Experiences from pre-operational HIRLAM 5.1.4 tests at FMI

By SIMO JÄRVENOJA, Finnish Meteorological Institute

12th June 2003

1 Introduction

The operational HIRLAM system at FMI was under a major upgrade in 2002 and in early 2003. First, migrating the operational system from the Cray T3E to the IBM eServer Cluster 1600 required a lot of resources (Eerola, 2003). Second, pre-operational tests of a new generation ISBA-based HIRLAM 5.1.4 system started. This report shortly describes the pre-operational suites and then summarizes some experiences from pre-operational test runs.

2 Renewing the operational HIRLAM

The operational HIRLAM at FMI has been based on the 4.6.2 system since November 1999, with some updates from later versions. As the Cray T3E at CSC was to be faced out in late 2002, migration to the new IBM eServer Cluster 1600 was necessary. The original aim was to upgrade the operational HIRLAM system at the same time with the computer change. Therefore, preparations for introduction of the HIRLAM 5.1.4 system were carried out in early 2002, and in summer 2002 pre-operational parallel tests on IBM were started. The HIRLAM 5.1.4 system includes several important new modules compared to the old operational 4.6.2 system, such as 3D-VAR analysis system, digital filter initialization, semi-Lagrangian advection and ISBA surface parameterization together with the related surface analysis scheme.

The HIRLAM system at FMI has consisted of two suites: the ATA suite with a 0.4° horizontal resolution and the nested ENO suite with a 0.2° resolution, both with 31 levels in the vertical, see e.g., Eerola (2002). The aim was to replace these two suites with one single 0.18°/40-level system. This suite was, however, unstable and suffered from frequent crashes, as described in Järvenoja (2003). Furthermore, it was found that temperature and wind fields were much noisier than those of the operational ATA suite. The reason for the noise was studied to a great extent, but the origin could not be found. The length scale of the noise was clearly larger than two grid lengths. Due to these problems the idea of one single new suite with a high resolution was abandoned and two different pre-operational suites were introduced: one with a 0.4° horizontal resolution and 31 levels, called ATB, and another with a 0.3° horizontal resolution and 40 levels in the vertical, called ATC. The same resolution used in ATA and ATB suites enables direct comparison of results, and the lower resolution (0.3°) makes the ATC suite more stable than the 0.18° system that suffered from crashes.

The three suites, the operational ATA, and pre-operational ATB and ATC are briefly described in the following:

ATA (Hirlam 4.6.2) :

- 0.4° horizontal resolution, 194 × 140 grid points
- 31 levels in the vertical
- non-linear normal mode initialization
- Eulerian, time step 3 min
- OI analysis
- simple surface parameterization

ATB (Hirlam 5.1.4) :

- 0.4° horizontal resolution, 194 × 140 grid points
- 31 levels in the vertical
- digital filter initialization
- Eulerian, time step 3 min
- 3D-VAR analysis and new surface analysis
- ISBA surface parameterization

ATC (Hirlam 5.1.4) :

- 0.3° horizontal resolution, 256 × 186 grid points
- 40 levels in the vertical
- digital filter initialization
- Semi-Lagrangian, time step 7.5 min
- 3D-VAR analysis and new surface analysis
- ISBA surface parameterization

The ATA and ATB suites use the same grid, and the area of ATC is just inside the ATA area so that the same ECMWF frames can be used as boundary conditions in all suites. The ATB suite differs from ATA only in the analysis method, initialization scheme and surface parameterization. The ATC suite differs from ATB in horizontal and vertical resolution, and advection scheme.

The pre-operational ATB suite was started in July 2002, and the ATC suite in September 2002, and both were continued until March 2003, when the ATC suite, with some further upgrades (Eerola, 2003), became operational with production name ATX.

3 Results from pre-operational tests

In the following, some results from the pre-operational tests will be presented for a selection of months. The emphasis in the evaluation has been the 2-metre temperature (T_{2m}) and mean-sea-level pressure (p_{msl}).

3.1 August

August 2002 was a very warm and dry month in northern Europe, dominated by high pressure weather, and the daily maximum temperature exceeded 25°C on many days, for example in Finland.

In the following, the 36 and 48 h biases of T_{2m} from field verification will be shown. Forecasts starting from 00 UTC analyses only are used so that 36 h forecasts are valid at daytime and those of 48 h at night-time. Figures 1 and 2 show the T_{2m} bias in 36 h ATA and ATB forecasts, respectively. In both cases the bias is computed against the T_{2m} analysis from the ATB suite. The ATA forecasts, based on the simple surface scheme (Fig. 1), show a cold bias of a few degrees at northern latitudes: North America, Greenland and northern Russia. On the other hand, there is a positive bias over continental Europe. The bias in the ISBA-based ATB forecasts (Fig. 2) is small, only of the order of ± 1°C and in a few limited areas only.

Figures 3 and 4 show the T_{2m} bias in 48 h ATA and ATB forecasts, respectively. There is a cold bias in the ATA forecasts (Fig. 3) at high latitudes also at night-time, while a positive bias of a few degrees dominates over most of Europe. The ATB forecasts (Fig. 4) show a slight positive bias e.g., over Greenland, in Scandinavia and Baltic countries, while a small negative bias can bee seen in southern Europe.

Figures 1-4 demonstrate that the ISBA-based ATB forecasts have a considerably smaller bias compared to ATA. The experience from operational ATA forecasts has shown that too warm night-time temperatures are predicted in summer. The introduction of the ISBA surface scheme now leads to a more accurate prediction of the diurnal temperature cycle.

3.2 Autumn

The autumn weather in northern Europe in 2002 was also dominated by high pressure. The active storms, low pressure systems, stayed far away over the Atlantic so that autumn 2002 did not provide a real test of typical, rapidly changing autumn weather for pre-operational suites. The quality of the T_{2m} forecasts in September and October was of the same order as in August: ISBA-based ATB and ATC suites were superior to the operational ATA suite (not shown).

The weather was variable with a few storms in November. Unfortunately, the IBM system was rather unstable with a lot breaks so that many forecasts with the pre-operational systems were lost. Therefore, no reliable statistics for November can be shown.

3.3 January

Winter 2002-2003 arrived early to northern Europe. For example, the permanent snow cover came to southern Finland in November, and the mean temperature for October - January was several degrees below the long-term average. January 2003 consisted of a cold high pressure period and a mild period with cyclone activity in northern Europe, and was therefore selected for comparison.

