Near-Surface Wind Observation Impact on Forecasts: Temporal Propagation of the Analysis Increment

JOËL BÉDARD

Centre pour l'étude et la simulation du climat à l'échelle régionale (ESCER), Department of Earth and Atmospheric Sciences, Université du Québec à Montréal, Montréal, and Data Assimilation and Satellite Meteorology Section, Environment and Climate Change Canada, Dorval, Québec, Canada

STÉPHANE LAROCHE

Data Assimilation and Satellite Meteorology Section, Environment and Climate Change Canada, Dorval, Québec, Canada

PIERRE GAUTHIER

Centre pour l'étude et la simulation du climat à l'échelle régionale (ESCER), Department of Earth and Atmospheric Sciences, Université du Québec à Montréal, Montréal, Québec, Canada

(Manuscript received 12 August 2016, in final form 5 December 2016)

ABSTRACT

This study examines the assimilation of near-surface wind observations over land to improve wind nowcasting and short-term tropospheric forecasts. A new geostatistical operator based on geophysical model output statistics (GMOS) is compared with a bilinear interpolation scheme (Bilin). The multivariate impact on forecasts and the temporal evolution of the analysis increments produced are examined as well as the influence of background error covariances on different components of the prediction system. Results show that Bilin significantly degrades surface and upper-air fields when assimilating only wind data from 4942 SYNOP stations. GMOS on the other hand produces smaller increments that are in better agreement with the model state. It leads to better short-term near-surface wind forecasts and does not deteriorate the upper-air forecasts. The information persists longer in the system with GMOS, although the local improvements do not propagate beyond 6-h lead time. Initial model tendencies indicate that the mass field is not significantly altered when using static error covariances and the boundary layer parameterizations damp the poorly balanced increment locally. Conversely, most of the analysis increment is propagated when using flow-dependent error statistics. It results in better balanced wind and mass fields and provides a more persistent impact on the forecasts. Forecast accuracy results from observing system experiments (assimilating SYNOP winds with all observations used operationally) are generally neutral. Nevertheless, forecasts and analyses from GMOS are more self-consistent than those from both Bilin and a control experiment (not assimilating near-surface winds over land) and the information from the observations persists up to 12-h lead time.

1. Introduction

In a continuous effort to improve short-term wind forecasts for different applications such as wind energy, airport operations, and road safety to name a few, recent studies proposed the assimilation of near-surface wind observations over land in numerical weather prediction (NWP) systems. Surface observations have smaller correlations with the flow aloft compared to integrated variables such as surface pressure. Their impact on analyses varies according to the atmospheric boundary layer (ABL) coupling with the upper air, which depends on atmospheric stability. Stationary background error covariances and simple balance relationships do not capture this well. Indeed, improvements from the assimilation of near-surface wind observations appear to be more significant within an assimilation system based on an ensemble Kalman filter (EnKF) rather than a three-dimensional variational data assimilation (3DVAR) system (Ancell et al. 2015) because EnKF samples flow-dependent (in time and space) background

Corresponding author e-mail: Joël Bédard, bedard.joel@gmail.com

DOI: 10.1175/MWR-D-16-0310.1

error covariances. Pu et al. (2013) also showed that the EnKF allows for a better spatial distribution of the analysis increment over complex terrain. The assimilation of near-surface wind observations with flow-dependent background error covariances has the potential to improve temperature, humidity and winds profiles (Hacker and Rostkier-Edelstein 2007), but results are sensitive to the way flow-dependent error statistics are handled by the assimilation system (Hacker et al. 2007; Pu et al. 2013). The impact of flow-dependent background error covariances on the assimilation of near-surface wind observations over land is encouraging, but there are still issues that need to be resolved. Although near-surface observations can constrain the state of the ABL, it is not yet fully understood why the wind observations only have a positive impact on very short-term local forecasts (Hacker and Snyder 2005; Rostkier-Edelstein and Hacker 2010; Zack et al. 2010, 2011; Ancell et al. 2011, 2015).

Near-surface wind observations sample finescale structures not explicitly resolved by NWP models. Many near-surface wind observations over land are available from the global observing system but were not used in most data assimilation systems until recently because of the discrepancy between the characteristics of the measured and forecasted variables. The model misrepresentation of surface characteristics generates so-called representativeness errors. Biases and representativeness errors limit the global influence of near-surface wind observations on operational forecast systems (Ingleby 2015). Winds over small islands, subgrid-scale headlands, and over land throughout the tropics are still excluded from the Met Office data assimilation system. Similarly, the NCEP Rapid Update Cycle (Benjamin et al. 2007, 2010) as well as the research group at University of Washington (Dirren et al. 2007; Ancell et al. 2011) use strict quality-control checks to reject observations likely affected by biases and representativeness errors. Other operational centers [e.g., Environment and Climate Change Canada (ECCC)] simply omit all wind observations from land stations to prevent degrading the near-surface wind analyses due to biases and representativeness errors.

Directly assimilating biased observations in a bias blind system undeniably produces biased analyses (Dee 2005). Still, removing model biases prior to the assimilation of surface observations only improves the analysis and nowcasting capability as the forecast quickly returns toward the model's own bias (Ancell et al. 2011; Ancell 2012). As an alternative to improve the model forecast representation in observation space, Bédard et al. (2013) proposed a geostatistical observation operator based on a multiple gridpoint approach called geophysical model output statistics (GMOS). As discussed in Bédard et al. (2015), this can alleviate issues associated with representativeness and systematic error. GMOS attributes higher weights to the most representative grid points and takes into account the natural on-site variability to better represent meteorological phenomena locally. It corrects the stationary and isotropic components of systematic and representativeness errors associated with local geographical characteristics (e.g., surface roughness or coastal effects). Using GMOS as the observation operator in data assimilation improves the consistency between background states and observations, which results in smaller innovations and analysis increments than what was obtained previously with a bilinear interpolation scheme used for in situ observations (Bilin).

The objective of this paper is to understand the multivariate impact of near-surface wind observations over land and assess the temporal propagation of their analysis increments. A description of the assimilation system is presented in the next section including the GMOS operator, the quality-control procedure, and the observation error characteristics for the surface wind data. Section 3 presents simple assimilation experiments to assess the impact of only assimilating near-surface wind observations. A large set of analyses were performed by using background states taken from a full assimilation experiment with ECCC's data assimilation system. These simplified assimilation experiments provide a framework to compare the impact of different assimilation strategies. Systematic model initial tendencies are used to assess the influence of the observation operator, the background error statistics, and different components of the NWP model (e.g., the boundary layer parameterization) on the temporal evolution of the analysis increment. Section 4 presents observing system experiments (OSEs) in which the cumulated impact of near-surface wind data over a 1-month period is assessed in an operational context. Special cases in which the assimilation of near-surface winds has a significant impact are examined. The paper ends by presenting conclusions in section 5.

2. Experimental framework

The experiments presented in this study are based on the global four-dimensional ensemble–variational data assimilation system (4DEnVar) developed at ECCC (Buehner et al. 2013, 2015). ECCC's global deterministic prediction system (GDPS), based on the Global Environmental Multiscale model (GEM; Côté et al. 1998; Charron et al. 2012; Zadra et al. 2014), has a relatively high resolution (0.35° × 0.23° latitude–longitude resolution: ~25 km), similar to the gridpoint spacing employed in other studies on the assimilation of

near-surface winds over land [e.g., 20 km in Benjamin et al. (2007); 27 km in Pu et al. (2013); 25 km in Ingleby (2015)]. The GDPS was chosen because it is well developed and it relies on well-calibrated flow-dependent background error covariances from the operational EnKF (Houtekamer et al. 2014). The higher-resolution regional system was not used because its data assimilation component is not fully cycled. Thus, it would not have been possible to assess the cumulative impact of observations. Moreover, using an incremental formulation (Laroche et al. 1999; Gauthier et al. 2007; Buehner et al. 2015), the global and regional analysis increments are generated on the same global lower-resolution grid with the same observations (only those located over the domain of the model) and the same error statistics because high-resolution flow-dependent background error statistics are still in development.

a. The data assimilation system

The operational 4DEnVar uses 4D ensemble-based background error statistics that comprise a stationary homogeneous (static) and a flow-dependent (dynamic) component to produce spatiotemporal analysis increments. The static component is estimated using the National Meteorological Center (NMC, now known as NCEP) method (Parrish and Derber 1992; Gauthier et al. 1998) and the flow-dependent component uses 256 ensemble members from ECCC's operational EnKF. The resolution of the analysis increment is determined by the resolution of the background error covariances (i.e., 800 × 400 grid, ~50-km horizontal grid spacing). As discussed in Laroche et al. (1999), if the innovations are computed with respect to the high-resolution background state, then the analysis increments can be computed at a coarser-resolution corresponding to the scales resolved by the background error covariances without any loss of information. The two background error components are blended equally below 40 hPa and gradually weighted toward the static error statistics above. The system can be run using only the static or the EnKF background error components. This feature is convenient to test the impact of the different background error components. More details on the 4DEnVar can be found in Buehner et al. (2015).

b. Geostatistical observation operator

Forecast values at the nearest model grid points of a surface weather station may not properly represent the observations, especially if the station is located on complex terrain or coastal sites. On the other hand, among the surrounding grid points, there are generally one or several points that are more representative of the observing site. Thus, GMOS uses a set of geostatistical weights to relate the most representative grid points with the observation site. From Bédard et al. (2013), the GMOS multipoint linear regression (H_{GMOS}), at the kth station, is formulated as follows:

$$H^{(k)}_{\mathrm{GMOS}}(x) = \sum_i (A^{(k)}_i x_i) + C^{(k)}, \tag{1}$$

where x is either the latitudinal (u) or longitudinal (v) modeled wind component (the regression is done independently for u and v), $C^{(\mathrm{k})}$ is the systematic error correction coefficient, and the subscript i is the index of the four closest grid points to the kth observing site in a 2 × 2 square pattern. The amplitude coefficients $A^{(k)}_i$ weight the surrounding grid points and adjust the forecast variability to best fit the observations. The wind speed and direction observations from SYNOP stations (available every 3 h) are transformed into wind components to compute two sets of statistical coefficients [$A^{(k)}_i$ and $C^{(k)}$] at each station according to Eq. (1). The least mean square error minimization algorithm is applied for u and v independently by comparing the observed wind components with corresponding short-range (3–12 h) wind forecasts produced twice a day by ECCC's GDPS. The forecasts used to train the observation operators and for the assimilation experiments are generated with the same model configuration.

