



---

---

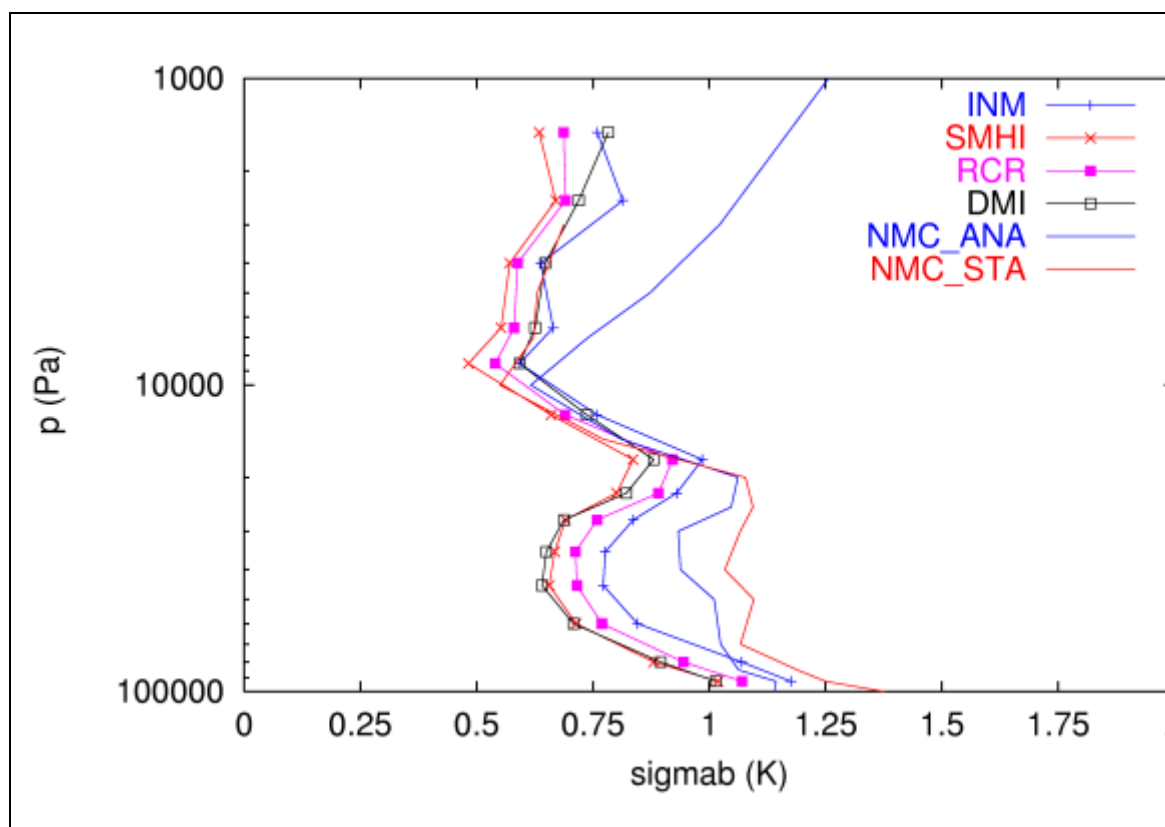
## Technical Report

No. 66, April 2006

---

---

### Diagnosis of error statistics in the HIRLAM 3D-VAR



Diagnosed and prescribed errors of temperature for sounding data

Beatriz Navascués, Magnus Lindskog, Xiaohua Yang and Bjarne Amstrup

This report is published by the HIRLAM-A Programme and the previous HIRLAM-6 Project. The Hirlam Technical Reports are intended for original scientific contributions and results supplied by the Hirlam Scientific staff. Every effort is taken to maintain a high standard.

The manuscript is reviewed independently by the Hirlam Project Leader and another, optionally anonymous, reviewer selected for his or her expertise in the field relevant to the Report. The reviews follow the same procedures as the major international Scientific Journals in the field.

The purpose of the Technical Reports is to communicate results relatively soon after compiling the manuscript. They are distributed to all Hirlam staff members and other research personnel and libraries in the Hirlam member institutes as well as to a large number of scientists and libraries in meteorological institutes and universities internationally. The distribution list contains in excess of 260 addresses and normally the Reports are printed in 300 copies. Extra copies are available on request to the Programme Manager, Jeanette Onvlee, KNMI, Postbox 201, 3730 AE De Bilt, Netherlands or [Onvlee@knmi.nl](mailto:Onvlee@knmi.nl).

# DIAGNOSIS OF ERROR STATISTICS IN THE HIRLAM 3D-VAR

Beatriz Navascués<sup>1</sup>, Magnus Lindskog<sup>2\*</sup>, Xiaohua Yang<sup>3</sup>, Bjarne Amstrup<sup>3</sup>

(1) Spanish Meteorological Institute (INM)

(2) Swedish Meteorological and Hydrological Institute (SMHI)

(3) Danish Meteorological Institute (DMI)

Manuscript received 25 November 2005, revised 23 February 2006.

## Abstract

This report summarizes the diagnostic study of 3D-VAR analyses at several HIRLAM operational centers, using the algorithm proposed recently by Desroziers et al. (2005). Four extensive data sets of archived statistic files from the DMI, INM, SMHI and FMI operational analyses have been used in the investigation. The diagnosed observation and background error standard deviations from these data sets are examined for conventional observation platforms and compared with specified error covariances. In addition, seasonal and latitudinal variations in the error variances are examined. The study shows that the partitioning of the innovation variance between observation and background errors depends on the prescribed error statistics in different operational centers, and it confirms a perception in the HIRVDA community of a general over-specification of the background error covariances and observation errors, at least for some observation types. It is anticipated that the new diagnosis tool used in this study will also be valuable for the future tuning of the error covariances specification in HIRLAM's variational data assimilation.

## 1. Introduction

Correctly specified background and observation error statistics is an important part of any data assimilation system. The standard deviations of observation errors utilised in the HIRLAM variational data assimilation system, HIRVDA, were originally obtained from the ECMWF 3D-Var (Courtier et al., 1998), but have afterwards been crudely revised (Lindskog et al., 2001). The observation errors are assumed to be uncorrelated for all conventional types of observations. The background error statistics are based on the so-called NMC method (Parrish and Derber, 1992; Rabier et al., 1998), in which statistics of forecast differences are used as an approximation to background error covariances. The correlations (usually referred to as structure functions) contains multivariate couplings between different variables. These couplings (or balances) are described either by using an analytical balance approach (Gustafsson et al., 2001) or using a statistical balance approach (Berre, 2000). Standard deviations of the forecast differences are rescaled with an empirically determined scaling coefficient to match the amplitude of the background errors.

The HIRVDA, like many other operational NWP systems, has for a long time lacked appropriate tools for tuning and diagnosing background and observation error standard deviations ( $\sigma_b$  and  $\sigma_o$ ) used in the variational data assimilation algorithm. In recent years several attempts have been made to apply the method suggested by Hollingsworth and Lönnberg (1986), in which correlations of innovations, calculated from the archived data sets of innovations of different variables and observation types, are extrapolated to zero distance, and used to assess whether the sum of background and observation error variances is correctly specified within the assimilation system. However, the success has been limited due to too low station densities in some parts of the domain and anisotropies in surface variables.

Recently, a new attractive method for diagnosing analysis and error variances ( $\sigma_b$  and  $\sigma_o$ ) presented by Desroziers et al. at the WMO data assimilation symposium in Prague, 2005, has been introduced in HIRVDA. The diagnosis software calculates and accumulates statistics of different combinations of innovations and analysis residuals for different observation types and variables, using output from the HIRVDA analyses (ACMA files). It has been applied to four extensive archived data sets from the DMI, INM, SMHI and FMI (RCR) operational 3D-VAR assimilations. Additionally, diagnosed observation and background error standard deviations have been obtained for different seasons using the archived RCR dataset. Using the INM statistics, latitudinal variation of error variances has also been examined. Results show that the partitioning of the innovation variances between observation and background errors depends on the prescribed error statistics in each of the operational HIRLAM suites, as a consequence of the employed methodology. In general, it is evident that the presently specified error variances in the HIRVDA systems do not provide a sufficiently good estimation of the observation and/or background error standard deviations, and some modifications are needed.

## 2. Methodology and experimental design

The innovation vector,  $\mathbf{d}^o_b$ , the deviation of the model state equivalent from the observations,  $\mathbf{y}$ , is routinely calculated in the very early stage of the analysis process at each observation position. Statistics of innovations,  $\mathbf{d}^o_b = (\mathbf{y} - \mathbf{H}\mathbf{x}^b)$ , provide a global check on the specification in the analysis of the observation errors covariance  $\mathbf{R}$ , and the covariance of background errors in observation space,  $\mathbf{H}\mathbf{B}\mathbf{H}^T$  (see e.g. Andersson, 2003). If observation and background errors are uncorrelated, and  $\mathbf{R}$  and  $\mathbf{H}\mathbf{B}\mathbf{H}^T$  are correctly specified in the analysis, the covariance of innovations  $E[\mathbf{d}^o_b \mathbf{d}^{o_b T}]$  should approximately equal  $\mathbf{R} + \mathbf{H}\mathbf{B}\mathbf{H}^T$ .

Desroziers et al. (2005) have proposed a method to additionally check separately the observation and background errors statistics used in the analysis. They make use of both innovations and the difference in observation space between analysis,  $\mathbf{H}\mathbf{x}^a$ , and background,  $\mathbf{d}^a_b = (\mathbf{H}\mathbf{x}^a - \mathbf{H}\mathbf{x}^b)$  to check the consistency of background errors in observation space. So, if  $\mathbf{H}\mathbf{B}\mathbf{H}^T$  is correctly assigned, then  $E[\mathbf{d}^a_b \mathbf{d}^{o_b T}] = \mathbf{H}\mathbf{B}\mathbf{H}^T$ . Following Desroziers et al. (2005), the consistency of observation errors covariance used in the analysis,  $\mathbf{R}$ , can be also checked by mean of statistics of the analysis residuals  $\mathbf{d}^o_a = (\mathbf{y} - \mathbf{H}\mathbf{x}^a)$  and innovations, because if  $\mathbf{H}\mathbf{B}\mathbf{H}^T$  and  $\mathbf{R}$  are correct,  $E[\mathbf{d}^o_a \mathbf{d}^{o_b T}] = \mathbf{R}$ .

We call  $\mathbf{H}\tilde{\mathbf{B}}\mathbf{H}^T = E[\mathbf{d}^a_b \mathbf{d}^{o_b T}]$  diagnosed background error covariance in observation space, and  $\tilde{\mathbf{R}} = E[\mathbf{d}^o_a \mathbf{d}^{o_b T}]$  diagnosed observation covariance matrices respectively. We then obtain that:

$$E[\mathbf{d}^o_b \mathbf{d}^{o_b T}] = E[\mathbf{d}^a_b \mathbf{d}^{o_b T}] + E[\mathbf{d}^o_a \mathbf{d}^{o_b T}] = \mathbf{H}\tilde{\mathbf{B}}\mathbf{H}^T + \tilde{\mathbf{R}}$$

The innovation variance is thus the sum of the diagnosed background and observation error variances.

As it is described in Desroziers et al. (2005) diagnosed background and observation error covariance matrices,  $\mathbf{H}\tilde{\mathbf{B}}\mathbf{H}^T$  and  $\tilde{\mathbf{R}}$ , may be seen as some adjusted covariance estimates, related to the exact  $\mathbf{B}^*$  and  $\mathbf{R}^*$  covariance matrices through the specified covariances of observation errors  $\mathbf{R}$ , and of background errors in observation space,  $\mathbf{H}\mathbf{B}\mathbf{H}^T$ :

$$\begin{aligned} \mathbf{H}\tilde{\mathbf{B}}\mathbf{H}^T &= \mathbf{H}\mathbf{B}\mathbf{H}^T (\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})^{-1} (\mathbf{H}\mathbf{B}^*\mathbf{H}^T + \mathbf{R}^*) \\ \tilde{\mathbf{R}} &= \mathbf{R} (\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})^{-1} (\mathbf{H}\mathbf{B}^*\mathbf{H}^T + \mathbf{R}^*) \end{aligned}$$

So, the method is implicitly assuming that the innovation variance equals the sum of the variances of the exact background and observation errors ( $\mathbf{H}\tilde{\mathbf{B}}\mathbf{H}^T + \tilde{\mathbf{R}} = \mathbf{H}\mathbf{B}^*\mathbf{H}^T + \mathbf{R}^*$ ), and the temporal evolution of innovations within the assimilation window is not taken into account. On the other hand, the usage of diagnosed background and observation error covariances would not change the weight given to observations and background respectively in the analysis provided

$$\begin{aligned} \mathbf{H}\tilde{\mathbf{B}}\mathbf{H}^T (\mathbf{H}\tilde{\mathbf{B}}\mathbf{H}^T + \tilde{\mathbf{R}})^{-1} &= \mathbf{H}\mathbf{B}\mathbf{H}^T (\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})^{-1} \\ \tilde{\mathbf{R}} (\mathbf{H}\tilde{\mathbf{B}}\mathbf{H}^T + \tilde{\mathbf{R}})^{-1} &= \mathbf{R} (\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})^{-1}. \end{aligned}$$

In this work, diagonal elements of the covariance matrix of innovations  $E[\mathbf{d}^o_b \mathbf{d}^{o_b T}]$ , of the statistical expectation of the cross product of differences between analysis and background in observation space and innovations,  $E[\mathbf{d}^a_b \mathbf{d}^{o_b T}]$ , and that of analysis residuals and innovations

$E[\mathbf{d}_a^o \mathbf{d}_b^{oT}]$  have been obtained. Taking the squared root of diagonal elements of  $\mathbf{H} \tilde{\mathbf{B}} \mathbf{H}^T$  and  $\tilde{\mathbf{R}}$ , the corresponding diagnosed background and observation error standard deviations are:

$$\sigma_b = \sqrt{\sum_{i=1}^p ((Hx^a)_i - (Hx^b)_i)(y_i - (Hx^b)_i) / p}$$

$$\sigma_o = \sqrt{\sum_{i=1}^p (y_i - (Hx^a)_i)(y_i - (Hx^b)_i) / p}$$

where  $p$  denotes the number of observations of a particular parameter from a particular observation type. Furthermore,  $y_i$  represents observation number  $i$  while  $(Hx^b)_i$  and  $(Hx^a)_i$  are the corresponding background and analysis state equivalents.

The error statistics diagnosing software is a program accumulating and calculating statistics of different combinations of innovations and residuals for different observation types and variables. The innovations and residuals are obtained from the resultant ASCII files when HIRVDA's `cmastat.x` is executed. The latter extracts analysis information from an ACMA-file, with namelist-variable `ldetail` set to `.true.`. The results of the diagnosis are  $\sigma_b$  and  $\sigma_o$  values as determined from the innovations and residuals of the respective observation types and variables. These values can then be compared to the prescribed HIRVDA statistics for  $\sigma_b$  and  $\sigma_o$ . The specified  $\sigma_b$  for all types of observations can be obtained by running BGOS (background errors in observation space), which is part of the HIRVDA-software. Then, a global check on the innovations variance, as the separate consistency diagnostics on background errors in observation space and on observation errors, can be done.

It should be pointed out that if one in the future utilizes the experimentally determined statistics to modify the HIRVDA pre-scribed  $\sigma_b$  and  $\sigma_o$  values, the statistics should be re-estimated based on the new prescribed  $\sigma_b$  and  $\sigma_o$  values. This is because the innovations and residuals are themselves dependent on the prescribed error statistics.

The method has been applied to four extensive operational data sets:

- SMHI archived statistics from the 11-month period between 20040601 and 20050501. The statistics are from the operational 3D-Var runs using the current reference HIRLAM structure functions which were derived with the NMC method using analytical balance approach from the archived SMHI operational data. As in the reference HIRLAM, a scaling factor of 0.6 is used. SMHI-HIRLAM runs with a 6 h assimilation cycle and a model domain with a 22 km horizontal grid-spacing and 40 vertical levels. The model domain is shown in Figure 1, labelled as "C22".
- DMI archived statistics for the one-year period between 20040901 and 20050831. DMI-HIRLAM also uses the Reference HIRLAM structure functions, with a scaling factor initially at 0.6, but 0.4 from June 2005 onwards. The DMI-HIRLAM runs with a complicated scheduling including a re-assimilation cycle, but the effective assimilation interval is 3 h. DMI-HIRLAM runs on a model domain with a 16 km

horizontal grid-spacing and 40 vertical levels. The DMI-HIRLAM model domain is shown in Figure 1, labelled as “T15”.

- FMI-HIRLAM statistics, archived at ECMWF, for the 13-month period between 20040401 and 20050531. FMI-HIRLAM uses reference HIRLAM as its operational system (RCR), thus the structure function in its 3D-VAR run is the same as in SMHI-HIRLAM. FMI-HIRLAM runs with a 3 h assimilation cycle. The RCR model domain has a 22 km grid-spacing and 40 vertical levels. The model domain is shown in Figure 1, labelled as “RCR”.
- INM archived statistics for the 3-month period between 20050301 and 20050531. INM-HIRLAM uses the structure function derived from a two-month period of operational FMI model forecasts archived at ECMWF, based on the NMC method and a statistical approach. The operational INM-HIRLAM runs with a 6 h assimilation cycle, on a model domain with a 17 km horizontal grid-spacing and 40 vertical levels. The model domain is shown in Figure 1, labelled as “ONR”.

It is worth noting here that apart from the above-mentioned differences in structure functions, assimilation interval and domain size, there are additional differences between the configurations of the four operational assimilation systems, such as those in data usage and observation cut-off limit, differences in horizontal diffusion, physical parameterisation and re-assimilation cycles utilizing late arriving observation data and ECMWF analyses and forecasts.

Using the archived FMI-RCR data at ECMWF, the seasonal evolution of the error standard deviations has also been investigated. The study is performed by applying the diagnosis to four subsets of data, each consisting of a three-month period representing spring (20040401-20040630), summer (20040701-20040930), fall (20041001-20041231) and winter (20050101-20050331) conditions, to obtain the seasonal estimations of  $\sigma_b$  and  $\sigma_o$ .

Furthermore, the latitudinal variation of the diagnosed  $\sigma_b$  and  $\sigma_o$  values has been studied using the INM statistics from spring 2005, for radiosonde innovations and residuals of temperature, specific humidity and wind components in four latitude bands, as detailed below.

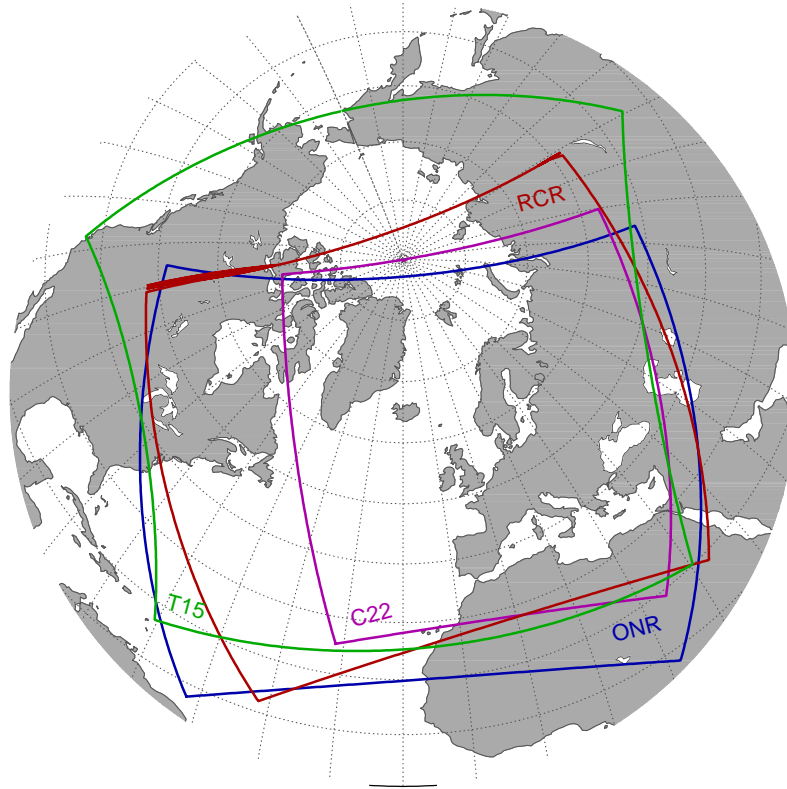


Figure 1: Operational model domains from which archived analyses data sets are diagnosed in this study. Shown in the plots are SMHI-C22 (magenta), DMI-T15 (green), FMI-RCR (red) and INM-ONR (blue).

