Diagnosis of residual covariances and tuning of background and observation errors in the ALADIN/HU 3DVAR

Gergely Bölöni (HMS)

31.07.2006

Contents

1 Introduction

Predefined statistics of background and observation errors are estimations of the real errors of the analysis system. These estimations are highly dependent on the sampling technique used while providing the error statistics and they not surely optimal for providing reliable analyses. A posteriori diagnostics can give important guidelines for the refinement of some error statistics (mostly variances) based on theoretical considerations. Several a posteriori validation techniques exist based on different theoretical backgrounds. One approach applied in ARPEGE/ALADIN is based on the assumption that the expected value of the cost function in the minimum is proportional to the number of observations used ($M[J_{min}] = p/2$, where $M[...]$ stands for the statistical expectation and p is the number of observations used in the analysis) (Talagrand 1998, Desroziers and Ivanov 2001, Chapnik et al. 2004, Sadiki and Fischer 2005). Another method was proposed recently by Desroziers et al. (2005) which is the theoretical basis for the work summarized in this report. This method is briefly explained in the next section, for more details we refer to the abovementioned paper.

2 Description of the method

In a linear analysis system where the $\mathbf{B}$ background and the $\mathbf{R}$ observation error covariance matrices are correct, the following equations are true (Desroziers et al. 2005):

$$
\begin{aligned}
M[\mathbf{d^o_b}(\mathbf{d^o_b})^T] &= \mathbf{HBH^T} + \mathbf{R} \\
M[\mathbf{d^a_b}(\mathbf{d^o_b})^T] &= \mathbf{HBH^T} \\
M[\mathbf{d^a_o}(\mathbf{d^o_b})^T] &= \mathbf{R} \\
M[\mathbf{d^a_b}(\mathbf{d^o_a})^T] &= \mathbf{HAH^T}
\end{aligned}
$$

where

$$
\begin{aligned}
\mathbf{d^o_b} &= \mathbf{y} - H(\mathbf{x_b}) \\
\mathbf{d^a_b} &= H(\mathbf{x_a}) - H(\mathbf{x_b}) = \ \mathbf{H}\delta\mathbf{x_a} \\
\mathbf{d^o_a} &= \mathbf{y} - H(\mathbf{x_a})
\end{aligned}
$$

are the residuals provided by the assimilation system, $\mathbf{H}$ is the linearized of the H observation operator and $M[...]$ stands for the statistical expectation. It can be checked if the equations above are really true for the used assimilation system, if one computes the residual covariances on the left hand side on archived residuals from this system and substitutes the predefined error characteristics to the right hand side. This diagnosis can be applied to the variances for instance. If the diagnostic equations are not fulfilled one can derive the misfit of the predefined error characteristics and possibly use this information to tune the assimilation system. In practice the variances can be estimated according to the left hand side of the second and third equations above:

$$\sigma_b^2 = \frac{1}{P_i}\sum_{i=1}^{P_i} d^a_{b\,i} d^o_{b\,i}$$
$$\sigma_o^2 = \frac{1}{P_i}\sum_{i=1}^{P_i} d^o_{a\,i} d^o_{b\,i}$$

where $d^o_{b\,i}$, $d^o_{a\,i}$ and $d^a_{b\,i}$ stand for the individual realizations of the first guess and analysis departures i being the index for these realisations.

3 Diagnostic results

First guess departures ($d^o_{b\,i}$) and analysis departures($d^o_{a\,i}$) of TEMP and AMDAR measurements were used from two sets of assimilation cycles (October-November 2005 and June 2006). These departures can be retrieved easily from the compressed ODB data base (CCMA) after the 3DVAR run. So one has just to collect these data bases in order to have a large enough sample and to dump the content using the ODB viewer. So far vertically averaged diagnostics were derived for temperature, humidity and wind separately according to the last equations above.

3.1 Main features

The day to day evolution of the residual statistics are changing very similarly for SNMC and ENSB. This is true for both periods and for all the variables but shown only for humidity (figures 1-2) for the period of summer 2006. This simply suggests that the two assimilation systems using different B matrices depend similarly on the actual weather situation in terms of first guess and analysis departures. Also a dependence on the network time can be observed. Namely at 06 and 18 UTC the estimated standard deviations (both σ_b and σ_o) are clearly smaller than at 00 and 12 UTC, which is probably due to the smaller amount of observational data and as a consequence the smaller average analysis and first guess departures at 06 and 18 UTC.

Figures 1-2 show also that the estimated σ_b is bigger for SNMC than for ENSB and the estimated σ_o is smaller for SNMC than for ENSB for all the variables. This implies through the corresponding covariances of residuals that the analysis is closer to the observations (and further from the background) if

SNMC is used than if ENSB is used. With other words, the SNMC statistics model the background as less reliable than the ENSB statistics in the analysis.

3.2 Estimation of the misfit

The a posteriori diagnostics explained above can be considered as a good estimate of the background and observation error standard deviations in an ideal assimilation system (Desroziers et al. 2005). In this section the estimation of the background error standard deviation will be discussed exclusively. The standard deviation misfit of the system can be obtained as a ratio (r) of the predefined (σ_p) and the estimated (σ_e) standard deviations:

$$r = \sigma_e/\sigma_p$$

This ratio can be used as a guidance for re-tuning the predefined standard deviations in the assimilation system (see also in a later section). The predefined values are given as a vertical profile on model levels for each variable separately (temperature, humidity, vorticity and divergence) in the ALADIN 3DVAR. As previous experiments gave promising results using an ensemble B matrix (Horvath 2005) we were concentrating on the estimation of the ENSB standard deviation misfit. As a first step, vertically averaged misfit ratios of each variable were computed. This means that the predefined standard deviation profiles were vertically averaged and the estimated standard deviations were computed including the first guess and analysis departures taken from the whole 3D domain (no vertical separation was done).

One important technical difficulty in the misfit estimation was related to the treatment of wind. The problem is that first guess and analysis departures are available for the wind components u and v but the predefined standard deviations are given for vorticity (ζ) and divergence (η), as these are the wind control variables in the ALADIN 3DVAR. In order to obtain comparable standard deviations, the predefined ζ and η standard deviations were converted to those of wind:

$$\sigma_b(uv) = \sqrt{-\tfrac{1}{2}\Delta^{-1}(\sigma_b^2(\zeta) + \sigma_b^2(\eta))}$$

where $\sigma_b(uv)$ corresponds to the average wind:

$$\sigma_b(uv) = \sqrt{\tfrac{(\sigma_b^2(u)+\sigma_b^2(v))}{2}}$$

The estimated standard deviations (σ_e) are obtained by an averaging over the investigated time period just like shown on figures 1-2. Tables 1 and 2 show the obtained misfit ratios for the different variables for the two separate periods. Also an average over the variables is computed in order to obtain one single misfit ratio for the whole analysis system. (The average over the variables are computed as the ratios of the average standard deviations over the variables and NOT as the average of the misfit ratios of the different variables!!! This latter solution gives considerably higher misfit ratios and i do not know which is the correct way to follow so it should be tested also in the future. Values in brackets in the tables stand for this latter solution.)

Variable	Predefined	Estimated	Missfit ratio
q	$2.27 * 10^{-4}$	$5.34 * 10^{-4}$	2.35
T	0.4917	0.7071	1.43
Wind	1.4840	1.9878	1.33
Average	**0.65**	**0.89**	**1.36 (1.70)**

Table 1: Misfit ratios for the period 26.10.2005 - 10.11.2005

Variable	Predefined	Estimated	Missfit ratio
q	$2.27 * 10^{-4}$	$5.82 * 10^{-4}$	2.56
T	0.4917	0.8010	1.62
Wind	1.4840	1.9203	1.29
Average	**0.65**	**0.9**	**1.38 (1.82)**

Table 2: Misfit ratios for the period 05.06.2006 - 20.06.2006

4 Tuning of the analysis system

The analysis system can be tuned by multiplying the predefined standard deviations with the misfit ratios. This can be done both for σ_b^2 and σ_o^2 in several ways of different complexity:

1. Use a uniform misfit ratio for the whole system
2. Use vertically uniform but variable dependent misfit ratios
3. Use height and variable dependent misfit ratios

In the ALADIN/HU 3DVAR, so far only the σ_b values were tuned following the first and second solutions. (It is rather easy technically to test the first solution. One has simply to set the REDNMC parameter to the estimated misfit ratio in the NAMVAR namelist of the 3DVAR run. The second solution requires some code modification in the subroutine SUEJBSTD which can be found under the "ald/var" project in the source code. The modification is to replace the multiplication of the standard deviations by REDNMC by a variable dependent multiplication. The third solution would need a variable and vertical level dependent multiplication in SUEJBSTD but it would make a sense only if the misfit ratios were estimated for each vertical model level.) Both the second an first solutions were tested in real assimilation cycling experiments over the summer 2006 period. It means a 16 day assimilation cycle (64 analysis steps) using approximately the same observational data base as the operational ALADIN/HU 3DVAR (SYNOP, TEMP, ATOVS AMSU-A and AMSU-B, MSG/AMV, AMDAR aircraft data.) Simulations of 48 hour production runs were started from 00 UTC after a 2 days warm up cycling. The impact of the tuning was tested through RMSE and BIAS score computations of the abovementioned simulated production runs.