GMOS combines a multiple gridpoint approach with statistical error corrections to take advantage of the correlation between resolved and unresolved scales. By extension, if the surrounding grid points are only weakly representative, the geostatistical coefficients will converge to zero and GMOS will give a reduced weight to the observation in the assimilation system (i.e., the observation impact on the analysis is reduced). Thus, GMOS implicitly characterizes the stationary part of the representativeness error of the observations. Consequently, quality-control criteria based on a difference between modeled and observed station heights (e.g., Pu et al. 2013) or based on large innovation variances [as in NCEP's Rapid Refresh (RAP) system] are not used in this study. More details on the use of GMOS in data assimilation can be found in Bédard et al. (2015).

c. Observation quality control

The quality control of the observations consists of a gross error check, a background check, and a variational quality control. Gross error and background checks are also performed on the training dataset because GMOS relies on the availability of good quality data to train its statistical coefficients (Bédard et al. 2013, 2015). Only the 5331 stations benefiting from a minimum of 2 months of data are considered because a GMOS operator based on a 2 × 2 gridpoint stencil requires a

minimum of 2 months of training data (Bédard et al. 2015). The background check is sensitive to the observation operator and it can bias the GMOS statistical coefficients. Therefore, the coefficients are computed and updated through an iterative process. From the 8 months of data in the training dataset available (November 2012–June 2013), 6 months are used to compute the statistical coefficients (training dataset) and 2 months are kept for cross validation to evaluate the statistical robustness of the GMOS operator at each site. The coefficients are first computed using the training dataset. Then they are used within the background check of the next iteration to filter out erroneous observations and they are recomputed. This last step is repeated until the GMOS coefficients converge. For most stations, the coefficients rapidly converge toward stable values (e.g., after two or three iterations). A maximum of 10 iterations is allowed to accommodate stations presenting fewer observations. Although the typical sample size varies between 500 and 1500 observation reports per station, 37 stations did not converge because the number of good observations is too small (some stations do not provide data for some periods of time, while many others only report irregularly). In such a case, a slight change in the coefficients may alter significantly the relative quantity of good data (from the few available). These stations are blacklisted because they cannot provide robust coefficients to the geostatistical operator. Still, because of a lack of observations at 205 of the sites, the estimated GMOS operator at these sites was not sufficiently robust to provide a better fit, as compared with Bilin, to the observations in the cross-validation dataset. Those sites, along with sites where the observation errors is above $5\,\mathrm{m\,s^{-1}}$ (147 stations) are blacklisted.

Although each experiment has its own quality control, the Bilin experiment uses the same observing stations as the GMOS experiment to ensure that all experiments benefit from the same number of observations to perform a fair evaluation. The quality control of the observing stations allows for a high data recovery (92.7%) and the resulting 4942 SYNOP stations are shown in Fig. 1. All stations presenting sufficient training data to obtain robust GMOS coefficients are used and all surface wind observations having wind speed above $1\,\mathrm{m\,s^{-1}}$ are assimilated, provided that they pass background and gross error checks (wind vanes have a poor accuracy at low wind speed; Bédard et al. 2015).

d. Observation error statistics

GMOS reduces the representativeness part of the observation error and removes the observation error bias. Bédard et al. (2015) showed that it is possible to diminish the global observation error statistics (σ_o) prescribed in the assimilation system when using GMOS rather than Bilin (from 1.98 to $1.79\,\mathrm{m\,s^{-1}}$). GMOS considers local representativeness errors independently for each observation site. In this case, the observation error statistics can also be site dependent. The observation error variances (σ_o^2) are computed at each site from innovation variances using the optimal background error variances to be used in the assimilation system [σ_b^2 for near-surface winds over land: as computed in Bédard et al. (2015)] following

$$\langle\{[\mathbf{y}-H(\mathbf{x}_b)]-\langle\mathbf{y}-H(\mathbf{x}_b)\rangle\}^2\rangle=\sigma_o^2+\sigma_b^2, \qquad (2)$$

where observations $\mathbf{y}$ are compared with the background state $\mathbf{x}_b$ using an observation operator H. Here, $\langle\,\rangle$ stands for the statistical average. Figure 2 shows the frequency distribution ($0.05\,\mathrm{m\,s^{-1}}$ bin intervals) of the computed site dependent observation error standard deviation (STD) using GMOS for the 4942 sites considered in this study.

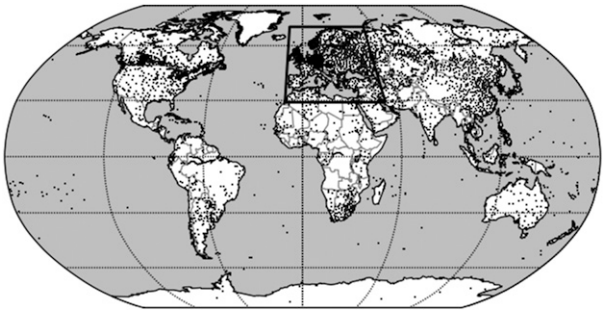

FIG. 1. Spatial distribution of the 4942 SYNOP sites considered in this study (black dots). The rectangle refers to the area where the upper-air evaluation is performed.

e. Forecast evaluation dataset

Observations and analyses are used to evaluate the forecast departure bias and STD. Wind observations from SYNOP stations are used to evaluate local near-surface wind forecasts. Radiosonde profiles and analyses are used to diagnose both the mass (humidity, temperature, and geopotential height) and wind fields at various vertical levels. Throughout this study, only scores for wind speed and geopotential height are presented because of the following: 1) bias scores for temperature and humidity are neutral, 2) the STD scores for temperature and humidity are similar to those for geopotential height, and 3) the bias and STD scores for wind components are similar to those for wind speed. While the near-surface wind evaluation is performed globally, upper-air diagnostics are only shown over Europe and neighboring countries where the SYNOP station density

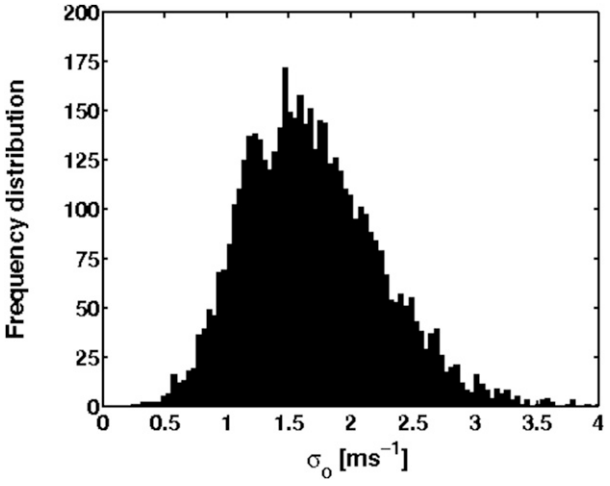

FIG. 2. Frequency distribution of the near-surface wind observation error STD using the GMOS operator for the 4942 SYNOP stations considered in this study.

and their impact are the highest (see Fig. 1). Figure 3 shows the 1487 SYNOP and 124 radiosonde stations located in this area, which covers latitudes ranging from 30° to 70°N and longitudes ranging from 10°W to 50°E.

3. Simplified assimilation experiments

The temporal propagation of the analysis increment and the multivariate impact of the analysis on the forecasts are evaluated through data assimilation experiments performed in a simplified and controlled environment. Only wind data from surface stations over land are assimilated. These noncycling experiments use the same 6-h background fields provided by a full assimilation cycle that did not assimilate the near-surface wind data over land (for an objective comparison). The different experiments are described in Table 1.

