

**Comparison of NMC and Lönnerberg-
Hollingsworth type of background error
statistics of 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
Sept-Oct 2004

Introduction

In every data assimilation technique, information about the state of atmosphere gained from observation in the observation point is spread in horizontal and vertical directions. This propagation of the information gained from observation is in the DA procedure mostly dictated by the **B** - background (usually 6h forecast) error covariance matrix. In a multivariate data assimilation formulation, **B** matrix is responsible not only for the propagation of information of the observed variable, but also about the propagation of information to other variables. In the ALADIN/HU multivariate data assimilation setup, the **B** matrix couples vorticity, divergence, temperature and surface pressure and specific humidity variables (more about couplings definition in Berre, 2000).

In general, for the success of DA procedure it is crucial to have dynamically and thermodynamically balanced **B** error covariance matrix. Nowadays, the most used **B** matrix in operational practice is computed by the NMC method, that approximates the forecast error from a set of differences between two forecast valid at the same time, but at different ranges (Parish and Derber, 1992). Modification of the standard NMC method (SNMC), to account for the need of the representing smaller scales in LAM is so-called lagged NMC statistics (Siroka et al.,). Lagged NMC statistics (LNMC) follows the same procedure as the standard variant, but for the shorter range forecast uses LBC data from the longer range forecast run and for IC uses the 24 hour forecast of the longer range forecast run. Thus, both of these variants of the NMC method are mainly model outputs.

An independent approach to the approximation of the forecast error covariances was proposed by Hollingsworth and Lönnberg (1986). They analysed the statistical structure of the forecast errors, by verifying the forecasts against the radiosonde data. As a by-product, this method provides estimates of the observation error covariance matrices for the observation type analysed.

During the research stay, both SNMC and LNMC error statistics were compared with the error statistics calculated using Lönnberg-Hollingsworth (LH) method. The analysis was for the time being concentrated only on the comparison of horizontal error covariances and correlations.

Preparation of the NMC horizontal grid-point covariances and auto-correlations

The Festat package was used to compute NMC statistics. The computation included 36-12h forecast difference samples prepared for both standard and lagged NMC calculation procedures collected in the period 02.05.2004 – 02.08.2004.

In order to compare NMC statistics and LH statistics (as LH covariances and correlations are calculated in grid-point space), a transformation of the NMC covariances from spectral to grid-point space had to be done. The inverse FFT was achieved by calling the 3Dvar setup code and switching on the option LCORCOSU for calculation of compactly supported covariances (i.e. correlations). For that purpose code in the suejbcosu.f subroutine had to be somewhat modified.

To compare NMC and LH correlations two ways of transforming spectral NMC covariances to grid-point correlations have been tested:

1. a) calculate the correlations in spectral space (achieved by calculating the normalised spectral density of the variance at one level that corresponds to spectral coefficients of the horizontal correlations under the homogeneity assumption, after Berre, 2000)
b) and then make inverse FFT to get the grid-point correlations
2. a) first perform the inverse FFT to get the grid-point covariances
b) calculate grid-point correlations from grid-point covariances by dividing the latter with the value of grid-point covariance at the origin

These two ways proved to yield the same result.

Preparation of the LH covariances and correlations

Calculation of LH statistics used a sample of first-guess departures stored in ODB format, collected in the period 02.05.2004-09.09.2004. The covariances of first guess departures are calculated using time series of departures for all pairs of stations. Then, they were classified into bins according to the distance between the pairs involved into the calculation and averaged for the given bin.

The LH calculation programme had to be modified in order to be able to calculate auto-correlations, beside auto-covariances. The calculation of specific humidity covariances and correlations demanded modification of the format of the input ascii file, i.e. mandaodb ODB viewer.

Results

Thus, LH and both NMC grid-point covariances and auto-correlations were compared. Variables chosen for the comparison were temperature and specific humidity. LH programme was used to calculate the involved values, using TEMP reports only, at levels 1000, 850, 700, 500, 300 and 200 hPa. Standard and lagged NMC statistics were calculated at model levels 37, 28, 23, 18, 12, 09 that approximately correspond to the former pressure levels.

Covariances

As mentioned before, LH statistics calculation verifies forecasts versus observation (first-guess departure) and as a side product provides the estimation of observation errors. Namely, only (co) variance value at the origin point is comprised of both observation and background error variances while values at any distance from the observation point are comprised only out of background errors. This is coming from the assumption that observation errors are correlated only at the same point and uncorrelated at different locations. In mathematical formalism it can be expressed as follows:

$$\text{cov}(e_k, e_l) = 1/N * \sum (e_{kj} - \underline{e}_k) * (e_{lj} - \underline{e}_l) ; j=1, N$$

where $\text{cov}(e_k, e_l)$ is covariance of first-guess departures, e_{kj} and e_{lj} are j -th elements of the errors (first-guess departures) at locations of the k -th and l -th station, and \underline{e}_k , \underline{e}_l their averages. If for simplicity we assume that errors are unbiased it follows:

$$\text{cov}(e_k, e_l) = 1/N * \sum (e_{kj}) * (e_{lj}) ; j=1, N$$

Decomposing first-guess departures into observation and background errors:

$$\text{cov}(e_k, e_l) = 1/N * \sum (e_{kjo} + e_{kjb}) * (e_{ljo} + e_{ljb}) ; j=1, N$$

that is

$$\text{cov}(e_k, e_l) = 1/N * \sum (e_{kjo} * e_{ljo} + e_{kjo} * e_{ljb} + e_{kjb} * e_{ljo} + e_{kjb} * e_{ljb}) ; j=1, N$$

By assumption that observation and background error are not correlated (not always true), parts 2 and 3 on the right side of that equation equal zero. It yields:

$$\text{cov}(e_k, e_l) = 1/N * \sum (e_{kjo} * e_{ljo} + e_{kjb} * e_{ljb}) ; j=1, N$$

Now, if the stations are the same, the covariance is composed of both the observation and background error variance. If the stations are not the same, assuming that observations errors are uncorelated between different stations, the first-guess departure covariance has contribution only from the background error covariance.

