Meteorological Institute
met.no
met.no report

no. 16/2005
HIRLAM, impact study

The Impact of Assimilating AMSU-A Observations over Sea Ice in HIRLAM 3D-Var

Vibeke Wauters Thyness, Frank Thomas Tveter and Harald Schyberg
Norwegian Meteorological Institute
Oslo, Norway

Title	Date
The Impact of Assimilating AMSU-A observations over Sea Ice in HRILAM 3D-Var	5.th October 2005
Section Section for Remote Sensing	**Report no.** no. 16
Author(s) Vibeke Wauters Thyness, Frank Thomas Tveter and Harald Schyberg	**Classification** ● Free ○ Restricted
	ISSN 1503-8025
	e-ISSN 1503-8025
Client(s) NA	**Client's reference** EVK3-2001-00116.

Abstract
The Norwegian Meteorological Institute has developed the HIRLAM 3-dimensional variational data assimilation system so that it is capable of assimilating AMSU-A radiance observations over sea ice. A 2 month long impact study of this implementation during early spring 2005 showed that assimilating AMSU-A observations over sea ice had a positive effect on the average mean sea level pressure forecasting skill and the temperature and wind profiles, especially in the Nordic regions.

Keywords
HIRLAM, ATOVS, impact study

Disiplinary signature	Responsible signature

Postal address P.O Box 43 Blindern N-0313 OSLO Norway
Office Niels Henrik Abels vei 40
Telephone +47 2296 3000
Telefax +47 2296 3050
e-mail: met.inst@met.no **internet:** met.no
Bank account 7694 05 00601
Swift code DNBANOKK

Contents

1 Introduction

In the IOMASA project, the Norwegian Meteorological Institute has developed a method for assimilating AMSU-A observations over sea ice in cooperation with the IOMASA project partners. AMSU-A (Advanced Microwave Sounding Unit) is a passive microwave instrument carried by the polar orbiting NOAA (National Oceanic and Atmospheric Administration) satellites. The usefulness of assimilating AMSU-A radiance observations in general has been demonstrated by many Numerical Weather Prediction centres. In the HIRLAM (High-Resolution Limited Area Model) model, the impact has been demonstrated using observations over open sea only (Schyberg et al., 2003). This report describes the work done at the Norwegian Meteorological Institute to assimilate AMSU-A observations over sea ice in HIRLAM 3D-Var.

An important component in the system for assimilating AMSU-A observations in HIRLAM 3D-Var, is the 'forward model', RTTOV (Saunders and Brunel, 2002). The forward model simulates the AMSU-A observations using a model representation of the atmosphere, and the data assimilation system then tries to optimally adjust the HIRLAM representation of the atmosphere so that the RTTOV simulated radiance observations better agree with the observed radiances.

For AMSU-A channels that are sensitive to the surface, the RTTOV forward model requires an estimate of the emissivity of the surface. The error in the forward model in channels that are sensitive to the surface emissivity, decreases as the emissivity estimate improves. Simple, yet very successful, AMSU-A emissivity models depending on surface windspeed are available for open sea (English and Hewison, 1998). Unfortunately, the nature of sea ice surfaces is such that emissivity depends on the history of the sea ice, and has proven to be more difficult to model in a simple way.

On the other hand, the sea ice emissivity properties have some stability in time, and first-year and multi-year ice have distinctly different emissivities. Thus, in this work we exploit daily sea ice charts showing first-year and multi-year sea ice concentration retrievals from an experimental version of the Ocean and Sea Ice Satellite Application Ficility (OSI SAF) processing chain. A natural basis for emissivity modelling would be to split the detected radiance into contributions from open water and various surface types in the footprint, assuming a constant surface temperature. Typical microwave emissivities for the sounding channels have been derived by Dr. Leif Toudal at the Danish Technical University (DTU). Further details of the method was given in Thyness et al., 2005a, and we also give some details below.

Another difficulty in setting up an assimilation scheme for AMSU-A over sea ice lies in identifying observations contaminated by clouds. RTTOV simulates AMSU-A observations in clear sky conditions, and when the forward model is used in cloudy conditions, the RTTOV model error increases, and usually a bias error occurs.

In the experiments presented in this report there is no masking of AMSU-A observations that are contaminated by cloud water. Instead, the Norwegian Meteorological Institute has developed a theory for optimal asymmetric quality control. This approach is based on the Bayesian risk of assuming that the observation error has a normal distribution, when it actually has a non-normal distribution. The method is described in Tveter, 2005, and Tveter and Thyness, 2005, and we also give some further details on the method below.

The first section in this report describes the AMSU-A emissivity estimation sea ice. The following section briefly describes the theory for observation quality control and data assimilation. The last section contains a detailed discussion of the impact of the assimilation of AMSU-A observations over sea ice.

2 Surface Emissivity over sea ice

As a starting point for using 'first-year' (FY) and 'multi-year' (MY) sea ice concentrations to estimate emissivities of the soundig channels, a method for estimating typical average surface emissivities within an AMSU-A footprint was developed by Toudal (2005). The first step was to identify areas completely covered by FY ice (for instance the Kara Sea) and areas completely covered by MY ice (for instance North of Greenland), and stable periods with minimal water vapour. Typical surface emissivities of FY and MY ice were then determined empirically (Toudal, 2005) for the selected study areas, using a simplified theory for microwave radiative transfer, explained below.

If the atmosphere is modelled as a single layer with temperature T_a, the radiative transfer equation for the total microwave radiation received at satellite altitude over an area with FY ice, converted to brightness temperature using Plancks law, T_b, can be written as

$$\begin{aligned} T_b &= \epsilon_{\mathrm{FY}} T_s (1-\alpha) + \alpha T_a + \alpha T_a (1-\alpha)(1-\epsilon_{\mathrm{FY}}) \\ &\quad + T_{sp} (1-\alpha)^2 (1-\epsilon) \end{aligned} \tag{1}$$

where ϵ_{FY} is the FY ice emissivity, T_s is the surface temperature, α is the absorption coefficient for the atmosphere and T_{sp} is the brightness temperature from space (2.7 K). We have here also assumed that the atmospheric attenuation can be reasonably approximated by an absorption factor α and that the water vapour load is minimal (so that the main contribution to the absorption is from oxygen, and not the water vapour). The downwelling atmospheric temperature, T_{dn}, can be modelled according to

$$T_{dn} = T_{sp}(1-\alpha) + \alpha T_a. \tag{2}$$

If we estimate T_{dn} from a radiative transfer model (MWMOD (Furhop et al. 1998)), and use a climatological value for T_a in Eqs. (1) and (2), we get

$$\epsilon_{\mathrm{FY}} = \frac{T_b - \alpha T_a - (1-\alpha)\alpha T_a - T_{sp}(1-\alpha)^2}{T_s(1-\alpha) - \alpha T_a (1-\alpha) - T_{sp}(1-\alpha)^2}$$

where the absorption factor α is given by

$$\alpha = \frac{T_{dn} - T_{sp}}{T_a - T_{sp}}.$$

This approach gives typical emissivities for areas with FY ice. The same procedure can be applied in areas with MY ice, to yield typical MY emissivities, ϵ_{MY}. Table 1 shows the typical emissivity values for AMSU-A for the two ice types found with this approach.

The typical emissivity values are used in combination with the ice concentration products from the OSI SAF to determine the surface emissivity in each AMSU-A footprint. Since the properties of the ice surface usually change slowly, information from recent passages of microwave radiometers can help determine the ice concentration in the AMSU-A footprints. The OSI SAF provides daily sea ice retrievals from the Arctic on a 10 km grid, based mainly on SSM/I (Breivik et al., 2001). The OSI SAF service has recently been extended to also produce estimates of FY and MY ice concentrations.

Over completely sea ice covered areas, the surface within each AMSU-A footprint is subdivided into concentrations of FY (c_{FY}) and MY (c_{MY}) sea ice from the OSI SAF products, with $c_{\mathrm{FY}} + c_{\mathrm{MY}} = 1$. The surface emissivity is estimated as $\epsilon = c_{\mathrm{FY}}\epsilon_{\mathrm{FY}} + c_{\mathrm{MY}}\epsilon_{\mathrm{MY}}$. Here ϵ_{FY} and

ϵ_{MY} are the typical emissivities for FY and MY sea ice, which are estimated separately for each AMSU channel using the method described above. More details on the implementation of the sea ice emissivity model in the preprocessing system at the Norwegian Meteorological Institute can be found in Thyness et al., 2005b.

3 Data assimilation and Observation Quality Control

Prior to data assimilation, a quality control is performed to reject observations which are contaminated by clouds in such a way that they are not well described by the forward model or contaminated with other types of gross errors. Such errors are generally non-Gaussian, and the observations would have a detrimental effect on the verification scores if used.

The observation quality control method used for the experiments presented in this report estimates the increment in the verification score when we assimilate each observation with non-normal error probability distribution while assuming that it has normal error distribution. If it is expected that the verification score will improve, the observation is used, otherwise it is discarded. The method is described in detail in Tveter, 2005. In short, the Bayesian risk is formulated so that it reflects the improvement in the impact which we would have if we used an observation with non-Normal error distribution, while the assimilation system assumes it has Normal error distribution. The risk increment is given as a function of the innovation, i.e. the departure between the observation (Y) and the first guess equivalent (HX_b).If the risk increment is less than zero, the observation has a positive effect on the analysis (i.e. the mean squared error verification score will improve). The innovation thresholds for 'valid' AMSU-A observations are determined so that each accepted channel gives a negative risk increment.