Two observation operators (Bilin and GMOS) were used to assimilate the near-surface wind observations. The experiments are evaluated against the control run (CNTRL) for which the forecasts are initiated from the 6-h background fields (i.e., no observations are assimilated). The experiments assimilated only 10-m wind speed and direction observations (after converting to zonal and meridional wind components) every 3 h from 4942 SYNOP stations globally distributed over land. The hybrid formulation of 4DEnVar was used first. Then, to test the impact of multivariate covariances from the different background error components, other experiments were performed using either the NMC or the EnKF background error statistics. Each experiment was conducted over a 1-month period (February 2011). The analyses and subsequent 48-h forecasts were produced twice daily at 0000 and 1200 UTC. The resulting forecasts were verified against the surface observations themselves and radiosonde profiles. Also, ECCC's digital filter was turned off to avoid filtering the analysis increments.

a. Evaluation against near-surface wind observations

GMOS is typically used to postprocess near-surface wind forecasts (e.g., in wind power applications). To

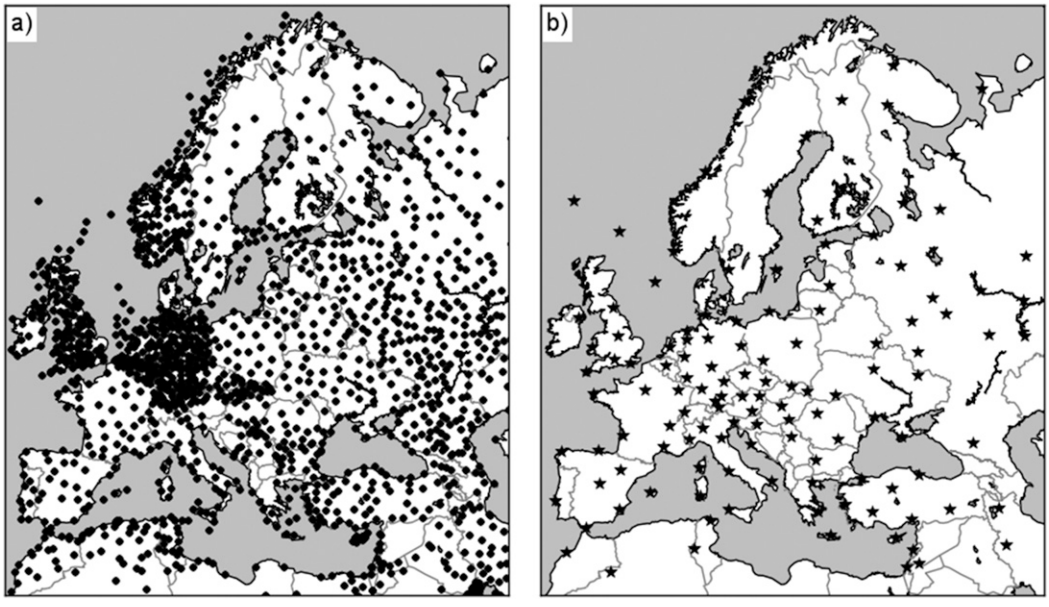

FIG. 3. (a) Spatial distribution of the 1487 SYNOP sites (black dots) located in the selected domain. This area is densely observed and (b) also includes 124 radiosonde stations used for upper-air evaluation purposes (black stars).

TABLE 1. Configuration of the seven simplified data assimilation performed. Each experiment is listed along with its own combination of near-surface wind observation operator, background, and observation error statistics prescribed to the data assimilation system as well as assimilated observations. It is also specified if the experiments are cycled or not.

Expt name	Cycling	Surface wind observation operator (over land)	Surface wind observation errors (over land)	Background error covariances	Observations assimilated
CNTRL	No	—	—	—	None
$\text{Bilin}_{\text{Hybrid}}$	No	Bilin	Homogeneous	Hybrid	SYNOP: Winds
$\text{Bilin}_{\text{EnKF}}$	No	Bilin	Homogeneous	EnKF	SYNOP: Winds
$\text{Bilin}_{\text{NMC}}$	No	Bilin	Homogeneous	NMC	SYNOP: Winds
$\text{GMOS}_{\text{Hybrid}}$	No	GMOS	Site dependent	Hybrid	SYNOP: Winds
$\text{GMOS}_{\text{EnKF}}$	No	GMOS	Site dependent	EnKF	SYNOP: Winds
GMOS_{NMC}	No	GMOS	Site dependent	NMC	SYNOP: Winds

show the improvement in the postprocessing brought by GMOS, Fig. 4 presents the evaluation results against near-surface wind observations for the CNTRL experiment, postprocessed using either Bilin or GMOS (square and circle symbols, respectively). Figure 4 also presents the forecast results for the $\text{Bilin}_{\text{Hybrid}}$ and $\text{GMOS}_{\text{Hybrid}}$ experiments (solid and dashed lines, respectively).

This comparison shows that using GMOS for postprocessing produces significantly lower wind speed departure bias and STD compared to Bilin. The use of GMOS results in a better agreement between model states and near-surface wind observations from the same stations as those used in the assimilation. For this reason and for consistency with the forecast postprocessing tools, GMOS is used to evaluate the surface wind forecasts against SYNOP observations.

Figure 4a also shows that the $\text{Bilin}_{\text{Hybrid}}$ experiment degrades wind speed biases up to 30-h lead time. This degradation is prevented when using GMOS and the bias scores are comparable to those from the CNTRL experiment. The bias score alone is not sufficient to evaluate the impact of the assimilation. However, the small analysis and forecast bias score differences between CNTRL and $\text{GMOS}_{\text{Hybrid}}$ (postprocessed using GMOS) indicate that GMOS performs an efficient bias correction. As the assimilation algorithm assumes that the background and observations are not biased, directly assimilating biased observations (e.g., using Bilin) have a detrimental impact on the analysis and forecasts: the resulting analysis is biased toward the observations and the forecast model drifts back to its own bias (Dee 2005). The relatively constant small bias depicted by the dashed line suggests that observations, forecasts, and analyses are more consistent when using GMOS for the assimilation and forecast postprocessing.

Figure 4b shows that the assimilation of near-surface wind observations is beneficial for the forecast error STD, but only during the first 6 h. This is consistent with the results of Ancell (2012) and Ancell et al. (2011, 2015). While the Bilin experiment reduces the near-surface wind speed STD by $0.07\,\text{m}\,\text{s}^{-1}$ (at 0 h) over the CNTRL experiment, the reduction with GMOS is $0.13\,\text{m}\,\text{s}^{-1}$.

To better understand why the forecast skill from the assimilation of near-surface wind observation vanishes quickly, the mean evolution of the forecast differences between CNTRL and the experiments ($\|\delta\mathbf{V}(t)\| = \sqrt{\delta u^2 + \delta v^2}$), is shown in Fig. 5. Results are presented for

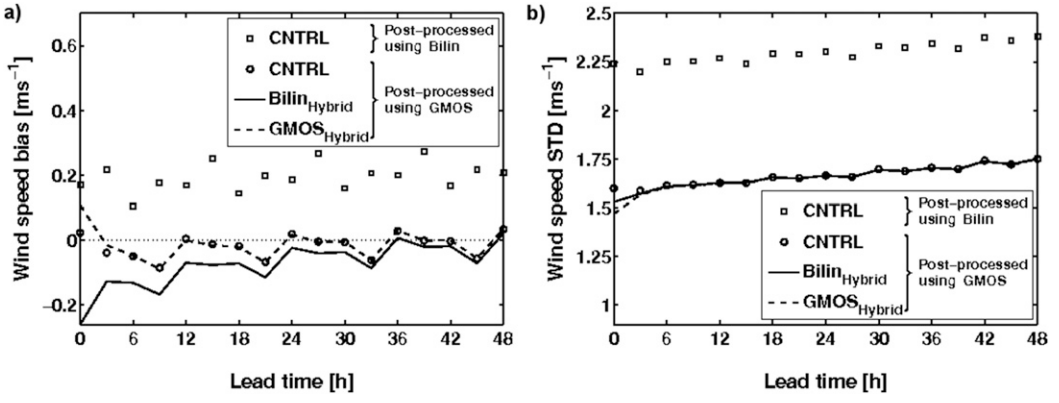

FIG. 4. (a) Wind speed departure bias and (b) STD as a function of forecast lead time for different experiments (CNTRL, Bilin, and GMOS). Note that the CNTRL experiment is postprocessed using both Bilin and GMOS in order to highlight the impact of the operator on postprocessing.

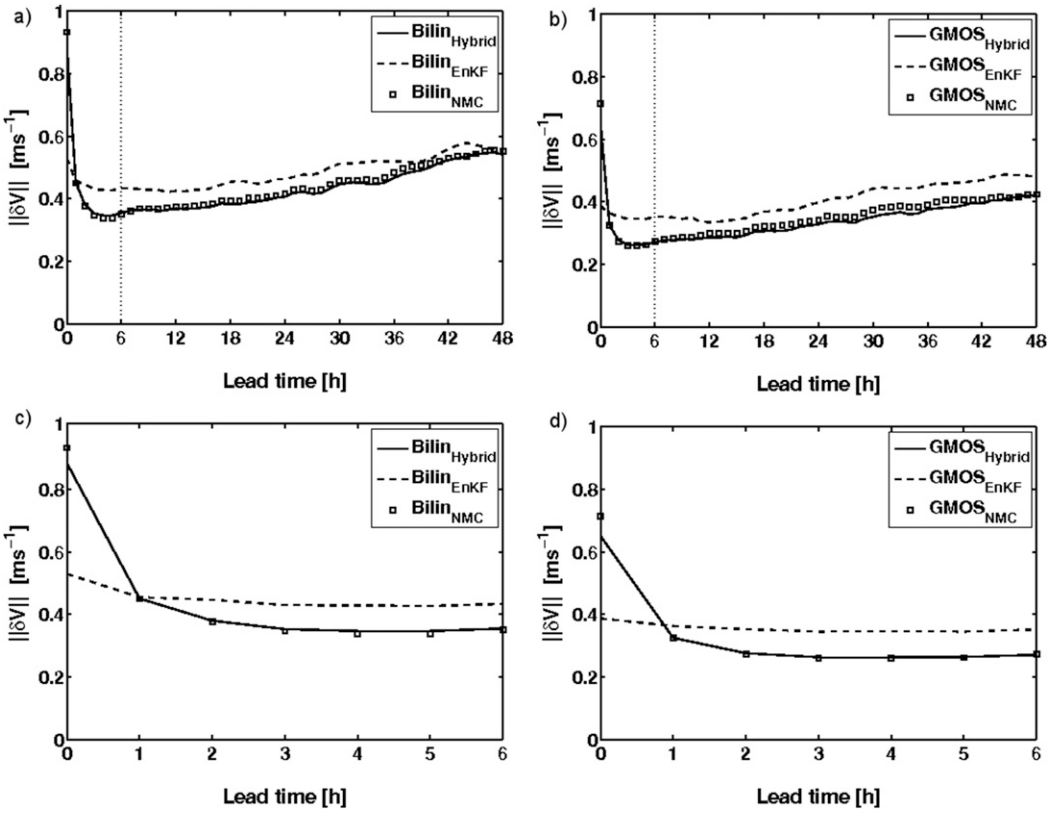

FIG. 5. Forecast differences ($\|\delta\mathbf{V}(t)\| = \sqrt{\delta u^2 + \delta v^2}$) between CNTRL and the experiments [(a),(c) Bilin and (b),(d) GMOS] using different background error covariances (Hybrid, NMC, and EnKF). Results are presented for the run launched at 0000 UTC 1 Feb. The results over the (top) whole 48-h forecast period and (bottom) only the first 6 h.

different observation operators (Bilin and GMOS) and background error statistics (Hybrid, NMC, and EnKF) over the February 2011 period. It represents the evolution of the analysis increment obtained when assimilating only near-surface winds. Figures 5a and 5b show the results for forecasts lead time ranging between 0 and 48 h, while Figs. 5c and 5d emphasize the first 6 h of the integration.