Temperature covariances

Temperature covariances of the LH statistics are presented at Figure 1. From the value at origin point, observation error should be subtracted from the total values by extrapolating the line that goes through the first and second (bin) value (if value at the origin is cosidered to be the value of the zero-th bin) to the origin. However at level 37, due to lack of observations, line could be extrapolated to the origin point using the second and third bin value, rather than first and second. Second, in the correlation study used aproach is to involve aproximation of equality of the observation and background error variances (thus, divide the values at the origin by 2).

LH covarinces show to be highly comaprable with the NMC statistics, with values mostly between the standard and lagged variants (Fig 1) in the upper and midle troposphere. At lower levels, specially level 37, LH covariances are of somewhat greater length scales, although standard deviations seem to be of similar values. This feature will be notable even better in the correlation space. Thus, to some extent, LH covariances confirm the shapes and horizontal length scales of the NMC temperature covariances. However, to judge the quality of the all three involved background error estimations, it is necessary to perform a verification procedure.

Specific humidity covariances

LH specific humidity covariances are well balanced with NMC covariances in the wide area of middle troposphere (Fig 2). However, at lowest level and at the levels near the tropopause, large differences between statistics can be noted, both in the variances a the origin point and the broadness of horizontal length scales.

Without the verification it is not easy to say which statistics gives the results that represent the real model error the best. However, quite huge LH covariances at great distances seem to be unrealistic.

Estimation of observation error variances

This estimation is done by the extrapolation procedure described above. It provided observation error estimates for the TEMP measurements of temperature and specific humidity (Table 1). At level 37, due to lack of significant number of observations, approximation of the equality of background and observation error variances was used. However at all levels, construction of bin calculation where all pairs of stations between 0 and 200 km are averaged as one 100 km distance bin value significantly contributes to the unreliability of error separation, because most of the pairs group around the upper boundary of the bin. The reliability of the separation procedure can be judged from the estimated $\sigma_b^2/(\sigma_b^2+\sigma_o^2)$ values. This ratio e.g. For humidity at 850 hPa equals 0.1, making observation variance 9 times bigger than background error variance, what seems to be rather unrealistic.

	Lev 09	Lev 12	Lev 18	Lev 23	Lev 28	Lev 37
$\sigma_o^2(T)[K^2]$	0.65	0.33	0.31	0.55	0.63	1.43
$\sigma_o^2(q)$ [g ² /kg ²]	5.2*10 ⁻¹¹	6.8*10 ⁻¹⁰	1.0*10 ⁻⁷	7.0*10 ⁻⁷	1.3*10 ⁻⁶	1.55*10 ⁻⁶

Table 1. Estimates of the TEMP observation error covariances at different levels of atmosphere by extrapolating method. Results do not have high level of reliability, because of small number of observations in the first bin and construction of the bin calculations that lead to a wide range of σ_b/σ_o values (see text).

Correlations

In case of NMC statistics, because of the assumption of the homogeneity (see e.g. Berre, 2000), correlations are just covariances normalized with the values at the origin point. In case of LH correlations, observation error covariances were subtracted from the value at the origin point to get the pure background error covariance. This was done by assuming $\sigma_b=\sigma_o$ (because of non very reliable results yielded from the separation procedure discussed above) i.e. dividing every station pair autocovariance value by a factor of 2 before the correlation calculation procedure.

Temperature correlations

Insight into temperature correlations (Fig 3) reveals that at higher levels LH estimates do not differ significantly from background error variance NMC estimates. However, differences grow with pressure and LH correlations at lowest levels are of much greater length scales

then both NMC variants, reaching quite unexpected values. In this way, temperature correlation is greater than 0.2 even at a distance of 1400 km. This might imply that the assumption of the equality of observation and background error variances is not a good one at low levels, where background error tends to be greater than the observation error. Moreover, it is possible that there is some factor in the model (e.g. surface parametrisations) that dominates the errors in such a way that it pushes them in the same direction over the wide area of the domain.

Specific humidity correlations

Specific humidity correlations (Fig 4) reveal even greater differences between LH and NMC statistics. The correlation functions are somewhat more comparable at middle tropospheric levels, where LH correlations are not so broad. However, at tropopause levels and the lowest levels of the model, the differences are extremely huge. If true, however, it might mean that some factors in the model are not well balanced, for case of humidity at high as well as low atmospheric levels.

Special features

It is interesting to note that at high levels temperature LH covariances and correlations gain negative values (Fig 5). The probable reason for that is found in the existence of high atmospheric disturbances that dominate dynamics of the high atmosphere. Structures of high-level disturbances were extensively analysed both theoretically (e.g. Charney 1947) and somewhat less in observational studies (e.g. Lim and Wallace 1991). If we consider that the wave length of the (temperature) Rossby wave is around 6000 km on average, we expect that the covariance will become negative for distances roughly more than quarter of the wavelength (up to three-quarters of the wavelength). This feature can qualitatively be seen at 200 hPa and 300 hPa plots of temperature covariance functions. It is however notable that for humidity there is no such negative dependence (not shown).

Temperature LH covariances at 1000 hPa are positive far away from the origin point (Fig 6). Covariances are higher than 0.2 K^2 up to distance of roughly 1800 km, what seems to be quite unrealistic. The same feature is visible in the correlation space as well as for humidity variable at low levels.

References:

- Berre, L., 2000: Estimation of synoptic and mesoscale forecast error covariances in a limited area model. *Mon. Wea. Rev.*, 128, 644-667.
- Charney, J. G., 1947: The dynamics of long waves in a baroclinic westerly current. *J. Meteor.*, 4, 135-163.
- Hollingsworth, A. and Lönnerberg, P., 1986: The statistical structure of short-range forecast errors as determined from radiosonde data. Part I: The wind field. *Tellus*, 38A, 111-136.
- Lim, G.-H. and Wallace, J. M., 1991: Structure and Evolution of Baroclinic Waves as Inferred from Regression Analysis. *J. Atmos. Sci.*, 48, 1718-1732.
- Parrish, D. F. and Derber, J. C., 1992: The National Meteorological Center's spectral statistical interpolation analysis system. *Mon. Weather Rev.*, 120, 1747-1763.
- Siroka, M., Fischer C., Cassé V., Brozkova R. and Geleyn, J.-F., 2001: The definition of mesoscale selective forecast error covariances for a limited area variational analysis. *Meteorology and Atmospheric Physics*, Meteor. Atmos. Phys. special issue of the SRNWP Workshop on high-resolution modeling, Offenbach, 25.-27.10.1999.