Figure 1 shows an example of the innovation probability distribution (grey circles), the estimated non-normal innovation distribution (dashed), the estimated normal components (grey) and the optimal normal innovation Approximation for noaa16, AMSU-A, channel 2 (31.4 GHz), based on 400.000 observations made in February and March 2005. The lower panel shows the risk increment as a function of the innovation. Note that the risk increment has been truncated to show details close to the x-axis. The AMSU-A channel 2 has a strong surface contribution and it is sensitive to cloud water contamination. The valid innovation range for this channel was $-28.9K$ to $25.9K$. The optimal normal approximation had a bias of $-10.4K$, with a bias of $-8.7K$ in the valid innovation data. Note how there is a second maximum at $-43.5K$ (probably caused by cloud water contamination or ice type misclassification), and that this gives a valid innovation range that is asymmetric with respect to the optimal normal approximation. Figure 2 shows the innovation and risk increment for noaa16, AMSU-A, channel 5 (53.6 GHz). The valid innovation range for this channel was $-2.1K$ to $2.1K$. The optimal normal approximation had a bias of $0.19K$, with a bias of $0.24K$ in the valid innovation data. Note that the bias in the normal approximation is not the same as the bias in the valid innovation data.

Tables 2 and 3 give the Quality control parameters used in HIRLAM 3D-Var for AMSU-A for the Noaa15 and Noaa16 satellites. The innovation range ($Y - HX_b$) and bias (μ) are determined using the method described above. Note that, due to a bug in the system, the error weights (R_{diagonal}) given to the different AMSU-A channels in this study are artificially high for the lower 5 channels (which are influenced by surface emissivity). The values used here were test values (based on qualified guessing) designed for an earlier experiment trial. Theory for determining more realistic error weights has been developed at the Norwegian Meteorological Institute, and is currently being tested.

4 Results

An impact study was performed over two months from late February until the end of April 2005. Version 6.2.1 of the HIRLAM forecast model was used. The model used digital-filtering initialisation, semi-Lagrangian advection scheme and physical parameterisations according to 'Savijärvi radiation', 'STRACO condensation', 'CBR turbulence scheme' and 'ISBA surface scheme' (Undén et al., 2002). Frames were taken from the ECMWF global model as lateral boundaries. The resolution was 20km. Version 6.2.1 of the HIRVDA (HIRlam Variational Data Assimilation) was used to assimilate observations in a 3 hour assimilation cycle with first guess at appropriate time (FGAT) enabled.

Further details on the modifications done in HIRVDA to assimilate AMSU-A over ice is given in Thyness et al. 2005a. The experiment trial assimilated AMSU-A over sea ice and conventional observations. The reference trial was identical to the experiment trial, except that it did not use the AMSU-A observations.

Note that we do not assimilate any AMSU-A at all in the reference. This means that the reference has less observations than the operational runs at the Norwegian Meteorological Institute, which do use AMSU-A over ocean. An alternative approach could have been to include the ocean AMSU-A data in both runs. With our chosen approach we probably obtain a clearer impact of the sea ice observations, as the effect of redundancy with other observations reducing the impact becomes less. Experiments with adding an observation system to a minimum set of observations always give more positive impact than the negative effect of removing the same observation system from a run using a full set of observations. Thus our experiment is not quite realistic in terms of the effect of adding sea ice AMSU-A data to the present operational system, but it serves to highlight the effect of these sea ice observations, and demonstrate that these observations contain information which the assimilation system can benefit from.

There was no observation cut-off (AMSU-A data typically has a half hour time delay), and typically 600 AMSU-A observations were available for use. About 60 % of the observations were discarded in the thinning, so that 200 were typically available for an assimilation cycle. Note that the geometry of the NOAA polar orbit causes the observation density to increase with latitude.

4.1 General results

Figure 3 shows the geographical distribution of the standard deviation in the $500hPa$ geopotential height (Z500) analysis between the experiment and reference trials, for the whole model area. The differences in the analysis are largest over the arctic sea ice, as we would expect. There are no visible differences over central Europe, which is as expected since there are many accurate conventional radiosonde observations in this area, and these are used in both the experiment and reference trials. Figure 4 shows the corresponding standard deviation in the Z500 48-hr forecast. We observe that perturbations in the experiment trial are advected inland from the sea during the forecast period.

Figure 5 (left panel) shows the standard deviation (Std) and (Bias) in MSLP between independent reference observations (the EWGLAM station list) and both the reference (no AMSU-A over sea ice) and experiment (AMSU-A over sea ice) forecasts over a period of two months. The right panel of figure 5 shows the same features, but the verifying observations are a subset of the EWGLAM stations situated in the Nordic countries. We observe that assimilating the AMSU-A observations improved the verification score, and more so for the northern region. There was no significant impact in the other surface parameters. Figure 6 shows how the

improvement in MSLP accumulated during the verification period, for all EWGLAM stations and the nordic subset respectively. We can see that there were in particular two periods (March 10-19, April 14-17) where the AMSU-A observations had large positive effects on the verification. Few observations (and long forecasts) explains the dip at the very end of the period. The AMSU-A instrument is designed to give temperature sounding information. Figure 7 shows the vertical profile of the standard deviation between independent reference observations from the EWGLAM list and the two trial forecasts. We observe that the AMSU-A observations improved the verification score in the temperature and geopotential profiles. There were also some improvements in the wind vector profile (not shown). AMSU-A observations had a neutral impact on the other parameters. Figure 8 shows the correlation between the error in the MSLP for the reference trial (without AMSU-A over sea ice), and the difference between the experiment and reference trials. We observe that the experiment (using AMSU-A) has much better forecasts for many of the cases where the reference experiment had a large error in MSLP.

4.2 Discussion in relation to weather situations

Inspection of the weather maps produced from the two runs show small differences most of the time. There are however periods of some differences which we shall examine in more detail.

The 500 hPa flow indicates how weather systems move. The upper flow in areas near the ice edge in the north Atlantic shows that the period from 7. to 20. March was particularly associated with air flow of varying strength from the ice towards the Northeast Atlantic. Upper advection from sea ice to North-East Atlantic was particularly strong on 8-14 March and from 18 to 20 March. As an example Fig. 9 shows the height of the 500 hPa surface on 12 March 00Z. This precides a peak in RMS error in the reference run, where we get improvements in the experiment run.

We generally see that these periods of upper air flow from the ice coincide with, or is before the periods seen on the timeseries plot with better scores for the experiment. In the rest of March the flow was weaker or more westerly in the region, coinciding with periods with less impact.

We will discuss a few comparisons of interest between the two runs. First we will study the situation on 14 March 00Z, where the RMS errors of the reference run is highest in the whole experiment period, and where we see large improvements in the experiment run.

We start looking at some differences in the forecast from 12 March. In the 48 hour forecast we observe that the added Arctic observations not only give differences between the runs near the Sea Ice, but also North of Scotland. In Figs 10 and 11 we find that both forecasts have a phase error in the location of the low pressure centre Northwest of Scotland. However, the low pressure centre is located closer to the analysed location and it's strength is more correctly depicted in the experiment run, something which could have been appreciated by a duty forecaster.

The 48 hrs forecasts valid for 14 March 00Z also show large differences in the Barents Sea area, closer to the actual location of the added observations. The same is the case for the 24 hrs forecast valid for the same time. In Fig 12 we compare the two runs in this region, and find quite large differences. There are not many conventional observing stations in the region of the low pressure system appearing in the experiment run between Norway and Spitzbergen, except for a ship observation off the coast of Norway and the island stations Bjørnøya and Hopen. These observations, in particular the two island stations, seem to give some support that the forecast from the experiment run is more realistic. It is worth noting that the pressure values for these stations show particularly large deviations from the reference run.

In Fig. 13 we present the corresponding analyses from the two runs (valid 14 March 00Z). Usually we find that the difference between the forecasts from the two runs increase with forecast time, but the assimilation brings them together, so the analyses are more similar than the forecasts. Thus, it is usually possible to use the analysis from either of the runs to verify the forecasts.

In this case however, the analyses seem to remain different, keeping the different features of the previous forecasts in each of the two runs. We find surprisingly large differences between the two analyses, and observe that the differences between the two 24 hrs forecasts are basically maintained in the analysis. There are areas where the difference between the two analyses is more than 10 hPa. This is probably related to the fact that there are so few conventional observations in the area, and they are not sufficient to bring the two forecasts together. The observations give support to the analysis of the experiment run.

Prior to being used in the analysis, the observations go through a quality control, where the observations that are too far away from the first guess 6 hrs forecast are considered erroneous and rejected. It is probable that in the reference run the Bjørnøya and Hopen station pressure values have been so far off from the first guess that they have been rejected in the analysis of the reference run. In this case, the difference between the MSLP in the reference run and the SYNOP station Bjørnøya was more than 10 hPa, which is of the same order as the observation error defined in the system.

In Fig. 14 we demonstrate that for Bjørnøya the reference run actually allow deviations from that observation station for a long period of time.

In Figs. 15 and 16 we show 48 hrs forecasts in the Barents Sea area valid at 18 March 00Z. Also at this time there was a peak in the model errors as verified against the EWGLAM stations, and again the experiment scored better. In this case the analyses of the two runs seem to converge, so it is possible to use the analysis from the reference run valid at the same time as a verification. Again we observe that both this analysis and SYNOP observations support the experiment run.

These case studies show that the benefitial impact of the added AMSU-A observations over sea ice can remain in the model for a long time and can have large effects on the simulated circulation pattern for instance in the Barents sea. This region and other parts of the ocean areas adjacent to the ice edge have few conventional observations, and the study shows the importance of adding satellite information in these areas. The study also shows that in these observation-sparse areas, single observations showing that the forecast is in error may not be sufficient for correcting the forecast in the assimilation system. There can be a problem with rejection by the quality control unless other observations in the region are added which support such single observations.

These examples show that in observation-sparse regions near the sea ice, added AMSU-A observations have a generally positive effect. The effect is related to synoptic situation, and the benefit is particularly seen in the EWGLAM verification statistics over Europe when there is a general upper level flow from the sea ice towards the North Atlantic. In such cases we also can get improvements far south of the region of added observations.

5 Conclusions

AMSU-A observations over sea ice were successfully assimilated in HIRLAM 3D-Var at the Norwegian Meteorological Institute. A new approach for estimating microwave sounding channel emissivity over sea ice was developed using information from the EUMETSAT OSI SAF.

