Dynamical Properties of MOS Forecasts: Analysis of the ECMWF Operational Forecasting System

S. VANNITSEM

Institut Royal Météorologique de Belgique, Brussels, Belgium

(Manuscript received 7 February 2008, in final form 13 May 2008)

ABSTRACT

The dynamical properties of ECMWF operational forecasts corrected by a (linear) model output statistics (MOS) technique are investigated, in light of the analysis performed in the context of low-order chaotic systems. Based on the latter work, the respective roles of the initial condition and model errors on the forecasts can be disentangled. For the temperature forecasted by the ECMWF model over Belgium, it is found that (i) the error amplification arising from the presence of uncertainties in the initial conditions dominates the error dynamics of the "free" atmosphere and (ii) the temperature at 2 m can be partly corrected by the use of the (linear) MOS technique (as expected from earlier works), suggesting that model errors and systematic initial condition biases dominate at the surface. In the latter case, the respective amplitudes of the model errors and systematic initial condition biases corrected by MOS depend on the location of the synoptic station. In addition, for a two-observables MOS scheme, the best second predictor is the temperature predicted at 850 hPa in the central part of the country, while for the coastal zone, it is the sensible heat flux entering in the evolution of the surface temperature. These differences are associated with a dominant problem of vertical temperature interpolation in the central and east parts of the country and a difficulty in assessing correctly the surface heat fluxes on the coastal zone. Potential corrections of these problems using higher-resolution models are also discussed.

1. Introduction

Model output statistics (MOS) techniques are used worldwide to improve meteorological operational forecasts. The MOS method consists of correcting the model outputs based on the information gathered from past forecasts (Wilks 2006). The most popular and approach is the use of a linear regression between the predictors coming from the forecasts and the observations (or predictands). This approach has been proven to be very successful for global (Glahn and Lowry 1972; Klein and Glahn 1974; Lemcke and Kruizinga 1988; Dallavalle et al. 2004; Hart et al. 2004; Taylor and Leslie 2005) as well as regional forecasts (Sokol 2003; Termonia and Deckmyn 2007; Cheng and Steenburgh 2007).

Corresponding author address: S. Vannitsem, Institut Royal Météorologique de Belgique, Ave. Circulaire, 3, 1180 Brussels, Belgium.
E-mail: svn@oma.be

DOI: 10.1175/2008WAF2222126.1

Recently, this MOS technique has been investigated from a dynamical point of view (Vannitsem and Nicolis 2008), in order to evaluate its ability in correcting initial condition and/or model errors. The analysis, applied within the context of low-order chaotic systems, has revealed a number of important features: (i) the MOS technique corrects systematic initial condition errors and part of the model errors; (ii) random initial condition errors are poorly corrected; and (iii) for model errors, the MOS correction depends on the mean of the phase-space velocity difference between the model and reality (mean of the tendency difference), and its covariance with the predictors. One consequence of these results is that the correction could be substantial only when model errors and/or systematic initial condition biases, are present. A second consequence is the fact that the model error correction will strongly depend on the choice of predictors, the better choice being the model observables that strongly correlate with the model error source (the tendency difference between the model and reality). These characteristics allow for an evaluation of the presence of model errors (and initial condition biases) in the forecasts.

The purpose of this note is to investigate the correction of the forecasts produced operationally at the European Centre for Medium-Range Weather Forecasts (ECMWF) in light of the results found in Vannitsem and Nicolis (2008), in order to clarify the respective roles of the initial and model errors on a specific operational example. More specifically, the MOS corrections of the temperature forecasts are analyzed at the surface and different pressure levels over Belgium. First, the analysis is performed on the full grid by correcting the forecasts using the past analyses (at different levels and at the surface) and second at several synoptic stations, for which the training is performed based on the synoptic observations (only for the 2-m temperature).

Global forecasts can obviously benefit from the improvement of atmospheric modeling. This should in principle reduce the model error that is otherwise present. One natural question in the present context is to know whether the correction provided by the postprocessing can be compensated for by the use of higher-resolution models (at global or regional scales). This aspect is also briefly investigated through the analysis of forecasts of the recent higher-resolution release of the ECMWF model as well as those produced by a high-resolution regional model used operationally over Belgium, the Aire Limitée Adaptation Dynamique Développement International (ALADIN) model.

In section 2, the MOS technique is briefly revisited in light of the analysis performed in Vannitsem and Nicolis (2008). Section 3 describes the forecasts used in the present work, and the results are presented in section 4. Finally, the main results are summarized in section 5.

2. The MOS technique

The MOS technique aims at correcting current forecasts based on statistical information gathered from past forecasts (see, e.g., Wilks 2006). In its most popular form, it is based on a linear relation between the reference (the observation of the truth) variables that we want to predict, X, and a set of model observables (or predictors), $\{V_i\,(t)\}$, at a certain lead time t:

$$X_C(t) = \alpha(t) + \sum_{i=1}^{n} \beta_i(t) V_i(t), \tag{1}$$

where $X_C(t)$ is the corrected forecast. The parameters $\alpha(t)$ and $\{\beta_i(t)\}$ are estimated using a set of past forecasts by minimizing a cost function:

$$J(t) = \sum_{k=1}^{K} [X_{C,k}(t) - X_k(t)]^2, \tag{2}$$

where $X_k(t)$ is the reference variable at time t at which the kth forecast is compared and K is the number of past predictions used to build the statistical relation (1). The minimization is performed by differentiating (2) against the parameters. This leads to the following relations for the parameters:

$$\alpha(t) = \langle X_k(t)\rangle - \sum_{i=1}^{n} \beta_i(t)\langle V_{i,k}(t)\rangle \quad \text{and} \tag{3}$$

$$\sum_i \beta_i(t)\langle V_{i,k}(t)V_{j,k}(t)\rangle + \alpha(t)\langle V_{j,k}(t)\rangle = \langle X_k(t)V_{j,k}(t)\rangle \quad \text{for} \quad j = 1, \ldots, n, \tag{4}$$

where $\langle\cdot\rangle$ refers to a statistical average over the ensemble of forecasts, K.

