Report for IOMASA Deliverable 2.1.1:

Pre-operational data stream and preparation for assimilation

Per Dahlgren, Tomas Landelius
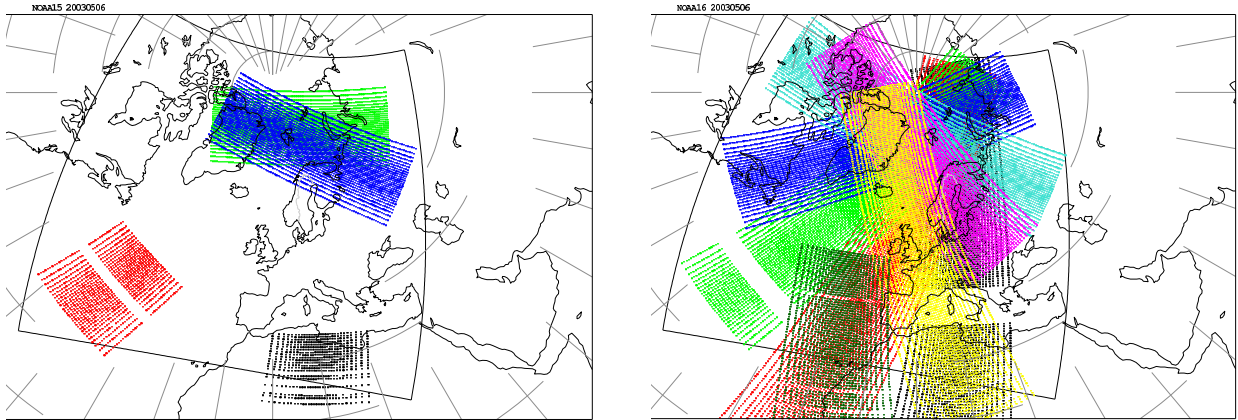
SMHI

Harald Schyberg

met.no

Contents

Figure 1: Data coverage of 20030506 for noaa15 (left) and 16 (right). Different colours shows different orbits

1 Introduction

This report describes the preparatory work in IOMASA for setting up a routine data stream for developing use of remote sensing data in numerical weather prediction (NWP) over Arctic regions. The work is undertaken at the Swedish Meteorological and Hydrological Institute (SMHI) and the Norwegian Meteorological Institute (met.no).

SMHI will work on AMSU-B moisture sounding data and the use of ice cover in the HIRLAM boundary layer scheme. Met.no will work with assimilation of AMSU-A temperature sounding data.

2 SMHI data processing

2.1 Humidity, AMSU-B

AMSU-A and AMSU-B data is received at SMHI via the EARS service (EUMETSAT ATOVS Retransmission Service) in level 1c format. The data thereafter has to be converted to the input format that is recognised by HIRLAM (i.e. BUFR). This is at SMHI done i two steps:

1. The level 1c files are converted to ASCII format. l1c → ASCII

2. The ASCII information is read and packed into BUFR format. ASCII → BUFR

The BUFR files are then transfered to the computational facilities, i.e. the computers where HIRLAM is run, via ftp.

