IDŐJÁRÁS

Quarterly Journal of the Hungarian Meteorological Service
Vol. 114, No. 1–2, January – June 2010, pp. 1–19

Diagnosis and tuning of background error statistics in a variational data assimilation system

Gergely Bölöni[1*] and **Kristian Horvath**[2]

[1]*Hungarian Meteorological Service*
P.O. Box 38, H-1525 Budapest, Hungary; E-mail: boloni.g@met.hu

[2]*Meteorological and Hydrological Service of Croatia*
Gric 3, Zagreb, HR-10000, Croatia; E-mail: horvath@cirus.dhz.hr

[*]*Corresponding author*

(Manuscript received in final form November 9, 2009)

Abstract — A very important source of information in data assimilation is the background, which is usually a short-range forecast of the numerical weather prediction (NWP) model valid at the assimilation time. The background is corrected by the atmospheric observations providing the analysis during the process of data assimilation. Errors of the background are estimated statistically and are taken into account with the aim of giving a proper weight to both the background and observations. Although, in theory, the background errors can be defined as the difference between the truth and the background, it is a great challenge to generate appropriate background error samples in practice for the computation of the background error statistics, because the true state is always unknown. A possible way to improve the background error representation is the development of the background sampling methods themselves. Another complementary approach is the "a posteriori" diagnosis or tuning of the already predefined background error statistics by a certain sampling technique. This paper presents two attempts for such a tuning within the ALADIN limited area model (LAM) and its 3d-var assimilation system used operationally at the Hungarian Meteorological Service (HMS). The first one is based on the comparison of the predefined statistics with those expected in a statistically optimal assimilation system. The second one is inspired by single observation experiments and it addresses the improvement of the multivariate statistical balance between humidity and the other analyzed variables.

Key-words: data assimilation, background errors, "a posteriori" tuning, multivariate statistical balance

1. Introduction

Data assimilation methods based on statistical optimality (*Gandin*, 1963; *Lorenc*, 1986; *Thépaut* and *Courtier*, 1991; *Bouttier* and *Courtier*, 1999) take into account the error of each available information type for weighting their contribution to the analysis $\mathbf{x}_a$, i.e., the initial condition for the forecast. This is done by solving the so-called BLUE (best linear unbiased estimation) analysis equation:

$$\mathbf{x}_a = \mathbf{x}_b + \mathbf{B}\mathbf{H}^T\left(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R}\right)^{-1}\left(\mathbf{y} - H(\mathbf{x}_b)\right). \quad (1)$$

Main available information are the atmospheric observations $\mathbf{y}$ and the background $\mathbf{x}_b$, which is usually a short-range forecast of the NWP model. The notation H stands for the so-called non-linear observation operator, which projects from the space of the model variables to that of the observed variables. The operator $\mathbf{H}$ is the linearized of H. The matrices $\mathbf{B} = \mathrm{E}\left(\boldsymbol{\varepsilon}_b \boldsymbol{\varepsilon}_b^T\right)$ and $\mathbf{R} = \mathrm{E}\left(\boldsymbol{\varepsilon}_o \boldsymbol{\varepsilon}_o^T\right)$ denote the background and observation error covariance matrices respectively, where E stands for the statistical expectation. The corresponding errors are defined as the difference from the true state of the atmosphere $\mathbf{x}_t$, i.e., $\boldsymbol{\varepsilon}_b = \mathbf{x}_b - \mathbf{x}_t$ can be written for the background errors and $\boldsymbol{\varepsilon}_o = \mathbf{y} - \mathbf{H}(\mathbf{x}_t)$ for the observation errors. In the lack of $\mathbf{x}_t$ the sampling of these errors is a real challenge in the field of data assimilation. For the sampling of background errors, many ideas have been developed, like the method of innovations (*Hollingsworth* and *Lönnberg*, 1986; *Lönnberg* and *Hollingsworth*, 1986) or the NMC (*Parrish* and *Derber*, 1992) and Ensemble (*Fisher,* 2003; *Belo Pereira* and *Berre*, 2006; *Stefanescu et al.,* 2006) methods. Also attempts have been taken to characterize LAM specific background errors with the lagged-NMC approach (*Siroká et al.*, 2003). With the application of the above sampling techniques the representation of the background errors has been improved in some aspects indeed. In this paper, however, we would like to point out, that as $\mathbf{x}_t$ is never known, it is impossible to compute perfect $\mathbf{B}$ and $\mathbf{R}$ matrices with any method, and it is always reasonable to tune certain elements of the predefined covariance matrices "a posteriori". We will concentrate on the tuning of the $\mathbf{B}$ matrix statistics primarily, which was applied to the 3d-var system of the ALADIN model (*Horányi et al.*, 1996) in our experiments. In Chapter 2 we describe a tuning experiment aiming to improve the background error variances based on statistical optimality criteria of analysis residuals. In Chapter 3 we show a tuning to improve the multivariate balance in the analysis based on the simplified analysis Eq. (1) and single observation tests.

2. Tuning of the background error statistics based on covariances of residuals

Based on the optimal estimation theory, several approaches exist for "a posteriori" validation and tuning. A technique applied earlier in ARPEGE/ALADIN was based on the assumption that the expected value of the variational cost function J at the minimum is proportional to the number of observations used, i.e.,

$E(J_{\min})=p/2$ hold, where p is the number of observations used in the analysis (*Talagrand*, 1998; *Désroziers* and *Ivanov*, 2001; *Chapnik et al.*, 2004; *Sadiki* and *Fischer*, 2005; *Fischer et al.*, 2005). Another method was proposed by *Désroziers et al.* (2006) later on, which is the theoretical basis for the tuning described in this chapter.

2.1. The method of tuning

It is explained in the above-mentioned paper that in a linear analysis system, where the **B** background and the **R** observation error covariance matrices are properly estimated, the following equations hold:

$$\begin{aligned}
&\mathrm{E}\left(\mathbf{d}_b^o \mathbf{d}_b^{oT}\right) = \mathbf{HBH}^T + \mathbf{R} \\
&\mathrm{E}\left(\mathbf{d}_b^a \mathbf{d}_b^{oT}\right) = \mathbf{HBH}^T \\
&\mathrm{E}\left(\mathbf{d}_a^o \mathbf{d}_b^{oT}\right) = \mathbf{R} \\
&\mathrm{E}\left(\mathbf{d}_b^a \mathbf{d}_a^{oT}\right) = \mathbf{HAH}^T
\end{aligned} \tag{2a}$$

where

$$\begin{aligned}
&\mathbf{d}_b^o = \mathbf{y} - \mathbf{H}(\mathbf{x}_b) \\
&\mathbf{d}_b^a = \mathbf{H}(\mathbf{x}_a) - \mathbf{H}(\mathbf{x}_b) \\
&\mathbf{d}_a^o = \mathbf{y} - \mathbf{H}(\mathbf{x}_a)
\end{aligned} \tag{2b}$$

are residuals of the model and the observations provided by the assimilation system. The notation **A** in Eq. (2a) stands for the analysis error covariance matrix. The optimality of the analysis system can be diagnosed if one computes the covariances of residuals on the left-hand side and substitutes the predefined error statistics to the right-hand side of Eq. (2a). The residuals on the left-hand side can be easily obtained as a by-product of the assimilation system. The second and the third equations in Eq. (2a) can be applied for instance for the diagnosis of the background and observation error standard deviations:

$$\sigma_{bd} = \sqrt{\frac{1}{P}\sum_{i=1}^{P} d_{bi}^a d_{bi}^o} \qquad \sigma_{od} = \sqrt{\frac{1}{P}\sum_{i=1}^{P} d_{ai}^o d_{bi}^o}\,, \tag{3}$$

where d_{bi}^a, d_{oi}^a, and d_{bi}^o stand for the individual i realizations of the background and analysis departures ($i=1,\ldots,P$). Then the misfit of the predefined standard deviations (σ_{bp} and σ_{op}) can be obtained as follows:

$$r_b = \frac{\sigma_{bd}}{\sigma_{bp}} \qquad r_o = \frac{\sigma_{od}}{\sigma_{op}}\,. \tag{4}$$

These misfit ratios can be used as guidance for tuning the predefined standard deviations in the assimilation system. One has to notice, that through Eq. (1), the Eq. (2b) residuals depend on σ_{bp} and σ_{op}, which means that ideally the tuning with the misfit ratios and the computation of the analysis (Eq. (1)) should be done iteratively until r_b and r_o converge. It has also been shown by *Désroziers et al.* (2006) in a simplified system, that the convergence can be reached in a few iterations.