Let us now turn to the mean square error (MSE) between the corrected forecast and the reference variable and focus only on MOS with $n = 1$ and $n = 2$, for which the analytical investigations have been performed in Vannitsem and Nicolis (2008). For $n = 1$, one assumes that the predictor of the MOS scheme is the same nominal variable as the forecasted one. The MSE can then be decomposed into three parts:

$$\langle [X_C(t) - X(t)]^2\rangle = \langle [V(t) - X(t)]^2\rangle - [\beta(t) - 1]^2\sigma_V^2(t) - [\langle X(t)\rangle - \langle V(t)\rangle]^2, \tag{5}$$

where the first term on the right-hand side is the MSE between the model solution and the true system variable and the two last terms are two negative corrections reducing the amplitude of the MSE. Obviously, the last term is associated with the correction of the mean of the forecast, which can be referred as a *drift correction* and is hereafter denoted DC, whereas the second term is the *variability correction* (hereafter VC). It can be rewritten as $-[\beta(t) - 1]^2\sigma_V^2(t) = -[\sigma_C(t) - \sigma_V(t)]^2$, where $\sigma_C(t)$ and $\sigma_V(t)$ are the standard deviations of the corrected forecast and the model variable, respectively. Note that if both the model and the true system are chaotic with sufficiently strong ergodic properties, $\sigma_C(t)$ will decrease and consequently the correction to the MSE will be mainly given by the variance of the model variable for long lead times (see Wilks 2006; Vannitsem and Nicolis 2008). A corollary of the decrease of the MOS forecast variance is the convergence of the forecast toward the climatological mean. Note that when only a systematic bias is corrected (corresponding to fix $\beta(t) = 1$ in 1), the VC correction term disappears in (5).

Usually, MOS schemes are not limited to a single model observable. The use of information from the model pertaining to (well chosen) additional variables that are different from the system variable under investigation may increase the MOS correction (see, e.g., Glahn and Lowry 1972; Sokol 2003). For two model observables, $V_1(t)$ and $V_2(t)$, the MSE between the corrected forecast and the reference solution can also be computed, assuming (without loss of generality) that the variable V_1 is the same nominal variable as X,

$$\langle [X_C(t) - X(t)]^2 \rangle = \langle [V_1(t) - X(t)]^2 \rangle - [\langle X(t) \rangle - \langle V_1(t) \rangle]^2 - \frac{1}{\sigma_1^2\sigma_2^2 - C(V_1, V_2)^2} \begin{Bmatrix} [\sigma_1^2\sigma_2 - \sigma_2 C(X, V_1)]^2 + \sigma_1^2[C(X, V_1)^2 - C(V_1, V_2)^2] \\ + 2C(X, V_1)C(V_1, V_2)[C(V_1, V_2) - C(X, V_2)] \end{Bmatrix}, \tag{6}$$

where the first term is the MSE between the model and the reference variables of interest, the second term is the DC, and the third term contains information on the variances and covariances between the different variables, here referred to as the VC, as was the case for the one-observable MOS scheme.

In Vannitsem and Nicolis (2008), the properties of the MOS forecasts for $n = 1$ and 2 presented above in the presence of both model and initial condition errors have been investigated in the context of low-order chaotic systems. If one decomposes the initial and model errors into random (i.e., the deviation from the mean) and systematic (i.e., the mean) parts, it has been demonstrated that systematic initial errors and model errors can be well corrected when the predictors are well chosen, while random initial condition errors cannot.

More specifically, the DC component of the correction is dominated for short times by the systematic initial condition error, $\langle X(0) \rangle - \langle V_1(0) \rangle$, and/or the systematic model error, $\langle F[\mathbf{X}(0)] \rangle - \langle G[\mathbf{V}(0)] \rangle$, where $F[\mathbf{X}(0)]$ and $G[\mathbf{V}(0)]$ refer to the true and model tendencies for the specific variable of interest evaluated initially at $\mathbf{X}(0)$ and $\mathbf{V}(0)$, respectively. For longer times, this component saturates toward a plateau.

For the VC component, a short-time correction of the random part of the initial conditions can be isolated but its amplitude is relatively small for small initial errors since it depends on the square of $\{\sigma_\epsilon^2 + C[X(0), \epsilon)]\}$, where σ_ϵ^2 is the variance of the initial errors and $C[X(0), \epsilon]$ is the covariance of the initial state with the initial condition error, ϵ. On the other hand, a substantial part of the variability of the model error can be removed provided that it is strongly correlated with the predictors of the MOS equations. For longer times, the VC component increases rapidly once the error dynamics enters into the nonlinear phases of growth. It has also been shown in Vannitsem and Nicolis (2008) that this latter phase is concommittant with both the convergence of the MOS forecast toward the climatological mean and the depletion of the MOS forecast variance. Note that as a corollary of the importance of the covariance between the model error and the MOS predictors on the amplitude of the VC correction, the choice of the (additional) model observables for short times is conditioned by their correlation with the model error.

It is important to emphasize here that the nature of the errors that can be corrected for by the MOS technique is different for the DC and VC components. The DC component is dominated by the systematic parts of the (model and initial) errors, while the VC part is dominated by the variability of these errors (predominantly by the variability of model errors). It is therefore clear that looking at the properties of the DC and VC parts allows us to gain useful information on the sources of the errors degrading the forecasts. In particular when (small) initial random errors constitute the major source of the uncertainty (no important model errors or systematic initial errors), no substantial corrections could be expected.

These two MOS schemes (with one and two variables) will be applied and analyzed for the data coming from the ECMWF model.

3. The dataset

The ECMWF model has been one of the leading operational forecasting systems providing medium-range forecasts since the end of the 1970s. Its forecasting performances are constantly increasing due to the continuous modeling effort performed at the center and the progressive improvement of the initial conditions. For our specific purposes, the MOS equations should in principle be built with an operational system whose properties remain invariant in time. Strictly speaking, this is not the case in such an operational environment since the system is regularly updated (sometimes several times a year). However, the new releases do not always impact dramatically the forecasts. In particular, we are able to isolate a recent period of about 4 yr during which the model has not experienced important modifications, as revealed by the almost constant quality of the forecasts. It covers the period from the end of 2001 up to the middle of 2006 (see the ECMWF Web site: http://www.ecmwf.int). At the end of 2006, the resolution of the model was dramatically increased leading to a strong modification of the forecasting system performance. To test the MOS equations on the ECMWF model, we therefore focus on the period from 1 December 2001 to 30 November 2005, over a domain covering Belgium. In addition, we will only focus here on the forecast cycle starting at 0000 UTC.

Many observables are critical for everyday operational forecasting purposes. One can quote precipitation, temperature at the surface or at 2 m, and cloud cover, among others. In the present work we will focus on one critical observable that can sometimes experience large errors even for short times, the 2-m temperature. The latter quantity is obtained in the ECMWF model by (complex) interpolation between the lowest model level and the surface, making use of the profile functions of the dry static energy ($gz + c_pT$ where g is the gravity, z is the height, c_p the specific heat at constant pressure, and T the temperature) deduced from the Monin–Obukhov similarity theory (IFS documentation, Cy25r1, Cy28r1, ECMWF). Note that these profiles depend on the (vertical) stability conditions of the air mass close to the ground.