2.1.1 Data coverage

At present, the EARS sends data received at Maspalomas (Canary Islands), Tromsø in northern Norway and Kangerlussuaq in Greenland.
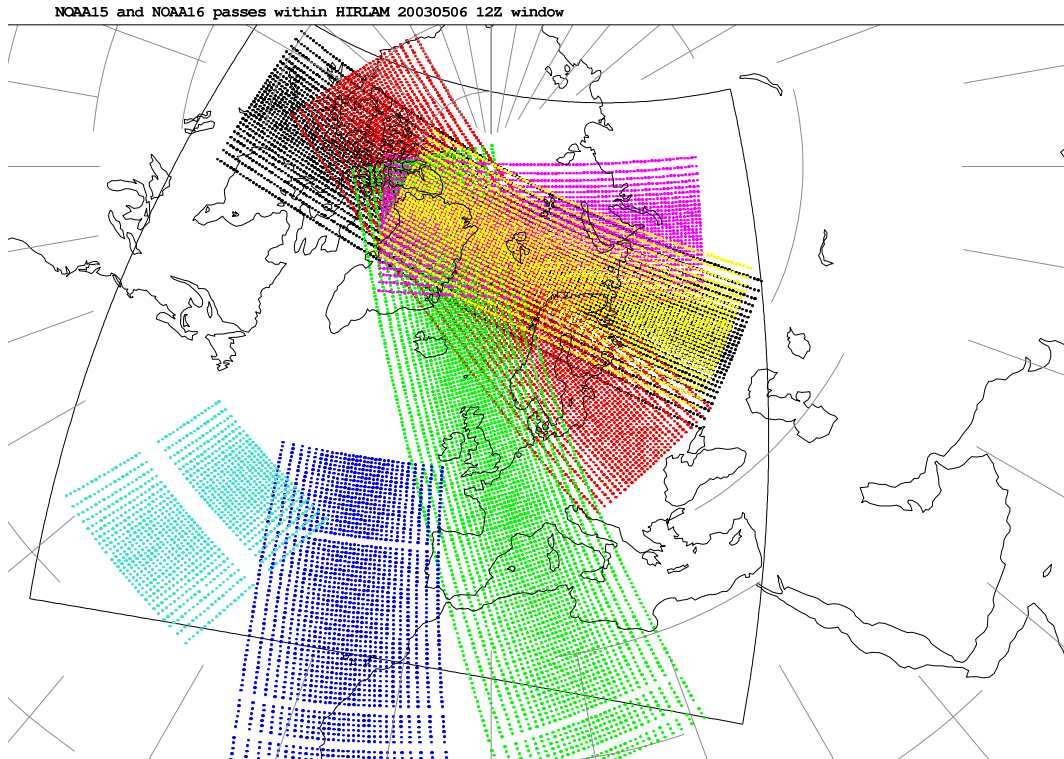
Figure 2: Data available for the HIRLAM forecast at 20030506 12:00 UTC. Different colours shows different orbits

Figures 1 and 2 are examples of the data coverage during 24 hours, and how much data that can be used as input to a specific HIRLAM forecast. Figure 2: the HIRLAM 12:00 Z window means 09:00 UTC → 15:00 UTC.

2.2 Ice concentration, OSI SAF

There exist already estimated ice concentrations, available from *www.dnmi.no*, with which we want to replace the existing climatological ice-fractions in the present HIRLAM-system. They are available on a polar-stereographic grid, with a resolution of 10 km. This analysis is incorporated in the HIRLAM surface analysis (Rodríguez et.al., 2003). To make a flexible system, easy to adopt to any HIRLAM setup, we have chosen to take the following steps:

- Download ice concentration field (630x450), given as a GRIB file from *www.dnmi.no*

- Reduce the information, by creating superobses on the scale of the current HIRLAM setup.

- Perform an analysis using the method of successive corrections.

- After the ice analysis, do not let the sea surface temperature exceed 0° C, if we have more than 5% sea ice in a grid square.

2.2.1 Creation of superobses

The creation of superobses, is done in the following way:

- First we estimate how dense the superobses should be, i.e. if HIRLAM has a resolution of about 40 km, it means that the superobses should also be at that scale.

- We step through the icefield on that resolution and store all points that has at least one point within a specified radius (R_1 = the largest influence radius, see below), that is not pure water or land.

- For those selected points we compute the "centre of gravity" for all (no-land) points of the original icefield, and the mean value of the ice fraction in these points (which position is converted to latitude and longitude), which thus defines the superobses on the desired scale. In a case of R_1 = 160 km and 40 km resolution in HIRLAM, the number of superobses is about 3300.

2.2.2 Analysis with successive corrections

This is a simple method, where the observation weights are defined as a function of distance between grid point and observation (r):

$$w(r) = \begin{cases} \frac{R^2 - r^2}{R^2 + r^2} & r \leq R \\ 0 & r > R \end{cases}$$

The weights are calculated iteratively for different values of R, here (40 km HIRLAM) we have used R_1 = 160 km and R_2 = 80 km.