### 3. Results

Using the diagnosis software developed in this study, the background and observation error standard deviations have been calculated for the above-mentioned operational data sets. The results obtained are presented below.

#### 3.1 Results of the different operational runs

First, the diagnostics of the three approximately one-year long operational data sets from SMHI, DMI and FMI, and the three spring months for INM, are shown together. For each conventional observation type (TEMP, AIREP, PILOT, SYNOP, SHIP and DRIBU) and variable (temperature, u-wind component and surface pressure) the square root of innovations variance is presented, followed by the estimated background and observation error standard deviations along with the prescribed area mean background error standard deviation, obtained through running BGOS with both analytical and statistical balance background constraints, and the assumed vertical profile of observation errors. In order to illustrate the weight given in the analysis to each observed variable, the ratios of observation error variance to innovation variance, and that of background error to innovation variance, are also shown.

Second, the diagnosed  $\sigma_b$  and  $\sigma_o$  values for different seasons for a 1-year period from 20040401 to 20050331, using the RCR dataset, are presented to demonstrate the annual variability of error variances.

Finally, latitudinal variation is shown through the obtained innovation variances and diagnosed  $\sigma_b$  and  $\sigma_o$  values for temperature, specific humidity and wind components derived from INM radiosonde statistics in spring 2005 for four latitude bands, 15N-30N, 30N-45N, 45N-60N and 60N-75N.

All Tables and Figures referred to within this section can be found attached at the end of this report.

### *Temperature*

Figures 2 and 3 show the vertical profile of the square root of innovation variances obtained from the statistics of radiosondes and aircrafts temperature observations in the different centers (INM, SMHI, DMI and FMI). Although the operational configurations are different and the time period is uneven (much shorter in case of INM), the magnitude of the innovations is very similar for all centers but there is a greater spread in the case of AIREP data, the smaller innovations corresponding to DMI operational dataset. AIREP temperature innovations are shown to be smaller than those for radiosondes, especially near the ground. However, it should be noted that low-level aircraft data are mostly concentrated in western Europe, whereas radiosonde stations are more scattered over the model domain.

Figures 4 and 5 present  $\sigma_b$  and  $\sigma_o$  values as derived from radiosonde temperature innovations and residuals statistics, respectively. In Figure 4 the vertical profile of the prescribed  $\sigma_b$  values at different centers is also represented. They are obtained by running the BGOS option of the HIRVDA software using analytical and statistical balance background constraint formulations, respectively. A value of 0.6 has been used as scaling factor to the original background error variances (in case of statistical balance) or standard deviations (in case of analytical balance), for the control variables that were tuned using the NMC method. The diagnosed  $\sigma_b$  values from all centers (Figure 4) appear smaller than the prescribed, and the differences between centers are larger than those observed in the innovation variance (Figure 2). Although the estimated  $\sigma_b$  value with INM statistics is the largest among all centers, it is still smaller than the area mean of that estimated using BGOS with the statistical background constraint. In Figure 5 the actually used radiosonde temperature observation errors have also been plotted. The obtained values show a better agreement among the different datasets. Within the troposphere, the INM  $\sigma_o$  is the smallest.

Figures 6 and 7 are the plots for AIREP temperature which correspond to Figure 4 and 5. Again the prescribed error standard deviations are seen to be larger than those estimated. The diagnosed  $\sigma_b$  value for temperature from TEMP and AIREP are not consistent, with smaller values for the AIREP data. It is probably worth noticing that the size of AIREP temperature innovations is also smaller than that for TEMP. Although the innovation magnitudes for AIREP temperature are very similar among INM, RCR and SMHI data, the partitioning between observation and background errors for INM differs from the other two centers.

Figures 8 and 9 present the ratios of observation error variance and background error variance to the innovation variance derived from the radiosonde and AIREP temperature statistics. Figure 8 could explain the differences observed in Figures 4 and 5 between INM and RCR

although the radiosonde temperature innovations are very similar. INM assumed error statistics give more weight to radiosonde temperature observations within the troposphere. All centers using analytical balance constraint show similar partitioning between background and observation errors. In case of AIREP data, the weight given to observations in the analysis is larger in the INM operational suite. This is also likely to be associated with the observed differences between RCR and INM in Figures 6 and 7.

### *Wind*

The same type of figures shown for temperature have been produced for wind components. Since similar features are observed in meridional and zonal winds, only results for the u-wind component are presented here. Figures 10 and 11 show the vertical profile of the square root of innovations variance obtained from the statistics of radiosonde and aircraft wind observations at each center, respectively. As in the case of temperature, the size of innovations is very similar for all centers. SMHI, INM and FMI have very similar AIREP wind innovations which are higher than those of DMI. Again the u-wind component innovations for AIREP are seen to be smaller than those for radiosonde.

Figures 12 and 13 present  $\sigma_b$  and  $\sigma_o$  values derived from radiosonde u-component statistics. The vertical profile of the prescribed  $\sigma_b$  value has also been plotted in the corresponding figure. The diagnosed  $\sigma_b$  value in all centers (Figure 12) looks much smaller than the prescribed ones in the statistical and analytical balance statistics. The differences between centers are larger than those observed in the innovation variance (Figure 10). The estimated  $\sigma_b$  value with INM statistics is the largest of all centers although the INM wind innovation size is very similar to that of FMI. On the other hand, the INM  $\sigma_b$  value is still smaller than the area mean of that estimated using BGOS with the statistical background constraint. The estimated  $\sigma_o$  values for RCR, SMHI and DMI agree fairly well but the INM  $\sigma_o$  value stands out as the smallest. It is observed that  $\sigma_b$  value estimated from BGOS in case of statistical balance is very large at upper levels. This feature is consistent with the large background error standard deviations at these levels mainly for unbalanced divergence, but also for vorticity, used in the INM operational analysis. As it has been already mentioned, these values have been found calculating background error statistics using RCR forecast differences. Different datasets of forecasts, based on INM HIRLAM configuration but at different vertical resolution, have also been used to derive background errors by NMC method. The obtained length scales and vertical profiles of vorticity and unbalanced divergence error standard deviations show to be much smaller than the values found for RCR forecast differences (not shown here).

Figures 14 and 15 show the diagnosed observation and background error standard deviation derived from AIREP wind statistics. Here the same type of features observed for radiosonde wind is observed. DMI statistics seems to be slightly different from SMHI and RCR results, but it has to be taken into account that the size of DMI wind innovations is clearly smaller.

Figures 16 and 17 present the ratios of observation error variance and background error variance to the innovation variance derived from the radiosonde and AIREP u-wind component statistics. They seem to justify the differences found in Figures 12-15 between INM and the rest of the centers. As in the case of temperature, both wind observations from radiosonde and aircraft reports have more influence in the INM upper air analysis than in the other centers.

Figures 18-21 show the u-wind component innovations size, diagnosed error standard deviations and ratios of background and observation error variances to innovation variance, obtained from PILOT data statistics. The vertical profile of innovation variances seems to be very similar to that of radiosondes, but larger near the ground. There is a good agreement between the different centers about the size of the innovations, but again the different background constraint used in the analysis leads to a different partitioning of innovation errors into background and observation error parts. On the other hand, it is observed that within the same center, wind background errors derived from TEMP and PILOT statistics are not consistent, probably due to the prescribed observation error statistics, which gives a larger error to wind observations from PILOT reports.

### *Surface pressure*

Tables 2-5 show  $\sigma_b$  and  $\sigma_o$  values for surface pressure as derived from innovations and residual statistics of SYNOP, DRIBU and SHIP data, respectively, as well as the values presently used in the reference HIRLAM. From these tables, diagnosed  $\sigma_b$  values within each center for different types of reports are seen to be very consistent, slightly larger over sea than over land. However, there are clear differences between the different centers. Estimated  $\sigma_o$  values present similar behaviour for the different centers: SYNOP surface pressure observation error is smaller than that of DRIBU, still smaller to that of SHIP reports. In case of SYNOP, the diagnosed background and observation error standard deviations are very similar among centers except for DMI, which presents the smallest  $\sigma_b$  values. However, innovation variance seems to be very similar for all SYNOP datasets, although slightly larger for INM and DMI statistics (not shown here). It has to be taken into account that probably SYNOP is the most sensitive dataset to the different model domains. Besides, the INM dataset covers only a short period in spring 2005. INM, FMI and SMHI analyses give a similar weight to the background and surface observations but DMI draws more to the background. From the results obtained, it seems that the area mean background error standard deviation produced by BGOS is not consistent with the obtained partitioning of the total innovation error into background and observation errors. Instead, the results suggest that the estimated background error standard deviation is very similar to the prescribed SYNOP observation error standard deviation. With respect to the size of  $\sigma_o$  values for sea reports, the estimated values from DMI and FMI data are smaller than the specified values, while from SMHI and INM statistics, the estimated values are more similar to the specified ones.

## **3.2 Seasonal variation of diagnosed background and observation standard deviation**

The estimates of  $\sigma_b$  and  $\sigma_o$  values for spring (20040401-20040630), summer (20040701-20040930), fall (20041001-20041231) and winter (20050101-20050331) conditions and presented in Figures 22-25 have been derived from radiosonde and AIREP temperature and wind RCR statistics.

### *Temperature*

Looking at the diagnosed background error standard deviation for temperature derived from radiosonde statistics (Figure 22), the differences between seasons appear rather small within the troposphere. For the stratosphere, an increase of  $\sigma_b$  values can be observed. However, the background error standard deviation derived from AIREP statistics (Figure 24), at mid-low levels presents a larger seasonal variability, reaching the minimum values in spring-summer. Again, the background error standard deviation for temperature obtained from AIREP looks

to be smaller than that diagnosed from TEMP. Meanwhile, the seasonal variation in the diagnosed observation errors (Figures 23 and 25) is more pronounced than that of the estimated background errors. The variation in diagnosed  $\sigma_b$  and  $\sigma_o$  is obviously associated with that of innovations caused by the seasonality of background errors. The temperature background error standard deviation values used in the analysis above 250 hPa have a monthly variation, but the amplitude of the prescribed annual oscillation seems to be very small. The results obtained can be explained by the slightly varying in the upper levels or constant below 250 hPa error statistics assumed in the analysis during the year, in particular the partitioning of innovation error into observation and background error parts, with observation errors larger than background errors, as was shown in Figures 8 and 9. Diagnosed observation errors are then more affected by innovations seasonality as a consequence of their higher values relative to those of background errors.

### *Wind*

The u-wind component presents very similar behaviour to that observed in temperature, as seen in Figures 26-29. Some seasonality appears in  $\sigma_o$  values due to the lack of annual variation of the prescribed error statistics. Hence, the true variability of background errors will probably be higher than that obtained in our diagnosis, mainly in the stratosphere.

### **3.3 Latitudinal variation of diagnosed background and observation standard deviation**

Latitudinal differences found in temperature and wind innovations using the spring INM dataset are much more pronounced than those observed between the different centers. However, specific humidity is the variable showing the largest variation with latitude.

#### *Temperature*

Figure 30 shows the vertical profile of radiosonde temperature innovations. Within the troposphere they present less variation than in the stratosphere. In mid-upper levels of the troposphere temperature innovations increase from south to north. In the stratosphere, they show to increase from north to south. The diagnosed background and observation error standard deviations are plotted in Figures 31 and 32, respectively.  $\sigma_o$  values seem to be very close in the troposphere and differ in the stratosphere where the estimated observation errors become much larger than the background errors, as is observed in Figure 33. Figure 31 shows a more comparable but opposite variation of  $\sigma_b$  values in the troposphere and the stratosphere. The observed decrease at mid-upper tropospheric levels seems to be a consequence of the balance imposed in the statistical background constraint, where the choice of the control variables, vorticity and residuals of divergence, temperature and specific humidity, allow the geopotential error variance to decrease with decreasing latitude.

#### *Specific humidity*

Figures 34-37 show the results found for specific humidity. Innovations decrease clearly from south to north, so does also the estimated observation error standard deviation, and to a lesser extent also the diagnosed  $\sigma_b$  values. Most of the prescribed specific total background error variance for specific humidity is due to the unbalanced specific humidity term that it is assumed constant in the model domain. On the other hand, specific humidity observation errors specified in the HIRLAM analysis are derived from those of relative humidity. Relative humidity observation errors in the HIRVDA code are specified by making use of an empirical

formula, which includes also dependency on temperature. Since the prescribed background error does not vary too much but prescribed observation errors do, the ratios of background and observation error to innovation variance do vary in the model domain as is seen in Figures 34 and 37. Therefore, the weight of specific humidity observations in the INM analysis decreases from north to south.

### **Wind**

The square root of innovation variance for radiosonde wind u-component is plotted in Figure 38. It seems to indicate an increase of wind errors from north to south, although there is a minimum in the lowest troposphere in the 45N-60N latitude band, a data rich area. This is in agreement with the low wind innovations found in AIREP statistics (see Figure 11). It could indicate either an overestimation of AIREP wind observation errors or an underestimation of radiosonde wind observation error. Diagnosed observation errors (Figure 40) are very consistent within the troposphere, where observation errors are much smaller than background errors (Figure 41) but they increase towards the equator in the upper levels. So, it seems that the true  $\sigma_b$  values should present an even larger increase towards the equator. It has to be taken into account that the wind analysis becomes less coupled to the mass analysis at decreasing latitude, where observations are sparse, especially for wind data (in particular when AMV (Atmospheric Motion Vectors) data are not used).

## **4. Conclusions**

The method proposed by Desroziers et al. (2005) has been demonstrated to be very useful in analysis diagnosis and our investigations have produced many interesting results which are insightful for our understanding of HIRLAM 3D-VAR analysis. The analysis diagnosis also provides an efficient tool which can be helpful for the tuning of HIRVDA analysis system.

The derived background and observation error standard deviations obtained from innovations and residuals slightly differ from one NWP center to another. However, some general conclusions can be drawn. Table 1 shows a summary of the consistency checks performed on innovations, on the background errors and on the observation errors used in the analysis to complement the plots that have been commented before.

The main findings of this study are:

1. The size of innovations is smaller than the sum of observation and background error variances assumed in the analysis within the troposphere. On the contrary, in the stratosphere the temperature innovations are larger than the sum of specified observation and background errors.
2. The diagnosed background error variances are found to be smaller than the prescribed in the HIRLAM 3D-VAR for all variables of the different observation types. In case of wind, the larger discrepancy between diagnosed and specified errors is observed in the background error statistics used in the INM operational suite.
3. Within the troposphere, the diagnosed observations errors are found to be smaller than the assumed in the HIRLAM 3D-VAR. The specified observation errors for aircraft observations are larger than the size of the corresponding innovations, indicating an overestimation of the values prescribed in the analysis. In the stratosphere, the

diagnosed observation errors are larger than the specified, but as it was said before, the size of innovations there is larger than the assumed in the analysis error statistics.

4. The partitioning of total error of innovations into background and observation parts varies from one center to another. The method applied to diagnose background and observation errors is based on time series of innovations and analysis residuals. It implicitly takes into account the weight given to observations and background respectively within the analysis to calculate the background and observation error standard deviation values. The size of innovations variance at all NWP operational suites is very similar, despite the fact that these statistics are associated with different time periods, model configurations, assimilation cycling, observations usage and background errors. The different background error formulation is likely the reason for the observed differences of diagnosed background and observation errors, between the INM results (making use of the statistical balance constraint) and those obtained at SMHI, RCR and DMI (based on the analytical background error formulation).
5. Seasonal and latitudinal variations have been observed in innovations. These dependences are markedly present in the stratosphere in the case of temperature and wind. The obtained monthly variation of innovations agrees qualitatively with the results found previously by Lindskog (2000). When the method proposed by Desroziers et al. (2005) is applied such dependence is found both for diagnosed values of observation and background errors. The diagnosed observation errors show the largest dependence on latitude, mostly in the upper troposphere and stratosphere, for temperature and wind, and in the lowest troposphere for humidity. The variations in the diagnosed observation errors with season and latitude are found at the levels where they are larger than the background errors. This might indicate that some tuning is needed to increase the weight given to observations in the analysis at these levels and to account for seasonality and latitudinal effects.

It is anticipated that the estimated vertical profiles of the innovation variance for the different observation types can be used to tune parameters of some quality control checks, such as the first-guess check performed in the screening step of HIRVDA. However, the tuning of the *a priori* probability of gross errors in the Variational QC seems to be more difficult if each operational suite is using different error statistics, because the fit of the analysis to observations will vary with the size of background error standard deviations and with cross-covariance relationships. The monitoring results could however provide guidance for re-tuning of pre-scribed error statistics within the HIRVDA minimization.

## 5. Acknowledgements

Nils Gustafsson, SMHI, participated in the initial discussion and planning of the current monitoring studies. Alberto Cansado and Jana Sánchez, INM, have prepared the dataset of innovations and analysis residuals of the INM HIRLAM operational suite used in this work. The authors want to thank Per Undén and an anonymous reviewer for their valuable suggestions that have helped to improve the original manuscript.

## 6. References

- Andersson E., 2003. Modelling of innovation statistics. *ECMWF seminar proceedings*, Reading, September 2003.
- Berre, L., 2000. Estimation of synoptic and meso scale forecast error covariances in a limited area model. *Mon. Wea. Rev.*, **128**, 644-667.
- Courtier, P., Andersson, E., Heckley, W., Pailleux, J., Vasiljević, 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. R. Meteorol. Soc.*, **124**, 1783-1808.
- Desroziers, G, L. Berre, B. Chapnic and P. Poli, 2005. Diagnosis of observation background and analysis error statistics in observation space. To appear in *Q. J. R. Meteorol Soc.*
- Gustafsson, N., Berre, L., Hörnquist, S., Huang, X.-Y., Lindskog, M., Navascués, B., Mogensen, K. S. and Thorsteinsson, 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. 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.
- Lindskog, M. , 2000. An estimate of the seasonal dependence of background error statistics in the HIRLAM 3D-Var. *HIRLAM Newsletter*, **35**, April 2000, pp. 71-86.
- Lindskog, M., Gustafsson, N., Navascués, B., Mogensen, K. S., Huang, X.-Y., Yang, X., Andræ, U., Berre, L., Thorsteinsson, S. and Rantakokko, J., 2001. Three-dimensional variational data assimilation for a limited area model. Part II: Observation handling and assimilation experiments. *Tellus*, **53A**, 447-468.
- Parrish, D.F. and Derber, J.C., 1992: The National Meteorological Centre's spectral statistical interpolation analysis system. *Mon. Wea. Rev.*, **120**, 1747-1763.
- Rabier, F. Mc Nally, A., Andersson, E., Courtier, P., Undén, P., Eyre, J., Hollingsworth, A. and Bouttier, F., 1998. The ECMWF implementation of three-dimensional variational assimilation (3D-Var). Part II: Structure functions. *Q. J. R. Meteorol. Soc.*, **124**, 1809-1830.

<b>Variable</b>	<b>Global check on innovations</b>	<b>Consistency of <math>\sigma_b</math></b>	<b>Consistency of <math>\sigma_o</math></b>
<b>Temperature (TEMP)</b>	$(\sigma_o^2 + \sigma_b^2)_{diag} > (\sigma_{o\_op}^2 + \sigma_{b\_bgos}^2)$ in the stratosphere $(\sigma_o^2 + \sigma_b^2)_{diag} < (\sigma_{o\_op}^2 + \sigma_{b\_bgos}^2)$ in the troposphere	$\sigma_{b\_diag.} < \sigma_{b\_bgos}$	$\sigma_{o\_diag.} < \sigma_{o\_op}$ in the troposphere $\sigma_{o\_diag.} > \sigma_{o\_op}$ in the stratosphere
<b>Temperature (AIREP)</b>	$(\sigma_o^2 + \sigma_b^2)_{diag} < (\sigma_{o\_op}^2 + \sigma_{b\_bgos}^2)$	$\sigma_{b\_diag.} < \sigma_{b\_bgos}$	$\sigma_{o\_diag.} < \sigma_{o\_op}$
<b>Wind (TEMP)</b>	$(\sigma_o^2 + \sigma_b^2)_{diag} < (\sigma_{o\_op}^2 + \sigma_{b\_bgos}^2)$	$\sigma_{b\_diag.} < \sigma_{b\_bgos}$	$\sigma_{o\_diag.} < \sigma_{o\_op}$ in the troposphere $\sigma_{o\_diag.} > \sigma_{o\_op}$ in the stratosphere
<b>Wind (AIREP)</b>	$(\sigma_o^2 + \sigma_b^2)_{diag} < (\sigma_{o\_op}^2 + \sigma_{b\_bgos}^2)$	$\sigma_{b\_diag.} < \sigma_{b\_bgos}$	$\sigma_{o\_diag.} < \sigma_{o\_op}$

Table 1: Summary of consistency checks on innovations, background and observation error standard deviations.  $(\sigma_o^2 + \sigma_b^2)_{diag}$  is the diagnosed innovation variance,  $(\sigma_{o\_op}^2 + \sigma_{b\_bgos}^2)$  is the sum of the observation error variance and the background error in observation space variance assumed in the analysis.  $\sigma_{b\_bgos}$  is the background error in observation space obtained using bgos,  $\sigma_{o\_op}$  is the observation error standard deviation value actually used in the analysis, and  $\sigma_{b\_diag.}$  and  $\sigma_{o\_diag.}$  are the values for background error and observation error standard deviation derived according to Desroziers et al. (2005).