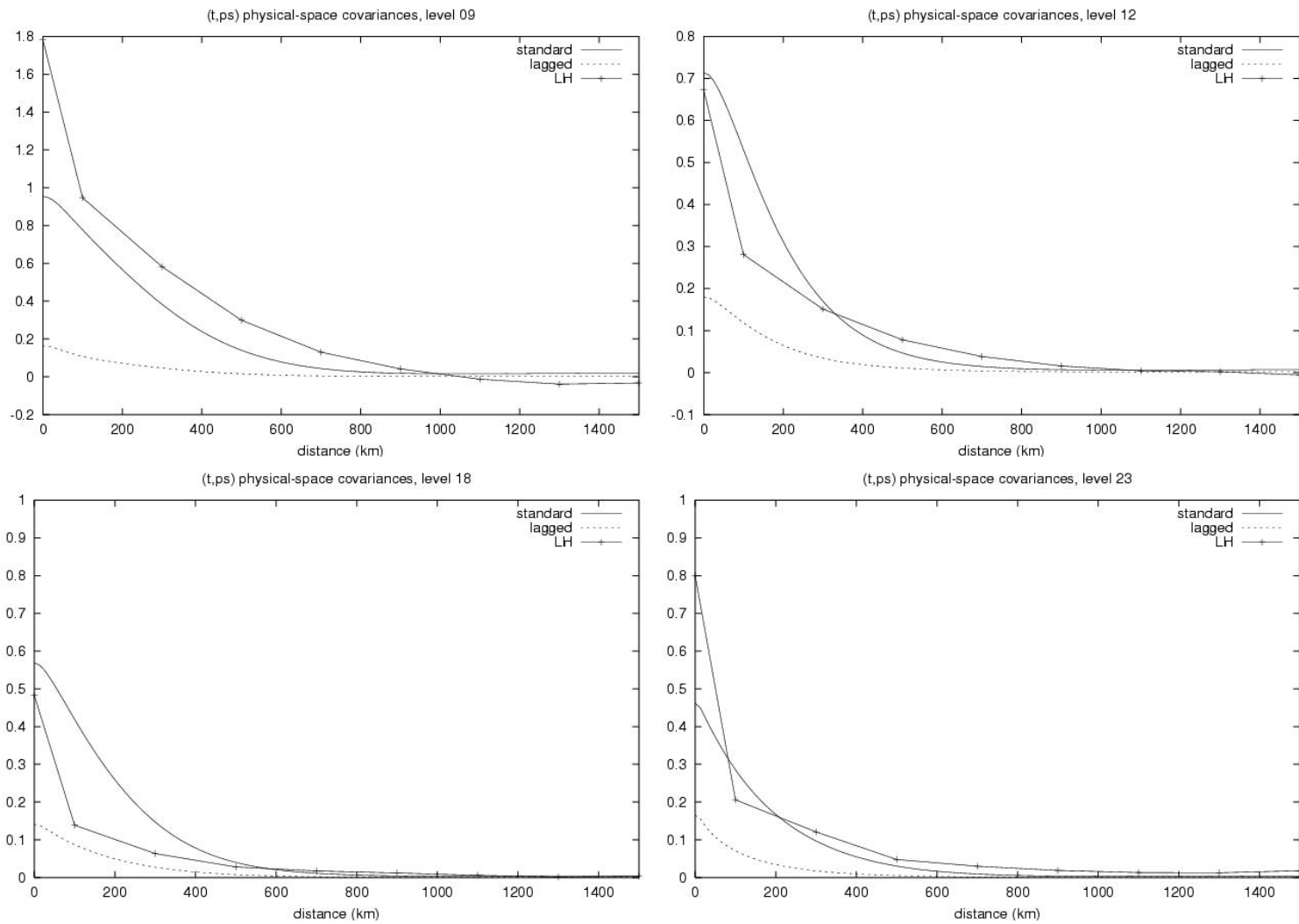


Figure 1. Comparison of the SNMC, LNMC and LH background error variance estimates for temperature variable at levels 09, 12, 18, 23, 28 and 37 approximately corresponding to 200hPa, 300hPa, 500hPa,