Figure 5 shows the average p_{msl} for January 2003, based on ATA analyses from 00 and 12 UTC. A large low pressure area covers the northern latitudes, with separate centres over the Atlantic south of Greenland and over the Arctic Sea north of Europe. Figures 6-8 demonstrate the p_{msl} bias in 48 h forecasts for January 2003 for ATA, ATB and ATC suites, respectively. The forecasts of each suite are verified against the own analysis. There is a positive bias of a couple of hPa at the northern latitudes, mainly over the cyclone centres, in the ATA forecasts (Fig. 6), while a negative bias is seen in southern Russia and over the Black Sea. The positive bias at the northern latitudes is smaller in ATB (Fig. 7) and ATC (Fig. 8) forecasts than in ATA forecasts. On the contrary, the negative bias in ATB and ATC forecasts is larger than in ATA and the area covered by the negative bias is larger and extends also to Scandinavia. A worrying thing is that the negative bias is more significant in ATC than in ATB. These suites differ only in the horizontal and vertical resolution, and in the time integration scheme.

The month of January was two-fold in terms of temperature in Scandinavia. Extremely cold weather dominated the early part, 1-11 January, when temperatures as low as -40°C were recorded in many places in Finland and northern Europe. This cold period was followed by a mild, synoptically active two-week period, 12-27 January, when cyclones were traveling across Scandinavia. Next, some verification statistics will be presented separately for these two periods.

Figures 9 and 10 demonstrate the observation verification scores of p_{msl} (left) and T_{2m} (right), using observations from Scandinavian stations only, for ATA (top), ATB (middle) and ATC (bottom) suites, and for the periods of 1-11 and 12-27 January, respectively. Figure 9, referring to the cold period of 1-11 January, reveals that there is practically no bias in ATA p_{msl} forecasts. On the other hand, there is a negative bias, which is growing as a function of the forecast length, in ATB and ATC forecasts. The negative biases of these suites are rather large: -1.5 hPa for ATB and of the order of -2.0 hPa for ATC in 48 h forecasts. The rms errors of the three different suites are similar, reaching 4 hPa in 48 h forecasts. This indicates that the standard deviation is smaller in the ISBA-based suites than in the ATA suite. The T_{2m} bias is positive for all suites during this cold period: about 2°C for ATA, 3.5°C for ATB and more than 4°C for ATC in 48 h forecasts. It is a known fact that the operational ATA fails to predict very low temperatures associated with sharp surface inversions. But, it was a surprise that the ISBA-based suites are still worse in predicting the very low temperatures. The rms errors are very large, exceeding the scale of the plots.

Figure 10, demonstrating the verification statistics for the mild period of 12-27 January, shows that the ISBA-based ATB and ATC suites have similar type of negative p_{msl} bias as in the cold period. The ATA routine has a slight positive bias. The verification scores for T_{2m} are completely different when compared to those of the cold period. There is a negative bias between -0.5 and -1°C in the ATA forecasts, whereas the ISBA-based suites ATB and ATC show a positive bias of less than 1°C. In mild conditions the biases are different: negative in case of the old surface scheme, positive in case of the ISBA scheme.

3.4 March

The pre-operational tests continued until mid-March. As in summer and autumn 2002, the ISBA-based ATB and ATC suites gave more accurate T_{2m} forecasts than the operational ATA in the first half of March 2003 (not shown). The ISBA-based suites have a larger p_{msl} bias than ATA over eastern and central Europe, as was seen also for January: the bias distributions for March are comparable to those of Figs. 5-7. Figures 11 and 12 depict the standard deviation, i.e., the random part of the 48 h p_{msl} forecast error, based on forecasts from 00 and 12 UTC analyses, during 1-15 March 2003 for ATA and ATC, respectively. The values of the standard deviations are of the same order for ATA and ATC over continental Europe. Over the Atlantic and over the Arctic Sea, in areas of cyclone tracks, the standard deviation is larger for ATC (Fig. 12) than for ATA (Fig. 11). This is slightly worrying. The reason might be in different versions of the CBR turbulence scheme used in ATA and ATC.

4 T_{2m} analysis problem in cold cases

The observation verification showed that the predicted 2-metre temperatures were several degrees too warm for the operational ATA suite as well as for pre-operational ATB and ATC suites in the cold period of 1-11 January. There were cases, when the 6 h T_{2m} forecast, i.e., the first guess for the T_{2m} analysis (ATB and ATC suites), deviated so much from the observation that the observation was rejected. In this kind of cases the observation has no impact on the analysis (no analysis increment): the first guess defines the analysis. This affects also the surface temperature analysis, because the T_{2m} analysis increment is further used in surface temperature analysis. It is therefore possible that in a cold period the surface temperature (and T_{2m}) can drift far from observations due to the inaccurate T_{2m} first guess and consequent rejection of T_{2m} observations in the analysis. The 6 hourly analyzed surface temperature is the strength of the ISBA scheme, but it is obvious that this benefit can be lost in cold cases. Furthermore, use of the T_{2m} analysis in the field verification causes distortion in the verification scores in such cases. This problem of rejected observations in the T_{2m} analysis can be clarified with Figs. 13 and 14, which illustrate the analysis-observation fit for T_{2m} on 4 January 2003 at 00 UTC. Figure 13 demonstrates the observed T_{2m}. Cold weather with temperatures lower than -30°C dominates Scandinavia and northern Russia, while higher temperatures are seen over Russia south of 60°N. Figure 14 shows the analysis-observation fit, i.e., the departure between the analyzed (ATB) and observed value of T_{2m}. The fit is good, the departure being close zero, over Russia south of 60°N, where the observed temperatures are above -20°C. In northern Russia, where observed temperatures are colder than -25°C, the analyzed T_{2m} can be as much as 15° higher than the observed T_{2m} due to rejection of observations in the analysis. The same feature can be seen over Scandinavia.

Too high predicted 2-metre temperatures in cold cases indicate a model deficiency, most probably associated with snow parameterization. Therefore, urgent attention should be paid to snow parameterization, in order to improve the predicted T_{2m} in cold cases.

5 Conclusions

Pre-operational HIRLAM 5.1.4 tests have been carried out for a lengthy period from August 2002 until March 2003. Forecasts from these test runs have been evaluated and compared to operational HIRLAM forecasts based on the 4.6.2 system.

Pre-operational tests have proved that the ISBA-based 5.1.4 HIRLAM system improves the quality of T_{2m} forecasts almost in all seasons when compared to the operational HIRLAM system with a simple surface scheme. This is a welcome feature because the reputation of HIRLAM has suffered from the poor T_{2m} predictions for years, especially in springtime. The only obvious weakness in the ISBA-based temperature forecasts is prediction of extreme cold temperatures. Winter 2002-2003 provided a real test for the model because the ISBA scheme was in the development and test phase never tested in such cold conditions (e.g., Järvenoja, 2002a, 2002b). Therefore, the poor performance of the ISBA scheme, poorer than that of the simple scheme, in extreme cold temperatures came as a surprise.