2.2. Estimation of the misfit ratios in our analysis system

The above method was applied in the ALADIN 3d-var system used at HMS (*Bölöni*, 2006) for the estimation of the misfit ratio for the background error standard deviations. The Eq. (2b) residuals were taken from two assimilation cycles run over an autumn (October–November 2005) and a summer (June 2006) period. In these assimilation cycles all the available observations were included, which are used operationally in the ALADIN 3d-var system at HMS (*Randriamampianina*, 2006). However, for the diagnosis of the background error standard deviations, only the residuals based on direct observations (i.e., radiosondes and aircrafts) were used. There was no iteration applied in the estimation for the sake of simplicity. In the fully spectral version of the ALADIN 3d-var system, predefined background error standard deviations are uniform horizontally and variable in the vertical, so they are available as one value for each model level and analyzed meteorological variables (i.e., temperature (T), specific humidity (q), vorticity (ζ), divergence (η)) (*Berre*, 2000). On the other hand, the diagnosed standard deviations can be obtained at the observation locations and for the observed quantities (T, q, and wind u, v components). In our estimation, a vertical averaging was applied both for the predefined and diagnosed values so that they became comparable independently from the height. For the comparison of predefined (ζ,η) and diagnosed (uv) wind error values, an average wind standard deviation ($\sigma_b(uv)$) was defined that can be computed as:

$$\sigma_{bd}(uv) = \sqrt{\frac{1}{2}\left(\sigma_{bd}^2(u) + \sigma_{bd}^2(v)\right)} \tag{5}$$

from the diagnosed u and v standard deviations values, and as:

$$\sigma_{bp}(uv) = \sqrt{-\frac{1}{2}\Delta^{-1}\left(\sigma_{bp}^2(\zeta) + \sigma_{bp}^2(\eta)\right)} \tag{6}$$

from the predefined ζ and η standard deviations, where Δ denotes the Laplace operator, i.e., the second space derivative in horizontal. *Table 1* shows the misfit ratios computed as explained above and also the predefined and diagnosed background error standard deviations both for the autumn and summer periods.

Table 1. Predefined, diagnosed background error standard deviations and misfit ratios (dimensionless) for specific humidity (kg/kg), temperature (K) and average wind speed (m s^{-1}) computed over the autumn (October 26 –November 10, 2005) and the summer (June 05 –20, 2006) periods

Variable	**Predefined**	**Autumn period**		**Summer period**	
		Diagnosed	**Misfit ratio**	**Diagnosed**	**Misfit ratio**
	(σ_{bp})	(σ_{bd})	(r_b)	(σ_{bd})	(r_b)
Specific humidity (q)	2.27×10^{-4}	5.34×10^{-4}	2.35	5.82×10^{-4}	2.56
Temperature (T)	0.4917	0.7071	1.43	0.8010	1.62
Wind (u,v)	1.4840	1.9878	1.33	1.9203	1.29
Average	**0.65**	**0.89**	**1.36**	**0.9**	**1.38**

Large misfit ratios for humidity suggest that the humidity background error standard deviations are the less accurate as predefined by the original **B** matrix in our assimilation system. The misfit ratios are slightly different for the two different time periods, which suggests a possible seasonal dependence of the accuracy in the predefined background error modeling.

2.3. The tuning experiments

The analysis system can be tuned by multiplying the predefined standard deviations with the misfit ratios introduced in the previous section. In our experiments such tuning was done only for the background error standard deviations in two steps with different complexity:

1. ENS1 uses a uniform misfit ratio for all the variables and vertical model levels. This means that the misfit ratio used was averaged over the vertical levels and over the variables q, T, and wind.

2. ENS2 uses a variable dependent misfit ratio still averaged over the vertical levels.

It is to be mentioned, that the predefined statistics were computed based on the Ensemble sampling method. Both the ENS1 and ENS2 tuning options were tested in real assimilation cycling experiments over the summer period. It means a 16-day assimilation cycle (64 analysis steps) using the operational observational dataset. Simulations of 48-hour production runs were started from 00 UTC after 2 days of warm up cycling. The impact of the tuning was measured through rmse and bias score (e.g., *Wilks*, 1995) computations of the above-mentioned production forecasts against surface and radiosonde observations. The reference assimilation cycle for the tuning experiments (ENS0) was the same as ENS1 and ENS2, except that no tuning of the background error standard deviations was applied, i.e., the predefined values were used. In experiment ENS1, a uniform misfit ratio, r_b=1.3 was used as a multiplication factor to increase the predefined background error standard deviations. This value was chosen based on the average misfit ratios shown in *Table 1*. The verification results show that the ENS1 tuning had no impact on the 2-meter parameters (not shown) but did improve the humidity forecasts on 700 hPa and the wind forecasts on 250 hPa (*Fig. 1b* and *d*). The impact on geopotential and temperature in the altitude was found to be rather neutral (*Fig. 1a* and *c*).

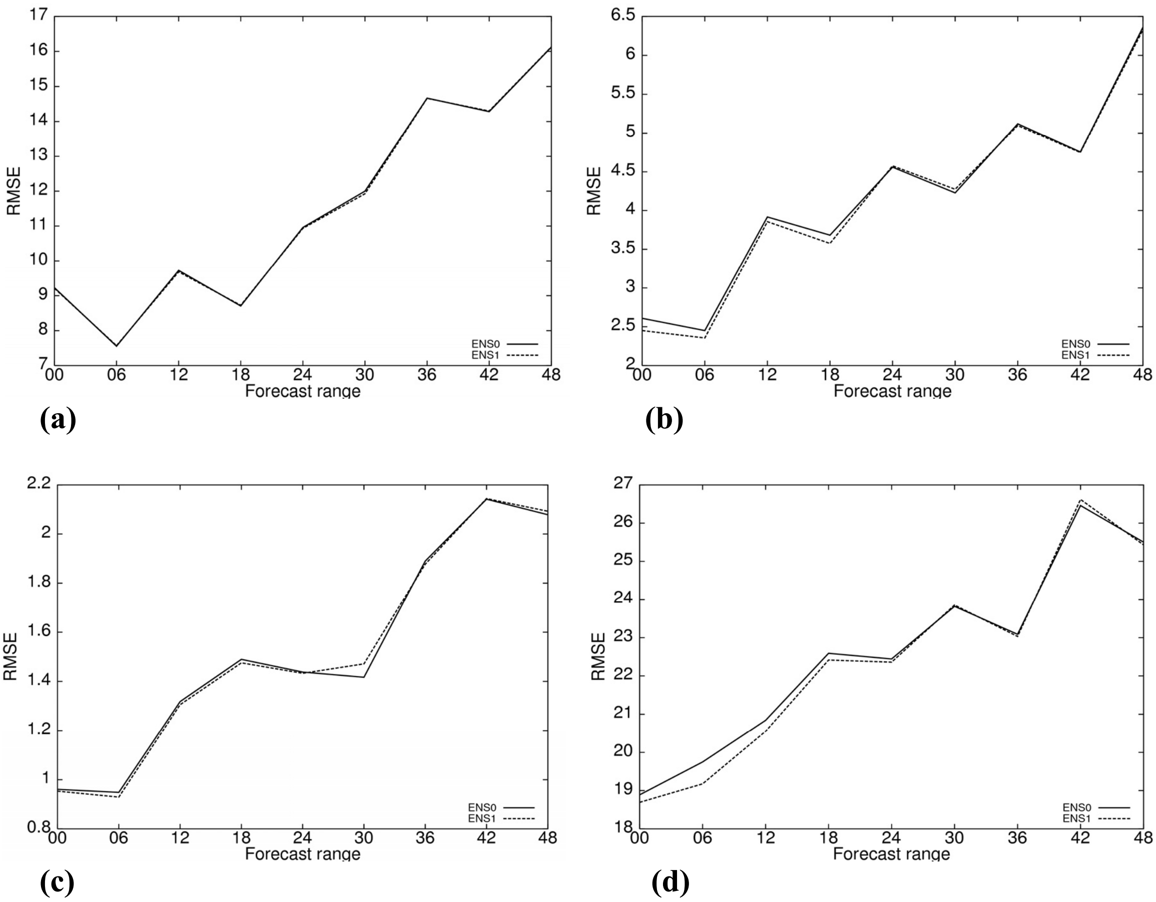

Fig. 1. Verification scores (rmse) computed for (a) geopotential at 500 hPa ($m^2 s^{-2}$), (b) wind speed at 250 hPa ($m s^{-1}$), (c) temperature at 850 hPa (K), (d) relative humidity at 700 hPa (%). Dashed: experiment ENS1, solid: reference ENS0. The scores are computed over the period June 7–20, 2006.