The analysis at grid point k, at iteration $j + 1$, a_k^{j+1}, is given by:

$$a_k^{j+1} = a_k^j + \frac{\sum_{i=1}^n w(r_{ik})\sigma_o^{-2}(y_i - a_i^j)}{\sum_{i=1}^n w(r_{ik})\sigma_o^{-2} + \sigma_b^{-2}}$$

The summations are done over the super-observation points (i), and $y_i - a_i^j$ is the departure from the background. σ_o and σ_b are the observed and background standard deviations, respectively. During the first iteration the background field is given by the first guess relaxed to a climatological field.

References:

Rodríguez, E., Navascués, B., Ayuso, J.J., Järvenoja, S., 2003: Analysis of surface variables and parameterisation of surface processes in HIRLAM. Part I: Approach and verification by parallel runs. *Hirlam Techn. Rep., No.58*
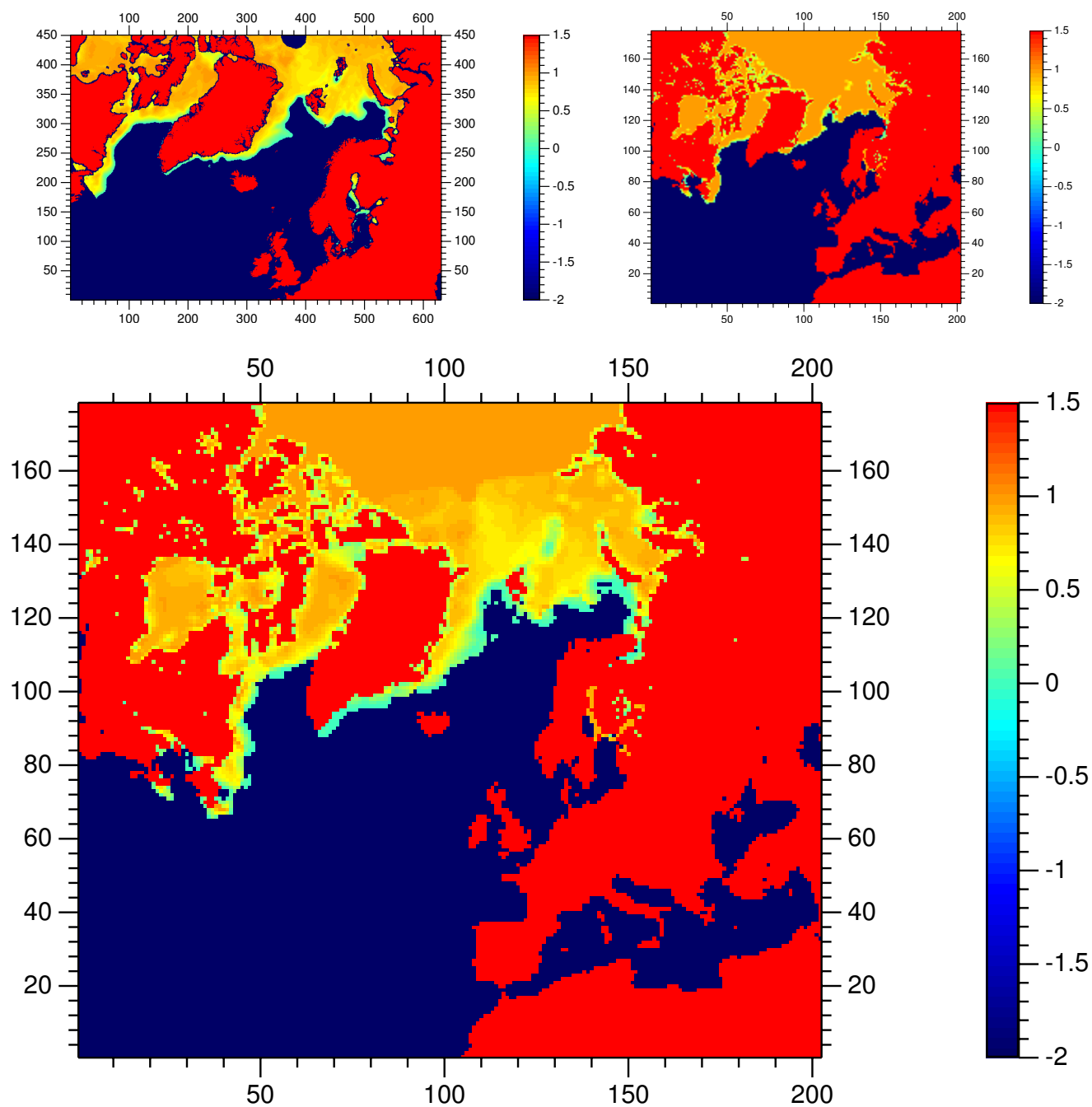
Figure 3: Top left: The icefield as given by SAF 2003 03 24 12. Top right: The corresponding icefield as given by original HIRLAM analysis. Bottom: The corresponding icefield as given by new HIRLAM analysis.