A quality control method developed specifically to handle asymmetric (cloud water) contamination was also developed. Assimilating AMSU-A observations over sea ice using the proposed approach had a positive effect on the MSLP and temperature profiles, especially in northern areas.

It is generally difficult to measure the impact of the added observations over the sea ice regions, beacuse there are almost no conventional observations which can be used for verification in the region.

When we measure the impact using the conventional observation network over Europe, we find that the impact varies with circulation pattern, and that the largest positive impact coincided with general upper flow from the sea ice towards the North Atlantic. This clearly shows that there is information contained in these observations which the assimilation system can benefit from.

In particular we saw that in areas in the Barents sea with sparse conventional observation coverage, the positive impact could remain in the model for a long period of time, and that in the absence of satellite data, problems of rejection by quality control can occur. This shows the importance of adding new data in data sparse regions, but also the importance of increasing the robustness of the assimilation system by having a certain density of observation information to avoid data rejection problems.

Even if the experiment period is limited in time, and the positive impact seem to originate from a few incidents within this period, this study strongly indicates benefits of adding AMSU-A observations over sea ice, and it should be planned to include this in operational model runs with HIRLAM.

6 Acknowledgements

This work was sponsored by the EU Fifth Framework Programme project IOMASA (Integrated Observing and Modelling of the Arctic Sea Ice and Atmosphere), project no EVK3-2001-00116. The authors are also grateful for the discussions with Dr. Leif Toudal.

7 References

Breivik, L.A, S. Eastwood, Ø. Godøy, H. Schyberg and R. Tonboe, 2001: Sea Ice products for EUMETSAT Sattelite Application Facility. Canadian Journal of Remote Sensing 27(5), 403-410.

English, S. and T.J.Hewison, 1998: A fast generic millimetre wave emissivity model. Proceedings of SPIE, vol 3503, 288-300.

Furhop, R, T.Grenfell, G. Heygster, K-P. Johnsen, M.Schlssel and co-authors, 1998: A Combined Radiative Transfer Model for sea ice, Open Ocean and Atmosphere. Radio Science 33(2), 303-316

Saunders, R. and P. Brunel, 2002: NWP SAF Technical Report NWPSAF-MO-TR-009. Available at http://www.metoffice.com/

Schyberg, Harald, Tomas Landelius, Sigurdur Thorsteinsson, Frank Thomas Tveter, Ole Vignes and others, 2003: Assimilation of ATOVS data in the HIRLAM 3D-Var system. HIRLAM Technical Report No. 60, April 2003. Available from HIRLAM-5 Project c/o Per Undén, SMHI, S-60176 Norrköping, Sweden.

Schyberg, H., V.W.Thyness and F.T. Tveter, 2005: AMSU-A assimilation over sea ice in HIRLAM 3D-Var: Impact studies for the period Februay-March 2005. Norwegian Meteorological Institute Note no 13/2005. Available from the Norwegian Meteorological Institute, http://www.met.no, P.O.Box 43 Blindern, NO-0313 Oslo, Norway.

Thyness, Vibeke, Frank Tveter and Harald Schyberg, 2005a: Assimilating AMSU-A over ice in HIRLAM 3D-Var: Initial method and some results. Norwegian Meteorological Institute Note no. 2/2005. Available from Norwegian Meteorological Institute, P.O.Box 43 Blindern, NO-0313 Oslo, Norway.

Thyness, V.W, F.T. Tveter, H. Schyberg 2005b: Ice Concentration Input for Assimilation of AMSU-A in HIRLAM 3D-Var. Norwegian Meteorological Institute Note no. 3/2005. Available from the Norwegian Meteorological Institute, http://www.met.no, P.O.Box 43 Blindern, NO-0313 Oslo, Norway.

Toudal, L. 2005. Personal correspondance.

Tveter, F.T, 2005: On optimal observation quality control theory for numerical weather prediction systems. Norwegian Meteorological Institute Note no. 7/2005. Available from the Norwegian Meteorological Institute, http://www.met.no, P.O.Box 43 Blindern, NO-0313 Oslo, Norway.

Tveter, F.T. and V.W. Thyness, 2005: Quality control for AMSU-A observations over sea ice. Norwegian Meteorological Institute Note no 12/2005. Available from the Norwegian Meteorological Institute, http://www.met.no , P.O.Box 43 Blindern, NO-0313 Oslo, Norway.

Undén, P, L. Rontu, H. Järvinen, D. Lynch, J. Calvo and co-authors, 2002: HIRLAM-5 Scientific Documentation. Available at http://hirlam.knmi.nl

Appendix: Methodological comments

MSE verification

The basis for our Mean Squared Error (MSE) verification is the penalty function $\mathrm{MSE}[t] = E[(\mathbf{X}[t] - \mathbf{X}_t[t])^2]$ where $\mathbf{X}[t]$ is a $+t$ hour forecast for the state of the atmosphere and $\mathbf{X}_t[t]$ is the true state of the atmosphere at the forecast time. However, the true state of the atmosphere is unknown. We use independent reference observations to estimate

$$\mathrm{RMS}[t] \quad = \quad \sqrt{\frac{1}{m}\sum_{i=1}^{m}(y_{r,i} - \mathbf{H}_{r,i}\mathbf{X}[t])^2} \tag{3}$$

where m is the number of independent reference observations, $\mathbf{H}_{r,i}$ is the forward operator for reference observation i, $\mathbf{X}[t]$ is a $+t$ hour forecast for the state of the atmosphere which is valid for the observation time. We assume further that the reference observations contain the same amount of information about each model state variable. It is then reasonable to expect that experiments with a small RMS also will have a small MSE.

Only forecasts that start at $00Z$ hours are used in the calculation of the RMS, and only reference observations at $00Z$, $06Z$, $12Z$ and $18Z$ are used in the verification. Only one parameter can be verified at a time (say the MSLP). The configuration that scores lowest in the MSE verification (using the same independent information), is considered to be better. In the figures, "m-o" indicates "model" minus "reference observations". Also, the vertical error figures show the +6, +24 and +48.

The atmospheric governing equations have chaotic properties that become visible if one applies a neutral change to the forecasting system, for instance by assimilating observations that do not improve nor degrade the forecast quality. In this case there will be changes from day to day, sometimes the original system scores better and sometimes it scores worse compared to the modified system. The reason for this is that the governing equations cause insignificant perturbations to grow rapidly, so that two equally probable states of the atmosphere may evolve into two very different weather situations, where one will score better than the other. We expect that our two systems will have comparable scores if one averages over a long enough period (we have then implicitly assumed that the "tails" of the probability distribution for the random perturbations are small). Much of the motivation for choosing a rather long verification period (4 months) was to increase the chances of having significant verification results.

The significance of any MSE comparison (using the null-hypothesis that two systems have equal quality) is difficult to determine since we do not know how much independent information we are using in our verification statistics. We may raise the question if the MSE verification always converges (as long as we verify over a long enough time-period). This is an interesting theoretical question, but we will not develop this discussion any further here.

We may wish to acquire an impression of how the MSE verification score over the 4 months accumulate. A fruitful approach is then to look at the contribution to the "area under the daily $\mathrm{RMS}[t]$ curve",

$$\phi = \sqrt{\sum_{t=6}^{48}(\mathrm{RMS}[t])^2}.$$

If we plot ϕ as a function of the day in the verification period for both the reference and our experiment, we can identify how any difference in final verification score was accumulated. If

there are just a few events with a large difference we may argue that the difference in verification score is less significant since if, say, the verification had stopped short of one of these events, the MSE verification score would have been very different.

Note that for the first couple of days in the period, there may only be short forecasts available, and the daily RMS will then tend to be lower. The opposite may be true for the last couple of days in the period, where only long forecasts may be available, and the daily RMS will tend to be higher.

List of Figures

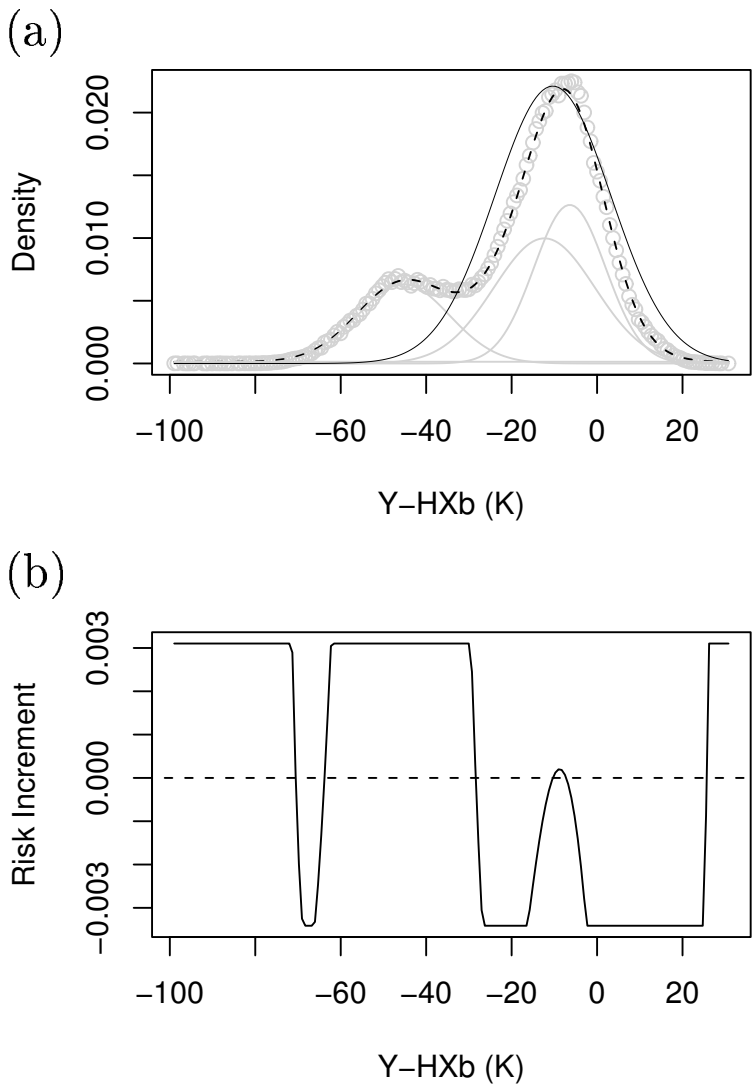

Figure 1: (a) Noaa16, AMSU-A, channel 2 (31.4 GHz) innovation probability distribution (grey circles), the estimated non-normal innovation distribution (dashed), the estimated normal components (grey) and the innovation in the optimal normal Approximation. (b) Bayes risk increment for noaa16, AMSU-A, channel 2.