In experiment ENS2, the misfit ratios computed from the autumn period were used. These are r_b=1.33 for the wind and mass variables (vorticity, divergence, surface pressure, and geopotential), r_b=1.43 for temperature, and r_b=2.35 for humidity, as shown in *Table 1*. The tuning had no impact on the 2-meter fields (not shown). Results for the atmospheric variables are displayed in *Fig. 2*. One can see that ENS2 gives clearly worse results than the reference ENS0 for temperature on 850 hPa and also somewhat for relative humidity on 700 hPa. There is a positive impact of the tuning for the wind on 250 hPa.

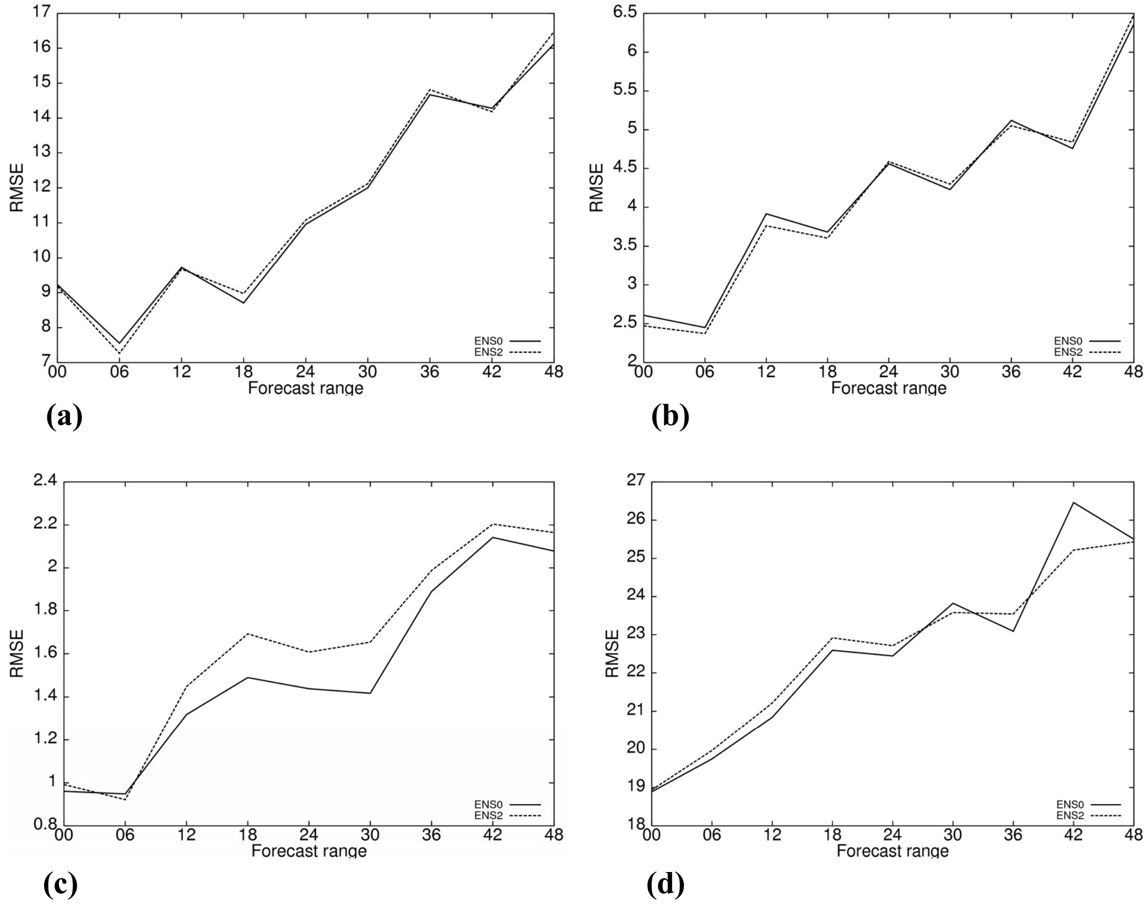

Fig. 2. Verification scores (rmse) computed for (a) geopotential at 500 hPa ($m^2\ s^{-2}$), (b) wind speed at 250 hPa ($m\ s^{-1}$), (c) temperature at 850 hPa (K), (d) relative humidity at 700 hPa (%). Dashed: experiment ENS2, solid: reference ENS0. The scores are computed over the period June 7–20, 2006.

The above results show that the uniform tuning (ENS1) provides better results than the variable dependent one (ENS2), which looks curious, because in theory Eq. (3) holds for each variable separately as well. In *Fig. 3* one can see the predefined and the tuned background error standard deviation profiles used in the experiments. It is obvious that the tuned profiles (ENS1 and ENS2) are quite similar to each other for wind. The difference is larger for temperature, and it is the largest for humidity. This suggests that the disappointing results of experiment ENS2 are mostly due to the large misfit ratios for humidity and temperature. A probable reason for the poorer result of the ENS2 experiment is that the misfit is averaged in the vertical, which can lead to an exaggerated tuning on those levels where the real misfit is drastically smaller than the average.

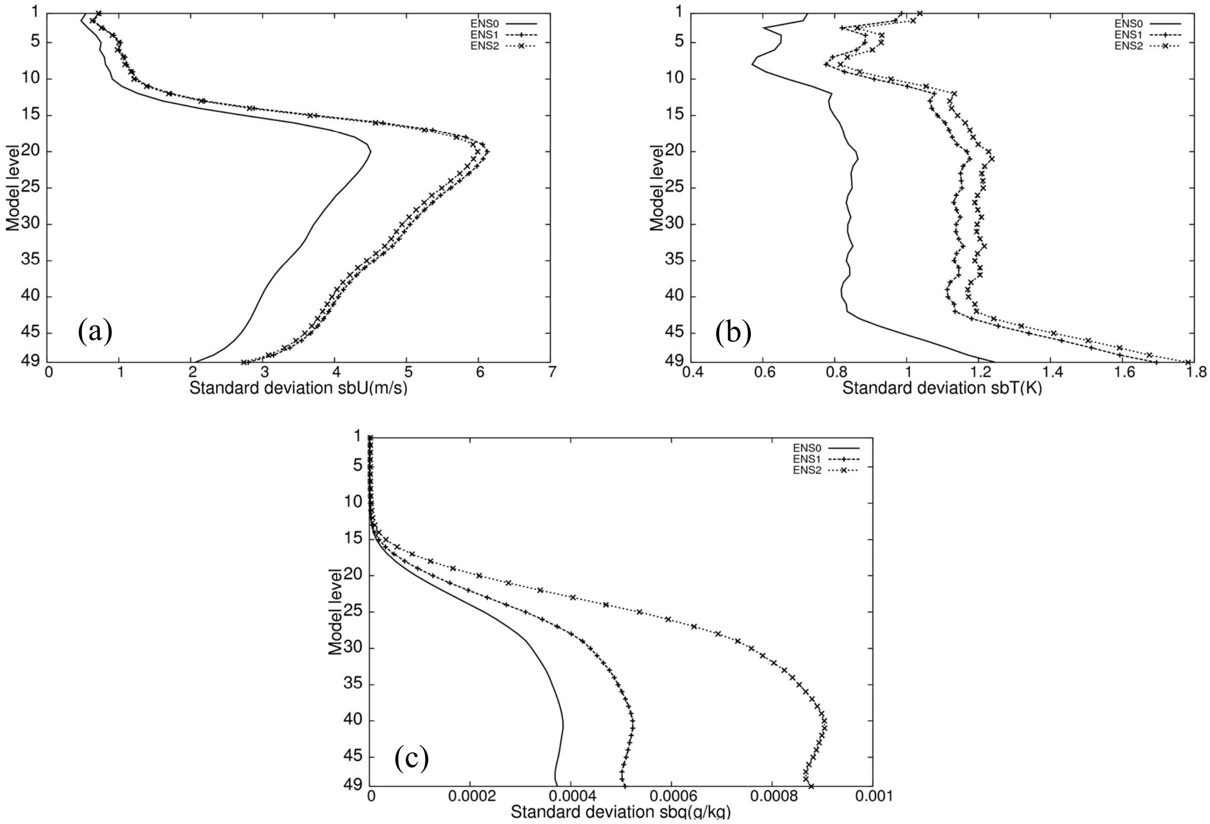

Fig. 3. Predefined (ENS0) and tuned (ENS1 and ENS2) profiles of total background error standard deviations of (a) wind, (b) temperature, (c) specific humidity.