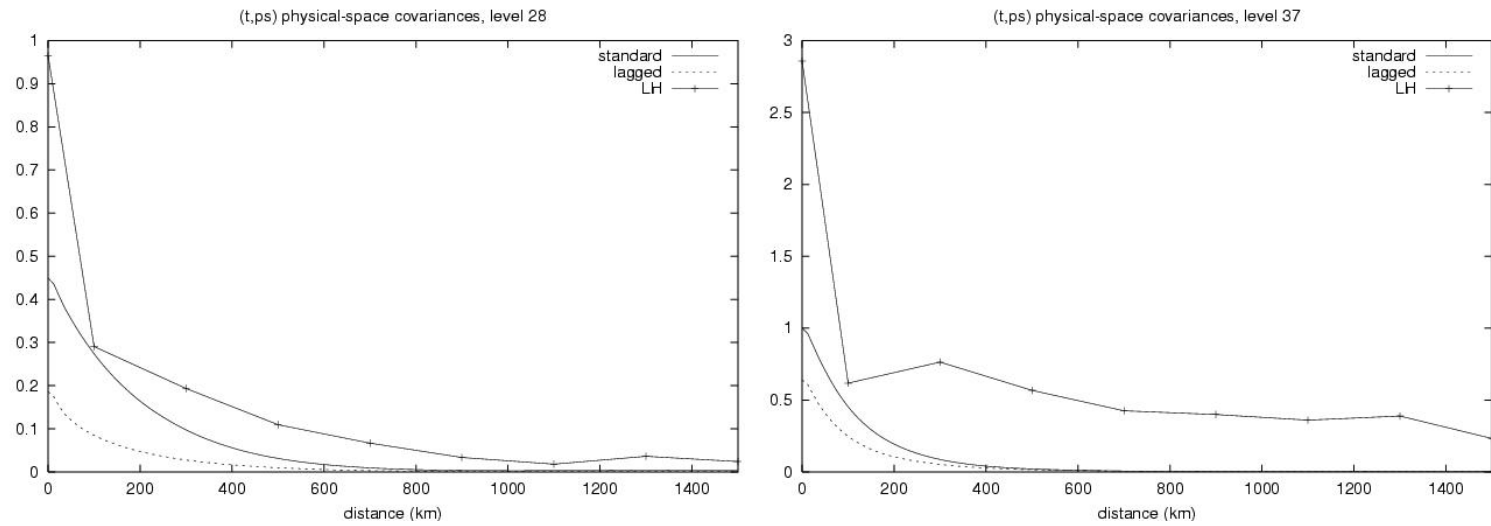


Figure 1.(cont) 700hPa, 850 hPa and 1000hPa. Note the need to involve separation of the observation and background error variances at the origin point for the case of LH statistics covariance function.

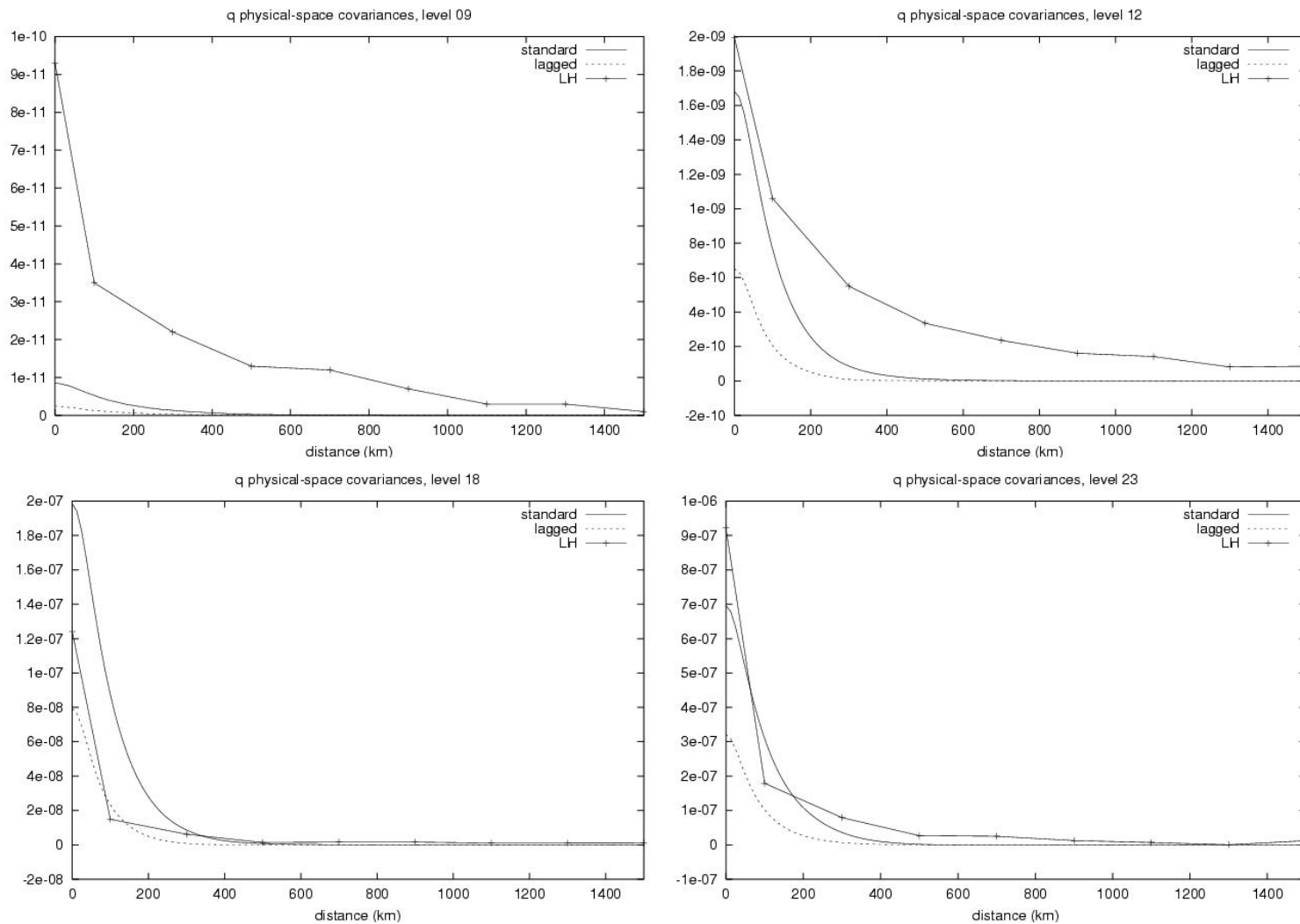


Figure 2. Comparison of the SNMC, LNMC and LH background error variance estimates for specific humidity variable at levels 09, 12, 18, 23, 28 and 37 approximately corresponding to 200hPa, 300hPa,

