Comparison of ensemble and NMC type of background error statistics for the ALADIN/HU model

Kristian Horvath
horvath@cirus.dhz.hr
Croatian Meteorological and Hydrological Service

supervised by
Bölöni Gergely
boloni.g@met.hu

RC LACE Internal Report
Oct 10 – Dec 5 2005

Summary

1. Introduction

In every data assimilation technique, information about the state of atmosphere introduced by the observations is spread in space. This propagation of the information gained from observation is in the data assimilation procedure mostly dictated by the **B** - background (usually 6h forecast) error covariance matrix.

Until a few years ago, many of the LAM data assimilation systems used NMC type of background error statistics. However, in last years there were a number of studies showing the benefit of using the ensemble type of **B** error covariance matrix in global models (e.g. Fisher, 1999; Belo Pereira and Berre, 2005). Namely, ensemble type of background error statistics inherently aligns with the fact that the spread of the ensemble model members due to change in initial conditions and subsequent model forecast errors are related and proportional. Moreover, it was shown (Berre et al., 2005), analytically and experimentally, that the NMC method relies on an inappropriate low-pass filter during each analysis step, in contrast with the analysis ensemble technique, which uses the exact high-pass filter of interest. Very recently, ensemble type of calculation was tested in LAM environment and Stefanescu et al. (2005) showed its impact in ALADIN/FR data assimilation set-up. Due to the positive effect shown, ensemble type of background error statistics was introduced in the operational ALADIN/FR set-up (Claude F., personal communication).

During this research stay, ensemble type of statistics was calculated from ARPEGE ensemble members. Upon that, it was diagnostically compared with the NMC type of statistics. Furthermore, comparison was done by the means of full-observation assimilation experiments, providing the verification scores for ensemble and standard NMC statistics for the chosen assimilation period.

2. Calculation of the ensemble statistics

The calculation of the ensemble statistics used 4 global ARPEGE ensemble analysis members that were produced by assimilating the perturbed observations. These members are denoted P31, P32, P36 and P37 further on in the text. All the members were downscaled to AL/HU resolution, and first-guesses were produced (+6h forecasts) throughout the chosen period Feb 04 – Mar 23 2002. The production starting time was chosen to be 06 UTC, giving forecasts at 12 UTC every day. In this way 4 sequences of +6h forecasts were available for the 48 days period.

Differences ALD_P31 - ALD_P32 and ALD_P36 - ALD_P37 (ALD denotes +6hr forecast) were produced; in this way a set of 96 differences was made, as an input to the festat program. This 96-forecast difference set was made in this way to be similar in length to a set of differences used for the calculation of the operational standard NMC statistics currently used in AL/HU.

Finally, background error statistics based on the above described ensemble forecast differences set was calculated.

Throughout the calculation, the operational AL/HU 49 level model version was used.

3. Diagnostic comparison of the ensemble and NMC type of statistics

It should be noted that ensemble and NMC type of statistics were calculated with different model versions. The NMC background error statistics were calculated with a 37 level AL/HU 28t1 model version with a 6.5 km horizontal resolution, while ensemble **B** was calculated with a 49 level AL/HU 28t3 model version at the same horizontal resolution. This fact should be taken into account when diagnostically comparing the different statistics. Also, it reflected the way the comparison was made, e.g. vertical profiles were always calculated on pressure levels, not on model levels as usual.
The period of sampling the differences needed for calculation of SNMC and LNMC statistics is May 2 – Aug 2 2004. Thus, there is a considerable difference in seasons for calculation of ensemble and NMC type of statistics that might have an influence on the comparison results.

3.1 Vertical profiles of standard deviation and correlation length-scales

In this part we will show the vertical profiles of standard deviations and correlation lengthscales. Beside ensemble (ENS_HU)and standard NMC vertical profiles the comparison includes the so-called lagged NMC (37 model levels) and French ensemble statistics (41 model levels).
Vertical profiles of standard deviations of vorticity, temperature, specific humidity and divergence are shown of Figure 1. For vorticity and temperature, ensemble standard deviation profiles are mostly between the standard and lagged NMC variants. The only exception is found at the middle levels for vorticity variable. On the other hand, for specific humidity ensemble standard deviation is much smaller, while for divergence somewhat bigger. The overall impression is that ensemble vertical profiles of standard deviation tend to be more constant with height. Maximas present around (below) tropopause in NMC statistics variants are much reduced in ensemble type of statistics.
French and Hungarian ensemble statistics are quite comparable for temperature and specific humidity, while some stronger differences are present in cases of vorticity and divergence. Ensemble vertical profile of standard deviation of temperature keeps the special feature of the AL/HU NMC statistics – it increases in the PBL with the maxima near ground. This feature is strictly AL/HU property, and is not present in AL/FR background error statistics.
Differences between vertical profiles of correlation lengthscales (Fig. 2) show rather irregular behaviour with respect to variable. The ensemble lengthscale of vorticity is smaller then both NMC variants, in contrast to specific humidity variable where ensemble lengthscales are bigger the NMC variants. Moreover, ensemble lengthscale of temperature is between standard and lagged NMC variants, while for divergence they are practically the same. AL/FR ensemble statistics has practically always greater lengthscales then the other statistics, probably due to somewhat lower resolution (9.5 km for AL/FR, 8 km for AL/HU).

3.2 Cross-covariance couplings

Cross-covariance couplings for ensemble and standard NMC statistics are shown on Figure 3. All the covariances are considerably reduced, especially for humidity variable. Structures of isolines in ensemble statistics stayed similar to NMC variants. Most pronounced differences are notable on average temperature (T-Pb) and humidity (q-Pb) covariances with vorticity balanced geopotential.

Average T-Pb covariances are characterised with two poles of opposite signs. The axes of the negative pole (if we imagine it as ellipsis) is tilted for approximately 45 in case of ensemble statistics, while in case of NMC statistics the axes is almost parallel with the abscise axes. This feature is similar for positive pole as well. It implies that in ensemble statistics vertical T-Pb correlations of other levels with the tropopause level are strongly reduced. In other words, ensemble vertical T-Pb correlations tend to be more local then in case of standard NMC statistics.

Average covariance of humidity at low levels and vorticity balanced geopotential in ensemble statistics tend to be much more complicated then the corresponding NMC covariance. Dominant feature of low level humidity coupling with balanced geopotential in case of ensemble statistics is a sequence of maximas and minimas, not present in NMC statistics to such an extent. At the moment, there is no clear understanding of this type of a covariance character.

3.3 Ratios of explained variables

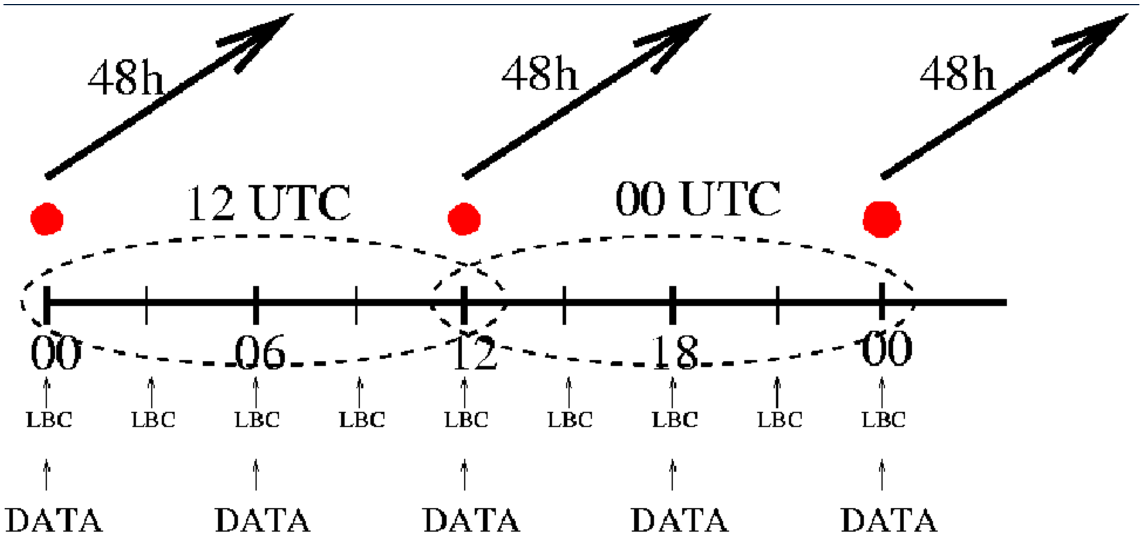
The structure functions of the ratios of explained variables are similar in both statistics (Fig. 4). Balanced geopotential tends to explain the other variables by the “large” scale couplings, while other couplings tend to be more expressed at ”middle” and “small” scales. All ratios are reduced in ensemble statistics, except the ratio of humidity explained by unbalanced divergence. The greatest reduction is present in the ratio of humidity explained by the balanced geopotential, where the reduction is of the order of factor 4. Other reductions are not reaching factor 1.5 (i.e. ensemble values are not smaller then the 2/3rd of the standard NMC value). In general, it seems that large-scale couplings are reduced more than the other ones.

3.4 Average vertical correlations

For average vertical correlations (Fig. 5) exceptionally, statistics files of the background error calculations performed on 49 level AL/HU model version were available. Differences are quite small except for temperature, where ensemble mean vertical correlations are much sharper. This is in line with the structural change of T-Pb covariance (Sec. 3.2), where smaller vertical correlations with tropopause level were noticed. Humidity correlations are somewhat sharper throughout all the troposphere. Divergence and vorticity correlations seem to be a bit broader at the lower and sharper in the upper levels, but the differences do not seem significant.

4. Full-observation experiments

A full observation experiment using SYNOP, TEMP, AIREP and AMSU-A satellite data was performed for the period 26 Oct 2005 – 10 Nov 2005. The experiment included the assimilation cycle with the standard NMC statistics operationally used at HMS and the newly calculated ensemble statistics. In the experiment 1, assimilation cycle was run with statistics as they are (REDNMC=1.0 for both statistics). In the experiment 2, background error standard deviation scaling factors were chosen subjectively (using experience in AL/FR)



Picture 1. Operational data assimilation cycle at HMS (plot provided by Bölöni Gergo), where "DATA" stands for

a) case: 12 UTC start production time - L+000 (00 UTC), L+003 (03 UTC), L+000 (06 UTC) and L+000 (06 UTC), L+003 (09 UTC), S+000 (12 UTC)
b) case: 00 UTC start production time - L+000 (12 UTC), L+003 (15 UTC), L+000 (18 UTC) and S+000 (18 UTC), S+003 (21 UTC), S+006 (24 UTC),

where L+0?? denotes long cut-off ARPEGE analysis and forecast and S+0?? short cut-off one.

to be:

- REDNMC(SNMC)=1.0
- REDNMC(ENSB)=1.5

The model set-up approximated the operational version run at HMS reasonably close as regards to the model, assimilation cycle and the initial and lateral boundary conditions (Pic. 1).

4.1 Experiment 1 – REDNMC(both statistics)=1.0

Thus, in the first experiment scaling factor REDNMC was kept equal to 1.0. Root mean square errors and biases for this experiment are shown on Figure 6. For geopotential ensemble scores are in general slightly negative compared to the SNMC– both for RMSE and bias. It is interesting to notice that in the upper troposphere and at tropopause level the ensemble analysis is shifted towards the background (compared to the SNMC), whilst in the lower and middle troposphere the ensemble analysis is shifted towards the observations. However, at +06 forecast range in the upper troposphere and at the tropopause level where ensemble analysis was shifted towards background, "ensemble forecast" is better (closer to the observations) then the "SNMC forecast". Similarly, in the lower and middle troposphere where ensemble analysis was shifted towards observations (compared to the SNMC analysis), ensemble +6h scores are worse (further away from the observations) then the SNMC +6h scores. This is a nice example of the nonlinearity of the model trajectories.

Relative humidity RMSE scores are positive, throughout the atmosphere. The shift of analysis towards observations at high levels seems to be a positive feature, because at those levels analysis knows to be even further away from observations then the background, due to vertical correlations (Gergo B., personal communication). Thus, although the standard deviation of humidity is reduced in ensemble statistics (what means shift of analysis towards background), vertical correlations overcome this effect and bring the analysis towards the observations. This implies that vertical correlations are significantly changed in ensemble and NMC type of statistics, and that tuning of the standard deviation of humidity in NMC statistics might be insufficient to achieve the proper cross-variable propagation. The bias scores for relative humidity overall seem to be a little bit worse.