Figure 5 shows that GMOS generates smaller analysis increments than Bilin. This can be observed for the experiments using static, flow-dependent, or hybrid background error covariances. In all cases, a significant amount of the Bilin increment vanishes in the first hours of forecast, while a greater fraction of the GMOS increment persists. In the case of the EnKF, the magnitude of the initial increment is reduced by 14% during the first forecast hour when using Bilin, and only 6% with GMOS.

The analysis increments from the NMC experiments experience an important reduction in the first hours of the integration (so-called forecast convergence). Most of the information from the observations is damped because the analysis increments are substantially reduced. This is indicative of having imbalanced analysis increments that only reach a balanced state after 3–6 h. In contrast in the EnKF experiments, the forecasts only slightly converge in the first hours of integration before diverging due to the inherent nonlinear perturbation growth. This suggests that the analysis increments are better balanced. By construction, the results for the hybrid case are intermediate, but are relatively similar to NMC results because the static error variances are larger than the flow-dependent variances from the EnKF near the surface (Bédard et al. 2015).

To understand the impact of the different background error statistics on the forecasts convergence, the momentum prognostic equation,

$$\frac{\partial \mathbf{V}}{\partial t} = \mathbf{T}_{\text{adv}} + \mathbf{T}_{\text{cori}} + \mathbf{T}_{\text{pg}} + \mathbf{T}_{\text{vd}} + \mathbf{T}_{\text{ob}} + \mathbf{T}_{\text{hd}}, \qquad (3)$$

is studied. The terms on the right-hand side of Eq. (3) are the tendencies from advection ($\mathbf{T}_{\text{adv}}$), Coriolis effect ($\mathbf{T}_{\text{cori}}$), pressure gradients ($\mathbf{T}_{\text{pg}}$), vertical diffusion ($\mathbf{T}_{\text{vd}}$),

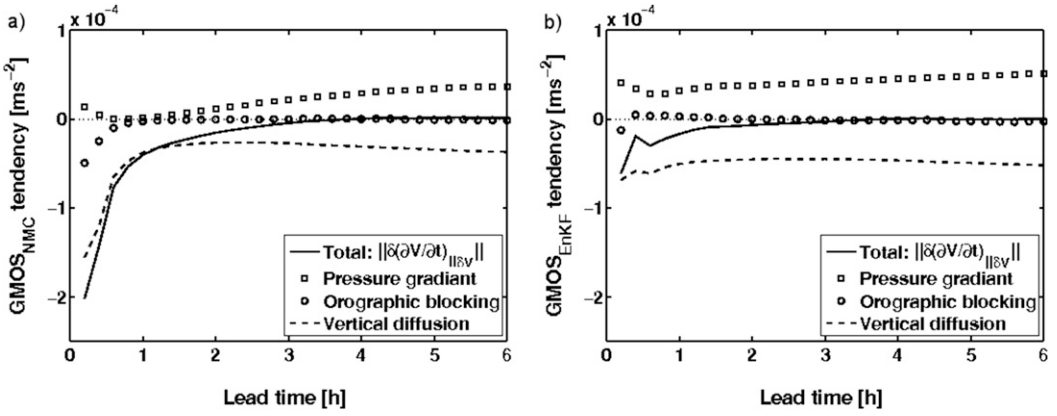

FIG. 6. Contribution of the main terms of Eq. (3) on evolution of the forecast difference between CNTRL and the experiments: (a) $\mathrm{GMOS_{NMC}}$ and (b) $\mathrm{GMOS_{EnKF}}$. Results for advection, Coriolis effect, and horizontal diffusion are omitted as their influence is small compared to the pressure gradient, the vertical diffusion, and orographic blocking.

orographic blocking ($\mathbf{T}_{\mathrm{ob}}$), and horizontal diffusion ($\mathbf{T}_{\mathrm{hd}}$) terms. Similar to Rodwell and Palmer (2007), the systematic tendencies in the first time steps of the integration are used to measure the degree of balance of the analysis increments and to assess the reduction of the forecast differences observed in Fig. 5. Considering that the analysis increment ($\delta\mathbf{V}$) is small, its evolution can be related to the tendency differences between two forecasts, one that uses the background ($\mathbf{V}_b$) as initial conditions (CNTRL) and the other using the analysis ($\mathbf{V}_a$). Given that $\delta\mathbf{V} = \mathbf{V}_a - \mathbf{V}_b$ at $t = 0$, then

$$\frac{\partial(\delta\mathbf{V})}{\partial t} \approx \frac{\partial\mathbf{V}_a}{\partial t} - \frac{\partial\mathbf{V}_b}{\partial t} = \delta\left(\frac{\partial\mathbf{V}}{\partial t}\right). \tag{4}$$

Combining Eqs. (3) and (4) allows attributing the influence of individual physical processes to the evolution of the analysis increment. This can explain the convergence of the forecasts during the first hours of model integration.

Each component of Eq. (3) is computed at every station and every model time step for the first 6-h forecasts of each experiment. The scalar projection of the differences of these tendencies onto $\delta\mathbf{V}$,

$$\left\| \delta\left(\frac{\partial\mathbf{V}}{\partial t}\right)_{\|\delta\mathbf{V}} \right\| = \frac{\delta\left(\frac{\partial\mathbf{V}}{\partial t}\right) \cdot \delta\mathbf{V}}{\|\delta\mathbf{V}\|}, \tag{5}$$

where the operator $\cdot$ denotes a dot product, is calculated and averaged over the 4942 stations. Equation (5) is used to show the parts of $\delta(\partial\mathbf{V}/\partial t)$ that project onto the analysis increment. Figure 6 shows the monthly mean influence of each component of the tendencies on $\delta\mathbf{V}$ [as computed in Eq. (5)] for the $\mathrm{GMOS_{NMC}}$ and $\mathrm{GMOS_{EnKF}}$ experiments. The advection, Coriolis effect, and horizontal diffusion terms are negligible compared to the other terms. Only the total tendency as well as pressure gradient, vertical diffusion, and orographic blocking tendencies are presented in Fig. 6.

Figure 6 shows that the $\mathrm{GMOS_{NMC}}$ case leads to substantially more vertical diffusion during the first model time steps than the $\mathrm{GMOS_{EnKF}}$ case. The vertical structure of the wind correlation with near-surface wind is similar on average for both cases (not shown). However, Bédard et al. (2015) showed that the EnKF wind error variances become small near the surface whereas the NMC near-surface wind error variances remain high. Thus, the $\mathrm{GMOS_{NMC}}$ case produces stronger wind increments near the surface. The formulation of the vertical diffusion scheme (Mailhot and Benoit 1982; Benoit et al. 1989) is sensitive to changes in the nominal wind speed at the model boundary (e.g., lowest prognostic level) as it is directly linked to the wind shear. Also, the orographic blocking scheme [a typical form drag formulated following Lott and Miller (1997); Zadra et al. (2003)] is locally more active in the $\mathrm{GMOS_{NMC}}$ experiment (Fig. 6). This tendency is sensitive to near-surface wind increments as it is proportional to the square of the wind speed. Consequently, the relatively large near-surface wind increments from the $\mathrm{GMOS_{NMC}}$ experiment generates more vertical diffusion and orographic drag than the $\mathrm{GMOS_{EnKF}}$ experiment.

Figure 6a also shows that the near-surface wind observations have limited influence on the pressure gradient force (square symbols) when using the static background error covariances. The NMC background error covariances include the coupling between wind and mass fields near the surface (i.e., Ekman spiral).

