ACCORD visit: Overview and Outcome

Time: 09.05.2022 – 27.05.2022

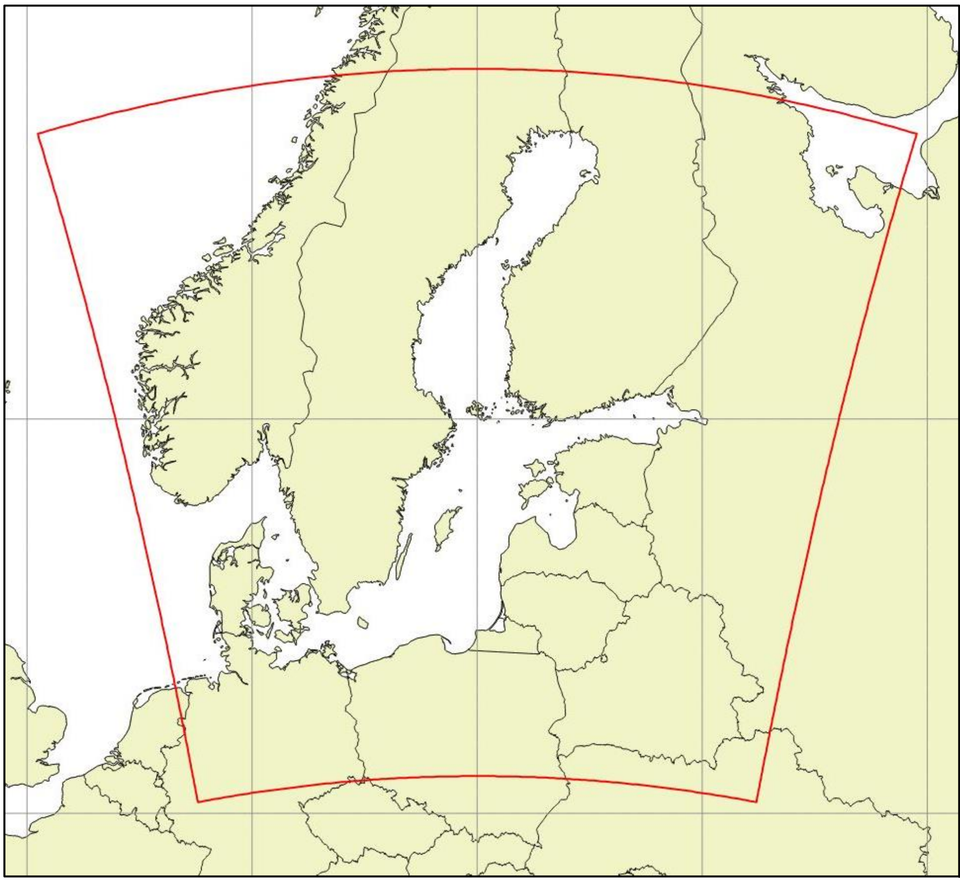
Location: Norwegian Meteorological Institute, Oslo

Visitor: Alina Lerner, Estonian Environmental Agency (ESTEA)

Hosts: Roger Randriamampianina, MET Norway

Background

The purpose of this visit was to proceed the learning of data assimilation in Harmonie-Arome system. I have had the visit to MET back in October-November'21, whereas the main activity was dedicated to learning of data assimilation basics, exploring Harmonie system structure at ECMWF computing resources, and derivation of structure functions for the domain chosen in ESTEA. By the May visit, the B matrix for HMEST25_2 domain (see Figure 1) were ready to use to conduct first data assimilation experiments.



```
'TSTEP'  => '75',
'NLON'   => '720',
'NLAT'   => '720',
'LONC'   => '20',
'LATC'   => '59',
'LON0'   => '20',
'LAT0'   => '59',
'GSIZE'  => '2500.',
'EZONE'  => '11',
```

***Figure 1**. Computational domain HMEST25_2 and its parameters*

Description of the visit

The work was divided into two parts. First, conduction of the experiments to ingest different observations to the data assimilation system for the HMEST25_2 domain. Second, the computation of the B matrix for the new high-resolution Estonian domain EST_HR (Figure 2).