For tuning No.1 the REDNMC=1.3 was used which means that global misfit was estimated as 1.3. The results are shown on figures 3 - 6. One can see a small but clear improvement of the wind field with compared to the reference (ENS0) experiment. The temperature field is also improved on most of the levels and there is a clear improvement in the relative humidity field too on 700 hPa.

For tuning No.2 the estimated misfit ratios from the 2005 autumn period were used (1.33 for surface pressure, geopotential, wind, vorticity and divergence, 1.43 for temperature and 2.35 for humidity). The impact of tuning No.2 is shown on figures 7 - 10. One can see that this tuning gives straightforwardly worse results than the experiment without any tuning. The reason for it is not clear so far (for me at least...). On the pictures 11-14 one can see the predefined and the tuned σ_b profiles of the ensemble B matrix used in the experiments. (These diagnostics are taken from the output listing file of the 3DVAR minimization. They are printed out in the routine SUEJBSTD after the tuning multiplications.) Looking at the pictures one can see that the tunings No.1 and No.2 result in quite similar σ_b profiles for mass and wind. The difference is bigger for temperature and the biggest difference can be observed in the humidity σ_b profiles. This suggests that the disappointing results of tuning No.2 are probably mostly due to the large misfit ratios for humidity and temperature. It would be interesting to test which of these two variables are responsible for the poor results. From another point of view it is very disappointing that the variable dependent tuning can not beat the averaged tuning of σ_b. One idea to explain this is that the misfit is not height dependent which can lead to an exaggerated tuning on some levels where in the reality the misfit is not very large.

Also it has to be mentioned that surface pressure, geopotential, wind, vorticity and divergence were all tuned with the same misfit ratio that was estimated for the mean wind. No experiments with tuning No.3 have been done yet.

Finally we show some results comparing the efficiency of the re-tuned (tuning No.1) ensemble background error statistics and the presently used operational NMC B matrix. Scores are shown on figures 15 - 18. The results are quite neutral for geopotential and wind. Improvement can be observed for temperature at mid-levels and for humidity on almost all the levels considering the experiment in which the tuned ensemble statistics were used. This suggests an operational use of the ensemble statistics in the future with tuning No.1 or possibly with some more efficient tuning. It is worth to mention here also that some tuning experiments were also performed for the autumn 2005 period (26.10.2005 - 10.11.2005) which showed an improvement as well using the tuned ensemble statistics compared to the NMC statistics. This latter experiment was done before the estimation of the misfit ratios so a REDNMC=1.5 tuning was used (tye of tuning No.1) based on the experience of our French colleagues who obtained this value using another a posteriori diagnostic technique, the so called J_{min} method (Sadiki and Fischer 2005).

5 Conclusions

It turned out to be technically easy to estimate σ_b and σ_o using the residual statistics proposed by the French colleagues. So far vertically averaged standard deviations were estimated, no vertical separation was applied. If one compares predefined and estimated standard deviations computed with the method of residual of covariances, one technical difficulty is that predefined values are given for vorticity and divergence and the estimated values can be obtained for u and v components. This problem was solved by converting the predefined vorticity and divergence standard deviations into those of wind components.

The residual statistics gave slightly different estimations of the standard deviations in different periods. Also, the estimations depend on the background error statistics used when creating the first guess and analysis departures. In our experiments NMC and ensemble statistics were both used and their standard deviations were re-estimated.

By comparison of predefined and estimated standard deviations, the misfit of the given analysis system can be derived. So far the misfit was estimated only for background errors but not for the observation errors. The estimated misfit can be used to re-tune the predefined values. Two different tunings were tried so far. A uniform one for all the levels and variables and a variable dependent one which is still uniform in the vertical. The uniform tuning gave promising results, it beats the un-tuned ensemble statistics in real assimilation experiments. Also this tuning gives slightly better results than the presently operational NMC B matrix in the ALADIN/HU 3DVAR system. The variable dependent tuning gave disappointing results. The reason is not really understood yet.

In the future a height dependent estimation of the standard deviations will be considered both for background and observation errors. If such estimations are available the misfit can be derived for each vertical model level and a height dependent tuning can be applied. The disappointing results with the variable dependent (vertically uniform) tuning should also be investigated in the future.

Based on the results summarized above the operational application of the ensemble statistics is expected in the future with the best possible tuning.

6 References

Chapnik, B., Desroziers, G., Rabier, F., and Talagrand, O., 2004: Properties and first application of an error-statistics tuning method in variational assimilation, *QJRM, 130, pp 2253-2275*

Desroziers, G., L. Berre, B. Chapnik and P. Polli, 2005: Diagnosis of observation background and analysis error statistics in observation space, *to appear in QJRM*

Desroziers, G, and Ivanov, S., 2001: Diagnosis of adaptive tuning of information error parameters in a variational assimilation, *QJRM 127, 1433-1452*

Sadiki, W. and C. Fischer, 2005: A posteriori validation applied to the 3D-VAR Arpege and ALADIN data assimilation systems, *Tellus, 57A, pp. 21-34*

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, pp 17-18, available at ECMWF*