Temperature RMSE scores tend to be rather mixed, while bias is clearly somewhat worse. One of the reasons that could be responsible for this feature is that the bias correction of the satellite data. Namely, a part of the bias correction accounts for correcting the bias of the radiative transfer (RT) model, which creates "model" radiances from the background field. Consequently, the bias correction depends on the quality of the background and by changing the statistics, the quality of the background was changed (Roger R., personal communication). Therefore, the bias correction needs to be recalculated for the assimilation cycle with ensemble statistics. We should also notice that temperature structure functions are rather strongly changed (compared to changes in other variables), e.g. regarding vertical auto-correlations (Fig. 5) or T-Pb cross-covariances (Fig. 3). Thus, it might simply be the feature of the statistics as well. Changes in the wind direction and wind speed are mostly neutral or slightly positive (e.g. wind direction RMSE).

4.2 Experiment 2 – REDNMC(ENSB)=1.5; REDNMC(SNMC)=1.0

Root mean square error and bias for this experiment are shown on Figure 7. With subjectively tuned ensemble standard deviations, the overall slightly negative impact untuned ensemble statistics had on geopotential (compared to the SNMC) is lost, and both statistics now show similar results. The only exception is a bias at high levels where ensemble results are still a bit worse.

The RMSE scores of relative humidity in experiment with tuned ensemble statistics are slightly worse when compared to the untuned ensemble statistics experiment, but still show a positive impact compared with standard NMC statistics experiment result. The bias is similar in tuned and untuned ensemble experiments (overall slightly worse then in SNMC). Effect of tuning on temperature RMSE is slightly positive when comparing the two ensemble statistics experiments, except in middle troposphere, where the effect slightly worse, while bias in general seems to be better in tuned then in the untuned ensemble experiment. However, in the middle troposphere (500hPa, 700hPa) it is still significantly worse then in the standard NMC experiment. In future, the experiments with the new bias correction for the satellite data will try to account for this negative feature.

5. Conclusions

Ensemble type of background error statistics was calculated from ARPEGE ensemble members and diagnostically compared with the NMC type of statistics. Furthermore, full-observation experiments were performed, enabling comparison between statistics in terms of verification scores.

Ensemble standard deviations of different variables are in general reduced compared to standard NMC statistics, the greatest reduction present regarding the specific humidity. The exception is divergence variable, where the ensemble standard deviation is increased compared to SNMC variant. Effect of ensemble statistics on correlation length-scales is quite variable. Probably the most noteworthy change is increase in the correlation length-scale of specific humidity. Cross-covariance couplings for ensemble statistics are considerably reduced, especially for humidity variable. Beside that, it seems that the most pronounced effect of ensemble statistics is a more local temperature-balanced geopotential coupling. Percentage of explained variables is reduced in ensemble type of statistics, except for the humidity variable explained by divergence. In general, it seems that the change mostly modified the large-scale couplings. The most significant difference in average vertical (auto)correlations is visible on temperature variable, where the correlations are strongly reduced.

The full-observation experiments showed the most prominent effects of ensemble statistics on relative humidity and temperature scores. The RMSE of relative humidity in ensemble statistics experiments is improved compared to the SNMC statistics, even a little bit more in the untuned experiment. Furthermore, the analysis of relative humidity at high levels is closer to observations in ensemble type of statistics. On the other hand, a strong negative effect on temperature bias is present in both ensebmle statistics experiments when compared to SNMC statistics experiments. It seems that this negative bias is the worst significant negative feature of the ensemble statistics and it needs to be investigated more. Some near future experiments with the new bias-correction for the satellite data will try to account for that.

Literature:

Belo Pereira, M., and L. Berre, 2005: The use of an ensemble approach to study the background error covariances in a global NWP model. Accepted in Mon. Wea. Rev.

Berre, L., Stefanescu, S.E., Belo Pereira, M., 2005: The representation of the analysis effect in three error simulation techniques. Accepted in Tellus.
Fisher, M., 1999: Background Error Statistics. derived from an ensemble of analyses. ECMWF Research Department Technical. Memorandum, **79**, 12pp.
Stefanescu, S.E., L. Berre, and M. Belo Pereira, 2005: The evolution of dispersion spectra and the evaluation of model differences in an ensemble estimation of error statistics for a limited area analysis. Accepted in MWR.

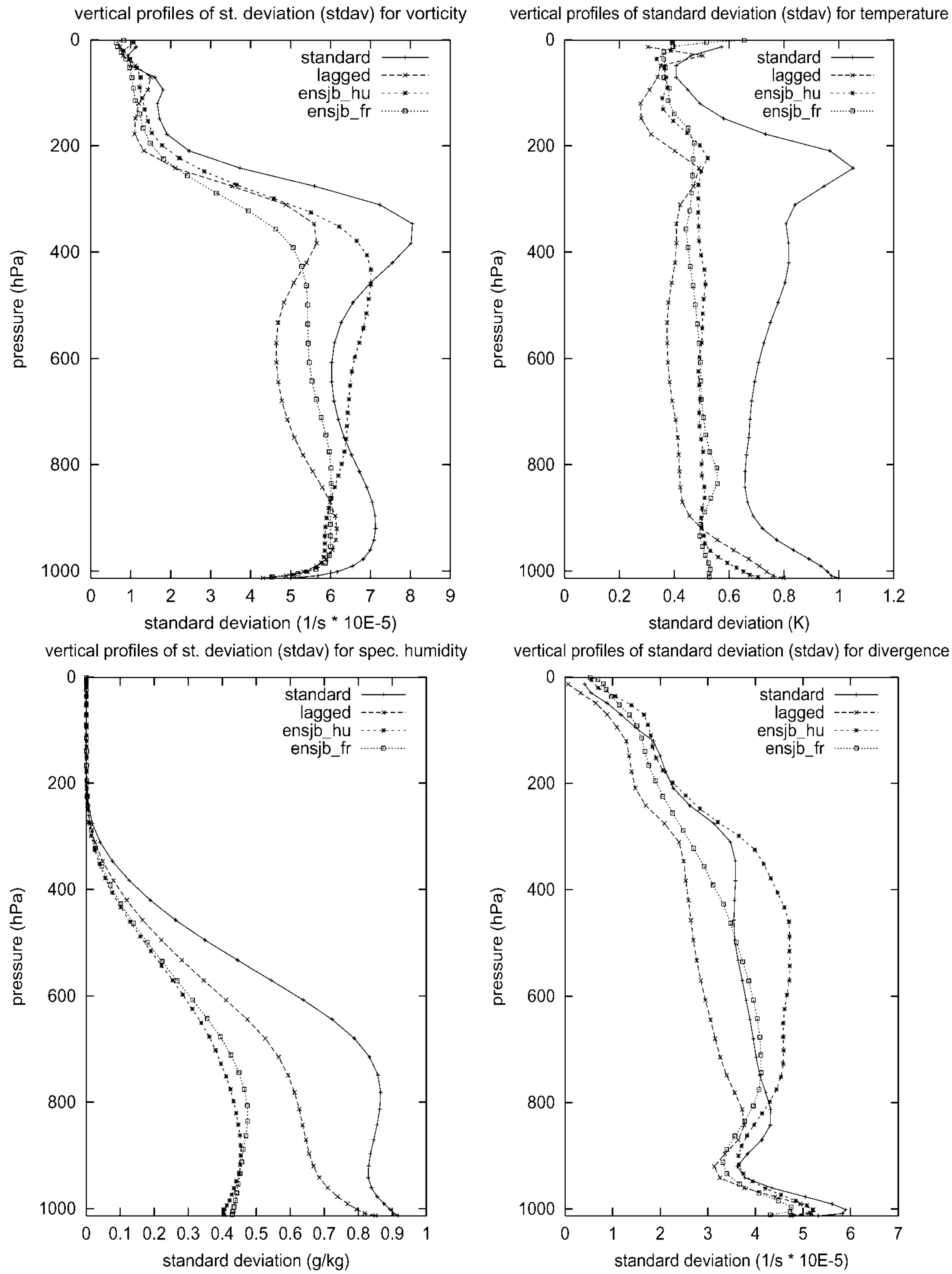


Figure 1. Vertical profiles of the standard deviation for balanced geopotential (upper-left UL), temperature (upper-right UR), specific humidity (lower-right LR) and divergence (lower-left LL).

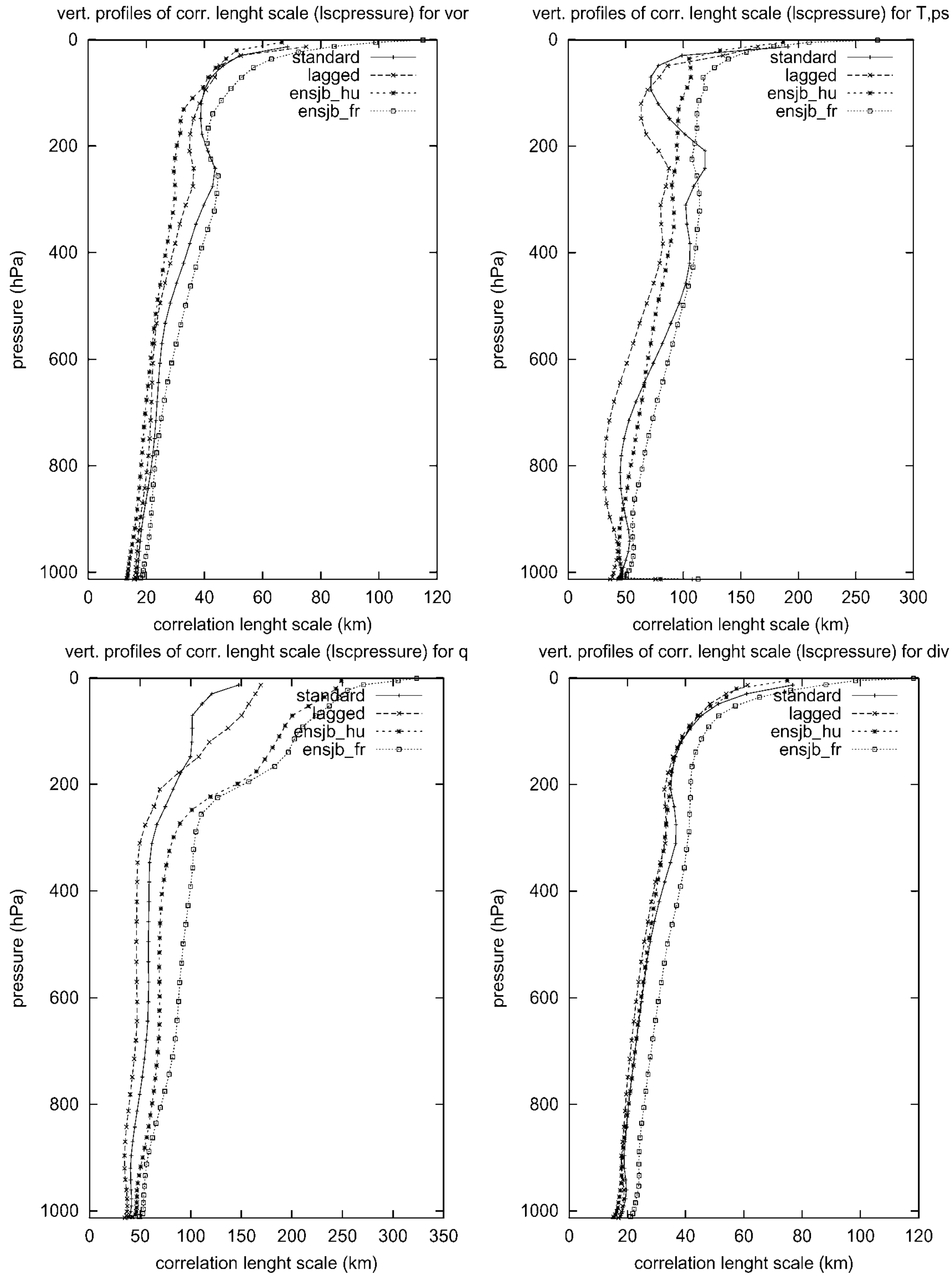


Figure 2. Vertical profiles of the correlation length-scales for balanced geopotential (UL), temperature (UR), specific humidity (LR) and divergence (LL).

