Satellite Bias Correction in the Regional Model ALADIN/CZ: Comparison of Different VarBC Approaches

PATRIK BENÁČEK

Numerical Weather Prediction Department, Czech Hydrometeorological Institute, and Atmospheric Physics Department, Charles University, Prague, Czech Republic

MÁTÉ MILE

Unit of Methodology Development, Hungarian Meteorological Service, Budapest, Hungary

(Manuscript received 12 October 2018, in final form 30 May 2019)

ABSTRACT

The bias correction of satellite radiances is an essential component of data assimilation system in numerical weather prediction (NWP). The variational bias correction (VarBC) scheme is widely used by global NWP centers, but there are still open questions regarding its use in limited-area models (LAMs). We present a study of key VarBC aspects in the limited-area 3D-Var system using the state-of-the-art NWP system ALADIN. Two basic VarBC applications are tested, specifically adopting bias coefficients from the global model ARPEGE and cycling bias coefficients independently in the LAM ALADIN (VarBC-LAM). The latter application is studied using daily update of bias coefficients with regards to static and dynamic settings of the VarBC stiffness. Extensive testing shows that the VarBC-LAM methods outperform the use of global coefficients from ARPEGE providing the better quality of the model first guess (3-h forecast), in the assimilation cycle with the largest normalized impact of 2%–3% for temperature and wind components in the midtroposphere. Compared to the global coefficients, there was little forecast impact between 24 and 48 h from using the VarBC-LAM coefficients. The various VarBC-LAM methods were comparable, but the CAM method may be most useful when an unexpected bias shows up.

1. Introduction

The ambition of data assimilation (DA) is to determine the most accurate and balanced initial conditions for numerical weather prediction (NWP) making optimal use of all available information and their uncertainty. Many factors make this initial value problem difficult to perform and the design of DA postulates standard assumptions in order to attain theoretical and practical solutions (see Bouttier and Courtier 1999). Usefulness of these solutions is reduced if the validity of these assumptions does not hold. One of the key demands is to ensure bias free DA. The output of the DA is an accurate analysis most often obtained by a variational DA method but viable alternatives also exists [e.g., the ensemble Kalman filter (EnKF) (Evensen 1994) with its variations]. For practical reasons, this study deals with the three-dimensional variational (3D-Var) DA method that is most often employed in the limited-area model

(LAM) systems. It minimizes so-called variational cost function:

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

where vector $\mathbf{x}_b$ is the background, $\mathbf{y}$ is the observation vector, and $\mathbf{x}$ is the state vector; $\mathbf{B}$ and $\mathbf{R}$ are the background and the observation error covariance matrices, respectively; and $h(\mathbf{x})$ is the forward operator providing an interpolation of model fields to the observation locations and simulation of the measured variables from the background parameters (e.g., a radiative transfer model). To perform unbiased estimation of $\mathbf{x}$, we assume that both $\mathbf{y}$ and $\mathbf{x}_b$ have only Gaussian, zero-mean errors that are mutually uncorrelated. If the model background or observations contain systematic errors (biases), the available data are not used optimally in the assimilation scheme and the analysis is biased (Dee 2005). Therefore, a bias correction is essential in order to get a positive impact in NWP from observations.

Corresponding author: Patrik Benáček, patrik.benacek@chmi.cz

The bias detection itself needs accurate and unbiased spatially and temporally collocated reference. The NWP centers use as reference the model forecast or an analysis as best estimate of the atmosphere to diagnose the bias in the observations. Because this reference is not completely unbiased, the bias thus diagnosed contains a mixture of observation and NWP model errors that cannot be distinguished (Dee 2005). Therefore, good knowledge of the source and nature of the biases is essential for its subsequent removal.

The background biases come mostly from inaccuracies of the NWP model like parameterization, spatial and temporal discretization, the imperfect use of boundary conditions, and unrepresented physical and dynamical processes (Lahoz et al. 2010; Torn and Davis 2012; Romine et al. 2013). The observation biases involve instrument error (e.g., poor calibration), approximations in forward operator and data processing such as the radiative transfer model or cloud detection (Auligné et al. 2007). This creates a complex bias depending on, for example, satellite-scan angle, air mass, and surface characteristics (Dee 2005; Auligné 2007).

Although the radiance bias correction is widely used in global DA systems, it has not been fully established in the regional DA because of difficulties in estimating the biases in limited-area domain. The bias correction quality is conditioned by a robust observation sample obtained under various meteorological conditions and scan-angle positions which is not fully assured within the limited domain. Especially polar-orbiting satellites provide a nonuniform observation sample on which general trends in the radiance bias cannot be completely captured. On the other hand, many operational centers were able to tackle these difficulties (e.g., by sharing VarBC with global system) and showed that satellite radiances are valuable components of their DA (e.g., Gustafsson et al. 2018).

In principle, there are two types of bias correction schemes used for satellite radiance observations in the (3D/4D) variational DA scheme: static and variational. Both schemes are represented by linear prediction model involving predictors derived from the model state and the observation geometry. The predictors are weighted by regression coefficients referred to as bias coefficients. While the static approach uses fixed bias coefficients in all consecutive analyses so that it is not capable to follow bias changes (Eyre 1992; Harris and Kelly 2001), the Variational Bias Correction (VarBC) scheme updates bias coefficients jointly with the model state within the variational assimilation system (Dee 2005). Since bias coefficients are adaptively updated during each minimization step, the VarBC system is capable to gain control over bias shifts consistently with all available information. The adaptive bias correction augmenting bias coefficients to the state vector can also be found within the EnKF method. There, the bias coefficients are analyzed jointly with the state vector using an ensemble of bias coefficients (Fertig et al. 2009).

In this vein, the choice of the variational bias correction seems obvious in both the LAM and global systems. In many NWP centers (e.g., Météo-France, Japan Meteorological Agency, Canadian Meteorological Centre), bias coefficients are adopted to LAM VarBC from their host global models assuming that the flow-dependent and limb-bias corrections are likely better characterized in the global system (Guidard et al. 2011). Kazumori (2014) suggested that bias coefficients from a global model can be used provided that (i) the same set of VarBC predictors and the same radiative transfer model are utilized; and (ii) differences between radiance biases detected in global and regional models are not considerably large. However, biases between global and regional systems can disagree due to a different level of the models' complexity (e.g., differences in model domain, physics or bigger resolution gap, vertical representation). In this context, Randriamampianina (2005) has shown for AMSU-A data that computing bias coefficients independently in the LAM system might lead to more stable scores in analysis and daily forecast compared to adopting the global bias coefficients.

Because the VarBC scheme was designed for global models, its application in a LAM system requires further attention. For example, Lindskog et al. (2012) suggested a revision of the LAM VarBC by rejecting collinear predictors and predictors peaking over the top of limited-area model. Moreover, they suggested doubling the adaptivity rate of bias coefficients leading to a proper spinup time of approximately a 1-month period (Campins et al. 2017). Randriamampianina et al. (2011) showed for polar-orbiting satellites, that the bias is hardly reduced in time due to the nonuniform data sample gathering different atmospheric conditions at different analysis times. To avoid this influence, they suggested a separate daily update of bias coefficients at each analysis time. Moreover, they removed the small and odd paths of satellites inside the LAM domain at particular analysis times because they are not representative of the entire model domain.

Lin et al. (2017) showed the relevance of maximizing the data coverage, the reduction of data latency, and the quality of bias correction in Rapid Refresh (RAP) hourly updating prediction system. Therefore, there is a need to include as much observations as possible holding the uncorrelated observation error structures

in recent DA systems. Cameron and Bell (2016) suggested an adaptivity approach in which the VarBC stiffness depends on a bias halving time parameter and the expected number of observations in the analysis cycle. Although this approach was originally designed for global models, the concept can be drafted mainly as a possible solution for solving lack of the uniform satellite dataset across analysis times.

Although, there are several configurations how to set up the VarBC scheme in LAMs, there is still no comprehensive study assessing their performance. In principle, two basic initializations of the VarBC are examined in a 3D-Var mesoscale NWP model system: adopting global bias coefficients and cycling of bias coefficients independently in the LAM system. The latter method is examined using the daily VarBC-cycling strategy, focusing on its adaptivity setting and the observation sample across analysis times. In this study, different experiments were conducted to answer the following questions: (i) How to initialize the bias coefficients in LAM to achieve the best forecast performance? (ii) What is the appropriate adaptivity setting in LAM to guarantee the best performance of the VarBC scheme? (iii) Can the VarBC stiffness be controlled with respect to the expected number of observations at each assimilation time? Can small and odd satellite paths be assimilated for better forecast performance? These hypotheses are studied in the regional model ALADIN (Termonia et al. 2018, Aire Limitée Adaptation dynamique Développement InterNational), which is operationally used at Czech Hydrometeorological Institute (Bučánek et al. 2015).

The structure of this paper is as follows: in section 2, we introduce the VarBC scheme. In section 3, attributes of the VarBC settings will be discussed with special emphasis on its application in the limited-area 3D-Var system. In section 4, we describe different configurations of the VarBC scheme examined in the ALADIN model. In section 5, we evaluate the initialization of bias coefficients, the forecast impact study, and the response of the VarBC configurations to an artificial bias. Finally, section 6 and section 7 provide a summary of the presented work and draw corresponding conclusions.