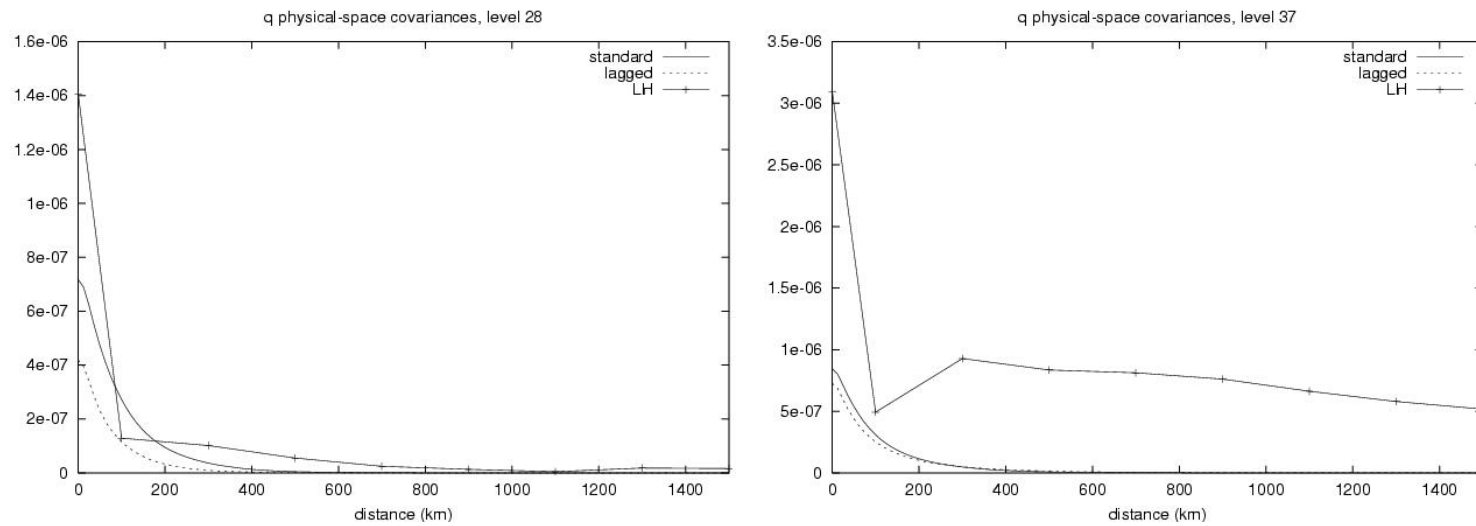


Figure 2.(cont) 500hPa, 700hPa, 850 hPa and 1000hPa. Note the need to involve separation of the observation and background error variances at the origin point for the case of LH statistics covariance function.

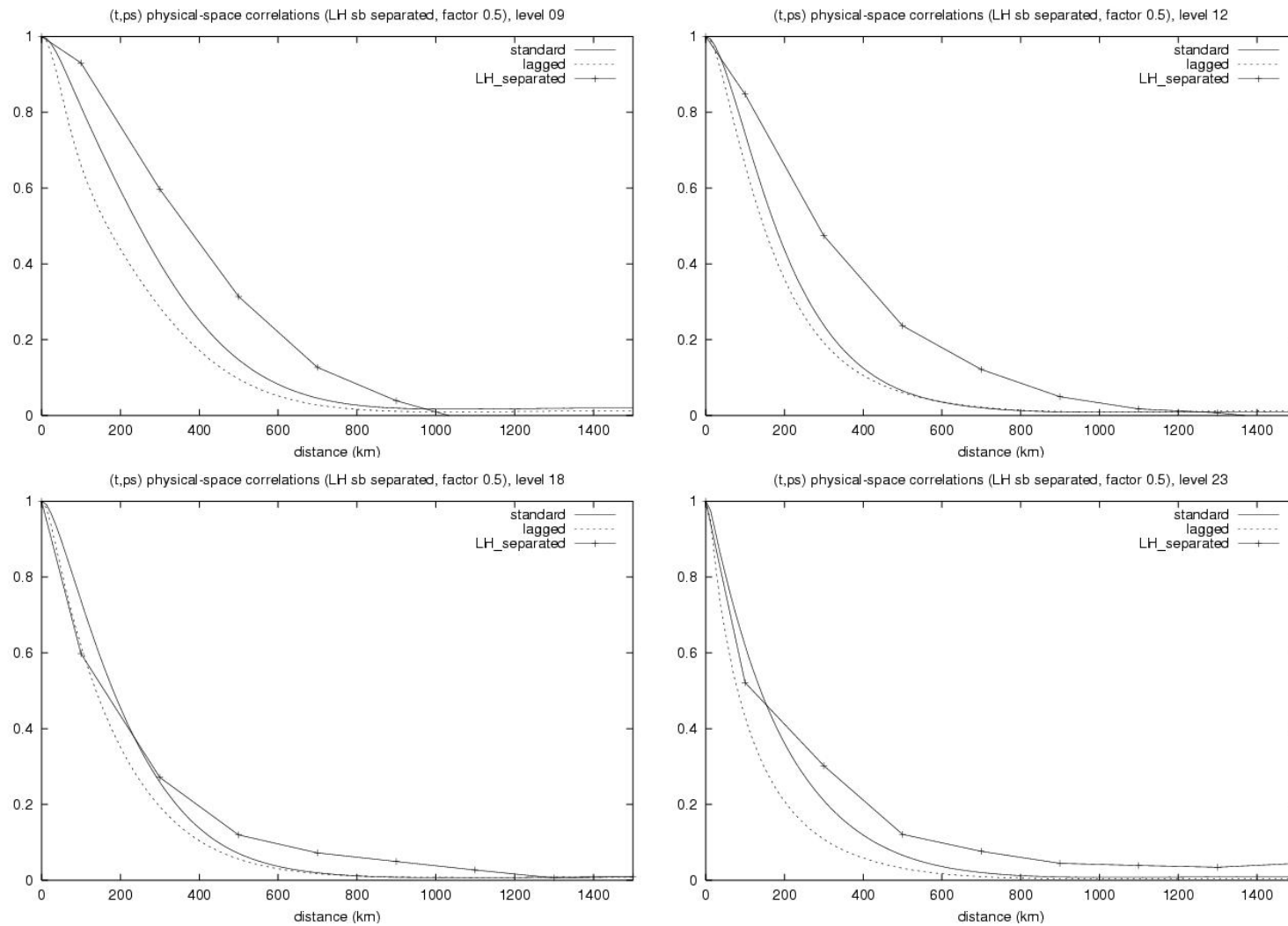


Figure 3. Comparison of the SNMC, LNMCM and LH background error variance estimates for temperature variable at levels 09, 12, 18, 23, 28 and 37 approximately corresponding to 200hPa, 300hPa, 500hPa,