Pre-operational tests revealed that the ISBA-based 5.1.4 model tends to have larger negative p_{msl} bias than the 4.6.2 model in some of the test months as shown in connection of results from January 2003. A surprise is also the possible bias dependence on the resolution: the negative bias is larger in a 0.3°/40-level model than in a 0.4°/31-level model.

Apart from problems in prediction of extreme cold temperatures and a larger negative p_{msl} bias in some months, the scores of the pre-operational test runs showed that the ISBA-based 5.1.4 system is better than or at least similar to the old operational 4.6.2 system. Due to more plus than minus points, the ISBA-based 5.1.4 HIRLAM system was accepted into operations in March 2003.

References

Eerola, K., 2002: The operational HIRLAM at the Finnish Meteorological Institute. *HIRLAM Newsletter*, **41**, 19-24.

Eerola, K., 2003: The operational HIRLAM at the Finnish Meteorological Institute. *HIRLAM Newsletter*, **43**, (in this issue).

Järvenoja, S., 2002a: ISBA tests in a Nordic area. Proceedings of the SRNWP/HIRLAM Workshop on Surface Processes, Turbulence and Mountain Effects, INM, Madrid 22-24 October 2001, 64-74.

Järvenoja, S., 2002b: ISBA tests in a Nordic area - an update. *HIRLAM Newsletter*, **41**, 63-73.

Järvenoja, S., 2003: Testing of the MC2 boundary treatment in HIRLAM. *HIRLAM Newsletter*, **43**, (in this issue).

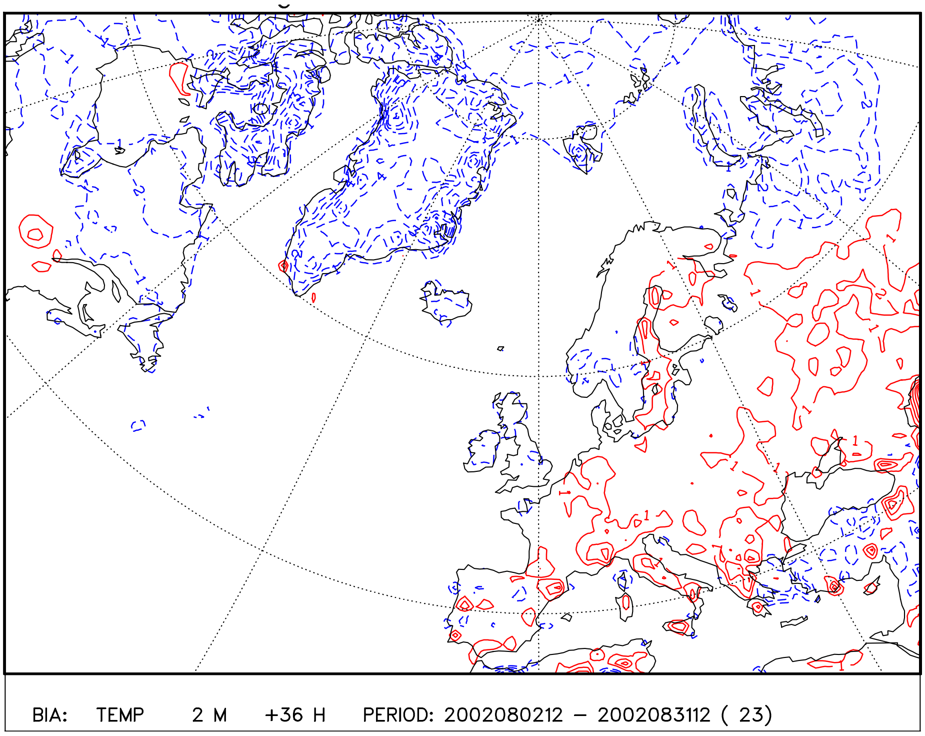

Figure 1: T_{2m} *bias in 36 h ATA forecasts (valid daytime) for August 2002. Contour interval: 1°C. The zero isoline not plotted, negative values indicated with dashed lines.*

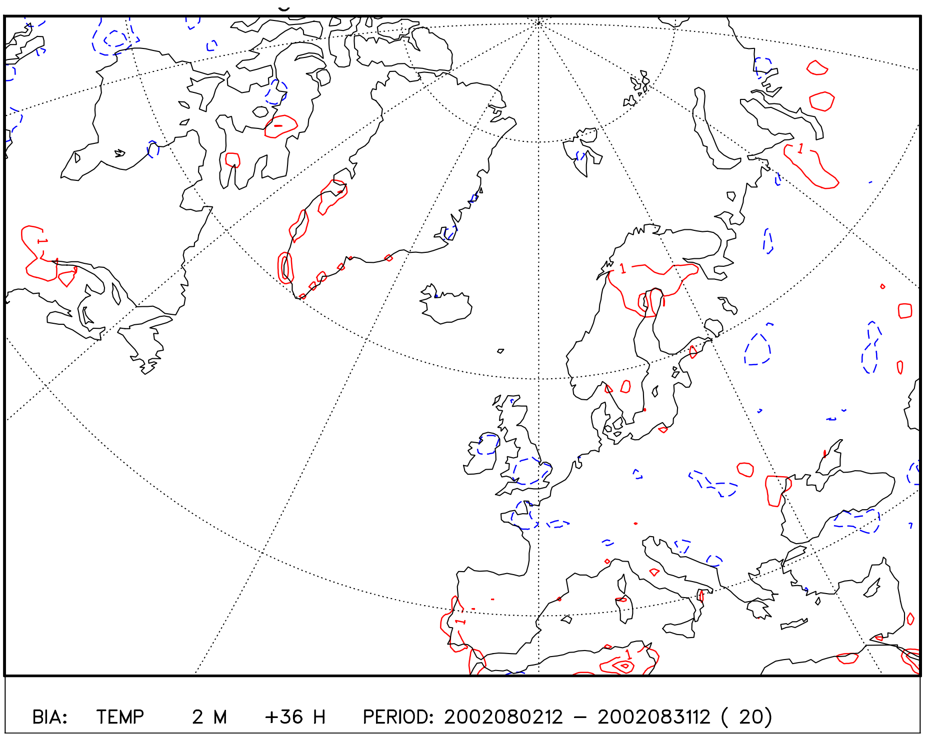

Figure 2: T_{2m} *bias in 36 h ATB forecasts (valid daytime) for August 2002. Contour interval: 1°C. The zero isoline not plotted, negative values indicated with dashed lines.*