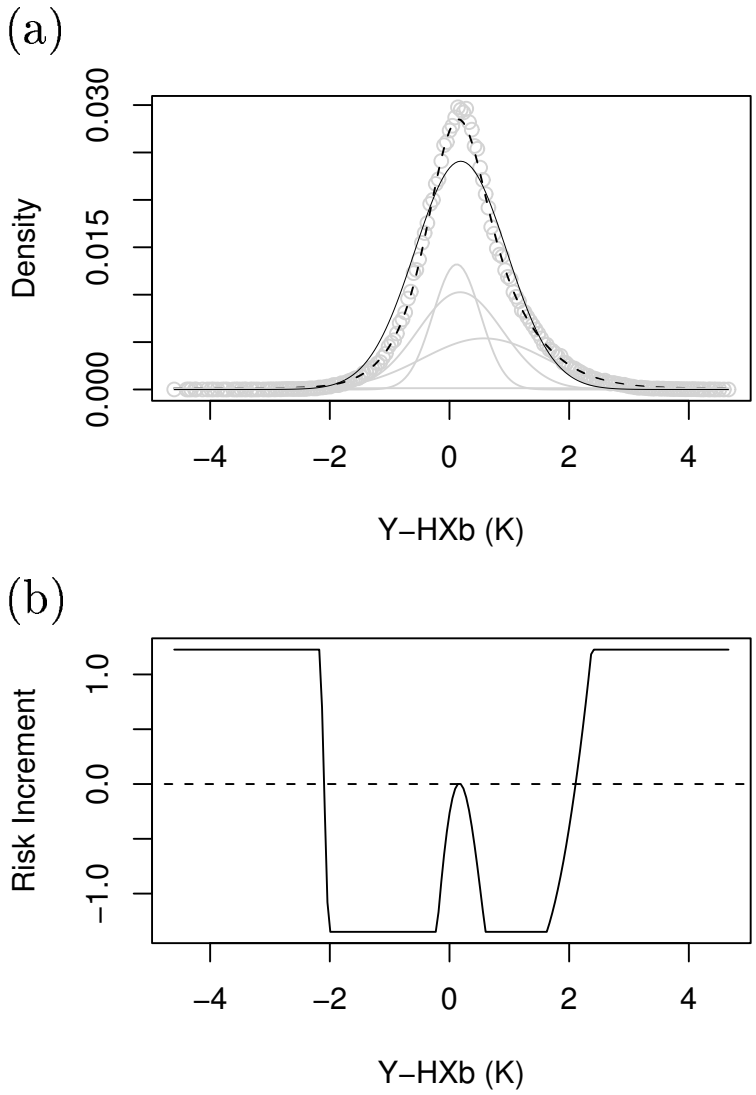

Figure 2: (a)Noaa16, AMSU-A, channel 5 (53.6 GHz) innovation probability distribution (grey circles), the estimated non-normal innovation distribution (dashed), the estimated normal components (grey) and the innovation in the optimal normal Approximation. (b) Bayes risk increment for noaa16, AMSU-A, channel 2.

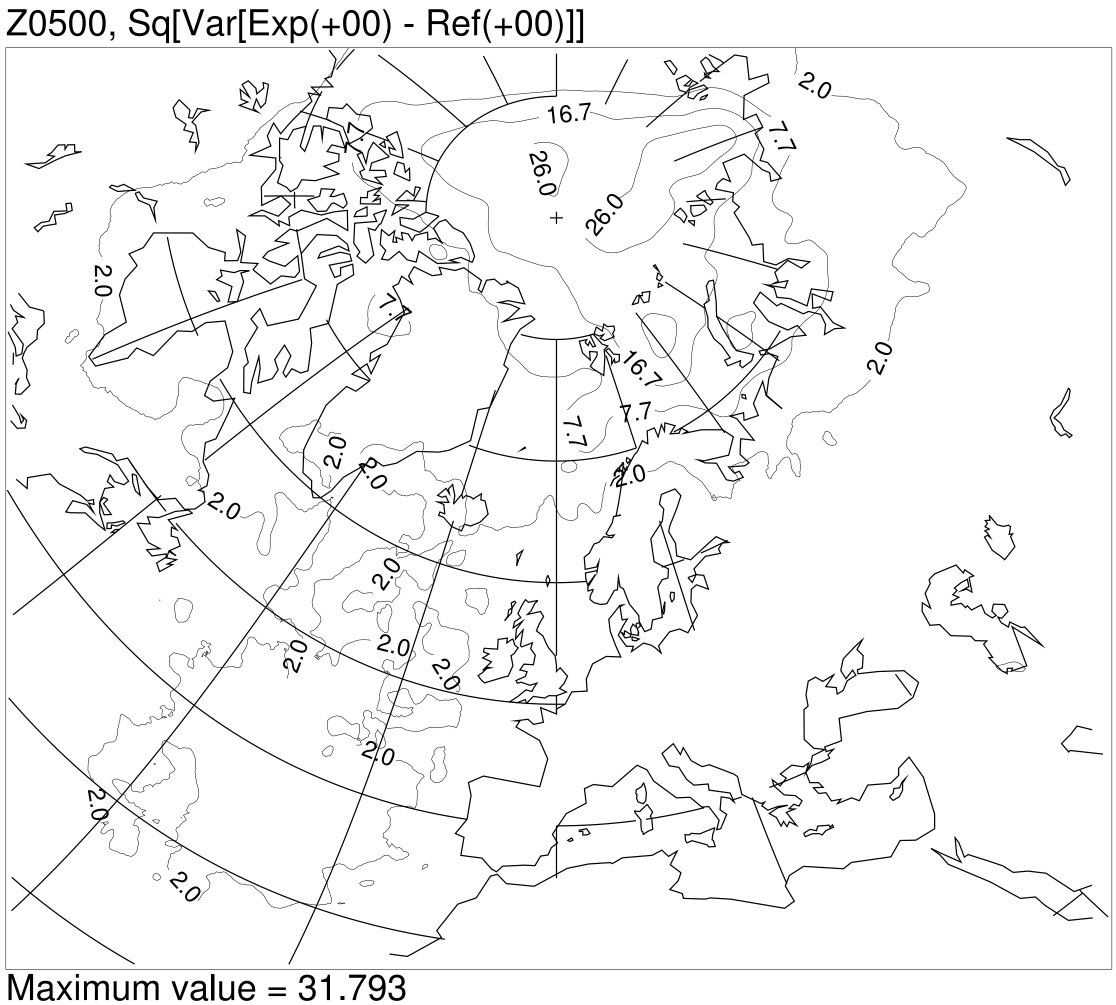

Figure 3: Geographical distribution of the standard deviation in the $500hPa$ geopotential height (Z500) analysis between the experiment and reference trials.

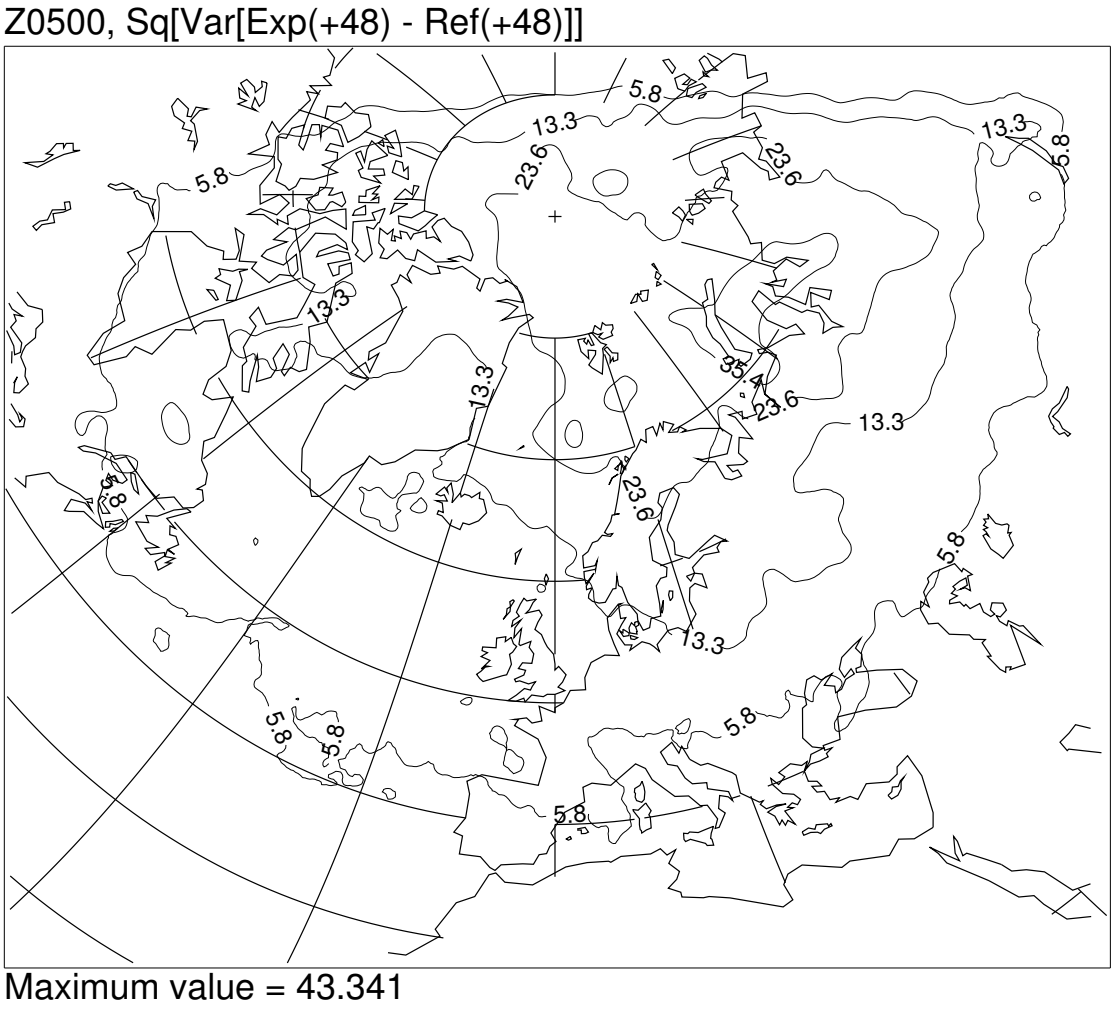

Figure 4: Geographical distribution of the standard deviation in the $500hPa$ geopotential height (Z500) 48-hr forecast between the experiment and reference trials.