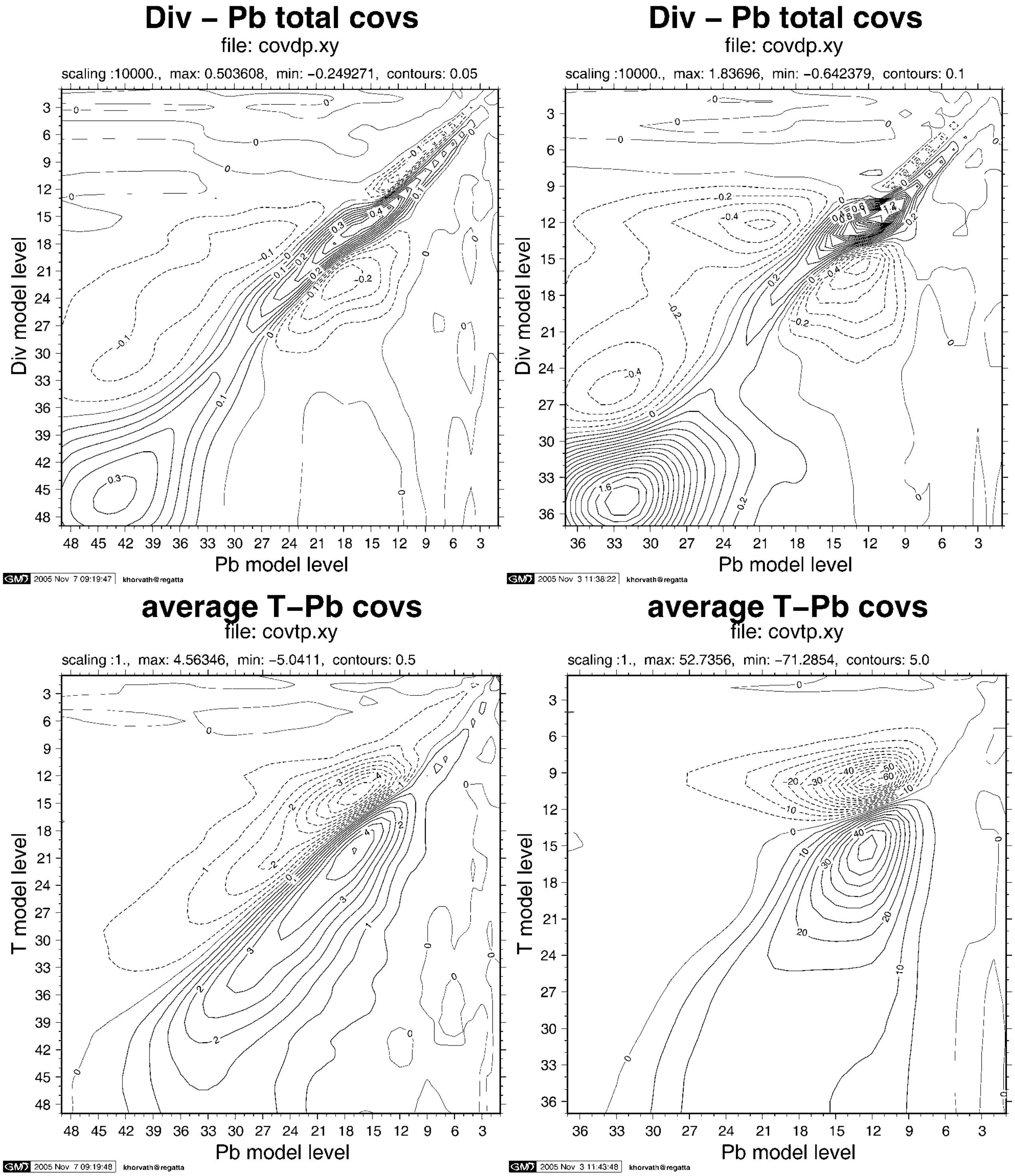


Figure 3. Ensemble (left) and standard NMC (right) mean vertical cross-covariance matrices between divergence and balanced geopotential (up) and temperature and balanced geopotential (down)

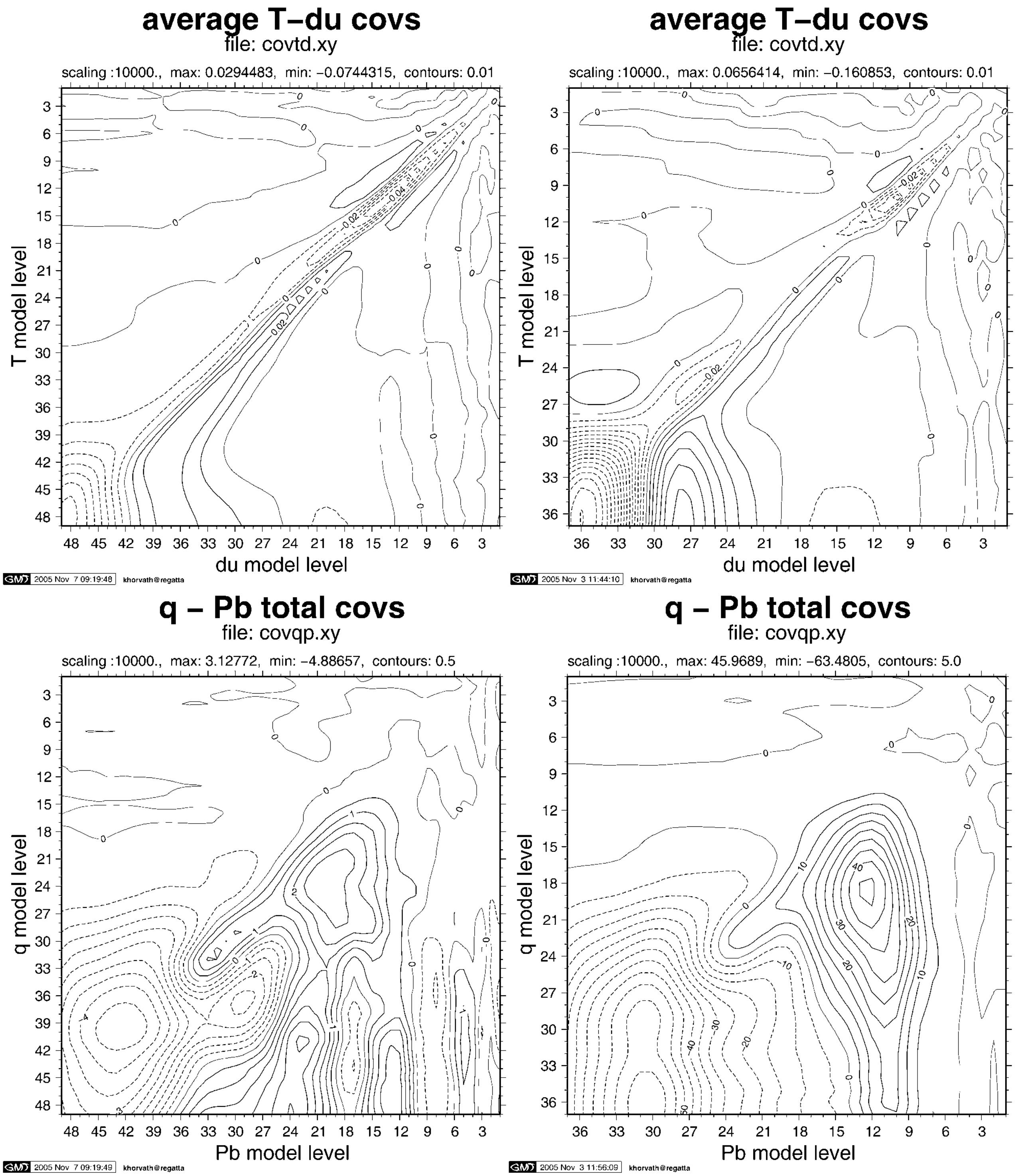


Figure 3. (continued) Ensemble (left) and standard NMC (right) mean vertical cross-covariance matrices between temperature and unbalanced divergence (up) and humidity and balanced geopotential (down)

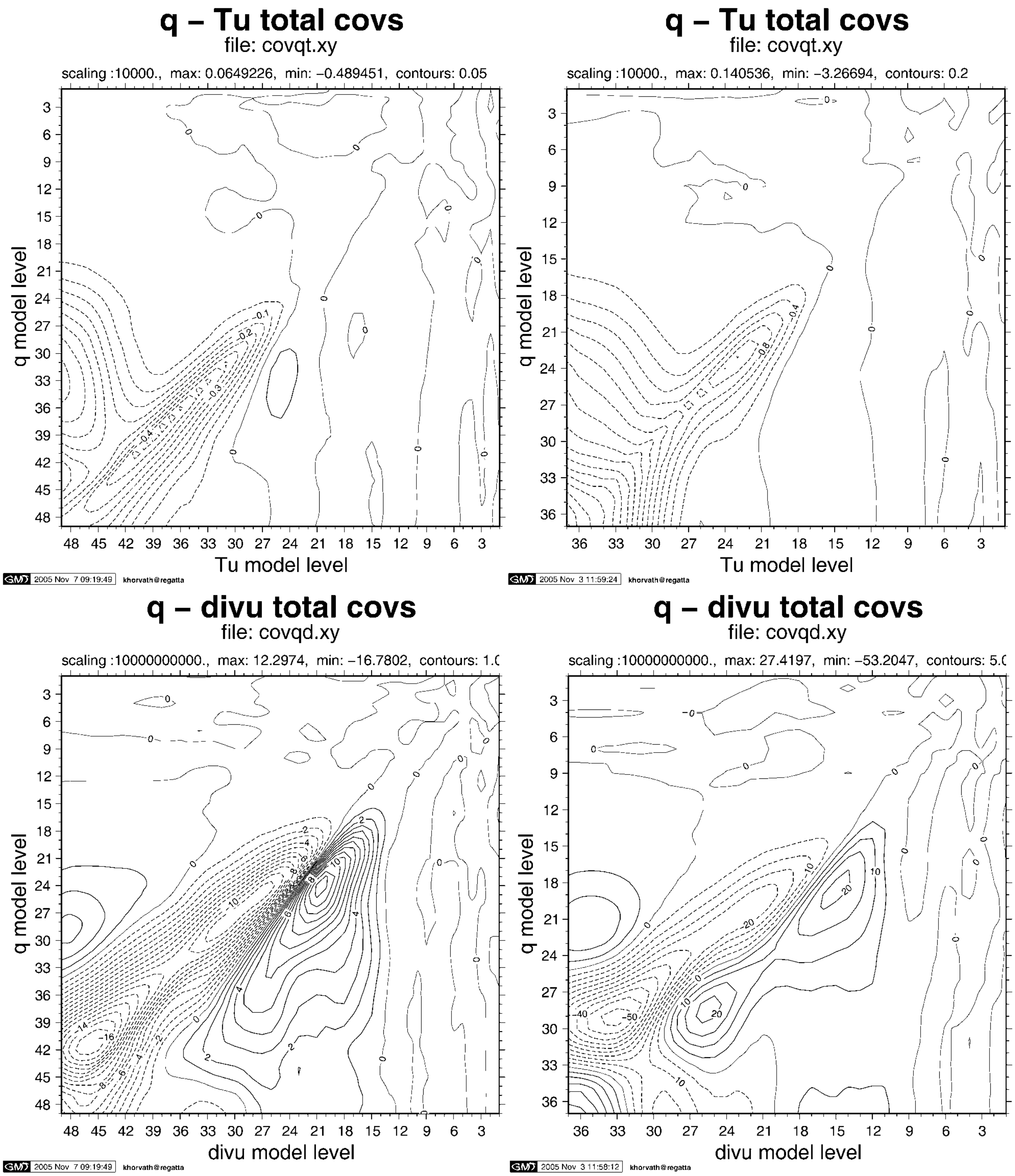


Figure 3. (continued) Ensemble (left) and standard NMC (right) mean vertical cross-covariance matrices between humidity and unbalanced temperature (up) and humidity and unbalanced divergence (down)