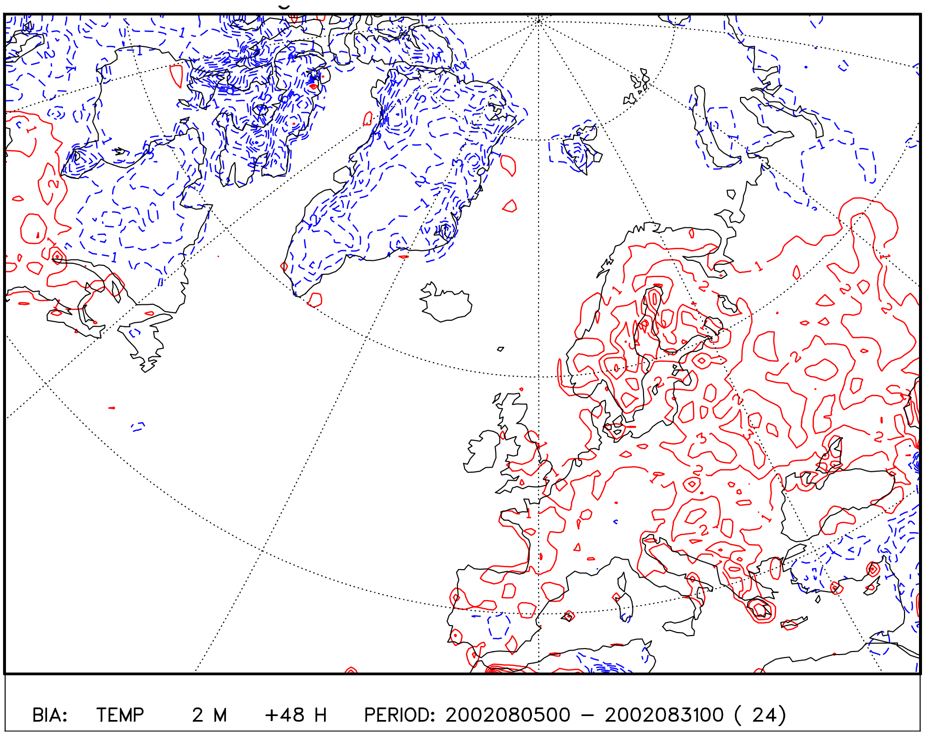

Figure 3: *T_{2m} bias in 48 h ATA forecasts (valid night-time) for August 2002. Contour interval: 1°C. The zero isoline not plotted, negative values indicated with dashed lines.*

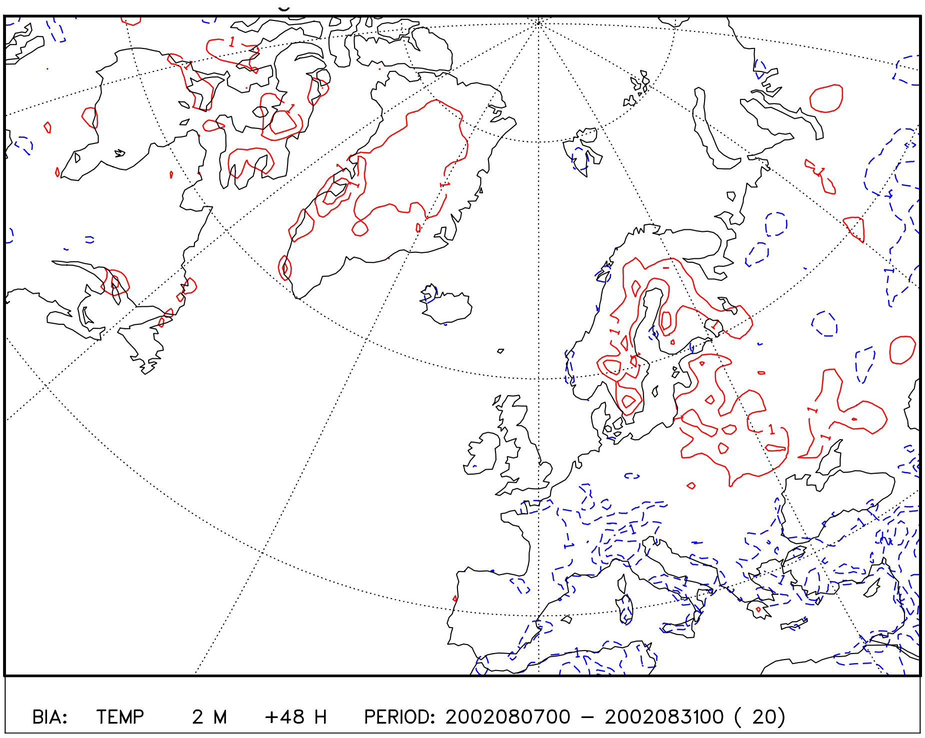

Figure 4: *T_{2m} bias in 48 h ATB forecasts (valid night-time) for August 2002. Contour interval: 1°C. The zero isoline not plotted, negative values indicated with dashed lines.*

