Data Assimilation at CHMI

progress report from September 2008

Alena Trojáková

Czech Hydrometeorological Institute
Na Šabatce 17, 143 06 Prague 4

Contents

1 Introduction

This report summarizes the first experiments with data assimilation performed at CHMI during the second and the third quarters of 2008. Based on Široká (2001) the first trial consists of adding 3DVAR analysis after upper-air blending procedure, so called BlendVar configuration. Concerning observation an impact of SYNOP and TEMP was studied so far.

2 First evaluation of the assimilation system

For the time being a single three-weeks period was studied. The first four days (March 26th - 29th 2008) were taken as warming up period and all evaluation were performed on period from 20080330 till 20080419. All experiments started from the same first guess at 20080325 00 UTC, which was computed by operational version of ALADIN/CE. Reference experiment (in more details described later) was based on the ALADIN/CE operational setting from 20080318. To mimic operational use both long cut-off analysis (for stand-alone assimilation cycle) and short cut-off analysis (for production) were kept in experimental settings.

For verifications 48 hours forecast from 00 UTC short cut-off analysis were provided for period from 20080330 till 20080419. The objective scores were evaluated by verification package VERAL. The BIAS, RMSE and STDE were computed from differences between the forecast and observation (SYNOP and TEMP). The significance tests described in Fisher (2001) of the objective verification scores were performed as well.

2.1 Experimental setting

Reference experiment was based on the ALADIN/CE operational setting from 20080318 (former parallel test AHO "ALAROminus3MT") with following general characteristics:

- cycle 32T1
- linear truncation E159x143, mean orography, 9km horizontal resolution and 43 vertical levels
- domain covers the same area as the formal LACE domain (309x277 grid points)
- 3h coupling interval, time step 360 s
- surface analysis (performed before upper-air one) is provided:
  - by CANARI surface analysis based on SYNOP reports for land
  - SST is taken from ARPEGE analysis
  - any other land soil variables which are not analyzed (like snow) are initialized from the ALADIN guess with the relaxation to the climatology as implemented within the CANARI configuration
- upper air analysis is provided:
  - by the digital filter spectral blending, long cut-off 6h cycle
  (filtering at truncation E47x42, no DFI in the next +6h guess integration)
  - digital filter spectral blending + incremental DFI initialization of short cut-off production analysis

Above mentioned ALADIN/CE setting will be referred as reference (or **Z02** experiment).

BlendVar configuration consists of adding 3DVAR just after blending and used 3DVAR has settings:

- B matrix was computed following the lagged NMC method (from period of October - December 2006)
- REDNMC=1
- SYNOP surface reports (geopotential used only)
- TEMP upper air reports (temperature, wind, specific humidity used only)

2.2 Analysis behavior of BlendVar

Overview of number of used observation and overall bias of given parameter can be obtained from time series on Fig 1. In order to evaluate analysis behavior of TEMP observation we have checked vertical profiles of monitoring statistics (mean and standard deviation). Obs-analysis statistics are mostly smaller than obs-guess ones, which means that analysis uses information coming from the TEMP data, see Fig 2.

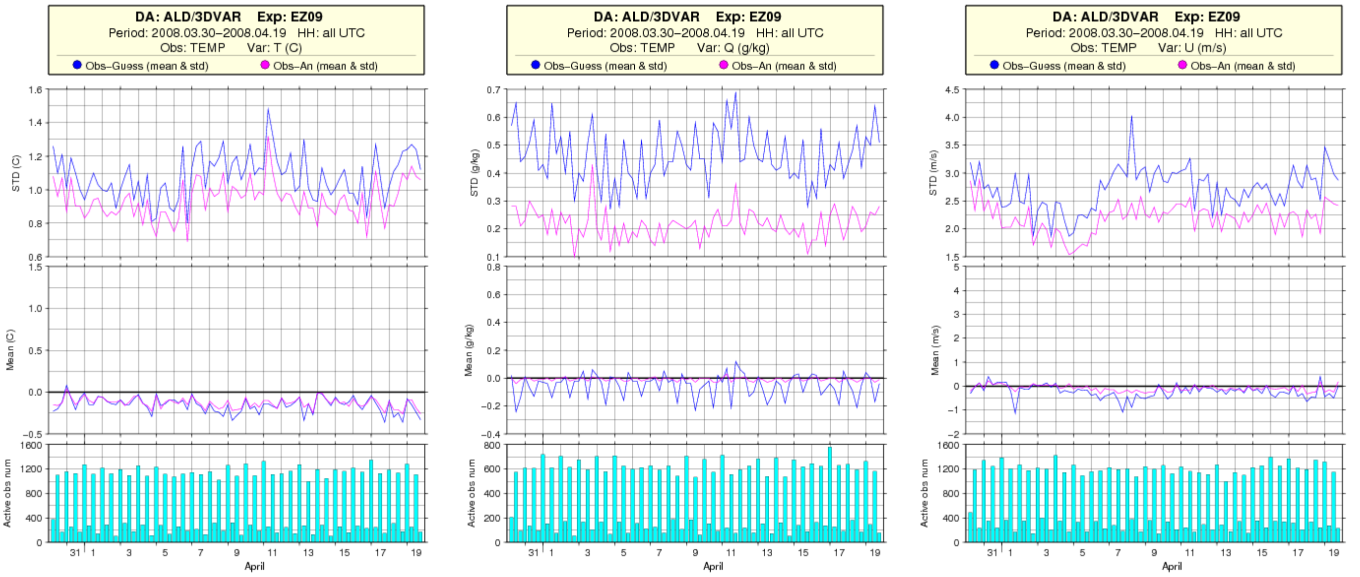

Fig. 1: Monitoring statistics, standard deviation, mean and number of active observation of TEMP observation: T (left), q (middle), u (right) from 20080330 till 20080419 of all (00,06,12,18) runs.