2. The VarBC implementation in limited-area models

a. Introduction of the method

Many observations include biases that should be removed before or during the assimilation procedure. Provided that observation errors follow Gaussian distribution, the observation bias $\mathbf{b}_o$ can be determined as follows:

$$\mathbf{b}_o = \langle \mathbf{y} - h(\mathbf{x}_t) \rangle, \qquad (2)$$

where $\langle \cdot \rangle$ represents a spatial/temporal average, $\mathbf{y}$ observation, and $h(\mathbf{x}_t)$-forward operator representing the true atmospheric state $\mathbf{x}_t$ at the observation space. Because the true state is unknown, it is often replaced by the model background $\mathbf{x}_b$ providing its best estimate but not completely unbiased (Dee 2005). Therefore, a nonzero bias determined by Eq. (2) contains a mixture of the observation, the forward operator, and the NWP model biases. To remove the bias, it is often used a parametric form of a bias prediction model $\hat{f}(\mathbf{x}_b, \boldsymbol{\beta})$ such that

$$\mathbf{b}_o = \langle \mathbf{y} - h(\mathbf{x}_b) - \hat{f}(\mathbf{x}_b, \boldsymbol{\beta}) \rangle = 0.$$

The bias is simulated as a linear combination of predictors $p_k(\mathbf{x}_t)$, $k = 1, \ldots, N_p$ depending on for example, the satellite scan position, model airmass thicknesses, or surface characteristics:

$$\hat{f}(\mathbf{x}_b, \boldsymbol{\beta}) = \beta_0 + \sum_{k=1}^{N_p} \beta_k p_k(\mathbf{x}_b), \qquad (3)$$

where β_0 represents a constant bias offset, and bias coefficients β_k, $k = 1, \ldots, N_p$ determine weight of the predictors within the bias prediction. The set of predictors is flexible for individual satellites, instruments, and channels (called satellite groups). Supposing N_g is the dimension of the satellite groups where each group has N_p predictors, then the total number of coefficients is $N_g \times N_p$. Although, the bias prediction model does not separate the origins of the bias, the set of predictors must be chosen very carefully to focus the bias correction scheme on removing the observation bias rather than the NWP model bias (Auligné et al. 2007). Moreover, too many predictors can provide unbiased analysis; however, there is also a risk of overfitting the bias prediction scheme. A selection of predictors has been studied for the static bias correction scheme by Eyre (1992); Harris and Kelly (2001), and more recently for the VarBC scheme by Auligné et al. (2007).

In the following text, we shall work in the context of variational DA where the bias prediction model is included in the cost function in Eq. (1). Thus in the VarBC scheme both the model state $\mathbf{x}$ and bias coefficients $\boldsymbol{\beta}$ are included in the augmented control vector. Provided the independence between the model state error and the bias coefficients error, the cost function becomes:

$$J(\mathbf{x},\boldsymbol{\beta}) = (\mathbf{x}_b - \mathbf{x})^{\mathrm{T}}\mathbf{B}^{-1}(\mathbf{x}_b - \mathbf{x})$$
$$+ (\boldsymbol{\beta}_b - \boldsymbol{\beta})^{\mathrm{T}}\mathbf{B}_{\boldsymbol{\beta}}^{-1}(\boldsymbol{\beta}_b - \boldsymbol{\beta})$$
$$+ [\mathbf{y} - h(\mathbf{x}) - \hat{f}(\mathbf{x},\boldsymbol{\beta})]^{\mathrm{T}}\mathbf{R}^{-1}[\mathbf{y} - h(\mathbf{x}) - \hat{f}(\mathbf{x},\boldsymbol{\beta})],$$
$$(4)$$

where $\boldsymbol{\beta}_b$ is the vector of background bias coefficients issued from the former analysis cycle and $\mathbf{B}_{\boldsymbol{\beta}}$ is the background bias coefficient error covariance matrix. In practice, $\mathbf{B}_{\boldsymbol{\beta}}$ is diagonal with the error variance of background bias coefficients $\sigma^2_{\beta_{b_j}}$, $j = 1, \ldots, N_g \times N_p$ (Dee 2004):

$$\sigma^2_{\beta_{b_j}} = \frac{\sigma^2_{o_j}}{N_{\mathrm{bg}_j}}, \quad (5)$$

where σ_{o_j} is the error standard deviation of the observations associated with β_j bias coefficient, and N_{bg_j} is a positive integer set by the user which determines the adaptivity rate of β_j in analysis cycle. For practical application, the same set of N_{bg_j} is used for all predictors and instruments channels (i.e., the parameter is set individually only for each satellite instrument).

b. The adaptivity of VarBC

The VarBC adaptivity represents a response of bias coefficients to sudden bias changes in the assimilation cycle. For simplicity, we use the cost function in Eq. (4) with only bias coefficients $\boldsymbol{\beta}$ in the control vector. Then the training of the bias correction is based on finding the bias coefficients minimizing the following cost function:

$$J(\boldsymbol{\beta}) = (\boldsymbol{\beta}_b - \boldsymbol{\beta})^{\mathrm{T}}\mathbf{B}_{\boldsymbol{\beta}}^{-1}(\boldsymbol{\beta}_b - \boldsymbol{\beta})$$
$$+ [\mathbf{y} - h(\mathbf{x}_b) - \hat{f}(\mathbf{x}_b,\boldsymbol{\beta})]^{\mathrm{T}}\mathbf{R}^{-1}[\mathbf{y} - h(\mathbf{x}_b) - \hat{f}(\mathbf{x}_b,\boldsymbol{\beta})].$$
$$(6)$$

Supposing that we have N measurements from one particular channel, and the bias prediction model includes only a constant bias correction [i.e., $\hat{f}(\mathbf{x}_b, \boldsymbol{\beta}) = \beta_0$]. Then the cost function in Eq. (6) is given by

$$J(\beta_0) = \frac{(\beta_0^b - \beta_0)^2}{\sigma^2_{\beta_b}} + \sum_{i=1}^{N}\frac{[y_i - h(x_b)_i - \beta_0]^2}{\sigma^2_o}, \quad (7)$$

where the error of background bias coefficients is given by Eq. (5). The best estimate of β_0 is found by minimizing its distance from the prior information β_0^b and best fits the current observations and the model background (the second term). Calculating the derivative of Eq. (7) with respect to β_0 and setting it equal to zero the equation becomes:

$$N_{\mathrm{bg}}(\beta_0 - \beta_0^b) = N\left\{\beta_0 - \frac{1}{N}\sum_{i=1}^{N}[y_i - h(x_b)_i]\right\}.$$

Note that background bias coefficient has the same weight as N_{bg} observations. In limiting case when $N_{\mathrm{bg}} = 0$ (or $N = 0$) the estimate of β_0 corresponds to the current bias estimate (or the background bias coefficient β_b). Thus if $N_{\mathrm{bg}} < N$, the VarBC scheme will adapt quickly from β_0^b toward the current bias. If $N_{\mathrm{bg}} > N$, the VarBC scheme gives the larger weight to β_0^b leading to the less adaptive bias correction scheme. Therefore, the VarBC adaptivity depends on setting of the N_{bg} parameter which is called the stiffness parameter of the VarBC scheme.

3. Application of VarBC in LAMs

a. Initialization of the VarBC

The VarBC bias coefficients can be either adopted from a global model at each analysis time (VarBC-global) or cycled independently in a regional model (VarBC-LAM). For the VarBC-global method, no initialization of bias coefficients is needed since up-to-date global bias coefficients are directly exploited to LAMs at each analysis time. The VarBC-LAM requires to be warmed up through a process called initialization. The initialization can be done in a passive mode (Liu et al. 2012; Auligné et al. 2007), that is, without impact of satellite measurements on the analysis, or in a full VarBC mode (Schwartz et al. 2012) [i.e., using Eq. (4)].

b. The availability of data in limited area

The satellite orbit determines the strategy of observing Earth from space as well as the observation availability over a limited area at a particular time. While the geostationary orbit satellite is always positioned above the same location of Earth providing a stable dataset in LAMs, the polar-orbiting satellite orbits around the Earth in sun-synchronous orbits providing a limited dataset in space and time. Supposing that only observations inside the model domain and closest to the analysis time are considered in the 3D-Var DA system, the polar satellites provide nonuniform observation sample within the analysis cycle. Here, we show an example of the *MetOp-B* (Meteorological Operational satellite; EUMETSAT polar system) data coverage within the 3-h analysis cycle in the ALADIN/CHMI (Fig. 1). Note that the satellite scans the whole domain at 0900 UTC but measurements hardly cover the full area at 1200, 1800, and 2100 UTC. The number of measurements can be even more reduced for particular channels by the observation screening
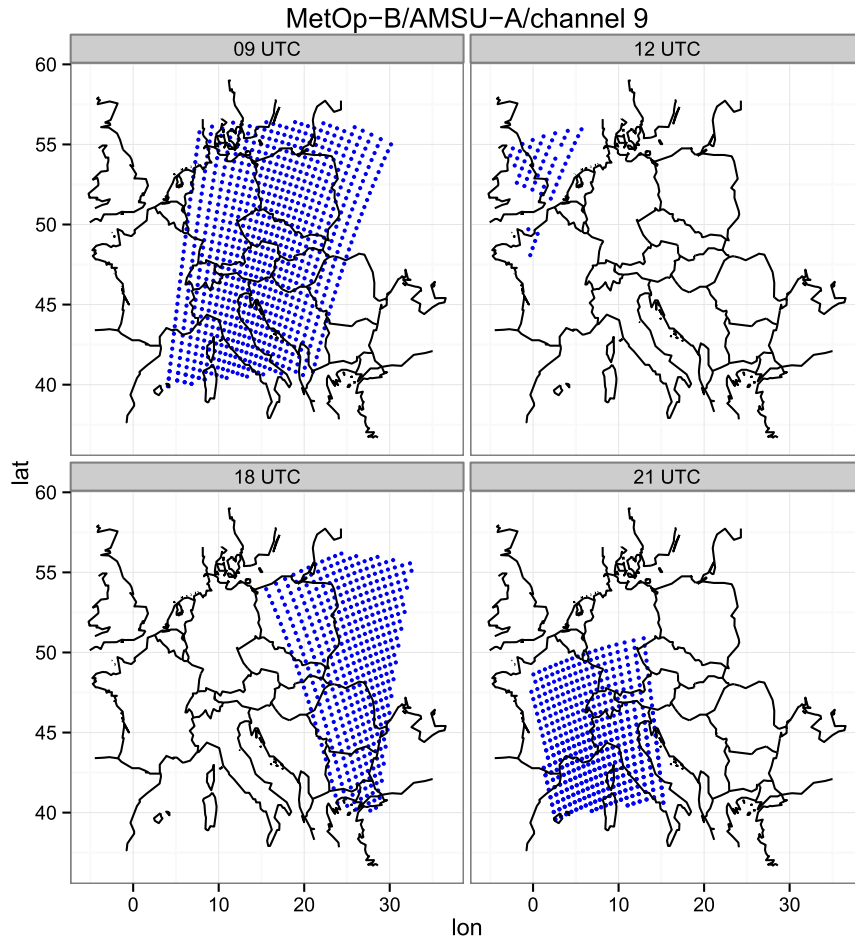
FIG. 1. The example of the *MetOp-B* data coverage in the ALADIN domain within 3-h assimilation cycle, taken for 1 Sep 2015.