Two main experiments are conducted. First, we evaluate the error evolution of the forecast on a 0.5° × 0.5° grid covering the region from the upper-left corner (52°N, 2°E) up to the lowest-right corner (49°N, 7°E), in which measurements at about 60 synoptic stations are currently available. The references in this case are the forecast analyses, assumed to be the best available representation of the truth. We next build the MOS equations using 2 yr of data from 1 December 2001 to 30 November 2003. These equations are developed separately for each season (December–February, DJF; March–May, MAM; June–August, JJA; September–November, SON), for each grid point, and for each forecast range. Each MOS equation is based on a set of about 180 forecasts. Once these equations are constructed, the MOS forecasts are made with first the training dataset and then a set of independent data, covering the period from 1 December 2003 to 30 November 2005.

The second experiment consists of developing the MOS equations and evaluating the MOS forecasts at different synoptic stations located over Belgium. Eight stations were selected in view of their good data quality (records without substantial interruptions): Uccle, Beauvechain, Elsenborn, Florennes, Deurne, Kleine Brogel, Koksijde, and Middelkerke. Some of them are located in the central part of the country (Uccle, Florennes, Beauvechain, and Deurne); Elsenborn and Kleine Brogel are located in the eastern part of Belgium, which is a more undulating region than the central part; and Middelkerke and Koksijde are located in the coastal zone.

As mentioned in the introduction, the use of higher-resolution models (with improved physics) is expected to provide better predictions than the model outputs discussed above. A first set of output data, which will also be used in the sequel, are the predictions provided by the new high-resolution ECMWF model, operational since September 2006 and interpolated on the same grid. A second set of output data are provided by the high-resolution regional ALADIN model, developed within the context of the ALADIN consortium, which is led by Météo-France (ALADIN International Team 1997). This regional model is currently integrated operationally over Belgium with a resolution of 7 km over a domain of about 1600 km × 1600 km. The initial and boundary conditions are provided by the runs of the operational regional ALADIN-France model of Météo-France (with 9.5-km grid spacing), interpolated on the Belgium grid. The boundary and initial conditions of the ALADIN-France model are provided by the global Action de Recherche Petite Echelle Grande Echelle (ARPEGE) model. We will make use of the operational forecasts (obtained with model release 28) that were stored for the stations mentioned above for the period 1 December 2003–30 November 2005.

4. Results

a. Results on the model grid

In this first experiment, the model and MOS forecasts are compared with the analyses used for initializing the model. These analyses are obviously not equivalent to the truth, but this is the best representation of the truth that is compatible with the dynamics of the model. This approach should clearly provide too optimistic results concerning the performances of the model since the comparison is made with a reference compatible with the dynamics of the model, but it allows for an investigation integrated over a specific region. In addition, this approach allows for a corrected forecast distributed on a regular grid and for providing improved spatial fields. In the present paper, we will not address the interesting aspect of the spatial properties of the correction (a question left for future work), but rather its temporal properties.

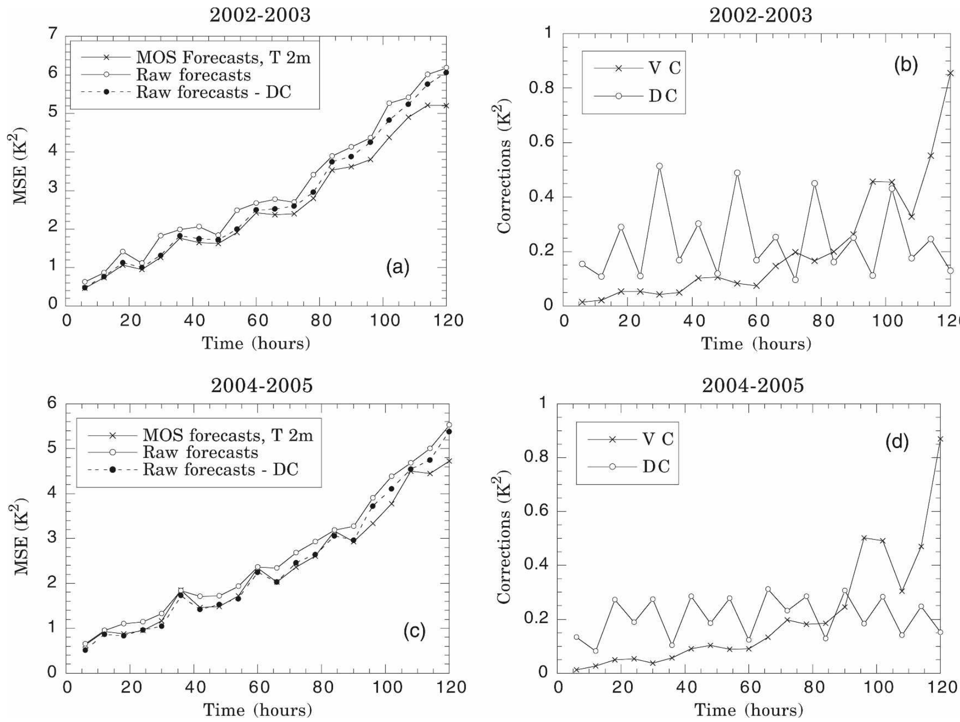

FIG. 1. MSE evolution for the raw and corrected ECMWF 2-m temperature forecasts, averaged over the grid covering Belgium, for (a) the training set (1 Dec 2001–30 Nov 2003) and (c) the verification set (1 Dec 2003–30 Nov 2005). The DC and VC corrections of the MOS forecasts are displayed in (b) and (d) for both sets, respectively. The reference is provided by the analyses on the grid covering Belgium.

Figure 1a shows the MSE evolution for the 2-m temperature over the whole grid covering Belgium for the 2-yr period used for training and for the corrected forecast based on a one-variable MOS equation (the model predictor used is the predicted 2-m temperature). An additional curve shows the MSE evolution when only the DC correction is subtracted. Clearly, the correction is quite substantial and is dominated by the DC correction for short times since both corrected curves are close to each other (crosses and full circles in Fig. 1). On the other hand, the VC term provides an additional correction that becomes progressively more substantial for longer lead times.

The two corrections, VC and DC, are represented in Fig. 1b. For short times, the DC correction is larger than the VC correction, indicating the predominance of a systematic drift of the mean over the whole grid. In addition, it fluctuates considerably in time, reflecting the diurnal variations of the quality of the model forecasts. On the other hand, besides the diurnal variation, the systematic drift seems to have rapidly reached an overall plateau after 6 h, suggesting that the short-time behavior discussed in Vannitsem and Nicolis (2008) is already over for DC. For VC, the growth is slow initially, and after a few days it increases rapidly, suggesting that the MOS forecasts progressively converge toward the climatological mean, as discussed in section 2.