3. Tuning of the multivariate balance

The ALADIN 3d-var is a multivariate analysis system provided by the statistical balance described by *Rabier et al.* (1998), *Courtier et al*. (1998), *Berre* (2000), and *Gustafsson et al.* (2001). In short, this statistical balance couples the analyzed variables in a meteorologically meaningful way by propagating a part of the increments of a given variable to those of another one. This insures the dynamical balance of the variables in the analysis. Experiments done at HMS showed that temperature observations have a large impact on the relative humidity analysis, which results in degradation of the relative humidity analysis and forecast verification scores in some weather situations. On the other hand, the humidity observations influence the temperature analysis in a very limited extent. This asymmetry in the multivariate balance brought us to study the analysis equation in a simplified framework and to propose a tuning of the error variances for a more symmetric balance.

3.1. Theoretical considerations

It can be shown that in a single point model (i.e., $\mathbf{H} = \mathbf{I}$, where $\mathbf{I}$ is the identity matrix), restricting to two variables only (T: temperature and Rh: relative humidity), the analysis Eq. (1) can be written as:

$$\delta T = \frac{\mathrm{cov}\left(\boldsymbol{\varepsilon}_{b,T}\boldsymbol{\varepsilon}_{b,Rh}\right)}{\sigma^2\left(\boldsymbol{\varepsilon}_{b,Rh}\right) + \sigma^2\left(\boldsymbol{\varepsilon}_{o,Rh}\right)}\Delta Rh\,, \tag{7}$$

where δ denotes the analysis increment (analysis minus background) and Δ stands for the observation increment or innovation (observation minus background). Based on *Daley* (1991) and *Hollingsworth* (1987), Eq. (7) can be decomposed as:

$$\delta Rh = \frac{\sigma^2\left(\boldsymbol{\varepsilon}_{b,Rh}\right)}{\sigma^2\left(\boldsymbol{\varepsilon}_{b,Rh}\right) + \sigma^2\left(\boldsymbol{\varepsilon}_{o,Rh}\right)}\Delta Rh \tag{8a}$$

$$\delta T = \frac{\mathrm{cov}\left(\boldsymbol{\varepsilon}_{b,T}\boldsymbol{\varepsilon}_{b,Rh}\right)}{\sigma^2\left(\boldsymbol{\varepsilon}_{b,Rh}\right)}\delta Rh\,, \tag{8b}$$

where (8a) is the univariate or filtering step, which provides humidity analysis increment from the humidity innovation and (8b) is the multivariate propagation step, which transforms the humidity analysis increment to temperature analysis increment. Multiplying and dividing the right-hand side of (8b) with $\sigma_b\left(\boldsymbol{\varepsilon}_T\right)$ and using the definition of the correlation, Eq. (8b) can be rewritten as:

$$\delta T = \mathrm{corr}\left(\boldsymbol{\varepsilon}_{b,T}\boldsymbol{\varepsilon}_{b,Rh}\right)\frac{\sigma\left(\boldsymbol{\varepsilon}_{b,T}\right)}{\sigma\left(\boldsymbol{\varepsilon}_{b,Rh}\right)}\delta Rh\,. \tag{8c}$$

In case of a temperature innovation, one can write the corresponding equations similarly to Eqs. (8a) and (8b) as follows:

$$\delta T = \frac{\sigma^2\left(\boldsymbol{\varepsilon}_{b,T}\right)}{\sigma^2\left(\boldsymbol{\varepsilon}_{b,T}\right) + \sigma^2\left(\boldsymbol{\varepsilon}_{o,T}\right)}\Delta T \tag{9a}$$

$$\delta Rh = \mathrm{corr}\left(\boldsymbol{\varepsilon}_{b,T}\boldsymbol{\varepsilon}_{b,Rh}\right)\frac{\sigma\left(\boldsymbol{\varepsilon}_{b,Rh}\right)}{\sigma\left(\boldsymbol{\varepsilon}_{b,T}\right)}\delta T\,. \tag{9b}$$

The asymmetry in the multivariate propagation step can be quantified as the ratio of T and Rh analysis increments induced by T and Rh innovations:

$$S = \left(\frac{\delta Rh}{\delta T}\right)_{\Delta T} \Big/ \left(\frac{\delta T}{\delta Rh}\right)_{\Delta Rh}\,. \tag{10}$$

The ratio **S** was computed within the ALADIN 3d-var system using single observations ($\Delta T = 1\,K$ and $\Delta Rh = 2\%$), taking the analysis increments right at the origin point (*Fig. 4*).

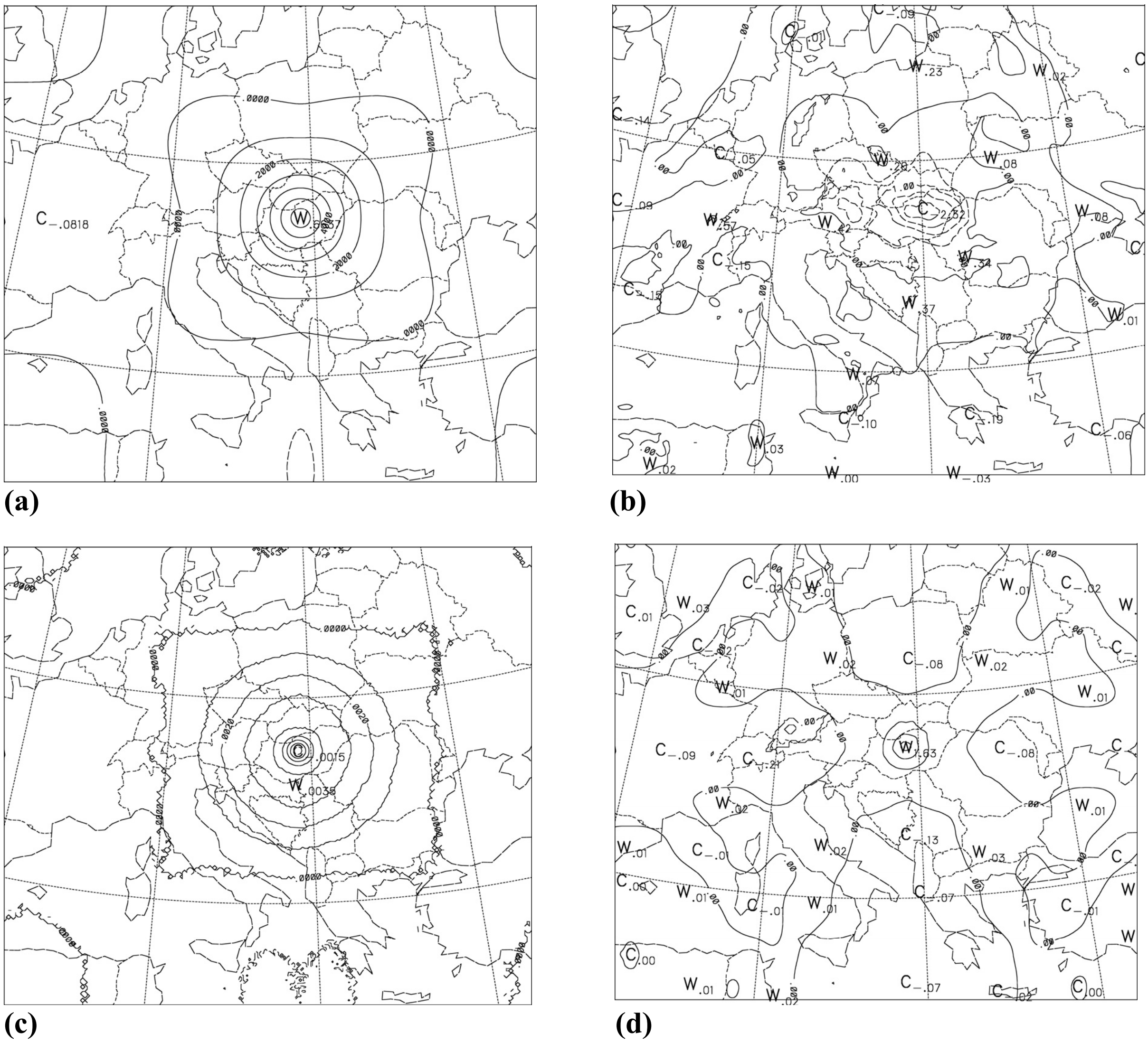

Fig. 4. Analysis increments due to single observations: (a) temperature increment induced by a temperature innovation, (b) relative humidity increment induced by a temperature innovation, (c) temperature increment induced by a relative humidity innovation, (d) relative humidity increment induced by a relative humidity innovation. The innovation values are $\Delta T = 1\,K$ and $\Delta Rh = 2\%$, observation error used are $\sigma(\boldsymbol{\varepsilon}_{o,Rh}) = 0.12\%$, $\sigma(\boldsymbol{\varepsilon}_{o,T}) = 1.1\,K$.