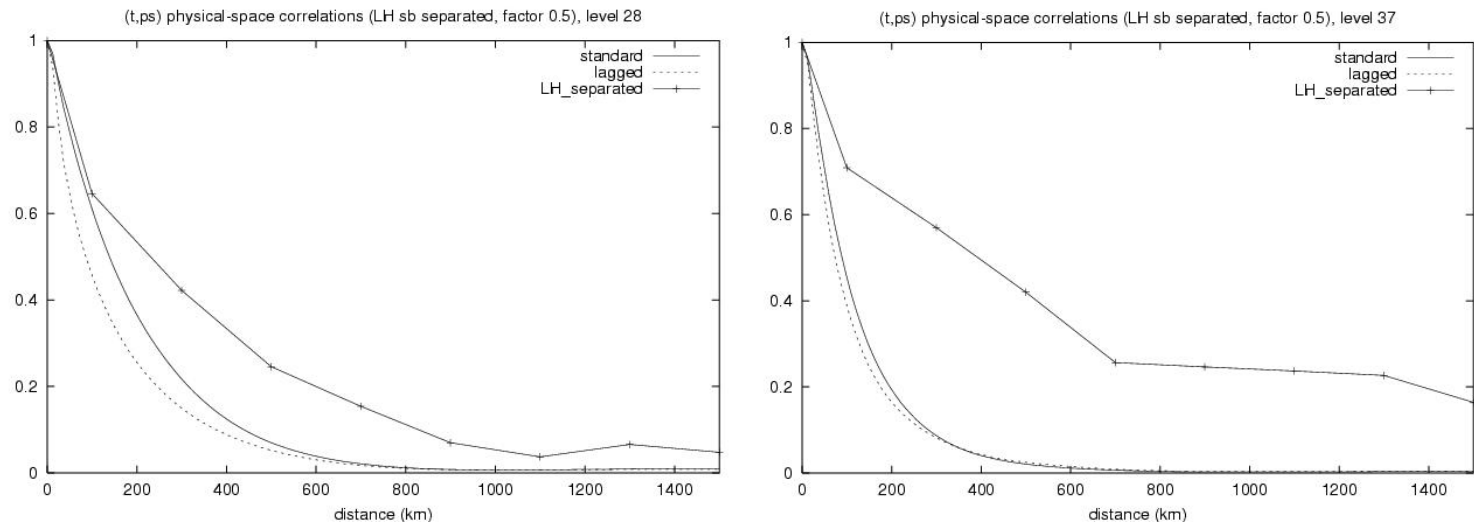


Figure 3.(cont) 700hPa, 850hPa and 1000hPa. Note the huge difference between the NMC and LH horizontal length-scales of correlation functions at low levels of atmosphere