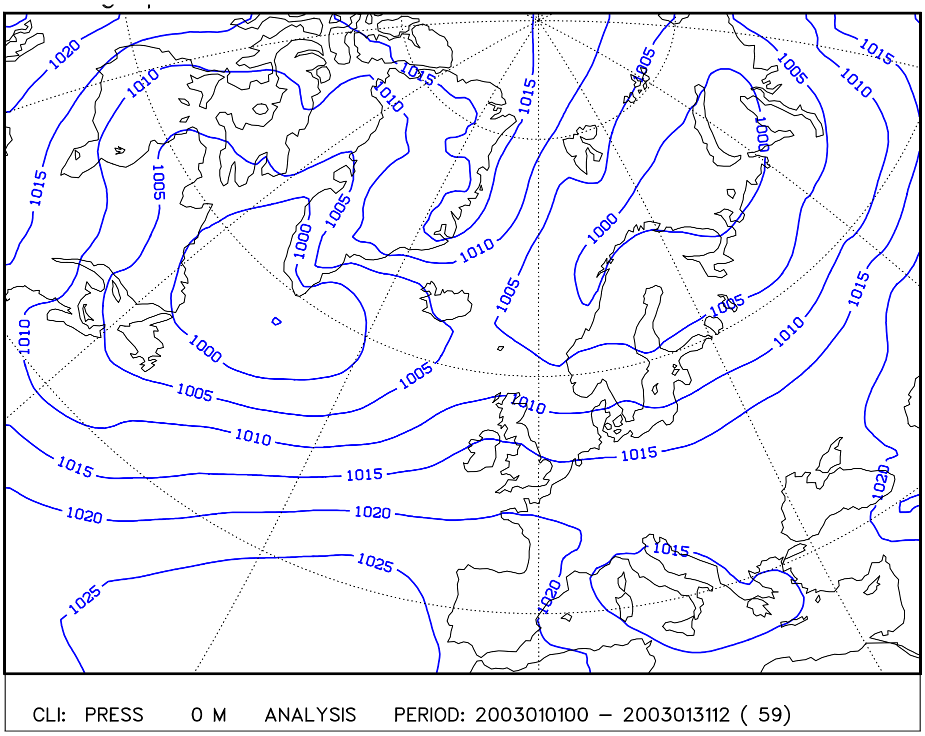

Figure 5: *Average p_{msl} for January 2003, as from ATA 00 and 12 UTC analyses. Contour interval: 5 hPa.*

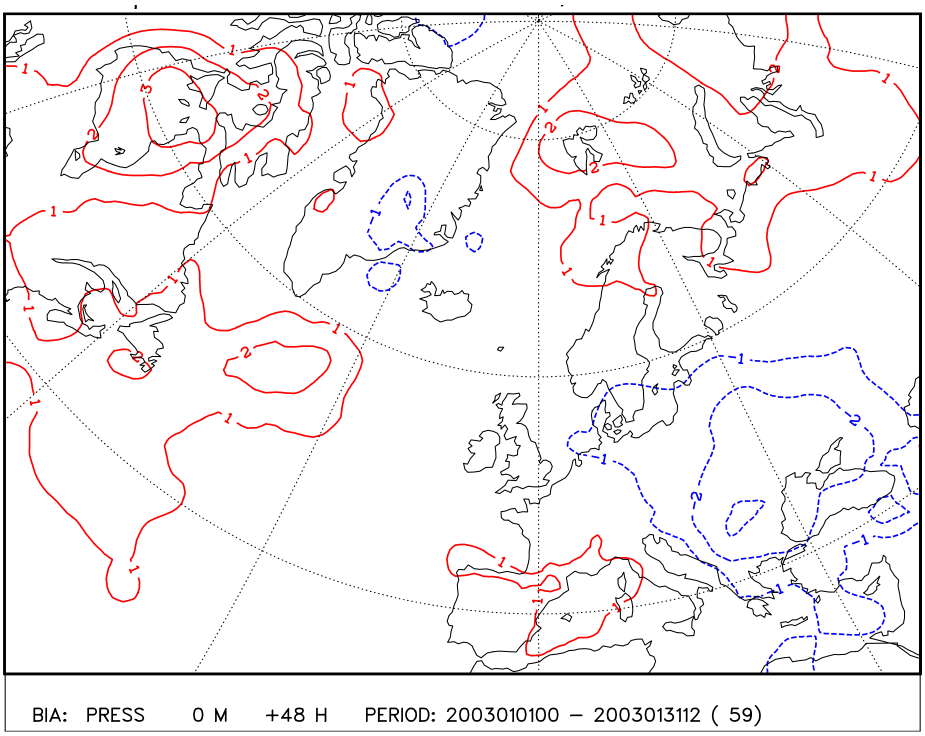

Figure 6: *P_{msl} bias in 48 h ATA forecasts (from 00 and 12 UTC analyses) for January 2003. Contour interval: 1 hPa. The zero isoline not plotted, negative values indicated with dashed lines.*

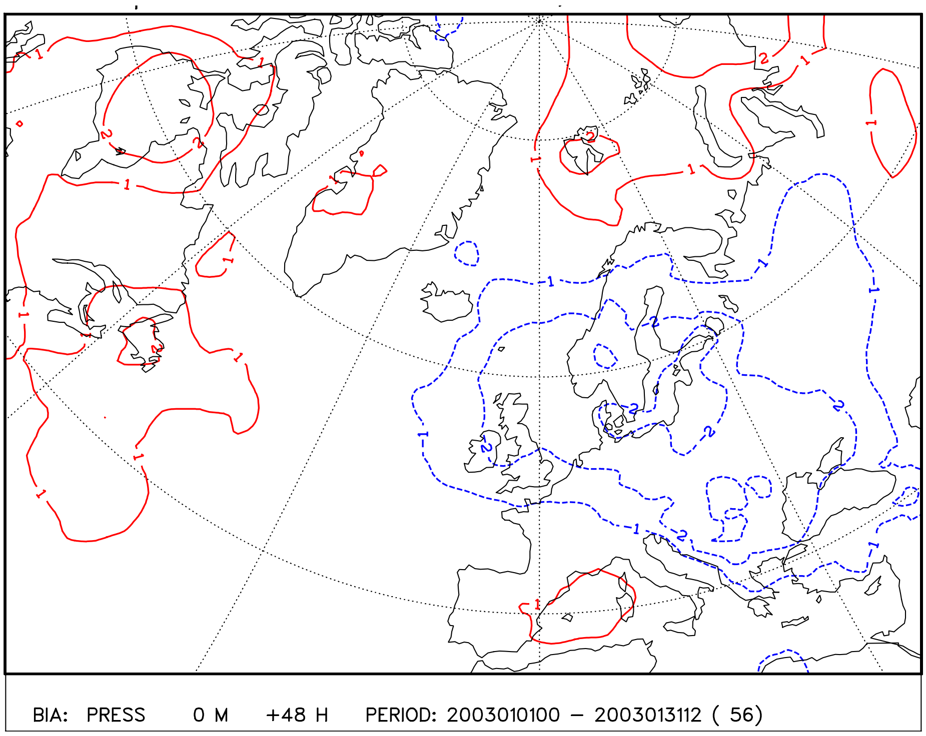

Figure 7: P_{msl} *bias in 48 h ATB forecasts (from 00 and 12 UTC analyses) for January 2003. Contour interval: 1 hPa. The zero isoline not plotted, negative values indicated with dashed lines.*