## RESULTS IN DIFFERENT CENTERS

PRESSURE	SIGB NMC REF	SIGB SYNOP EXP	SIGB SHIP EXP	SIGB DRIBU EXP	SIGO SYNOP REF	SIGO SYNOP EXP	SIGO SHIP REF	SIGO SHIP EXP	SIGO DRIBU REF	SIGO DRIBU EXP
925	1.77	0.70	0.69	0.73	0.84	0.71	1.68	1.65	1.38	1.32
800		0.74			0.96	1.31				
600		0.55			1.2	1.18				

Table 2: SMHI surface pressure  $\sigma_b$  and  $\sigma_o$  values for different observation types [hPa], as function of pressure. 'EXP' as obtained from statistics and 'REF' as pre-scribed in HIRVDA (in the case of  $\sigma_b$  when using a scaling factor of 0.6 for the values obtained with the NMC method, using analytical balance).

PRESSURE	SIGB NMC REF	SIGB SYNOP EXP	SIGB SHIP EXP	SIGB DRIBU EXP	SIGO SYNOP REF	SIGO SYNOP EXP	SIGO SHIP REF	SIGO SHIP EXP	SIGO DRIBU REF	SIGO DRIBU EXP
975	1.77	0.55	0.58	0.58	0.84	1.03	1.68	1.33	1.38	1.08
925		0.63	1.02	0.93	0.96	1.41		2.61		1.85
800		0.75			1.2	1.76				
600		0.76				1.10				

Table3: DMI surface pressure  $\sigma_b$  and  $\sigma_o$  values for different observation types [hPa], as function of pressure. 'EXP' as obtained from statistics and 'REF' as pre-scribed in HIRVDA (in the case of  $\sigma_b$  when using a scaling factor of 0.6 for the values obtained with the NMC method, using analytical balance).

PRESSURE	SIGB NMC REF	SIGB SYNOP EXP	SIGB SHIP EXP	SIGB DRIBU EXP	SIGO SYNOP REF	SIGO SYNOP EXP	SIGO SHIP REF	SIGO SHIP EXP	SIGO DRIBU REF	SIGO DRIBU EXP
975	1.77	0.60	0.67	0.69	0.84	0.65	1.68	1.48	1.38	1.17
925		0.70	1.23	1.37	0.96	0.84		2.84		2.51
800		0.70			1.2	1.70				
600		0.60				1.21				

Table4: RCR surface pressure  $\sigma_b$  and  $\sigma_o$  values for different observation types [hPa], as function of pressure. 'EXP' as obtained from statistics and 'REF' as pre-scribed in HIRVDA (in the case of  $\sigma_b$  when using a scaling factor of 0.6 for the values obtained with the NMC method, using analytical balance).

PRESSURE	SIGB NMC REF	SIGB SYNOP EXP	SIGB SHIP EXP	SIGB DRIBU EXP	SIGO SYNOP REF	SIGO SYNOP EXP	SIGO SHIP REF	SIGO SHIP EXP	SIGO DRIBU REF	SIGO DRIBU EXP
925	1.71	0.80	0.89	0.84	0.84	0.82	1.68	1.5	1.38	1.12
800		0.84			0.96	1.9				
600		0.71			1.2	1.07				

Table5: INM Surface pressure  $\sigma_b$  and  $\sigma_o$  values for different observation types [hPa], as function of pressure. 'EXP' as obtained from statistics and 'REF' as pre-scribed in HIRVDA (in the case of  $\sigma_b$  when using a scaling factor of 0.77 for the values obtained with the NMC method, using statistical balance).

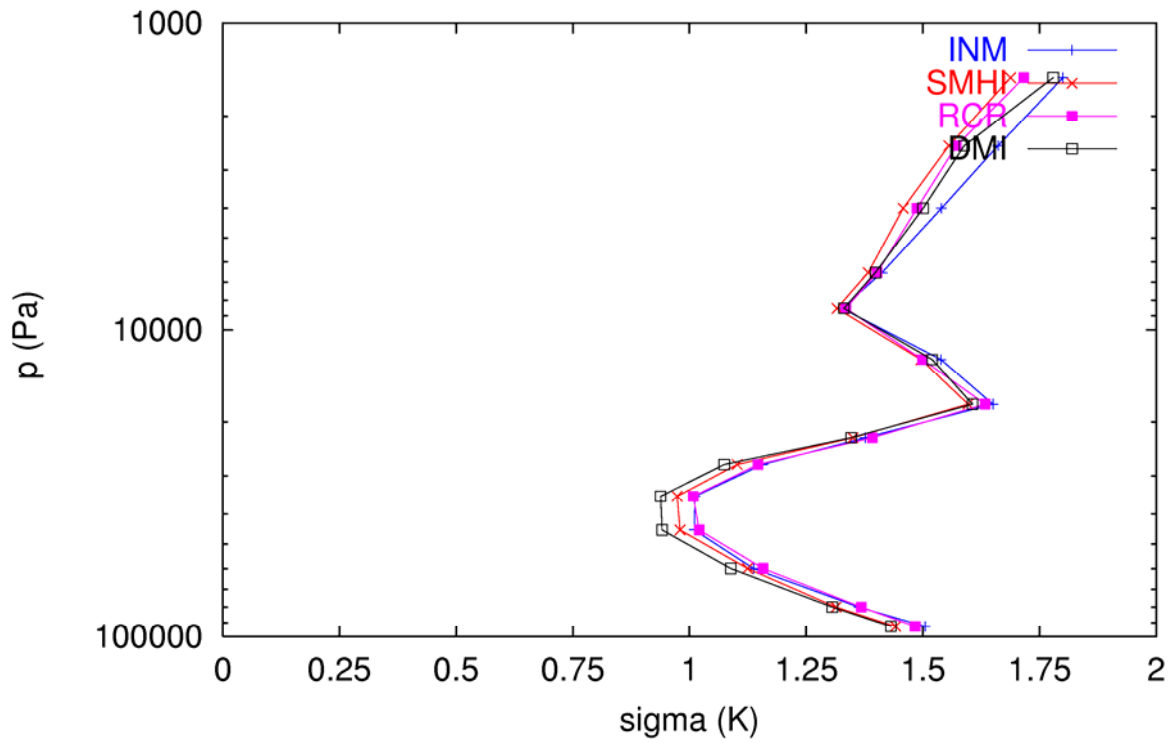


Figure 2: Square root of the sum of estimated observation and background temperature error variances (K) from TEMP statistics obtained at the different centers SMHI, DMI, FMI (RCR) and INM.

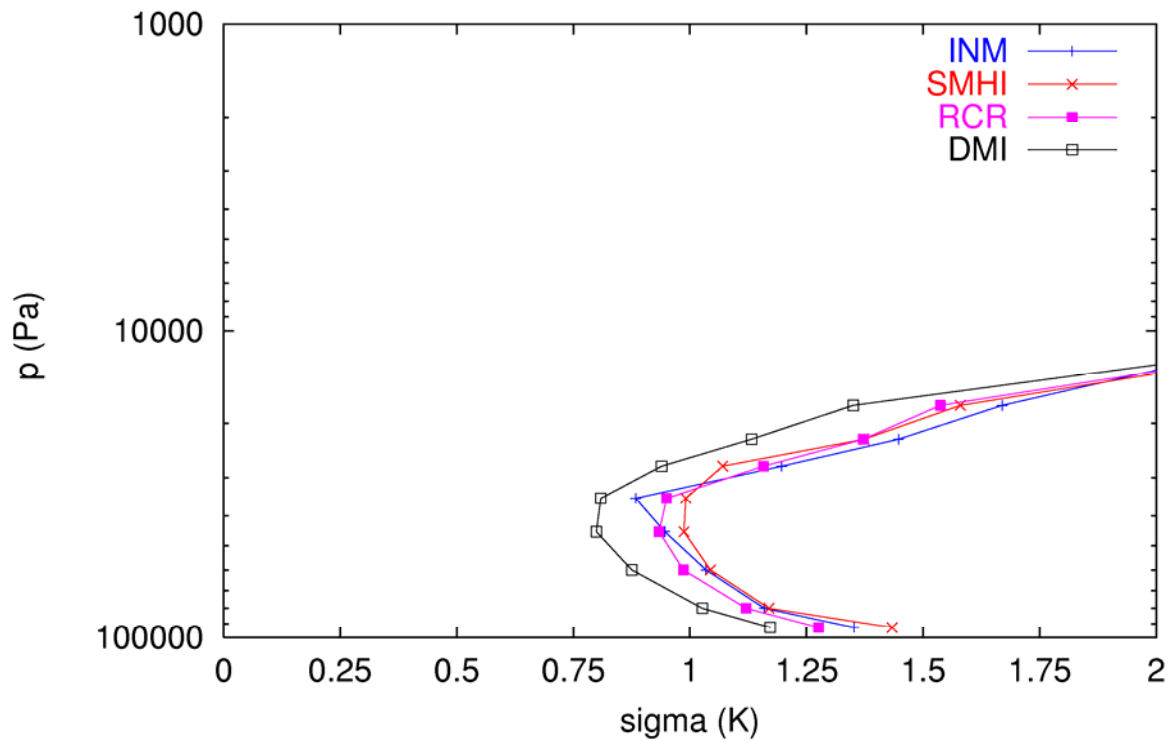


Figure 3: Square root of the sum of estimated observation and background temperature error variances from AIREP statistics obtained at the different centers SMHI, DMI, FMI and INM .

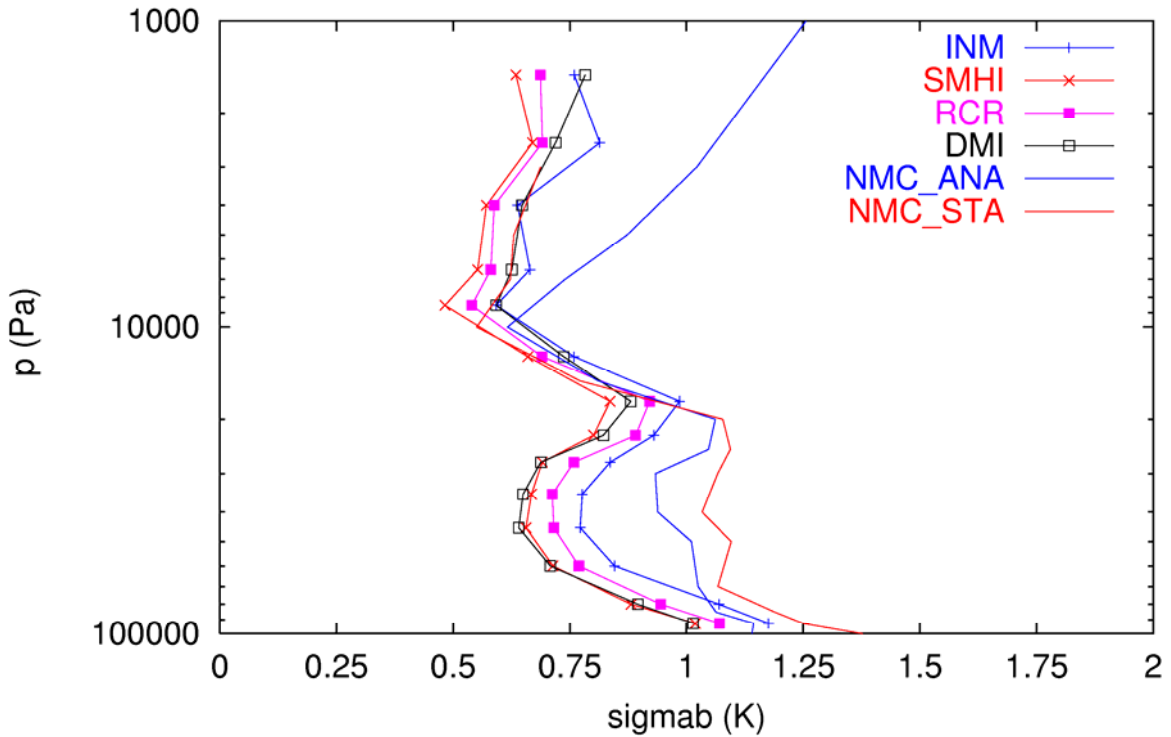


Figure 4: Estimated background error standard deviation for temperature from TEMP statistics at the different centers SMHI, DMI, RCR and INM and the area mean temperature background error standard deviation obtained from the assumed error statistics using a sample of random perturbations to the control variable for both statistical and analytical background constraint formulations (labeled as NMC\_ANA and NMC\_STA, solid lines).

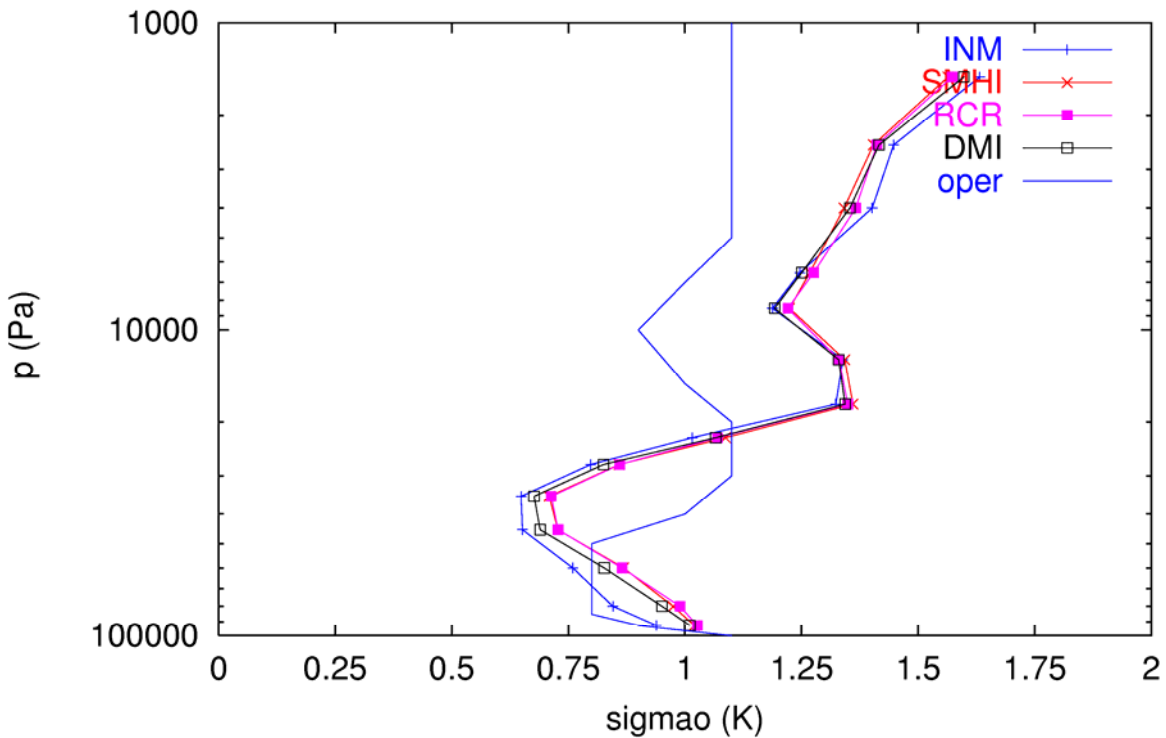


Figure 5: Estimated observation error standard deviation for temperature (K) from TEMP statistics at the different centers SMHI, DMI, FMI and INM and the assumed vertical profile of observation errors in the operational suites (solid line).

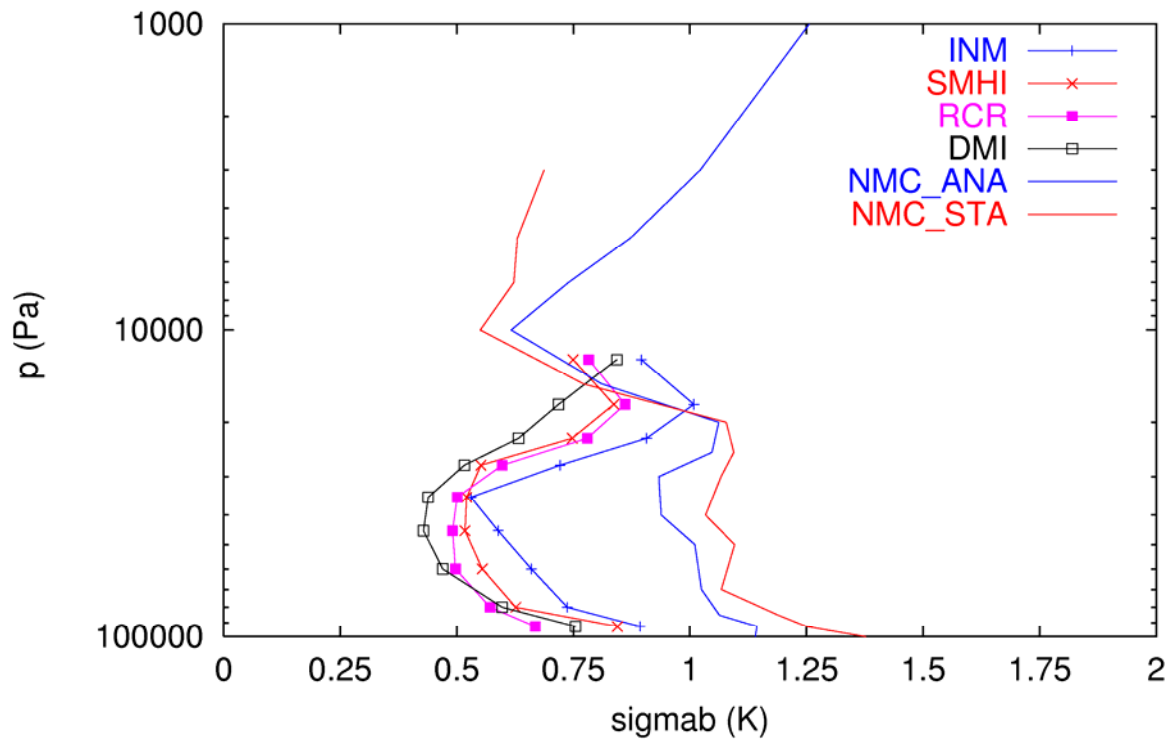


Figure 6: Estimated background error standard deviation for temperature from AIREP statistics at the different centers SMHI, DMI, RCR and INM and the area mean temperature background error standard deviation obtained from the assumed error statistics at SMHI, DMI and FMI and using a sample of random perturbations to the control variable (labeled as NMC, solid line).

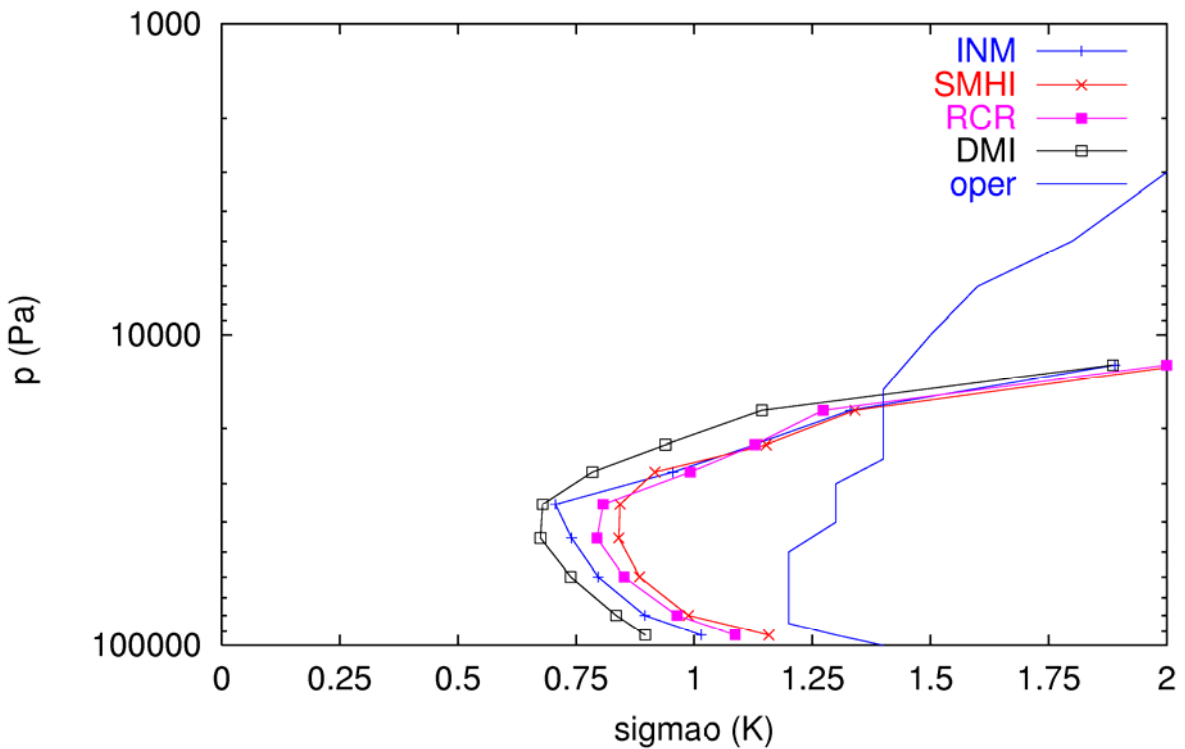


Figure 7: Estimated observation error standard deviation for temperature (K) from AIREP statistics at the different centers SMHI, DMI, FMI and INM and the assumed vertical profile of observation errors in the operational suites (solid line).