3 Met.no data processing

The preparatory work at met.no has consisted of two tasks. Firstly, in planning IOMASA it was realized that NWP model data of a selected set of weather parameters would be of great help in interpreting SSM/I, QuikSCAT and other remote sensing data over the Arctic. This could for instance help build a better understanding of how the ice surface emissivities change, and two years of NWP output (2003-2004) will be provided in support of the tasks undertaken by the other partners in the project.

Secondly, as a preparation for assimilating AMSU-A data over ice, it is necessary to collect statistics for the AMSU observations versus the NWP model. This is done by co-locating the observations with NWP model data valid at the same place and time. Ice information will also be co-located as auxiliary data. A near-real-time processing chain for producing this information is described below and will be set up.

3.1 Storage of model fields (two year dataset)

3.1.1 NWP models and data availability

For its operational forecasting, met.no presently runs several regional NWP models for short-range (up to $2\frac{1}{2}$ days) forecasting. These models are versions of the HIRLAM model run on 3 different domains at 20, 10 and 5 km horizontal resolution. This is supplemented with numerical forecasts from the European centre ECMWF up to 10 days range. The regional models use boundary values from the ECMWF global model.

For Arctic applications as in this project, the 20 km model is the best one to use because it covers large parts of the Arctic. For IOMASA it is suggested to store the following fields

- near-surface temperature

- near-surface wind

- total (vertically integrated) cloud water

- total (vertically integrated) water vapour

- time-accumulated precipitation

In general the best and most consistent meteorological analysis fields available is obtainable from the ECMWF reanalysis project which the ECMWF member states have access to. These data are produced off-line covering a long time period in the past in a consistent fashion, for instance is the model version kept the same in the production.

These data are not so easily accessible, so in this project we have chosen to use HIRLAM data. At met.no data from the HIRLAM operational models are stored at three levels: First all model data output are stored on easily accessible "operational" disks for 2-3 days, then, secondly, a selected subset of data is stored for a few months on a rolling "short-term" disk archive. Finally all data are stored permanently (or more precisely until the decay of the tape) on a tape robot archive, from where it is more cumbersome to extract the data. It turns out that not all of the data required for IOMASA are available on the met.no short-term-archive, so the

best procedure is to set up a daily routine to extract the necessary data from the "operational" disks. Some reprocessing will be necessary to extract data from the start of 2003 until the start of the daily storage routine.

The HIRLAM 20 model has been operational at met.no since 20 March 2003. It has, however, been run in experimental mode before that, so data are archived at the met.no tape archive since December 2002. The model domain of HIRLAM 20 is 468 by 378 points on a rotated spherical grid with approximately 20 km horizontal mesh width, and this domain is shown in Fig. 4. New forecasts are started every 6 hours.

3.1.2 Notes on use of the data

The forecast quality of the HIRLAM 20 model is documented to be of significantly improved quality compared to earlier versions of HIRLAM in use at met.no, even if it is still somewhat behind the quality of the ECMWF model at the same forecast ranges.