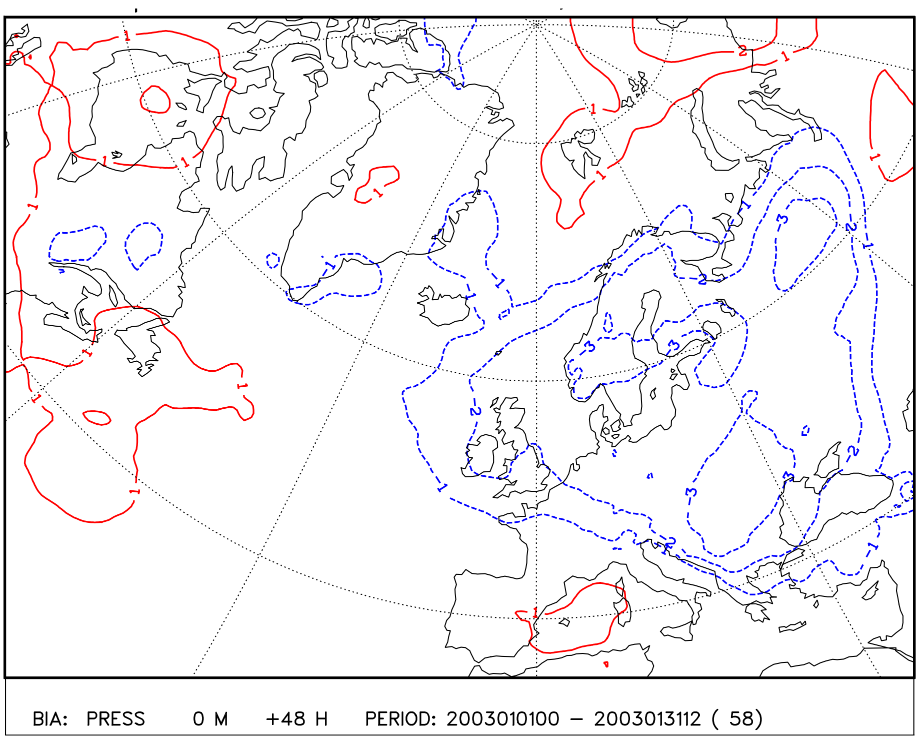

Figure 8: P_{msl} *bias in 48 h ATC forecasts (from 00 and 12 UTC analyses) for January 2003. Contour interval: 1 hPa. The zero isoline not plotted, negative values indicated with dashed lines.*

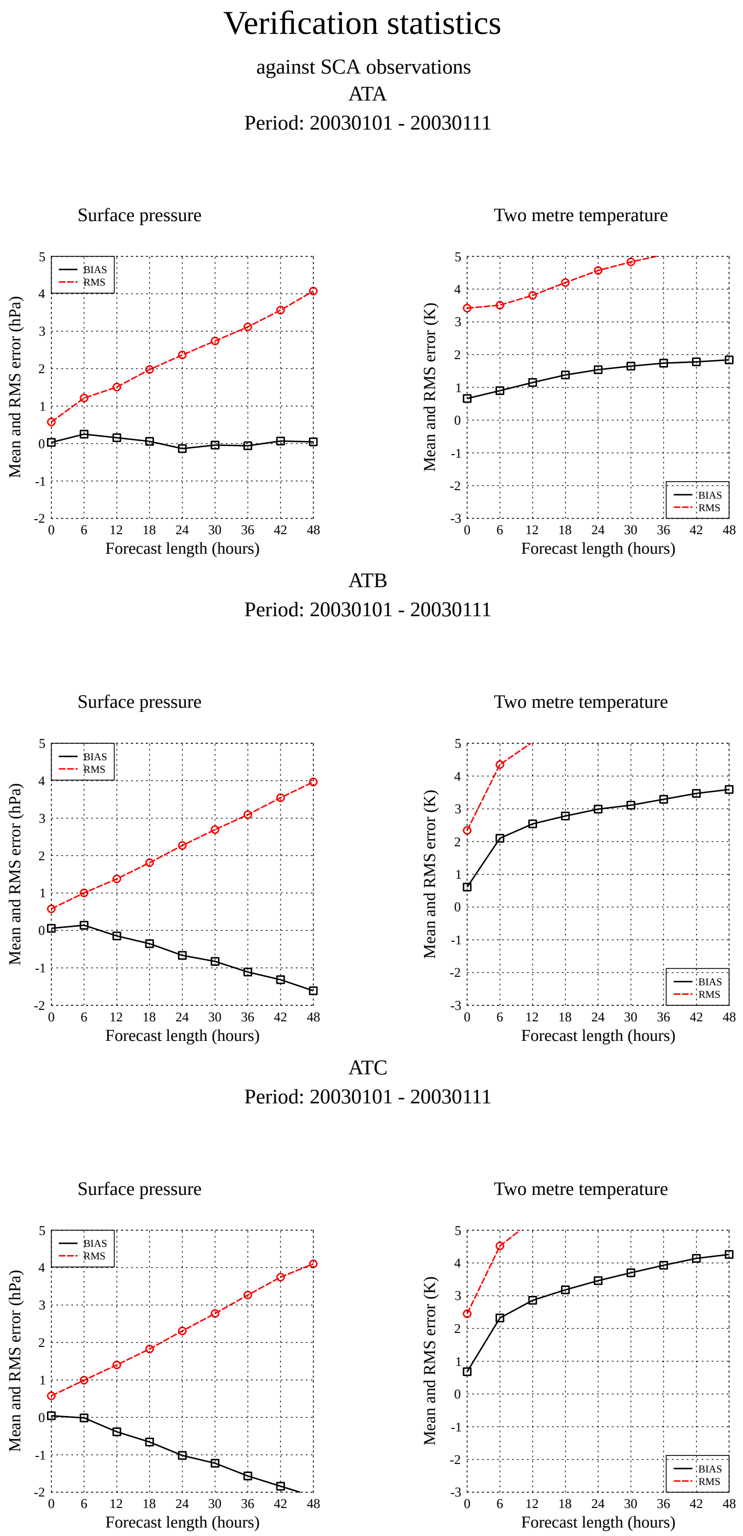

Figure 9: *Observation verification scores of* p_{msl} *(left) and* T_{2m} *(right) for the cold period of 1-11 January 2003, using Scandinavian stations only: ATA (top), ATB (middle) and ATC (bottom). Bias is indicated with squares and rms error with circles.*