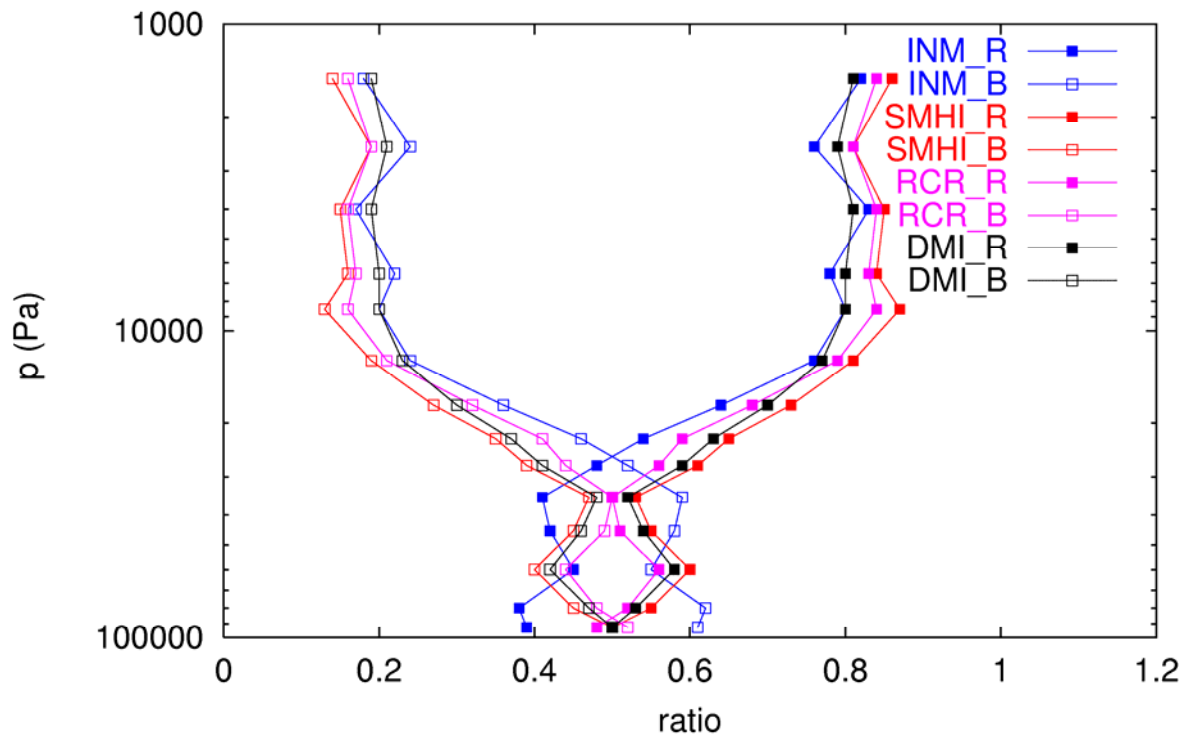


Figure 8: Observation to total error variance (filled squares) and background to total error variance (squares) ratios for temperature obtained from TEMP statistics at the different centers SMHI, DMI, FMI and INM.

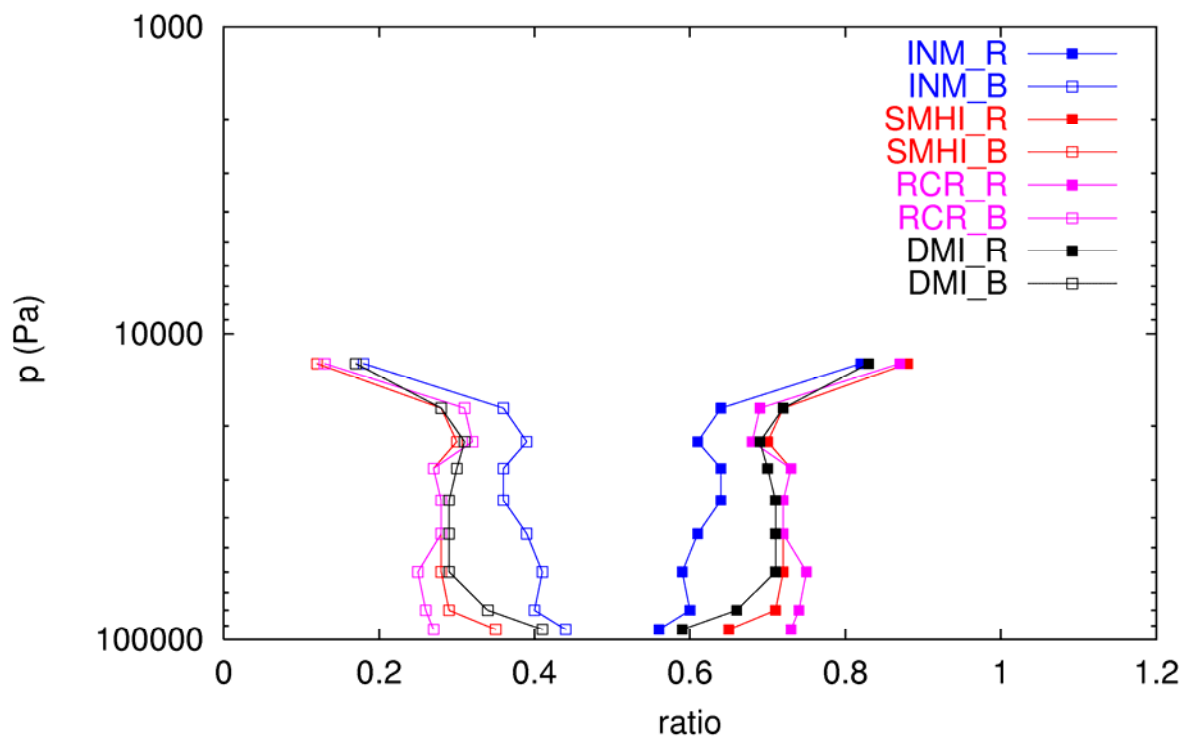


Figure 9: Observation to total error variance (filled squares) and background to total error variance (squares) ratios for temperature obtained from AIREP statistics at the different centers SMHI, DMI, FMI and INM.

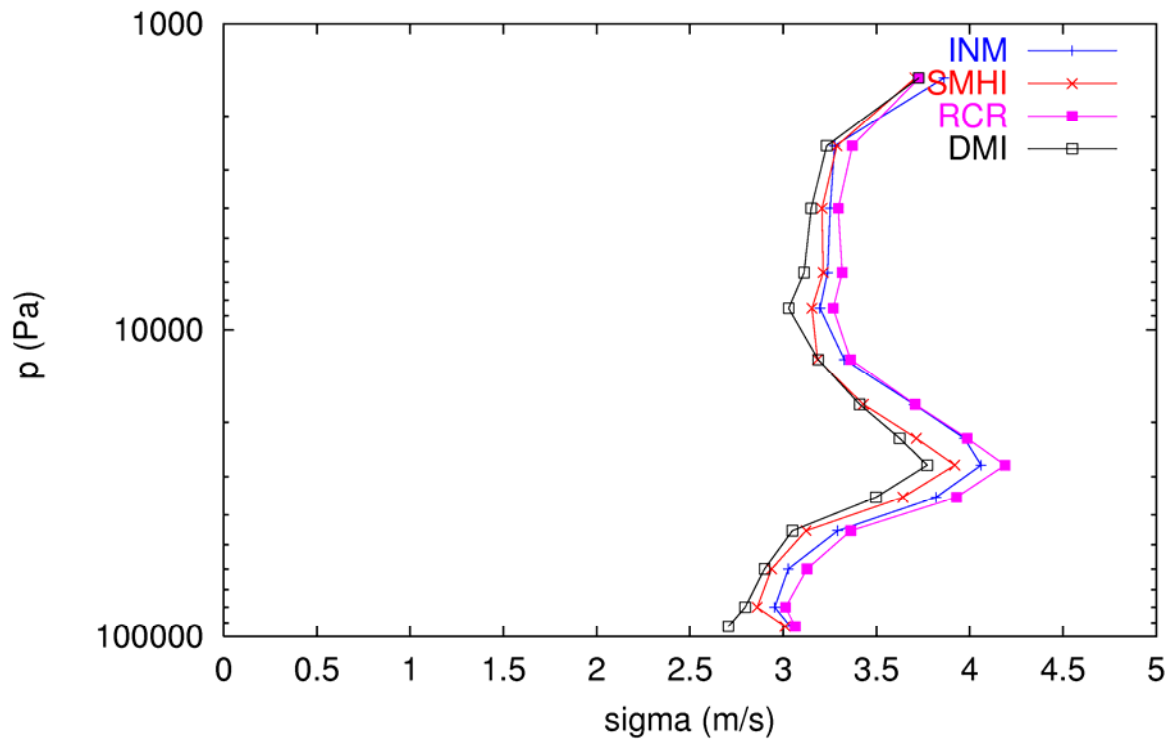


Figure 10: Square root of the sum of estimated observation and background wind u-component error variances (m/s) from TEMP statistics obtained at the different centers SMHI, DMI FMI and INM .

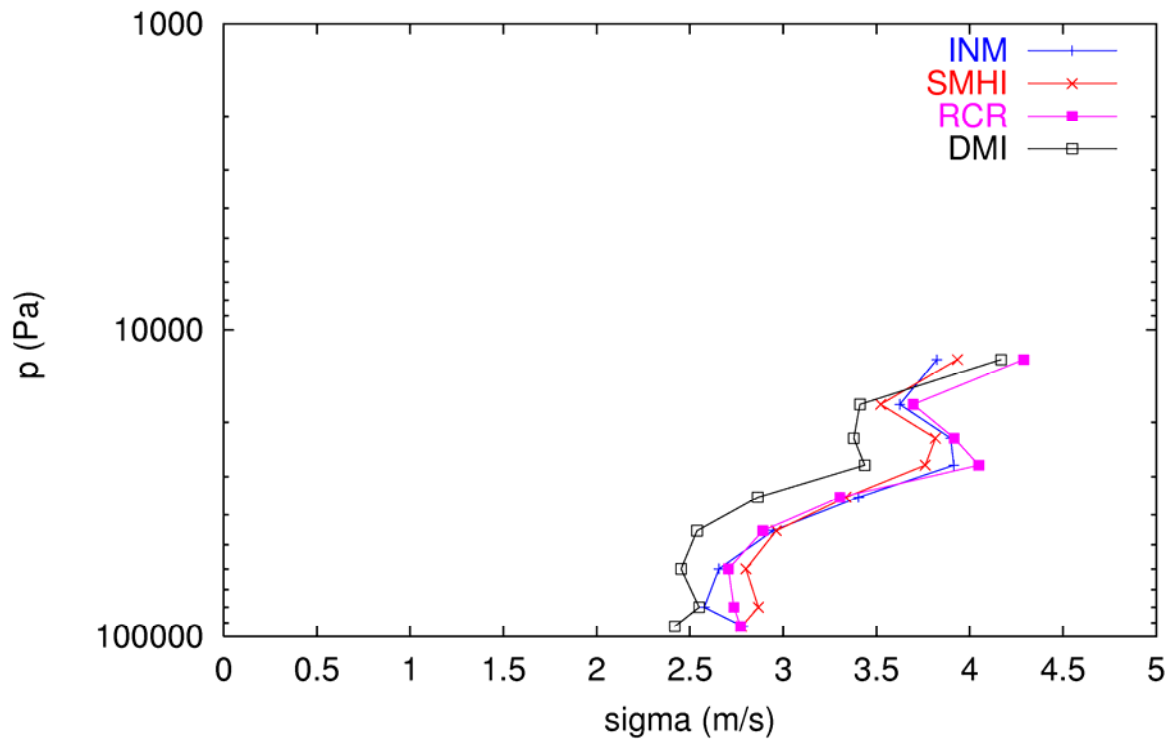


Figure 11: Square root of the sum of estimated observation and background wind u-component error variances (m/s) from AIREP statistics obtained at the different centers SMHI, DMI FMI and INM .

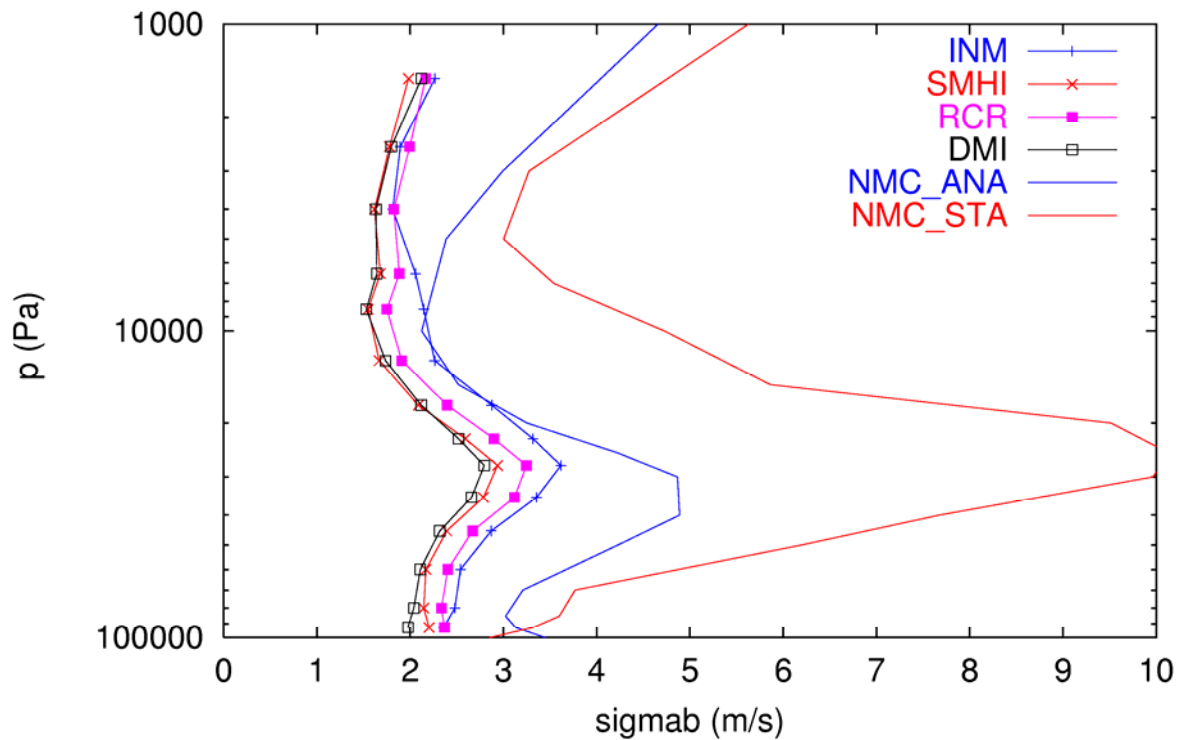


Figure 12: Estimated background error standard deviation for wind u-component (m/s) from TEMP statistics at the different centers SMHI, DMI, RCR and INM and the area mean wind background error standard deviation obtained from the assumed analytical and statistical balance error statistics using a sample of random perturbations to the control variable (labeled as NMC\_ANA and NMC\_STA, solid lines).

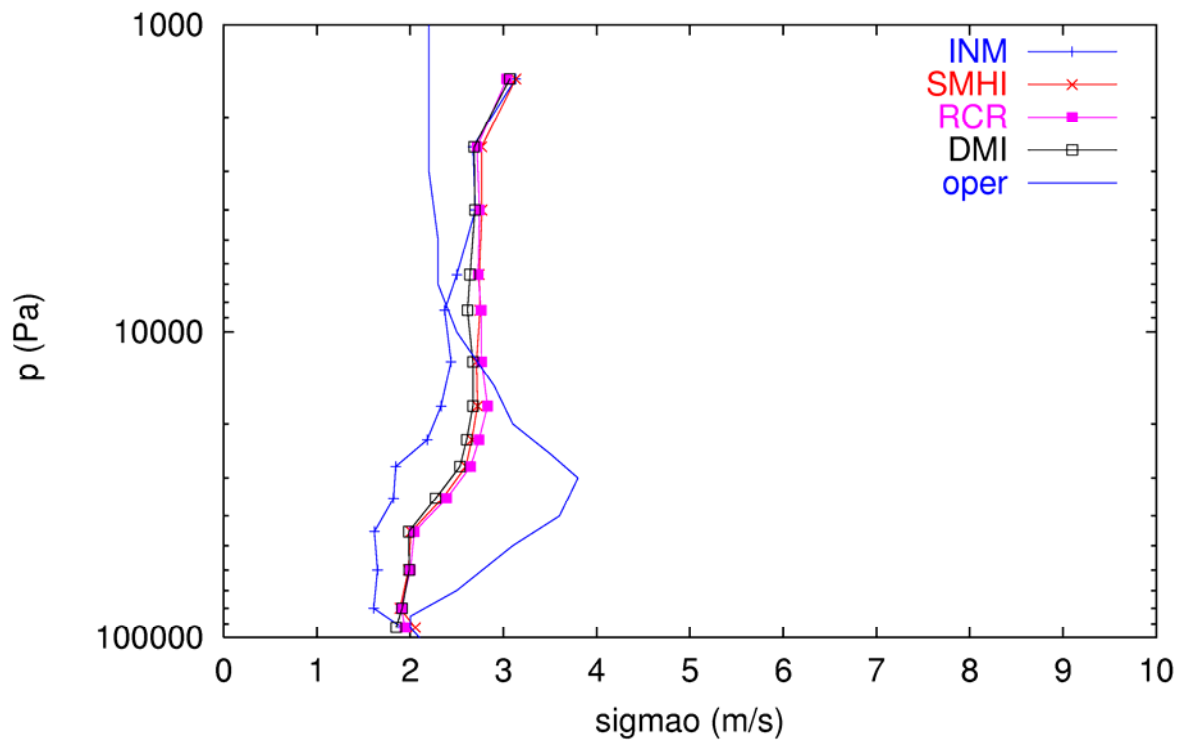


Figure 13: Estimated observation error standard deviation for wind u-component (m/s) from TEMP statistics at the different centers SMHI, DMI, FMI and INM and the assumed vertical profile of observation errors in the operational suites (solid line).