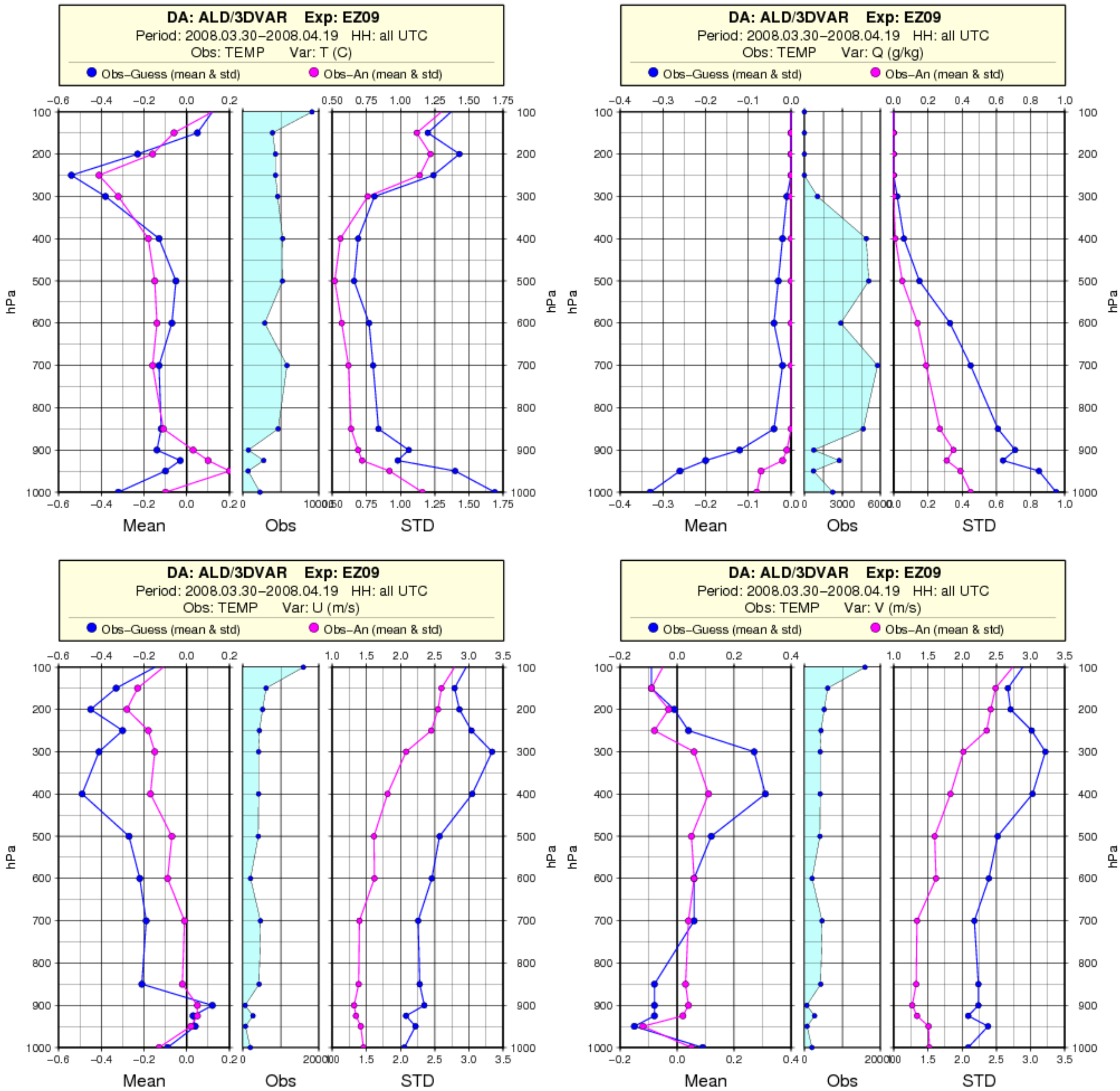

Fig. 2: Monitoring statistics, mean and standard deviation of T (upper-left), q (upper-right), u (bottom-left) and v (bottom-right) from 20080330 till 20080419 of all (00,06,12,18) runs.

2.3 Verification against observation

On following figure VERAL scores (against SYNOP and TEMP observation) are plotted. There is apparent small positive impact of BlendVar for the first +6H at least for most of the variables.