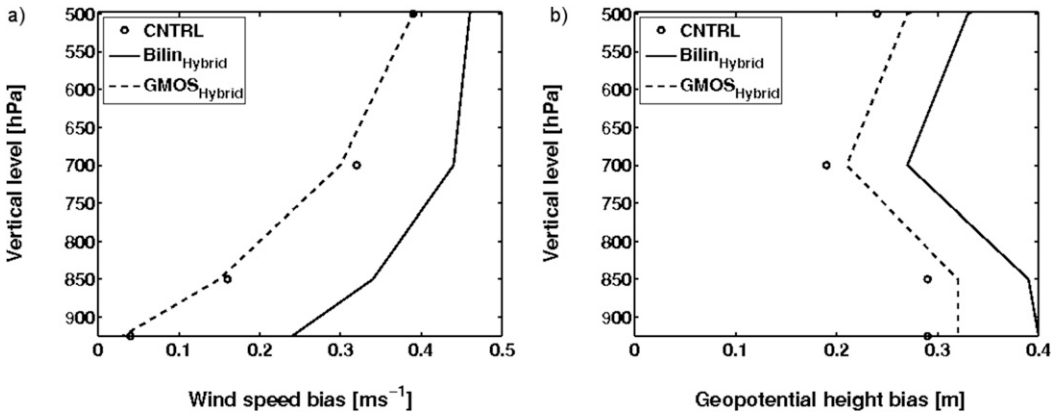

FIG. 7. The 24-h forecast departure bias from hybrid runs as evaluated against radiosonde observations over Europe for (a) wind speed and (b) geopotential height.

However, single-observation experiments suggest that the large-scale homogeneous and isotropic mass increments from nearby stations are not always consistent with each other due to the inhomogeneous and non-isotropic nature of the near-surface flow characteristics (not shown). In such cases, the interaction between nearby mass increments is likely to reduce the effective surface pressure increments, as observed in Fig. 6a. In this case, the vertical diffusion dominates during the first hour of model integration. The large vertical diffusion and orographic blocking tendencies damp the unbalanced analysis increments (dashed line and circle symbol, respectively) to reach a balance with the pressure gradient tendency. It explains why the forecast difference between CNTRL and the GMOS$_{NMC}$ experiment decreases sharply over time in Fig. 5. On the other hand, the EnKF error statistics are flow dependent. It is more likely to provide coherent surface wind and pressure cross correlations for nearby stations. Figure 6b shows that GMOS$_{EnKF}$ experiment produces an effective pressure gradient force (square symbol) that balances the small vertical diffusion (dashed line) and orographic blocking (circle symbol) tendencies right from the start and the increments persist over time [$||\delta(\partial \mathbf{V}/\partial t)\text{Proj}|| \approx 0$]. Consequently, most of the information extracted from the observations is propagated within the forecast when using the EnKF background error covariances, as shown in Fig. 5 (dashed lines).

Overall, the results show that large near-surface background error variances can generate a significant amount of vertical diffusion and orographic drag. Unless counterbalanced by proper pressure gradient forces, the atmospheric boundary layer parameterization schemes (i.e., vertical diffusion and orographic blocking) cause the local wind increments to be quickly damped. The use of flow-dependent background error statistics providing coherent multivariate correlations and variances is thus crucial to produce sustainable analysis increments in the lower troposphere. The general benefit of hybrid data assimilation (Buehner 2005) is not questioned here. However, it appears that the homogeneous and isotropic component of the background error statistics is particularly ill-adapted for the assimilation of near-surface wind observations over land, which limits their impact on forecasts.

b. *Evaluation against upper-air observations*

Independent radiosonde profiles from 124 stations (see Fig. 3b) are used to diagnose the forecast errors of both wind speed and mass fields at different levels. As mentioned earlier, the upper-air diagnostics are performed over Europe because this area possesses the highest SYNOP station density. Figures 7 and 8 present vertical profiles of observation minus forecast departure bias and STD for the February 2011 period. Figure 7 only presents the 24-h lead time biases because results from the 12-, 36-, and 48-h lead times are similar. On the other hand, Fig. 8 displays the forecast departure STD at all lead times to fully appreciate the temporal evolution of the information brought in by near-surface wind observations. Since the focus of this study is the lower troposphere, only results from the surface up to 500 hPa are presented.

The assimilation of near-surface wind using Bilin significantly degrades wind speed and geopotential height biases (Fig. 7), as well as the geopotential height STD scores at all vertical levels (Fig. 8b). Differences in Fig. 8a are not statistically significant, except at 48 h where the GMOS wind speed STD score outperforms those from the Bilin and CNTRL experiments. Bilin$_{Hybrid}$ negative impact on the geopotential height

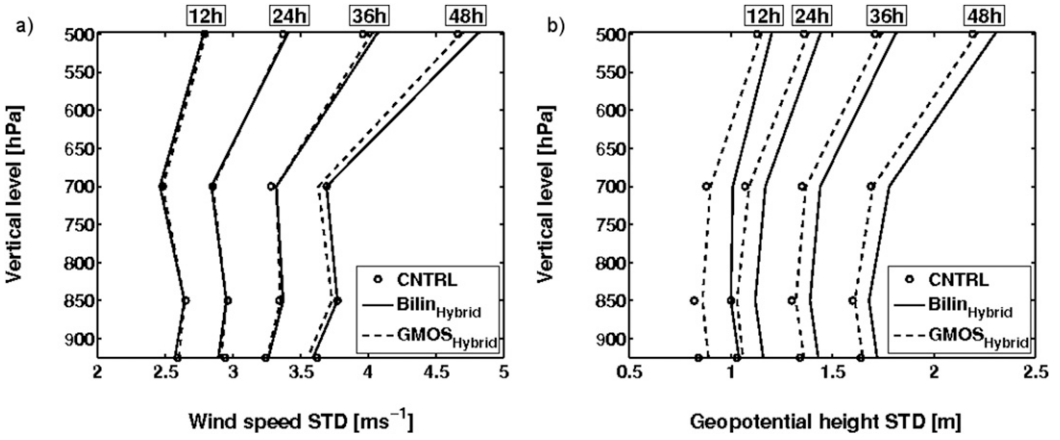

FIG. 8. Forecast departure STD from hybrid runs as evaluated against radiosonde observations over Europe for (a) wind speed and (b) geopotential height. (from left to right) Results are shown from 12-, 24-, 36-, and 48-h forecasts, respectively.

field is already present in the analyses. It is likely related to the introduction of wind observations with significant biases and representativeness errors that can degrade the surface pressure fields through the multivariate background error statistics. These results are consistent with previous studies (e.g., Benjamin et al. 2007, 2010; Ingleby 2015) and it explains why most operational centers do not assimilate near-surface wind observations over land.

Figures 7 and 8 also show that GMOS$_{Hybrid}$ results are generally neutral when compared to the CNTRL experiment. It suggests that using GMOS to account for representativeness errors and biases allows assimilating near-surface wind observations without negatively impacting the upper-air wind biases and the surface pressure field (and thus geopotential height bias and STD) through the multivariate background error covariances. Consequently, the degradation previously observed (when using Bilin) can be avoided with the use of the GMOS operator.

4. Experiments with the operational system

OSEs were performed to evaluate the cumulative impact of assimilating SYNOP wind data with the two different observation operators (Bilin and GMOS) in an operational context. The experiments use ECCC's operational GDPS (including 4DEnVar). The main characteristics of these experiments are presented in Table 2.

In these experiments, all observations assimilated in the GDPS were used, which include those from radiosondes, aircraft, wind profilers, land stations, ships, buoys, scatterometers, atmospheric motion vectors, satellite-based radio occultation, microwave, and infrared satellite sounders/imagers. Wind observations from the 4942 SYNOP stations over land are also assimilated in GMOS$_{OSE}$ and Bilin$_{OSE}$ experiments. Although the use of flow-dependent error statistics is beneficial for the assimilation of near-surface wind data, previous results from Buehner (2005) showed that the assimilation system more generally benefits from the hybrid representation of the background error covariances. Therefore, the operational 4DEnVar configuration using hybrid covariances was selected (Buehner et al. 2015). In total, three OSEs were performed. Each experiment was carried out over a 5-week period (24 January–28 February 2011). The experiments are evaluated over the February 2011 period only. The 48-h forecasts were produced twice a day (at 0000 and 1200 UTC) and were verified against surface stations, radiosonde profiles, and analyses from the same forecast system (i.e., the so-called own analyses).

a. *Evaluation against near-surface wind observations*

Near-surface wind observations from the same 4942 SYNOP stations as those used in the assimilation were used to diagnose the observation impact on local short-term wind forecasts. Figure 9 presents the wind speed forecast departure bias and STD as a function of forecast lead time for the February 2011 period after being postprocessed with GMOS. Figure 9a clearly shows that Bilin has a detrimental effect on biases that persist up to 12-h lead time. This degradation is prevented when using GMOS such that GMOS$_{OSE}$ results are comparable to those from the CNTRL$_{\mathbf{OSE}}$ experiment.

Figure 9b shows that CNTRL$_{OSE}$ near-surface wind departure STD rapidly decreases, and then grows smoothly (after a 3-h forecast lead time). As CNTRL$_{OSE}$ does not assimilate near-surface winds over land, information on near-surface winds is mostly inferred from upper-air wind observations and from observations

TABLE 2. Configuration of the three OSEs performed. Each experiment is listed along with its own combination of near-surface wind observation operator, background and observation error statistics prescribed to the data assimilation system, as well as assimilated observations. It is also specified if the experiments are cycled or not.

Expt name	Cycling	Surface wind observation operator (over land)	Surface wind observation errors (over land)	Background error covariances	Observations assimilated
$CNTRL_{OSE}$	Yes	—	—	Hybrid	All
$Bilin_{OSE}$	Yes	Bilin	Homogeneous	Hybrid	All + SYNOP winds
$GMOS_{OSE}$	Yes	GMOS	Site dependent	Hybrid	All + SYNOP winds

describing the mass fields (e.g., temperature, humidity, surface pressure) through the background error correlations at ~50-km resolution. In such case, the analysis cannot depict finescale local effects as observed by near-surface wind observations and it is up to the model to develop them during the first hours of forecast, which explains the error reduction between 0 and 3 h. This error reduction is avoided when assimilating near-surface wind observations (full and dashed lines).