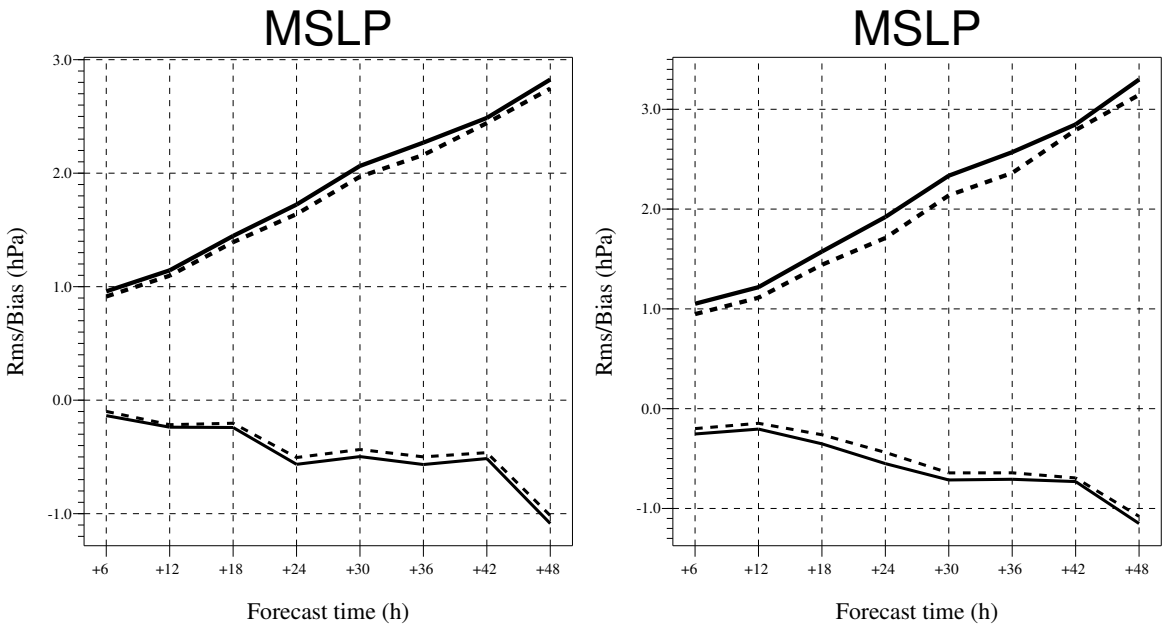

Figure 5: Comparison of the standard deviation (Std) and bias (Bias) in MSLP between independent observations and two trials, one with AMSU-A observations over sea ice (dashed) and one without (solid). Left: all EWGLAM stations. Right: Nordic subset of the EWGLAM stations

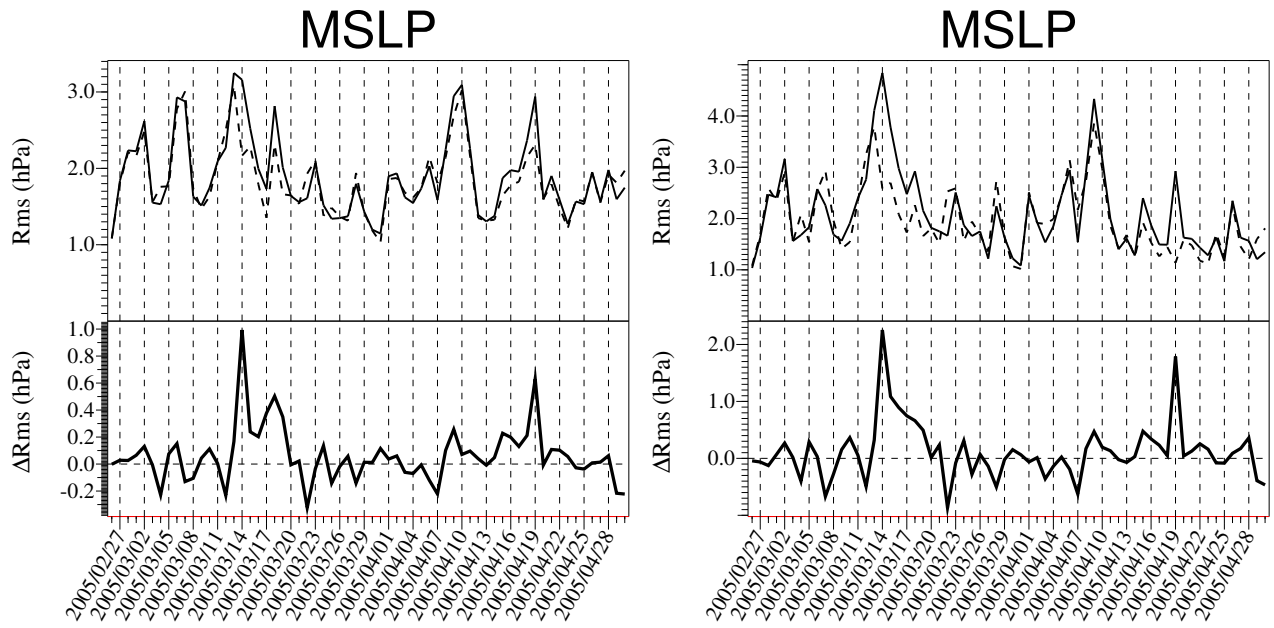

Figure 6: Daily contribution to the MSLP RMSE verification for the two trials, one with AMSU-A observations over sea ice (dashed) and one without (solid). The upper panels show daily RMSE for the two trials. The lower panels show the difference between the two trials (a positive ΔRms indicates that the AMSU-A observations had a positive effect on the verification results). Left: All EWGLAM stations. Right: Nordic subset of the EWGLAM stations

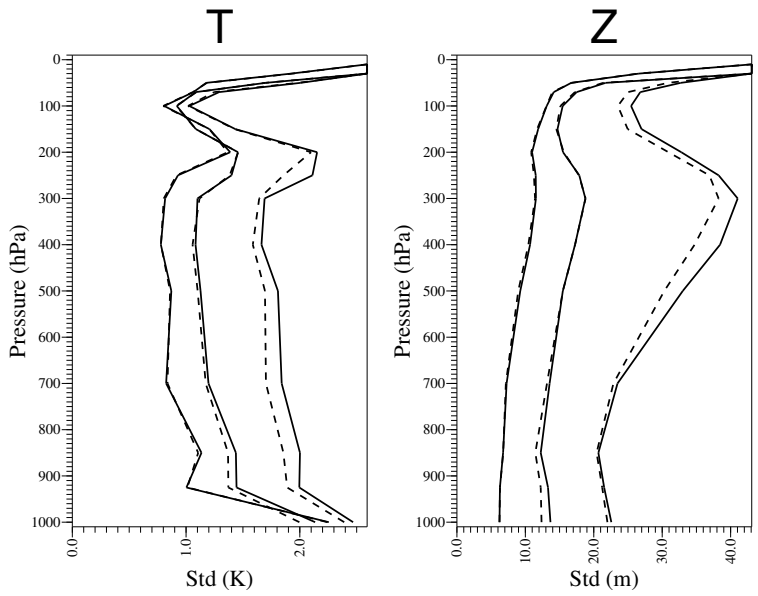

Figure 7: Comparison of the vertical profile of the standard deviation between independent reference observations and forecasts (6-hr, 24-hr and 48-hr) from two trials, one with AMSU-A observations over sea ice (dashed, experiment trial) and one without (solid, reference trial).

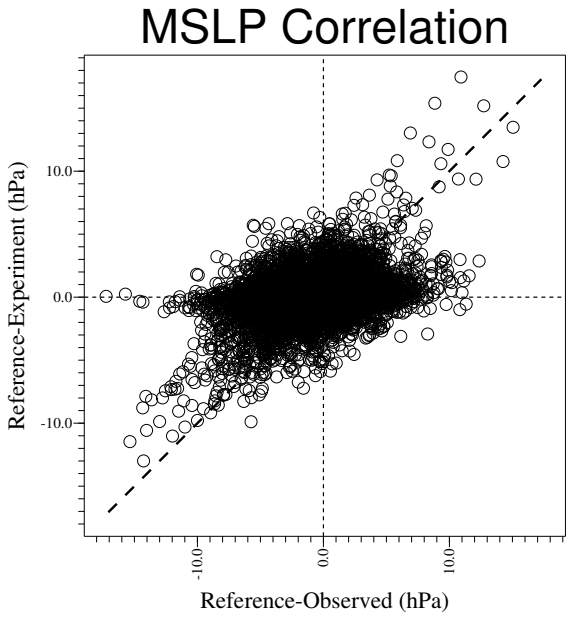

Figure 8: Correlation in MSLP between reference trial (no AMSU-A over sea ice) forecast minus corresponding experiment trial (AMSU-A over sea ice) forecast and reference trial forecast - observed reference observations.