7 Figures

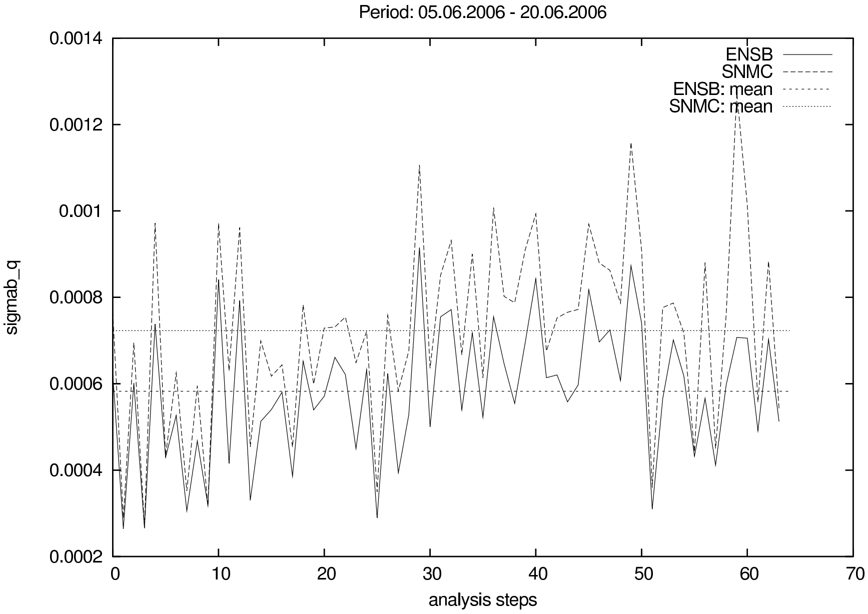

Figure 1: σ_b / humidity

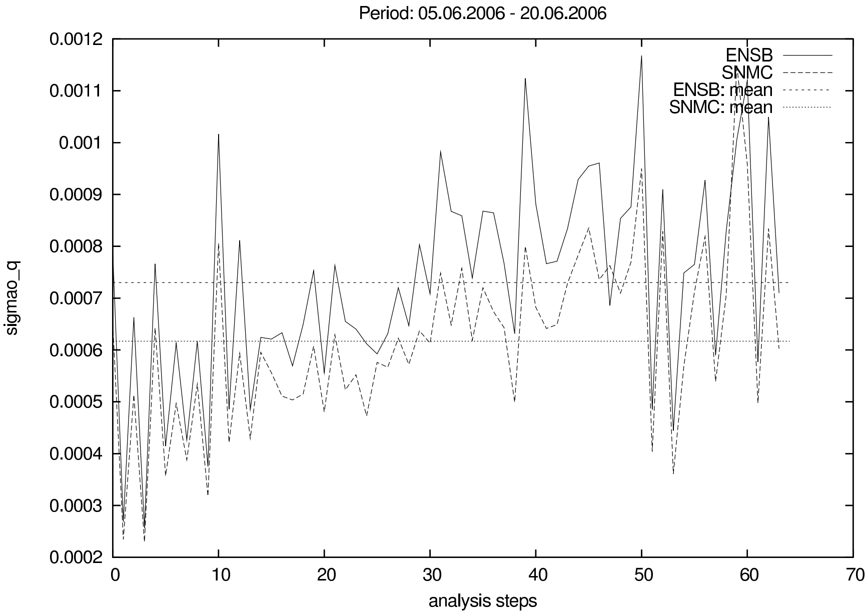

Figure 2: σ_o / humidity

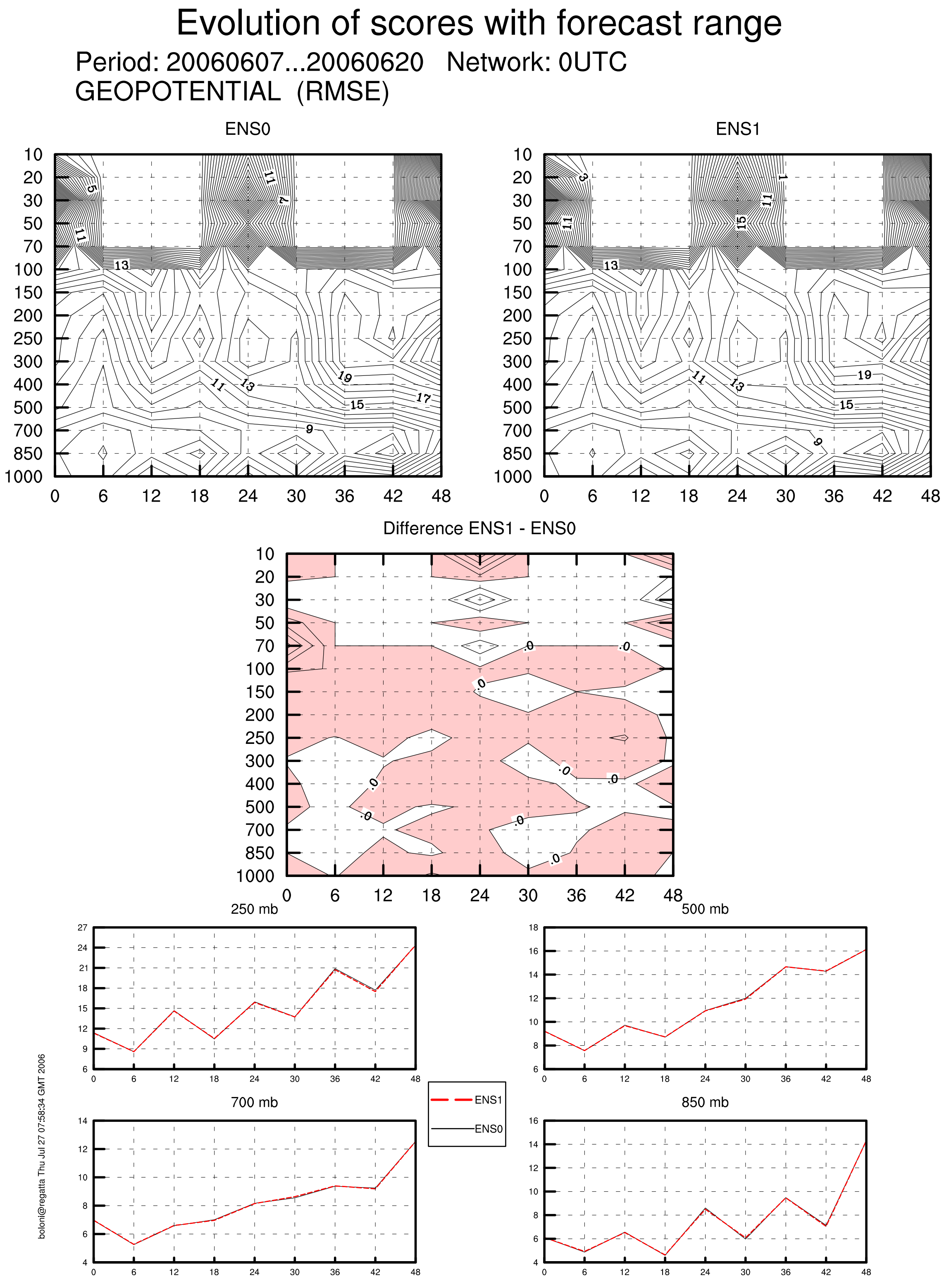

Figure 3: RMSE scores: geopotential ENS0: no tuning, ENS1: tuning No.1

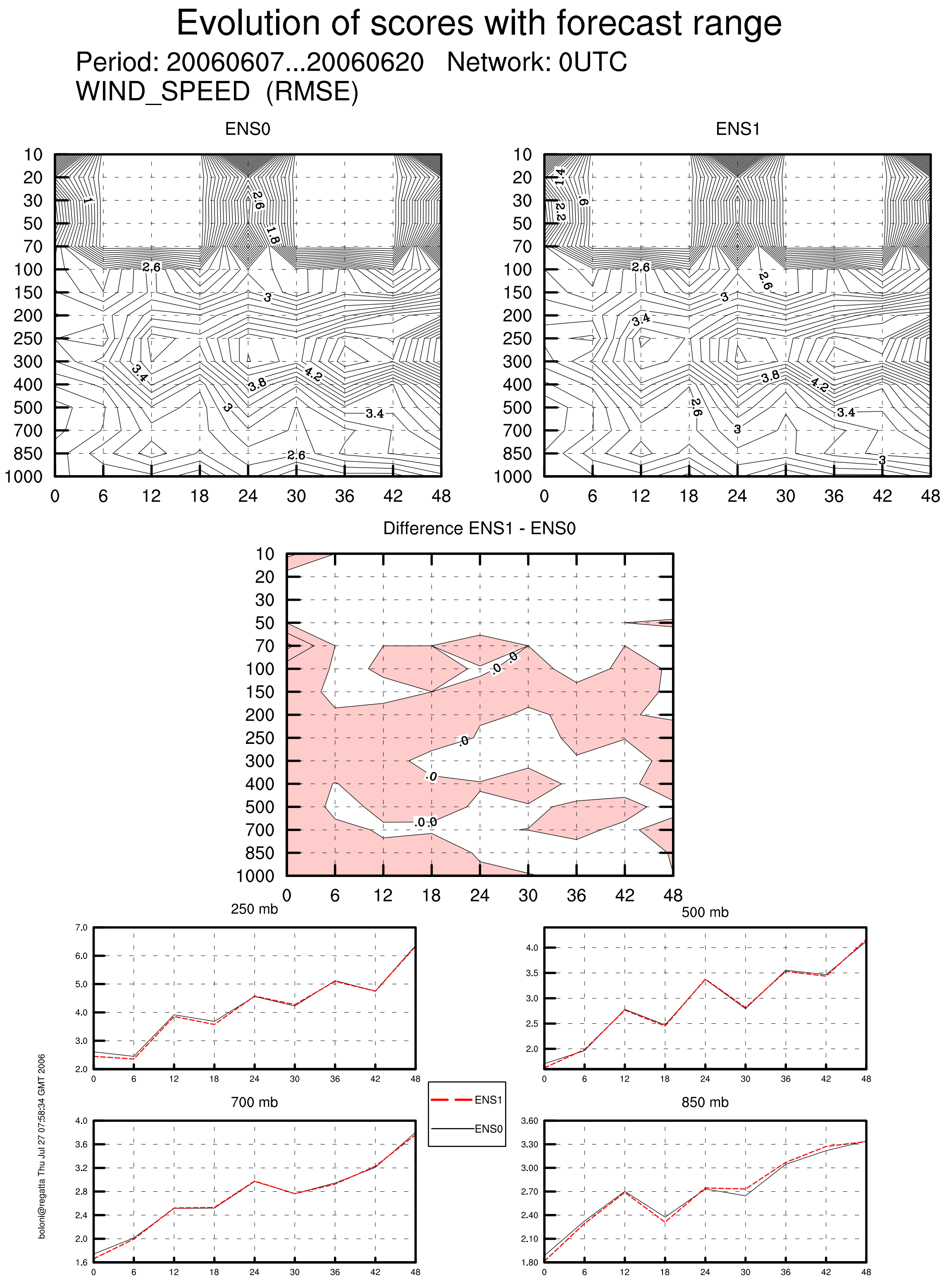