involving, for example, quality checks, data thinning or permanently incorrect data blacklisting (Järvinen and Undén 1997).

c. Data sample issues in LAMs

To predict bias with all its dependencies, it is essential to have a miscellaneous data sample obtained under different meteorological conditions and/or from different satellite scan positions. Even if such data sample is ensured by global models, polar-orbiting satellite coverage is nonuniform and highly variable in LAMs (see paragraph above). Assuming that the expected number of observations at each analysis time is N_{avg}, and the sample of observation-minus-background (OMB) residuals is independent with the standard deviation σ_o. Then the sample variance of the bias is given by (Wilks 2011, p. 122):

$$\sigma_{b_o}^2 = \frac{\sigma_o^2}{N_{avg}} k, \tag{8}$$

where k is a variance inflation factor to account for autocorrelation or other effects that might affect the results (Geer 2016).

The sample variance of bias is compared between the regional and global model systems gathering dataset for different scales of computing area. Figure 2 shows the time evolution of the bias (the mean of OMB residuals) calculated for the particular MHS channel 5 during the 1-month analysis cycle. At each analysis time, the bias is calculated for the global ARPEGE system, with the full sample of 8000 measurements (ARP8000) and random sample of 300 measurements (ARP300), and for the regional ALADIN system with the full sample of 300 measurements (ALD300). Note that ARP8000 provides a stable bias in time with its sample variance $\sigma_{b_o}^2$ of approximately 0.002 K. The sample of 300 measurements increases $\sigma_{b_o}^2$ to approximately 0.01 K in ARP300 while the variance increases significantly up to 0.18 K in ALD300. This effect is most likely due to spatially and temporally correlated sample of OMB residuals in the
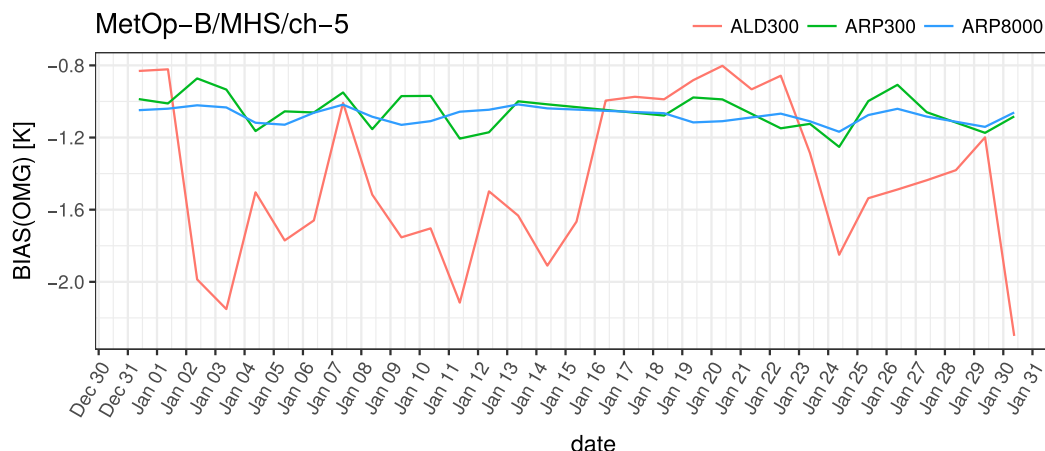
Fig. 2. The bias evolution in global ARPEGE (1200 UTC) and regional ALADIN (0900 UTC) models for particular MHS channel 5 on *MetOp-B* during January 2016. At each analysis time, the bias is calculated for ARPEGE with the full sample of 8000 measurements (ARP8000) and random sample of 300 measurements (ARP300), and for ALADIN with the full sample of 300 measurements (ALD300).

LAM leading to higher variance inflation factor k of approximately 18 for this particular channel. As a result, two things should be considered when estimating the VarBC stiffness in LAMs, namely the expected number of observations (N_{avg}) and the variance inflation factor (k).

d. Revision of the adaptivity of the VarBC

Regarding data sample issues, the stiffness parameter should be revised to constraint appropriately bias coefficient estimates in LAMs. In the VarBC scheme, the counterpart to Eq. (8) is using the stiffness parameter variance inflation factor $N_{bg} = N_{avg} \times k$, which controls the adaptivity of bias coefficients in the LAM analysis cycle. Wilks (2011, p. 127) characterized the variance inflation factor by the lag-1 autocorrelation coefficient; however, this approach is impractical for the VarBC scheme application (e.g., when a new satellite is sent into orbit). Cameron and Bell (2016) proposed the stiffness parameter $N_{bg_j}, j = 1, \ldots, N_g \times N_p$ such that the variance inflation factor is characterized by a bias halving time parameter $\mathbf{n}_h$:

$$\mathbf{N}_{bg} = \max(\mathbf{N}_{avg}, \mathbf{N}_{min}) \underbrace{\left(\frac{1}{2^{1/\mathbf{n}_h} - 1}\right)}_{\mathbf{k}}, \qquad (9)$$

where $\mathbf{N}_{min}$ is a minimum number of observations used as a safety mechanism to constraint $\mathbf{N}_{avg}$. The bias halving time parameter $\mathbf{n}_h$ represents the number of steps in the assimilation cycle needed to halve the bias assuming the cost function [Eq. (6)] and the independent data sample. Although these assumptions are not fully met by the DA system, this approach gives more intuitive instructions as to how to control the variance of bias coefficients in the LAM analysis cycle.

In the LAM system, the bias estimate is highly variable in time. It comprises small-scale day-to-day features varying in daily and weekly cycle (see Fig. 2), as well as, it can comprise large-scale biases, varying in time scale longer than 10 days (not shown). The large-scale stationary biases are included in the observation, the NWP model, and the radiative transfer model. The day-to-day features result from local errors in the NWP model fields depending on season, weather regime, and time of the day. The aim of the VarBC scheme is to remove the slowly varying biases prior to assimilation and maintain the small-scale bias features. The bias halving time parameter allows one to set the bias correction scheme with respect to the time scale. However, finding appropriate bias halving time parameter is a difficult task. For a too long $\mathbf{n}_h$, the bias correction is impractical when it takes a long time to correct instrument bias changes. For a too short $\mathbf{n}_h$ the day-to-day features in the NWP model error can be misinterpreted as the instrument bias.

4. Experimental design

a. The model ALADIN

This study was performed using the regional model ALADIN operated at the Czech Hydrometeorological Institute (CHMI). The model domain covers mostly central Europe (Fig. 1), with horizontal grid size of 4.7 km and 87 vertical levels from surface up to 0.1 hPa. Vertical levels are unevenly distributed, having the most dense and sparse coverage near the surface and above 50 hPa, respectively. The 3D-Var DA system (Fischer et al. 2005) is applied with a 3-h forward intermittent cycle used for this study. The forward operator includes the radiative transfer model RTTOV-10

(Saunders et al. 1999) to simulate radiances from the model variables. Surface analysis is formed by an optimal interpolation method using 2-m measurements from land surface stations. The assimilation (production) cycle uses ±1.5-h assimilation window. The ALADIN system is coupled in space and time to the global system Action de Recherche Petite Echelle Grande Echelle (ARPEGE) providing lateral-boundary conditions.

Conventional observations used in this study are radiosondes (TEMP), aircraft reports (AMDAR, MODE-S/MRAR), land surface stations (SYNOP), and winds from atmospheric motion vectors (AMVs). In terms of satellite observations, we use radiances from Advanced Tiros Operational Vertical Sounder (ATOVS) on polar satellites *NOAA-18*, *NOAA-19*, and *MetOp-B*, Infrared Atmospheric Sounder Interferometer (IASI) on *MetOp-B*, and Spinning Enhanced Visible and Infrared Imager (SEVIRI) on board the geostationary satellite *Meteosat-11*. The ALADIN model shares the same radiance processing for ATOVS and IASI as ARPEGE. In the current configuration, radiance observations from ATOVS (Gérard et al. 2003), IASI (Guidard et al. 2011), and SEVIRI (Montmerle et al. 2007) are employed with the same channel selection and data preprocessing (e.g., quality control, cloud detection, observation error variances) as described in the references with the following exceptions: For ATOVS, the land surface emissivity dynamical approach is activated (Karbou et al. 2006). For AMSU-A, the channels 11–13 peaking above 10 hPa are removed due to the sparse model resolution in this region (Lindskog et al. 2012; Lin et al. 2017). For AMSU-B/MHS, measurements are used in a higher resolution in ALADIN using maximum 90 field-of-views (FOVs) instead of 30 FOVs in ARPEGE. For IASI, channels 323–1271 assimilated over sea are blacklisted in ALADIN due to a small portion of seas in the model domain. The spatial data thinning length in ALADIN is 80 km for ATOVS, 70 km for SEVIRI (Montmerle et al. 2007), and 80 km for IASI (Guidard et al. 2011). In ARPEGE, the assimilation density is 139 km for ATOVS, 100 km for IASI, and 125 km for SEVIRI.

b. The VarBC configurations in ALADIN

We evaluate the performance of the VarBC-LAM and VarBC-global configurations in the framework of the limited-area model ALADIN.