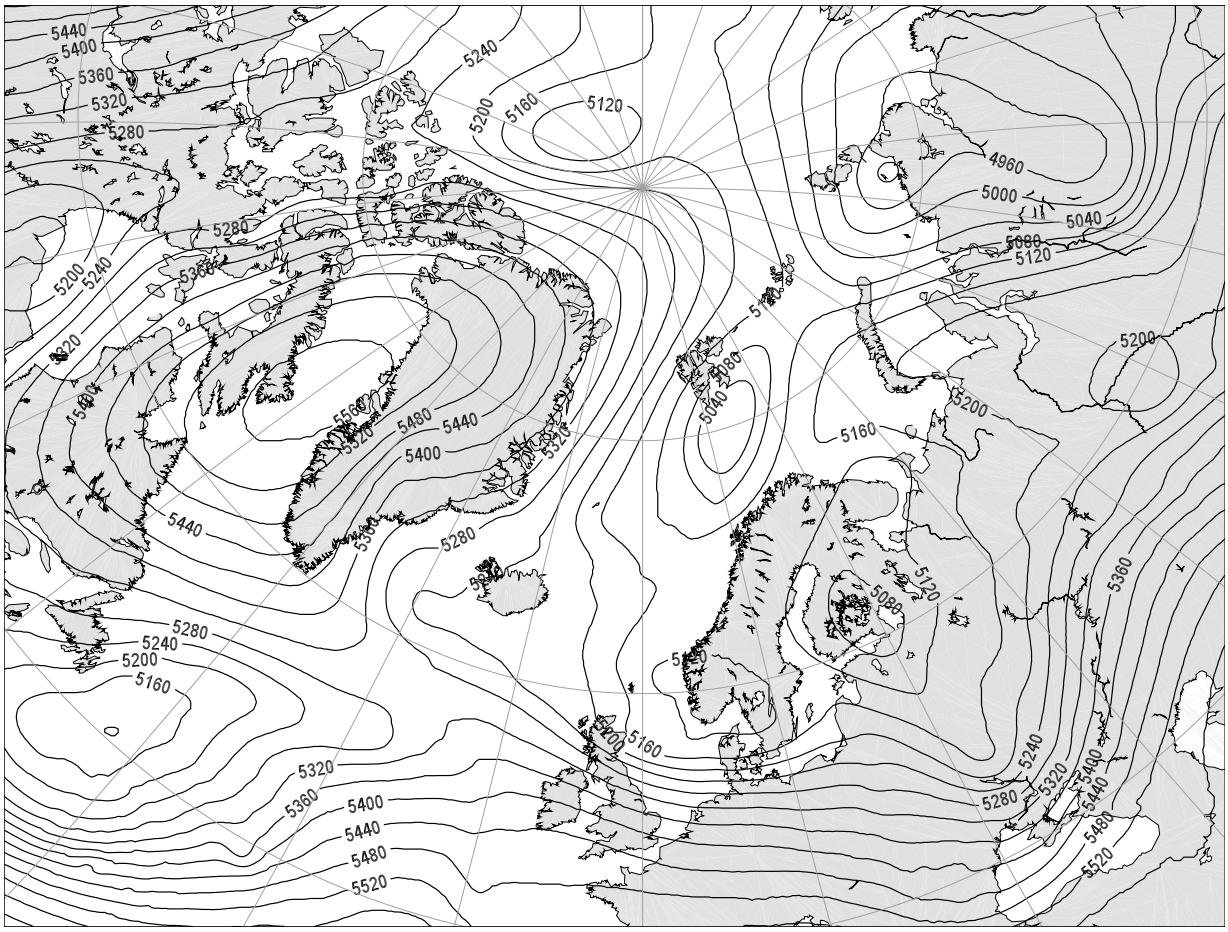

Figure 9: 500 hPa surface height analysis for the reference run valid 12 March 00Z.

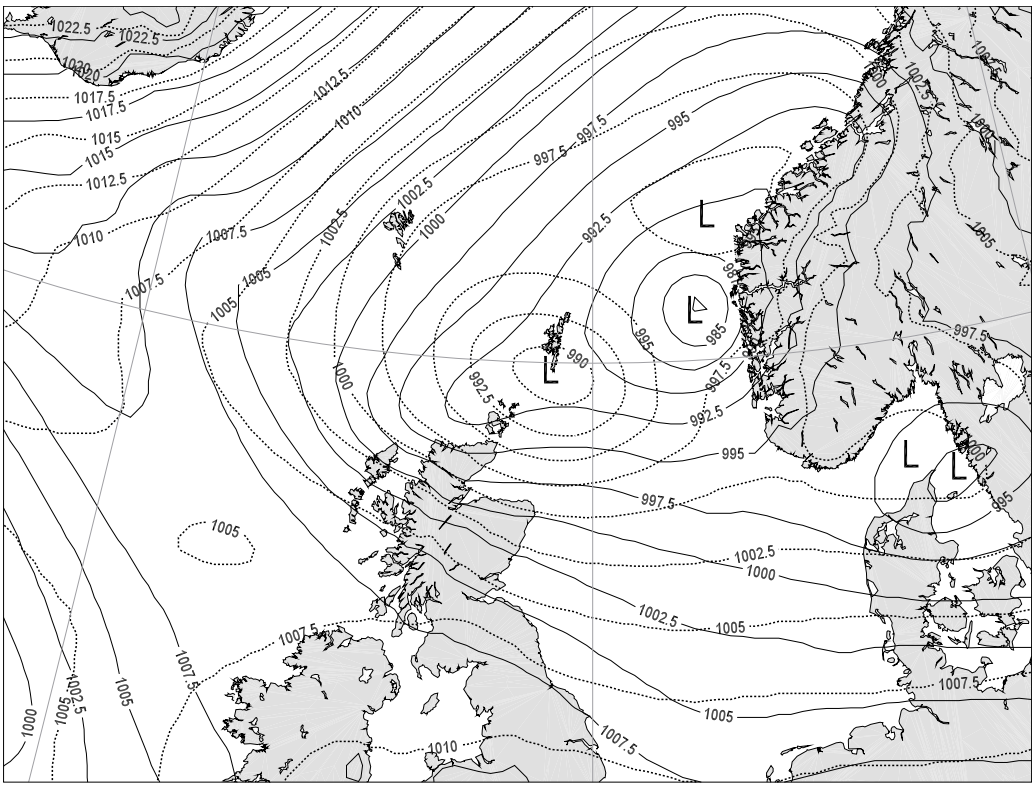

Figure 10: Mean sea level pressure (MSLP) 48 hrs forecast from the reference run, valid 14 March 00Z (solid lines). Corresponding MSLP analysis from the reference run in dashed lines.

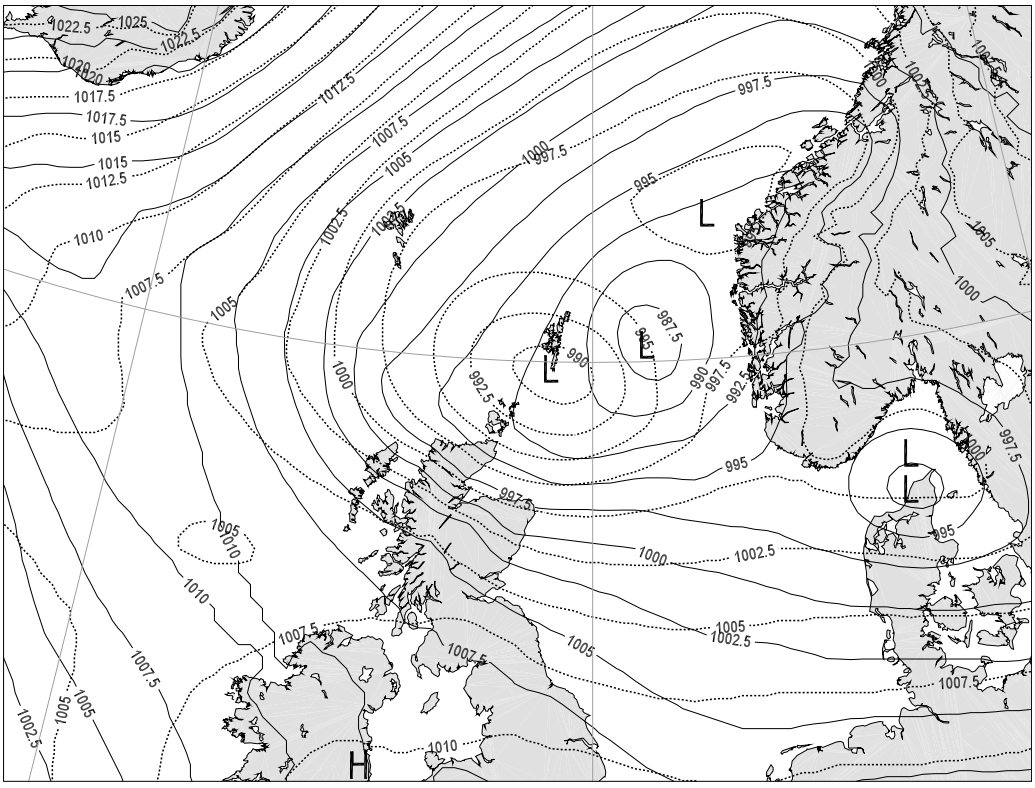

Figure 11: Mean sea level pressure 48 hrs forecast from the experiement run, valid 14 March 00Z (solid lines). Corresponding MSLP analysis from the reference run in dashed lines.