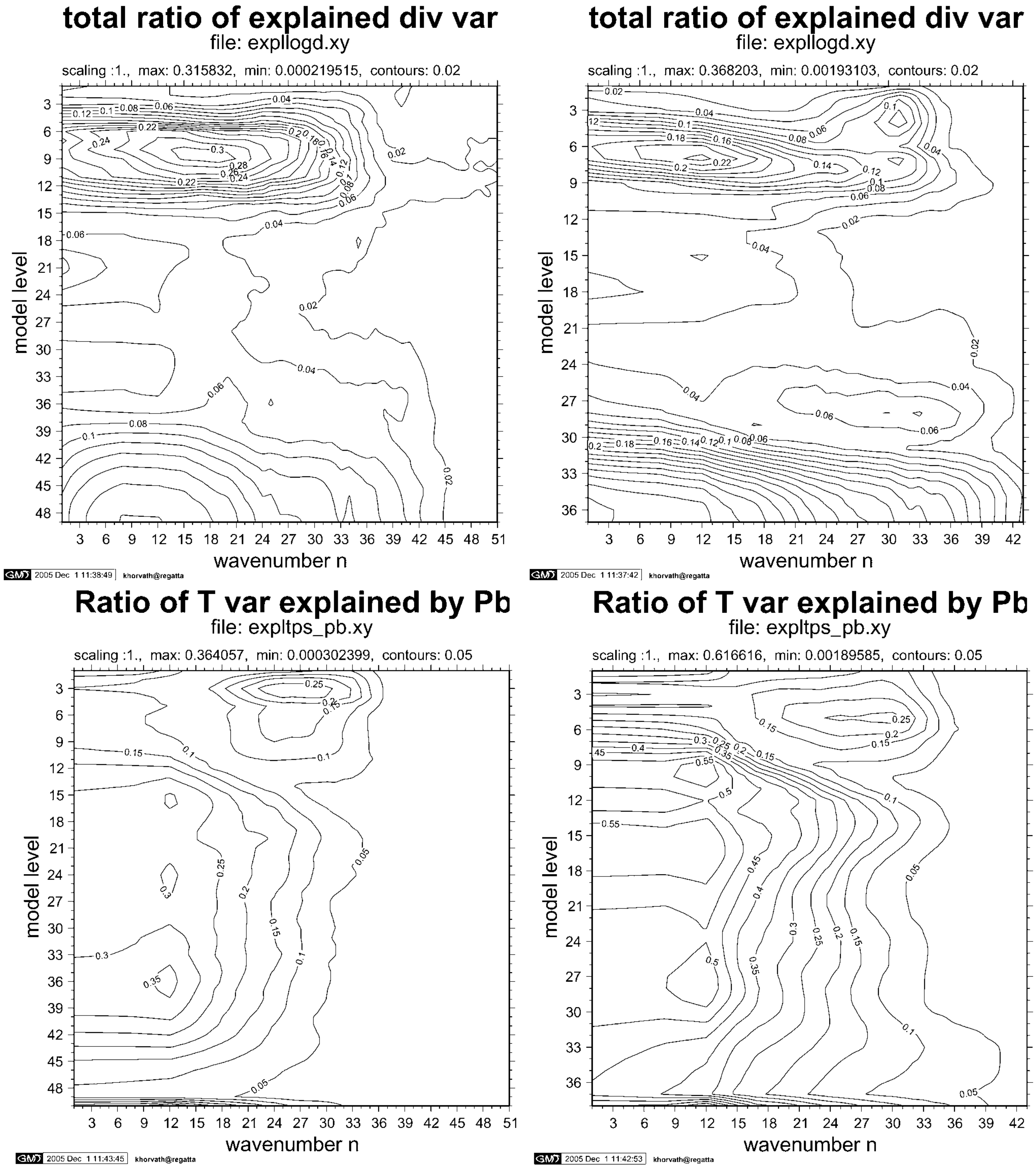


Figure 4. Ensemble (left) and standard NMC (right) percentages of the variance of divergence (up) and temperature (down) explained by balanced geopotential.

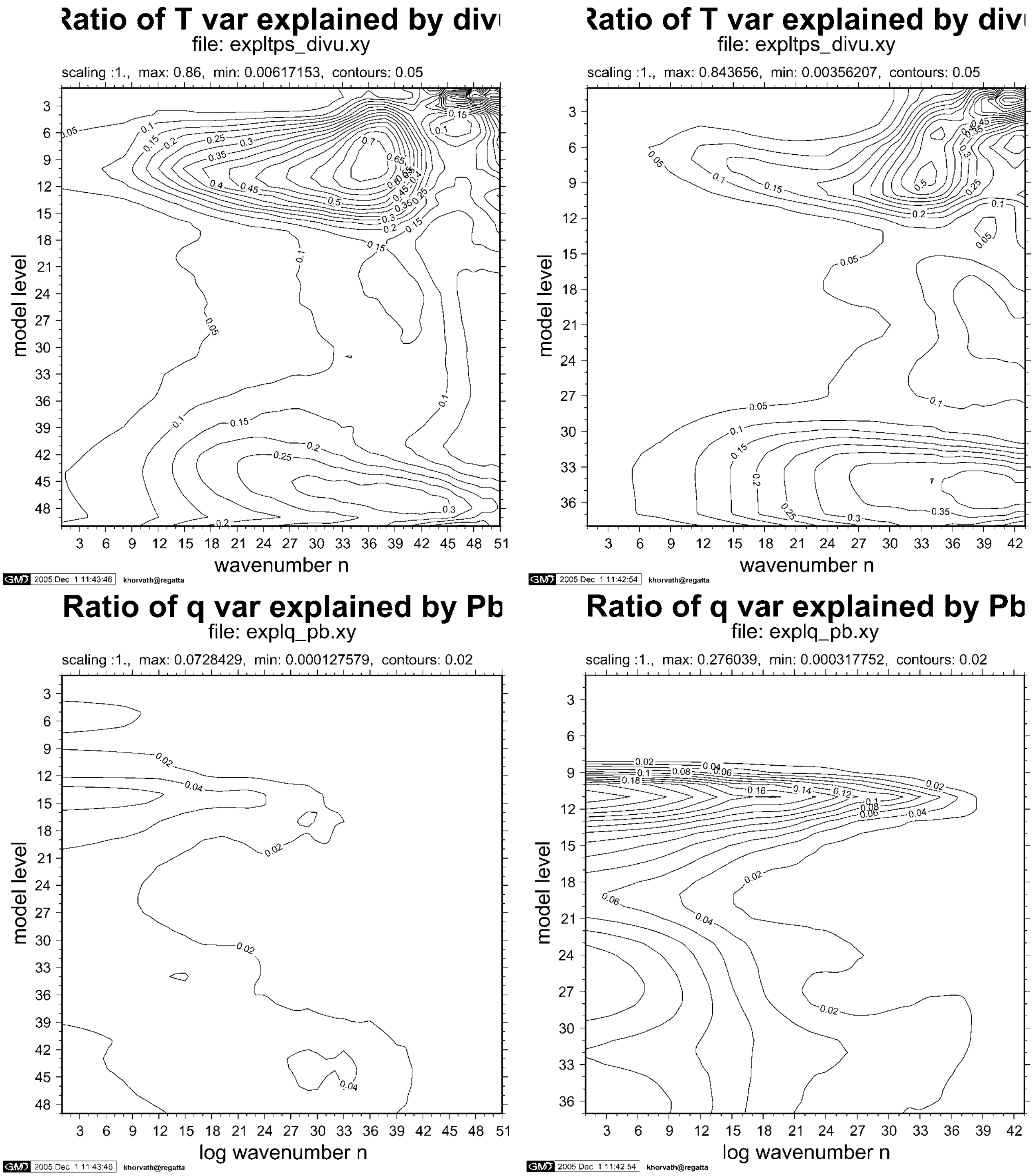


Figure 4. (continued) Ensemble (left) and standard NMC (right) percentages of the variance of temperature explained by unbalanced divergence (up) and specific humidity explained by balanced geopotential (down).

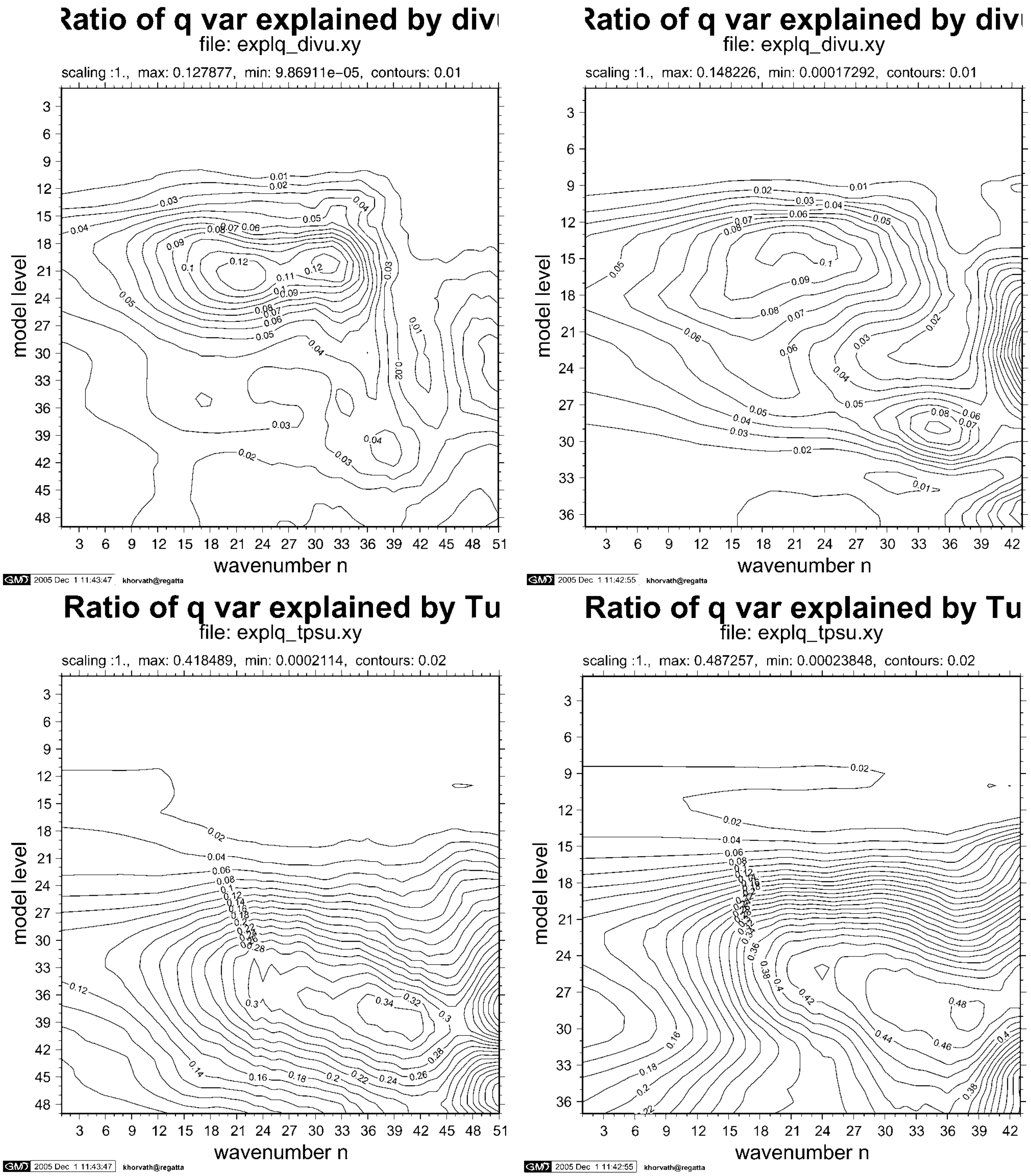


Figure 4. (continued) Ensemble (left) and standard NMC (right) percentages of the variance of specific humidity explained by unbalanced divergence (up) and unbalanced temperature (down).