Figure 4: RMSE scores: wind speed ENS0: no tuning, ENS1: tuning No.1

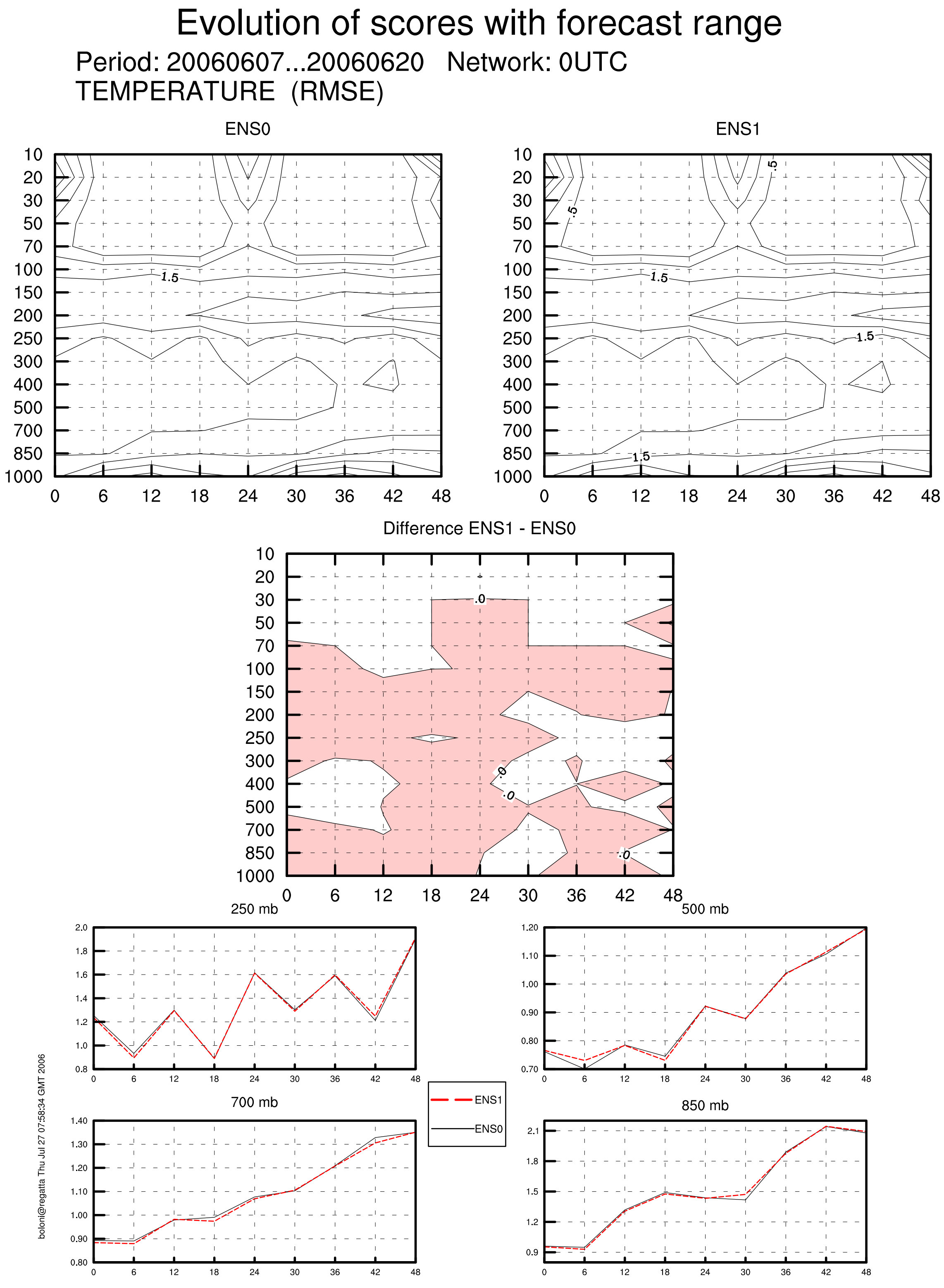

Figure 5: RMSE scores: temperature ENS0: no tuning, ENS1: tuning No.1

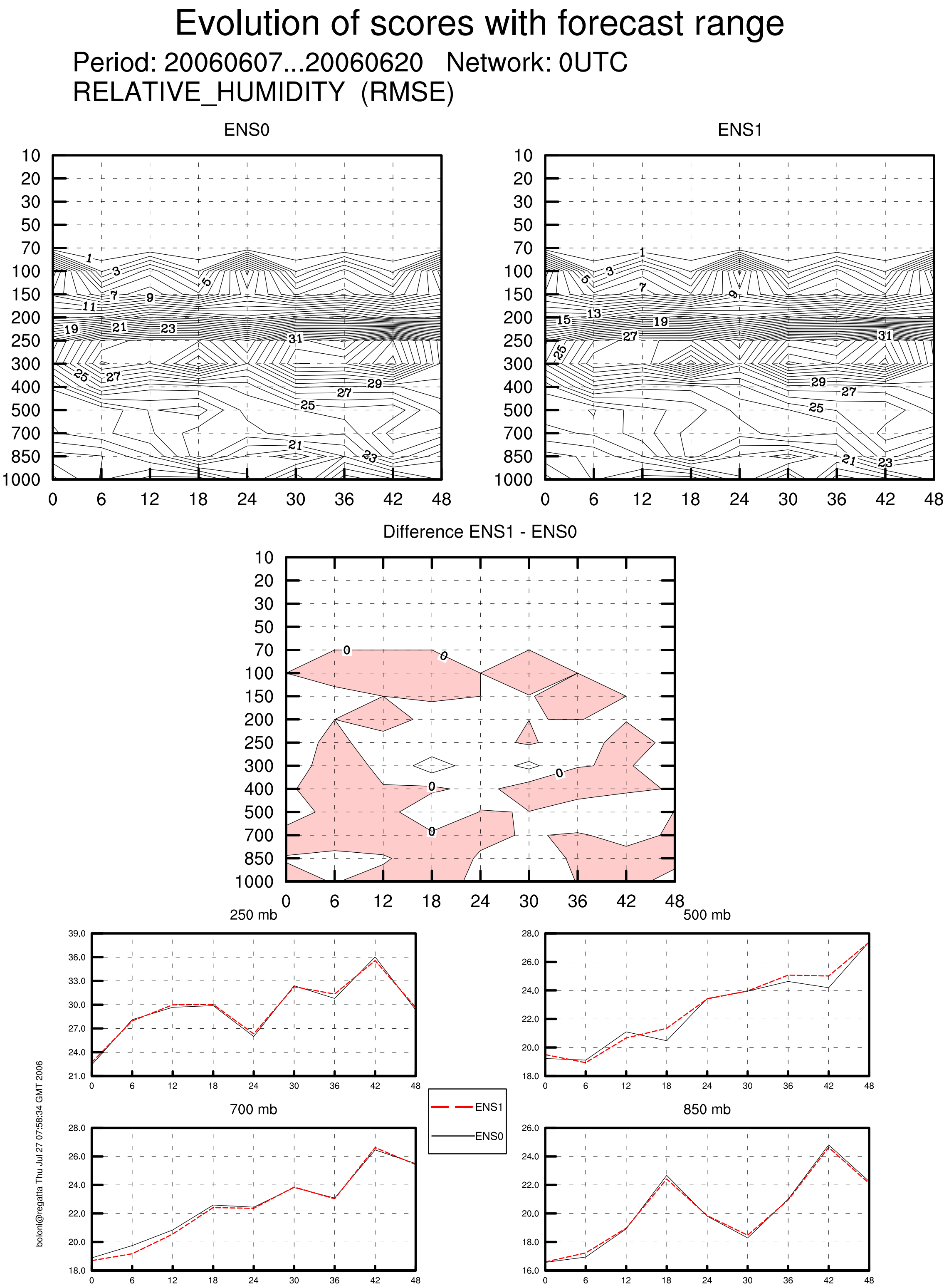

Figure 6: RMSE scores: relative humidity ENS0: no tuning, ENS1: tuning No.1

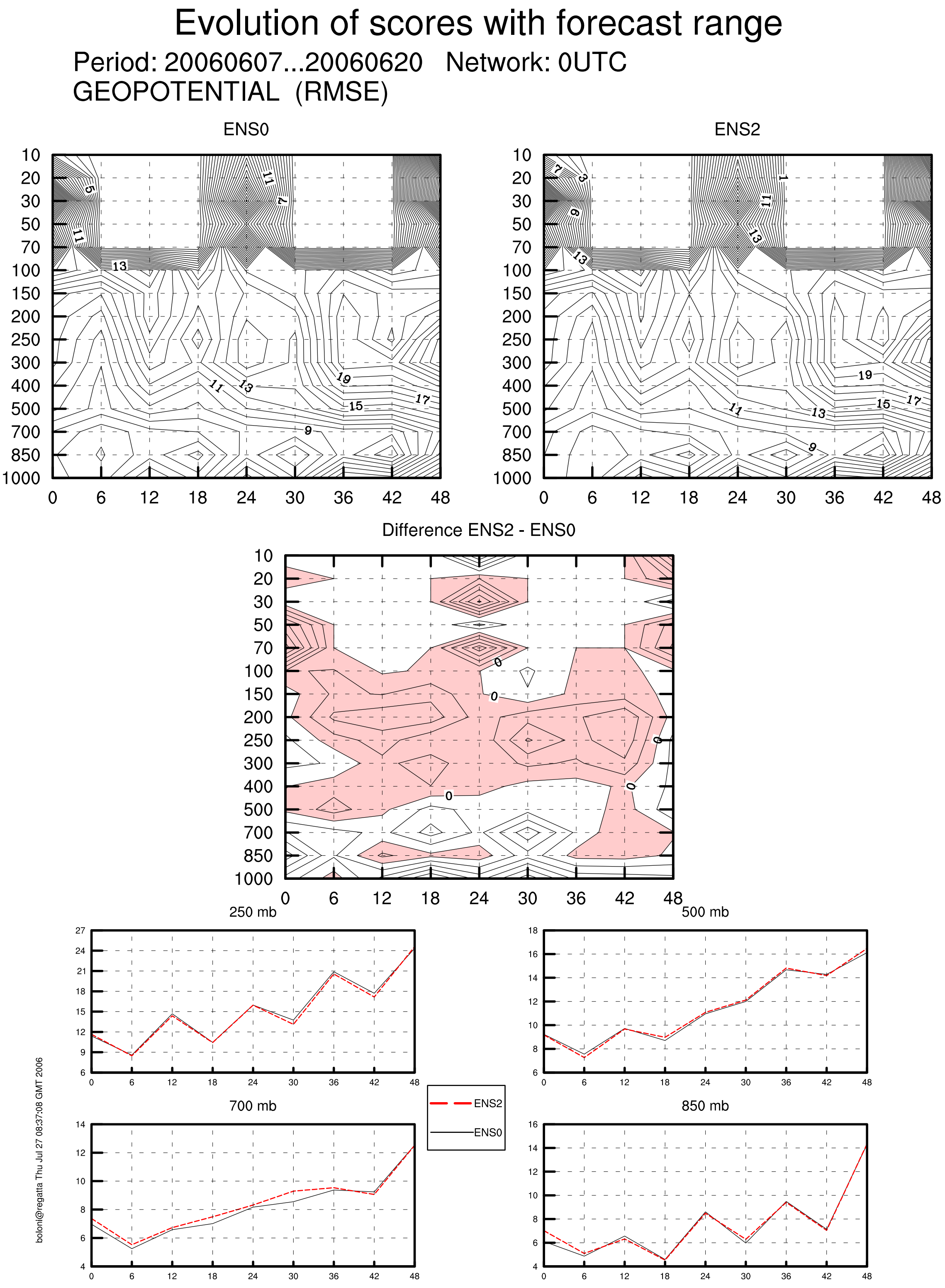