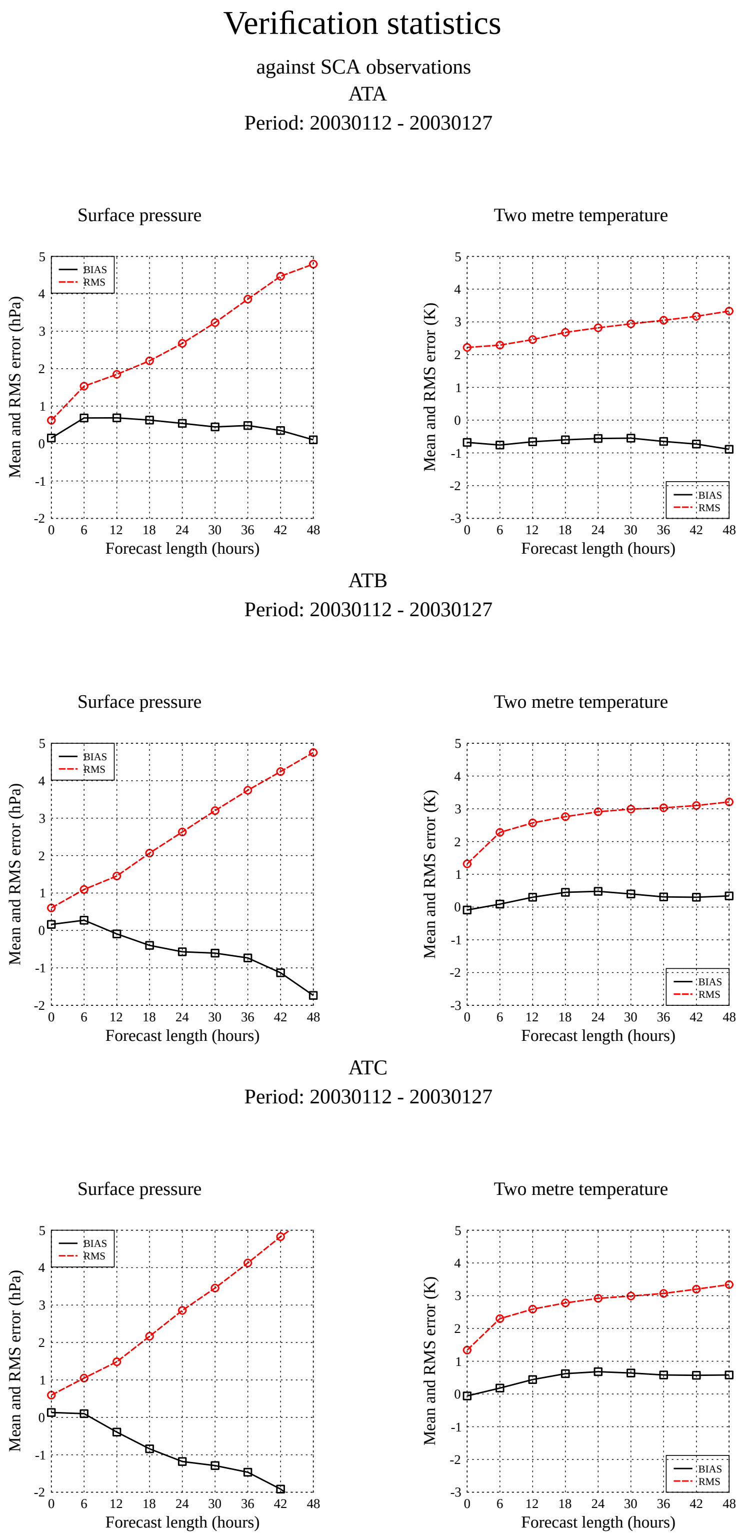

Figure 10: *Observation verification scores of* p_{msl} *(left) and* T_{2m} *(right) for the mild period of 12-27 January 2003, using Scandinavian stations only: ATA (top), ATB (middle) and ATC (bottom). Bias is indicated with squares and rms error with circles.*

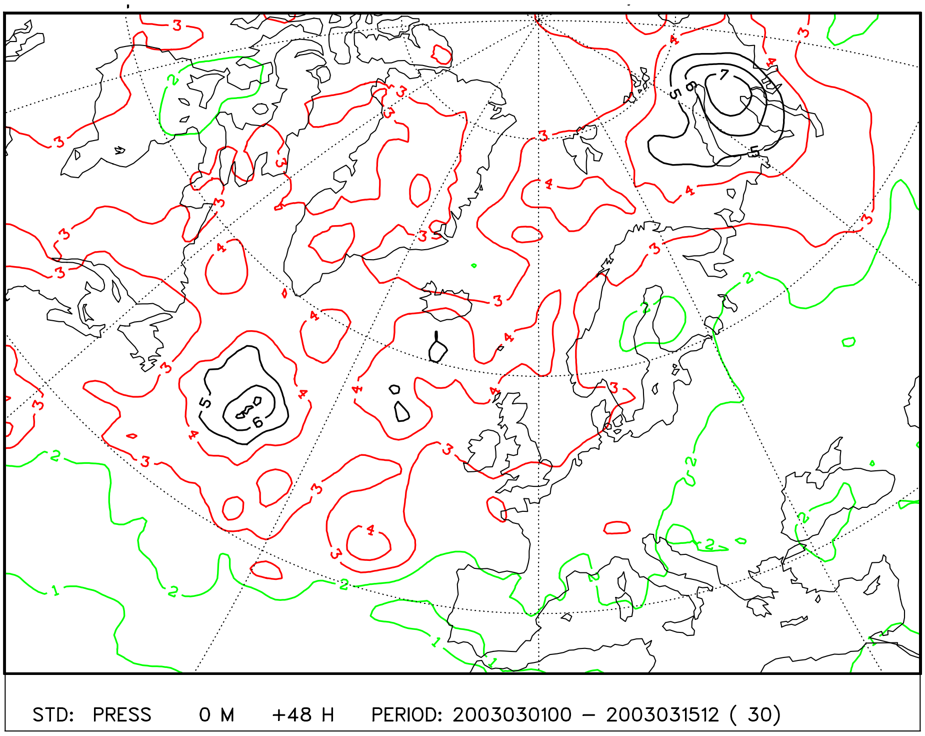

Figure 11: *Standard deviation of* p_{msl} *in 48 h ATA forecasts for 1-15 March 2003. Contour interval: 1 hPa.*