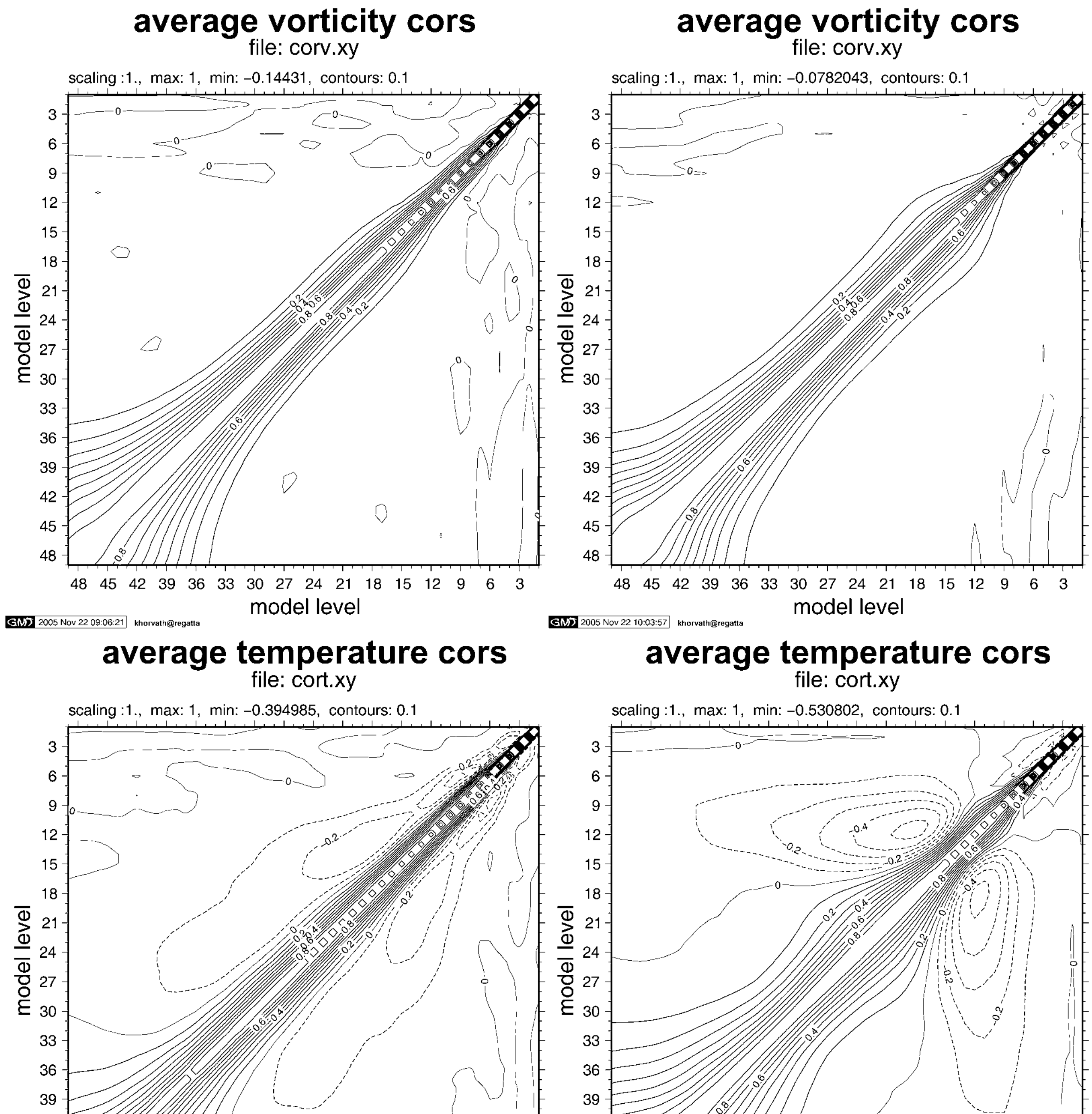


Figure 5. Ensemble (left) and standard NMC (right) average vertical correlations for vorticity (up) and temperature (down)

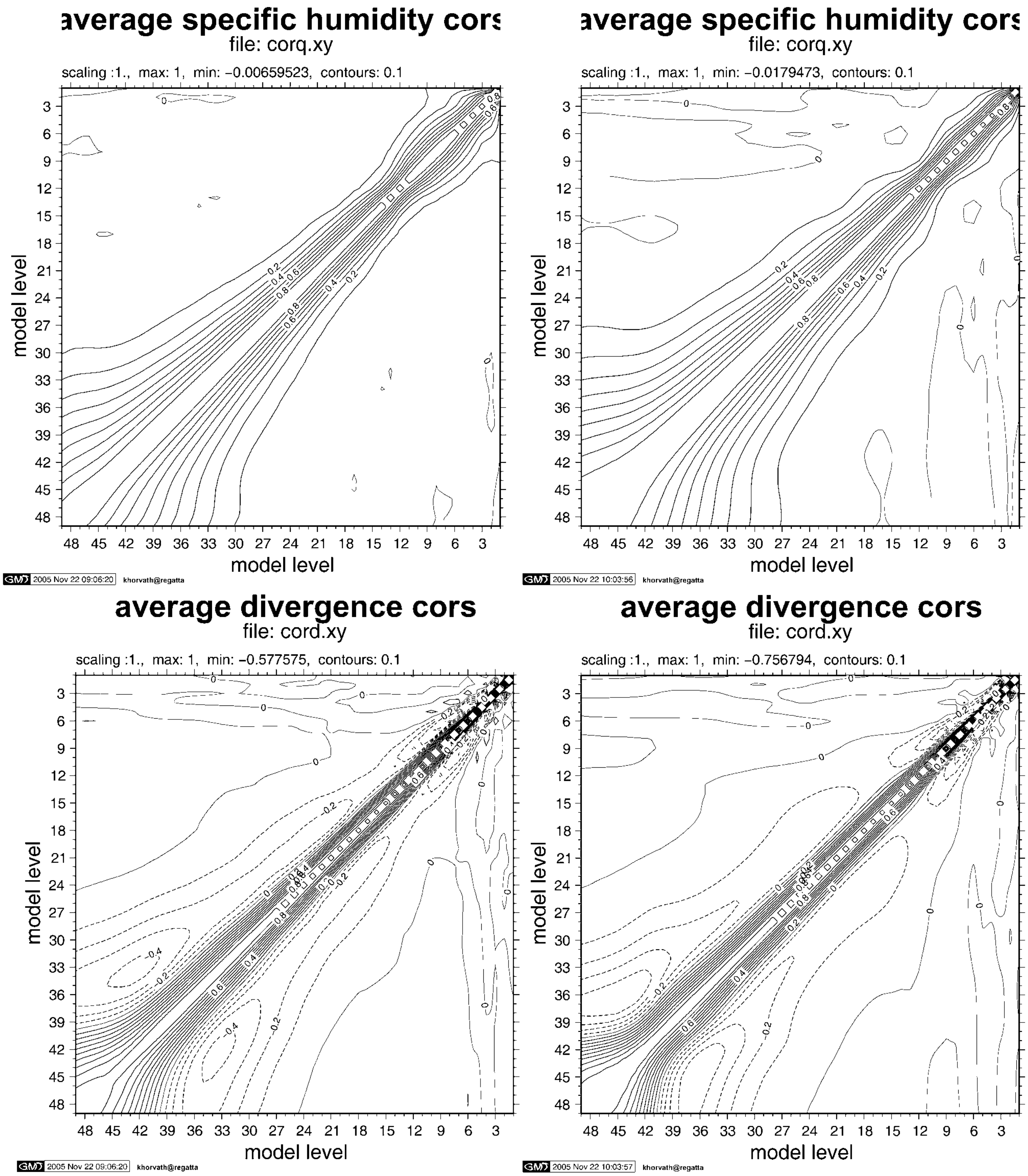


Figure 5. (continued) Ensemble (left) and standard NMC (right) average vertical correlations for specific humidity (up) and divergence (down)

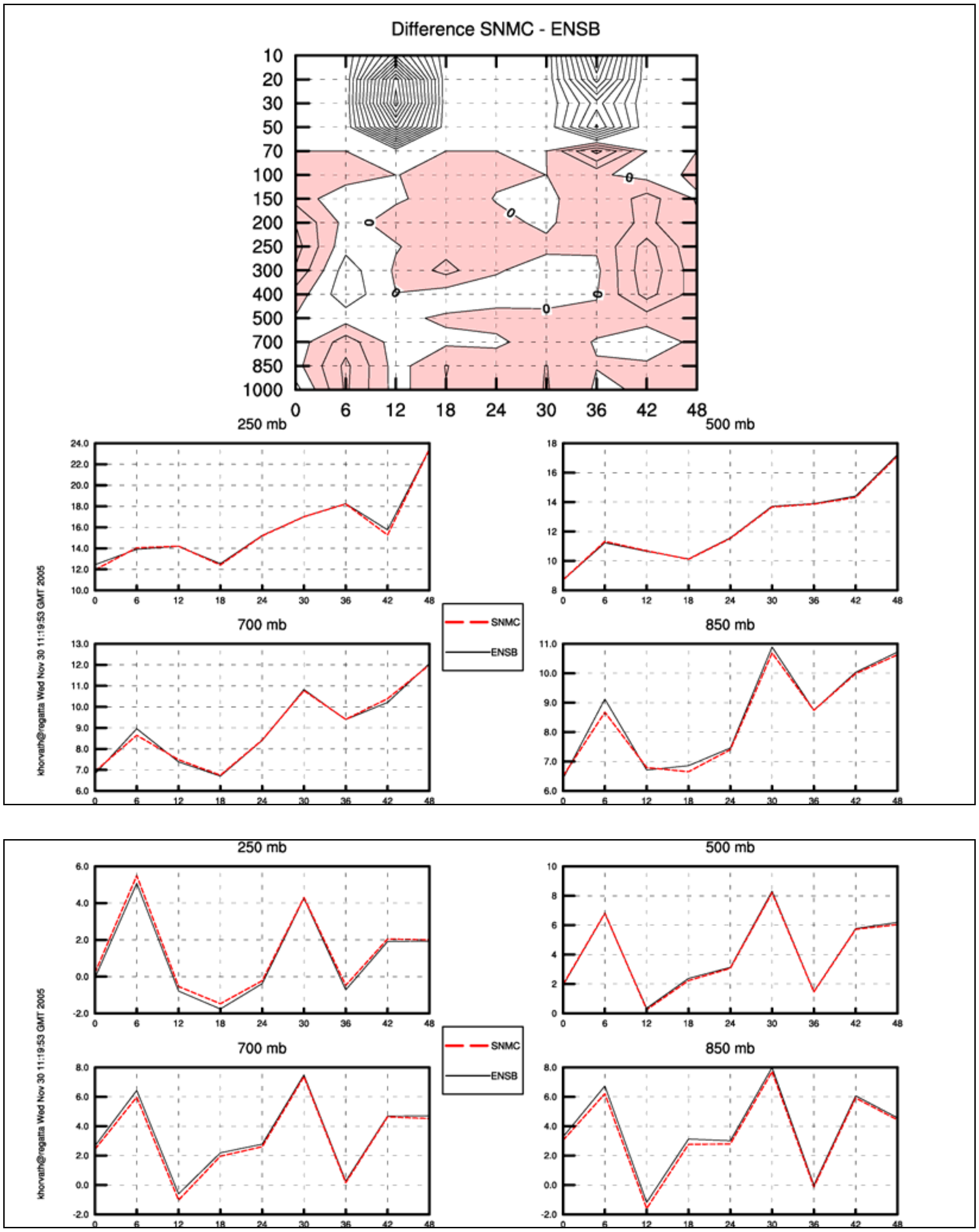


Figure 6. Root mean square error (up) and bias (down) of geopotential in full observation experiment using REDNMC=1.0 for both ensemble and standard NMC statistics.