Figure 7: RMSE scores: geopotential ENS0: no tuning, ENS2: tuning No.2

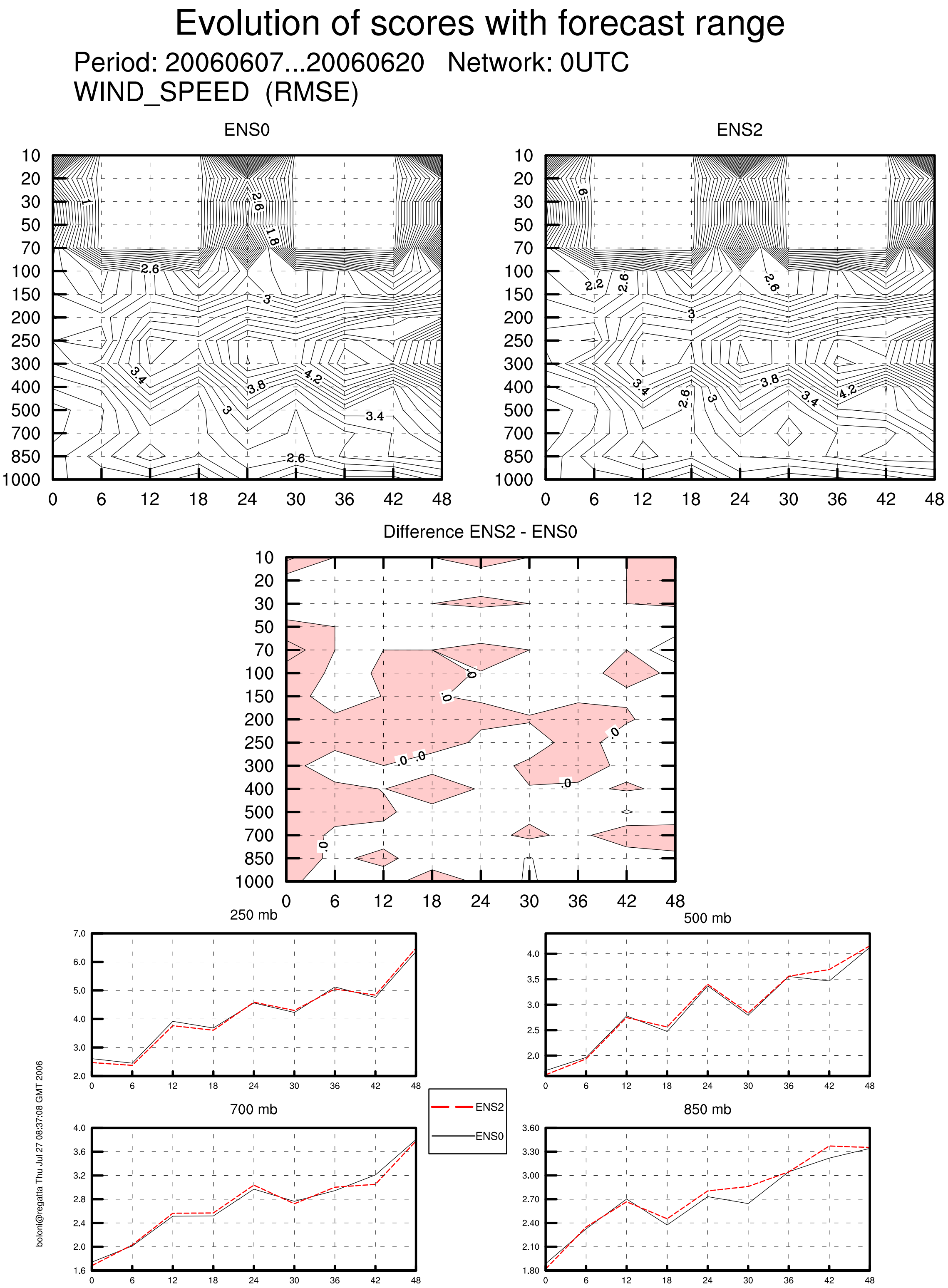

Figure 8: RMSE scores: wind speedENS0: no tuning, ENS2: tuning No.2

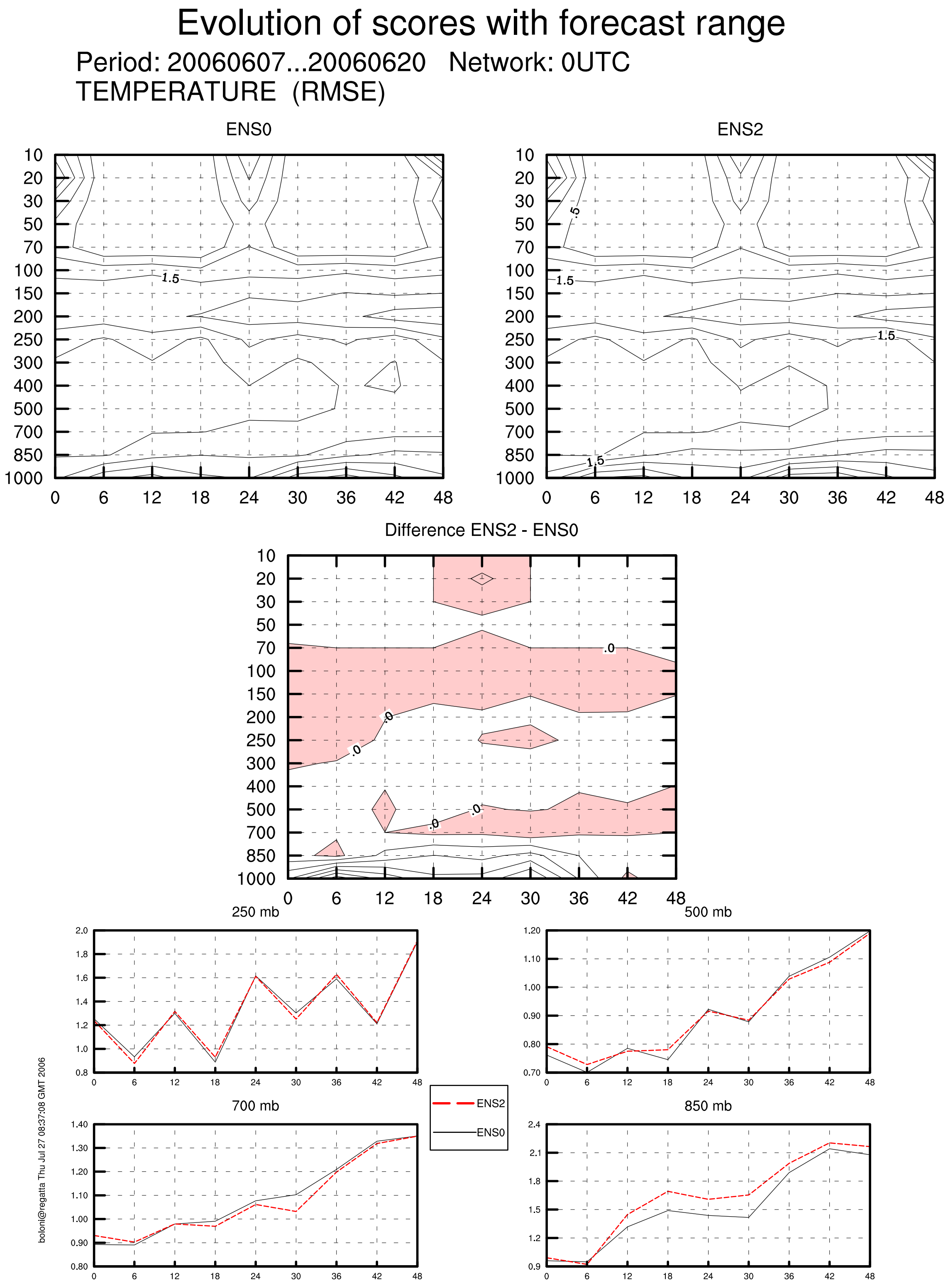

Figure 9: RMSE scores: temperature ENS0: no tuning, ENS2: tuning No.2

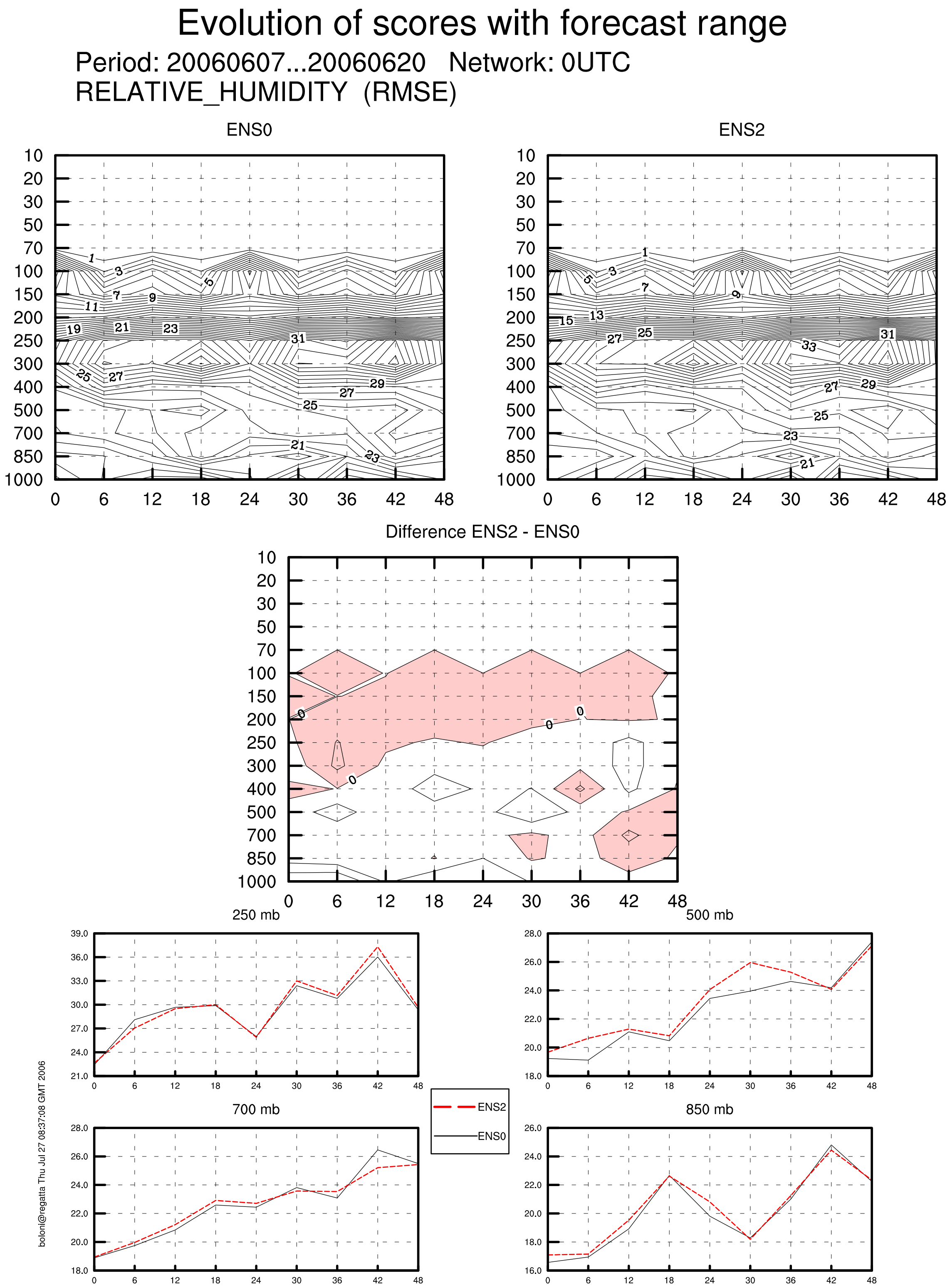