In VarBC-global, bias coefficients are adopted from the global model ARPEGE at each analysis time without any initialization period. To assure a consistent set of global bias coefficients with the LAM system, the same set of predictors is used in the ARPEGE and ALADIN systems (see Table 1). Following Kazumori (2014), we estimate

TABLE 1. List of selected predictors in the ARPEGE/ALADIN models.

Predictor No.	Predictor
0	Constant
1	1000–300-hPa thickness
2	200–50-hPa thickness
3	Skin temperature
4	Total column water
5	10–1-hPa thickness
6	50–5-hPa thickness
8	Nadir-viewing angle
9	Nadir-viewing angle **2
10	Nadir-viewing angle **3
11	Nadir-viewing angle **4
15	Land or sea ice mask
16	View angle (land)
17	View angle **2 (land)
18	View angle **3 (land)

bias differences between the ARPEGE and ALADIN systems. Figure 3 shows that biases are similar for the tropospheric-peaking AMSU-A channels 5–10, while larger differences are detected for the AMSU-A channels 11–13 and IASI channels 16–109 (not shown) peaking above 10 hPa. These high-peaking channels are removed from the ALADIN DA system as described above.

In VarBC-LAM, daily update of bias coefficients is used (Randriamampianina et al. 2011) so that there are 8 sets of bias coefficients, one set for every 3 h, which evolve independently in the ALADIN. The bias coefficients are initialized from 15 October to 30 November 2015 using the passive VarBC mode. We examine the performance of different VarBC stiffness parameter N_{bg} settings currently used in LAM systems: 5000 (NBG5000) as in ARPEGE (V. Guidard 11 January 2018, personal communication,) and 2000 (NBG2000) suggested by Lindskog et al. (2012). In addition, we employ the Cameron and Bell (2016) approach (CAM) described in Eq. (9). The VarBC-global method is referred to as ARP in our experiments.

1) CHANNEL BLACKLISTING

According to Randriamampianina et al. (2011), the instrument channels providing $N_{avg} \leq N_{min}$ are rejected from the active DA system at given analysis times. We study N_{min} based on a response of different VarBC-LAM configurations to an artificial bias in the passive VarBC mode. The bias is set to the observation standard deviation σ_o. Figure 4 shows the percentage bias correction after 15 days period with respect to N_{avg}. The N_{min} is found as the minimum value of N_{avg} that provides at least 75% bias correction. Note that the minimum number of observation is approximately 300 ($N_{min} = 300$) for NBG5000 and 100 ($N_{min} = 100$) for NBG2000, while the CAM method allows assimilating
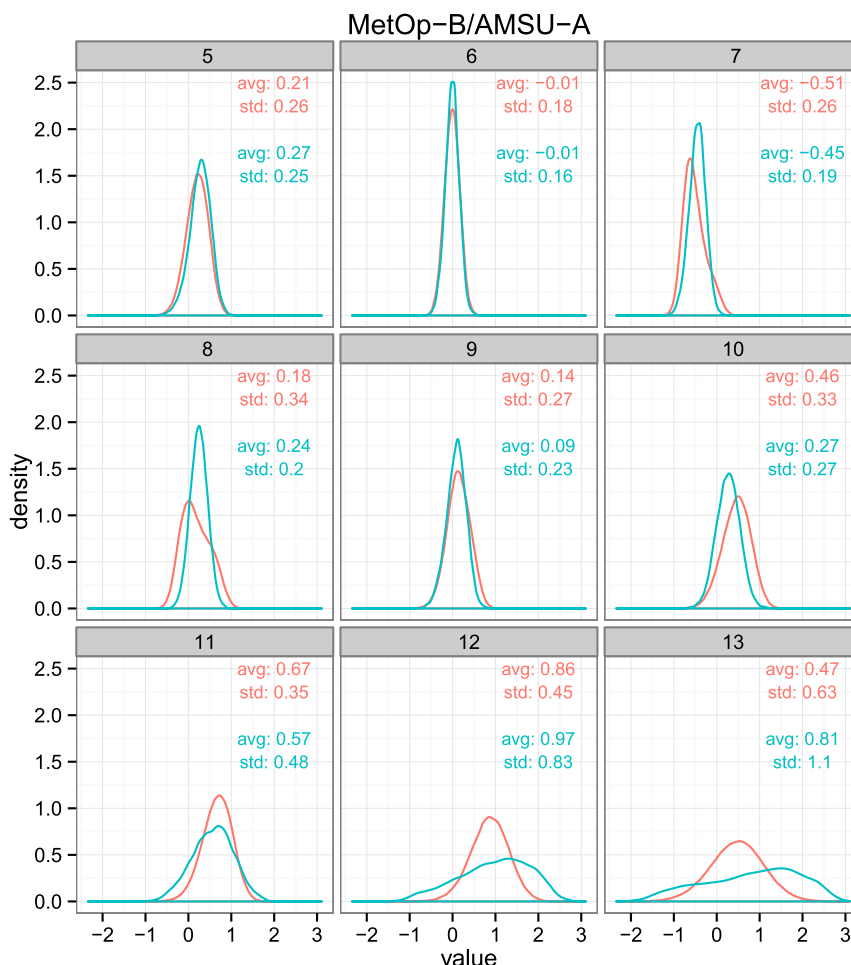
FIG. 3. Density function of the observation-minus-background differences for the AMSU-A channels on *MetOp-B*. The differences are gathered separately for global ARPEGE (red) and regional ALADIN (blue) models during January 2016.

even 50 observations ($N_{min} = 50$) with the same bias correction performance. Also note that in case of $N_{avg} \geq 250$, the performance of CAM and NBG2000 is comparable, while NBG5000 still underfits the bias correction and appears to be the least effective.

In this study, the VarBC configurations ARP, NBG5000, and NBG2000 are evaluated with N_{min} set to 100 observations. The CAM method is evaluated with two different N_{min} settings 100 (CAM100) and 50 (CAM50). The CAM50 method allows to increase the amount of satellite observations used in the DA system as well as to evaluate the forecast impact of the small paths of satellite in the model domain. The channel selection for different VarBC methods and analysis times is described in Table 2.

2) CAM INFLATION FACTOR

In the CAM method, the same bias halving time parameter n_h is set for all satellite groups used in the

DA system. The parameter n_h is set to 15 ($n_h = 15$) for the passive VarBC initialization period, and to 5 ($n_h = 5$) for the active DA period. Supposing the daily VarBC-cycling strategy, we need approximately 15 days to compute half the value of bias coefficients during the VarBC initialization, and approximately 5 days to compute half the value of the bias coefficients in the active DA system. There are two reasons for using 5 days for the active DA: 1) to overcome the influence of the NWP model biases on time scales on the order of 5–10 days (Dee and Todling 2000); 2) supposing that the expected number of ATOVS observations for full domain coverage is 300 per channel, the VarBC stiffness corresponds to NBG2000 (see Fig. 5). This allows a fair comparison of the dynamic and fixed VarBC stiffness settings in this study. For simplicity, the expected number of observations is approximated by the current observation number $N_{avg} \approx N$ in Eq. (9), which is feasible given the quasi-uniform
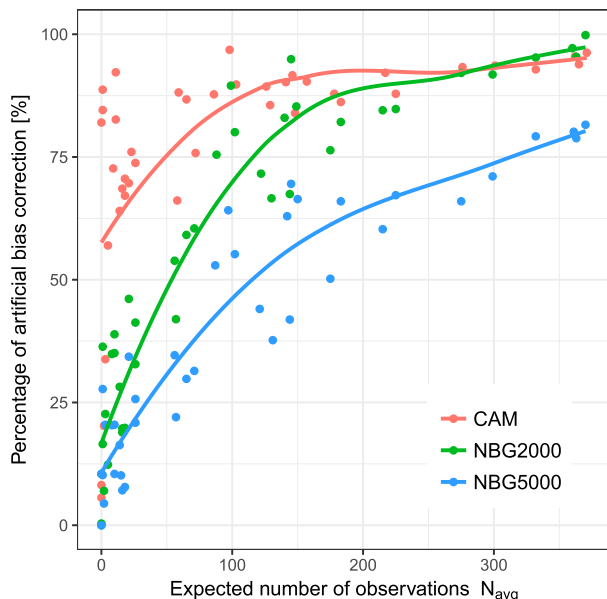
FIG. 4. The percentage correction of the artificial bias depending on the expected number of observations with respect to various VarBC-LAM configurations. Each point represents a percentage bias correction for a particular instrument/channel after 15-day spinup period in the passive VarBC mode. The performance of VarBC-LAM methods is represented by a moving average line.

observation sample at particular analysis times (see Fig. 5a). In case of small paths of the satellite at particular analysis times (Fig. 5b), the VarBC stiffness in CAM is constraint by N_{min} to stabilize the adaptivity.

3) EXPERIMENTAL DESIGN

The design of the experiments is summarized in Table 3. We run 0–48-h forecasts initialized at 0000 and 0600 UTC from assimilation run covering 1 December 2015 to 31 January 2016 (including 124 forecasts). The particular initialization times are selected to assure the uniform amount of radiosonde observations for forecast impact studies. There is a different number of TEMP measurements between 0000 and 1200 UTC (approximately 32 observations) and 0600 and 1800 UTC (appropriately 6 observations) in the ALADIN domain.