To clarify the evolution of the MOS forecast, we have computed the two first moments of the forecasts. These are illustrated in Fig. 2 for winter averaged over the region of interest (the whole grid). As expected, the mean is better in the MOS forecast than in the model forecast. For the variance, the quality of the MOS forecast is close to the one of the analyses for short times, but degrades for longer times. As was already mentioned in section 2, this behavior reflects the convergence of the MOS forecasts toward the climatological mean. In Vannitsem and Nicolis (2008), this depletion (arising after about a day) has been shown to be associated with the increasing role played by the nonlinear terms on the error dynamics in low-order chaotic systems. For the data discussed here, this suggests that the nonlinear terms start to play a role in the error dynamics only after 1 day (or so). This result is consistent with the estimate of the duration of validity of the linearized dynamics in the ECMWF model of the order of 1 day (Gilmour et al. 2001).

The same conclusions are reached when using the MOS equations on the independent dataset covering the period from 1 December 2003 to 30 November 2005

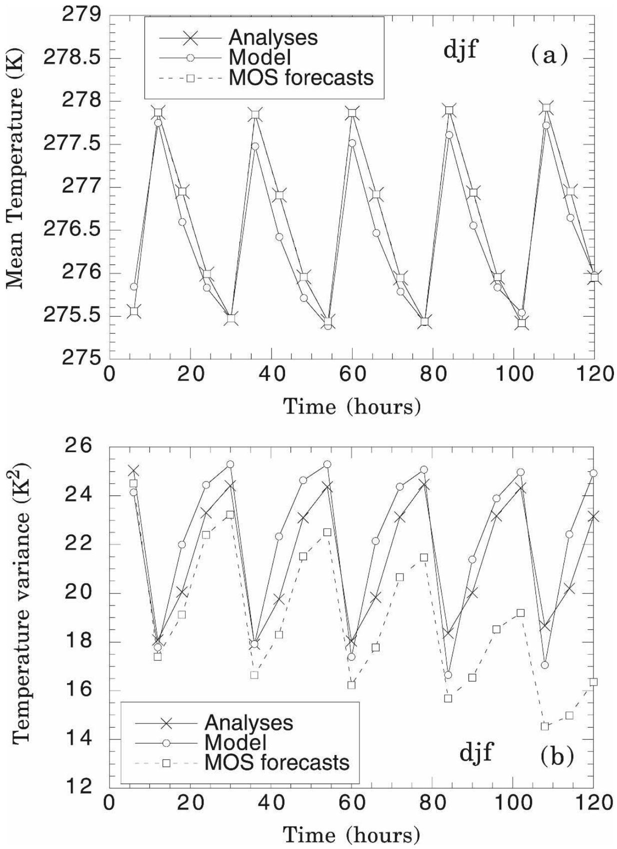

FIG. 2. (a) Evolution of the mean 2-m temperature, averaged during two winter periods of the training set and on the grid covering Belgium, for the analyses, the model forecasts, and the one-observable MOS forecasts. (b) As in (a), but for the temperature variance.

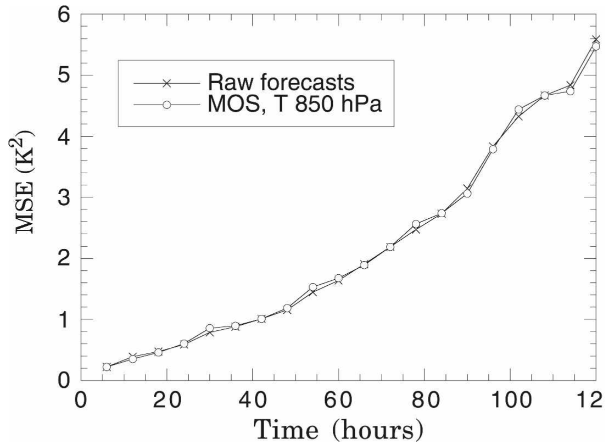

FIG. 3. MSE evolution, averaged over the whole grid covering Belgium, for the temperature at 850 hPa forecasted by the model (crosses) and corrected by the MOS equations (open circles). The reference is provided by the analyses on the grid covering Belgium.

(Figs. 1c and 1d), except that the corrections are slightly smaller than for the training set. This result could of course be expected since the MOS correction is optimized for the training set, but it suggests further that small modifications to the forecasting system (which are present between the training and verification periods) do not significantly alter the ability of the MOS equations to provide a better result than the raw forecasts. This contrasts with the result obtained when these MOS equations are used in the ECMWF model output covering the period from 1 December 2006 to 30 November 2007, for which the correction is not better than the model forecast itself. This reflects the major modifications made in the ECMWF model release provided in 2006, which cannot be accommodated by the MOS equations built from a previous forecasting period.

When additional variables are used (temperature at different levels, sensible heat flux, soil temperature, latent heat flux, relative humidity, wind velocity, etc.), the gain is not substantial, except for temperature at 850 hPa for which a small overall improvement is visible. This point will be discussed more extensively in the next section when studying the results for the synoptic stations.

Note that the application of the same experiment but for temperature at higher altitudes does not show any substantial improvements in the forecasts, as illustrated in Fig. 3 for temperature at 850 hPa. This very interesting feature suggests that if model errors are playing a role in the error evolution, their impact is quite small as compared to the (random) initial condition errors since no corrections are obtained from the MOS equations, as discussed in section 2. This further reveals that in the bulk of the atmosphere, the impacts of model errors seem much less important than the initial condition errors, which justifies the large effort toward the improvement of the data assimilation schemes and the data collection in the atmosphere. This result does not of course imply that the model is perfect at upper levels and does not advocate for neglecting the potential gain that can be reached by improving the model dynamics and physics at these levels.

b. *Results at synoptic stations*

Figure 4 displays the same quantities as in Figs. 1a and 1b but evaluated at three representative synoptic stations (Uccle, Elsenborn, and Middelkerke). In this case, the raw forecast is simply provided by the model forecast obtained at the closest grid point of the grid. The initial amplitude of the error is now larger than on the grid. This feature should be first related to the fact