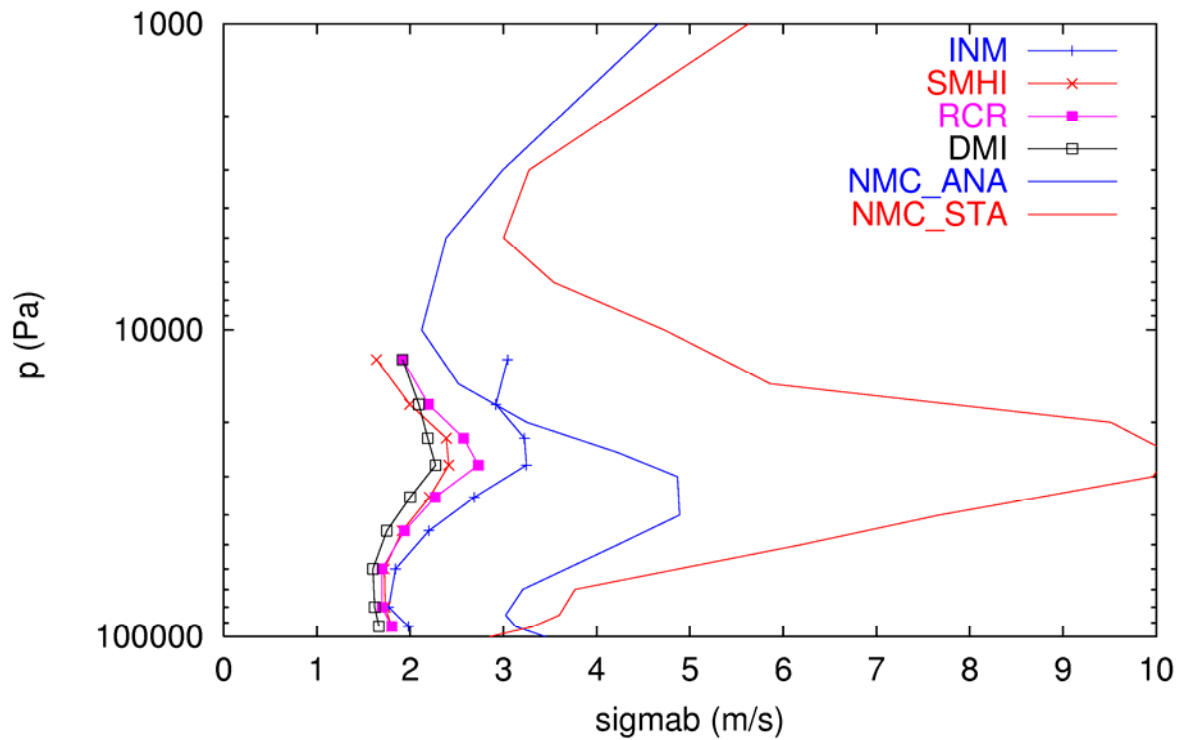


Figure 14: Estimated background error standard deviation for wind u-component (m/s) from AIREP statistics at the different centers SMHI, DMI, RCR and INM and the area mean wind background error standard deviation obtained from the assumed analytical and statistical balance error statistics using a sample of random perturbations to the control variable (labeled as NMC\_ANA and NMC\_STA, solid lines).

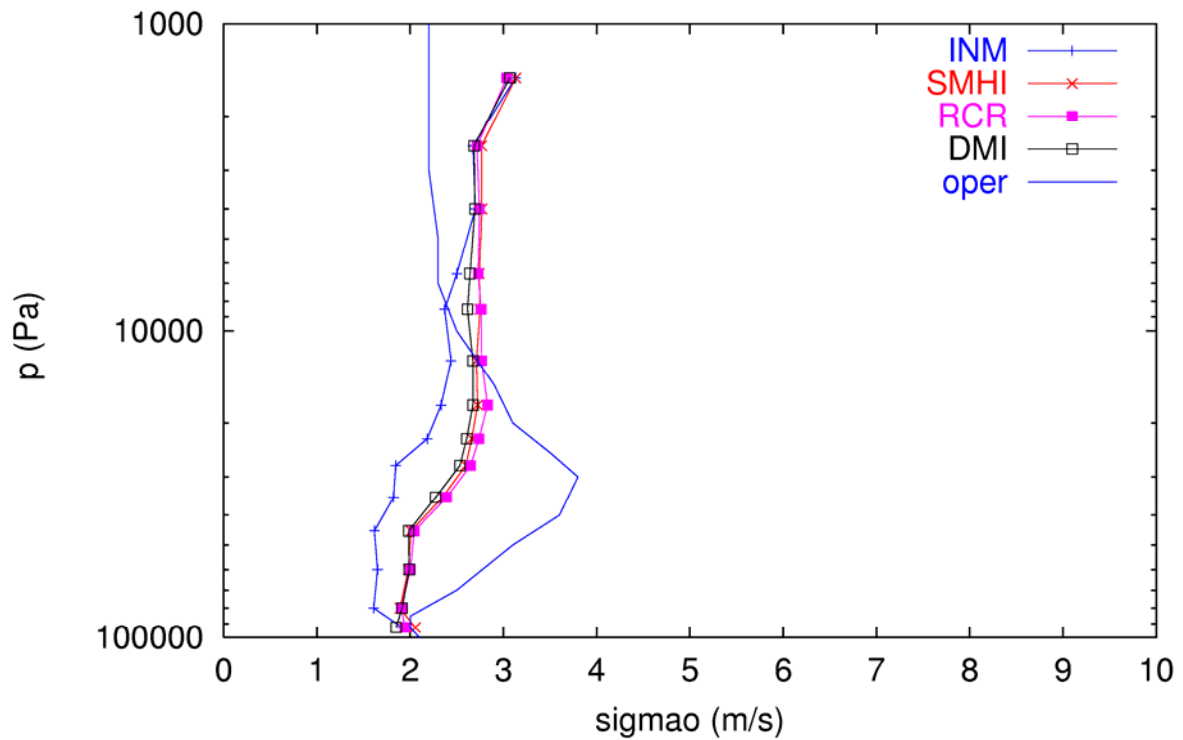


Figure 15: Estimated observation error standard deviation for wind u-component (m/s) from AIREP statistics at the different centers SMHI, DMI, FMI and INM and the assumed vertical profile of observation errors in the operational suites (solid line)

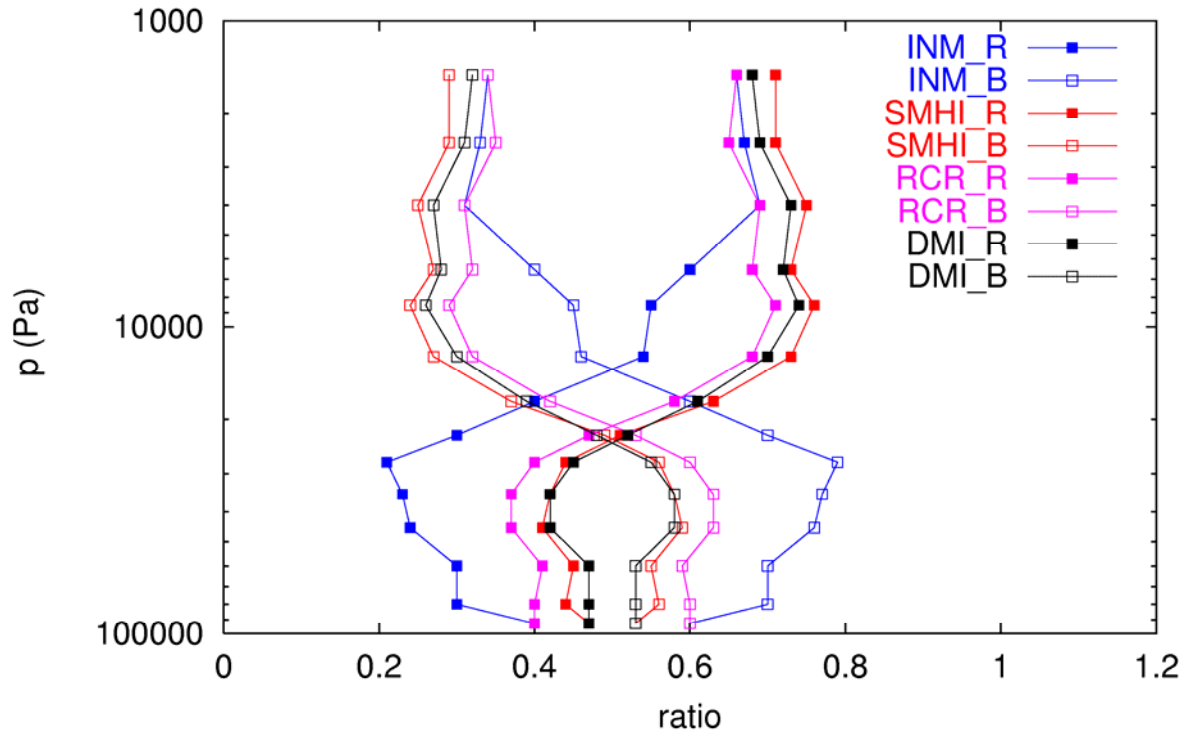


Figure 16: Observation to total error variance (filled squares) and background to total error variance (squares) ratios for wind u-component obtained from TEMP statistics at the different centers SMHI, DMI, FMI and INM.

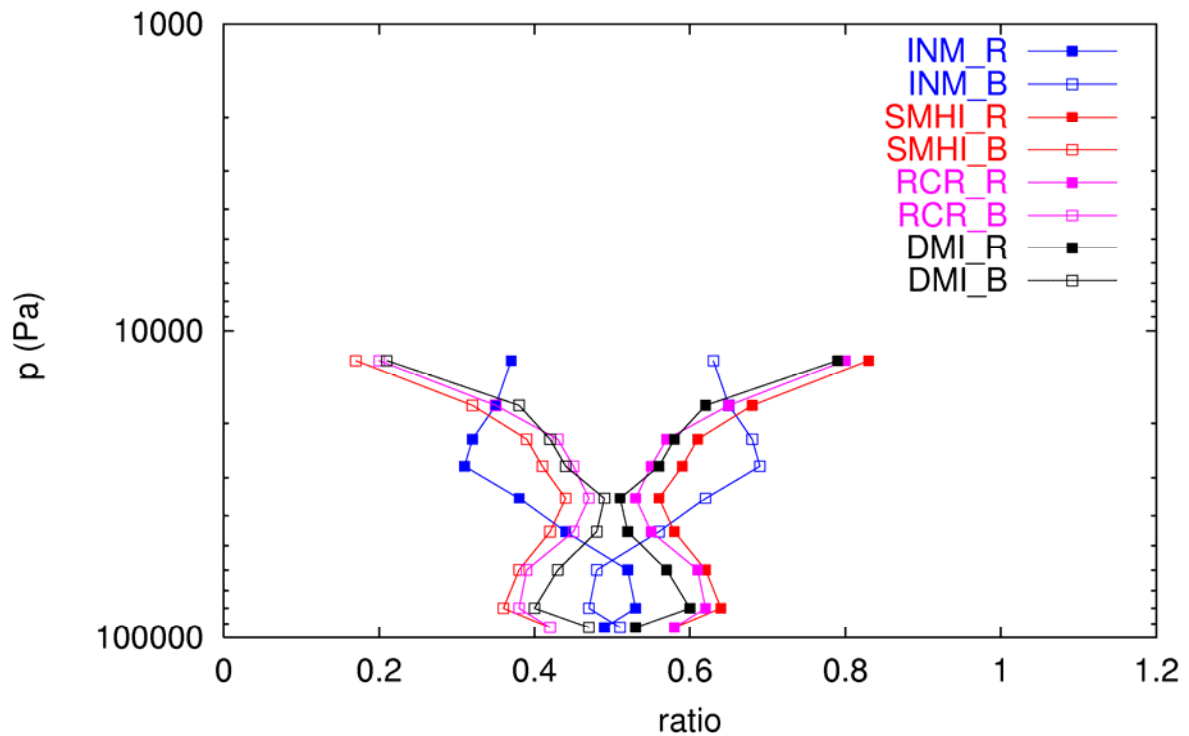


Figure 17: Observation to total error variance (filled squares) and background to total error variance (squares) ratios for wind u-component obtained from AIREP statistics at the different centers SMHI, DMI, FMI and INM.

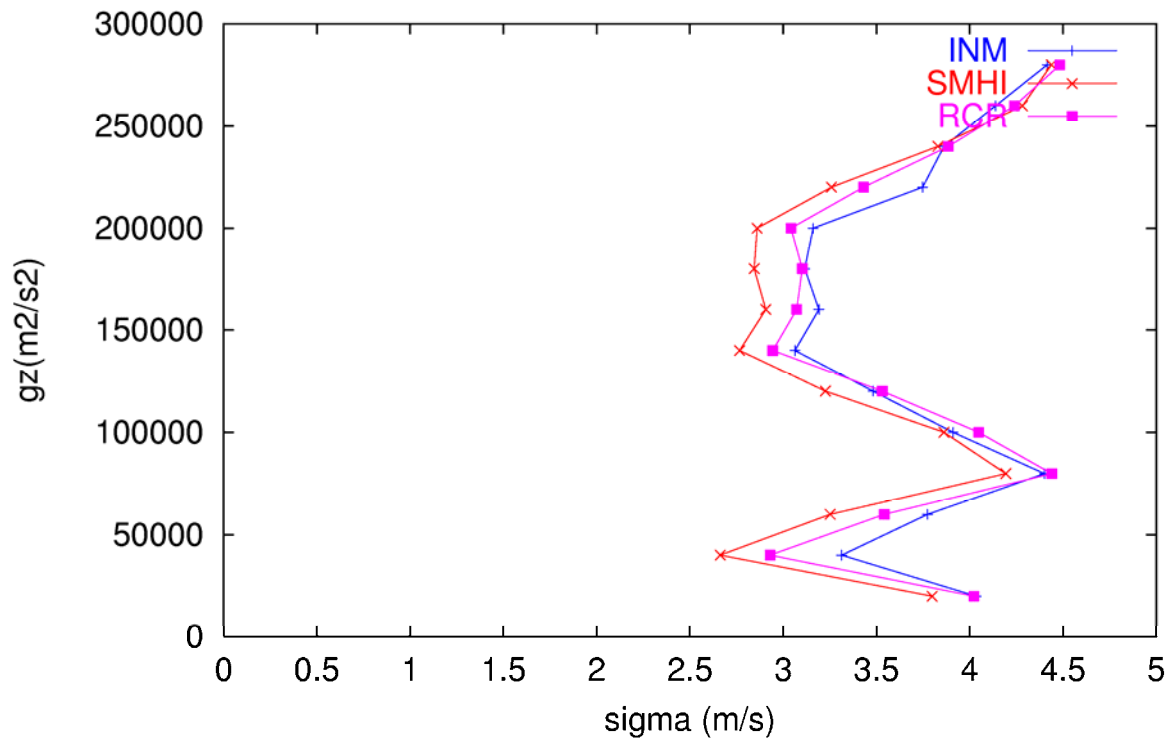


Figure 18: Square root of the sum of estimated observation and background wind u-component error variances (m/s) from PILOT statistics obtained at the different centers SMHI, FMI and INM.

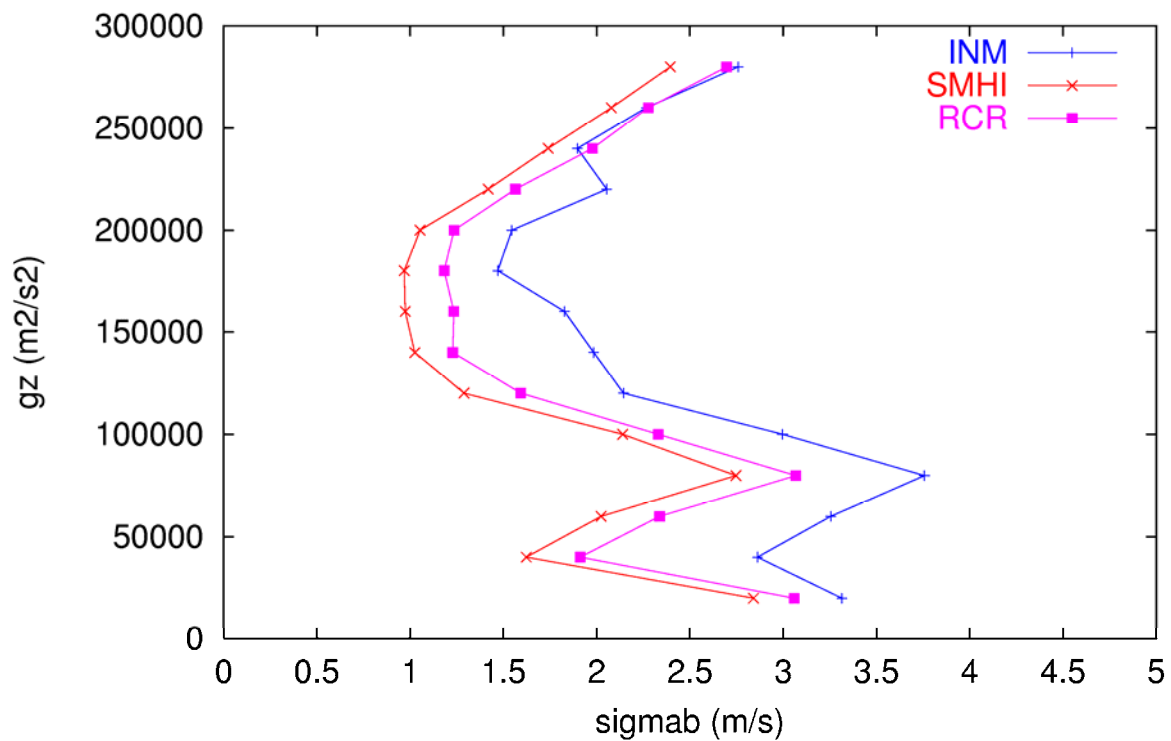


Figure 19: Estimated background error standard deviation for wind u-component (m/s) from PILOT statistics at the different centers SMHI, RCR and INM.

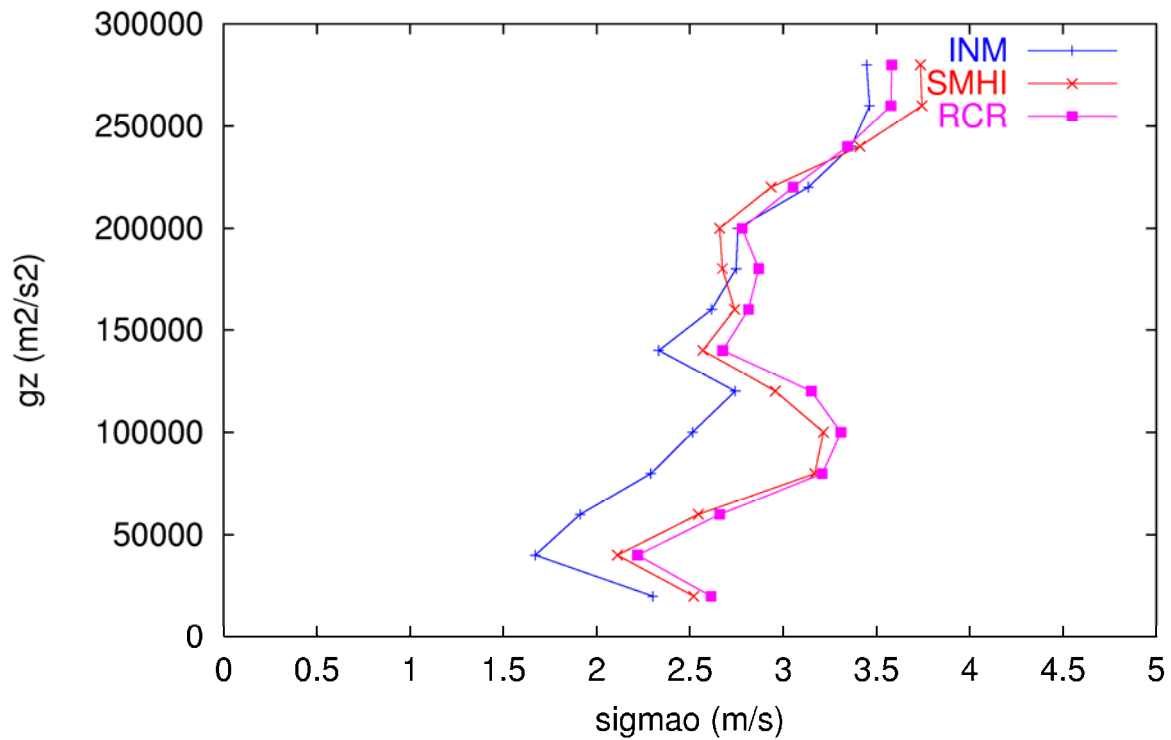


Figure 20: Estimated observation error standard deviation for wind u-component (m/s) from PILOT statistics at the different centers SMHI, RCR and INM.