- **Z02** - reference (Blending)
- **Z09** - BlendVar

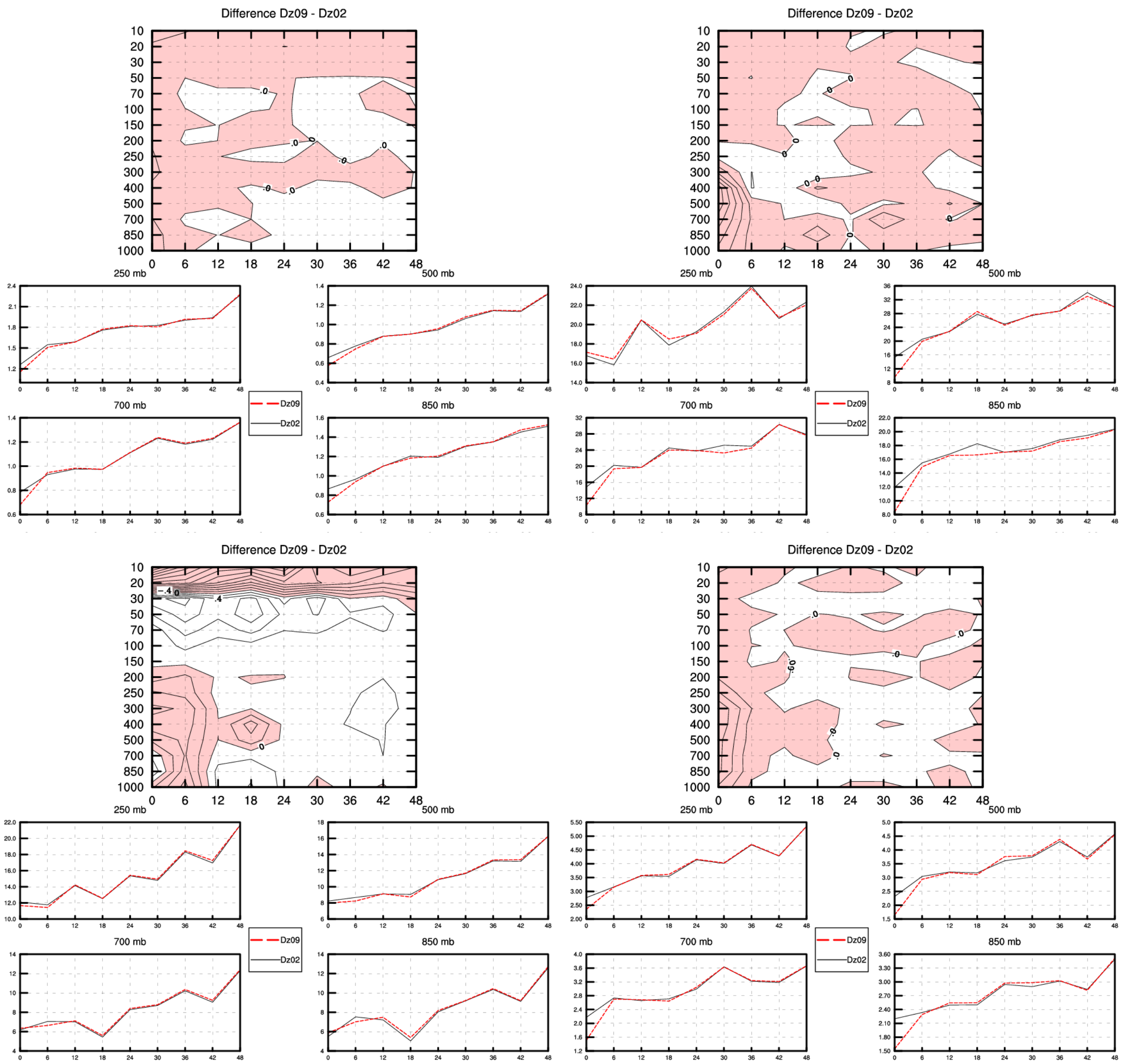

Fig. 3: RMSE of T (upper-left), RH (upper-right), ϕ (bottom-left) and wind speed (bottom-right) from 20080330 till 20080419 of 00 UTC runs. Red areas denote positive impact of BlendVar usage, isolines every 0.1 K, 1%, 0.2 m^2s^{-2} and 0.2 m/s.

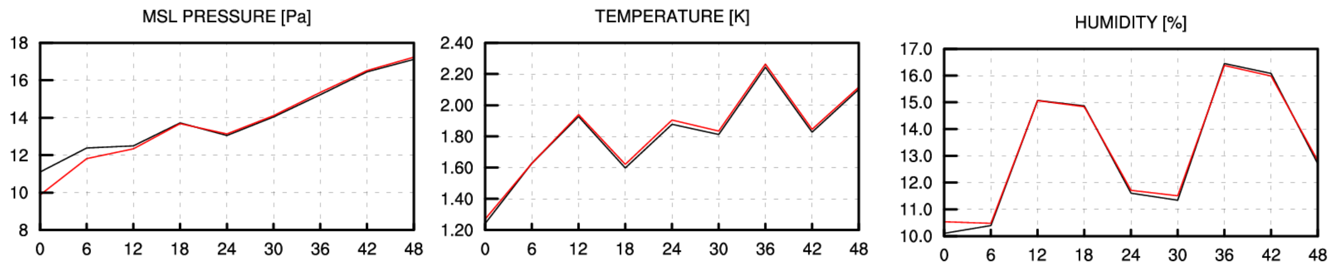

Fig. 4: RMSE of $MSLP$ (left), $T2m$ (middle) and $RH2m$ (right) from 20080330 till 20080419 of 00 UTC runs. BlendVar in red and reference experiment in black color.

Significance test proved mostly improvement, but pointed some degradation mainly near the surface, for complete list see Table 1.

Parameter	Forecast	Significance at 90%	Parameter	Forecast	Significance at 90%
MSLP	+00H	**better**	RH 850 hPa	+36H	**better**
	+06H	**better**	T 850 hPa	+00H	**better**
	+12H	**better**	V 850 hPa	+00H	**better**
RH2m	+00H	worse	ϕ 700 hPa	+06H	**better**
	+06H	worse	RH 700 hPa	+00H	**better**
	+30H	worse		+30H	**better**
T2m	+00H	worse	T 700 hPa	+00H	**better**
	+18H	worse	V 700 hPa	+00H	**better**
	+24H	worse	ϕ 500 hPa	+00H	**better**
	+30H	worse		+06H	**better**
V10m	+00H	worse	RH 500 hPa	+00H	**better**
	+12H	worse	T 500 hPa	+00H	**better**
ϕ 1000 hPa	+06H	**better**	V 500 hPa	+00H	**better**
	+12H	worse		+24H	worse
	+18H	worse	ϕ 400 hPa	+00H	**better**
RH 1000h Pa	+00H	**better**	RH 400 hPa	+00H	**better**
T 1000h Pa	+00H	**better**		+48H	worse
	+12H	worse	T 400 hPa	+00H	**better**
	+36H	worse		+06H	**better**
V 1000h Pa	+00H	**better**	V 400 hPa	+00H	**better**
ϕ 850h Pa	+00H	worse	ϕ 300 hPa	+00H	**better**
	+06H	**better**	RH 300 hPa	+00H	**better**
	+12H	worse	T 300 hPa	+00H	**better**
	+18H	worse		+36H	**better**
RH 850h Pa	+00H	**better**	V 300 hPa	+00H	**better**
	+18H	**better**	ϕ 250 hPa	+00H	**better**
			T 250 hPa	+00H	**better**
			V 250 hPa	+00H	**better**

Table 1: List of parameters and forecast ranges where Z09 performs better/worse than Z02 in terms of RMSE scores against observation with significance 90 % two side confidence interval significance test.

2.4 Summary

Overall scores are very encouraging, worrying is small but still significant degradation of some near the surface parameters.

3 Tuning of error statistics

On the basis of linear estimation theory Desroziers *et al* (2006) proposed simple diagnostics which should be fulfilled in an optimal analysis. For any subset of observations i with p_i observations one can compute diagnosed value of observation and background error

$$(\sigma_i^o)^2 = \sum_{j=1}^{p_i} \frac{(y_j^o - y_j^a)(y_j^o - y_j^b)}{p_i}$$

$$(\sigma_i^b)^2 = \sum_{j=1}^{p_i} \frac{(y_j^a - y_j^b)(y_j^o - y_j^b)}{p_i}$$

An advantage of proposed method is that the diagnostic is quite simple and cheap to compute. Those diagnostics could provide an information on imperfectly known observation and background error statistics.