Figure 9b also shows that the assimilation of near-surface wind observations is beneficial for very short-term local wind predictions. However, this positive impact decreases in time and is only significant up to 6 h because the vertical diffusion and orographic blocking schemes damp the surface wind increments during the first hours of the model integration. Still, the $GMOS_{OSE}$ and $Bilin_{OSE}$ experiments, respectively, improve the fit of near-surface wind analyses to the observations (STD) by 0.16 and 0.10 m s^{-1} over the $CNTRL_{OSE}$ experiment. The fact that GMOS was used in the verification of the forecasts of the BILIN experiment explains the large difference at 0 h between the GMOS and BILIN experiments. However, there is little difference between the STD scores between the two experiments beyond 3 h.

As expected, such improvements are small compared to the improvement obtained from using GMOS as a postprocessing module (Fig. 4) because the background states and analyses carry information from all assimilated observations. Still, in the OSE context, the fact that near-surface wind observations can further improve low-level wind analyses and very short-term forecasts (in the presence of all operationally assimilated observations) can be helpful for applications that rely on very short-range forecasts.

b. *Upper-air evaluation*

The forecast of both wind and mass fields is evaluated at different levels over Europe and adjacent countries. The 124 radiosonde stations presented in Fig. 3 are used to compute the upper-air scores against observations. Analyses from the experiments (i.e., the so-called own analyses) are also used to assess the forecast. Unlike the forecast verifications against observations, the score against own analyses cannot be used to diagnose short-range forecast skills because the forecast errors have a strong imprint of the analysis error. Nevertheless, it is useful to assess how consistent they are with each other. Balance and higher consistency in itself is desirable, but does not indicate that the forecasts are more accurate. As pointed out by Rodwell and Palmer (2007), imbalanced initial conditions should experience a rapid spin-down associated with relaxation, through diffusion, to

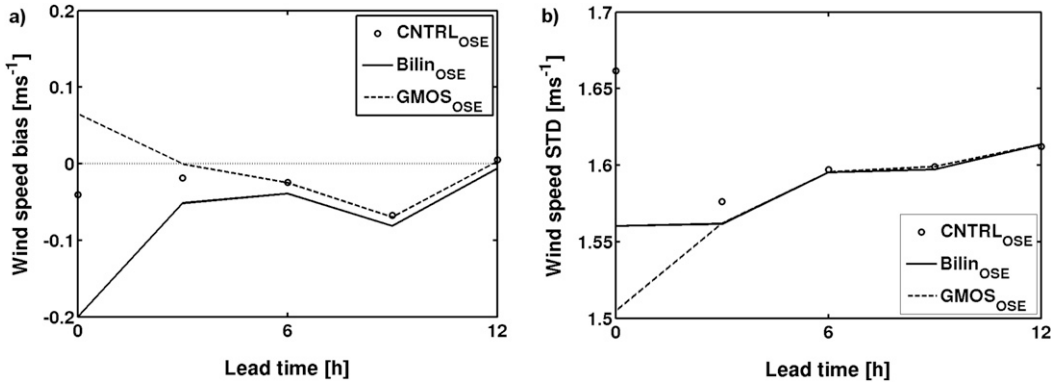

FIG. 9. Wind speed (a) departure bias and (b) STD as a function of forecast lead time for different experiments ($CNTRL_{OSE}$, $Bilin_{OSE}$, and $GMOS_{OSE}$). Note that the GMOS operator is also used for postprocessing in all experiments.

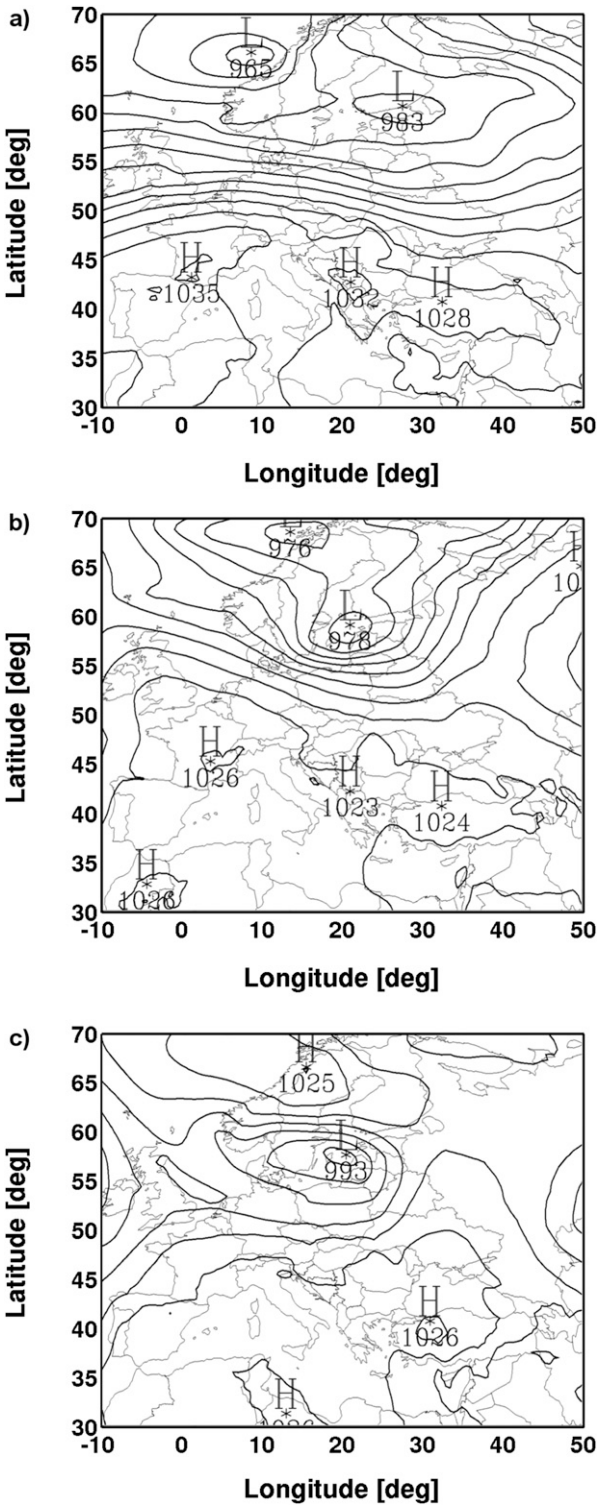

FIG. 10. Sea level pressure (hPa) over Europe valid at 0600 UTC (a) 5 Feb, (b) 8 Feb, and (c) 11 Feb.

the model's own balance. Well-balanced cases, where the analysis increment is propagated in time, should provide forecasts that are more consistent with analyses, as the information persists in the system, than unbalanced cases, where the increments can be diffused by, for example, the ABL parameterizations.

The weather over the area of interest (Fig. 3) in February 2011 is examined to interpret the verification scores. A careful inspection of the meteorological conditions indicates that three low pressure systems moved across the evaluation area during the first two weeks of February. Figure 10 shows the sea level pressure for those three cases (at 0600 UTC 5, 8, and 11 February). More specifically, in each case the depression developed over the Norwegian Sea and moved over Scandinavia and the Baltic Sea before hitting Russia. Then a large anticyclone formed over Russia and northern Europe and remained quasi stationary during the second half of February.

The scores against upper-air observations from both experiments ($GMOS_{OSE}$ and $Bilin_{OSE}$) indicate that near-surface wind observations have a neutral impact on short-term tropospheric forecasts (not shown). As opposed to the verification scores against radiosondes, the verification scores against analyses were calculated for the whole area of interest without putting emphasis on more densely observed regions. They provide a complementary perspective that gives a more unified view of the impact to capture the influence of near-surface wind observations on forecasts over regions not well observed. The forecast evaluation against own analyses shows that the $GMOS_{OSE}$ forecasts and analyses are slightly more self-consistent than those from the $CNTRL_{OSE}$ experiment and $Bilin_{OSE}$ produces forecasts and analyses that are less self-consistent (not shown). Although the differences are not statistically significant, results are coherent with results shown in Fig. 5 where a significant part of the increment vanishes in the first hours of the forecasts from the Bilin experiment, while a greater fraction of the increment persists in the forecasts from the GMOS experiment.