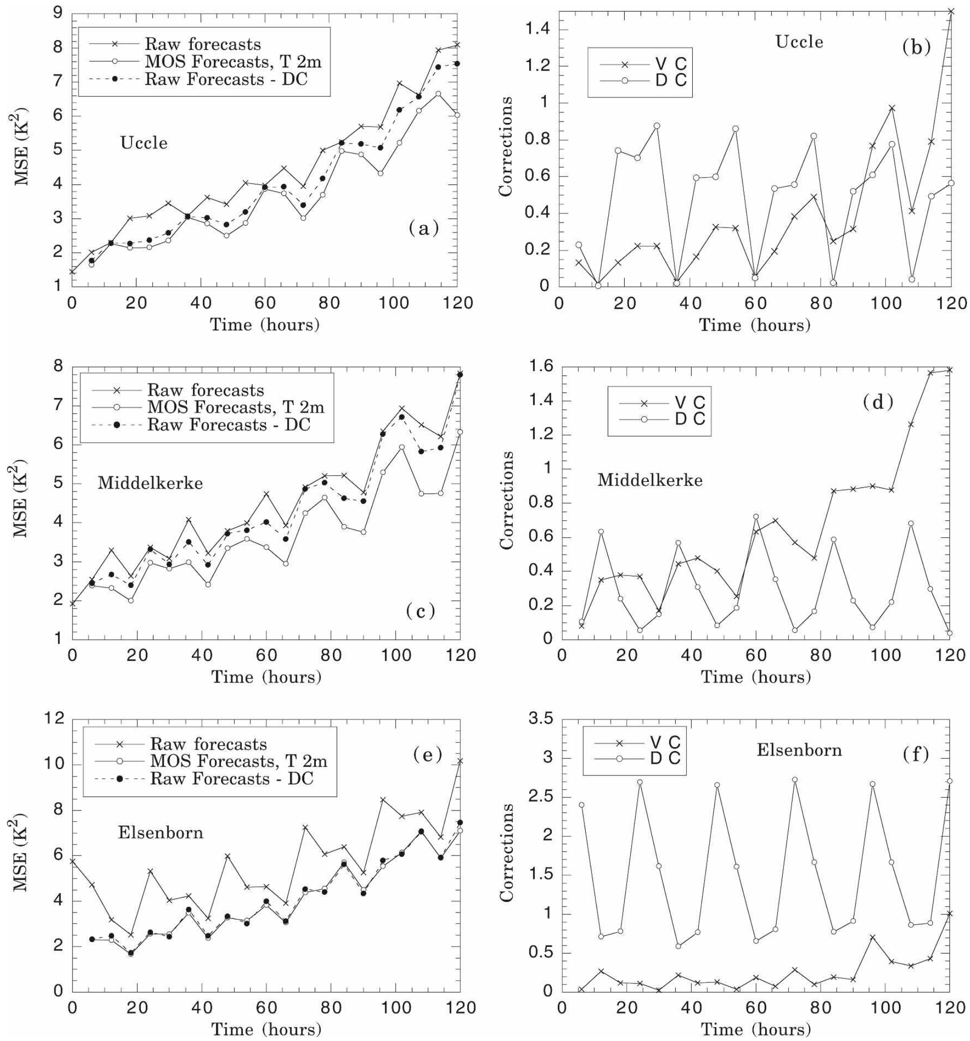

FIG. 4. MSE evolution for the raw and corrected ECMWF 2-m temperature forecasts, evaluated at three synoptic stations (a) Uccle, (c) Middelkerke, and (e) Elsenborn, for the training dataset (1 Dec 2001–30 Nov 2003). (b), (d), (f) As in (a), (c), (e), but for the DC and VC corrections.

that the station is not located at a specific grid point of the model and, second, that the consistency of the "well balanced" analyses with the possible natural solution of the model is now lost. Instead, the comparison is made with natural observations that are not living in the same phase space as the model.

The results displayed in Fig. 4 suggest that large corrections can be obtained by using the MOS equations (cf. the continuous lines in each panel). The nature of the dominant corrections for Elsenborn and Uccle is the drift correction, which is even larger for Elsenborn. This feature is further emphasized by comparing the MOS-corrected curve (open circles) and the DC-corrected curve (full circles) for which only the DC correction has been subtracted.

For Middelkerke (located in the coastal zone), the picture is different since the correction based on one-variable MOS equations can for some lead times be dominated by the VC correction. Note that this effect can be even more important when looking season by season. For instance, in summer, the VC correction is more important than the drift correction for all lead times, indicating that the variability is less well represented during this season for Middelkerke (not shown). These results highlight the fact that the nature of the correction of the forecasts on a small country like Belgium strongly depends on the climatological properties (distance from the coast, local orography, etc.).

The application of the MOS equations to the verification dataset (Fig. 5) shows improvements similar to the ones obtained with the training dataset (Fig. 4).

To disentangle the respective roles of the model and

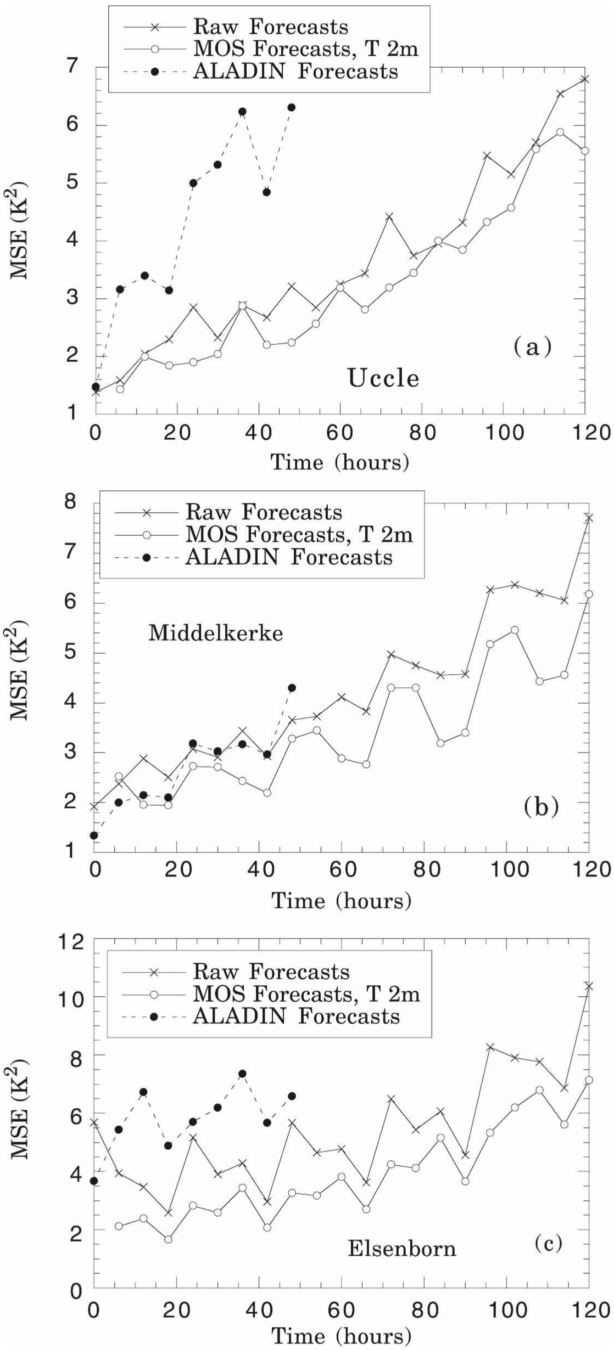

FIG. 5. MSE evolution for the raw and corrected ECMWF 2-m temperature forecasts, evaluated at three synoptic stations: (a) Uccle, (b) Middelkerke, and (c) Elsenborn, for the verification dataset (1 Dec 2003–30 Nov 2005). The full dots represent the MSE of the ALADIN forecasts.

initial condition errors on the forecasts and the MOS corrections, the amplitude of the systematic initial error is estimated [evaluation that was not possible with the investigation performed on the grid of the previous section since $\mathbf{V}(0)$ and $\mathbf{X}(0)$ are identical in this case and equal to the analyses]. Indeed, this quantity enters into the behavior of the DC correction as discussed in section 2. To this aim, we compute the difference between the analysis used for the forecast and the observation (of course, some measurement errors in the observations are also present but they will be considered to be negligible). For the whole period from December 2003 to November 2005, the systematic errors are equal to -0.48, -0.07, and 1.82 K for Uccle, Middelkerke, and Elsenborn, respectively. For Elsenborn, the amplitude of the systematic error is very large, it is smaller for Uccle, and negligible for Middelkerke, implying different impacts on the amplitude of the DC correction. The results shown in Fig. 4 for DC reflect the influence of this systematic error.