Figure 10: RMSE scores: relative humidity ENS0: no tuning, ENS2: tuning No.2

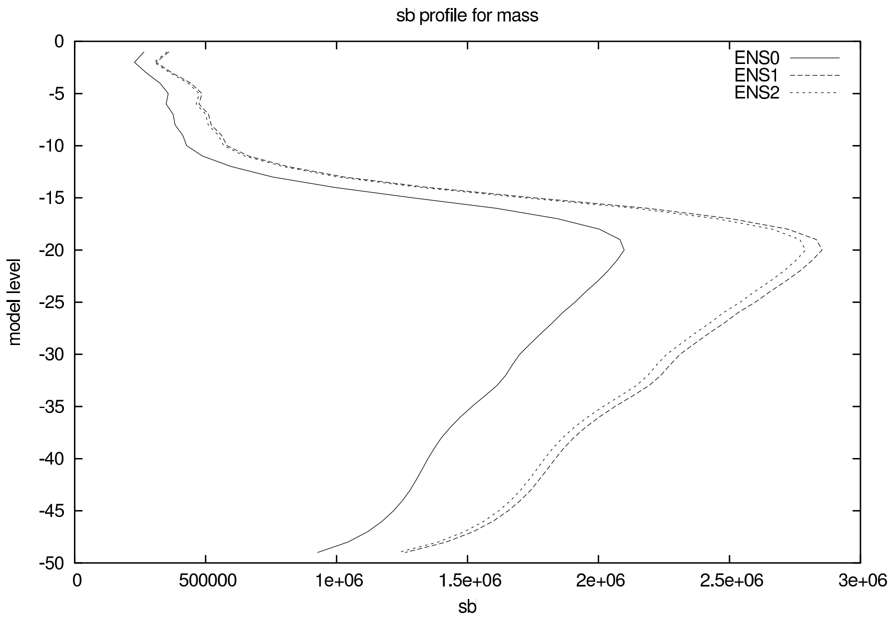

Figure 11: σ_b profiles / mass

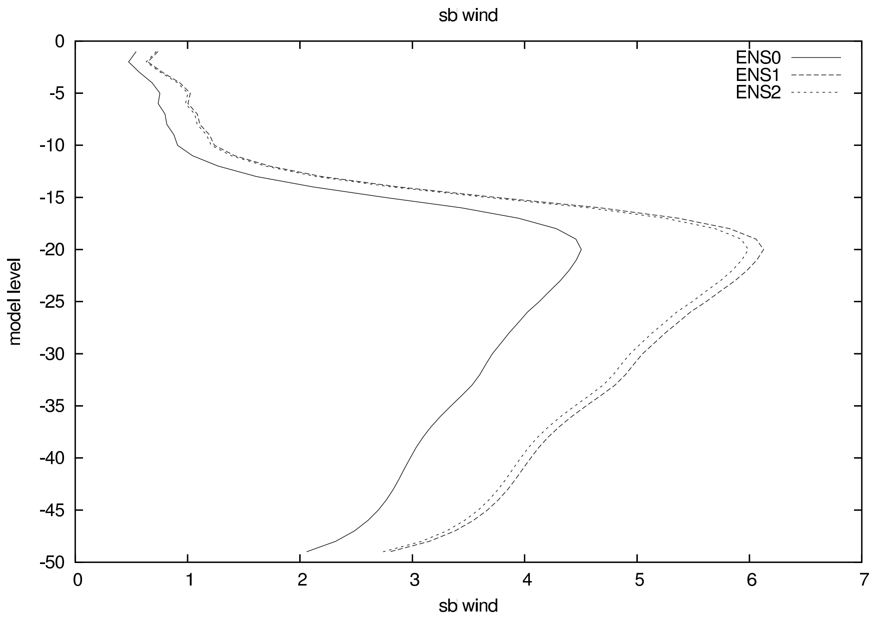

Figure 12: σ_b profiles / wind

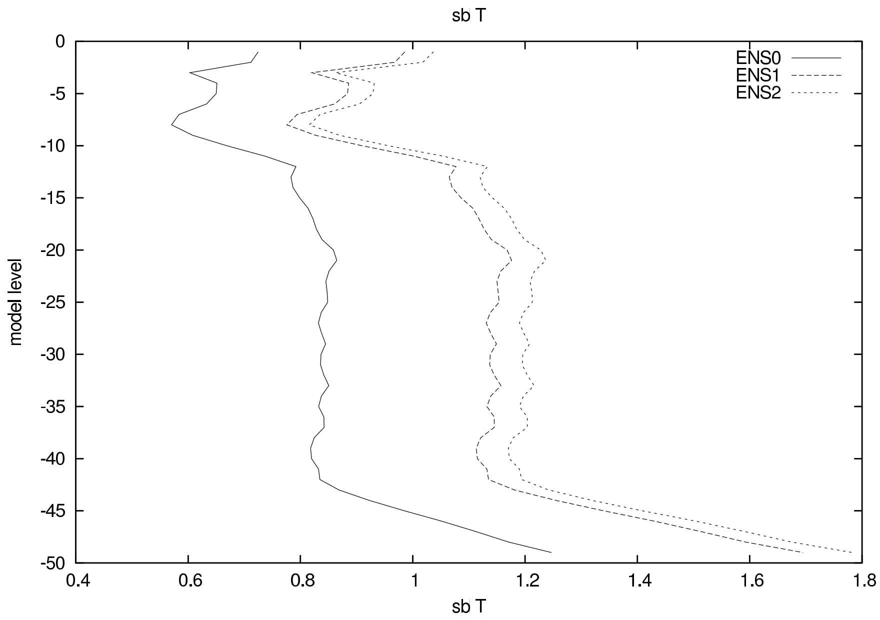

Figure 13: σ_b profiles / temp

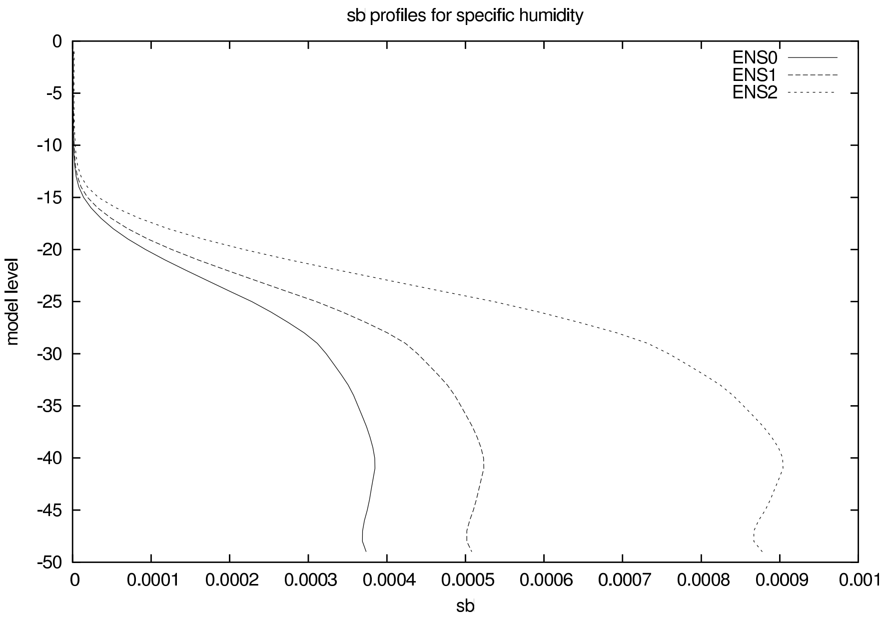