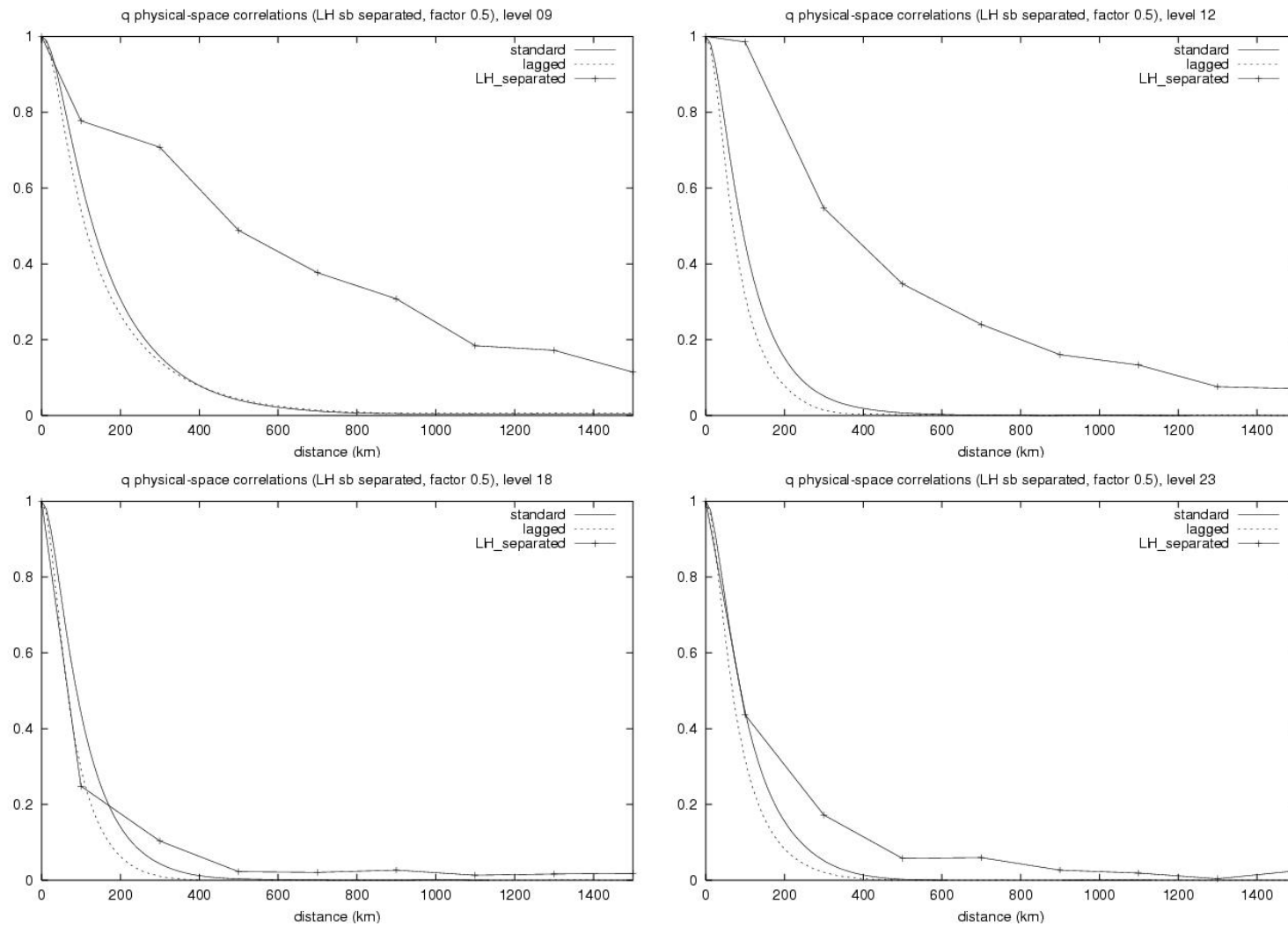


Figure 4. Comparison of the SNMC, LNMC and LH background error correlation functions for specific humidity variable at levels 09, 12, 18, 23, 28 and 37 approximately corresponding to 200hPa, 300hPa,

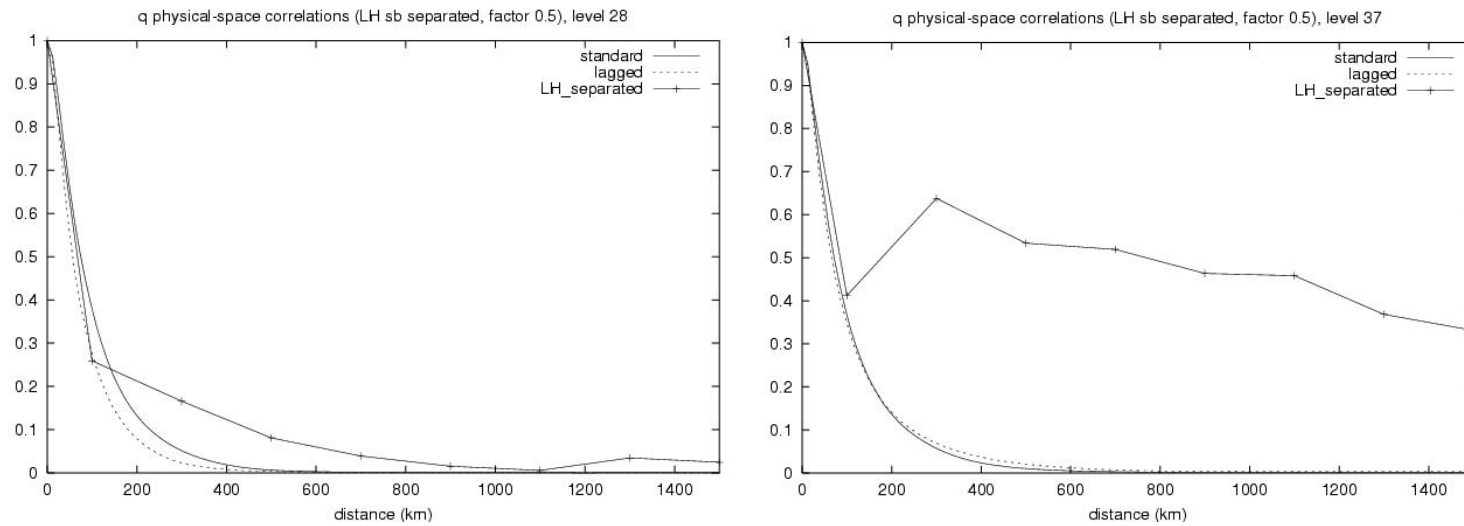


Figure 4.(cont) 500hPa, 700hPa, 850hPa and 1000hPa. Note the huge difference between the NMC and LH horizontal length-scales of correlation functions at high, but as well as low levels, while the agreement is better in the middle troposphere