Diagnosed values of observation error were computed for analyzes of given experiment and compared with prescribed observation errors currently used in the model. The statistic were computed as an average for the evaluation period of 21 days from 20080330 00 UTC till 20080419 18 UTC. We have noticed that some departures $(y_j^o - y_j^a)$ and $(y_j^o - y_j^b)$ have different signs, which means that analysis is state not between guess and observation state and it can be considered as probably wrong analysis behavior. Those points were skipped from the statistic computations.

Predefined σ_o for TEMP observations are vertically divided into following 19 layers: 1000, 950, 900, 850, 800, 700, 600, 500, 400, 300, 250, 200, 150, 100, 70, 50, 30, 20, 10 hPa (see uti/bator/bator_init.F90). Figure 5. shows comparison of obtained diagnostics and prescribed observation errors for TEMP observations, default observation errors are generally larger (except for q) than diagnosed ones.

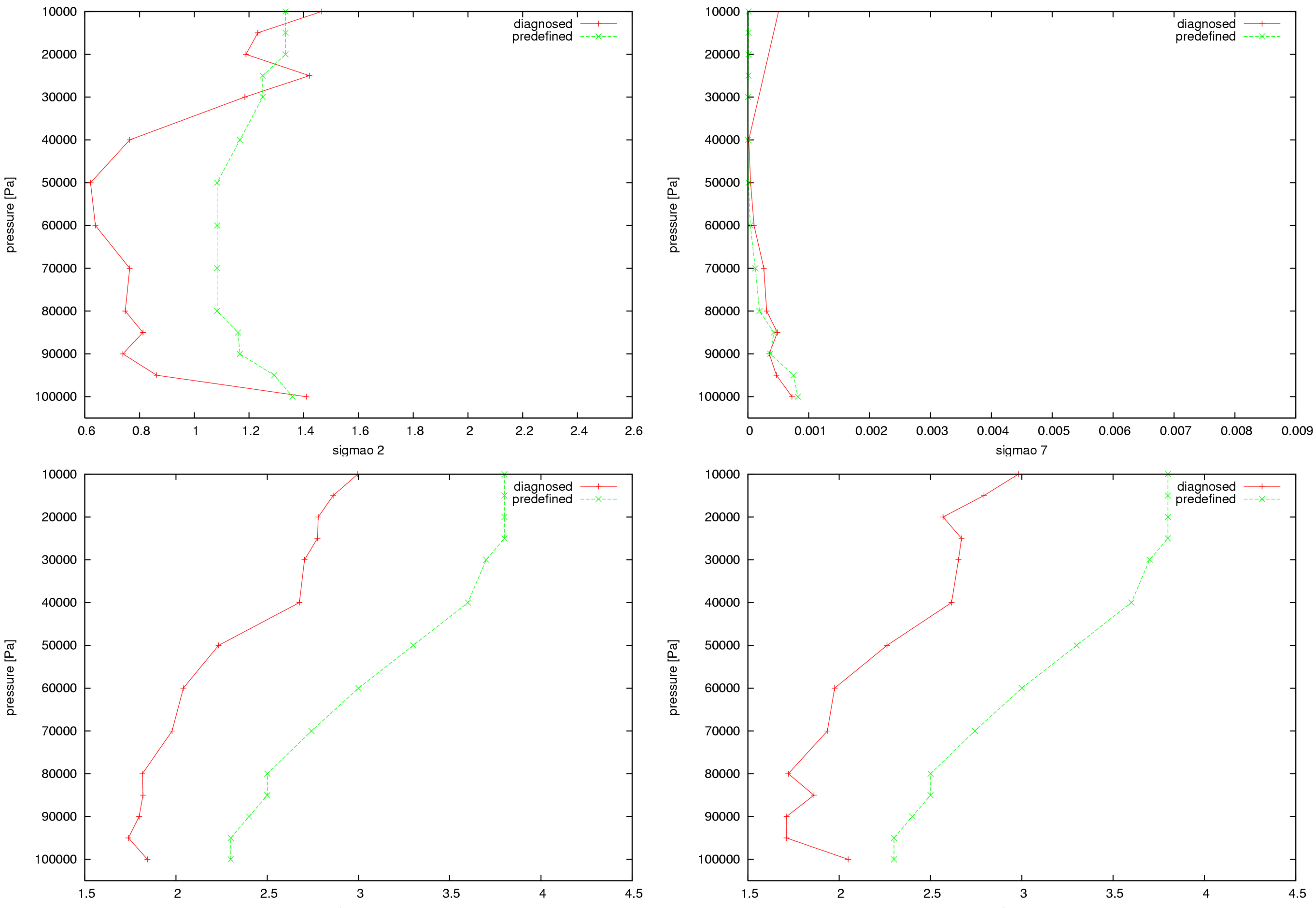

Fig. 5: Diagnosed (as red full line) and default σ_o (as green dashed line) for T (upper-left), q (upper-right), u (bottom-left) and v (bottom-right) of TEMP observations.

Predefined σ_b can be obtained from minimization listing, where vertical profile of standard deviation are printed for vorticity, unbalanced divergence, wind, mass, temperature (balanced and unbalanced) and specific humidity. Diagnostics are based on departures of analyzed parameters, so for the moment temperature, specific humidity and wind components are considered only. Based on Bölöni (2006) diagnosed standard deviation of wind components were recomputed to average wind using following formula

$$\sigma_b(uv) = \sqrt{\frac{\sigma_b^2(u) + \sigma_b^2(v)}{2}} \tag{1}$$

The diagnostics of error statistics can be used for estimation of misfit ratio r of predefined σ_p and diagnosed σ_d

$$r = \frac{\sigma_p}{\sigma_d} \tag{2}$$

to perform vertical dependent tuning (multiplication of misfit ratios in Ald/var/suejbstd.F90) of temperature, specific humidity and wind standard deviations. The open question is if tuning can be performed on some parameters only, while remaining parameters are untouched or is better to try some kind of uniform tuning, e.g. by REDNMC factor. For the first trial we have tried vertically dependent tuning of temperature, specific humidity and wind of TEMP observation only.

Following figures show comparison of obtained diagnostics and prescribed background standard deviations for TEMP observations.