The spatial and temporal distribution of the verification scores against analyses is studied to assess the difference between the experiments ($GMOS_{OSE}$ and $Bilin_{OSE}$) and $CNTRL_{OSE}$. The geographical distribution of the score differences is shown in Fig. 11. Results for 12-h forecasts of 10-m wind speed are presented for the February 2011 period. It shows that $GMOS_{OSE}$ and $Bilin_{OSE}$ scores are neutral over southwestern Europe (light gray shades) because this region is already densely observed by radiosonde stations (see Fig. 3) and characterized by low synoptic activity during the evaluation period. This is consistent with the neutral upper-air

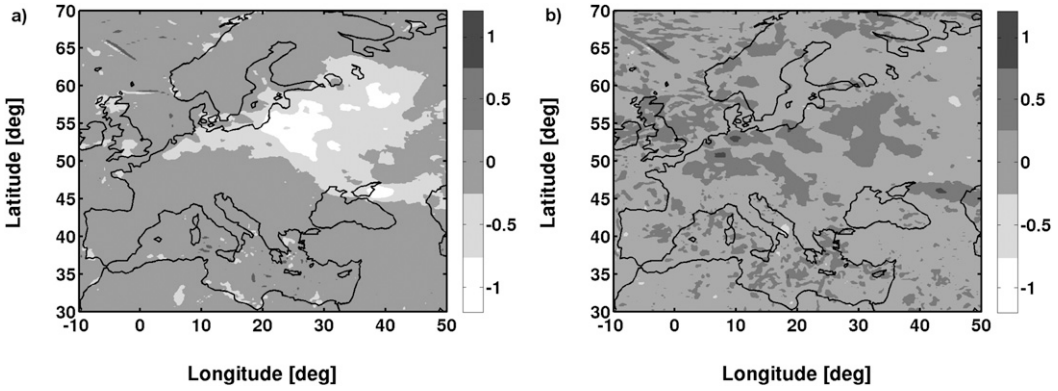

FIG. 11. Mean 12-h forecast departure STD (against own analyses) differences (control minus experiment) over Europe. Results for 10-m wind speed ($m\,s^{-1}$) are averaged for the month of February 2011. Results for the (a) $Bilin_{OSE}$ and (b) $GMOS_{OSE}$ experiments. Positive (negative) values are represented by dark (light) colors. A positive value (dark gray) indicates that the experiment is better than $CNTRL_{OSE}$, while the light gray color indicates neutral results.

evaluation in this most densely observed region (e.g., western Europe). Figure 11a indicates that $Bilin_{OSE}$ forecasts and analyses are less self-consistent (-0.5 to $-1\,m\,s^{-1}$) than those from $CNTRL_{OSE}$ for the area impacted by synoptic activity during the evaluation period (white shades). In this case, the use of Bilin is detrimental to the assimilation system, probably because of biases and representativeness errors from coastal sites, as suggested by Ingleby (2015) and Bédard et al. (2015). On the other hand, Fig. 11b shows that $GMOS_{OSE}$ forecasts and analyses are generally more self-consistent ($0.5\,m\,s^{-1}$) than those from $CNTRL_{OSE}$ (dark gray shades). Figure 3 shows that there are few radiosondes in this area and one has to rely on verification against analyses, which shows a positive impact from near-surface wind observations over less densely observed areas for the $GMOS_{OSE}$ case, as suggested in Dong et al. (2011). There are many SYNOP stations in the area that can provide useful information to the assimilation system when accounting for biases and representativeness errors with the GMOS operator.

Figure 12 presents a Hovmöller diagram to assess the wind speed forecasts and analyses self-consistency as a function of longitude and time (averaged over latitudes 30°–70°N). It presents the differences between the experiments ($Bilin_{OSE}$ or $GMOS_{OSE}$) and $CNTRL_{OSE}$ 12-h forecast departure STD against own analyses. The Hovmöller diagram shows that results are generally neutral except during the first two weeks of February between longitudes 20° and 40°E where three light-colored bands oriented from upper left to lower right depict negative impacts ($-0.5\,m\,s^{-1}$) moving from west to east for $Bilin_{OSE}$ on 5, 8, and 11 February (Fig. 12a). For the $GMOS_{OSE}$ case (Fig. 12b), a dark band oriented from upper left to lower right shows positive impacts ($0.5\,m\,s^{-1}$) between longitude 30° and 40°E on 8 February. For both experiments, the SYNOP wind observations impact coincides with the depressions described previously.

These results suggest that the analysis increments are propagated by the NWP model and allow near-surface wind observations over land to have an impact on forecasts downstream of the three low pressure systems considered. For those specific cases, $GMOS_{OSE}$ produces forecasts that verify better against analyses from the same NWP system (produced using SYNOP wind observations along with all observations assimilated operationally).

5. Summary and conclusions

The long-term goal of this study is to evaluate the improvements to short-term tropospheric forecasts and wind nowcasts from assimilating near-surface wind observations over land using an improved observation operator (GMOS) in ECCC's 4DEnVar. Specifically, the multivariate impact of near-surface wind observations on analyses and forecasts, the spatiotemporal propagation of the information in the NWP system and the influence of different components of the assimilation and prediction systems were examined.

First, simplified experiments were performed in which only near-surface winds were assimilated using the same background states and using either the Bilin or GMOS observation operator. The results show that Bilin significantly degrades surface and upper-air fields. GMOS, on the other hand, produces smaller increments that are in better agreement with the model state confirming that

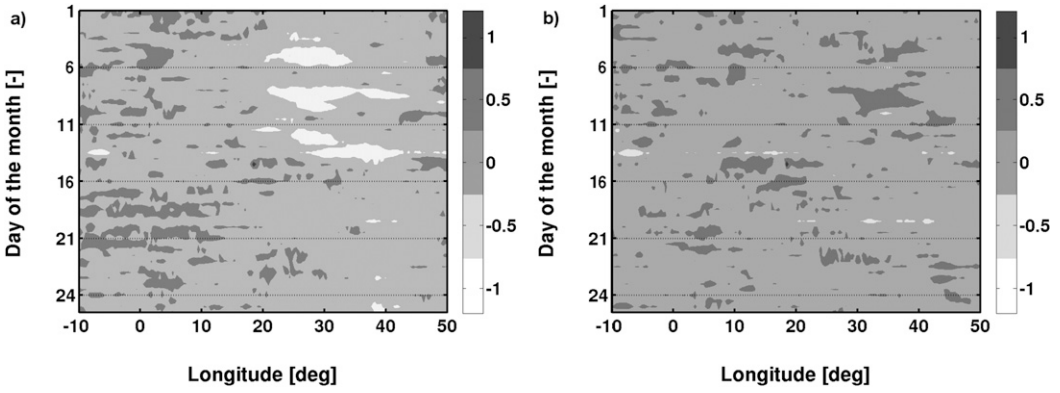

FIG. 12. Hovmöller diagram presenting the 12-h forecast departure STD (against own analyses) differences (control minus experiment) over Europe. Results for 10-m wind speed ($m\,s^{-1}$) are presented through February 2011 for different longitude bands over Europe. Results for the (a) $Bilin_{OSE}$ and (b) $GMOS_{OSE}$ experiments. Positive (negative) values are represented by dark (light) colors. A positive value (dark gray) indicates that the experiment is better than $CNTRL_{OSE}$, while the light gray color indicates neutral results.

GMOS is a more suitable observation operator. It leads also to better short-term near-surface wind forecasts than Bilin and does not deteriorate the upper-air forecasts. The information persists longer in the system with GMOS, although the local improvements do not propagate beyond 6-h lead time. Initial model tendencies reveal that the spatiotemporal propagation of the information is limited by the quality of the background error covariances. The vertical structure of the stationary error statistics produces large near-surface wind increments that cause the forecast model to generate significant amounts of vertical diffusion and orographic drag. Its homogenous and isotropic characteristics are also ill adapted and the interaction between nearby observing stations reduces the overall surface pressure increments. The resulting increments are therefore unbalanced and they are quickly damped by the large vertical diffusion and orographic blocking tendencies. On the other hand, the flow-dependent error statistics from the EnKF have a vertical structure that can generate small wind increments near the surface. This avoids shocking the NWP system and producing strong transients dissipated through vertical diffusion and orographic drag. The flow-dependent background error statistics of the EnKF modify both wind and mass fields in a more coherent way, generating pressure gradient forces that balance the diffusive forces from the ABL parameterization schemes. The system is initially better balanced and allows the increment to propagate further in time within the forecast. The use of flow-dependent background error statistics is critical to produce sustainable impacts on the atmosphere by means of coherent correlations between wind and mass fields as well as appropriate variance structure in the vertical.

OSEs in which wind data from SYNOP stations over the globe are assimilated, along with all other observations, were carried out to assess the impact of near-surface wind observations in ECCC's operational context. The evaluation against upper-air data suggests that SYNOP wind observations have a neutral impact on short-term tropospheric forecasts. Conversely, the evaluation against own analyses shows that the forecasts and analyses issued from the $GMOS_{OSE}$ experiment are more self-consistent with than those from the $CNTRL_{OSE}$ and $Bilin_{OSE}$ experiments for 12-h forecast lead times. This suggests that GMOS allows for a better use of near-surface wind observations and the analysis increments are propagated in time.

Near-surface wind observations have a limited influence aloft because the flow is decoupled at the top of the ABL (Bédard et al. 2015). Also, a large quantity of observations is already assimilated in the current system (12 million observations per day). These may be factors limiting the observation impact on the full 3D NWP model, and improvements on short-term tropospheric forecasts are modest especially over already well-observed areas. However, because of their relatively low cost and their use for monitoring, safety, and climatological needs, near-surface wind observations are abundant (~30 000 stations over the globe). The results of this study suggest that the use of GMOS renews the possibility of providing useful information on low-level flow from the assimilation of near-surface wind observations. Operational centers could start assimilating near-surface wind observations over land using GMOS to take advantage of the short-term wind forecasts improvements without degrading surface and upper-air fields. Such features can be useful in a NWP system using

short data assimilation windows. This is the scope of a future study where ~20 000 stations will be used in the Rapid Refresh (RAP) system currently employed at the National Centers for Environmental Prediction (NCEP), and thereby result in larger forecast improvements.

With the objective of making a better use of near-surface wind observations over land and improving their impact on short-term tropospheric forecasts, future work must focus on the improvement of background error statistics. With the improvement of flow-dependent background error covariances and the reduction of the static error component in the hybrid schemes, the impact of near-surface wind observations may propagate further in time and space.

Acknowledgments. The authors acknowledge the contributions of Jean Côté (UQAM's ESCER Centre and ECCC), Claude Girard (ECCC), Paul Vaillancourt (ECCC), and Ayrton Zadra (ECCC) for their suggestions on computing the momentum equation components. The authors wish to thank Mark Buehner (ECCC), Joshua Hacker (NCAR), and three anonymous reviewers for their suggestions and comments. The first author also thanks the Natural Sciences and Engineering Research Council of Canada (NSERC), Hydro-Québec research institute (IREQ), and ECCC Atmospheric Science and Technology Directorate for their financial support.