5. Results

The evaluation of the VarBC configurations is divided into three parts. First, we compare bias coefficients initialized by the different VarBC configurations during the initialization period from 15 October to 30 November 2015. Supposing that bias predictors are standardized and nondimensional, the relative contribution of each predictor to the bias correction is compared (Auligné et al. 2007). Second, the VarBC configurations are evaluated with respect to SYNOP, TEMP, and AMDAR observations during the validation period from 1 December 2015 to 31 January 2016, via the following diagnostics: (i) check of the analysis and the background with respect to the conventional observations; and (ii) check of the forecast with respect to the conventional observations. Finally, a response of the VarBC-LAM methods on an artificial bias is evaluated in the active DA system during a 1-month period in December 2015.

The diagnostics are measured by root-mean-square (RMS) error and mean (BIAS) of observation-minus-model residuals. According to Geer (2016), the statistical significance is based on the one-sample t test of the mean of paired differences $\overline{d} = 1/n \sum_{i=1}^{n} \text{RMS}_i^{\text{exp}} - \text{RMS}_i^{\text{ref}}$ calculated across n forecasts with 95% confidence interval. No inflation factor is used here assuming negligible time-autocorrelations between the paired differences. Given the very small impact, we calculate a normalized reduction in RMS by $\overline{d}/\overline{\text{RMS}_{\text{ref}}} \times 100$ (%) (Benjamin et al. 2004; Geer 2016) where the VarBC-global (ARP) experiment is used as reference.

Based on Benjamin et al. (2010), we separated the forecast error into three vertical layers: 1000–800 hPa (dominated by boundary layers and surface

TABLE 2. Satellite instruments/channels assimilated in ALADIN with regards to analysis times: active (●), blacklisted (×), active only for CAM50 (○).

Satellite	Instrument	Channel	Analysis time [UTC]							
			0000	0300	0600	0900	1200	1500	1800	2100
NOAA-18	AMSU-A	5–10	×	○	●	×	×	●	○	×
	AMSU-B	3, 4, 5	×	○	●	×	×	●	○	×
NOAA-19	AMSU-A	5–7, 9, 10	○	●	×	×	●	×	×	×
	MHS	4, 5	○	●	×	×	●	×	×	×
MetOp-B	AMSU-A	5–10	×	×	×	●	×	×	●	●
	MHS	3, 4, 5	×	×	×	●	×	×	●	●
	IASI	109–299	×	×	×	●	×	×	●	●
Meteosat-11	SEVIRI	2, 3	●	●	●	●	●	●	●	●
	SEVIRI	4, 6, 7	○	○	○	○	○	○	○	○

Fixed and dynamic VarBC adaptivity setting for MHS/Ch4/MetOp−B
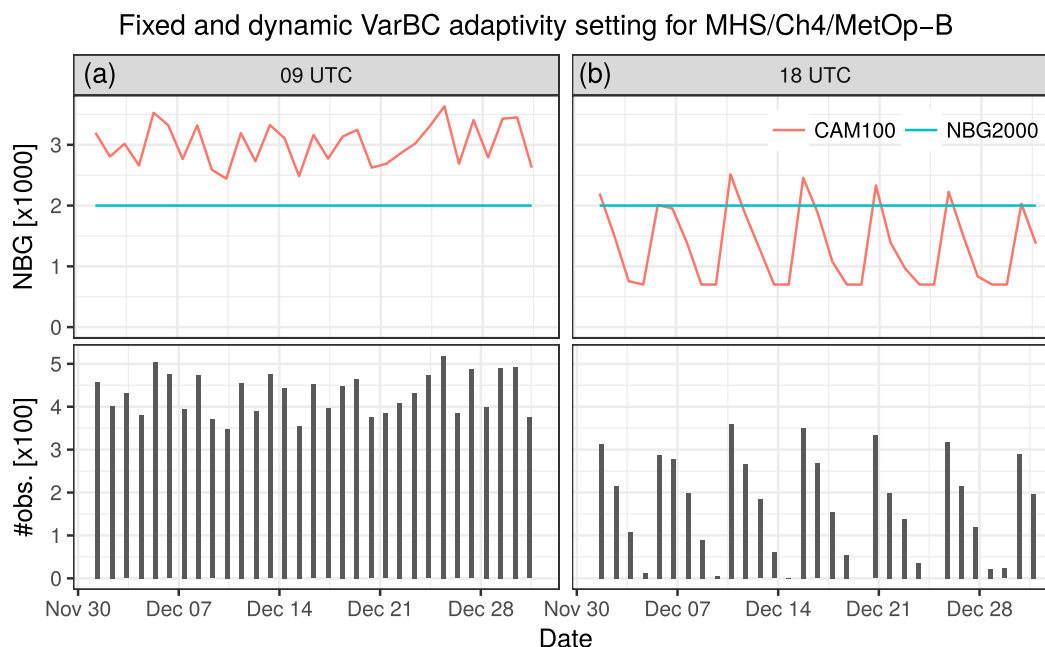

FIG. 5. (top) The stiffness parameter changes with assimilation cycle for CAM and NBG2000 methods with regards to (bottom) the number of observations. The stiffness parameter calculation is shown for the MHS channel 4 on *MetOp-B* at (a) 0900 and (b) 1800 UTC analysis times during December 2015.

effects), 800–400 hPa (middle troposphere), and 400–100 hPa (upper troposphere to lower stratosphere including tropopause and upper-level jet maxima).

a. VarBC initialization

Figure 6 presents a comparison of the mean of bias coefficients calculated after their initialization from 15 October to 30 November 2015. Note that bias coefficients are comparable between all VarBC-LAM methods NBG5000, NBG2000, and CAM50. Compared to ARP, significant differences are detected especially for the global-offset and airmass predictors due to distinct characteristics of the ARPEGE and ALADIN models. For AMSU-A, significant differences are detected for predictors p_2, p_5, and p_6 describing layer thicknesses between 200 and 1 hPa (Fig. 6a). This is most likely due to the sparser vertical resolution in ALADIN at these heights compared to ARPEGE. For MHS, significant differences are detected for the predictor p_1 describing the layer thickness between 1000 and 300 hPa (Fig. 6b). This difference is likely due to an influence of specific weather regimes in the regional model during the winter initialization period. Another possible reason could be a higher resolution of troposphere in the mesoscale ALADIN, and its more accurate description of the model physics compared to ARPEGE. The geometric bias correction, which is represented by the

satellite scan-angle predictors p_8, p_9, and p_{10}, proves to be model independent.

b. Analysis and first-guess control

We measure the BIAS and RMS of observation-minus-first guess (OMG) and observation-minus-analysis (OMA) differences with respect to all conventional observations assimilated in ALADIN (Strajnar et al. 2015).

Figure 7 shows the normalized RMS change of OMA and OMG using different VarBC configurations where the VarBC-global (ARP) represents the reference experiment. Although the OMA difference is not a measure of analysis quality, this diagnostic indicates the change of analysis when different bias correction methods are used. Note that using VarBC-LAM methods improves the analysis fit compared to the global bias coefficients, especially

TABLE 3. Trials used for initialization (passive VarBC-mode) and evaluation (active DA) of the studied VarBC methods in ALADIN.

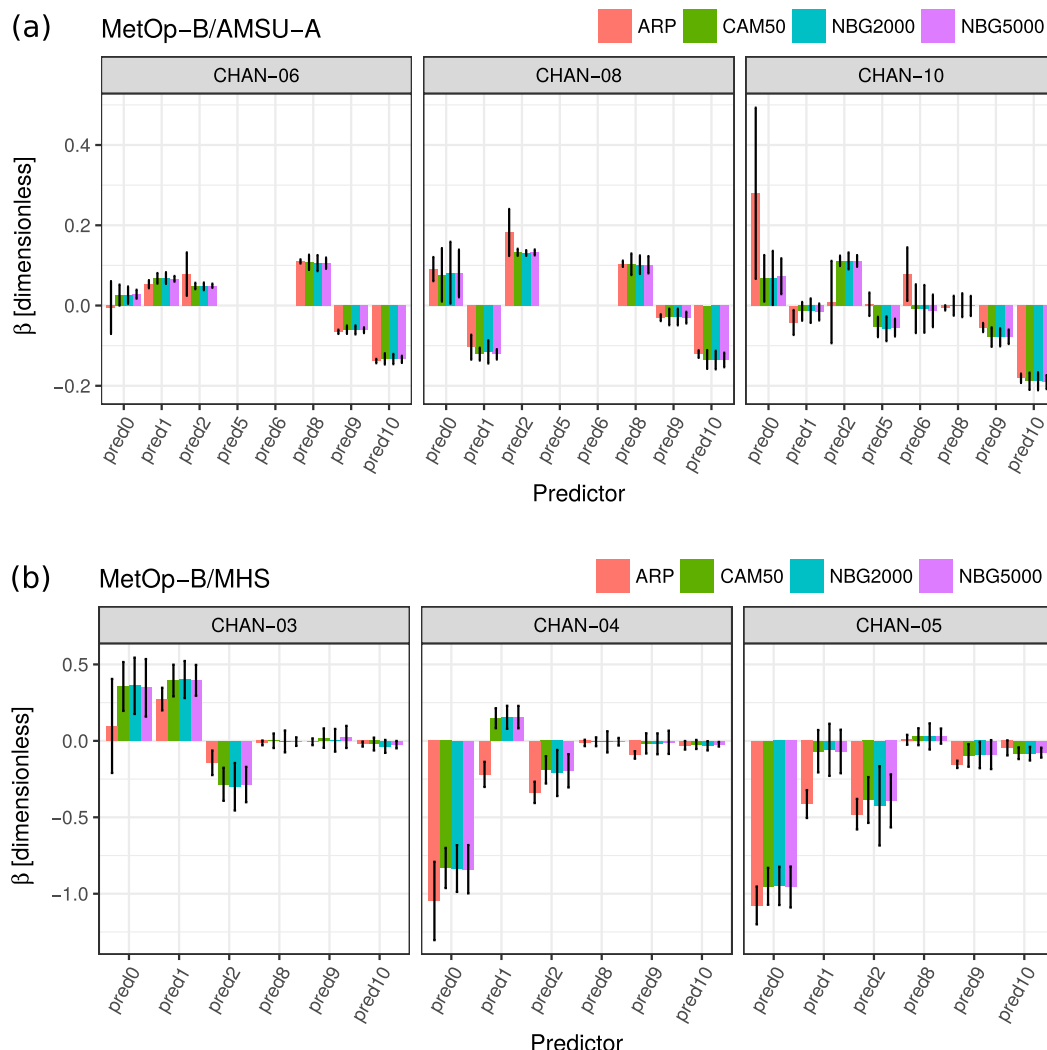| Experiment	Method	N_{bg}	n_h	N_{min}
	NBG5000	5000	—	100
Initialization	NBG2000	2000	—	100
(15 Oct–30 Nov 2015)	CAM50	Eq. (9)	15	50
	ARPEGE	5000	—	100
	NBG5000	5000	—	100
Evaluation	NBG2000	2000	—	100
(1 Dec 2015–31 Jan 2016)	CAM100	Eq. (9)	5	100
	CAM50	Eq. (9)	5	50