On the other hand, the variability corrections do not contain a substantial signature of the initial condition error but do of the model error that could be corrected by the MOS technique. The VC correction is larger for short times for Middelkerke than for Uccle and Elsenborn, suggesting that the modeling error is more important in the coastal zone than in the interior of the country and is probably of another origin. This aspect can be partially answered for through the investigation of the MOS techniques based on a larger number of predictors.

When more variables are used, some improvements can be gained, as indicated in Fig. 6a for Uccle. In particular, the use of the temperature forecasts at 850 hPa improves considerably the MOS forecasts. As discussed in Vannitsem and Nicolis (2008), this feature reflects that a substantial "model error" is present and is associated with this specific model observable. This quite vague assertion could be understood as follows. First, the temperature at 2 m is evaluated by interpolating the temperature obtained at the ground and at the lowest model level (see section 3 for more details). This interpolation is of course an approximation of the true temperature at 2 m, which could lead to large model errors. The temperature at 850 hPa seems to provide a correction indicating that model errors can be related to the vertical discretization and interpolation scheme, discussed in section 3. A modification of this scheme or of the vertical discretization could provide a better representation of the 2-m temperature.

The above conjecture is supported by the analysis of the MSE evolution obtained with the higher-resolution model whose use started in September 2006. In Fig. 7 the MSE for 2007 (1 December 2006–30 November 2007) is compared with the one obtained for the period 2004–05 (1 December 2003–30 November 2005). The change of (vertical) resolution seems to have improved

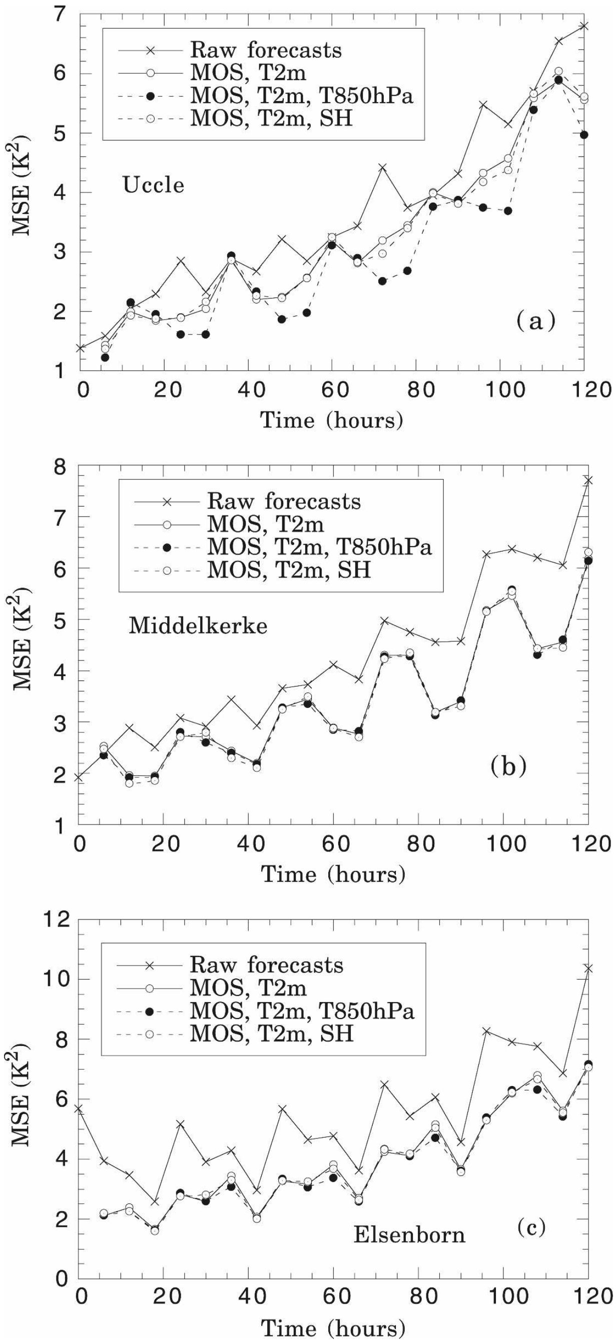

FIG. 6. As in Fig. 5, but with additional MOS schemes, namely two two-observable MOS schemes, one with the temperature at 850 hPa as a second predictor and one with the sensible heat flux (SH).

the quality of the forecasts in the sense that the temporal evolution of the MSE for 2007 is closer to the one obtained by correcting the model in 2004–05. Interestingly, the out-of-phase feature between the error evolution of the raw forecast and the MOS forecasts is now partly removed. Of course, this is not a definitive answer since there is also a large variability in the year-to-year performances of the model, but it is an encouraging result.

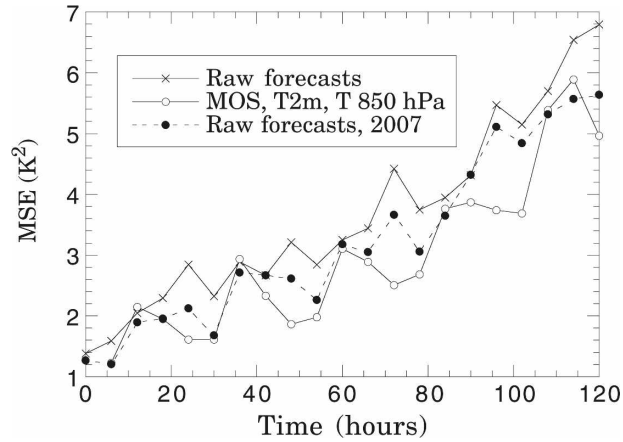

FIG. 7. MSE evolution during the verification period for Uccle. The results displayed correspond to the raw forecasts, the two-observables MOS scheme with the temperature at 850 hPa as a second predictor, and the raw forecasts obtained with the high-resolution ECMWF model operational since September 2006.

