasts in tropical and Arctic regions. *Meteor. Atmos. Phys.*, **121**, 79–98, https://doi.org/10.1007/s00703-013-0251-y.

Courtier, P., J.-N. Thépaut, and A. Hollingsworth, 1994: A strategy for operational implementation of 4D-Var, using an incremental approach. *Quart. J. Roy. Meteor. Soc.*, **120**, 1367–1387, https://doi.org/10.1002/qj.49712051912.

Dee, D. P., 2004: Variational bias correction of radiance data in the ECMWF system. *Proc. ECMWF Workshop on Assimilation of High Spectral Resolution Sounders in NWP*, Vol. 28, Reading, United Kingdom, ECMWF, 97–112.

——, 2005: Bias and data assimilation. *Quart. J. Roy. Meteor. Soc.*, **131**, 3323–3343, https://doi.org/10.1256/qj.05.137.

——, and S. Uppala, 2009: Variational bias correction of satellite radiance data in the ERA-Interim reanalysis. *Quart. J. Roy. Meteor. Soc.*, **135**, 1830–1841, https://doi.org/10.1002/qj.493.

Derber, J. C., and W.-S. Wu, 1998: The use of TOVS cloud-cleared radiances in the NCEP SSI analysis system. *Mon. Wea. Rev.*, **126**, 2287–2299, https://doi.org/10.1175/1520-0493(1998)126<2287:TUOTCC>2.0.CO;2.

Drüe, C., W. Frey, A. Hoff, and T. Hauf, 2008: Aircraft type-specific errors in AMDAR weather reports from commercial aircraft. *Quart. J. Roy. Meteor. Soc.*, **134**, 229–239, https://doi.org/10.1002/qj.205.

Eyre, J. R., 2016: Observation bias correction schemes in data assimilation systems: A theoretical study of some of their properties. *Quart. J. Roy. Meteor. Soc.*, **142**, 2284–2291, https://doi.org/10.1002/qj.2819.

Gao, F., X. Y. Zhang, N. A. Jacobs, X.-Y. Huang, X. Zhang, and P. P. Childs, 2012: Estimation of TAMDAR observational error and assimilation experiments. *Wea. Forecasting*, **27**, 856–877, https://doi.org/10.1175/WAF-D-11-00120.1.

Harris, B. A., and G. Kelly, 2001: A satellite radiance-bias correction scheme for data assimilation. *Quart. J. Roy. Meteor. Soc.*, **127**, 1453–1468, https://doi.org/10.1002/qj.49712757418.

Huang, X.-Y., and Coauthors, 2009: Four-dimensional variational data assimilation for WRF: Formulation and preliminary results. *Mon. Wea. Rev.*, **137**, 299–314, https://doi.org/10.1175/2008MWR2577.1.

Isaksen, L., D. Vasiljevic, D. P. Dee, and S. Healy, 2012: Bias correction of aircraft data implemented in November 2011.

ECMWF Newsletter, No. 131, ECMWF, Reading, United Kingdom, 6–7.

Jacobs, N. A., D. J. Mulally, and A. K. Anderson, 2014: Correction of flux valve–based heading for improvement of aircraft wind observations. *J. Atmos. Oceanic Technol.*, **31**, 1733–1747, https://doi.org/10.1175/JTECH-D-13-00175.1.

Le Dimet, F.-X., and O. Talagrand, 1986: Variational algorithms for analysis and assimilation of meteorological observations: Theoretical aspects. *Tellus*, **38A**, 97–110, https://doi.org/10.1111/j.1600-0870.1986.tb00459.x.

Lin, Y., and K. E. Mitchell, 2005: The NCEP stage II/IV hourly precipitation analyses: Development and applications. *19th Conf. on Hydrology*, San Diego, CA, Amer. Meteor. Soc., 1.2, https://ams.confex.com/ams/Annual2005/techprogram/paper_83847.htm.