We will also mention here two effects which in some cases might affect the data quality. All operational NWP models have so called "spin-up" problems related to the forcing of the model by observations at initial time. This particularly affects variables related to the water cycle, for instance clouds, moisture and precipitation. To reduce the effect of such problems, we have selected to use 6-hours forecasts instead of initial analyses to provide the fields. Typically a 6 hours forecast has a quality which has not degraded much relative to an analysis. There will, however, still be a spin-up problem for the precipitation data provided, which is accumulated precipitation for the last six hours, that is from analysis time up to 6 hour forecast time.

In a limited area model such as HIRLAM one should also be aware of the treatment of the lateral boundaries which are forced by the ECMWF model. The 8 outermost points at the boundaries are forced to have the same values as in the ECMWF model, so the information here does not actually come from HIRLAM. Then there is a border relaxation zone which extends 10 grid points inwards of that. In that zone a combination of information from HIRLAM itself and the forcing ECMWF model is used. Sometimes spurious phenomena will appear in this relaxation zone due to inconsistencies in the model formulations of the inner and outer model.

3.1.3 Documentation of storage format

GRIB (gridded binary format) is a quasi-standard format for storing fields at operational weather centres. GRIB coding and decoding code is therefore widespread, and documentation as well as routines for dealing with the format are freely available on the world wide web. Detailed information about GRIB and decoding software can be found at several places on Internet, for instance the sites below.

A guide to GRIB:

```
http://www-imk.fzk.de/imk2/kasima/aktuelles/grib/index.html
```

GRIB library of UCAR:

```
http://dss.ucar.edu/libraries/grib/
```

At this site, we find the following routine, which have been in use for unpacking:

```
http://www.dss.ucar.edu/libraries/grib/c_routines/unpackgrib.c
```

Information on reading GRIB files and some code:

```
http://wesley.wwb.noaa.gov/reading_grib.html
```

A NASA page with links to documentation and read software:

```
http://polar.gsfc.nasa.gov/data_stuff/formatPages/GRIB.html
```

ECMWF GRIB library:

```
http://www.ecmwf.int/products/data/software/
```

WMO GRIB documentation:

```
http://www.wmo.ch/web/www/WDM/Guides/Guide-binary-2.html
```

A Norwegian page with some info:

```
http://noserc.dnmi.no/formkonv/index.html
```

3.1.4 File content

Note that what is stored are 6 hrs forecasts, so on the GRIB files the "reference time" (start of forecast) is 6 hrs before the time for the data. The forecast lead time (always 6hrs) is stored in a time range parameter on the file. We will provide one file for each day (containing the fields at 4 termins at 6 hours intervals). Table 1 shows the definitions of the stored fields and values of various corresponding GRIB parameters.

3.1.5 Coordinate system

The HIRLAM 20 model grid is defined in terms of rotated spherical coordinates. The conversion to and from geographical latitude and longitude requires a "g" array which defines the grid. The definition of the grid and values of various parameters are given in Table 2.

Coordinate conversions can be done with the following FORTRAN code, converting between lon and lat positions in radians contained in n-dimensional arrays x and y and grid positions contained in the same arrays. The conversions makes use of a subroutine called "sph2rot.f" (enclosed in the appendix), which converts between geographical lat/lon and rotated spherical lat/lon.

```
C The rotated spherical grid for HIRLAM20 is defined by
C a g array as follows:
      data g /-46.5,-36.5,  0.2,  0.2,  0., 65./
      pi180 = 3.141592654/180.

C Convert input from degrees to radians:
      xwr = g(1)*pi180
      ysr = g(2)*pi180
```

Table 1: Fields contained in output files

Quantity	wmo	typ	lev	par	scal	field	unit
2m temperature	11	105	2.	0	1.	't2 '	(K)
x wind component	33	105	10.	0	1.	'u10 '	(m/s)
y wind component	34	105	10.	0	1.	'v10 '	(m/s)
Cloud water path	76	200	0.	0	1.	'cwpath'	(kg/(m*m))
Water vapour path	54	200	0.	0	1.	'wvpath'	(kg/(m*m))
6 hrs accumulated precipitation	61	105	0.	0	1.	'prt '	(mm)
- where:							
wmo = wmo table number							
typ = level type							
lev = level value							
par = 1 if field is a parameter (constant) field							
scal = scaling factor							
field = field name (character*6)							
(- for more detailed explanation, see GRIB documentation)							