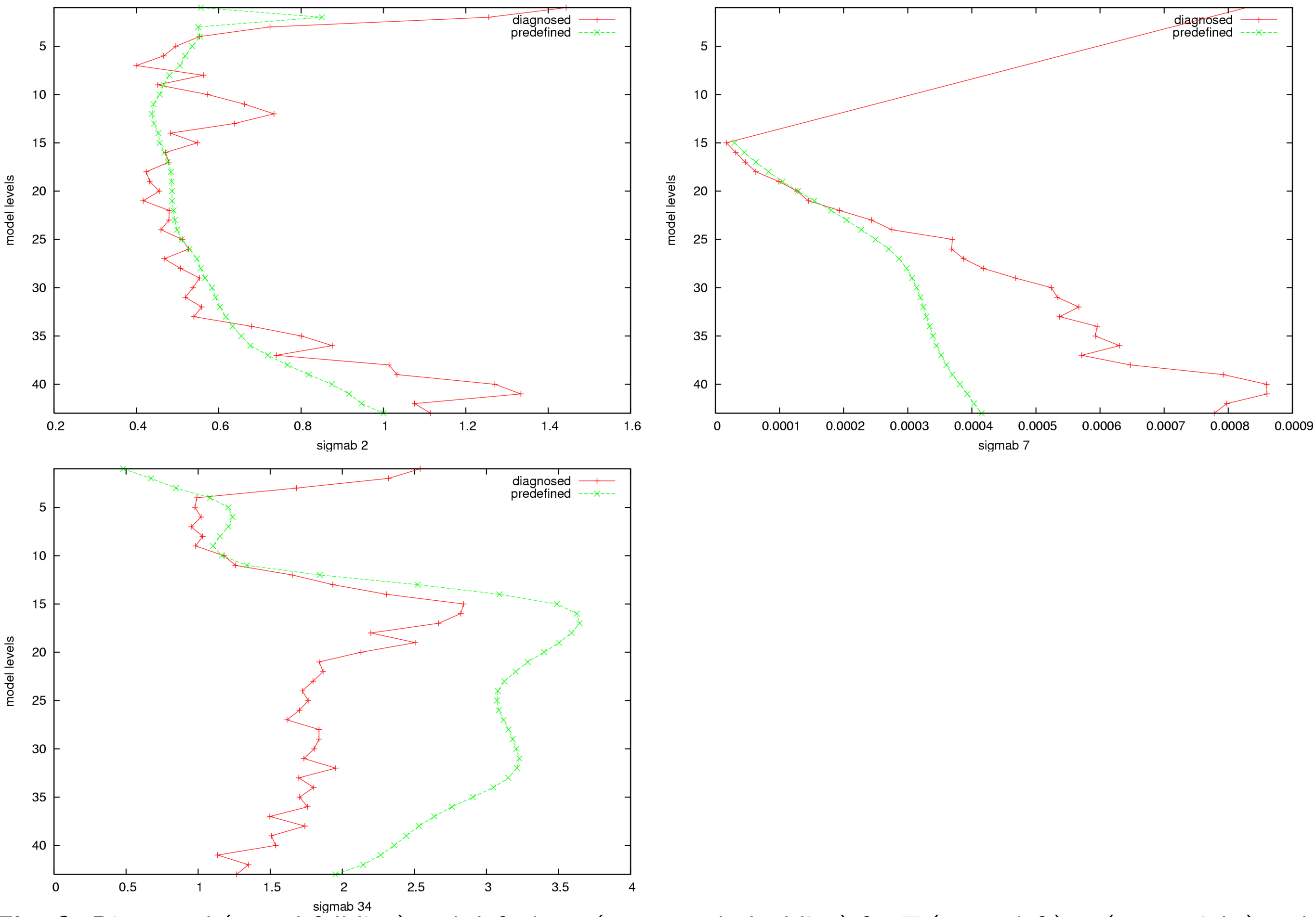

Fig. 6: Diagnosed (as red full line) and default σ_b (as green dashed line) for T (upper-left), q (upper-right) and wind (bottom-left) of TEMP observations.

3.1 Impact of error statistics tuning

In order to evaluate impact of error statistics tuning, next experiment **Z10** was performed. Diagnosed standard deviations were used instead of default ones.

- **Z10** - BlendVar with diagnosed σ_o and σ_b

The same diagnostics of observation errors were performed and figures Fig. 7 and Fig. 8 show better agreement of predefined and diagnosed standard deviations.