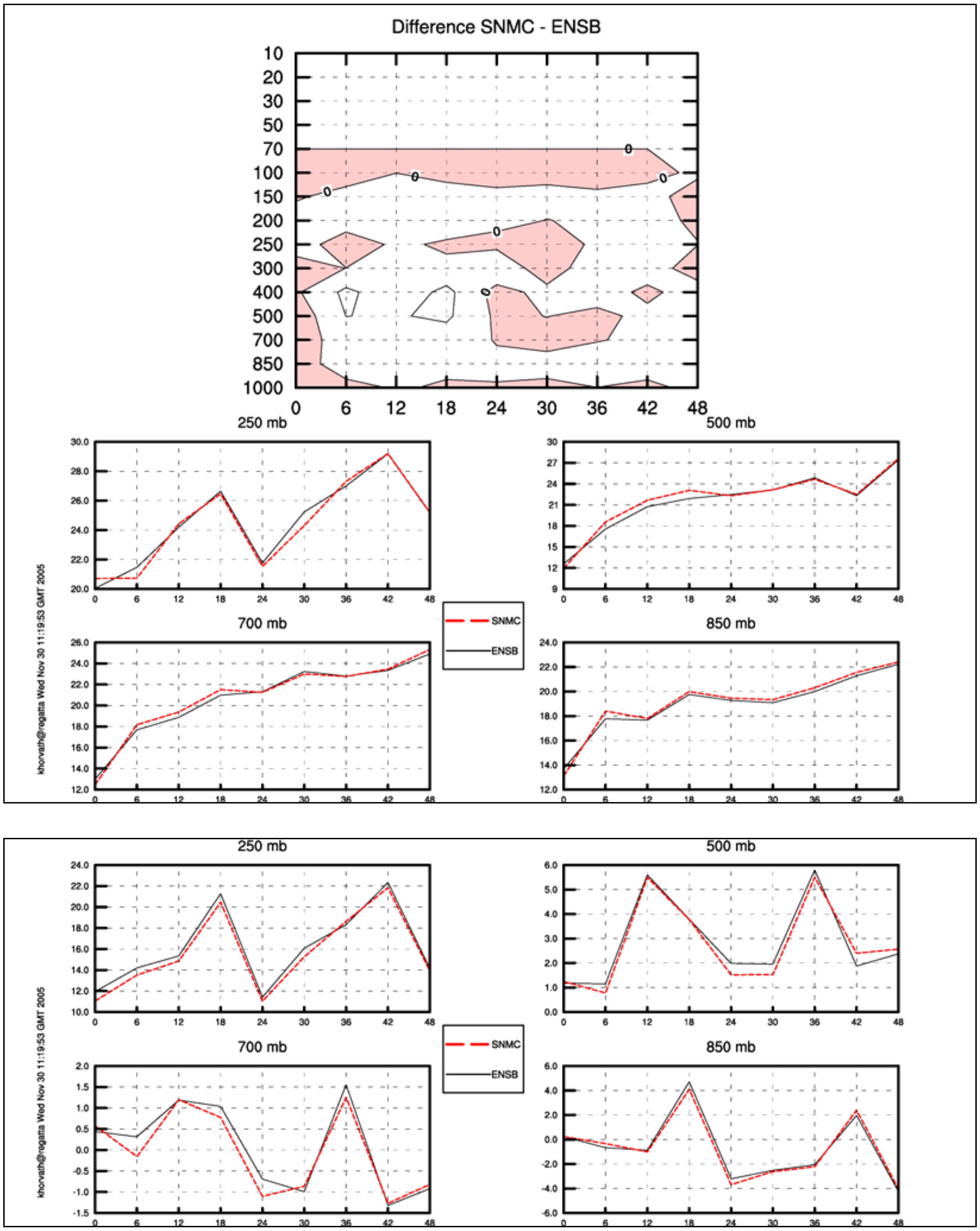


Figure 6. (continued) Root mean square error (up) and bias (down) of relative humidity in full observation experiment using REDNMC=1.0 for both ensemble and standard NMC statistics.

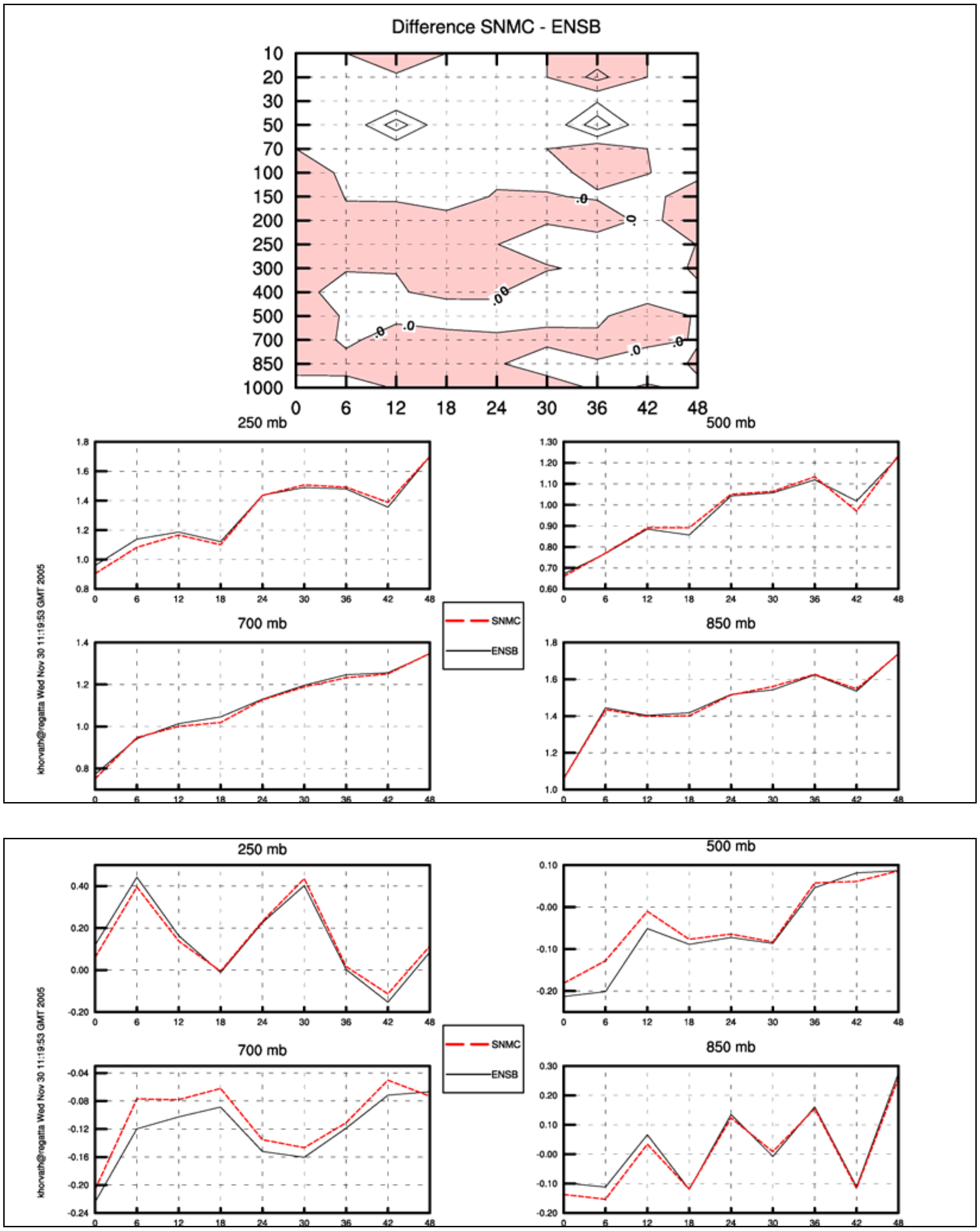


Figure 6. (continued) Root mean square error (up) and bias (down) of temperature in full observation experiment using REDNMC=1.0 for both ensemble and standard NMC statistics.

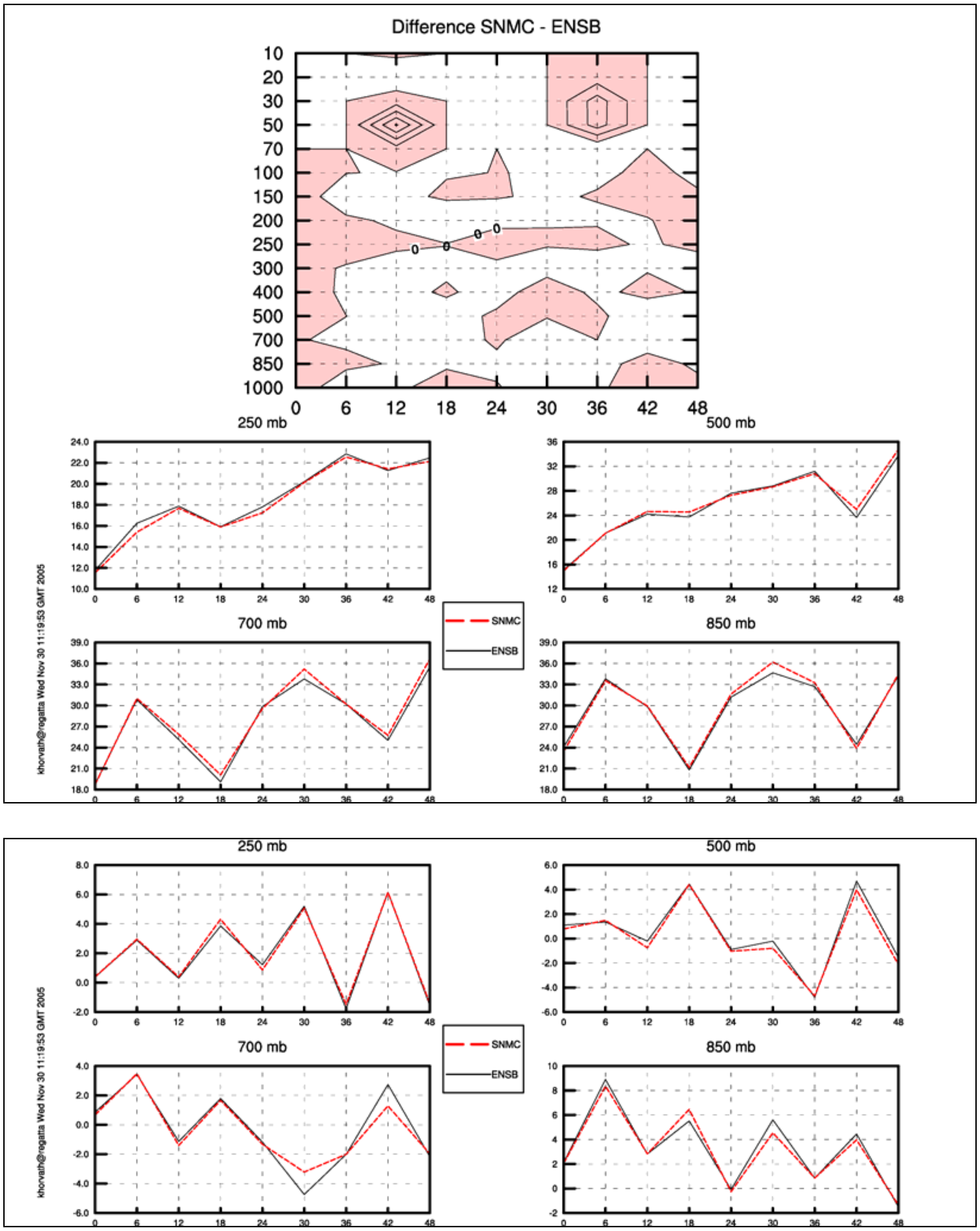
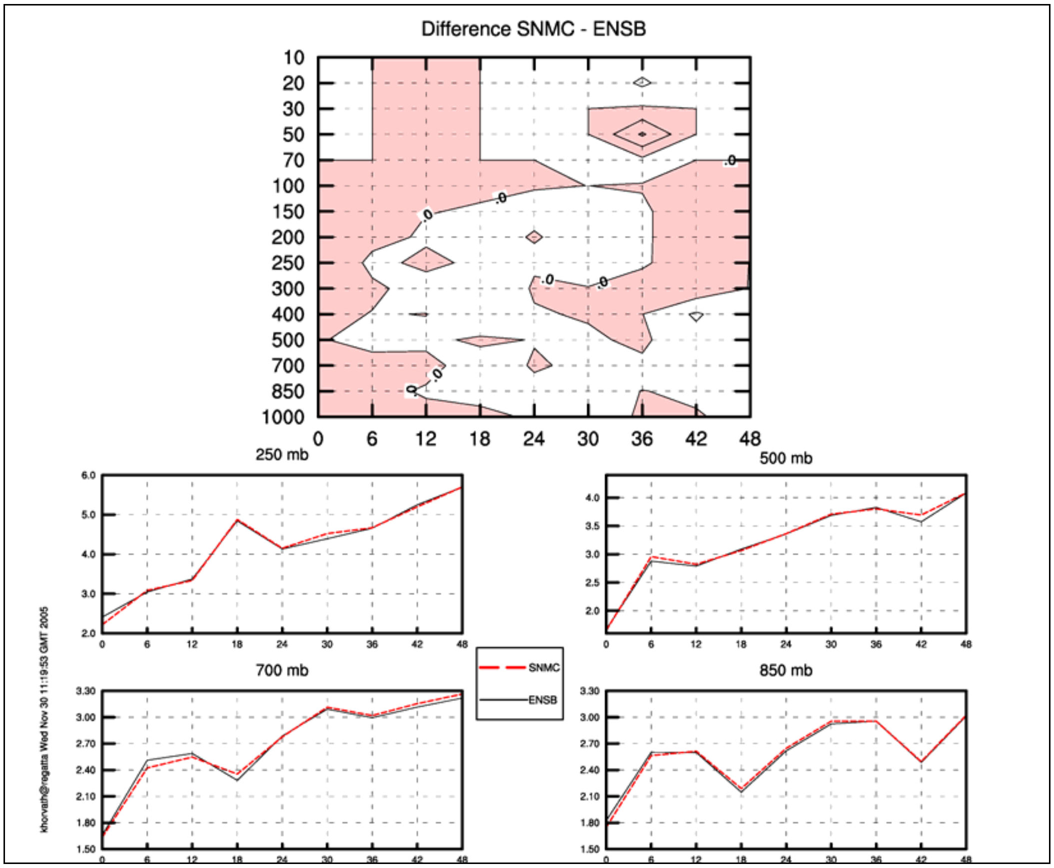


Figure 6. (continued) Root mean square error (up) and bias (down) of wind direction in full observation experiment using REDNMC=1.0 for both ensemble and standard NMC statistics.