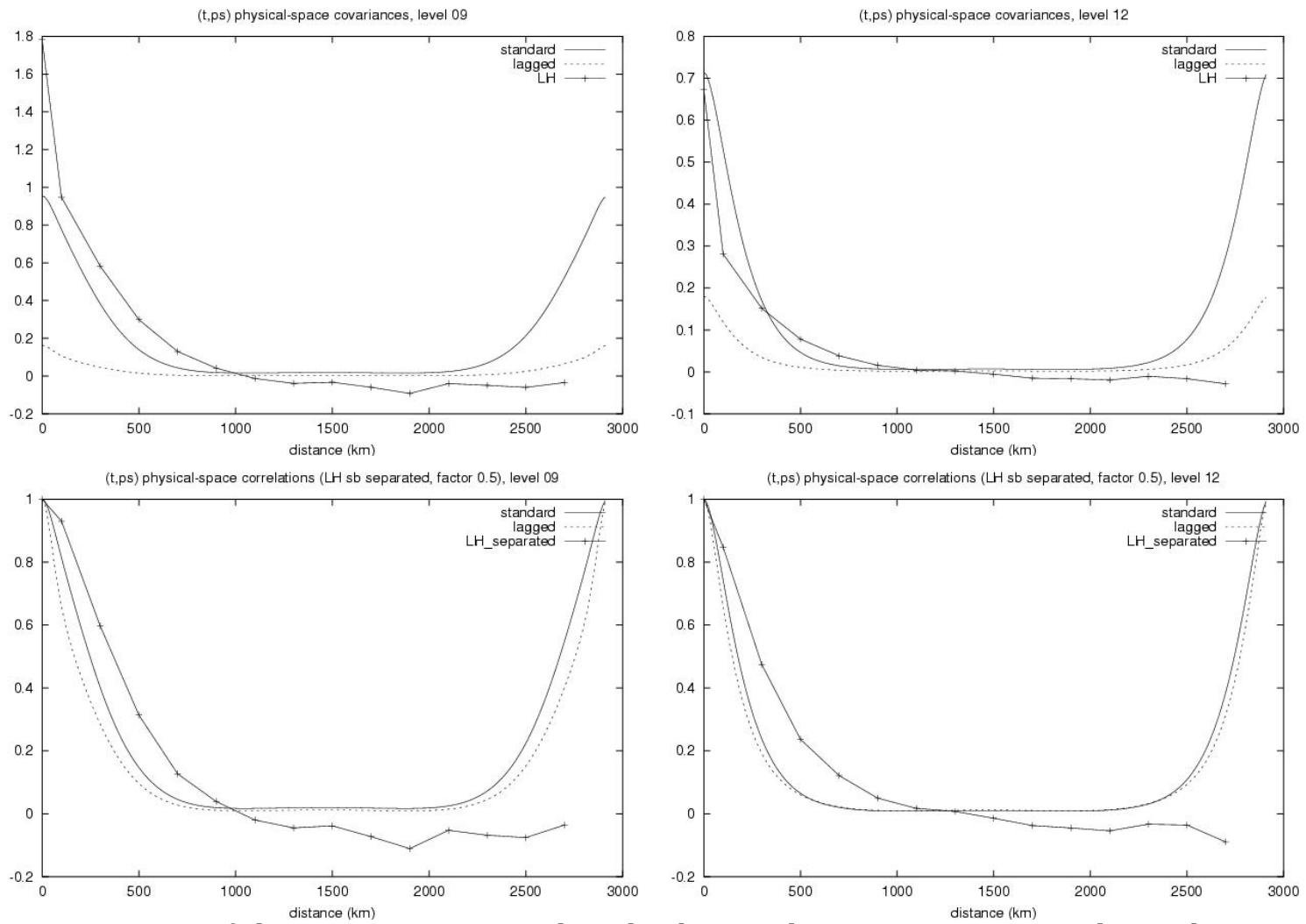


Figure 5. Comparison of the SNMC, LNMCM and LH background error covariance and correlation functions for temperature at levels 09 and 12 approximately corresponding to 200hPa and 300hPa

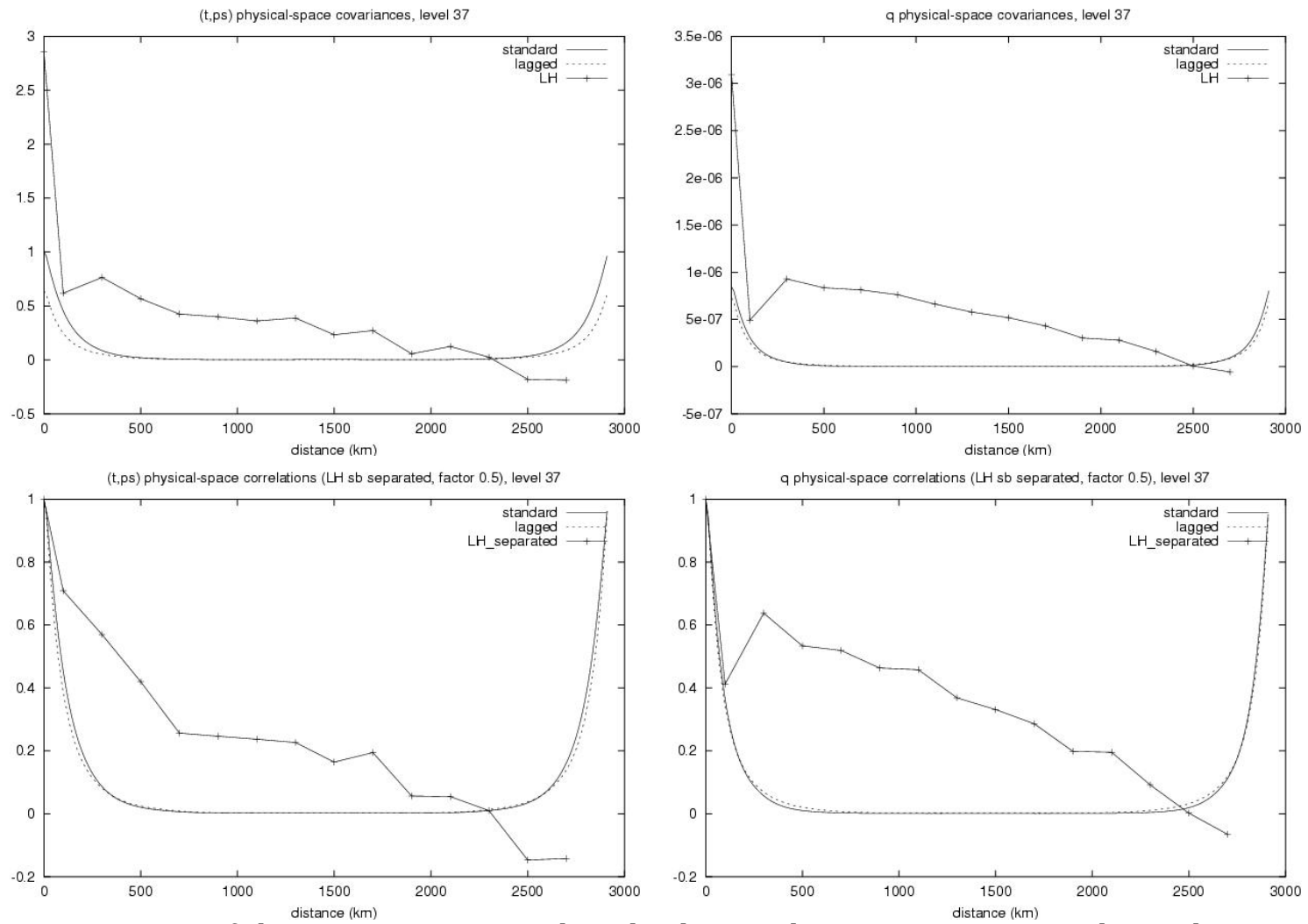


Figure 6. Comparison of the SNMC, LNM and LH background error covariance and correlation functions for temperature at level 37 approximately corresponding to 1000hPa level.