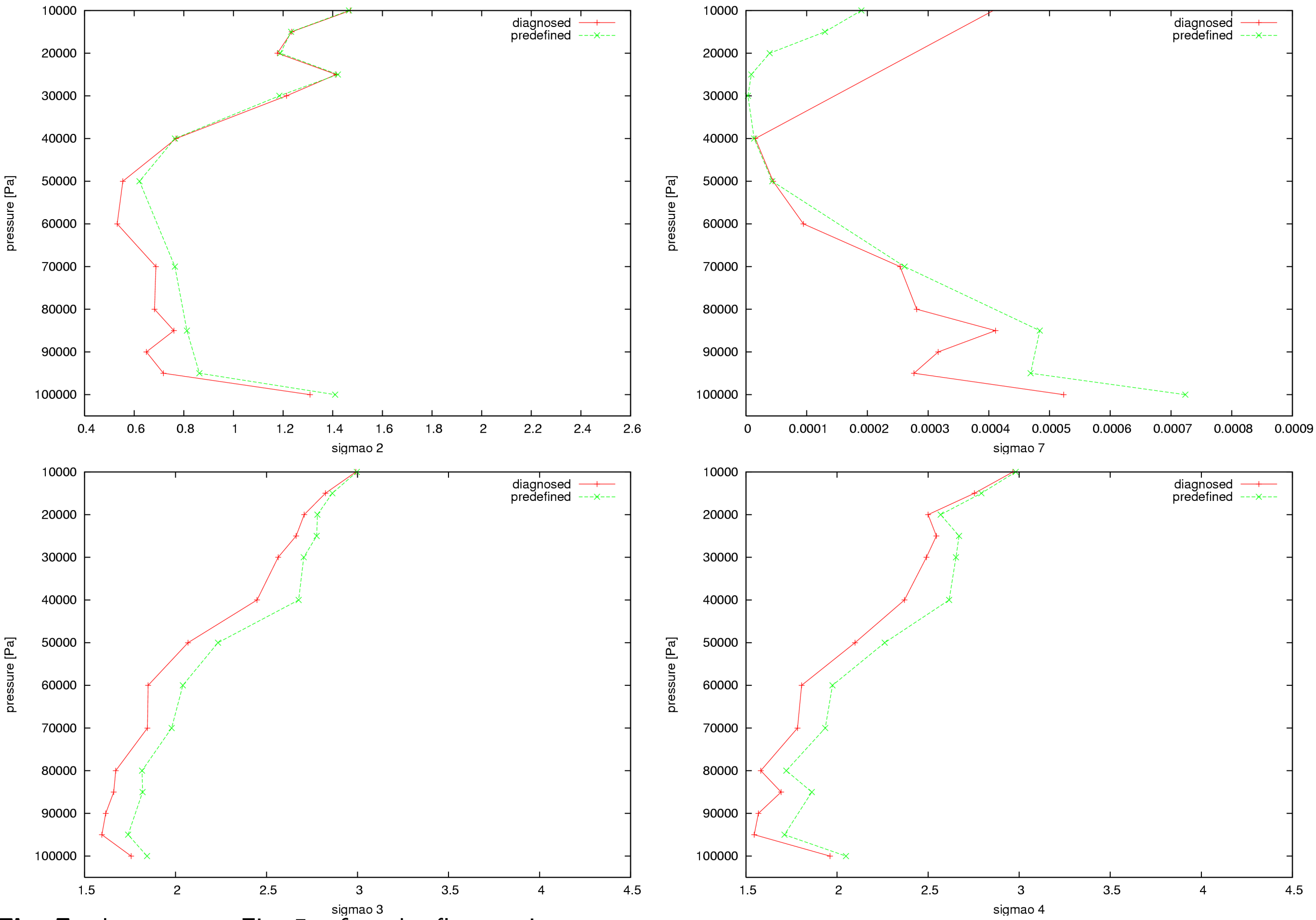

Fig. 7: the same as Fig. 5, after the first tuning

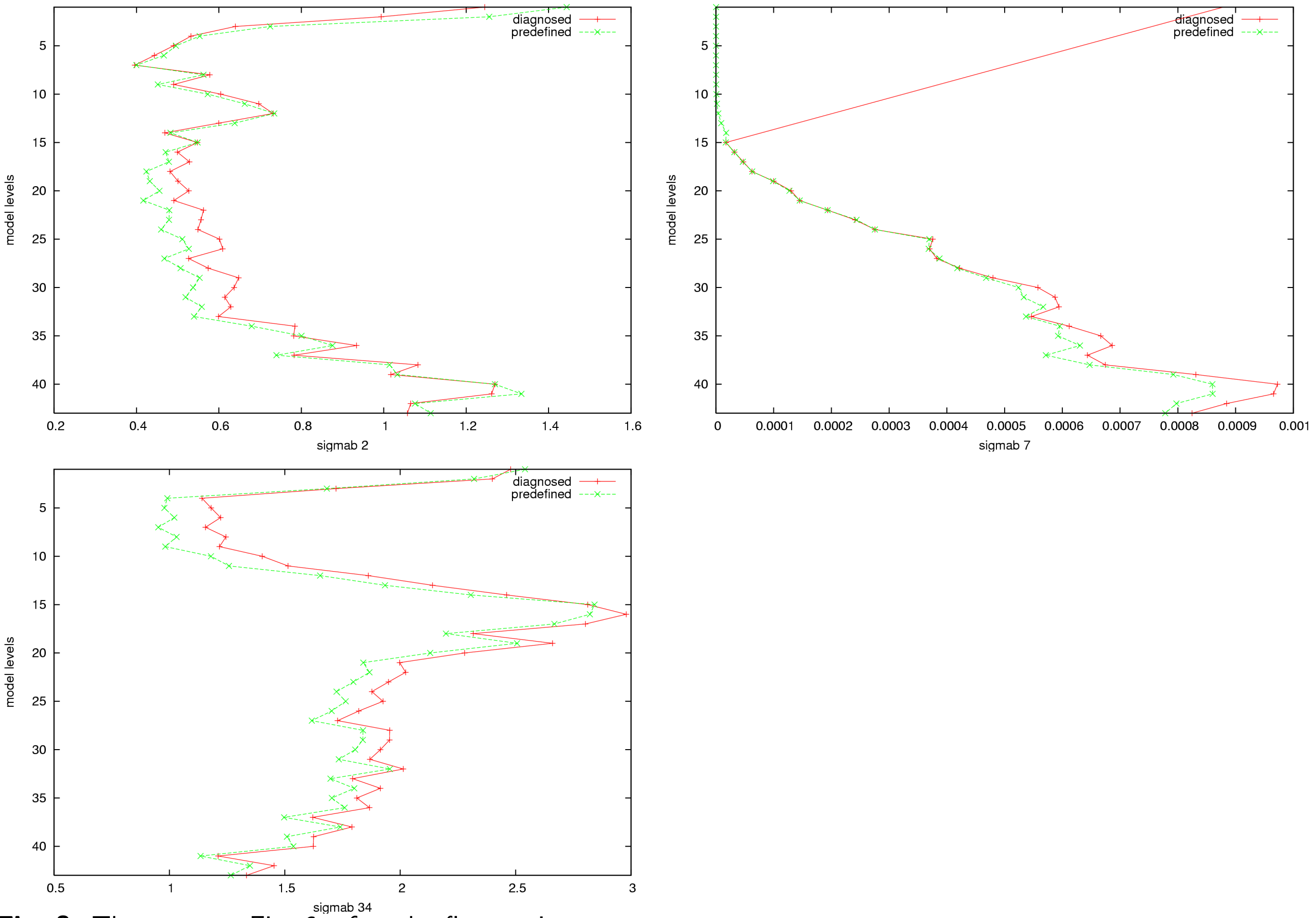

Fig. 8: The same as Fig. 6, after the first tuning

3.2 Analysis behavior

The number of observation is almost the same, and from vertical profiles of monitoring statistics (mean and standard deviation) show better better fit to observation. Standard deviation of obs-analysis are smaller than with default setting of standard deviation. see Fig 9.