FIG. 6. The mean of bias coefficients for particular (a) AMSU-A and (b) MHS channels calculated for the VarBC-global and VarBC-LAM configurations in December 2015. Error bars represent the standard error of the mean of bias coefficients.

for temperatures between 100 and 800 hPa with respect to AMDAR and TEMP observations. The decrease of OMG differences for VarBC-LAM indicates a better quality of first guess (3-h forecast) with respect to most of the observations. Note that VarBC-LAM methods improve at 95% significance level the RMS of first guess with the largest normalized impact of 3% for temperature and 2% for wind components between 400 and 800 hPa. There is no statistically significant impact of VarBC-LAM methods on the first-guess BIAS score (not shown).

The VarBC-LAM configurations NBG5000, NBG2000, and CAM100 are comparable at 95% significance level in terms of the first-guess quality (Fig. 7). Regarding the CAM50 method, including small data samples has no

significant impact on analysis but a significant negative impact on first guess, with the largest normalized RMS impact of 1% for specific humidity between 800 and 1000 hPa. Assimilation of small data samples has no significant impact on the first-guess BIAS score (not shown).

c. Forecast impact study

The forecast quality is measured by RMS and BIAS of forecast-minus-observation differences with respect to SYNOP, TEMP, and AMDAR observations.

The use of VarBC-LAM methods improves the RMS of the short-time (6–18 h) forecast at 95% significance level with respect to the global bias coefficients (see Fig. 8). The improvement is detected for the RMS of
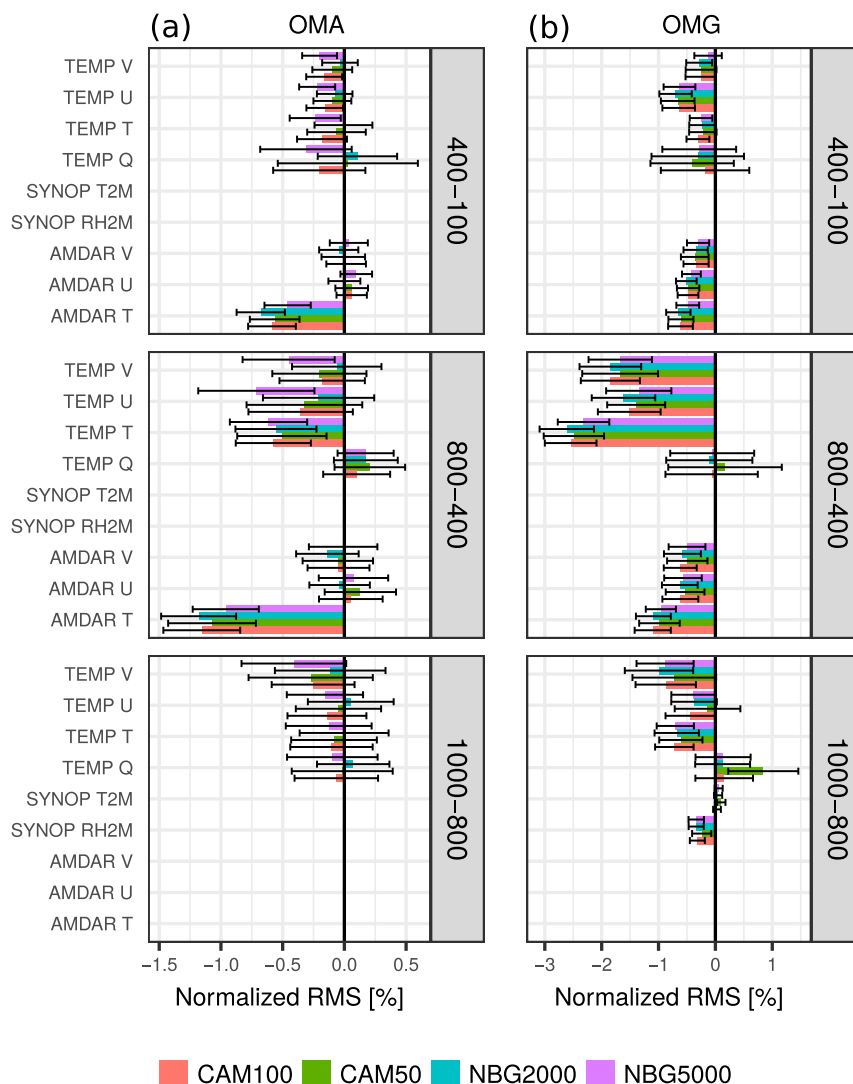
Fig. 7. The normalized RMS of (a) OMA and (b) OMG residuals are evaluated for different VarBC-LAM methods with respect to particular conventional observations from 1 Dec 2015 to 31 Jan 2016. The VarBC-global method (ARP) is used as a reference. Improvement (degradation) of the VarBC-LAM methods with respect to the reference is represented by negative (positive) values. Error bars represent 95% significance level.

temperature (6–18-h forecast) with the largest normalized impact of 0.6% between 400 and 800 hPa and wind speed (6-h forecast) with the largest impact of 0.5% between 400 and 800 hPa. The impact of regional versus global VarBC coefficients on the 24–48-h forecast is comparable since the 95% confidence intervals include zero. There is no significant difference in the forecast BIAS scores between VarBC-global and VarBC-LAM methods (not shown).

The VarBC-LAM methods NBG5000, NBG2000, and CAM100 have comparable impact at 95% significance level on 6–48-h forecast quality in terms of RMS (see

Fig. 8) and BIAS (not shown). The forecast impact of small data samples in CAM50 is evaluated with respect to CAM100 used as a reference (not shown). The CAM50 method has a negative RMS impact at 95% significance level on temperature (6–18-h forecast) with the largest normalized impact of 0.3% between 800 and 1000 hPa. There is no significant impact of CAM50 on the forecast BIAS score (not shown).

Since no instrument bias changes were recorded by the EUMETSAT User Notification Service (UNS) during the validation period, the forecast performance of various VarBC-LAM configurations is comparable.
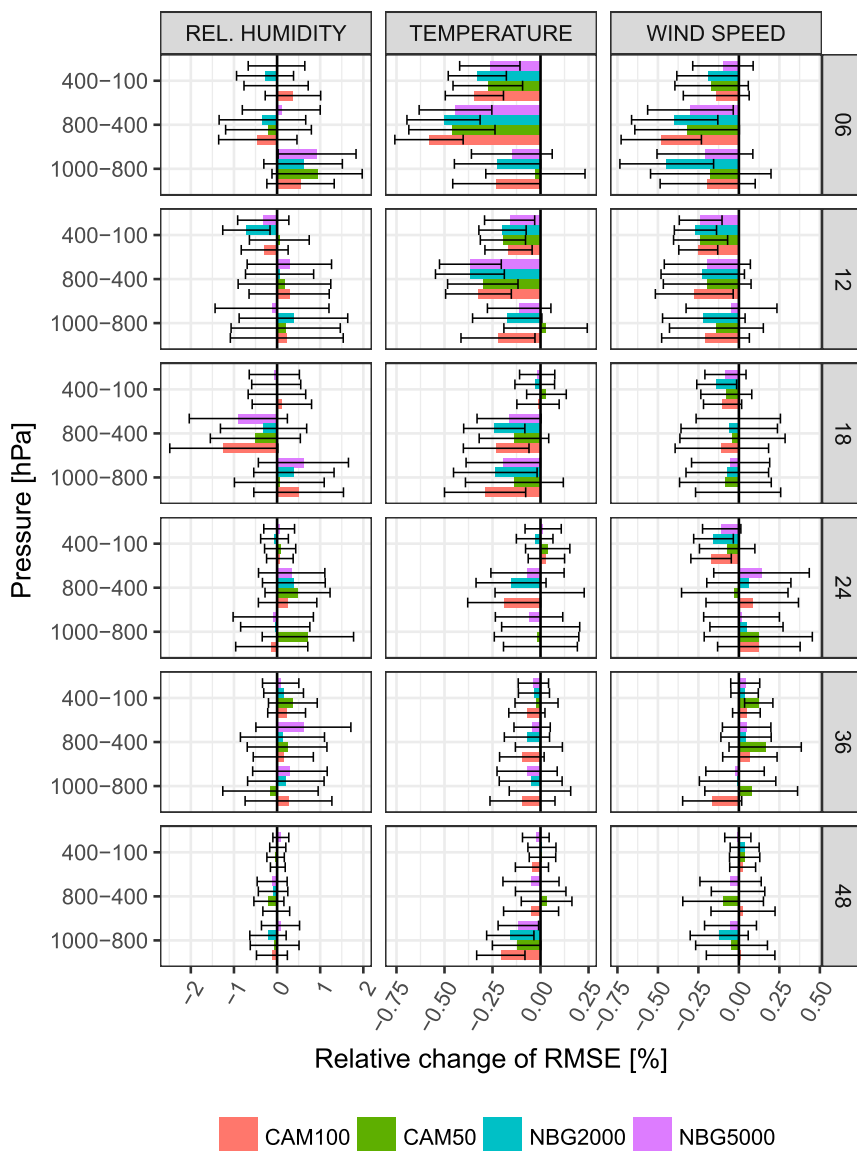
FIG. 8. The normalized RMS impact of various VarBC-LAM methods NBG5000, NBG2000, CAM100, and CAM50 on the 6–48-h forecast for humidity, temperature, and wind speed. The VarBC-global (ARP) is used as a reference. The forecast scores are evaluated from 1 Dec 2015 to 31 Jan 2016 with respect to AMDAR and TEMP measurements separated into three vertical layers. Error bars represent 95% significance level.