Note that the single observation experiments were run using NMC background error statistics sampled from differences of 36- and 12-hour forecasts of the operational ALADIN model over the period May–July 2004. Substituting the analysis increments shown in *Fig. 4* into Eq. (10), $S \cong 3688$ was found. It is easier to interpret this value if one writes up Eq. (10) using the Eqs. (8c) and (9b), which gives:

$$\mathrm{S} = \left(\frac{\sigma\ (\boldsymbol{\varepsilon}_{\mathrm{b,Rh}})}{\sigma\ (\boldsymbol{\varepsilon}_{\mathrm{b,T}})} \right)^2 \cong 3688\,. \tag{11}$$

This shows that the symmetry of the multivariate coupling depends only on the variance ratios of the two variables. Taking the square root of the above equation one gets:

$$\sqrt{\mathrm{S}} = \frac{\sigma\ (\boldsymbol{\varepsilon}_{\mathrm{b,Rh}})}{\sigma\ (\boldsymbol{\varepsilon}_{\mathrm{b,T}})} \cong 60 \frac{\%}{\mathrm{K}}\,. \tag{12}$$

According to the above equation, a 1 K temperature error is associated to a 60% error in relative humidity and vice versa, which points out the rather asymmetric behavior of the multivariate propagation steps taking into account typical changes of temperature and humidity in the troposphere. The proposal for the tuning is thus to decrease the above ratio of standard deviations, which can be done either by increasing $\sigma(\boldsymbol{\varepsilon}_{\mathrm{b,T}})$ or by decreasing $\sigma(\boldsymbol{\varepsilon}_{\mathrm{b,Rh}})$. Examining the univariate steps in case of the T (*Fig. 4a*) and Rh innovations (*Fig. 4d*), one can see that the Rh analysis increment is about the 80% of the Rh innovation, while the T analysis increment is about the 60% of the T innovation. This means that the trust in the humidity background is indeed very small, which was the basis for assuming that $\sigma(\boldsymbol{\varepsilon}_{\mathrm{b,T}})$ is well chosen and $\sigma(\boldsymbol{\varepsilon}_{\mathrm{b,Rh}})$ is overestimated. As a consequence, it was decided to run tuning experiments by reducing the humidity background error standard deviations. An additional remark is that beside reducing humidity background error standard deviations, we considered to reduce also the humidity observation error standard deviations in order to keep Eq. (8a) univariate filtering step unchanged.

It should be clarified here, that the ALADIN 3d-var uses specific humidity (q) instead of relative humidity as analysis control variable, which also implies that the $\sigma(\boldsymbol{\varepsilon}_{\mathrm{b,q}})$ specific humidity background error standard deviations were chosen for tuning instead of $\sigma(\boldsymbol{\varepsilon}_{\mathrm{b,Rh}})$. The reason for using Rh in the above elaboration was that the asymmetry of the multivariate balance was observed first in relative humidity scores, and also the reader might judge relative humidity changes in the atmosphere easier than those of specific humidity.

Another feature that requires explanation here is that Rh single observation increments in *Figs. 4b* and *4d* show some obvious anisotropy in spite of the fact that the ALADIN horizontal background error structure is designed to be isotropic by origin (*Berre*, 2000). The clue for this seeming contradiction is that the anisotropy is not included by the structure functions directly but rather by the computation Rh from q. Namely, to derive and plot relative humidity increments, first the $q \rightarrow Rh$ computation is performed both on the background and analysis fields, and then their difference is taken. As the $q \rightarrow Rh$ computation is a non-linear function of specific humidity and temperature, the resulting relative humidity increment will not keep the isotropic structure even if both specific humidity and temperature increments alone are isotropic.

3.2. The method of tuning

The extent of reduction of humidity errors was determined on the basis of an error estimation proposed by *Hollingsworth* and *Lönnberg* (1986) (HL method), which is independent from the NMC method. In short, the HL method estimates the observation and background errors based on the first equation of Eq. (2a), where the innovations are composed by the difference of radiosonde observations and model forecasts. By sorting the innovations in vertical and horizontal distance categories, one can estimate a vertical profile of observation and background error standard deviations and horizontal correlation length scales. This independent method was used as a guideline for estimating new background and observation standard deviations for humidity within our analysis system. The HL method was applied for the same period as the one used for the NMC sampling (May 2–August 2, 2004). The HL standard deviations were first computed on the standard levels of the radiosonde observations, and then they were interpolated on the model levels for the comparison with the NMC statistics.

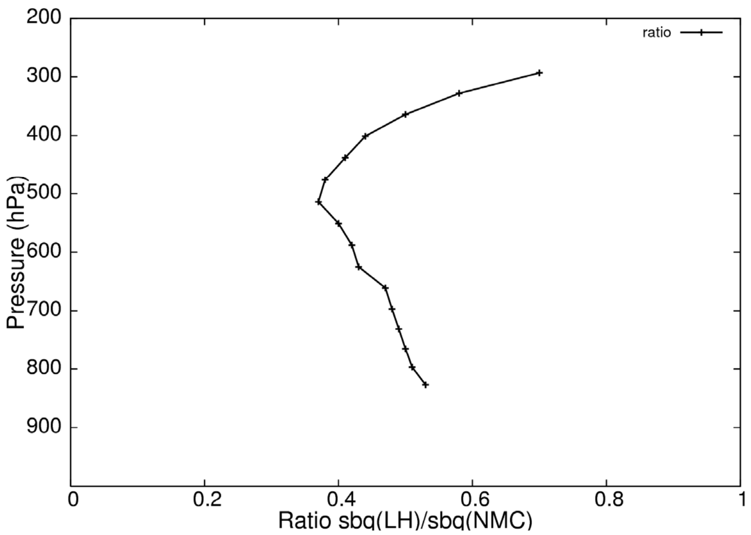

Fig. 5. Vertical profile of specific humidity background error standard deviation ratios obtained by the HL and NMC methods.

In *Fig. 5* the ratio of HL and NMC background error standard deviations are shown for specific humidity. Levels under 850 hPa were not considered for the tuning due to the small amount of radiosonde observations. Despite the strong vertical variability, the ratio is smaller than 1 for all the vertical levels, which means that the HL estimation of $\sigma(\boldsymbol{\varepsilon}_{b,q})$ is smaller than the NMC one in the whole troposphere. This is in accordance with the notion of $\sigma(\boldsymbol{\varepsilon}_{b,q})$ being overestimated in the NMC statistics, which was also suggested by the theoretical considerations described earlier. Thus, it was decided to tune the $\sigma(\boldsymbol{\varepsilon}_{b,q})$ and $\sigma(\boldsymbol{\varepsilon}_{o,q})$ profiles used in the ALADIN 3d-var experiments according to the ratio profile shown in *Fig. 5*. Additionally, in some of the experiments $\sigma(\boldsymbol{\varepsilon}_{b,q})$ was drastically reduced (multiplied by a factor of 0.005) above 250 hPa in order to prevent to propagate high altitude humidity increments to the lower troposphere. This latter modification was based on the experiences at ECMWF (*Andersson et al.*, 1998).