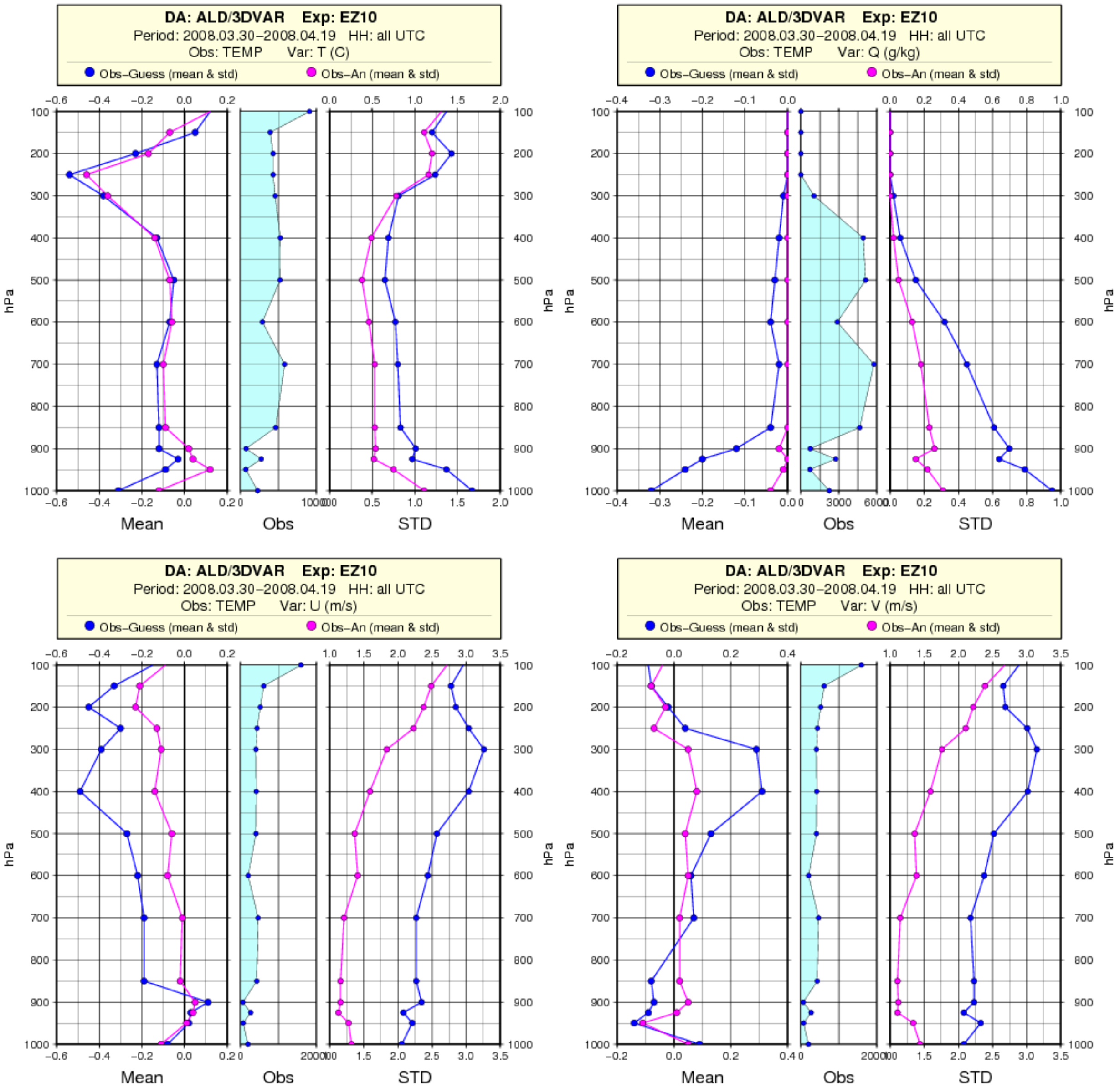

Fig. 9: the same as Fig. 2., after the first tuning

3.3 Verification against observation

VERAL scores were computed for the evaluation period from 20080330 till 20080419 of 00 UTC production runs. RMSE scores are plotted on Fig. 10. RMSE differences are very small, up to 0.1 K for T, 1% for RH, 0.2 m^2s^{-2} for ϕ and 0.2 m/s for wind.

- **Z09** - BlendVar with default σ_o and σ_b
- **Z10** - BlendVar with diagnosed σ_o and σ_b