1 .Data assimilation experiments

In the beginning, I tried to make a Harmonie cycle 46 setup with data assimilation for HMEST25_2 domain on *ecgate-cca*. I had one winter month and one summer month difference files and B matrices. The background error statistics implementation method was adapted from CARRA reanalysis setup. This method got B of a day through interpolation between January and July statistics to the date chosen. The MetCoOp observation files of 7.05.2022 and 8.05.2022 and variation bias correction files were retrieved from the archive on Cirrus and used to conduct the data ingestion experiments. The observation types included were: SYNOP, TEMP, AMDAR, SCATT, GNSS ZTD, Mode-S, Radar.

These experiments kept on failing in Bator_conv task. The failure happened in bator_lecture at line 304: *ABORT! 1 ERR : bator_lecobsoul0 : OBsoul incorrect OBSOUL.conv*. After some attempts, we found out the problem lies behind the wrong format of the OBSOUL.gpssol file containing GNSS ZTD observations. The content of the OBSOUL file produced by GNSStoOBSOUL in my system looked as follows:

```
20220507  3

 17 1  110    52.14000   -4.57000  'ABEPMETO'  20220507  30000  134.0 1 11111 0 128  15.0 0.12000E-01 2.39400     -3290

 17 1  110    52.79000   -4.74000  'ADARMETO'  20220507  30000   94.0 1 11111 0 128  15.0 0.12000E-01 2.40200     -3740

 17 1  110    52.15000    1.60000  'ALDBMETO'  20220507  30000   12.0 1 11111 0 128  15.0 0.12000E-01 2.47900     -1850

…
```

Correctly formatted OBSOUL file (provided by Frank Guillaume) should look as follows:

```
   20220507        3

 17  1       110  52.14000   -4.57000  'ABEPMETO'  20220507  30000  1.34000E+02     1    11111       0  128  1.50000E+01
0.12000E-01  2.39400E+00      -3290

 17  1       110  52.79000   -4.74000  'ADARMETO'  20220507  30000  9.40000E+01     1    11111       0  128  1.50000E+01
0.12000E-01  2.40200E+00      -3740

 17  1     110 52.15000   1.60000 'ALDBMETO' 20220507 30000 1.20000E+01    1  11111     0 128 1.50000E+01 0.12000E-
01  2.47900E+00      -1850

…
```

Ideally, I would need to add some changes in GNSStoOBSOUL script to make it produce the well-formatted OBSOUL.gpssol file. As we had a shortage of time (and already spent too many days understanding the error before we asked Frank Guillaume), it was decided to switch to the cycle 43 to eliminate the problem. In addition, for the first time I switched from *ecgate-cca* to Atos.

I have set up the data assimilation experiment on *aa* using the usual average B matrix. There OBSOUL.gpssol had the right format, but I needed to put these observations to Passive, since

I did not have a whitelist for GNSS ZTD, yet. I followed the instructions by Jana Sanchez Arriola from AEMET to do that.

The script scr/GNSStoOBSOUL uses:

INFILE=OBSOUL.gnss

OUTFILE=OBSOUL.gpssol

and runs the package pregpssol to create the output file. The routine pregpssol.F90 calls read_list_gpssol.F90 to read the whitelist. I had two options to avoid this step:

a) to write OBSOUL.gnss in OBSOUL.gpssol format and avoid to run GNSStoOBSOUL in Oulan;

b) to change pregpssol routines commenting out the call to read_list_gpssol.F90 from pregsssol.F90.

I did the first option. Then, the following lines were added to the blacklist src/blacklist/hirlam_blacklist.b.conv_data_selection:

```
if (OBSTYP = synop) then
 ( ...)
 if (VARIAB = apdss) then
 fail(EXPERIMENTAL);
 endif;
endif;
```