Note that the improvement obtained with the 850-hPa temperature forecasts (as a second predictor) cannot be generalized for all seasons and all regions. In the central part of Belgium, it essentially provides large improvements during spring and summer. This effect in spring and summer has also been observed at other stations located in the same region (e.g., Florennes, Beauvechain, and Deurne) and, to a much less extent, in the eastern part of Belgium (e.g., Elsenborn, Kleine Brogel); see also Fig. 6c for Elsenborn (for the whole year). Interestingly, in the coastal zone, this (second) predictor is not the best one but rather it is the sensible heat flux that provides the best corrections at the synoptic stations of Middelkerke and Koksijde, suggesting that the model errors present in the coastal zone should be related to the evaluation of the surface temperature, which depends on the sensible heat flux. Figure 6b shows the impact of the sensible heat flux for Middelkerke for the whole year. A small improvement is visible (in the afternoon) as compared to the one obtained with the temperature at 850 hPa. This effect is much more substantial in summer, with an improvement of the order of 10%–20% (in the afternoon of the first 2 days of the forecast) as compared to the result obtained with the temperature at 850 hPa (not shown). Sensible heat flux was also considered in Termonia and Deckmyn (2007) as one of the main sources of errors in the ALADIN model running over Belgium. This property can be well understood since the coastal zone is a region interfacing two different surfaces, whose flux

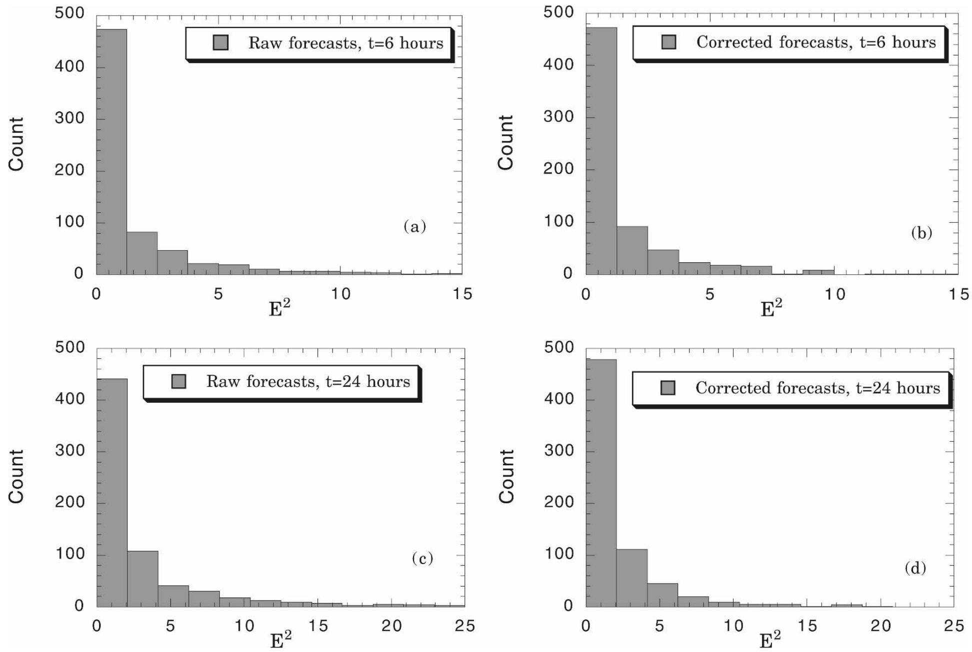

FIG. 8. Histogram of the instantaneous squared error (K^2) for Uccle: (a) the raw forecasts at 6-h lead time, (b) the corrected MOS forecasts at 6-h lead time, (c) the raw forecasts at 24-h lead time, and (d) the corrected MOS forecasts at 24-h lead time.

properties are drastically different. Modeling these features is therefore an important challenge that is not at all obvious within the context of a global atmospheric model whose grid spacing is of the order of a few tens of kilometers (Kalnay 2003).

Finally, Fig. 8 displays the histograms of the squared errors (E^2) obtained for the one-variable MOS experiment whose average is shown in Fig. 5a (Uccle). In Figs. 8a and 8c, the histograms of the raw forecasts at two different lead times (6 and 24 h) are displayed, while Figs. 8b and 8d show the histograms for the corrected forecasts. The tails of the histograms are less populated in the cases of the corrected forecasts, indicating the positive impact of the MOS correction on the bad forecasts. Note that at $t = 60$ h, the histograms of the raw and corrected forecasts are almost indistinguishable, indicating that the MOS technique has a neutral impact for this lead time (not shown). This is also reflected in the MSE evolution in Fig. 5a.

c. Comparison with a high-resolution regional forecasting model

Over the past few decades, significant efforts have been devoted to the development of regional models whose resolution is much higher than the global models. This change in the scale of description is accompanied by a finer description of the processes acting at smaller scales. This feature should obviously improve the quality of the model and hopefully of the forecasts. However, to get a very powerful forecasting system, one also needs good initial condition quality and a regional model that is not strongly affected by the lateral boundaries (e.g., Anthes et al. 1985; Mesinger 1996; Vannitsem and Chomé 2005; Alexandru et al. 2007). Within this context, one important question is whether these high-resolution models can compete with global ones. We have tested this point by evaluating the quality of the forecasts of the ALADIN model that was running for operational purposes over Belgium during the same period of time (1 December 2003–30 November 2005).

Figure 5 displays the results obtained at the three representative stations with both the global ECMWF model and the regional release 28 ALADIN model (full dots, dashed line). Note that the use of a more recent release (29), in which a better representation of low clouds is implemented (Brozkova et al. 2006), improves the MSE by an amount up to 1 K^2 for lead times beyond a day (P. Termonia 2008, personal communication). This reflects that the physics of the ALADIN model currently running over Belgium has been considerably improved as compared with release 28.

Two important features should be emphasized here: First, the initial condition error of the regional model estimated at the station location is not much better than the one obtained with the global model. This feature can be explained partly by the fact that the initial conditions of the ALADIN model are here based on an interpolation over the 7-km grid of the initial state provided for the ALADIN-France model. Second, the error amplification is faster for Uccle and Middelkerke, suggesting the presence of a rapid error amplification whose origin is different than that of the global model. Two possible reasons for this latter behavior can be advanced: (i) The model has an intrinsic instability that is faster than the one associated to the global model and (ii) errors arising from the coupling with a large-scale model at the boundaries of the regional model rapidly affect the regional solution. Concerning the first reason, it is indeed possible that fast instabilities associated with small-scale convective or gravity wave processes affect the predictions, in which case extreme care should be put into determining the initial conditions in order to compete with global-scale models. Concerning the second possible reason, several works have indicated the necessity of carefully defining the model coupling at the lateral boundaries of the regional system to get a good prediction (e.g., Termonia 2003), but also to have a sufficiently large domain of integration in order to avoid the strong influence of the one-way nesting procedure used in this context (Vannitsem and Chomé 2005, and references therein). The current domain of the ALADIN model integrated over Belgium is not very large (1600 km $\times$ 1600 km), indicating that large potential errors can arise from this operational setting. In summary, both effects are probably acting in the present case with a local performance that is apparently less good than the global-scale model.