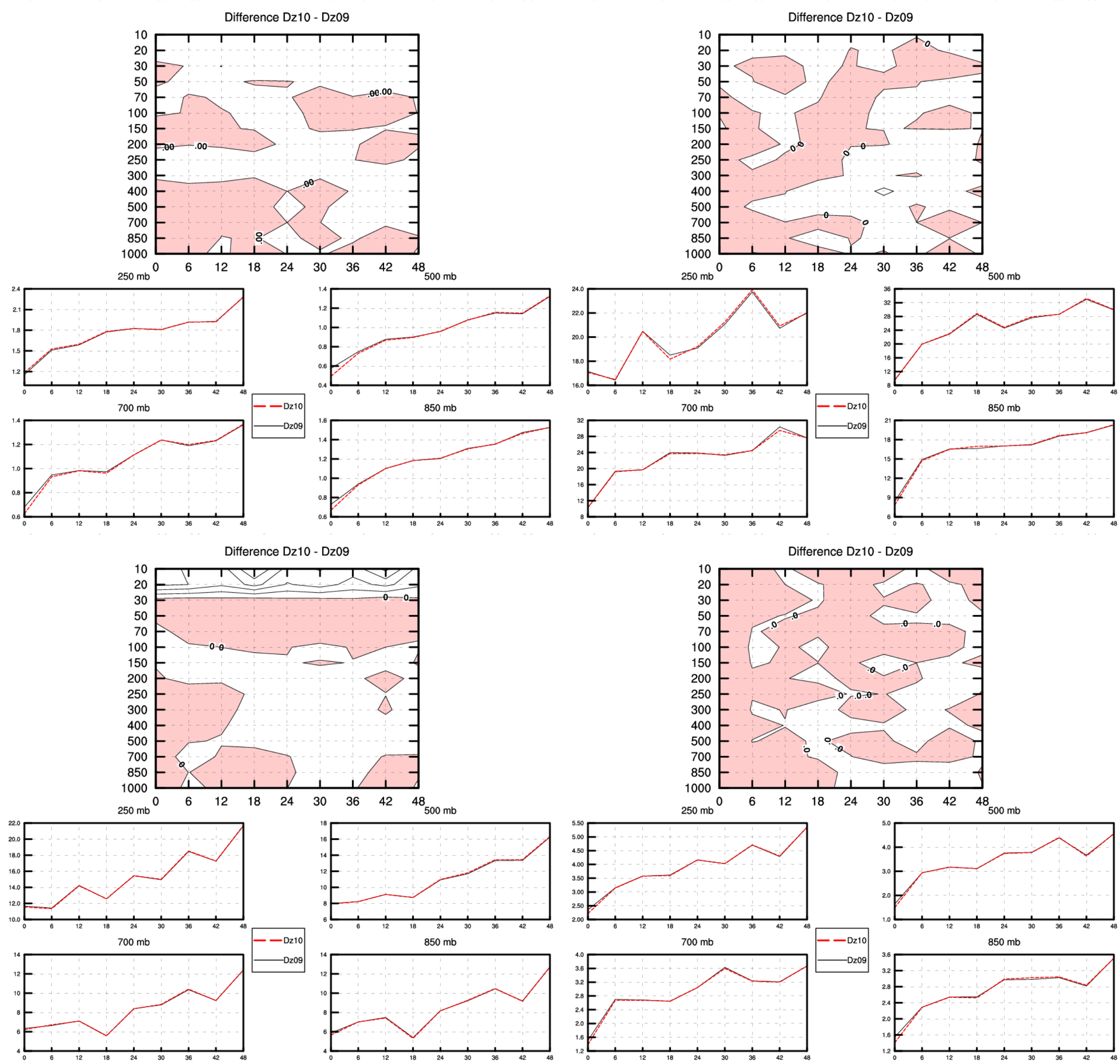

Fig. 9: RMSE of T (upper-left), RH (upper-right), ϕ (bottom-left) and wind speed (bottom-right) from 20080330 till 20080419 of 00 UTC runs. Red areas denote positive impact of σ_o and σ_b tuning.

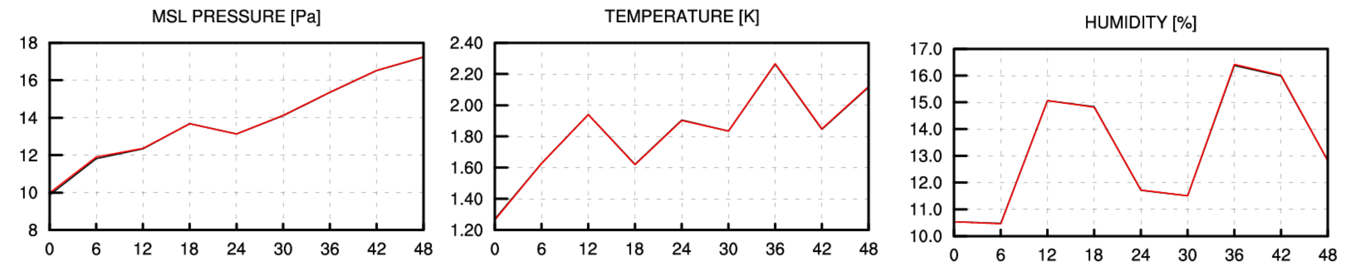

Fig. 10: RMSE of $MSLP$ (left), $T2m$ (middle) and $RH2m$ (right) from 20080330 till 20080419 of 00 UTC runs. Experiment **Z10** in red and **Z09** in black color.

Results of significance test are summarized in Table 2. There prevail improvements for the analysis time, but there are degradation of MSLP and mainly above 400 hPa as well.

3.4 Summary

The first tuning didn't bring clearly positive impact, maybe more tuning iteration should be performed and/or tuning of standard deviation of surface geopotential should be considered.

Parameter	Forecast	Significance at 90%	Parameter	Forecast	Significance at 90%
MSLP	+00H	worse	T 500 hPa	+00H	**better**
	+06H	worse		+06H	**better**
ϕ 1000 hPa	+00H	**better**	V 500 hPa	+00H	**better**
RH 1000 hPa	+00H	**better**	ϕ 400 hPa	+00H	**better**
	+18H	**better**		+06H	**better**
T 1000 hPa	+00H	**better**	RH 400 hPa	+24H	worse
V 1000 hPa	+00H	**better**		+30H	worse
	+42H	worse		+36H	worse
ϕ 850 hPa	+00H	**better**	T 400 hPa	+00H	**better**
	+18H	**better**	V 400 hPa	+00H	**better**
RH 850 hPa	+00H	**better**	ϕ 300 hPa	+00H	**better**
T 850 hPa	+00H	**better**		+06H	**better**
V 850 hPa	+00H	**better**	T 300 hPa	+00H	worse
ϕ 700 hPa	+00H	**better**		+06H	worse
	+06H	worse	V 300 hPa	+00H	**better**
	+36H	worse	ϕ 250 hPa	+00H	**better**
RH 700 hPa	+42H	**better**		+06H	**better**
T 700 hPa	+00H	**better**	T 250 hPa	+00H	worse
	+06H	**better**		+06H	worse
V 700 hPa	+00H	**better**	V 250 hPa	+00H	**better**
ϕ 500 hPa	+30H	worse			

Table 2: List of parameters and forecast ranges where Z10 performs better/worse than Z09 in terms of RMSE scores against observation with significance 90 % two side confidence interval significance test.

4 Conclusions and future plans

Overall scores of the first BlendVar experiments are very encouraging, worrying is small but still significant degradation of some near the surface parameters. Concerning optimal BlendVar setting, study of standard deviation tuning should be more elaborated (second iteration can be tried, tuning of standard deviation of surface geopotential can be performed, comparison with uniform tuning via REDNMC factor), an impact of B matrix (more update one, including ALARO physics and/or LAGGED statistics including surface analysis) and effect of IDFI initialization can be considered. Concerning verification issues, the case studies and more detailed verification of precipitation is planned. Only afterwards evaluation of different periods and impact of other observations (SYNOP T2m,RH2m, SEVIRI, AMDAR) will be studied.

References

G. Bölöni, 2006: Diagnosis of residual covariances and tuning of background and observation errors in the ALADIN/HU 3DVAR *internal report*

G. Desroziers, L. Berre, B. Chapnik and P. Poli, 2006: Diagnosis of observation, background and analysis error statistics in observation space *QJRMS*, **131**, 3385 - 3396

M. Fisher, 2001: Statistical Significance Testing of Forecast Experiments *ECMWF Research Department Memorandum*

M. Široká, 2001, Report on experiments with ALADIN/LACE 3DVAR: combination of 3DVAR with lagged statistics and blending by DFI *RC LACE internal report*