REFERENCES

Ancell, B. C., 2012: Examination of analysis and forecast errors of high-resolution assimilation, bias removal, and digital filter initialization with an ensemble Kalman filter. *Mon. Wea. Rev.*, **140**, 3992–4004, doi:10.1175/MWR-D-11-00319.1.

——, C. F. Mass, and G. J. Hakim, 2011: Evaluation of surface analyses and forecasts with a multiscale ensemble Kalman filter in regions of complex terrain. *Mon. Wea. Rev.*, **139**, 2008–2024, doi:10.1175/2010MWR3612.1.

——, E. Kashawlic, and J. Schroeder, 2015: Evaluation of wind forecasts and observation impacts from variational and ensemble data assimilation for wind energy applications. *Mon. Wea. Rev.*, **143**, 3230–3245, doi:10.1175/MWR-D-15-0001.1.

Bédard, J., W. Yu, Y. Gagnon, and C. Masson, 2013: Development of a geophysic model output statistic module for improving short-term numerical wind predictions over complex sites. *Wind Energy*, **16**, 1131–1147.

——, S. Laroche, and P. Gauthier, 2015: A geo-statistical observation operator for the assimilation of near-surface wind data. *Quart. J. Roy. Meteor. Soc.*, **141**, 2857–2868, doi:10.1002/qj.2569.

Benjamin, S. G., W. R. Moninger, S. R. Sahm, and T. L. Smith, 2007: Mesonet wind quality monitoring allowing assimilation in the RUC and other NCEP models. *22nd Conf. on Weather Analysis and Forecasting/18th Conf. on Numerical Weather Prediction*, Park City, UT, Amer. Meteor. Soc., P1.33. [Available online at https://ams.confex.com/ams/22WAF18NWP/techprogram/paper_124829.htm.]

——, B. D. Jamison, W. R. Moninger, S. R. Sahm, B. 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, doi:10.1175/2009MWR3097.1.

Benoit, R., J. Coté, and J. Mailhot, 1989: Inclusion of a TKE boundary layer parameterization in the Canadian regional finite-element model. *Mon. Wea. Rev.*, **117**, 1726–1750, doi:10.1175/1520-0493(1989)117<1726:IOATBL>2.0.CO;2.

Buehner, M., 2005: Ensemble-derived stationary and flow-dependent background error covariances: Evaluation in a quasi-operational NWP setting. *Quart. J. Roy. Meteor. Soc.*, **131**, 1013–1043, doi:10.1256/qj.04.15.

——, J. Morneau, and C. Charette, 2013: Four-dimensional ensemble-variational data assimilation for global deterministic weather prediction. *Nonlinear Processes Geophys.*, **20**, 669–682, doi:10.5194/npg-20-669-2013.

——, and Coauthors, 2015: Implementation of deterministic weather forecasting systems based on ensemble-variational data assimilation at Environment Canada. Part I: The global system. *Mon. Wea. Rev.*, **143**, 2532–2559, doi:10.1175/MWR-D-14-00354.1.

Charron, M., and Coauthors, 2012: The stratospheric extension of the Canadian global deterministic medium-range weather forecasting system and its impact on tropospheric forecasts. *Mon. Wea. Rev.*, **140**, 1924–1944, doi:10.1175/MWR-D-11-00097.1.

Côté, J., S. Gravel, A. Méthot, A. Patoine, M. Roch, and A. Staniforth, 1998: The operational CMC/MRB Global Environmental Multiscale (GEM) model. Part I: Design considerations and formulation. *Mon. Wea. Rev.*, **126**, 1373–1395, doi:10.1175/1520-0493(1998)126<1373:TOCMGE>2.0.CO;2.

Dee, D. P., 2005: Bias and data assimilation. *Quart. J. Roy. Meteor. Soc.*, **131**, 3323–3343, doi:10.1256/qj.05.137.

Dirren, S., R. D. Torn, and G. J. Hakim, 2007: A data assimilation case study using a limited-area ensemble Kalman filter. *Mon. Wea. Rev.*, **135**, 1455–1473, doi:10.1175/MWR3358.1.

Dong, J., M. Xue, and K. Droegemeier, 2011: The analysis and impact of simulated high-resolution surface observations in addition to radar data for convective storms with an ensemble Kalman filter. *Meteor. Atmos. Phys.*, **112**, 41–61, doi:10.1007/s00703-011-0130-3.

Gauthier, P., M. Buehner, and L. Fillion, 1998: Background-error statistics modelling in a 3D variational data assimilation scheme: Estimation and impact on the analyses. *Proc. ECMWF Workshop on the Diagnostics of Assimilation Systems*, Reading, United Kingdom, ECMWF, 131–145.

——, M. Tanguay, S. Laroche, S. Pellerin, and J. Morneau, 2007: Extension of 3DVAR to 4DVAR: Implementation of 4DVAR at the Meteorological Service of Canada. *Mon. Wea. Rev.*, **135**, 2339–2354, doi:10.1175/MWR3394.1.

Hacker, J. P., and C. Snyder, 2005: Ensemble Kalman filter assimilation of fixed screen-height observations in a parameterized PBL. *Mon. Wea. Rev.*, **133**, 3260–3275, doi:10.1175/MWR3022.1.

——, and D. Rostkier-Edelstein, 2007: PBL state estimation with surface observations, a column model, and an ensemble filter. *Mon. Wea. Rev.*, **135**, 2958–2972, doi:10.1175/MWR3443.1.

——, J. L. Anderson, and M. Pagowski, 2007: Improved vertical covariance estimates for ensemble-filter assimilation of near-surface observations. *Mon. Wea. Rev.*, **135**, 1021–1036, doi:10.1175/MWR3333.1.

Houtekamer, P. L., X. Deng, H. L. Mitchell, S. J. Baek, and N. Gagnon, 2014: Higher resolution in an operational ensemble Kalman filter. *Mon. Wea. Rev.*, **142**, 1143–1162, doi:10.1175/MWR-D-13-00138.1.

Ingleby, B., 2015: Global assimilation of air temperature, humidity, wind and pressure from surface stations. *Quart. J. Roy. Meteor. Soc.*, **141**, 504–517, doi:10.1002/qj.2372.

Laroche, S., P. Gauthier, J. St-James, and J. Morneau, 1999: Implementation of a 3D variational data assimilation system at the Canadian Meteorological Centre. Part II: The regional analysis. *Atmos.–Ocean*, **37**, 281–307, doi:10.1080/07055900.1999.9649630.

Lott, F., and M. J. Miller, 1997: A new subgrid-scale orographic drag parametrization: Its formulation and testing. *Quart. J. Roy. Meteor. Soc.*, **123**, 101–127, doi:10.1002/qj.49712353704.

Mailhot, J., and R. Benoit, 1982: A finite-element model of the atmospheric boundary layer suitable for use with numerical weather prediction models. *J. Atmos. Sci.*, **39**, 2249–2266, doi:10.1175/1520-0469(1982)039<2249:AFEMOT>2.0.CO;2.

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

Pu, Z., H. Zhang, and J. Anderson, 2013: Ensemble Kalman filter assimilation of near-surface observations over complex terrain: Comparison with 3DVAR for short-range forecasts. *Tellus*, **65A**, 19620, doi:10.3402/tellusa.v65i0.19620.

Rodwell, M. J., and T. N. Palmer, 2007: Using numerical weather prediction to assess climate models. *Quart. J. Roy. Meteor. Soc.*, **133**, 129–146, doi:10.1002/qj.23.

Rostkier-Edelstein, D., and J. P. Hacker, 2010: The roles of surface-observation ensemble assimilation and model complexity for nowcasting of PBL profiles: A factor separation analysis. *Wea. Forecasting*, **25**, 1670–1690, doi:10.1175/2010WAF2222435.1.

Zack, J., E. J. Natenberg, S. Young, G. V. Knowe, K. Waight, J. Manobianco, and C. Kamath, 2010: Application of ensemble sensitivity analysis to observation targeting for short-term wind speed forecasting in the Washington–Oregon region. Lawrence Livermore National Laboratory Tech. Rep. LLNL-TR-458086, 65 pp. [Available online at http://computation.llnl.gov/projects/starsapphire-data-driven-modeling-analysis/LLNL-TR-458086.pdf.]

——, ——, G. V. Knowe, K. Waight, J. Manobianco, D. Hanley, and C. Kamath, 2011: Observing system simulation experiments (OSSEs) for the Mid-Columbia basin. Lawrence Livermore National Laboratory Tech. Rep. LLNL-TR-499162, 17 pp. [Available online at https://e-reports-ext.llnl.gov/pdf/515298.pdf.]

Zadra, A., M. Roch, S. Laroche, and M. Charron, 2003: The subgrid scale orographic blocking parametrization of the GEM model. *Atmos.–Ocean*, **41**, 155–170, doi:10.3137/ao.410204.

——, and Coauthors, 2014: Improvements to the Global Deterministic Prediction System (GDPS) (from version 2.2.2 to 3.0.0), and related changes to the Regional Deterministic Prediction System (RDPS) (from version 3.0.0 to 3.1.0). Tech. Note, Canadian Meteorological Centre, 88 pp. [Available online at http://collaboration.cmc.ec.gc.ca/cmc/CMOI/product_guide/docs/lib/op_systems/doc_opchanges/technote_gdps300_20130213_e.pdf.]