3.3. The tuning experiments

The tuning experiments were run with the following three settings:

1. EX1Q: modified $\sigma(\boldsymbol{\varepsilon}_{b,q})$ between 850 and 250 hPa according to the HL/NMC ratio, reduced $\sigma(\boldsymbol{\varepsilon}_{b,q})$ above 250 hPa.
2. EX2Q: same as in EX1Q but also $\sigma(\boldsymbol{\varepsilon}_{o,q})$ is reduced in the same heights and in the same extent as $\sigma(\boldsymbol{\varepsilon}_{b,q})$.
3. EX3Q: same as EX2Q except that there was no tuning applied above 250 hPa.
4. REFQ: reference run using the predefined $\sigma(\boldsymbol{\varepsilon}_{o,q})$ and $\sigma(\boldsymbol{\varepsilon}_{b,q})$ values.

As a first step, the computation of Eq. (10) was repeated with the tuned error standard deviations based on single observation experiments again. It was proven that the tuning indeed reduced the asymmetry in the expected way as $\sqrt{S} = 35\%/K$ was found with the EX1Q and $\sqrt{S} = 33\%/K$ was obtained with the EX2Q settings. As a second step, complete assimilation cycles were run for a two-week period (August 29–September 10, 2004) with all the above experimental settings in order to see the impact of the tuning in a real operational-like context. The assimilation cycles were run with a 6-hour frequency, using surface, radiosonde, aircraft, and satellite (ATOVS AMSU-A and B) observations. Production forecasts (up to 48 hours) were run from the 00 UTC analyses. In general, the tuning had a smaller impact in these real cycling experiments than expected in terms of verification scores, which was surprising after the firm effect in the single observation experiments. This can be explained by the fact that the single observation experiments are much closer to the simplified system Eqs. (7)–(10) than the complex assimilation cycling experiments, where also the spatial correlations may play an important role. The largest impact was found for the experiment EX2Q, which will be referred in the further comparison against the reference run REFQ. In *Figs. 6* and *7*, bias and rmse scores for the analyses are shown for the EX2Q and REFQ runs. The tuning resulted in a better (worse) fit to the observations concerning the temperature (humidity) analysis rmse. The impact on the wind analysis is rather mixed.

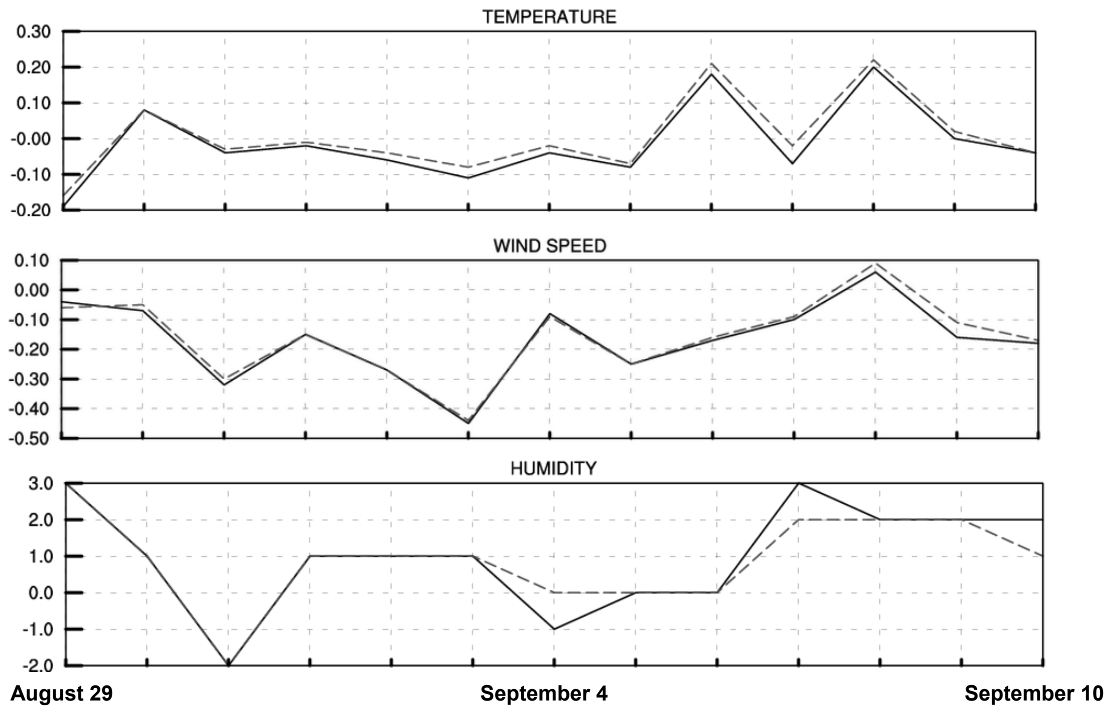

Fig. 6. Evolution of analysis bias scores for the experiments EX2Q (solid line) and the reference REFQ (dashed line). From top to bottom: temperature at 850 hPa (K), wind speed at 500 hPa (m s^{-1}), relative humidity at 700 hPa (%). The scores are computed over the period August 29 – September 10, 2004.

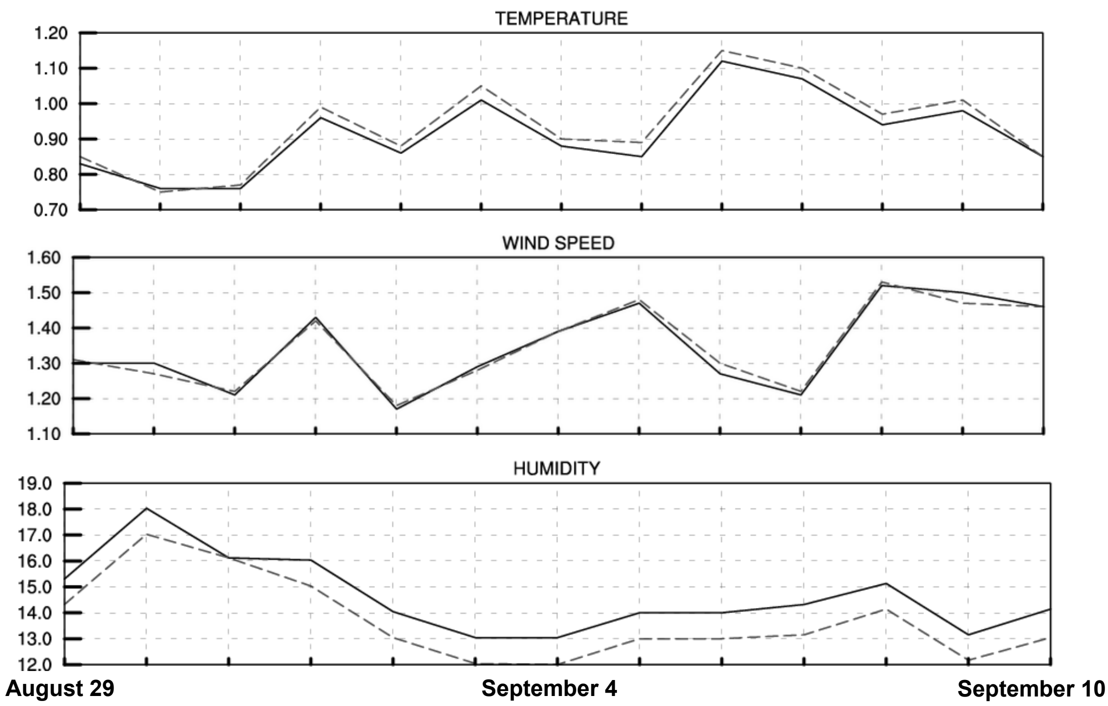

Fig. 7. Evolution of analysis rmse scores……→

In our explanation the good fit of temperature comes indeed from the improved multivariate analysis, namely from a larger multivariate propagation of the humidity increments into those of temperature through the process described by Eq. (8b). On the other hand, the increased analysis departure of humidity may have two reasons: (i) multivariate humidity increments induced by those of temperature and wind became too small by excessively decreasing $\sigma(\boldsymbol{\varepsilon}_{b,q})$; and (ii) the univariate step of the humidity analysis Eq. (8a) was degraded by reducing $\sigma(\boldsymbol{\varepsilon}_{b,q})$, in spite of our compensation by reducing $\sigma(\boldsymbol{\varepsilon}_{o,q})$ at the same time.