However, some differences can be expected in the event of a failure of a satellite instrument. In this case the best VarBC configuration should immediately react to the instrument bias changes, as described in the next paragraph.

d. VarBC response to artificial bias

A response of VarBC-LAM methods NBG5000, NBG2000, and CAM100 to the artificial bias 0.2 K is studied for a particular AMSU-A channel 6 on *NOAA-19*

in the active DA system. The control experiment (CTRL), without activating the artificial bias, is used as a reference.

Figure 9 presents the performance of the VarBC-LAM methods on the artificial bias correction with respect to 0300 UTC (Fig. 9a) and 1200 UTC (Fig. 9b) analysis times. Note that NBG2000 and CAM100 methods reduce the artificial bias equally effectively in approximately 1-month period. A residual bias of approximately 15% (0.03 K) still remains for NBG5000 after this spinup
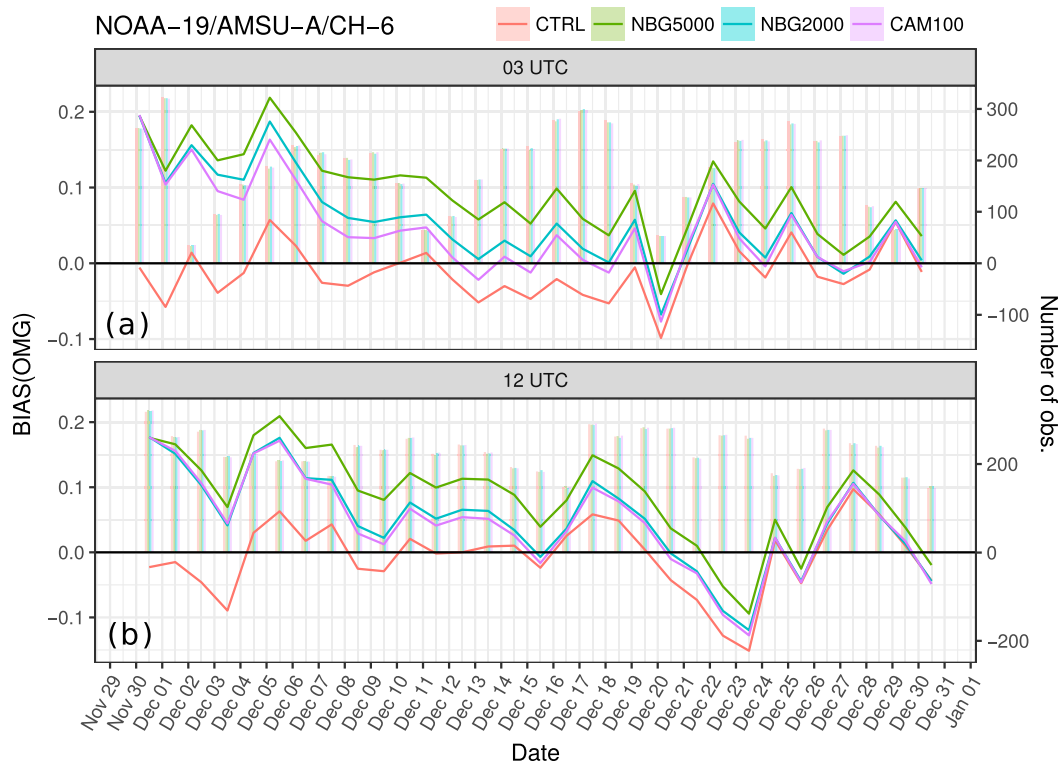
FIG. 9. Reaction of various VarBC-LAM methods on the artificial bias 0.1 K set for the AMSU-A channel 6 on *NOAA-19*. Time-evolution of the mean of OMG residuals is monitored at particular analysis times (a) 0300 and (b) 1200 UTC when the satellite provides a different number of observations (bar plots) in the ALADIN domain.

period. A little advantage of CAM100 is detected at 0300 UTC analysis time when the number of satellite measurements changes significantly from 20 to 300 for the particular channel. Note that the CAM100 method provides faster adaptation compared to NBG2000 and NBG5000, especially in the first half of the tested period. This underlines the fact that NBG2000 and NBG5000 performs well for larger number of measurements while the performance in bias correction decreases in case of small data samples (see Fig. 4).

6. Summary and discussion

Application of the VarBC scheme was studied in the limited-area model ALADIN with the daily update of VarBC bias coefficients (Randriamampianina et al. 2011). The VarBC scheme was initialized by (i) global bias coefficients adopted from the ARPEGE model at each analysis time (VarBC-global); and (ii) bias coefficients estimated/cycled independently in the LAM (VarBC-LAM). The latter method was studied with various stiffness parameter N_{bg} settings: 5000 (NBG5000), 2000 (NBG2000), and using the Cameron and Bell (2016) approach (CAM).

The daily VarBC update strategy requires a minimum expected observation sample N_{min} that must be available at each analysis time to ensure a proper response of the VarBC in case of sudden bias changes. Therefore instrument channels providing $N_{avg} \leq N_{min}$ were removed from the DA system. In ALADIN, we set N_{min} to 100 observations for the VarBC configurations ARP, NBG5000, NBG2000, and CAM100. Additionally, we set N_{min} to 50 observations for CAM50 to study the forecast impact of small and odd satellite paths.

The VarBC-LAM bias coefficients were initialized in the ALADIN DA system during a one-and-half month period. The global versus regional VarBC bias coefficients were comparable with respect to the geometric bias correction. On the other hand, significant differences were detected for bias coefficients associated with constant bias-offset and airmass predictors. The differences are most likely conditioned by specific weather regimes in the winter period. The second reason is probably a different vertical resolution of the ALADIN and ARPEGE models in stratosphere and boundary layers as well as more accurate description of the LAM ALADIN physics. It is worth noting that adopting global bias coefficients to LAMs is conditioned by ensuring consistency of the VarBC scheme and the biases

between the regional and global models (Kazumori 2014). This may not be the case for each NWP mesoscale models. To ensure consistency in forecast biases, it would require one to employ for example a unified, multiscale physics package, which is very hard to achieve, especially when LAM models are tailored for high resolutions. This reinforces the need to deploy the VarBC-LAM method.

The forecast impact of different VarBC configurations was evaluated during a two-month winter period. In this case, the NWP model error of 500-hPa geopotential is generally higher due to larger amplitude of Rossby waves (Geer 2016). Thus we also expected higher differences in forecast performance between the VarBC configurations. Supposing that VarBC-LAM bias coefficients are most likely conditioned by weather regimes, a longer experimental period would be desirable to overcome this effect. In summary, the NBG2000 and CAM100 provided the best performance of the VarBC scheme compared to global bias coefficients and other VarBC-LAM methods. The use of VarBC-LAM methods improved the short-term 6–18-h forecast, with the largest normalized impact of 3%. This magnitude is comparable with other forecast impact studies performed in mesoscale NWP models evaluating [e.g., the impact of radiance (Lin et al. 2017) or Mode-S MRAR (Strajnar et al. 2015) observations]. On the other hand, VarBC-LAM methods had little impact beyond the 24-h forecast compared to the global bias coefficients in this study.

Regarding the CAM method, it should be noted that the bias halving time parameter $\mathbf{n}_h$ was derived empirically in this study ($\mathbf{n}_h$ = 5). This setting may not be optimal for the entire DA system, and further tuning can lead to a better performance.

7. Conclusions

The present study was designed to assess various VarBC configurations in the regional ALADIN/CHMI model, including a 1.5 month initialization and 2 month evaluation period. A fair comparison of VarBC-LAM and VarBC-global initialization methods was presented and a proper VarBC stiffness setting was found for the limited-area model with respect to the forecast performance and the response of VarBC to sudden bias changes.

Regarding the initialization of the VarBC scheme, the best forecast performance was achieved by cycling of bias coefficients independently in the LAM. The VarBC-LAM methods significantly improved the RMS of the model first guess (3-h forecast) with respect to the global bias coefficients, with the largest normalized

impact of 3% for temperature and 2% for wind components in middle troposphere. The positive impact of VarBC-LAM was also detected on the 6–18-h forecast for the RMS of temperature and wind speed in middle troposphere, with the largest normalized impact of approximately 0.5%. The impact of global versus regional VarBC coefficients on the 24–48-h forecast was comparable at the 95% significance level.

The comparison of various VarBC-LAM adaptivity settings appeared that NBG2000 and CAM100 provided the best performance of the VarBC scheme in the LAM. Although, all the VarBC-LAM methods had comparable impact on the 0–48-h forecast at the 95% significance level, significant differences were detected in terms of the VarBC response to artificial bias. In case of uniform satellite data coverage, the NBG2000 and CAM100 methods removed the artificial bias comparably effective in the one-month period while NBG5000 still left a residual bias of approximately 15%. The best use of the CAM100 method was for situations with nonuniform satellite data coverage. The dynamic VarBC stiffness setting adjusted bias coefficients with regards to $\mathbf{N}_{avg}$ ensuring a stable response of the VarBC scheme to bias changes across instrument channels and analysis times.