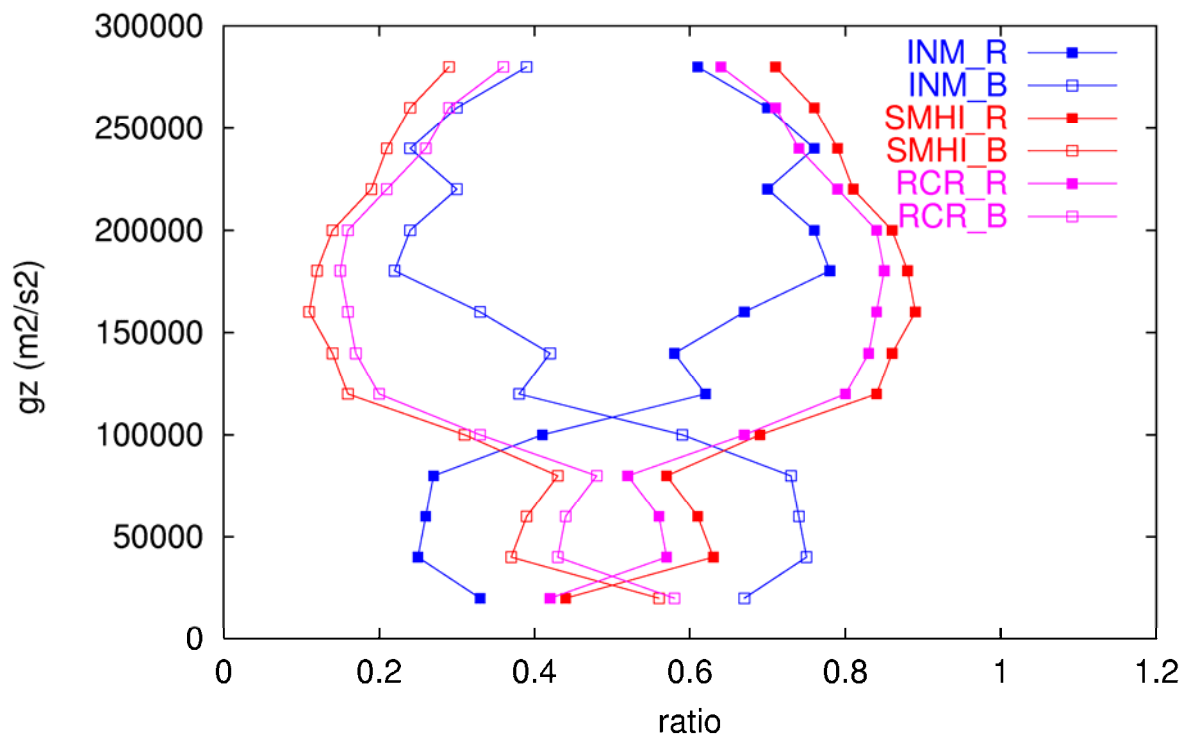


Figure 21: Observation to total error variance (filled squares) and background to total error variance (squares) ratios for wind u-component obtained from PILOT statistics at the different centers SMHI, FMI and INM.

## SEASONAL VARIATION OBTAINED FROM RCR STATISTICS

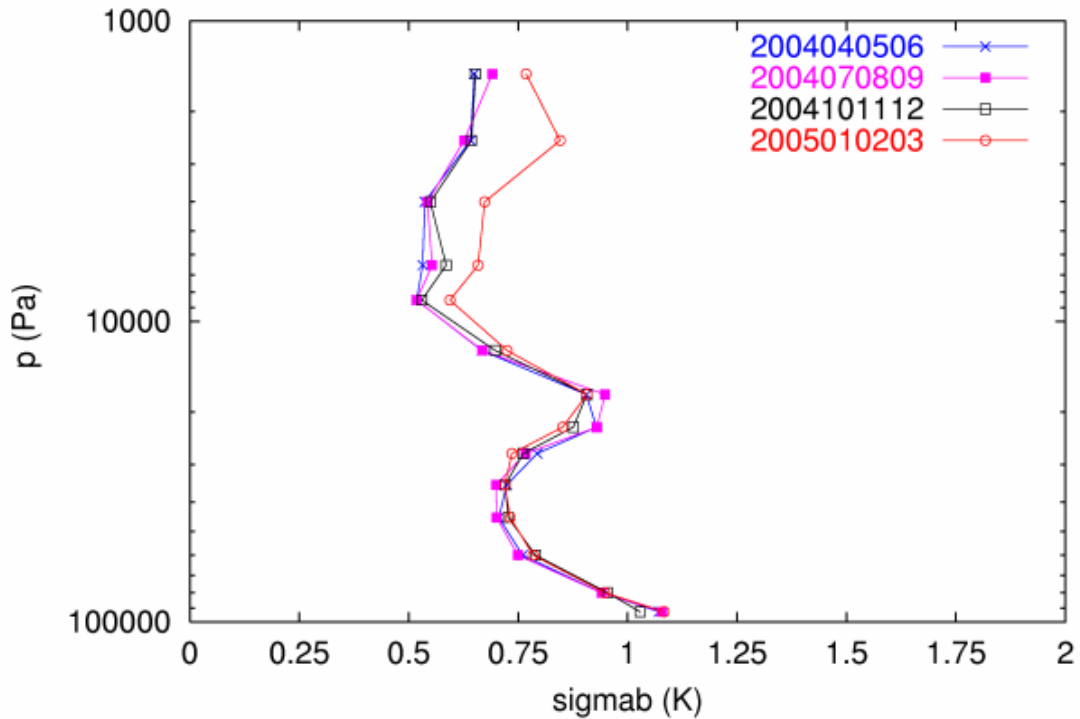


Figure 22: Temperature background error standard deviation (K) estimated from RCR radiosonde statistics in different seasons: spring (crosses), summer (filled squares), autumn (squares) and winter (circles).

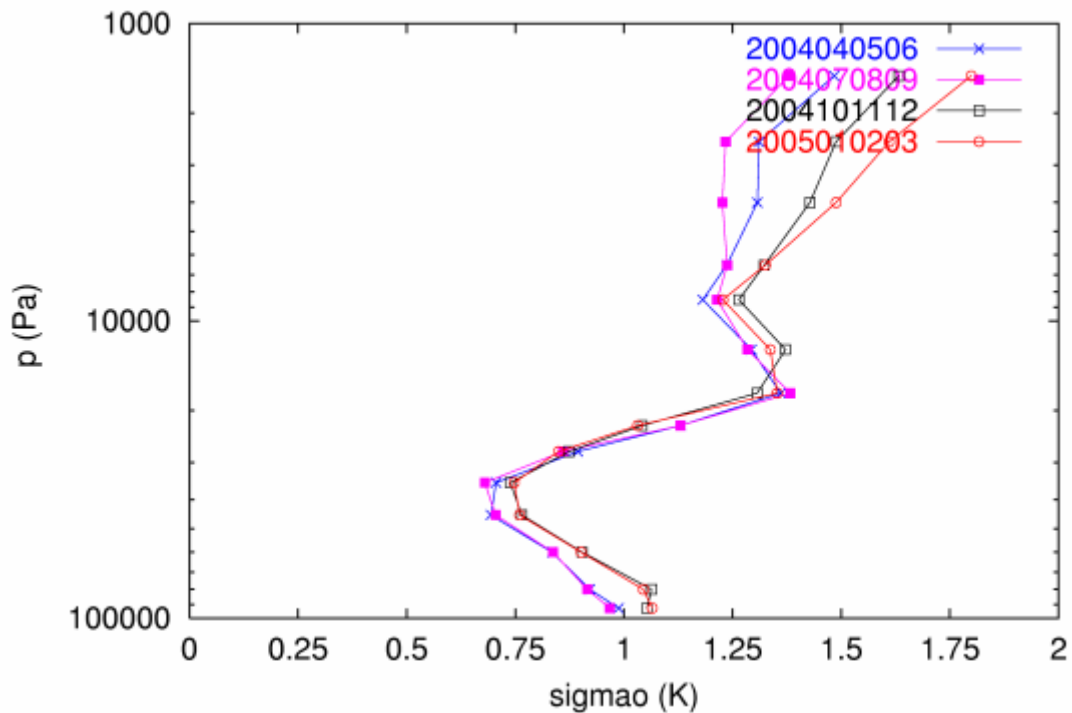


Figure 23: Temperature radiosonde observation error standard deviation (K) estimated from RCR run statistics in different seasons: spring (crosses), summer (filled squares), autumn (squares) and winter (circles).

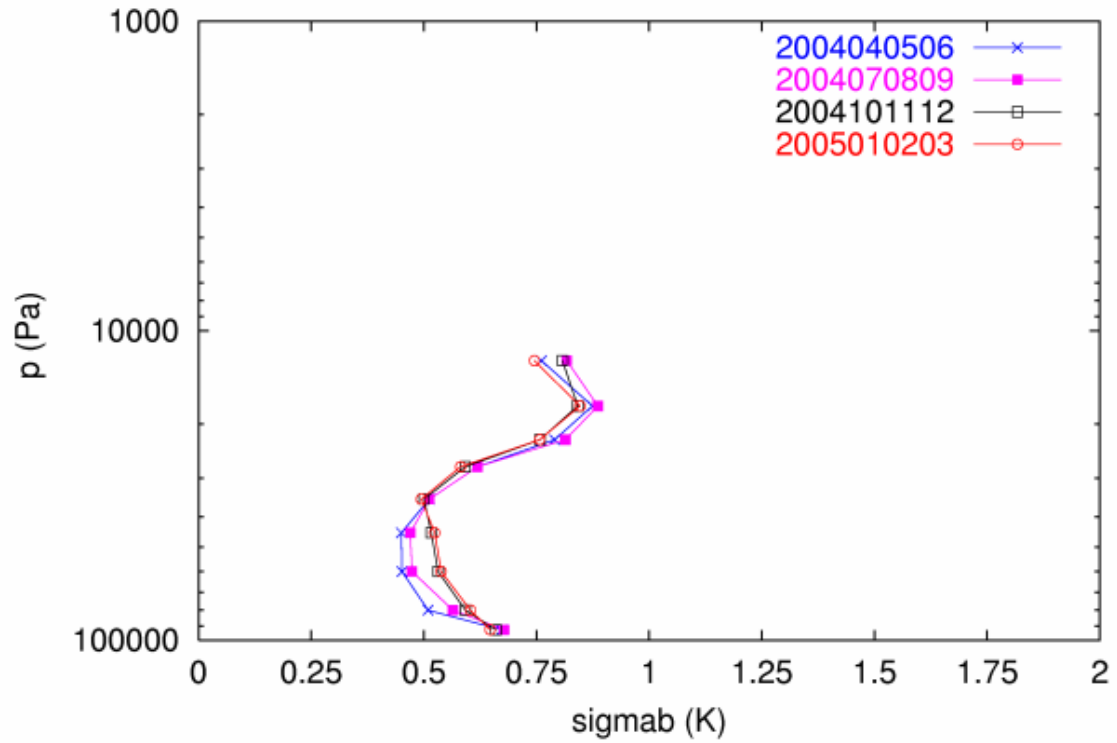


Figure 24: Temperature background error standard deviation (K) estimated from RCR AIREP statistics in different seasons: spring (crosses), summer (filled squares), autumn (squares) and winter (circles).

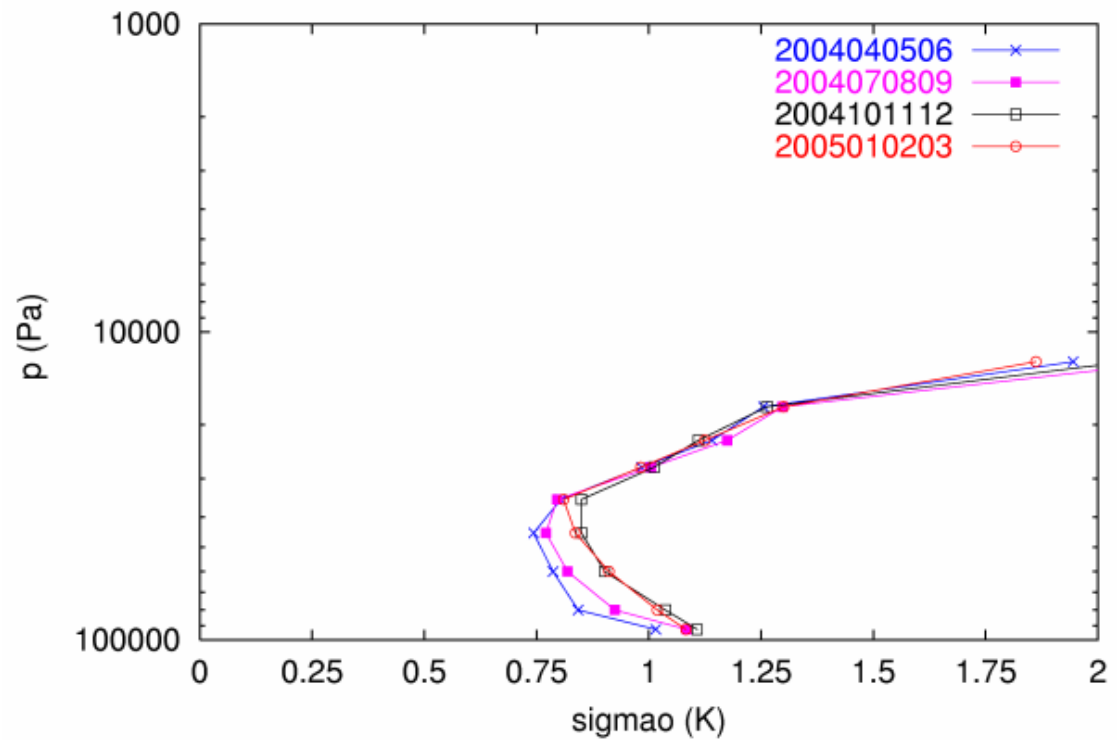


Figure 25: Temperature AIREP observation error standard deviation (K) estimated from RCR run statistics in different seasons: spring (crosses), summer (filled squares), autumn (squares) and winter (circles).

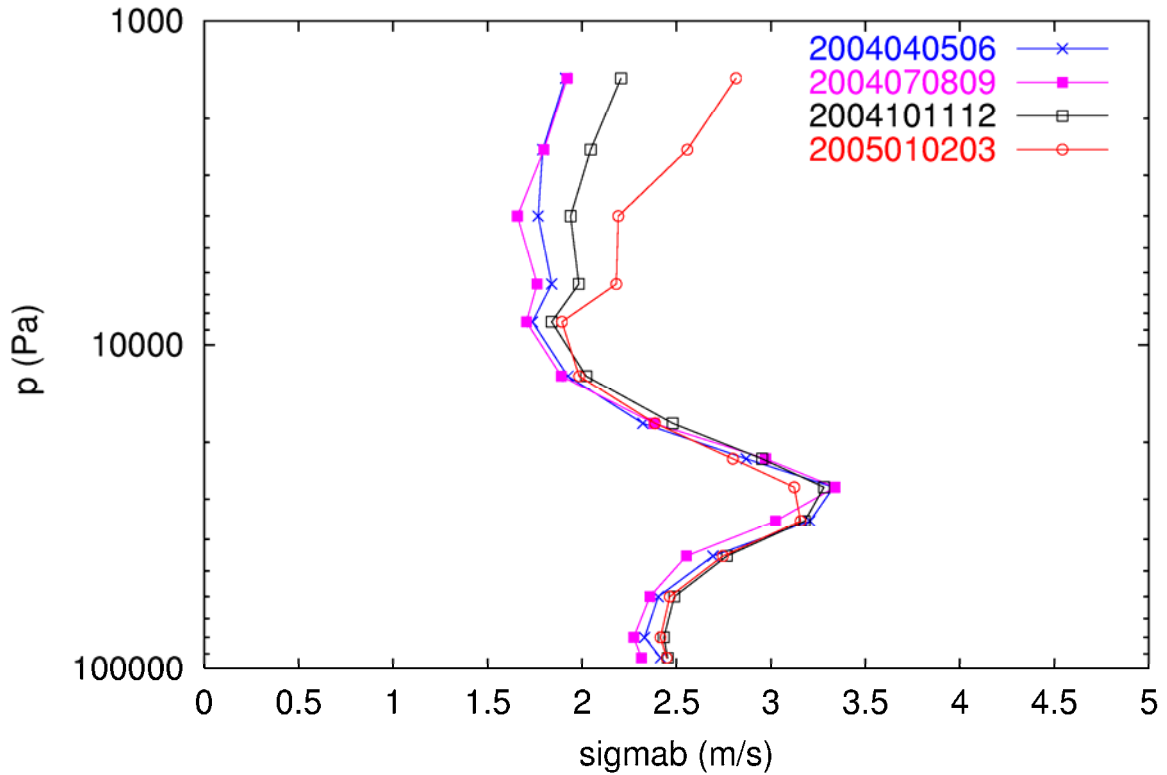


Figure 26: Wind u-component background error standard deviation (m/s) estimated from RCR radiosonde statistics in different seasons: spring (crosses), summer (filled squares), autumn (squares) and winter (circles).

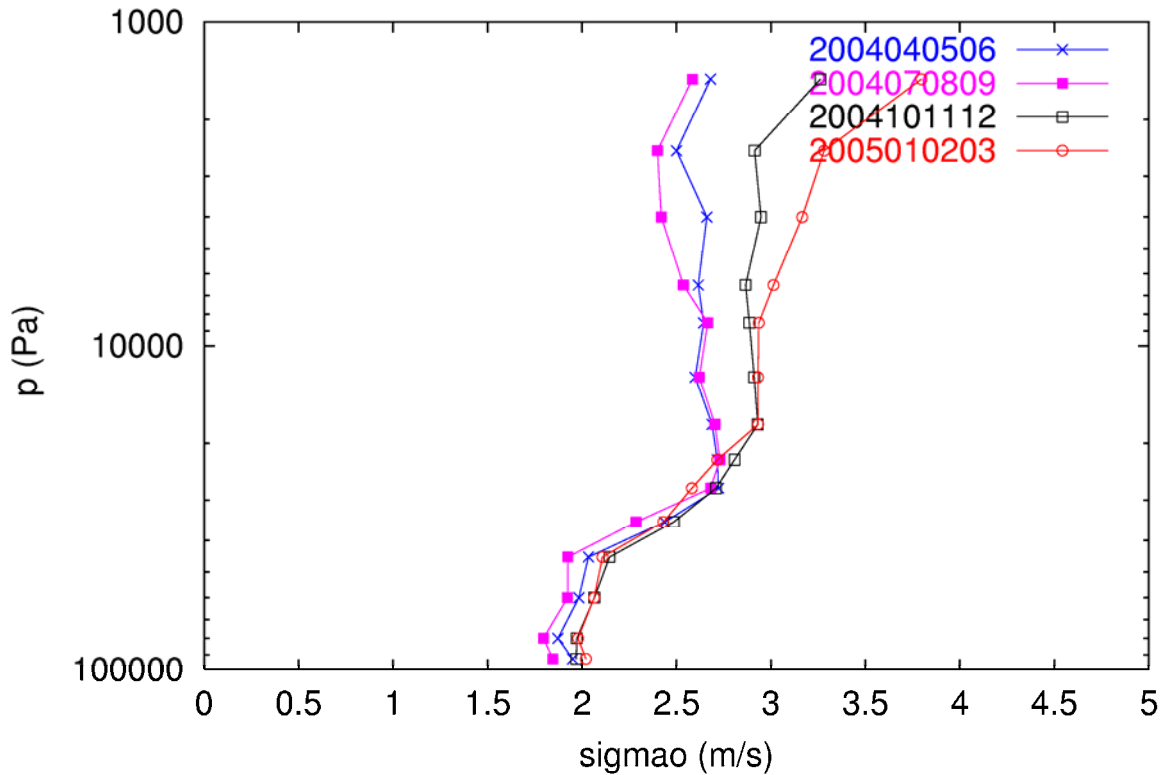


Figure 27: Wind u-component radiosonde observation error standard deviation(m/s) estimated from RCR statistics in different seasons: spring (crosses), summer (filled squares), autumn (squares) and winter (circles).