Figs. 8 and *9* show that the positive impact on temperature decreases quickly with the forecast range. Nevertheless, humidity 6-hour forecast scores are improved by the performed tuning both in terms of bias and rmse. The impact on wind is kept mixed for this forecast range, while the overall effect of the tuning on longer forecast ranges is rather neutral (not shown).

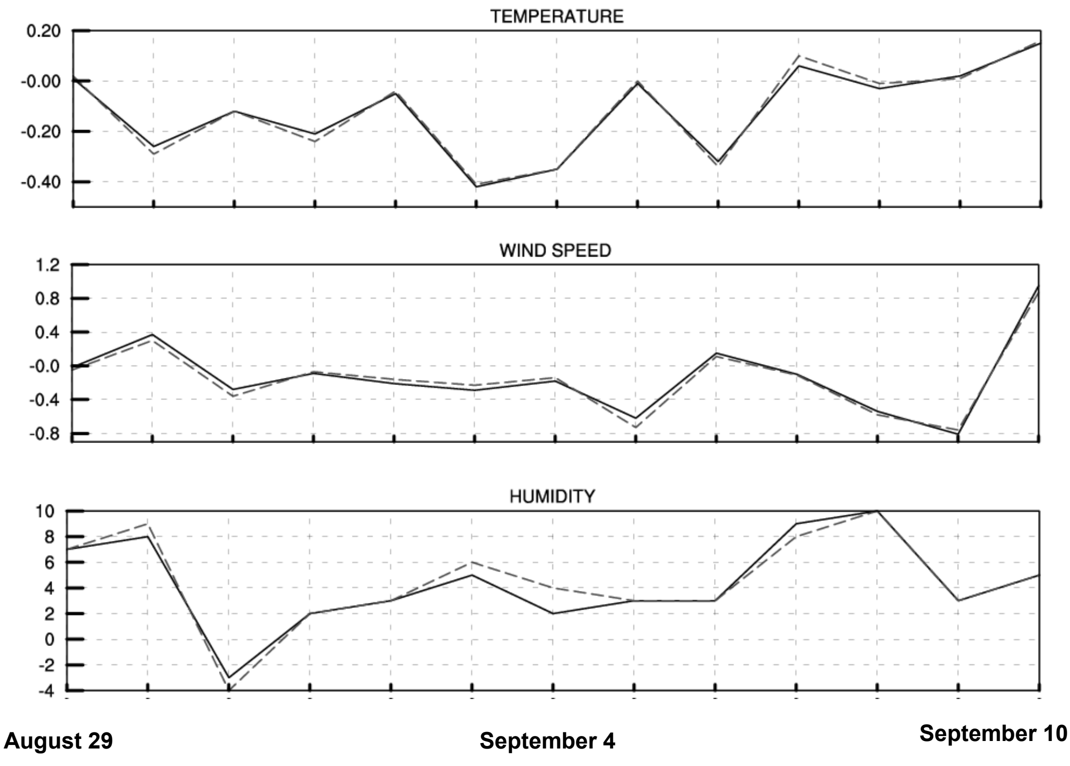

Fig. 8. Evolution of 6-hour forecast bias scores for the experiments EX2Q (solid line) and the reference (dashed line). From top to bottom: temperature at 850 hPa (K), wind speed at 500 hPa (m s^{-1}), relative humidity at 700 hPa (%). The scores are computed over the period August 29 – September 10, 2004.

Fig. 7. Evolution of analysis rmse scores for the experiments EX2Q (solid line) and the reference REFQ (dashed line). From top to bottom: temperature at 850 hPa (K), wind speed at 500 hPa (m s^{-1}), relative humidity at 700 hPa (%). The scores are computed over the period August 29 – September 10, 2004.

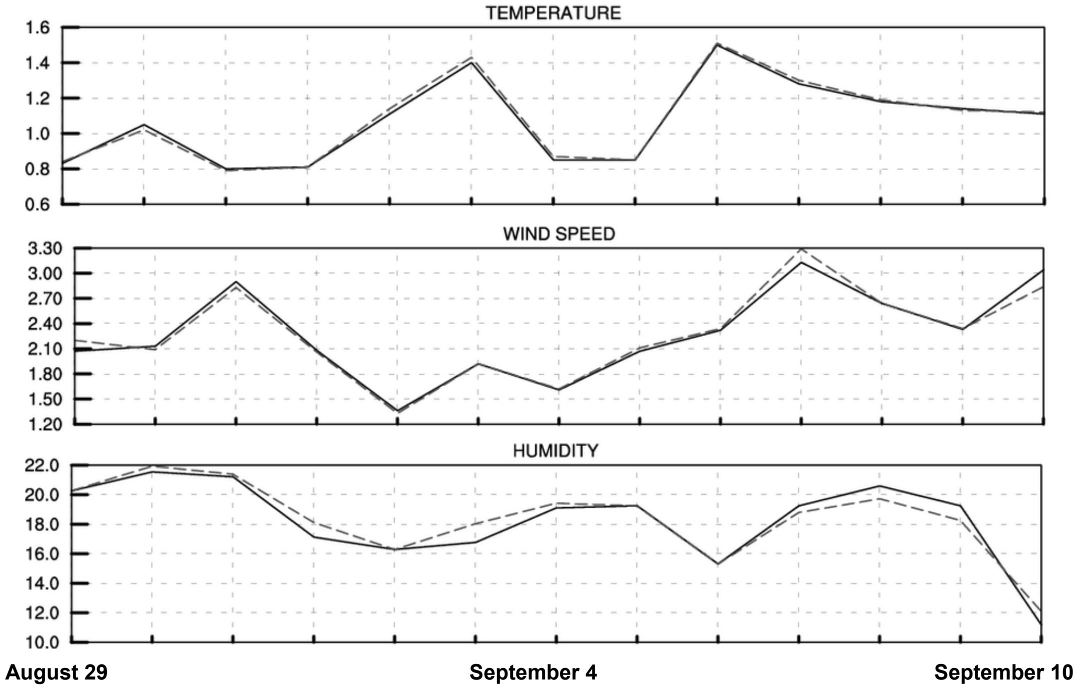

Fig. 9. Evolution of 6-hour forecast rmse scores for the experiments EX2Q (solid line) and the reference (dashed line). From top to bottom: temperature at 850 hPa (K), wind speed at 500 hPa (m s^{-1}), relative humidity at 700 hPa (%). The scores are computed over the period August 29 –September 10, 2004.

4. Discussion

Two attempts to improve the Hungarian version of the ALADIN 3d-var analysis system were described. Both of them involve the tuning of the background and observation error standard deviations "a posteriori", but on a different theoretical basis.

The first tuning shown is based on the statistical optimality criteria derived by *Désroziers et al.* (2006) and aims to tune the errors for the full set of the background variables. The tuning was applied both in a variable dependent way and also as a uniform tuning for all the variables. A clear positive impact of the uniform tuning was found, while in the variable dependent case the verification showed a degradation of the forecasts. This seems to be a contradiction, as the more complex tuning had a worse impact than the simpler one. A probable explanation of this feature is that no vertical dependence of the tuning was included in any of the experiments, but a vertically averaged tuning ratio was used instead, being a considerable limitation of the approach. It is assumed that the degrading effect of the vertical averaging appears more intensively in the variable dependent tuning if the background error of the given variable shows a large vertical dependence (i.e., in the case of temperature and humidity). The uniform tuning has been introduced in the operational ALADIN 3d-var data assimilation system of HMS after the experiments.

The second tuning shown is based on the simplified BLUE analysis equations, proposing a modification only for the humidity errors. On the basis of single observation experiments, it was shown that the predefined error statistics produce an asymmetric multivariate balance between humidity and the other variables. Namely, humidity increments due to wind and temperature innovations are excessive compared to temperature and wind increments due to humidity innovations. On the basis of the simplified BLUE Eqs. (7) – (10), it was derived that the asymmetry can be decreased by reducing the background error standard deviations of humidity, which was demonstrated in a set of single observation experiments. To estimate the decrease in the humidity background error standard deviations the HL method was used (*Hollingsworth* and *Lönnberg*, 1986), which suggested a decrease of approximately half of the predefined standard deviation values on average. Full-observation experiments were performed with the modified multivariate propagation step, as well as experiments where the observation error standard deviations of humidity were also decreased to the same extent in order to keep the univariate analysis step unchanged. In real assimilation cycling experiments, the overall impact of the tuning was found to be quite moderate. This was probably caused by additional effects of spatial correlations of background errors, which were not taken into account in the simplified equations used as the basis of the tuning. However, a positive impact on 6-hour humidity forecasts was found with the application of the tuning procedure, suggesting that methodology and humidity tuning performed have a potential to improve humidity scores of numerical weather prediction models, at least for the shorter forecast ranges.