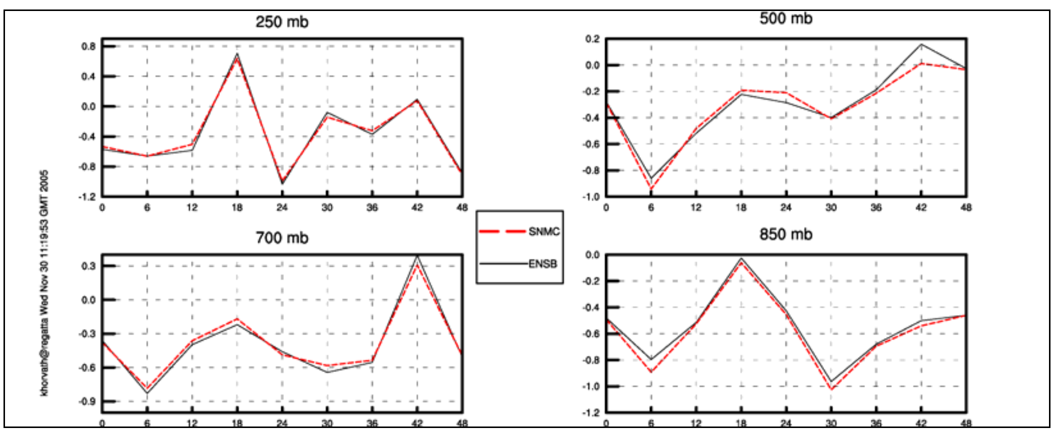
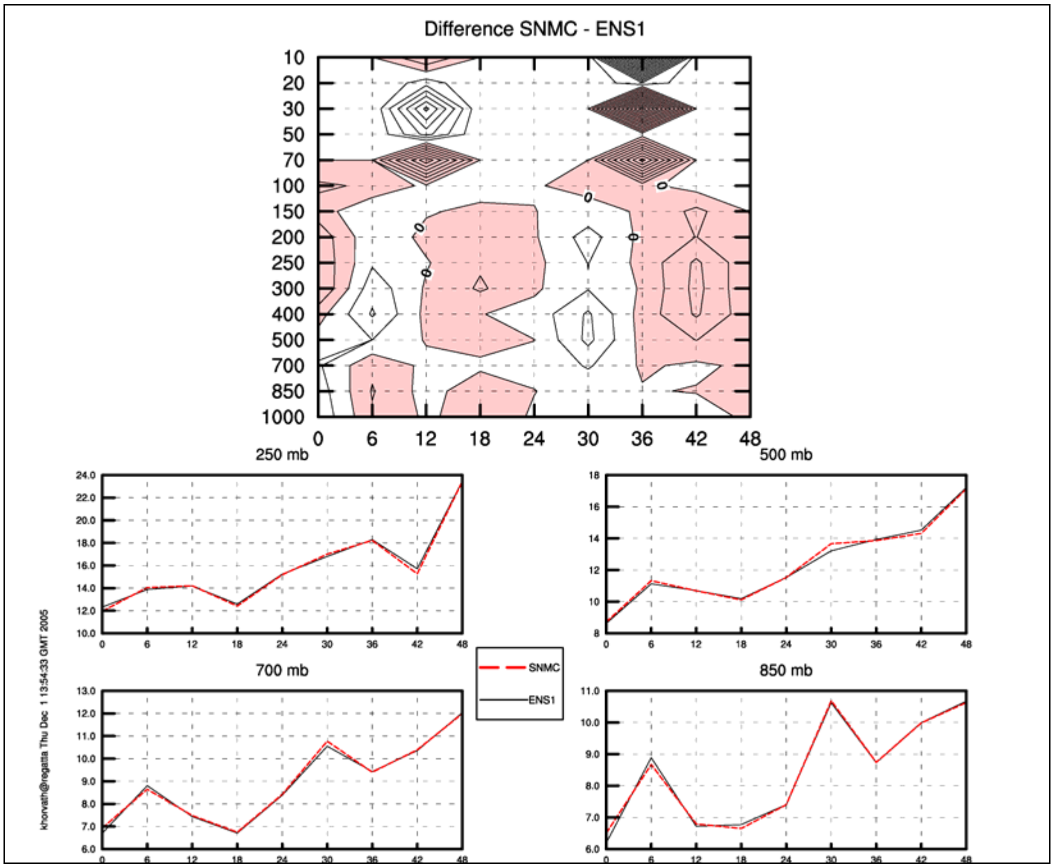


Figure 6. (continued) Root mean square error (up) and bias (down) of wind speed in full observation experiment using REDNMC=1.0 for both ensemble and standard NMC statistics.



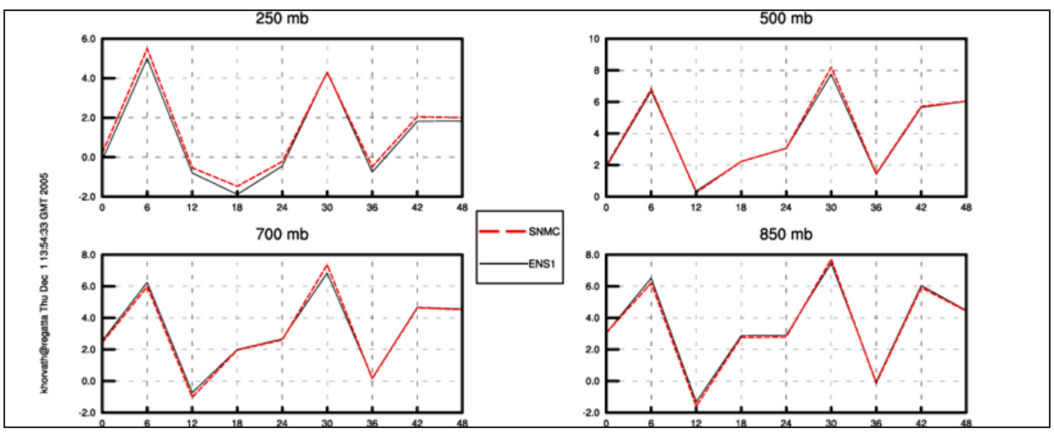


Figure 7. Root mean square error (up) and bias (down) of geopotential in full observation experiment using REDNMC(ensemble)=1.5 and REDNMC(SNMC)=1.0.

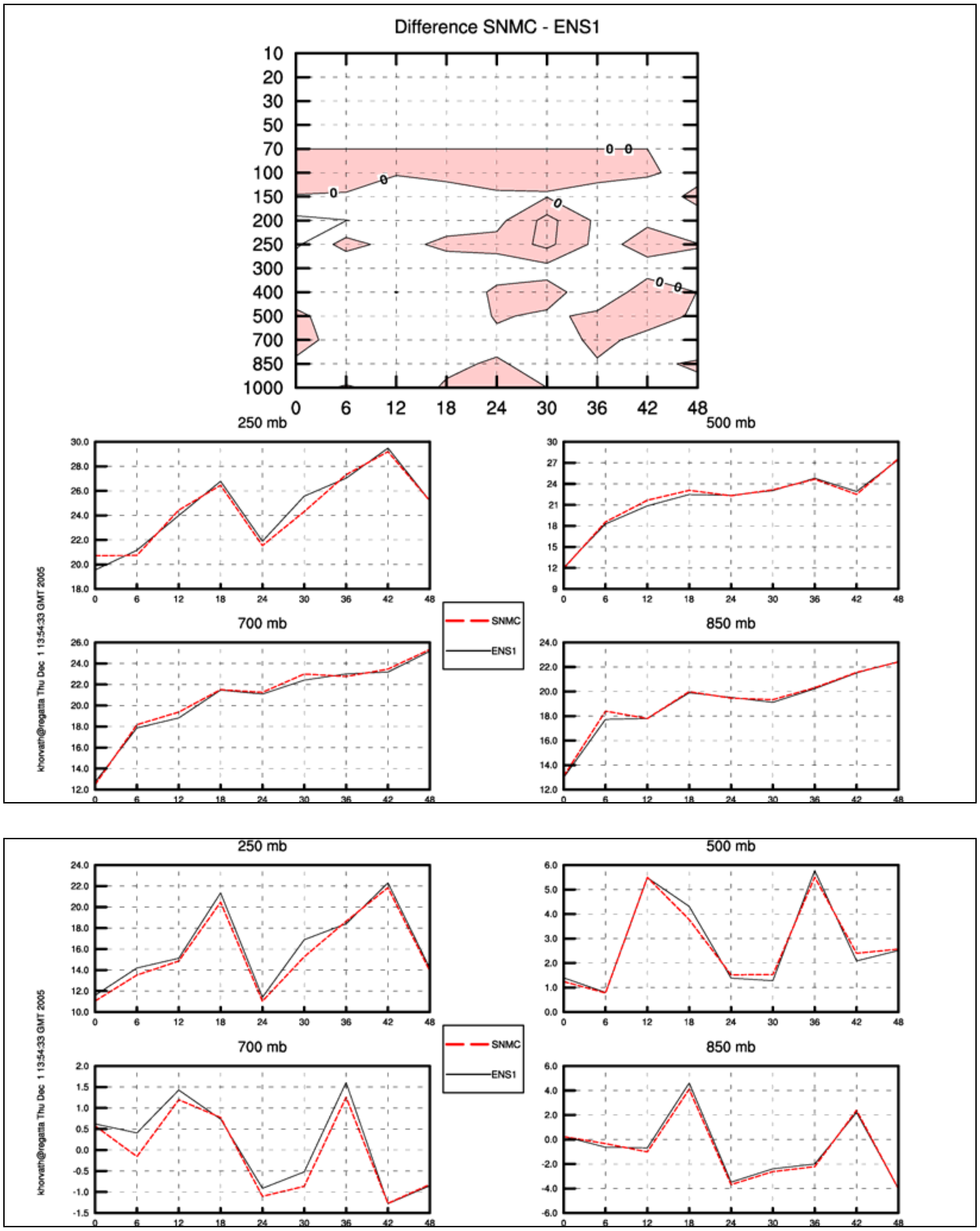
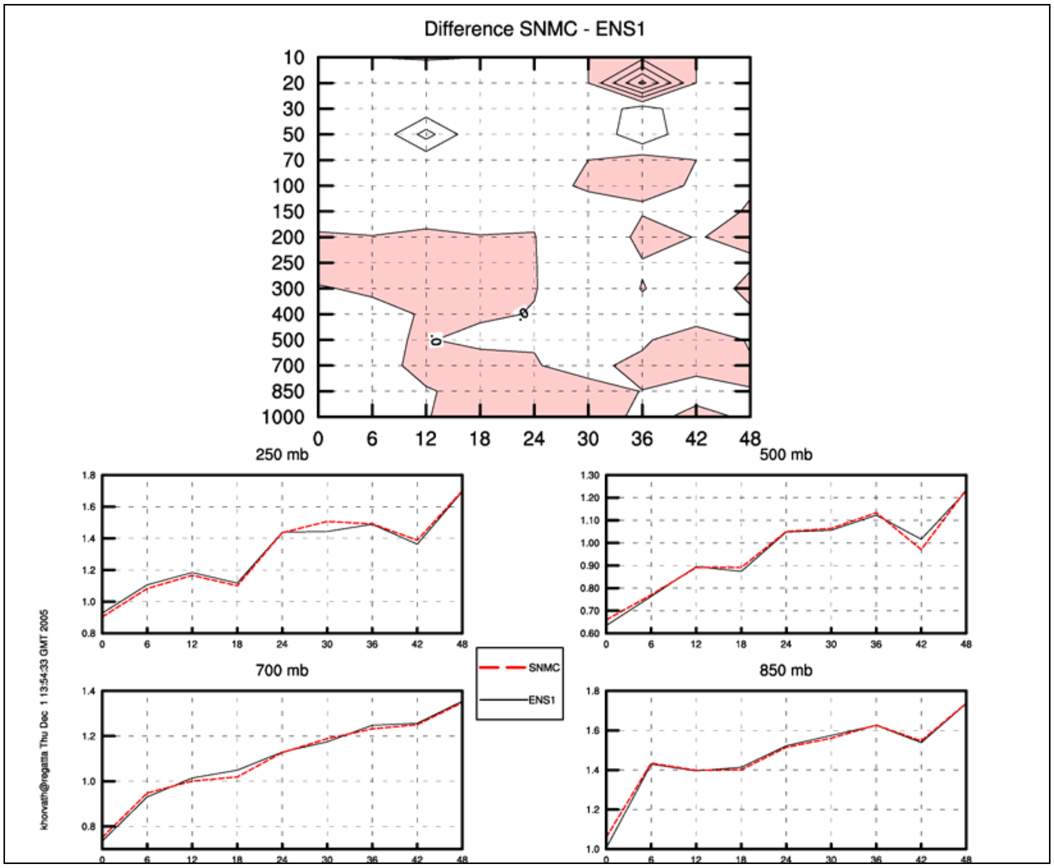