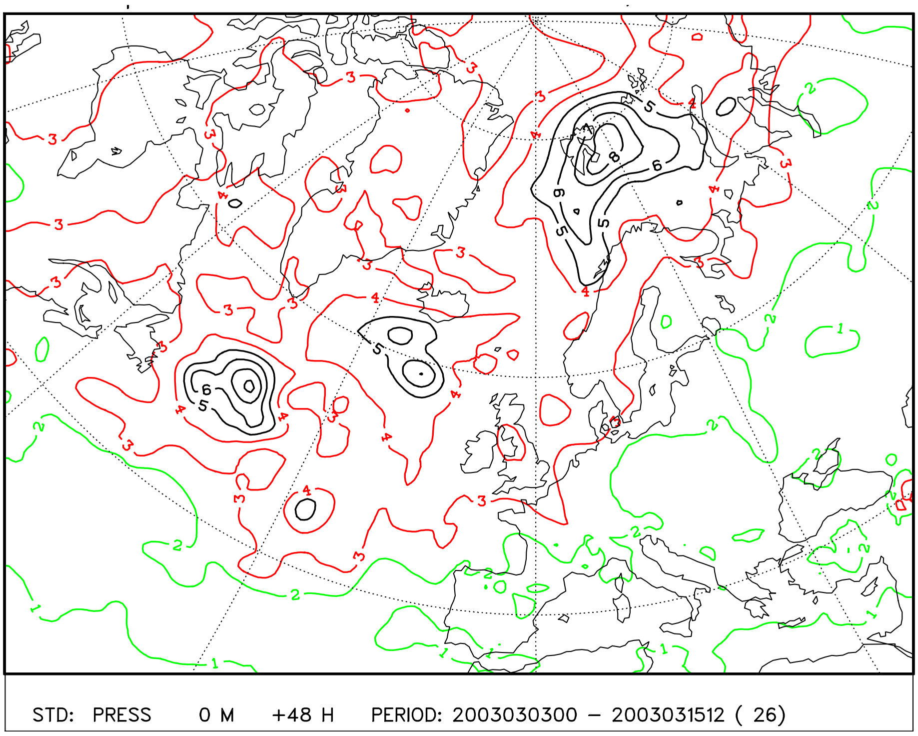

Figure 12: *Standard deviation of* p_{msl} *in 48 h ATC forecasts for 1-15 March 2003. Contour interval: 1 hPa.*

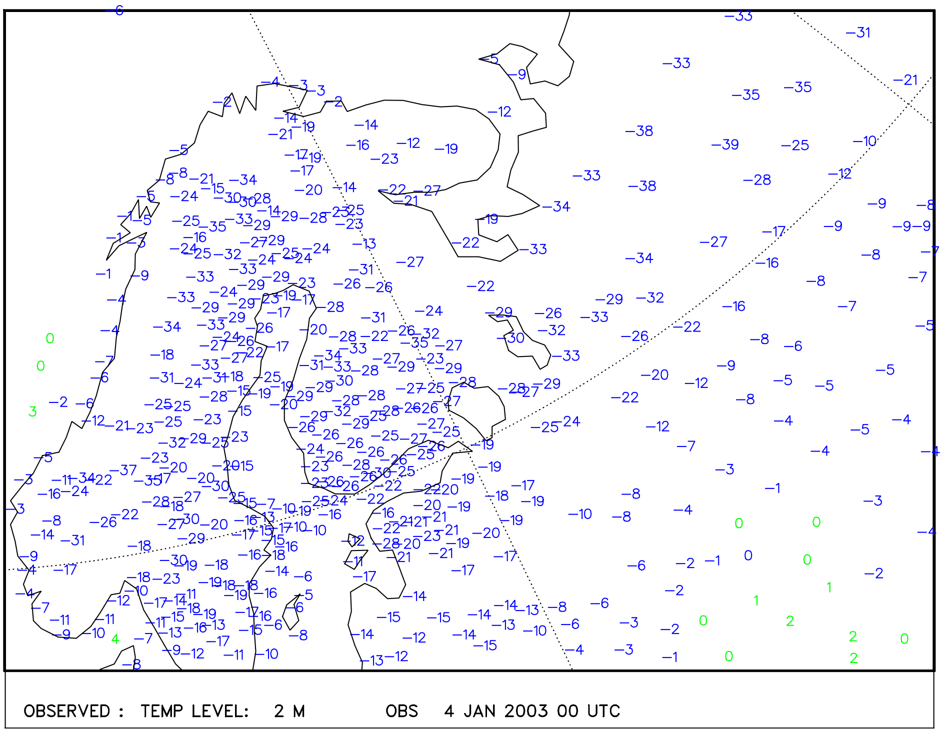

Figure 13: *Observed T_{2m} in northern Europe on 4 January 2003 at 00 UTC.*

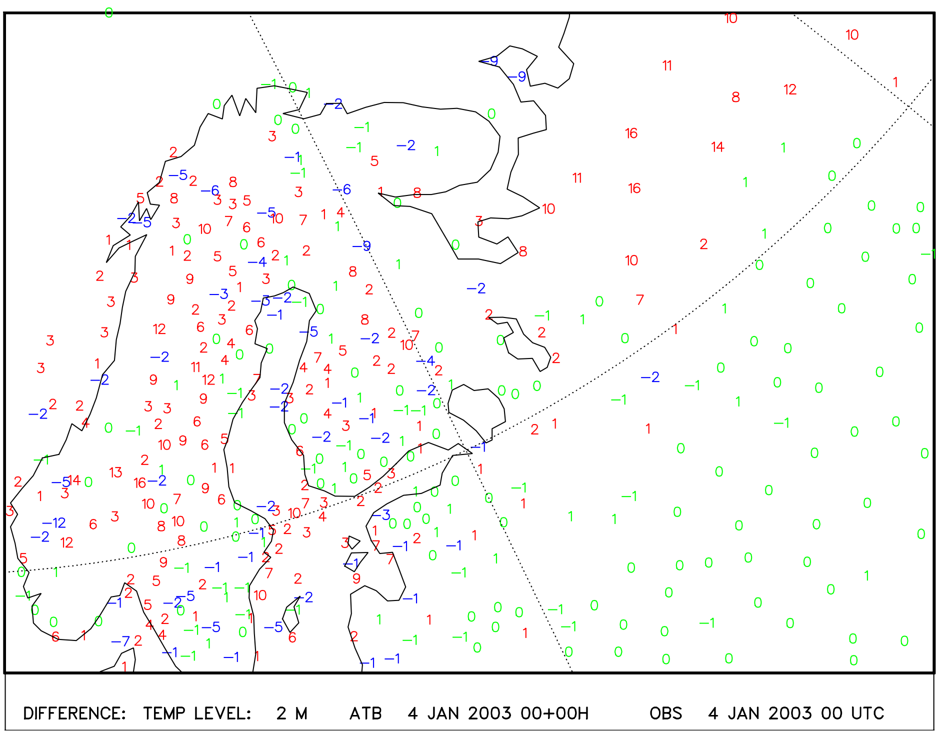

Figure 14: *Analysis-observation fit (ATB analysis minus observation) for T_{2m} in northern Europe on 4 January 2003 at 00 UTC.*