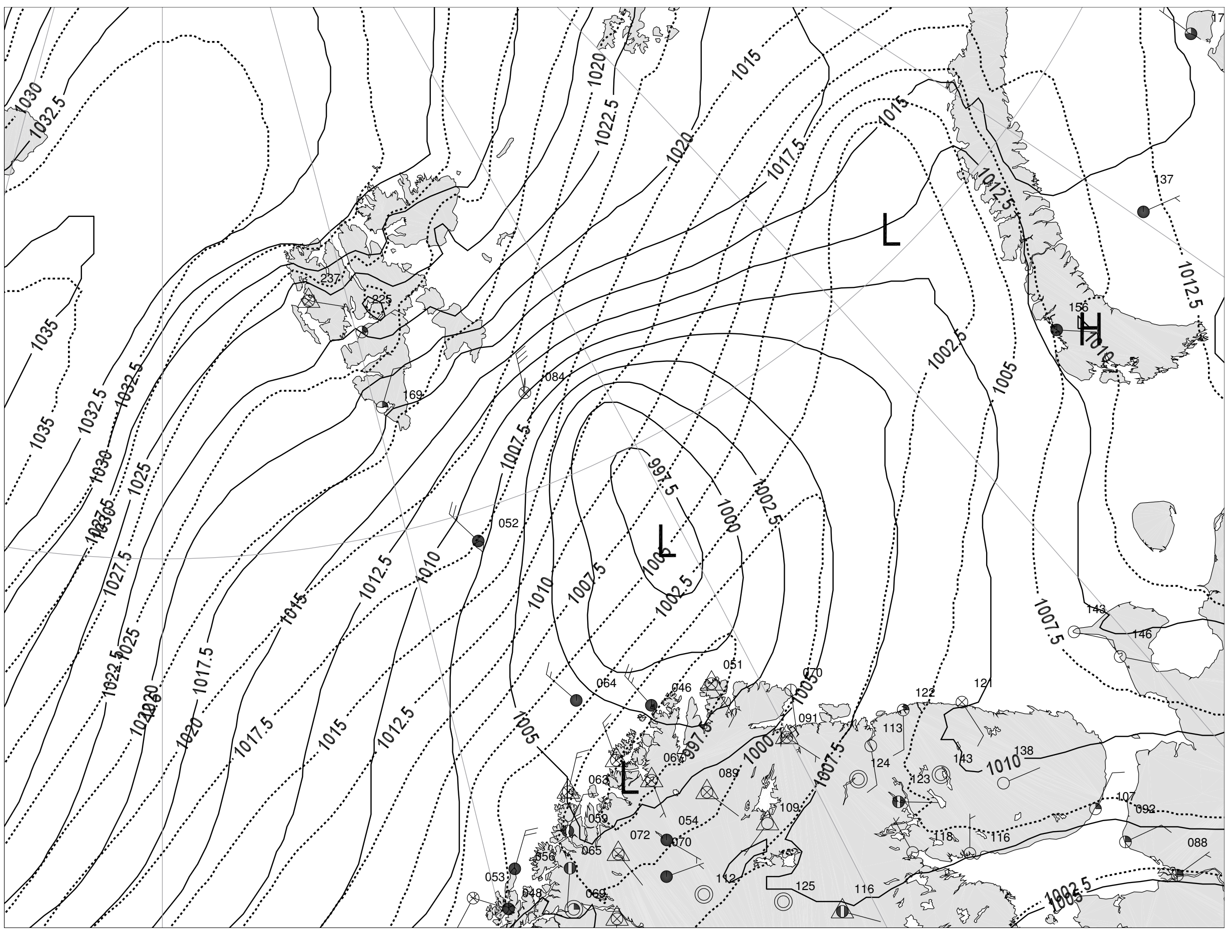

Figure 12: MSLP 24 hrs forecasts valid 14 March 00Z. The experiment run in solid lines, the reference run with dashed lines. Some SYNOP observations taken at the time are also shown (following the convention for weather maps, the numbers are the 2 digits before the one after the decimal point in the pressure hPa value).

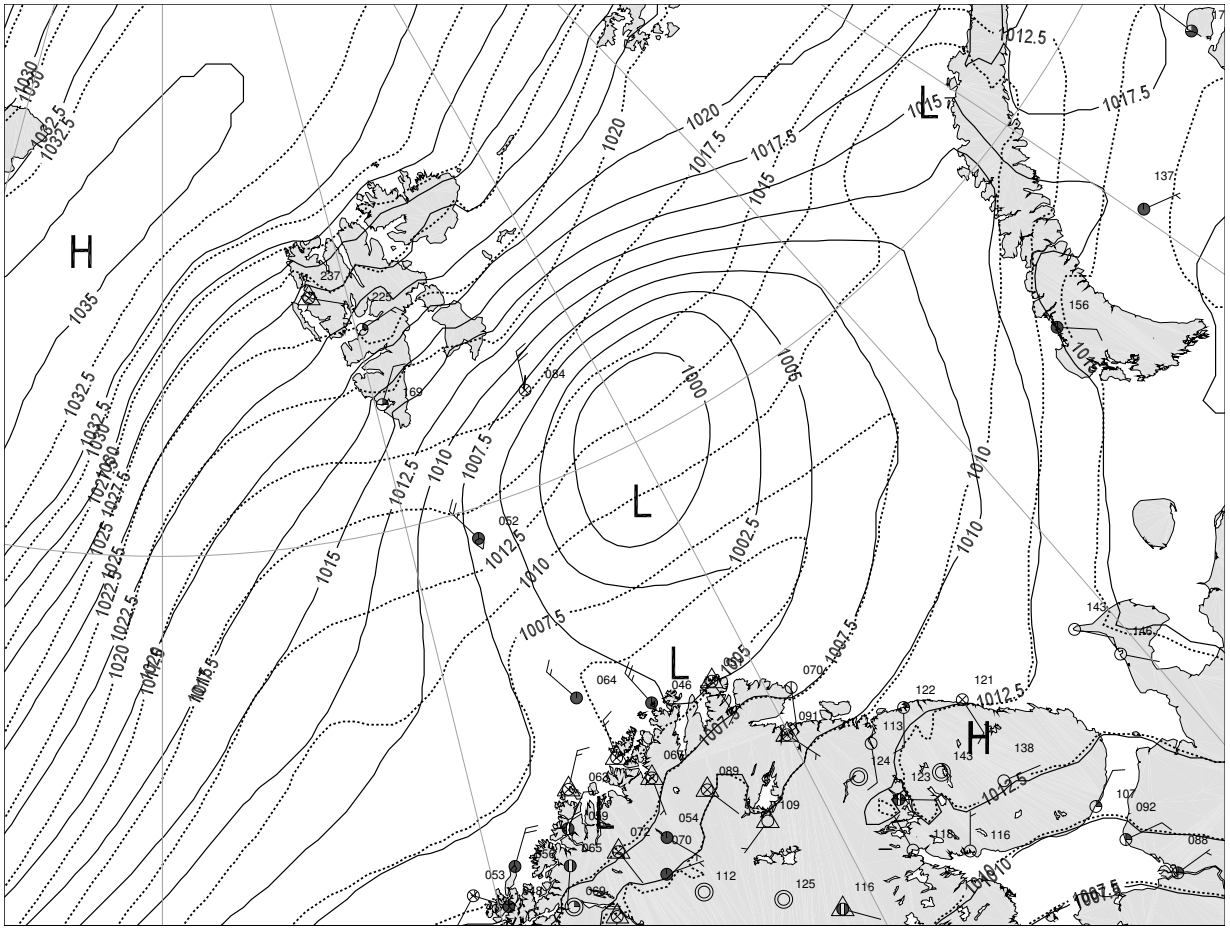

Figure 13: MSLP analyses valid 14 March 00Z. The experiment run in solid lines, the reference run with dashed lines. Some SYNOP observations taken at the time are also shown.

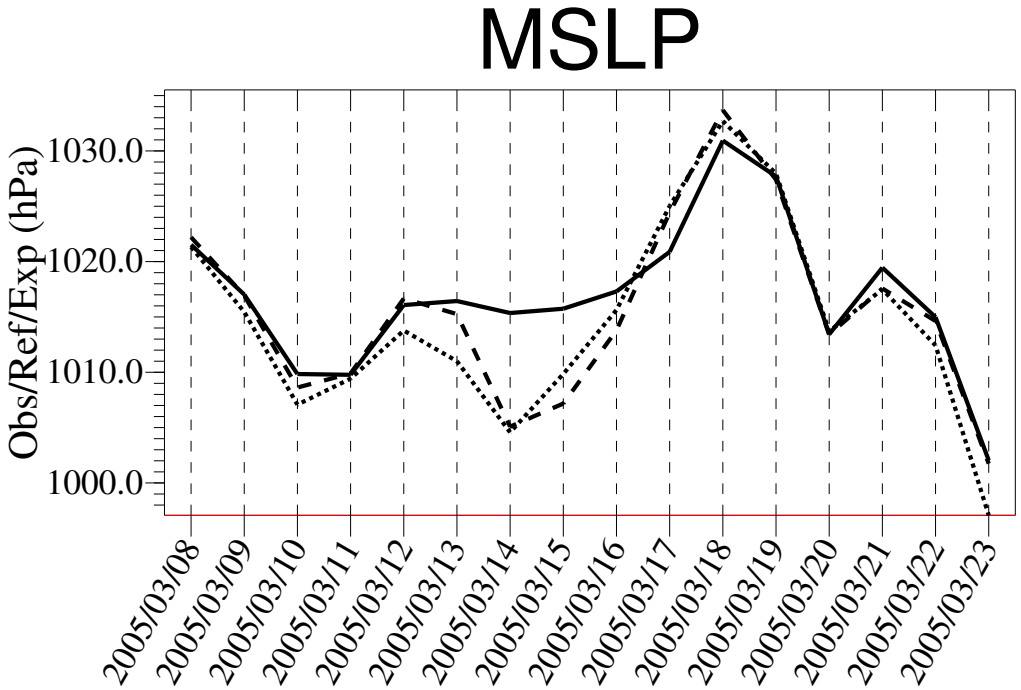

Figure 14: MSLP time series for SYNOP station Bjørnøya.The experiment run in dashed line, the reference run with solid line, SYNOP observation in dotted line.