Table 2: Parameters describing the rotated spherical grid

grid	nx	ny	g(1)	g(2)	g(3)	g(4)	g(5)	g(6)	Description
2001	468	378	-46.5	-36.5	0.2	0.2	0.	65.	0.20 deg rot (hirlam)
Array for definition of spherical (rotated) grid.									
grid = grid number									
nx,ny = horizontal dimensions									
g = grid definition parameters									
spherical grids: g(i)=lon1,lat1,dlon,dlat,xcen,ycen									

```
    dx = g(3)*pi180
    dy = g(4)*pi180
    xcen= gr(5)*pi180
    ycen= gr(6)*pi180


C Spherical rotated coordinates from lat-lon:
    icall= 1
    CALL sph2rot(icall,n,x,y,xcen,ycen,ierror)
    do j = 1, n
        x(j) = (x(j) - xwr)/dx + 1.
        y(j) = (y(j) - ysr)/dy + 1.
    enddo
```

```
C Lat-lon from spherical rotated
     do j = 1,n
        x(j) = xwr + (x(j)-1.)*dx
        y(j) = ysr + (y(j)-1.)*dy
     enddo
     icall= -1
     CALL sph2rot(icall,n,x,y,xcen,ycen,ierror)
```

Rotating velocities to North and east components: The x and y wind components are defined as components along the rotated spherical grid axes. For conversion to east and north component a similar transform must be performed.

The code mentioned here and more detailed documentation will be provided electronically upon request.

3.2 Processing for comparison of NWP model data and AMSU data

Co-location files containing AMSU observations, HIRLAM profiles and surface data at AMSU footprints as well as ice information from a processing chain based on the OSI SAF (see below) will be produced routinely as new AMSU observations become available in near-real-time. HIRLAM 20 short-range forecasts and analyses will be interpolated in time and space to to observation time and location of each observation.

The collection of a dataset of co-location files over a time period is the first step towards development and improvement of assimilation of AMSU data over ice, because it allows building statistics of model deviations and to assess the data quality. It will also be a useful dataset for testing of improved emissivity models for the sounding channels, because it will include ice information which can be used to assist in emissivity estimation. The data flow in the co-location chain in development is shown in Fig. 5.

Met.no has set up reception facilities for the EARS retransmission service. Routines for converting the AMSU-A data to BUFR format via ASCII (in the same way as done at SMHI as described above) have been prepared. Met.no also has a local reception antenna which has been the sole source of near-real-time ATOVS data up to now. The EARS retransmission now means much better near-real-time coverage of ATOVS over Arctic, and will be of great value in this project. An impression of the coverage was shown in the SMHI section.

We will as mentioned above also use ice data from an experimental extended ice processing chain as input for extracting ice data at the AMSU footprints. This chain will be based on that in use in the EUMETSAT Ocean and Sea Ice (OSI) Satellite Application Facility (SAF) hosted by DMI and met.no. The ice data here comes from SSM/I observations and in the future possibly also from QuikSCAT and AVHRR. Initially ice concentration and multi-year ice concentration will be provided on the AMSU footprints. Later on more information of interest for emissivity calculations in the SAF grid may be provided and co-located with the AMSU observations.

A Appendix: Subroutine for conversion between latitude/longitude and rotated spherical coordinates