Figure 14: σ_b profiles / hum

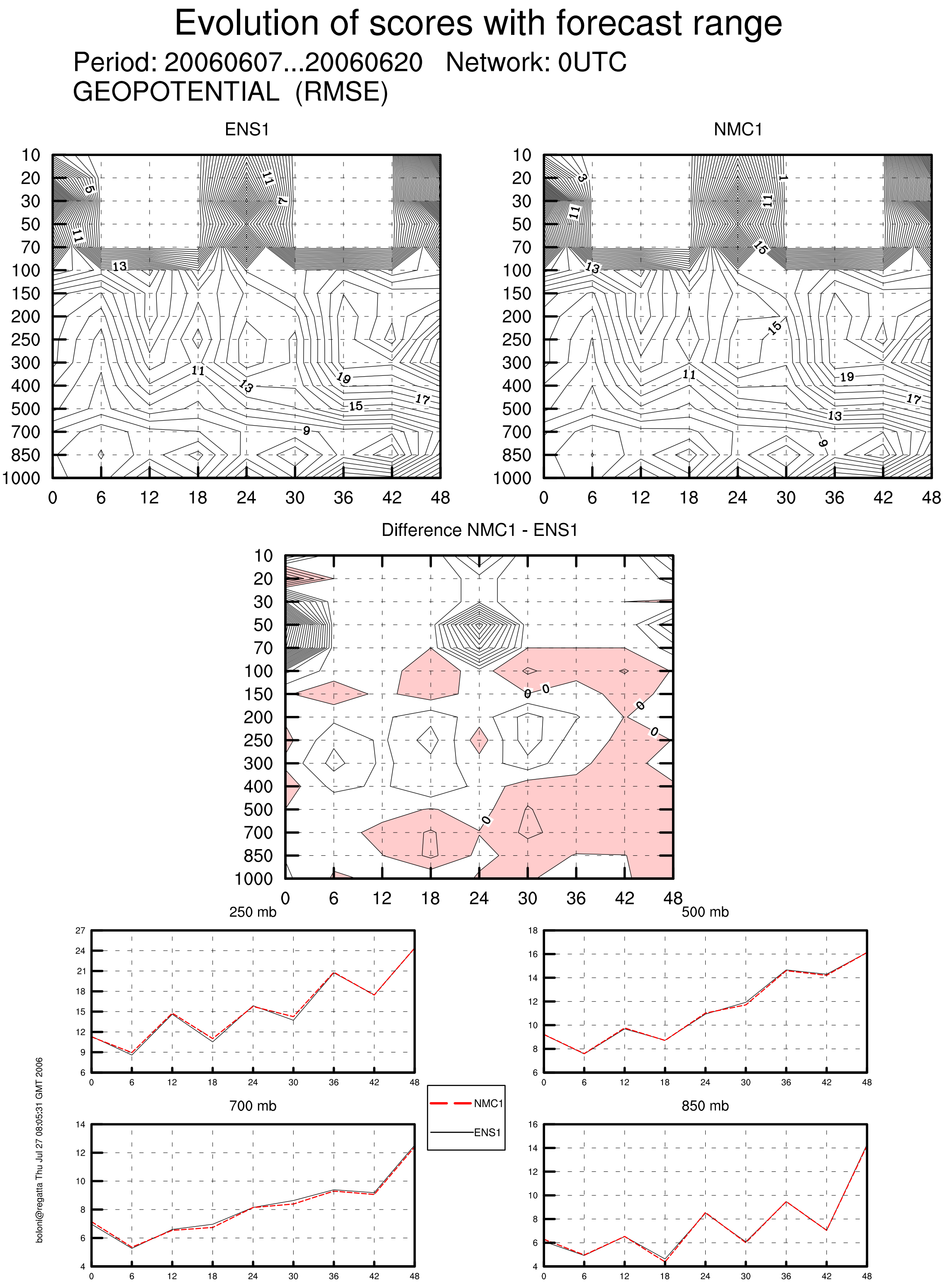

Figure 15: RMSE scores: geopotential NMC1: presently operational NMC B matrix, ENS1: tuning No.1

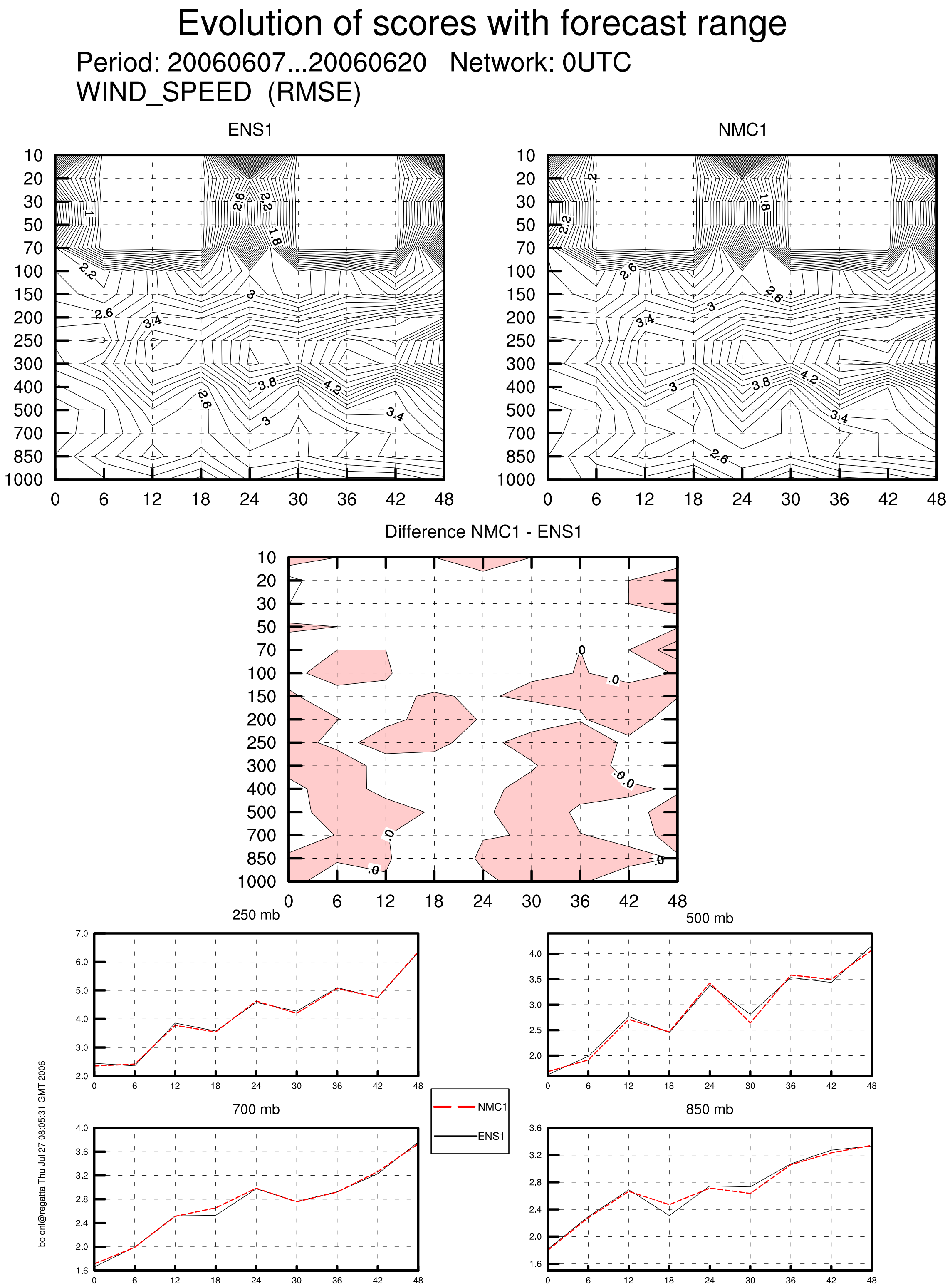

Figure 16: RMSE scores: wind speed NMC1: presently operational NMC B matrix, ENS1: tuning No.1