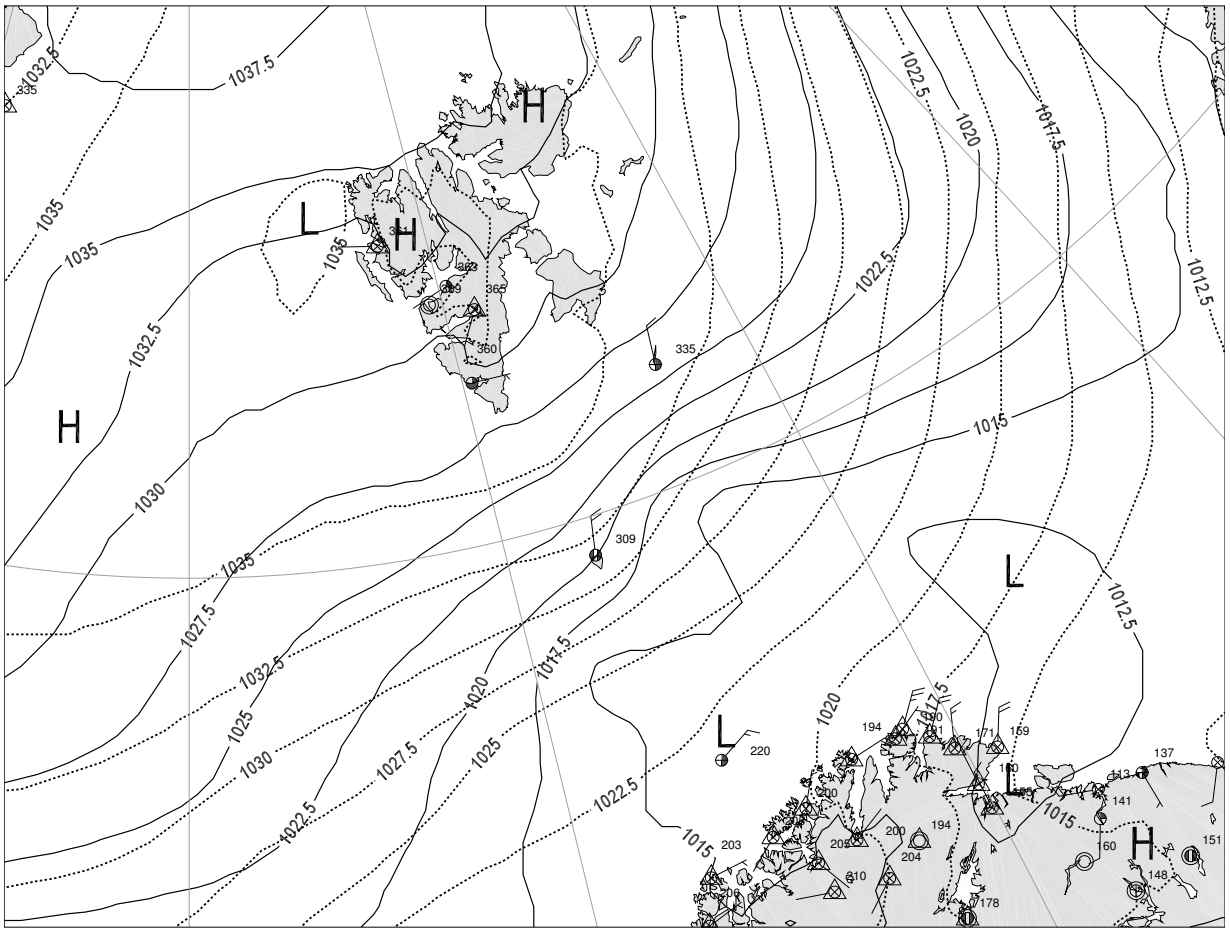

Figure 15: MSLP 48 hrs forecast from the reference run, valid 18 March 00Z (solid lines). Corresponding MSLP analysis from the reference run in dashed lines. Some SYNOP observations taken at the time are also shown.

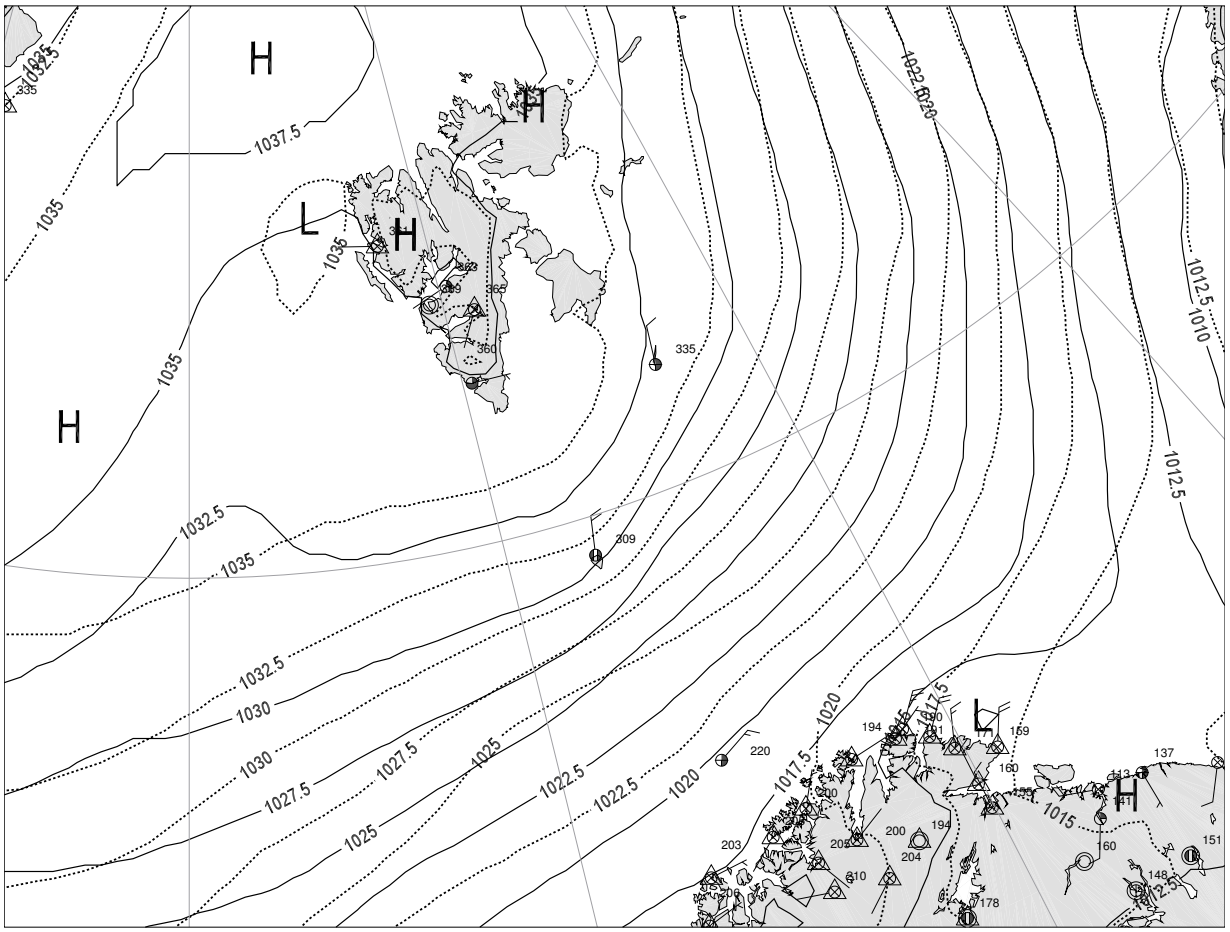

Figure 16: MSLP 48 hrs forecast from the reference run, valid 18 March 00Z (solid lines). Corresponding MSLP analysis from the reference run in dashed lines.Some SYNOP observations taken at the time are also shown.

Table 1: Emissivity values for the AMSU-A instrument.

Ch	ϵ_{FY}	ϵ_{MY}
1	0.971	0.874
2	0.970	0.829
3	0.928	0.796
4	0.928	0.796
5	0.928	0.796
6	0.928	0.796
7	0.928	0.796
8	0.928	0.796
9	0.928	0.796
10	0.928	0.796
11	0.928	0.796
12	0.928	0.796
13	0.928	0.796
14	0.928	0.796
15	0.913	0.744

Table 2: Data quality control parameters (in K) for Noaa15 AMSU-A used in the presented impact study.

Ch	Frequency[GHz]	min $(Y - \mathbf{H}\mathbf{X}_b)$	max $(Y - \mathbf{H}\mathbf{X}_b)$	μ	R_{diagonal}
1	23.8	-22.65	23.66	-8.74	400000
2	31.4	-26.68	21.83	-11.92	400000
3	50.3	-36.14	15.56	-7.29	400000
4	52.6	-7.38	6.28	-0.93	400000
5	53.6	-1.89	2.59	0.19	400000
6	54.4	-1.93	1.05	-0.39	0.35
7	54.9	-1.07	1.73	0.33	0.35
8	55.5	-1.25	1.74	0.39	0.70
9	57.3	-1.90	1.58	-0.02	14.0
10	57.3	-2.95	1.87	-0.26	300.0

Table 3: Data quality control parameters (in K) for Noaa16 AMSU-A used in the presented impact study.

Ch	Frequency[GHz]	min $(Y - \mathbf{H}\mathbf{X}_b)$	max $(Y - \mathbf{H}\mathbf{X}_b)$	μ	R_{diagonal}
1	23.8	-20.73	26.06	-6.42	400000
2	31.4	-28.90	25.86	-10.35	400000
3	50.3	-13.65	15.16	-5.15	400000
4	52.6	-5.64	6.30	-0.82	400000
5	53.6	-2.11	2.11	0.19	400000
6	54.4	-1.42	0.80	-0.26	0.35
7	54.9	-1.33	1.74	0.27	0.35
8	55.5	-1.58	2.18	0.36	0.70
9	57.3	-2.24	1.81	-0.11	14.0
10	57.3	-3.23	2.54	-0.34	300.0