This FORTRAN subroutine performing the necessary transformations will be provided upon request. The routine converts between arrays with any number "n" of lat/lon pairs in the rotated HIRLAM grid and real geographical lat/lon values. A file with this routine can be delivered upon request. The value of the parameter "icall" determines the direction of the transformation.

```
      subroutine sph2rot(icall,n,x,y,xcen,ycen,ierror)
c
c   conversion between spherical (x=xsph,y=ysph) and spherical rotated
c   (x=xrot,y=yrot) coordinates. (xcen,ycen) is the position of the
c   rotated equator/greenwich in terms of (longitude,latitude).
c   all values are given in radians.
c
      integer icall, n, ierror
      real    x(n), y(n), xcen, ycen
c
      ierror = 0
c
      zsycen = sin(ycen)
      zcycen = cos(ycen)
c
      if (icall.eq.1) then
c
c   compute spherical rotated coordinates as function of
c   spherical coordinates
c
      do j = 1,n
         xsph = x(j)
         ysph = y(j)
         zxmxc  = xsph - xcen
         zsxmxc = sin(zxmxc)
         zcxmxc = cos(zxmxc)
         zsysph = sin(ysph)
         zcysph = cos(ysph)
         zsyrot = zcycen*zsysph - zsycen*zcysph*zcxmxc
         zsyrot = max(zsyrot,-1.0)
         zsyrot = min(zsyrot,+1.0)
         yrot   = asin(zsyrot)
         zcyrot = cos(yrot)
         zcxrot = (zcycen*zcysph*zcxmxc +
```

```fortran
     +                zsycen*zsysph)/zcyrot
          zcxrot = max(zcxrot,-1.0)
          zcxrot = min(zcxrot,+1.0)
          zsxrot = zcysph*zsxmxc/zcyrot
          xrot   = acos(zcxrot)
          if (zsxrot.lt.0.0) xrot = -xrot
          x(j) = xrot
          y(j) = yrot
       enddo
c
       elseif (icall.eq.-1) then
c
c  compute spherical coordinates as function of
c  spherical rotated coordinates
c
       do j = 1,n
          xrot = x(j)
          yrot = y(j)
          zsxrot = sin(xrot)
          zcxrot = cos(xrot)
          zsyrot = sin(yrot)
          zcyrot = cos(yrot)
          zsysph = zcycen*zsyrot + zsycen*zcyrot*zcxrot
          zsysph = max(zsysph,-1.0)
          zsysph = min(zsysph,+1.0)
          ysph   = asin(zsysph)
          zcysph = cos(ysph)
          zcxmxc = (zcycen*zcyrot*zcxrot -
     +               zsycen*zsyrot)/zcysph
          zcxmxc = max(zcxmxc,-1.0)
          zcxmxc = min(zcxmxc,+1.0)
          zsxmxc = zcyrot*zsxrot/zcysph
          zxmxc  = acos(zcxmxc)
          if (zsxmxc.lt.0.0) zxmxc = -zxmxc
          xsph = zxmxc + xcen
          x(j) = xsph
          y(j) = ysph
       enddo
c
       else
c
       ierror=1
c
       endif
```
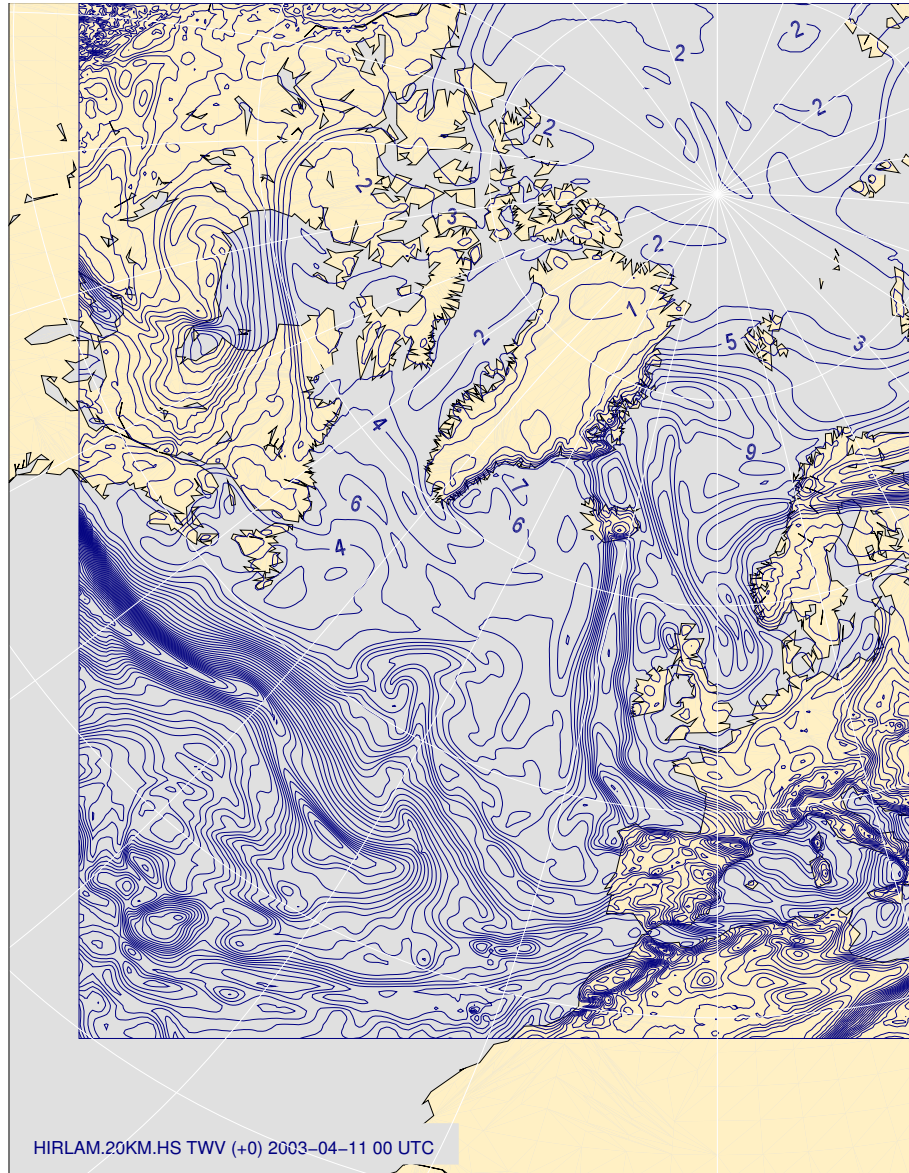
```
c
      return
      end
```