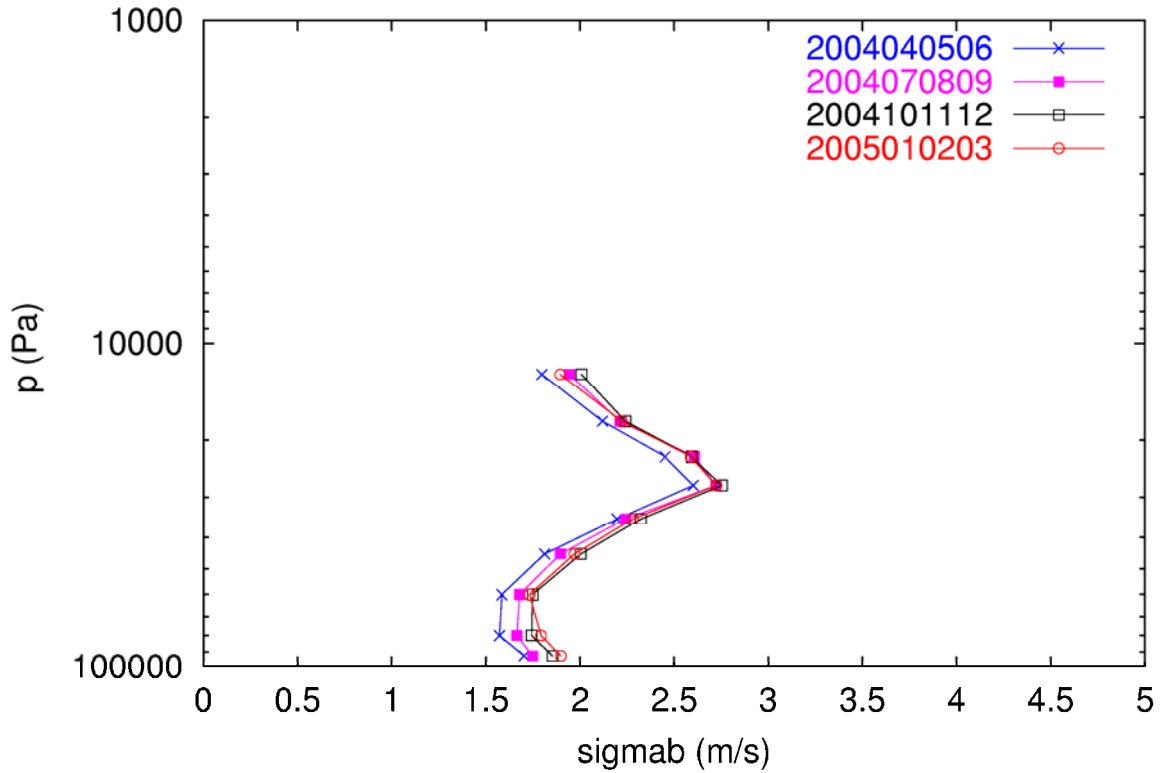


Figure 28: Wind  $u$ -component background error standard deviation (m/s) estimated from RCR AIREP statistics in different seasons: spring (crosses), summer (filled squares), autumn (squares) and winter (circles).

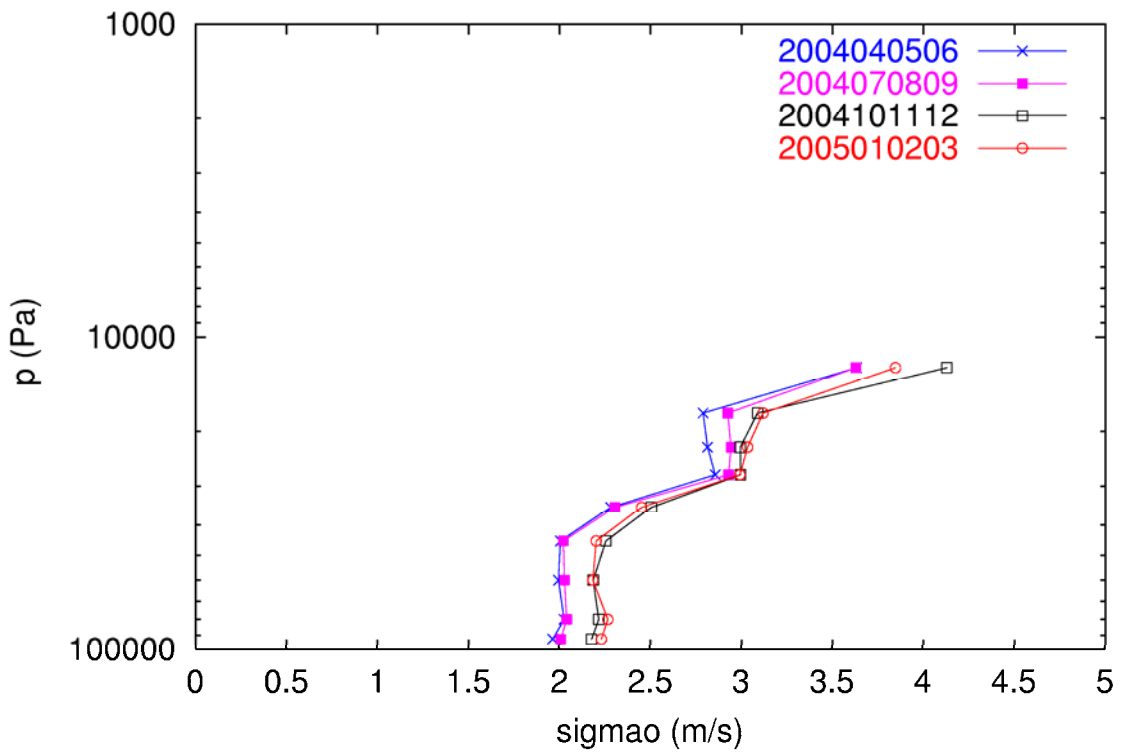


Figure 29: Wind  $u$ -component AIREP observation error standard deviation (m/s) estimated from RCR statistics in different seasons: spring (crosses), summer (filled squares), autumn (squares) and winter (circles).

### LATITUDINAL VARIATION OBSERVED IN INM STATISTICS

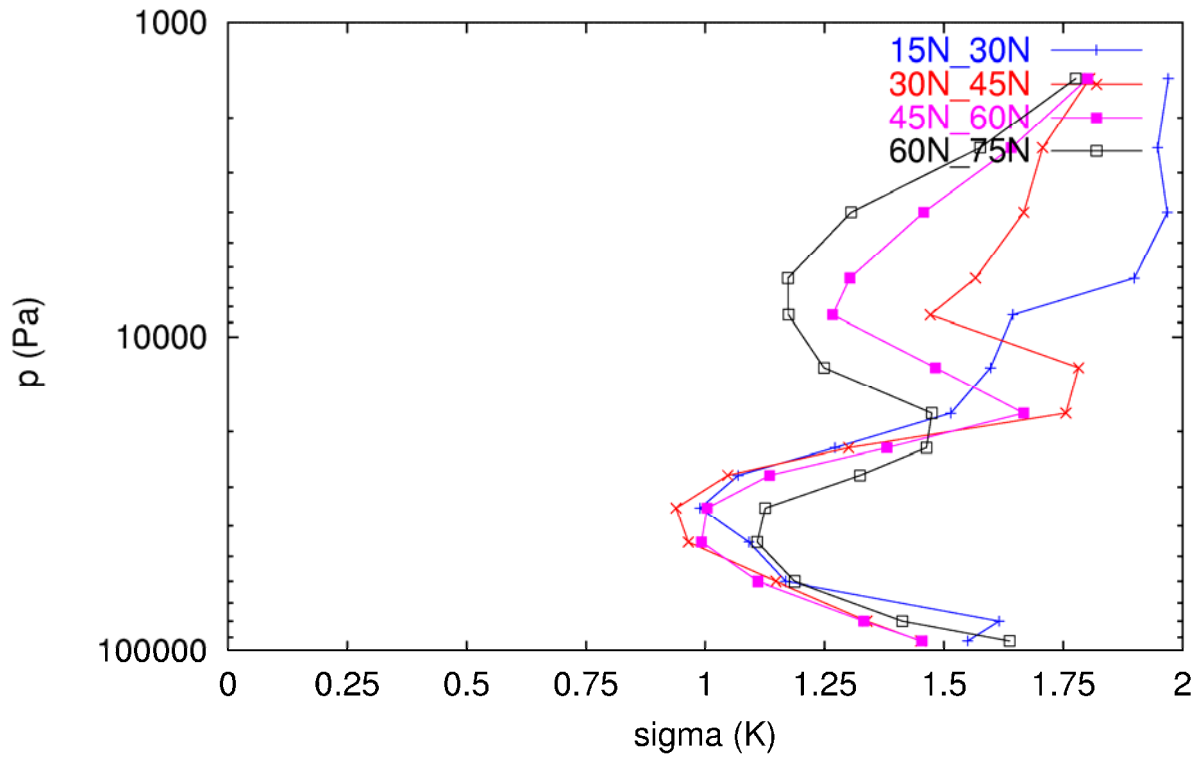


Figure 30: Square root of the sum of estimated observation and background temperature error variances (K) from TEMP statistics obtained in different latitudinal bands.

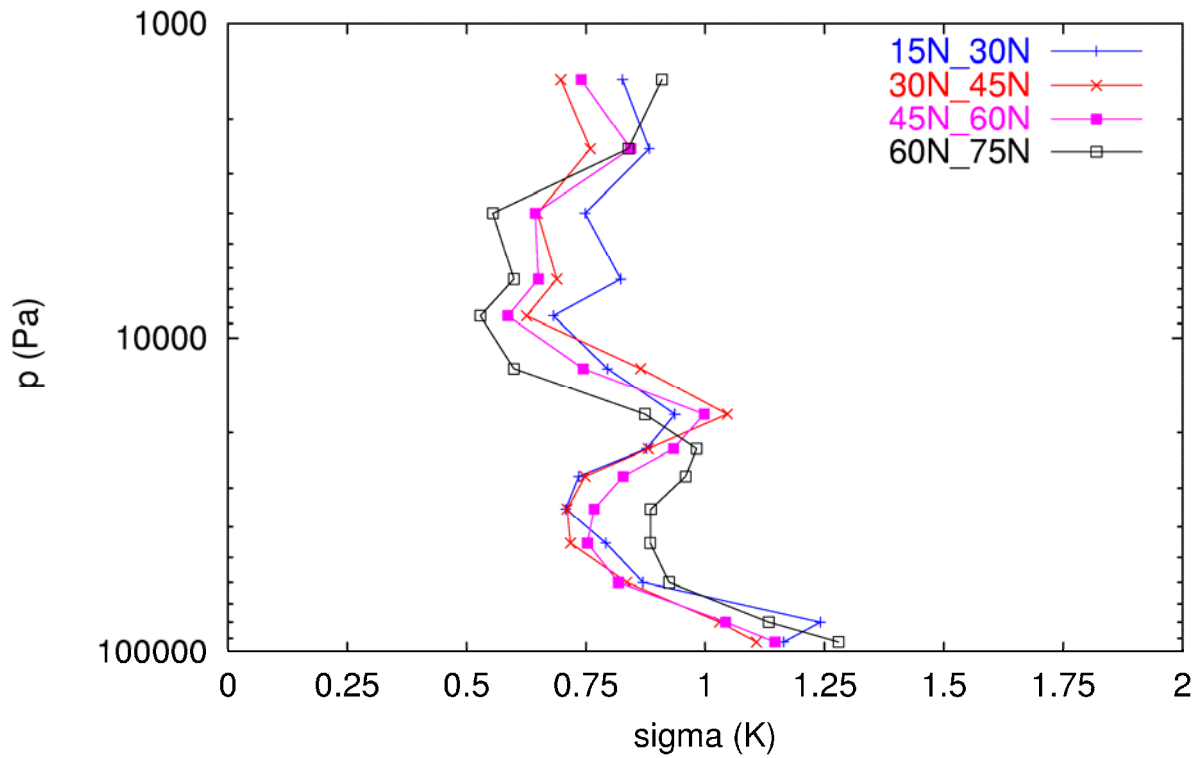


Figure 31: Estimated temperature background error standard deviation in different latitudinal bands.

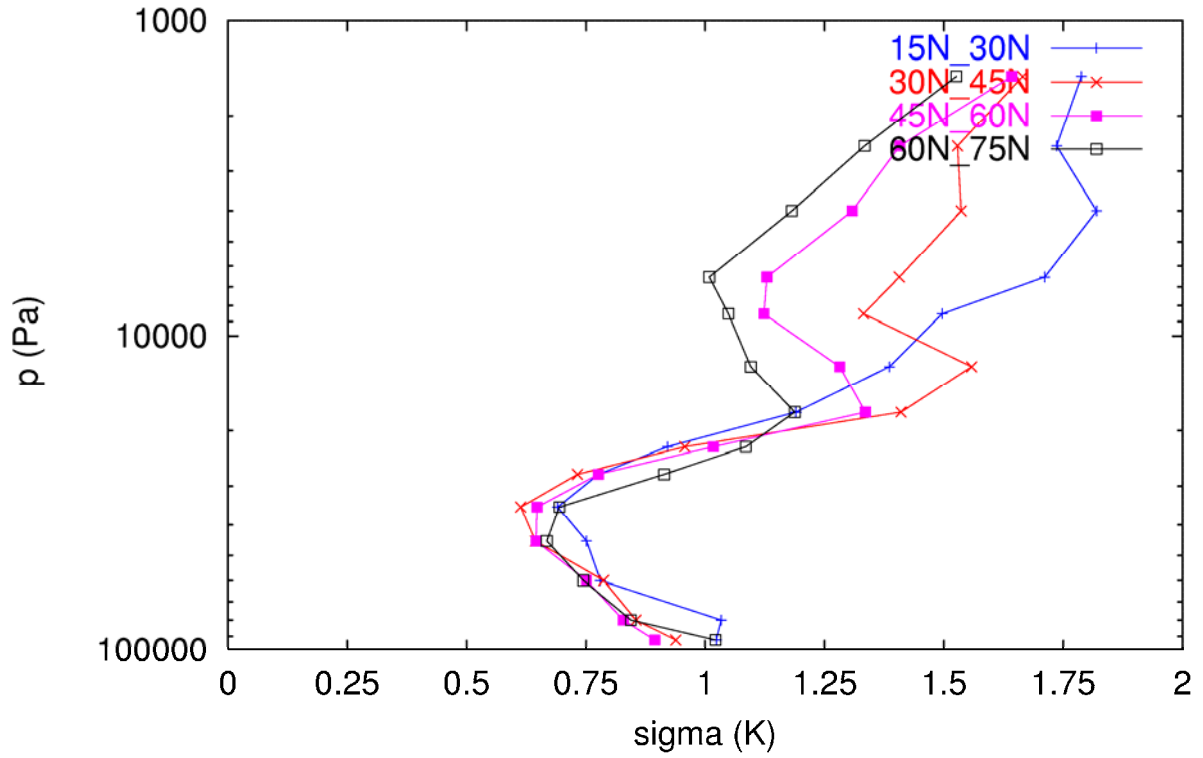


Figure 32: Estimated temperature observation error standard deviation in different latitudinal bands.

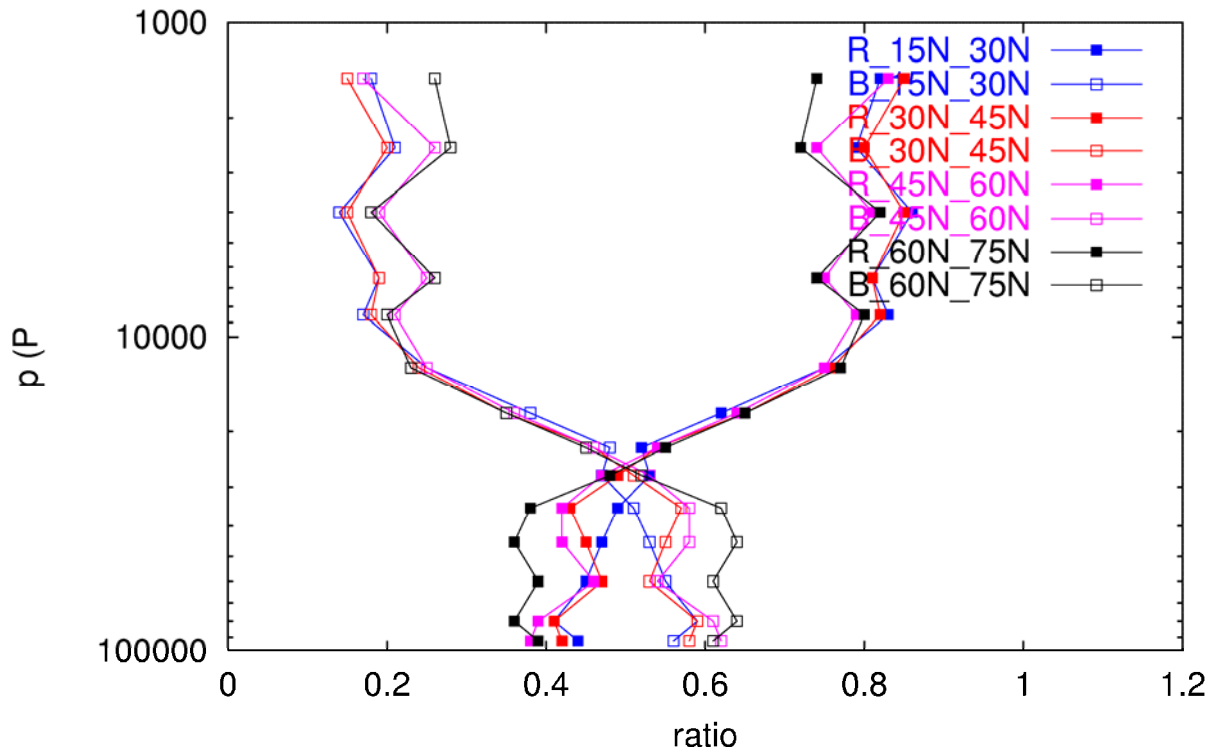


Figure 33: Ratios of temperature background (squares) and observation (filled squares) error variances to innovation variance in different latitudinal bands.

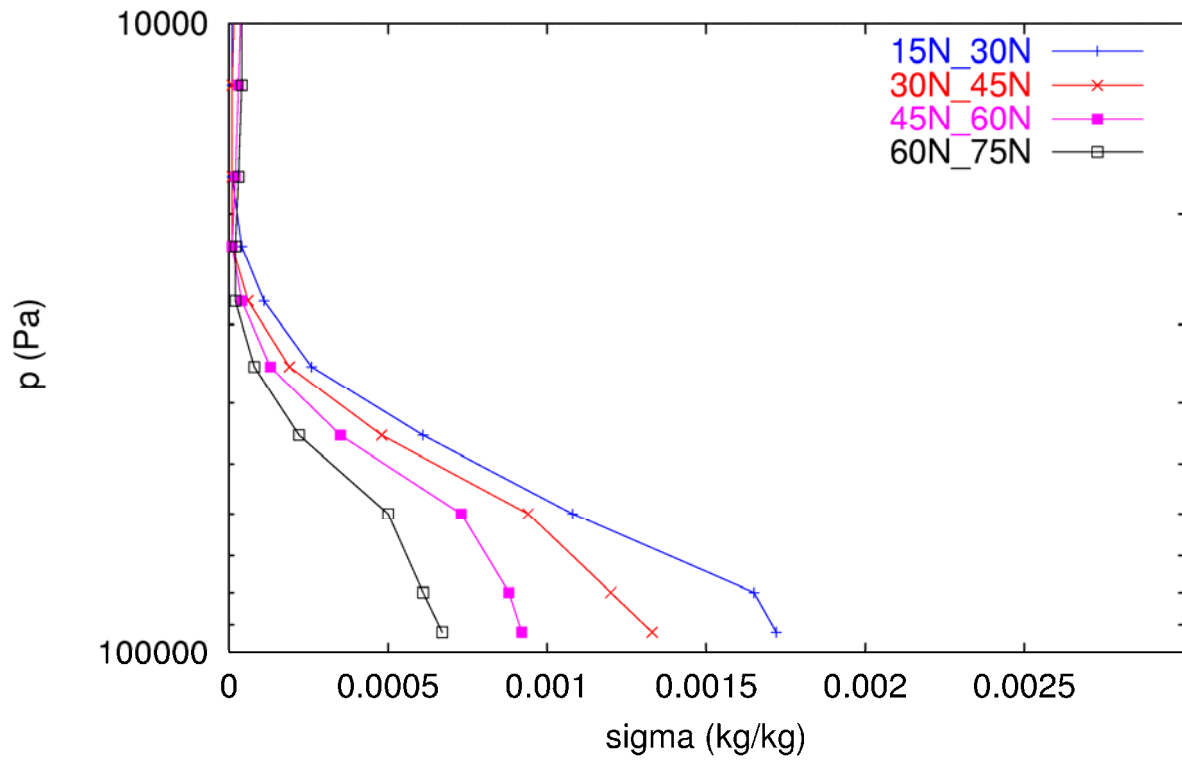


Figure 34: Square root of the sum of estimated observation and background specific humidity error variances (kg/kg) from TEMP statistics obtained in different latitudinal bands.