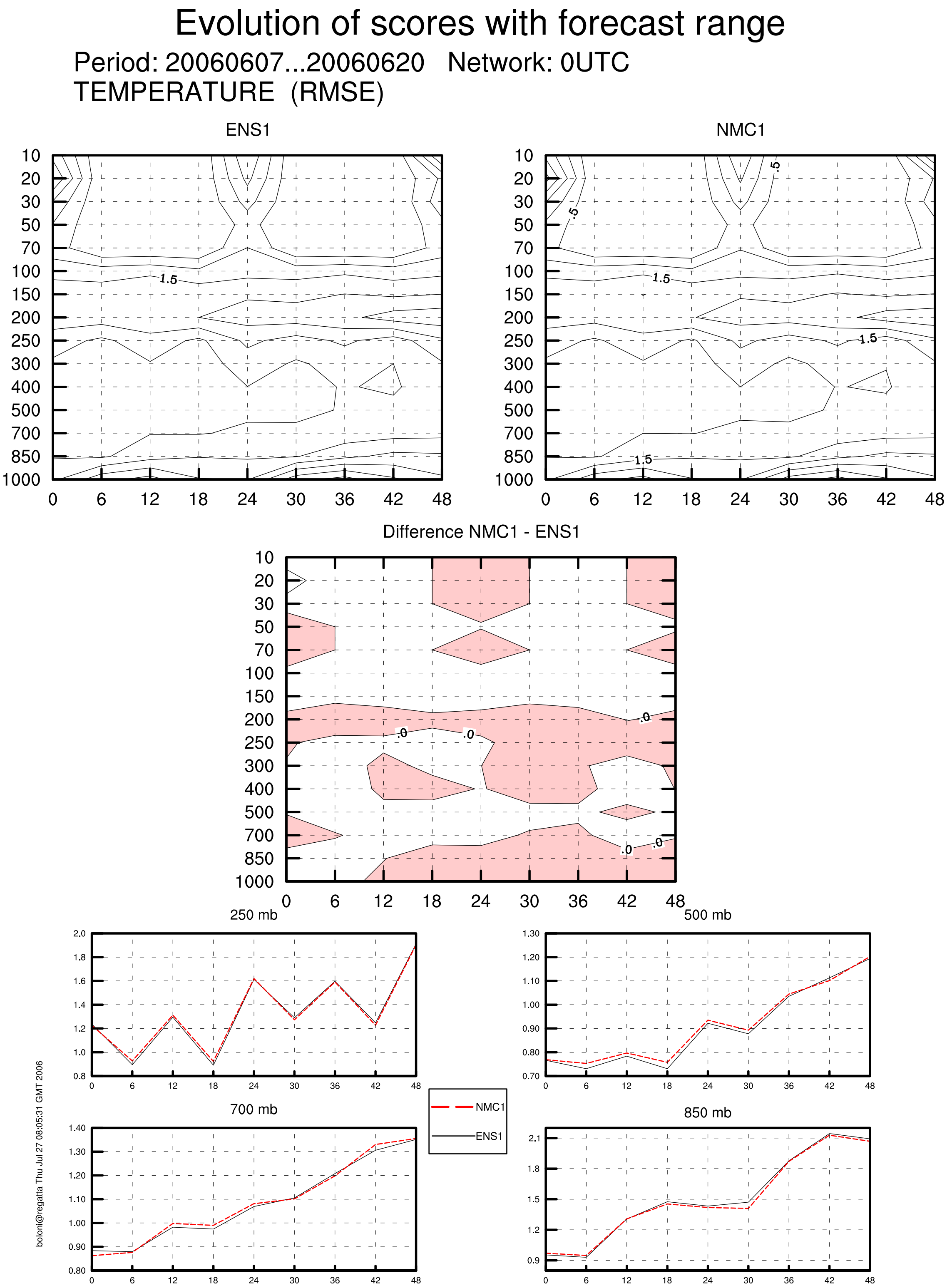

Figure 17: RMSE scores: temperature NMC1: presently operational NMC B matrix, ENS1: tuning No.1

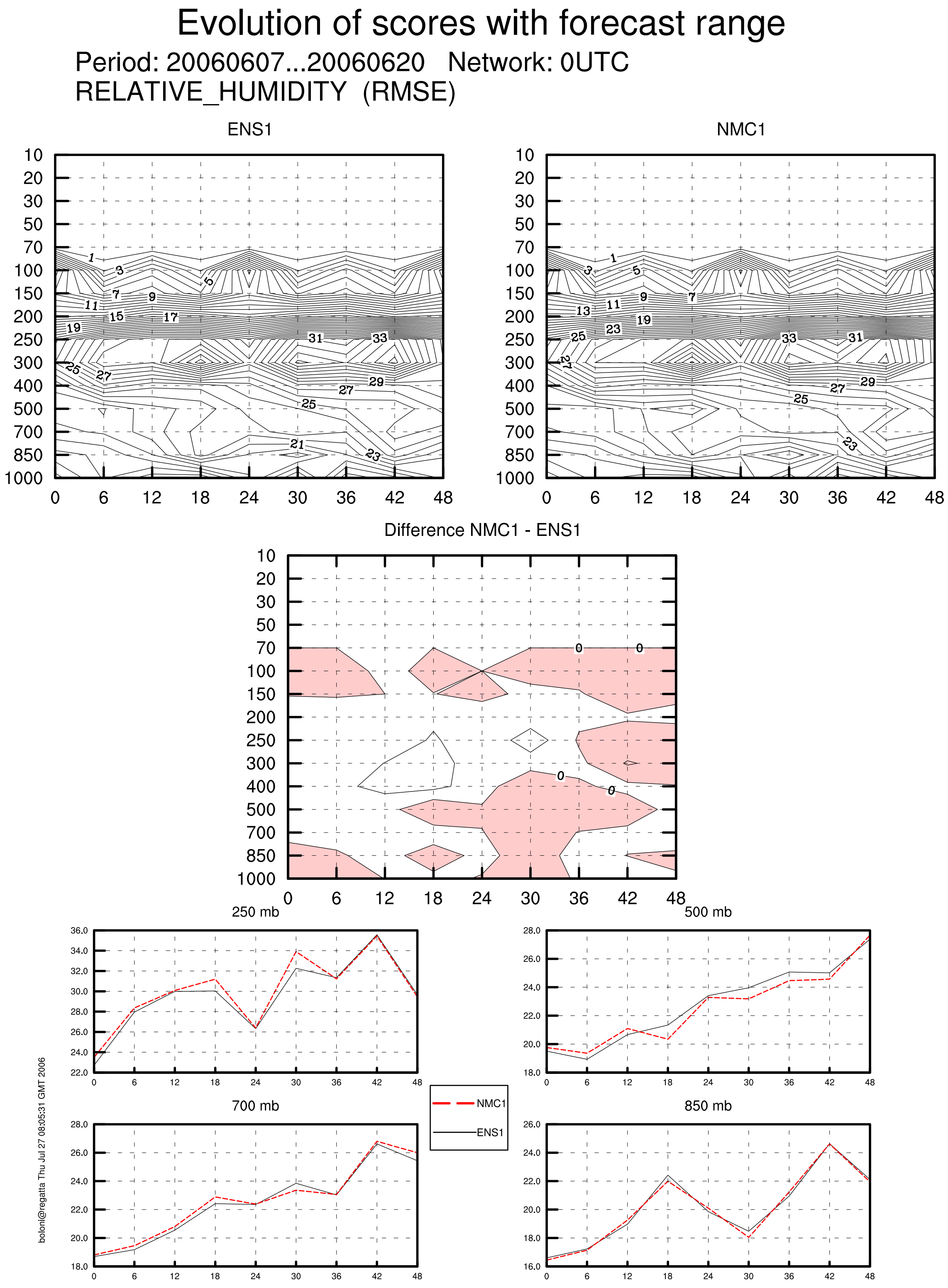

Figure 18: RMSE scores: relative humidity NMC1: presently operational NMC B matrix, ENS1: tuning No.1