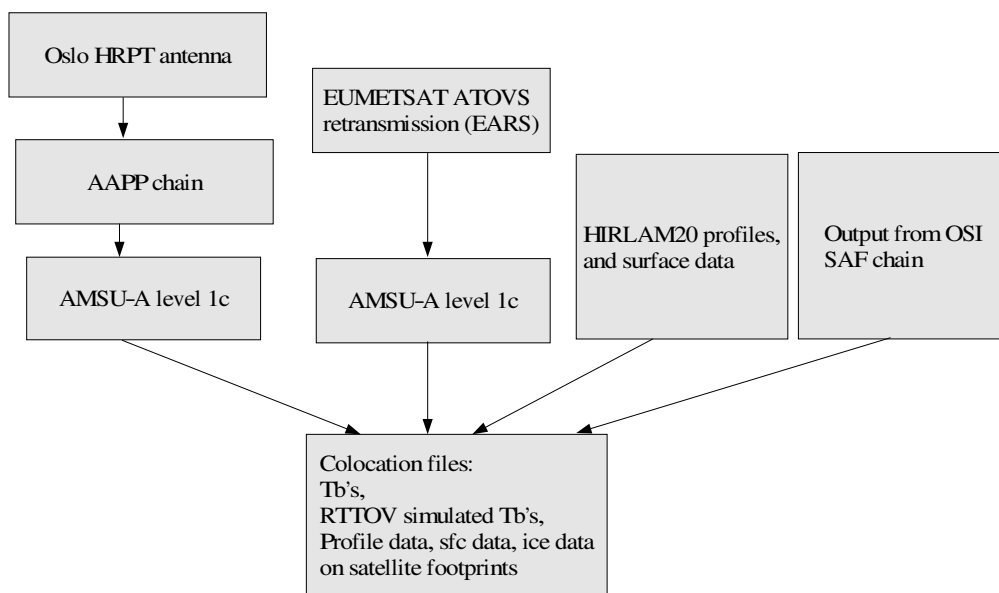Figure 4: The met.no HIRLAM 20 area (with a total water vapour field plotted for illustration purpose).

Figure 5: Setup of the co-location process.