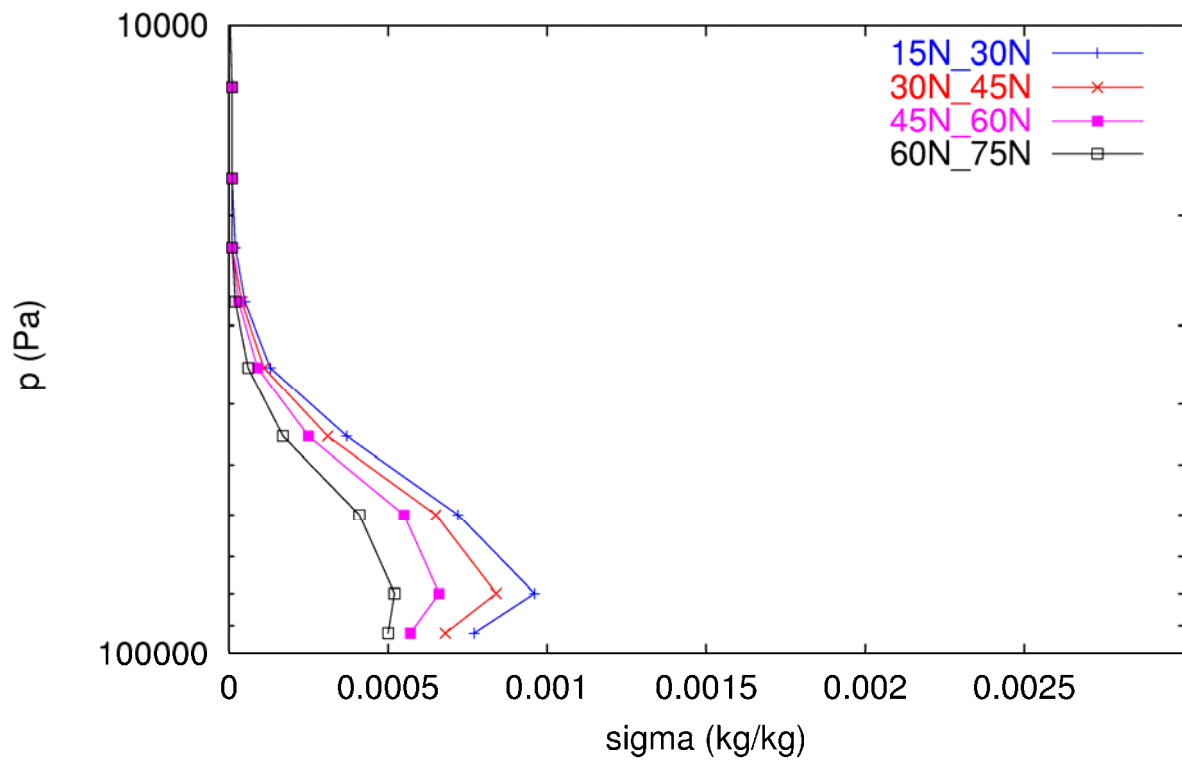


Figure 35: Estimated specific humidity background error standard deviation (kg/kg) from TEMP statistics obtained in different latitudinal bands.

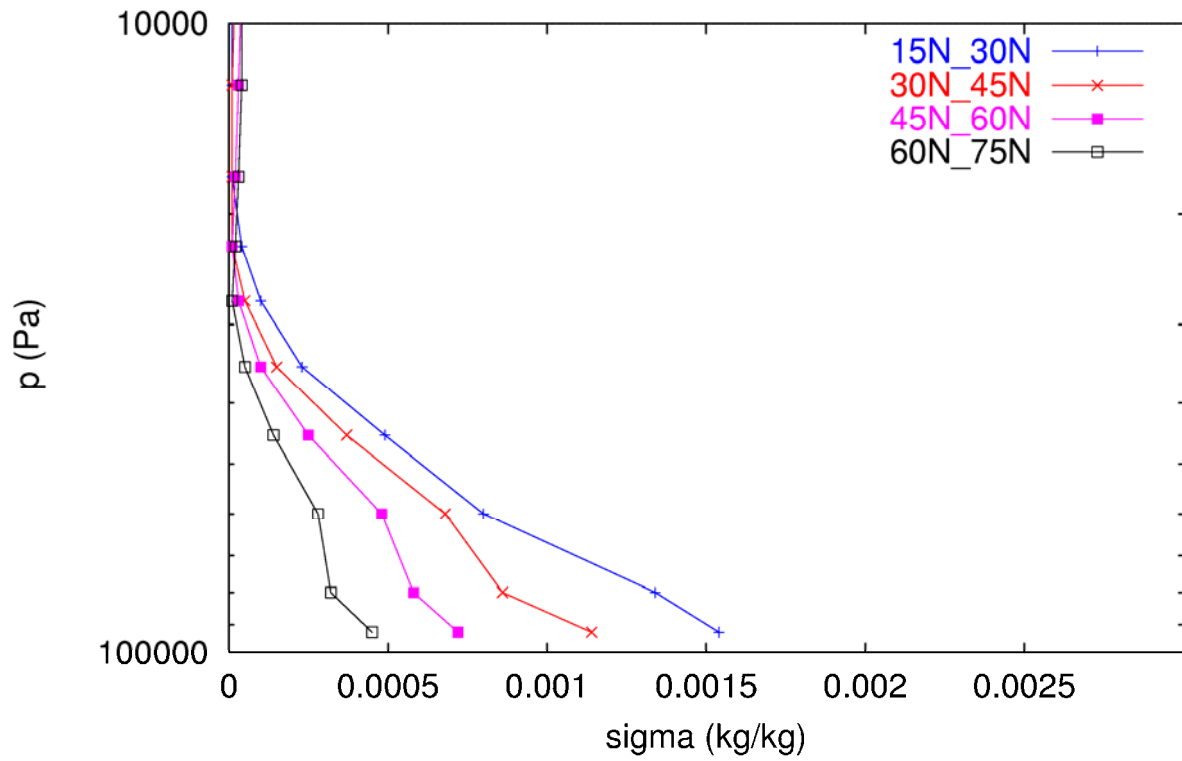


Figure 36: Estimated specific humidity observation error standard deviation (kg/kg) from TEMP statistics obtained in different latitudinal bands.

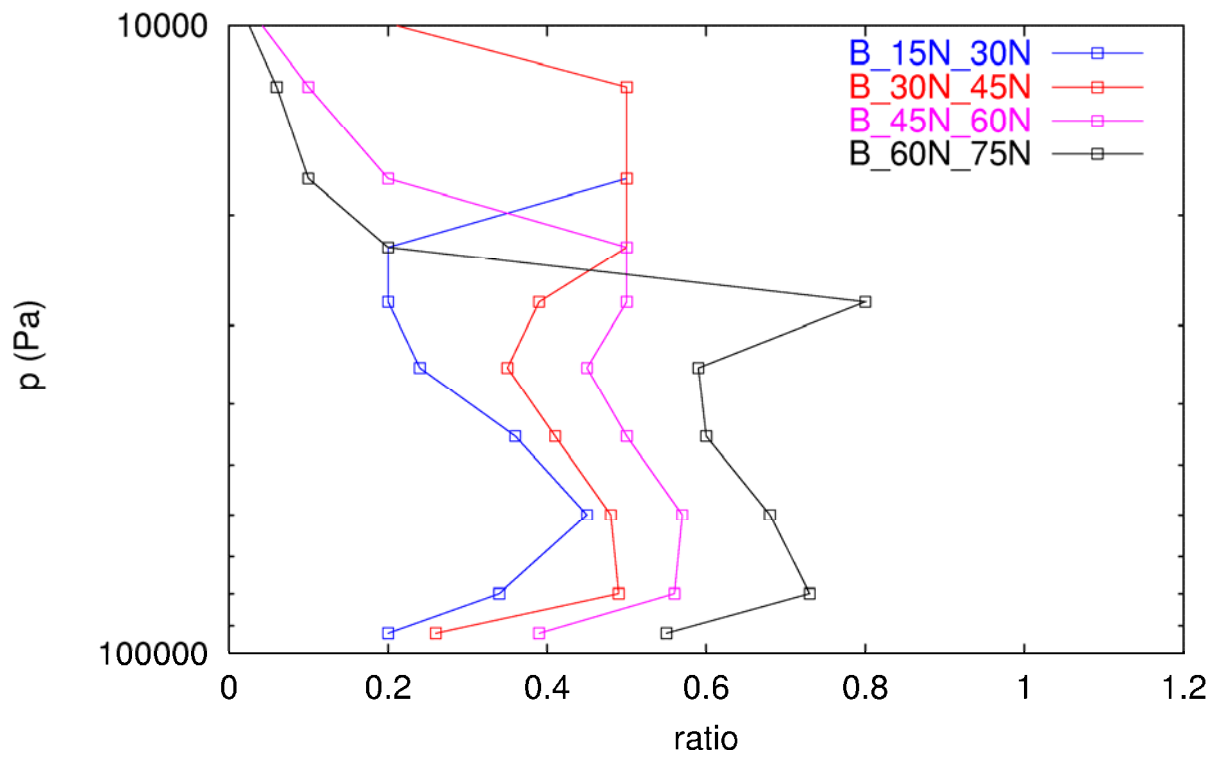


Figure 37: Ratio of specific humidity background error variance to innovation variance in different latitudinal bands.

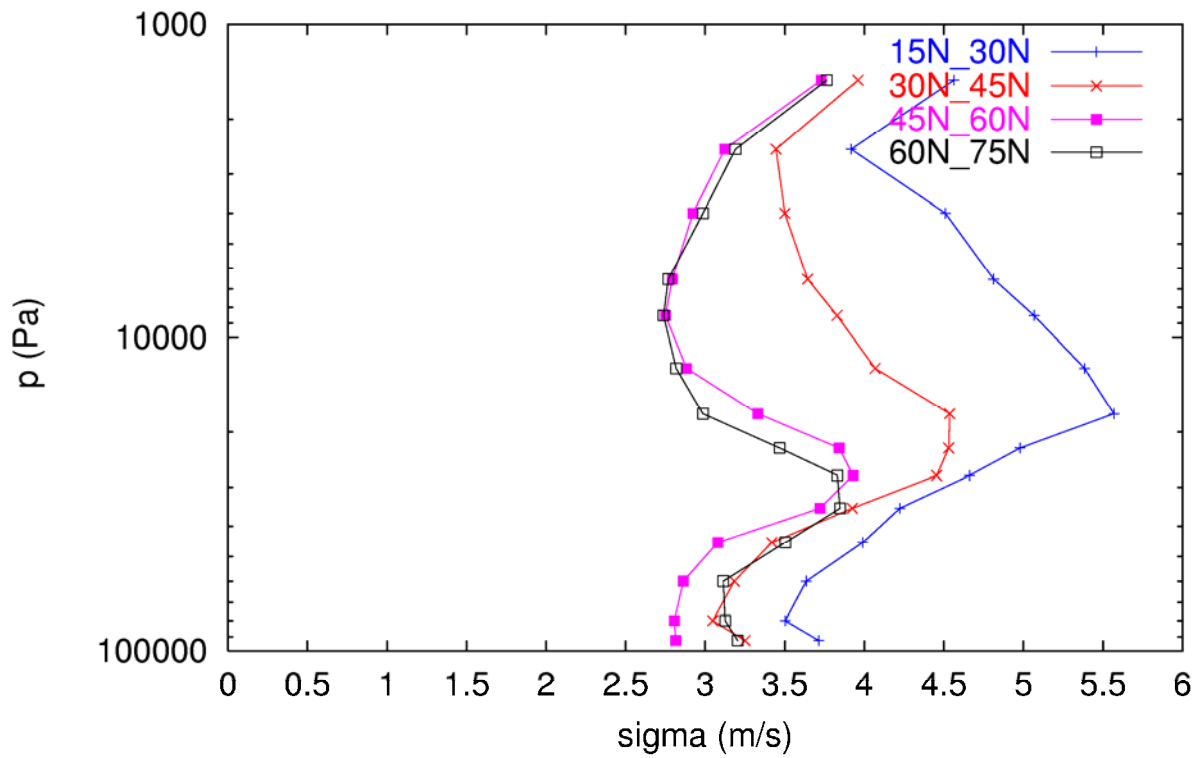


Figure 38: Square root of the sum of estimated observation and background wind u-component error variances (m/s) from TEMP statistics obtained in different latitudinal bands.

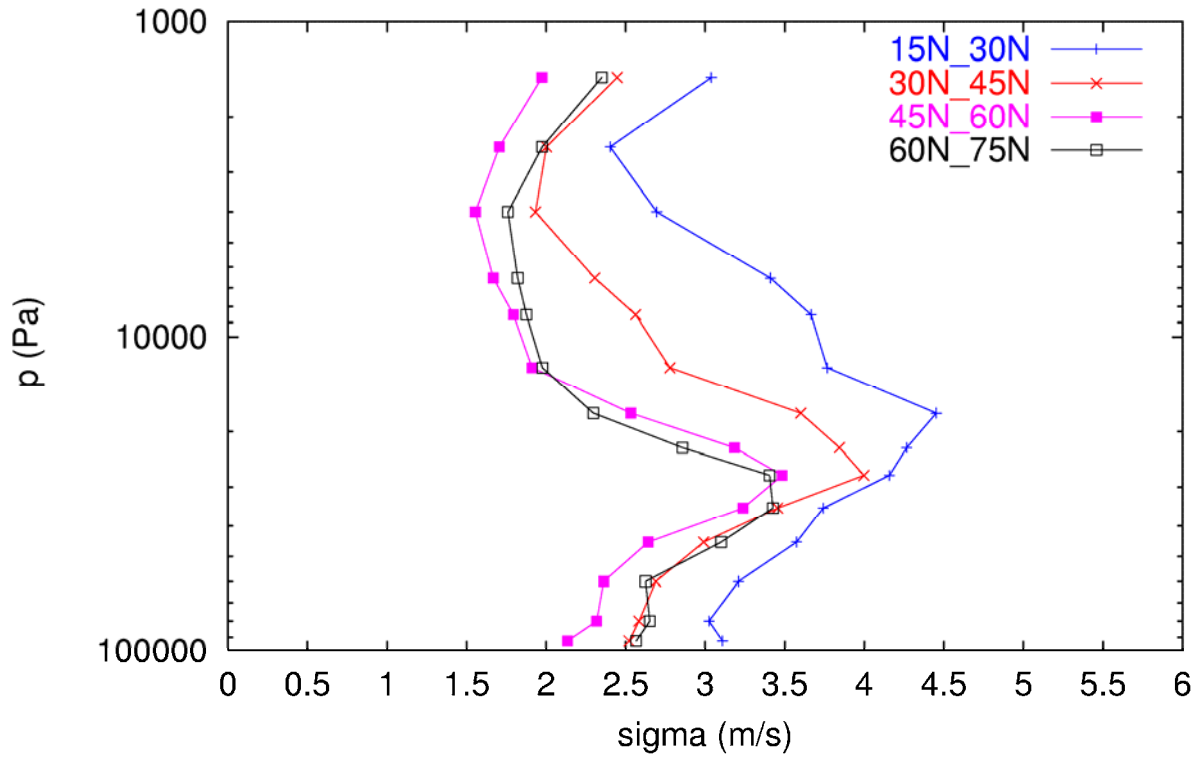


Figure 39: Estimated wind u-component background error standard deviation (m/s) from TEMP statistics in different latitudinal bands.

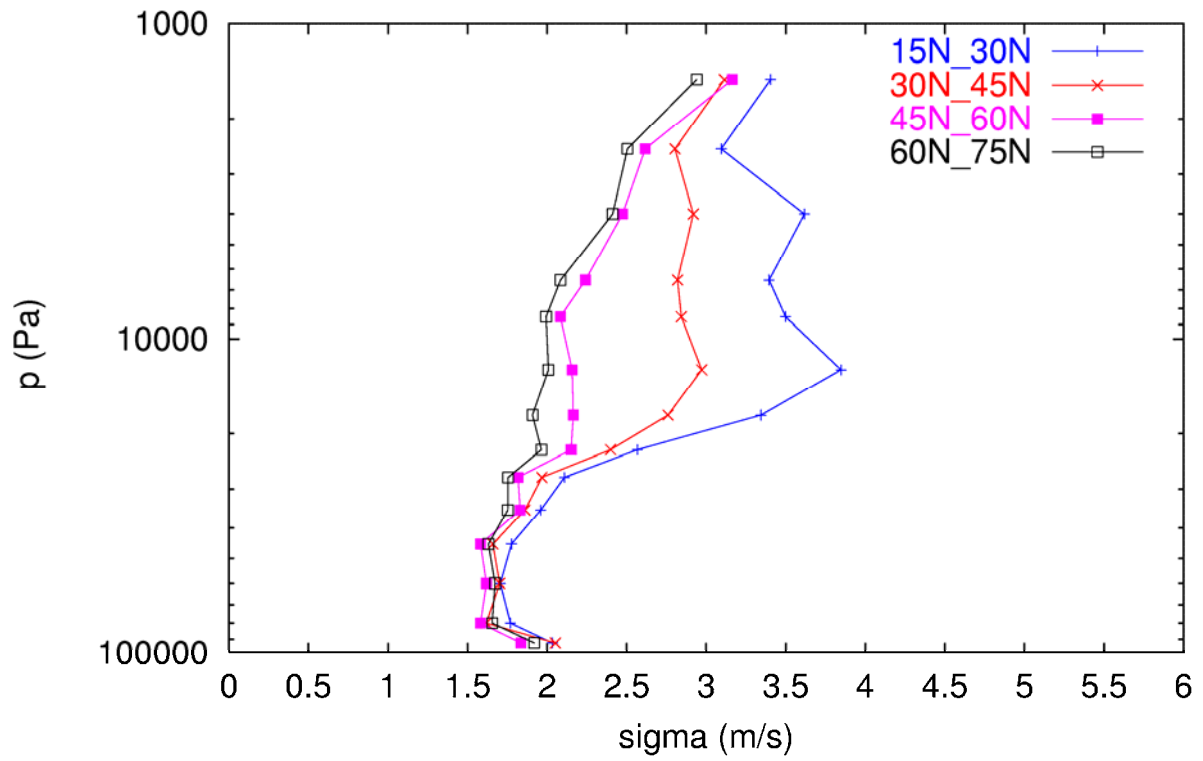


Figure 40: Estimated wind u-component TEMP observation error standard deviation (m/s) in different latitudinal bands.

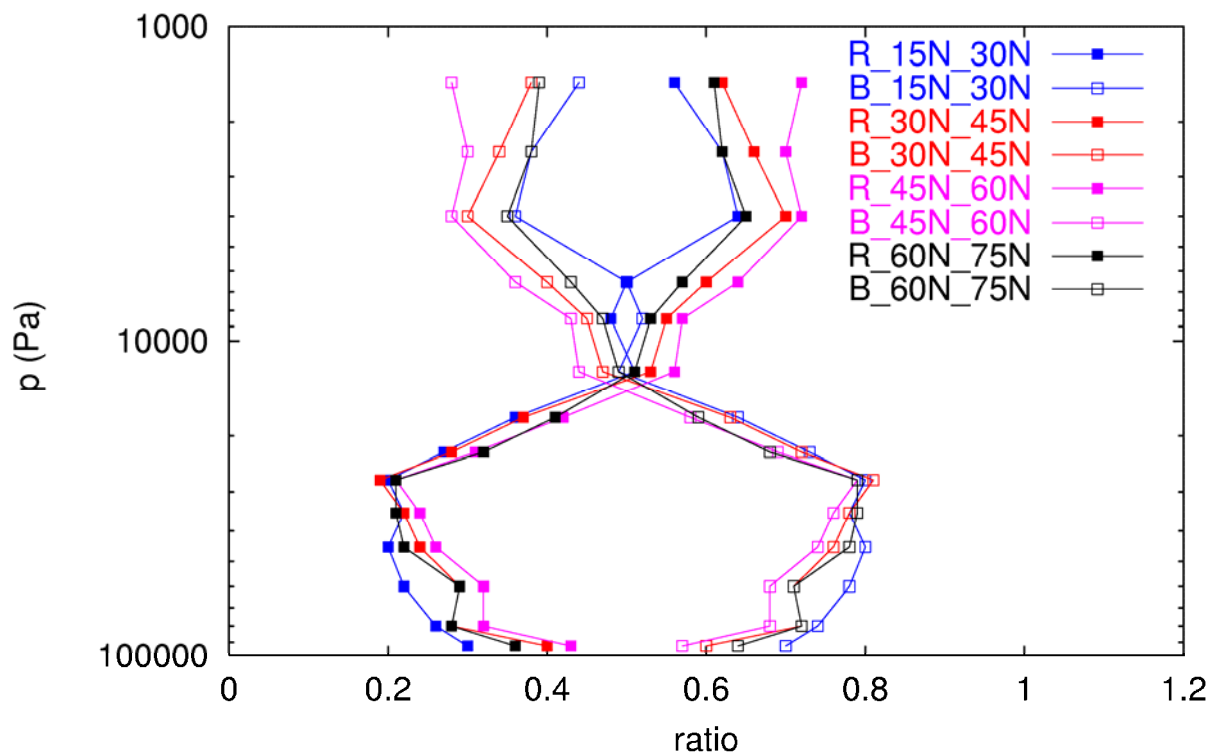


Figure 41: Ratios of wind u-component background (squares) and observation (filled squares) error variances to innovation variance in different latitudinal bands.