Note that the two tuning approaches described in the paper were performed independently from each other, mainly because they were motivated by independent problems. However, the two approaches are surely in a strong interaction, as both of them account for the modification of background error standard deviations. One can notice that the two tuning approaches propose to modify the humidity background error standard deviations in the opposite direction. This probably comes from the important fact that the tuning by residual covariances (Chapter 2) was based on predefined statistics sampled with the Ensemble method, while in the tuning of the multivariate balance (Chapter 3), the predefined statistics were sampled by the NMC method. Namely, background error standard deviations sampled by the Ensemble method are found to be smaller than those sampled by the NMC method (*Belo Pereira* and *Berre*, 2006; *Stefanescu et al.*, 2006), which may explain the seemingly opposing guidance on humidity tuning by the two tuning experiments. The change in the sampling method was due to better performance of the ALADIN 3d-var system of HMS using the Ensemble method, which entailed also its operational use. It is, thus, desirable to repeat the tuning of the multivariate balance within the Ensemble sampling framework in the future and to see its interaction with the tuning based on residuals covariances.

Acknowledgements**—**The authors would like to thank first the great help and encouragement given by *Loïk Berre* during the work related to the tuning of the multivariate balance. We also appreciate very much the advices of *Roger Randriamampianina* and *Claude Fischer* in several issues during our experimentations. The work of *Helga Tóth* is appreciated concerning the implementation of the HL diagnostic tool. We are very grateful for *Edit Adamcsek* for her useful suggestions regarding the text and formulas. This work was supported by the Hungarian National Scientific Foundation (OTKA T049579) and Ministry of Science, Technology and Sports of Republic of Croatia (MZOS project 0004001), as well as scholarship grants awarded by the Regional cooperation for Limited Area modeling in Central Europe and Hungarian Scholarship Board.

References

Andersson, E., Haseler, J., Undén, P., Courtier, P., Kelly, G., Vasiljevic, D., Brankovic, C., Gaffard, C., Hollingsworth, A., Jakob, C., Janssen, P., Klinker, E., Lanzinger, A., Miller, M., Rabier, F., Simmons, A., Strauss, B., Viterbo, P., Cardinali, C., Thépaut, *J-N.,* 1998: The ECMWF implementation of three-dimensional variational assimilation (3D-Var). III : Experimental results. *Q J Roy Meteor Soc 124*, 1831-1860.

Belo Pereira, M. and *Berre, L.,* 2005: The use of an ensemble approach to study the background error covariances in a global NWP model. *Mon Weather Rev 134*, 2466-2489.

Berre, L., 2000: Estimation of synoptic and meso scale forecast error covariances in a limited are amodel. *Mon Weather Rev 128*, 644-667.

Bouttier, F. and *Courtier, P.*, 1999: Data assimilation concepts and methods. *ECMWF Lecture Series (Data assimilation and use of satellite data http://www.ecmwf.int/newsevents/training/lecture_ notes/pdf_files/ASSIM/Ass_cons.pdf)*, 72.

Bölöni, G., 2006: Development of a variational data assimilation system for a limited area model at the Hungarian Meteorological Service. *Időjárás 110*, 309-327.

Chapnik, B., Désroziers, G., Rabier, F., Talagrand, O., 2004: Properties and first application of an erro- statisticstuning method in variational assimilation. *Q J Roy Meteor Soc 130*, 2253-2275.

Courtier, P., Andersson, E., Heckley, W., Pailleux, J., Vasiljevic, D., Hamrud, M., Hollingsworth, A., Rabier, F., and *Fisher, M.,* 1998: The ECMWF implementation of three dimensional variational assimilation (3D-Var). Part I: Formulation. *Q J Roy Meteor Soc 124*, 1783-1808.

Daley, R., 1991: *Atmospheric Data Analysis.* Cambridge University Press, Cambridge, 457 pp.

Désroziers, G. and *Ivanov, S.,* 2001: Diagnosis of adaptive tuning of information error parameters in a variational assimilation. *Q J Roy Meteor Soc 127*, 1433-1452.

Désroziers, G., Berre, L., Chapnik, B. and *Poli, P.,* 2006: Diagnosis of observation, background and analysis error statistics in observation space. *Q J Roy Meteor Soc 131*, 3385-3396.

Fischer, C., Montmerle, T., Berre, L., Auger, L., Stefanescu, S., 2005: An overview of the variational assimilation in the Aladin/France numerical weather prediction system. *Tellus 131*, 3477-3492.

Fisher, M., 2003: Background error covariance modelling. *ECMWF Seminar on Recent Developments in Data Assimilation for Atmosphere and Ocean*, 45-63.

Gandin, L.S., 1963: *Objective Analysis of Meteorological Fields.* Gidrometeorologicheskoe Izdatelstvo, Leningrad. English translation by Israeli Program for Scientific Translations, Jerusalem, 1965.

Gustaffson, N., Berre, L., Hörnquist, S., Huang, X.-Y., Lindskog, M., Navascués, B., Mogensen, K.S., and *Thorsteinnson, S.,* 2001: Three-dimensional variational data assimilation for a limited area model. Part I: General formulation and the background error constraint. *Tellus 53A*, 425-446.

Hollingsworth, A., 1987: Objective analysis for numerical weather prediction. Short and medium-range numerical weather prediction. *Proc. of the WMO/IUGG NWP Symp., Tokyo.* Meteorological Society of Japan, 11-59.

Hollingsworth, A. and *Lönnberg, P.*, 1986: The statistical structure of short-range forecast errors as determined from radiosonde data. Part I. The wind field. *Tellus 38A*, 111-136.

Horányi, A., Ihász, I. and *Radnóti, G.*, 1996: ARPEGE/ALADIN: A numerical weather prediction model for Central-Europe with the participation of the Hungarian Meteorological Service. *Időjárás 100*, 277-301.

Lorenc, A.C., 1986: Analysis methods for numerical weather prediction. *Q J Roy Meteor Soc 114*, 205-240.
Lönnberg, P. and *Hollingsworth, A.,* 1986: The statistical structure of short-range forecast errors as determined from radiosonde data. Part II: The covariance of height and wind errors. *Tellus 38A*, 137-161.
Parrish, D.F. and *Derber, J.C.*, 1992: The National Meteorological Centre's spectral statistical interpolation system. *Mon Wea Rev,* 1747-1763.
Rabier, F., Andersson, E., McNally, A., Courtier, P., Undén, P., Eyre, J., Hollingsworth, A., Bouttier, F., 1998: The ECMWF implementation of three-dimensional variational assimilation (3D-Var). II: Structure functions. *Q J Roy Meteor Soc 124*, 1809-1829.
Randriamampianina, R., 2006: Impact of high resolution observations in the ALADIN/HU model. *Időjárás 110*, 329-349.
Sadiki, W. and *Fischer, C.,* 2005: A posteriori validation applied to the 3D-VAR ARPEGE and ALADIN data assimilation system. *Tellus* 57A, 21-34.
Siroká, M., Fischer, C., Cassé, V., Brozková, R. and *Geleyn, J-F.,* 2003: The definition of mesoscale selective forecast error covariances for a limited area variational analysis. *Meteor Atmos Phys* 82, 227-244.
Stefanescu, S.E., Berre, L., and *Belo Pereira, M.,* 2006: The evolution of dispersion spectra and the evluation of model differences in an ensemble estimation of error statistics for a limited area analysis. *Mon Wea Rev 134*, 3456-3478.
Talagrand, O., 1998: A posteriori evaluation and verification of analysis and assimilation algorithms. *Workshop on Diagnosis of Data Assimilation Systems,* 2-4 November 1998, ECMWF, 17-18.
Thépaut, J.-N. and *Courtier, P.,* 1991: Four-dimensional data assimilation using the adjoint of a multilevel primitive equation model. *Q J Roy Meteor Soc 117*, 1225-1254.
Wilks, D.S., 1995: *Statistical Methods in the Atmospheric Sciences*. Academic Press, San Diego, California, 467 pp.