Figure 7. (continued) Root mean square error (up) and bias (down) of relative humidity in full observation experiment using REDNMC(ensemble)=1.5 and REDNMC(SNMC)=1.0.



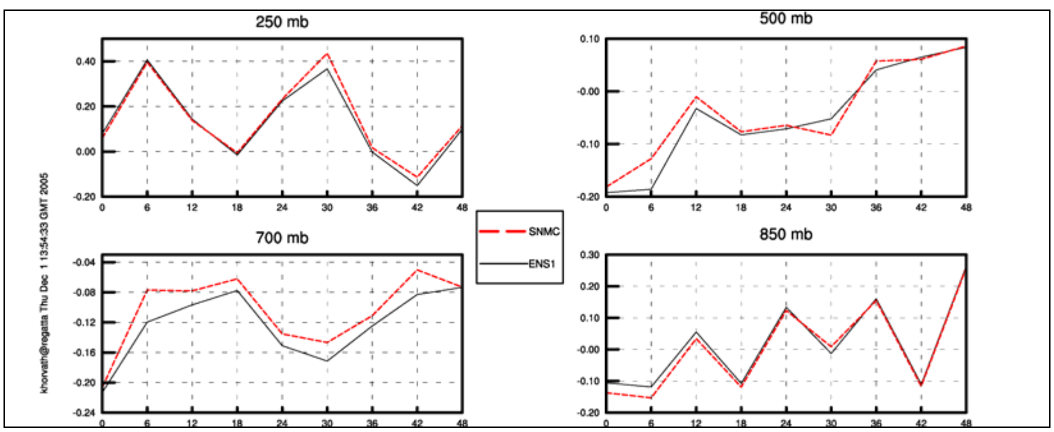
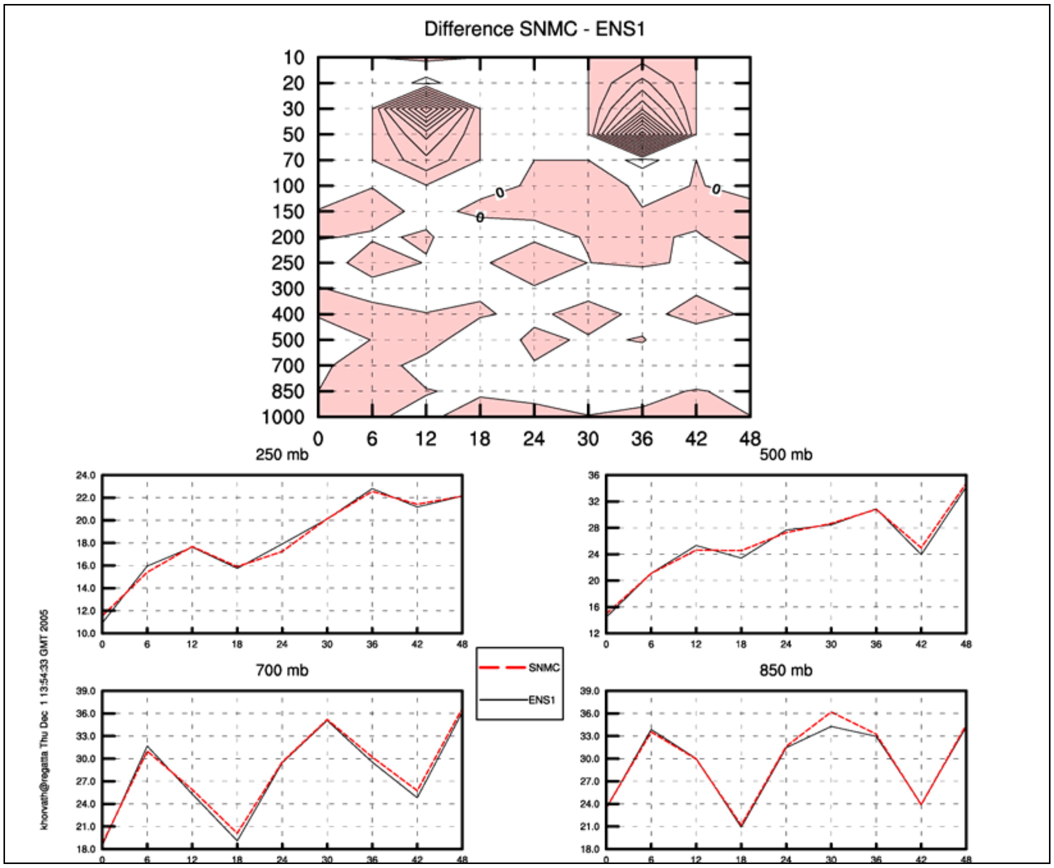


Figure 7. (continued) Root mean square error (up) and bias (down) of temperature in full observation experiment using REDNMC(ensemble)=1.5 and REDNMC(SNMC)=1.0.



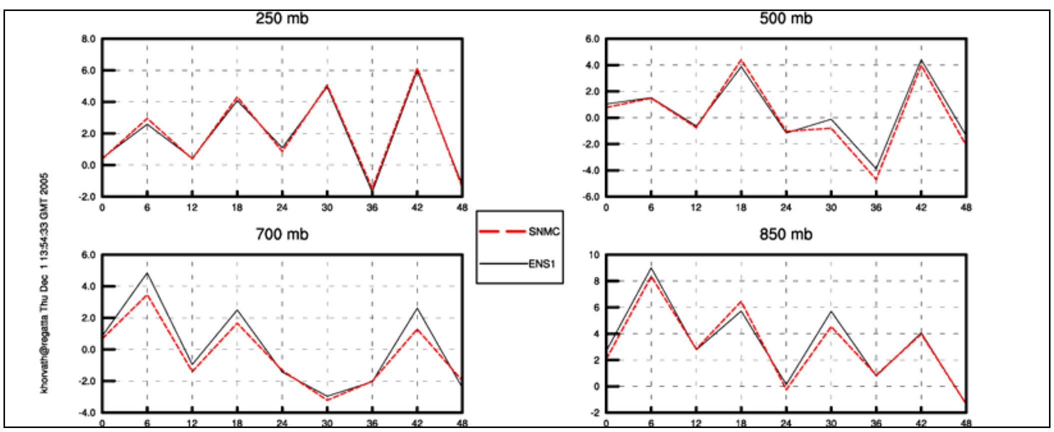


Figure 7. (continued) Root mean square error (up) and bias (down) of wind direction in full observation experiment using REDNMC(ensemble)=1.5 and REDNMC(SNMC)=1.0.

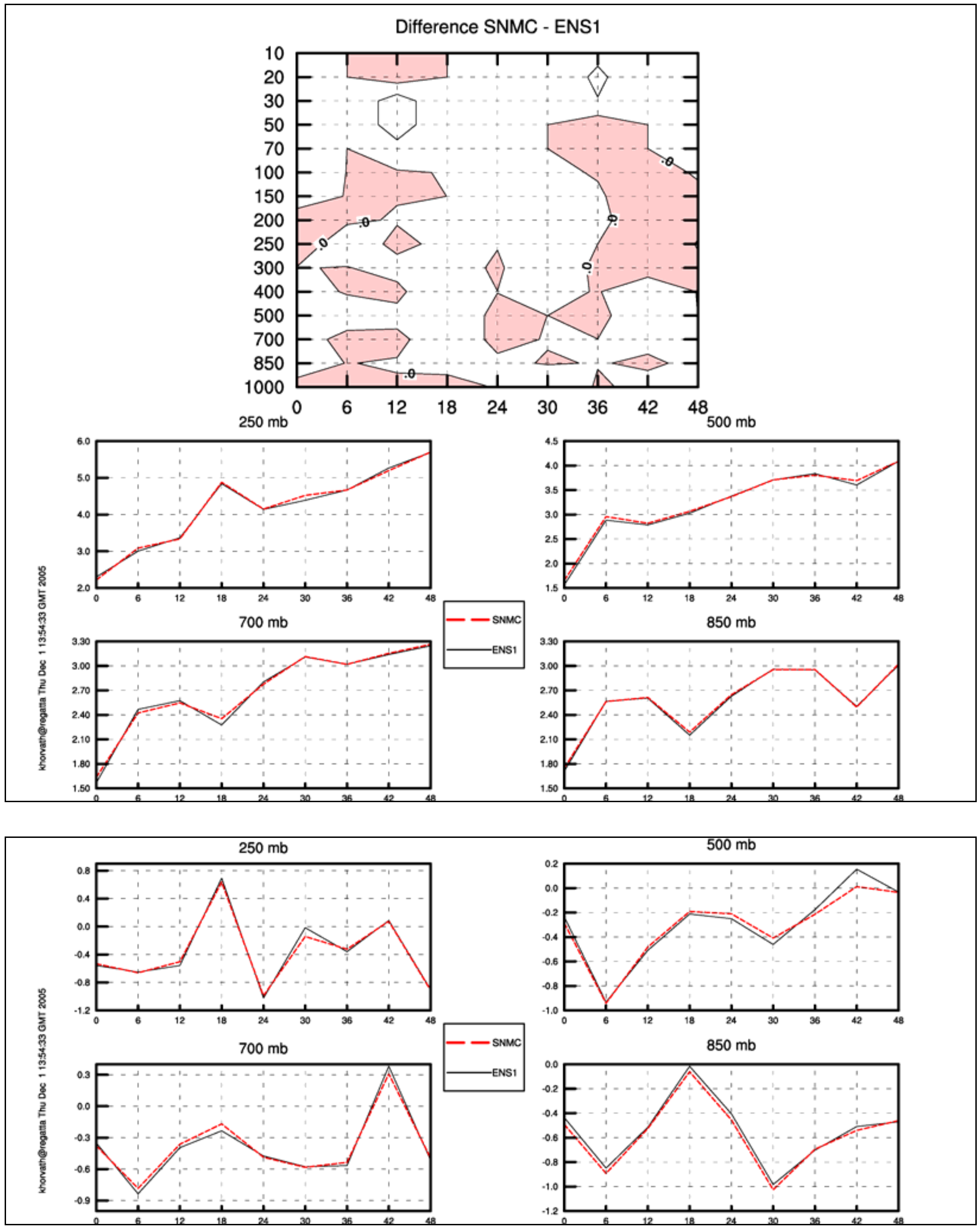


Figure 7. (continued) Root mean square error (up) and bias (down) of wind speed in full observation experiment using REDNMC(ensemble)=1.5 and REDNMC(SNMC)=1.0.