Moninger, W. R., S. G. Benjamin, B. D. Jamison, T. W. Schlatter, T. L. Smith, and E. J. Szoke, 2010: Evaluation of regional aircraft observations using TAMDAR. *Wea. Forecasting*, **25**, 627–645, https://doi.org/10.1175/2009WAF2222321.1.

Parrish, D. F., and J. C. Derber, 1992: The National Meteorological Center's Spectral Statistical Interpolation analysis system. *Mon. Wea. Rev.*, **120**, 1747–1763, https://doi.org/10.1175/1520-0493(1992)120<1747:TNMCSS>2.0.CO;2.

Petersen, R. A., 2016: On the impact and benefits of AMDAR observations in operational forecasting—Part I: A review of the impact of automated aircraft wind and temperature reports. *Bull. Amer. Meteor. Soc.*, **97**, 585–602, https://doi.org/10.1175/BAMS-D-14-00055.1.

Roberts, N. M., and H. W. Lean, 2008: Scale-selective verification of rainfall accumulations from high-resolution forecasts of convective events. *Mon. Wea. Rev.*, **136**, 78–97, https://doi.org/10.1175/2007MWR2123.1.

Sato, Y., 2007: Introduction of variational bias correction technique into the JMA global data assimilation system. CAS/JSC WGNE Research Activities in Atmospheric and Oceanic Modeling, Rep. 37, 2 pp., http://bluebook.meteoinfo.ru/uploads/2007/individual-articles/01_Sato_Yoshiaki_sato-varbc.pdf.

Schaefer, J. T., 1990: The critical success index as an indicator of warning skill. *Wea. Forecasting*, **5**, 570–575, https://doi.org/10.1175/1520-0434(1990)005<0570:TCSIAA>2.0.CO;2.

Schwartz, C. S., Z. Liu, and X.-Y. Huang, 2015: Sensitivity of limited-area hybrid variational-ensemble analyses and forecasts to ensemble perturbation resolution. *Mon. Wea. Rev.*, **143**, 3454–3477, https://doi.org/10.1175/MWR-D-14-00259.1.

Shao, H., and Coauthors, 2016: Bridging research to operations transitions: Status and plans of community GSI. *Bull. Amer. Meteor. Soc.*, **97**, 1427–1440, https://doi.org/10.1175/BAMS-D-13-00245.1.

Skamarock, W. C., and Coauthors, 2008: A description of the Advanced Research WRF version 3. NCAR Tech. Note NCAR/TN-475+STR, 113 pp., https://doi.org/10.5065/D68S4MVH.

Wang, X., D. M. Barker, C. Snyder, and T. M. Hamill, 2008a: A hybrid ETKF–3DVAR data assimilation scheme for the WRF Model. Part I: Observing system simulation experiment. *Mon. Wea. Rev.*, **136**, 5116–5131, https://doi.org/10.1175/2008MWR2444.1.

——, ——, ——, and ——, 2008b: A hybrid ETKF–3DVAR data assimilation scheme for the WRF Model. Part II: Real observation experiments. *Mon. Wea. Rev.*, **136**, 5132–5147, https://doi.org/10.1175/2008MWR2445.1.

World Meteorological Organization 2003: Aircraft Meteorological Data Relay (AMDAR) reference manual. WMO Rep. 958, WMO, 84 pp., https://www.wmo.int/pages/prog/www/GOS/ABO/AMDAR/publications/AMDAR_Reference_Manual_2003.pdf.

Zhu, Y., J. C. Derber, A. Collard, D. P. Dee, R. Treadon, G. Gayno, and J. A. Jung, 2014: Enhanced radiance bias correction in the National Centers for Environmental Prediction's Gridpoint Statistical Interpolation data assimilation system. *Quart. J. Roy. Meteor. Soc.*, **140**, 1479–1492, https://doi.org/10.1002/qj.2233.

——, ——, R. J. Purser, B. A. Ballish, and J. Whiting, 2015: Variational correction of aircraft temperature bias in the NCEP's GSI analysis system. *Mon. Wea. Rev.*, **143**, 3774–3803, https://doi.org/10.1175/MWR-D-14-00235.1.