Assimilating measurements from the small and odd satellite paths deteriorated forecast performance. Including additional instrument channels with $50 < \mathbf{N}_{avg} < 100$ (CAM50) into the DA system had a significant negative impact on the first guess of specific humidity in boundary layers at the 95% significance level, with the largest normalized RMS impact of 1% with respect to CAM100. A significant degradation was also detected for the short-term (6–18 h) forecast of temperature in boundary layers, with the largest normalized RMS impact of 0.3%.

In future years, it is planned to assimilate the particular polar-orbiting satellite instruments into the operational ALADIN/CHMI model together with the VarBC scheme. Based on the results of this paper, the VarBC scheme should be implemented with: cycling the bias coefficients independently in the LAM system and using the daily cycling strategy together with the dynamic VarBC stiffness settings proposed by Cameron and Bell (2016). The future development of the LAM VarBC scheme could be focused on: tuning of the bias halving time parameter $\mathbf{n}_h$ with respect to the best forecast performance, and setting $\mathbf{n}_h$ parameter individually for each predictor or groups of predictors.

REFERENCES

Auligné, T., 2007: An objective approach to modelling biases in satellite radiances: Application to AIRS and AMSU-A. *Quart. J. Roy. Meteor. Soc.*, **133**, 1789–1801, https://doi.org/10.1002/qj.145.

——, A. P. 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.

Benjamin, S. G., B. E. Schwartz, E. J. Szoke, and S. E. Koch, 2004: The value of wind profiler data in U.S. weather forecasting. *Bull. Amer. Meteor. Soc.*, **85**, 1871–1886, https://doi.org/10.1175/BAMS-85-12-1871.

——, B. D. Jamison, W. R. Moninger, S. R. Sahm, B. E. Schwartz, and T. W. Schlatter, 2010: Relative short-range forecast impact from aircraft, profiler, radiosonde, VAD, GPS-PW, METAR, and mesonet observations via the RUC hourly assimilation cycle. *Mon. Wea. Rev.*, **138**, 1319–1343, https://doi.org/10.1175/2009MWR3097.1.

Bouttier, F., and P. Courtier, 1999: Data assimilation concepts and methods. Meteorological Training Course Lecture Series, ECMWF, 59 pp., https://www.ecmwf.int/sites/default/files/elibrary/2002/16928-data-assimilation-concepts-and-methods.pdf.

Bučánek, A., A. Trojáková, and R. Brožková, 2015: The BlendVar assimilation scheme at CHMI. *Meteor. Bull.*, **68**, 180–185.

Cameron, J., and W. Bell, 2016: The testing and planned implementation of variational bias correction (VarBC) at the Met Office. Met Office, 21 pp., https://cimss.ssec.wisc.edu/itwg/itsc/itsc20/papers/11_01_cameron_paper.pdf.

Campins, J., A. J. Sánchez, M. V. Diez Muyo, F. J. Calvo Sánchez, and B. Navascués, 2017: Assimilation of ATOVS and GNSS ZTD data in the HARMONIE-AROME model configuration run at AEMET. *ALADIN-HIRLAM Newsletter*, No. 8, 40–50, https://repositorio.aemet.es/bitstream/20.500.11765/6833/1/Assimilation_Campins.pdf.

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*, Reading, United Kingdom, ECMWF, Vol. 28, 97–112.

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

——, and R. Todling, 2000: Data assimilation in the presence of forecast bias: The GEOS moisture analysis. *Mon. Wea. Rev.*, **128**, 3268–3282, https://doi.org/10.1175/1520-0493(2000)128<3268:DAITPO>2.0.CO;2.

Evensen, G., 1994: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.*, **99**, 10 143–10 162, https://doi.org/10.1029/94JC00572.

Eyre, J., 1992: A bias correction scheme for simulated TOVS brightness temperatures. ECMWF Tech. Memo.186, ECMWF, 35 pp.

Fertig, E., and Coauthors, 2009: Observation bias correction with an ensemble Kalman filter. *Tellus*, **61A**, 210–226, https://doi.org/10.1111/j.1600-0870.2008.00378.x.

Fischer, C., T. Montmerle, L. Berre, L. Auger, and S. E. Ştefănescu, 2005: An overview of the variational assimilation in the ALADIN/France numerical weather-prediction system. *Quart. J. Roy. Meteor. Soc.*, **131**, 3477–3492, https://doi.org/10.1256/qj.05.115.

Geer, A. J., 2016: Significance of changes in medium-range forecast scores. *Tellus*, **68A**, 30229, https://doi.org/10.3402/tellusa.v68.30229.

Gérard, E., F. Rabier, D. Lacroix, and Z. Sahlaoui, 2003: Use of ATOVS raw radiances in the operational assimilation system at Météo-France. *Proc. ITSC XIII*, Ste. Adéle, Canada, Météo-France,18–29.

Guidard, V., N. Fourrié, P. Brousseau, and F. Rabier, 2011: Impact of IASI assimilation at global and convective scales and challenges for the assimilation of cloudy scenes. *Quart. J. Roy. Meteor. Soc.*, **137**, 1975–1987, https://doi.org/10.1002/qj.928.

Gustafsson, N., and Coauthors, 2018: Survey of data assimilation methods for convective-scale numerical weather prediction at operational centres. *Quart. J. Roy. Meteor. Soc.*, **144**, 1218–1256, https://doi.org/10.1002/qj.3179.

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.

Järvinen, H., and P. Undén, 1997: Observation screening and background quality control in the ECMWF 3D-Var data assimilation system. ECMWF Tech. Memo. 236, ECMWF, 34 pp.

Karbou, F., É. Gérard, and F. Rabier, 2006: Microwave land emissivity and skin temperature for AMSU-A and -B assimilation over land. *Quart. J. Roy. Meteor. Soc.*, **132**, 2333–2355, https://doi.org/10.1256/qj.05.216.

Kazumori, M., 2014: Satellite radiance assimilation in the JMA operational mesoscale 4DVAR system. *Mon. Wea. Rev.*, **142**, 1361–1381, https://doi.org/10.1175/MWR-D-13-00135.1.

Lahoz, W., B. Khattatov, and R. Menard, 2010: *Data Assimilation: Making Sense of Observations.* Springer-Verlag, 718 pp.

Lin, H., S. S. Weygandt, S. G. Benjamin, and M. Hu, 2017: Satellite radiance data assimilation within the hourly updated rapid refresh. *Wea. Forecasting*, **32**, 1273–1287, https://doi.org/10.1175/WAF-D-16-0215.1.

Lindskog, L., M. Dahlbom, S. Thorsteinsson, P. Dahlgren, R. Randriamampianina, and J. Bojarova, 2012: ATOVS processing and usage in the HARMONIE reference system. *HIRLAM Newsletter*, Vol. 59, HIRLAM, 33–43.

Liu, Z., C. S. Schwartz, C. Snyder, and S. Ha, 2012: Impact of assimilating AMSU-A radiances on forecasts of 2008 Atlantic tropical cyclones initialized with a limited-area ensemble Kalman filter. *Mon. Wea. Rev.*, **140**, 4017–4034, https://doi.org/10.1175/MWR-D-12-00083.1.

Montmerle, T., F. Rabier, and C. Fischer, 2007: Relative impact of polar-orbiting and geostationary satellite radiances in the Aladin/France numerical weather prediction system. *Quart. J. Roy. Meteor. Soc.*, **133**, 655–671, https://doi.org/10.1002/qj.34.

Randriamampianina, R., 2005: Radiance-bias correction for a limited area model. *Quart. J. Hungarian Meteor. Service*, **109** (3), 143–155.

——, T. Iversen, and A. Storto, 2011: Exploring the assimilation of IASI radiances in forecasting polar lows. *Quart. J. Roy. Meteor. Soc.*, **137**, 1700–1715, https://doi.org/10.1002/qj.838.

Romine, G. S., C. S. Schwartz, C. Snyder, J. L. Anderson, and M. L. Weisman, 2013: Model bias in a continuously cycled assimilation system and its influence on convection-permitting forecasts. *Mon. Wea. Rev.*, **141**, 1263–1284, https://doi.org/10.1175/MWR-D-12-00112.1.

Saunders, R., M. Matricardi, and P. Brunel, 1999: An improved fast radiative transfer model for assimilation of satellite

radiance observations. *Quart. J. Roy. Meteor. Soc.*, **125**, 1407–1425, https://doi.org/10.1002/qj.1999.49712555615.

Schwartz, C. S., Z. Liu, Y. Chen, and X.-Y. Huang, 2012: Impact of assimilating microwave radiances with a limited-area ensemble data assimilation system on forecasts of typhoon morakot. *Wea. Forecasting*, **27**, 424–437, https://doi.org/10.1175/WAF-D-11-00033.1.

Strajnar, B., N. Žagar, and L. Berre, 2015: Impact of new aircraft observations Mode-S MRAR in a mesoscale NWP model. *J. Geophys. Res. Atmos.*, **120**, 3920–3938, https://doi.org/10.1002/2014JD022654.

Termonia, P., and Coauthors, 2018: The ALADIN system and its canonical model configurations AROME CY41T1 and ALARO CY40T1. *Geosci. Model Dev.*, **11**, 257–281, https://doi.org/10.5194/gmd-11-257-2018.

Torn, R. D., and C. A. Davis, 2012: The influence of shallow convection on tropical cyclone track forecasts. *Mon. Wea. Rev.*, **140**, 2188–2197, https://doi.org/10.1175/MWR-D-11-00246.1.

Wilks, D. S., 2011: *Statistical Methods in the Atmospheric Sciences*. 3rd ed. International Geophysics Series, Vol. 100, Academic Press, 704 pp.