Improvements of the regional forecasts should therefore be strongly related with the use of a regional data assimilation system and the use of lateral boundaries that do not greatly affect the regional output, in particular by increasing the domain size. Consequently, studies on the impact of boundary conditions should be carefully conducted in order to define the best possible domain, also taking into account the computational constraints of the operational environment.

5. Conclusions

A recent investigation on the dynamical properties of the MOS forecasts within the context of low-order chaotic systems has allowed the respective roles of the initial and model errors on the forecasts to be disentangled (Vannitsem and Nicolis 2008). The present work is a follow-up of this analysis with the aim at understanding what are the respective roles of the model and initial condition errors within the context of a realistic operational weather prediction system, in this case the ECMWF forecasting system.

The analysis has revealed several interesting results that can be summarized as follows. (i) The error amplification arising from the presence of uncertainties in the initial conditions dominates the error dynamics in the free atmosphere (at least for temperature). One of the consequences of this result is that for these levels of the model (and this type of variable), forecast improvements should result from an improvement of the initial conditions. (ii) Forecast errors at 2 m can be partly corrected for by the use of the MOS technique, suggesting the presence of either systematic initial condition errors or model errors (or both). In the coastal zone, this initial systematic bias is negligible, indicating that the MOS forecast is mainly correcting model errors. (iii) The analysis of the two-observable MOS schemes reveals that the best second predictor is the temperature predicted at 850 hPa in the central part of the country and the sensible heat flux (entering into the evaluation of the surface temperature) for the coast. This suggests that for 2-m temperatures away from the coast, the vertical interpolation scheme can partly cause the model error and the correction of this problem should follow from increasing the resolution or through modifications of the interpolation scheme. On the other hand, for the coastal zone, the source of the model error seems, rather, to be associated with the description of the surface fluxes in the region interfacing the sea and the continent. The latter result is to be expected since the global models are not able to give an accurate description of the specific dynamics occurring over a few kilometers in coastal zones (Kalnay 2003). (iv) The comparison with a regional model reveals that the global model still provides competitive forecasts even locally. This (old) regional model release (integrated on a quite small domain) gave poor results in view of the legitimate expectations from a very high-resolution system. It further suggests that when using regional models, extreme care should be taken when constructing the initial conditions and in the choice of the model lateral boundaries.

Further investigations should be performed for other variables (surface wind and pressure, cloud cover, precipitation, etc.). Additional complications are expected, especially for precipitation whose highly erratic nature can lead to the use of more specific MOS techniques than the linear approach used so far in this work.

Acknowledgments. This work has benefited from enlightening discussions with C. Nicolis and P. Termonia. The two anonymous reviewers are acknowledged for their constructive comments. This work is partly supported by the Belgian Federal Science Policy Program under Contract MO/34/017.

REFERENCES

ALADIN International Team, 1997: The ALADIN project: Mesoscale modelling seen as a basic tool for weather forecasting and atmospheric research. *WMO Bull.,* **46,** 317–324.

Alexandru, A., R. de Elia, and R. Laprise, 2007: Internal variability in regional climate downscaling at the seasonal scale. *Mon. Wea. Rev.,* **135,** 3221–3238.

Anthes, R. A., Y.-H. Kuo, D. P. Baumhefner, R. M. Errico, and T. W. Bettge, 1985: Predictability of mesoscale motions. *Advances in Geophysics,* Vol. 28B, Academic Press, 159–202.

Brozkova, R., M. Derkova, M. Bellus, and A. Farda, 2006: Atmospheric forcing by ALADIN/MFSTEP and MFSTEP oriented tunings. *Ocean Sci.,* **3,** 1–24.

Cheng, W. Y. Y., and W. J. Steenburgh, 2007: Strengths and weaknesses of MOS, running-mean bias removal, and Kalman filter techniques for improving model forecasts over the western United States. *Wea. Forecasting,* **22,** 1304–1318.

Dallavalle, J. P., M. C. Erickson, and J. C. Maloney III, 2004: Model output statistics (MOS) guidance for short-range projections. Preprints, *20th Conf. on Weather Analysis and Forecasting and 16th Conf. on Numerical Weather Prediction,* Seattle, WA, Amer. Meteor. Soc., 6.1. [Available online at http://ams.confex.com/ams/pdfpapers/73764.pdf.]

Gilmour, I., L. A. Smith, and R. Buizza, 2001: On the duration of the linear regime: Is 24 hours a long time in weather forecasting? *J. Atmos. Sci.,* **58,** 3525–3539.

Glahn, H. R., and D. A. Lowry, 1972: The use of model output statistics (MOS) in objective weather forecasting. *J. Appl. Meteor.,* **11,** 1203–1211.

Hart, K. A., W. J. Steenburgh, D. J. Onton, and A. J. Siffert, 2004: An evaluation of mesoscale-model-based model output statistics (MOS) during the 2002 Olympic and Paralympic Winter Games. *Wea. Forecasting,* **19,** 200–218.

Kalnay, E., 2003: *Atmospheric Modeling, Data Assimilation and Predictability.* Cambridge University Press, 341 pp.

Klein, W. H., and H. R. Glahn, 1974: Forecasting local weather by means of model output statistics. *Bull. Amer. Meteor. Soc.,* **55,** 1217–1227.

Lemcke, C., and S. Kruizinga, 1988: Model output statistics forecasts: Three years of operational experience in the Netherlands. *Mon. Wea. Rev.,* **116,** 1077–1090.

Mesinger, F., 1996: Improvement in quantitative precipitation forecasts with the Eta Regional Model at the National Centers for Environmental Prediction: The 48-km upgrade. *Bull. Amer. Meteor. Soc.,* **77,** 2637–2649.

Sokol, Z., 2003: MOS-based precipitation forecasts for river basins. *Wea. Forecasting,* **18,** 769–781.

Taylor, A. A., and L. M. Leslie, 2005: A single-station approach to model output statistics temperature forecast error assessment. *Wea. Forecasting,* **20,** 1006–1020.

Termonia, P., 2003: Monitoring and improving the temporal interpolation of lateral-boundary coupling data for limited-area models. *Mon. Wea. Rev.,* **131,** 2450–2463.

——, and A. Deckmyn, 2007: Model-inspired predictors for model output statistics. *Mon. Wea. Rev.,* **135,** 3496–3505.

Vannitsem, S., and F. Chomé, 2005: One-way nested regional climate simulations and domain size. *J. Climate,* **18,** 229–233.

——, and C. Nicolis, 2008: Dynamical properties of model output statistics forecasts. *Mon. Wea. Rev.,* **136,** 405–419.

Wilks, D. S., 2006: *Statistical Methods in the Atmospheric Sciences.* Academic Press, 464 pp.