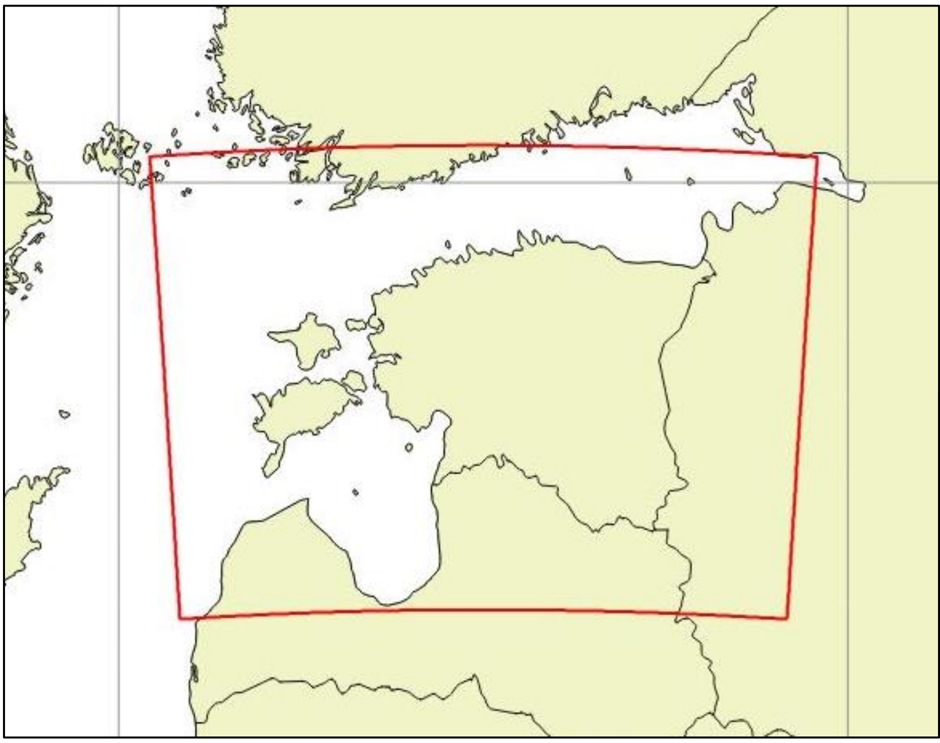
The next problem, which is still not resolved, is connected to the ingestion of the radar observations. I kept on getting the error in the Screening task. I was advised to leave only one country's observations and raise the number of processors used. I left only Sweden and raised nprocx to 16 and nprocy to 20 in Env_submit. It did not help. I will keep on trying to solve the problem.

2. Setting up EST_HR and B computation

The new high-resolution Estonian domain EST_HR has been set up for future work on creating the Harmonie-based nowcasting system for Estonia. The choice is made based on the idea to include a lot of sea to potentially use Seviri observations. I-zone is enhanced from default 8 points to 40 points in order to provide a smoother transition between the values inside the EST_HR (500 m resolution) and lateral boundary conditions derived from IFS (9 km resolution). 90 vertical levels of Météo-France are used in the model setup (MF_90).

The two-step EDA method is used to derive the structure functions for EST_HR. The perturbations of four ensemble members' boundary conditions are used to provide the differences. Differences from 2 weeks of every season are combined in the end to compute the averaged B using Festat standalone.

Numerous crashes of Femars took place during the downscaling computation due to OOM problem. So, it is advisable to enhance the memory used in this task in Env_submit from the default 16000 MB to, for example, 24000 MB.



```
'TSTEP'  => '15',
'NLON'  => '1024',
'NLAT'   => '750',
'LONC'  => '25',
'LATC'   => '58.6',
'LON0'  => '25',
'LAT0'   => '58.9',
'GSIZE'  => '500.',
'EZONE' => '11',
```

Figure 2. *Computational domain EST_HR and its parameters*

As the result, the preliminary annually-averaged B matrix for EST_HR domain was derived from downscaling (1st step of derivation is completed).

Foreseen next steps

Further work should be done in terms of setting up the HMEST25_2 assimilation system. First, the problem with radar observations ingestion should be resolved – excluding the wind files and manipulating memory usage are the means to try first. Then, more observation types need to be included in the system in a passive state. For example, more satellite data (AMV, Seviri) and crowdsource data (Netatmo and/or SMAPS). Then, variational bias correction is required to be accomplished in order to produce the stations whitelist for HMEST25_2.

As was mentioned before, EST_HR domain is set up to develop NWP-based nowcasting with data assimilation for Estonia. It is going to be my PhD project to accomplish this task. This was suggested by Roger Randriamampianina to make a research regarding finding the best suitable background error statistics application in such conditions. I am going to consider annually-averaged B and B as a function of the season/day/hour. Potentially, when OOPS becomes available, the flow-dependent B is also a good thing to develop for a nowcasting system.

Another important part of setting up this system is assimilating as many as possible high-resolution observations with the right (sub-hour) timeliness. In addition to the conventional and satellite observations, we could use non-conventional ones. I have started the initiative to send more GNSS ZTD data to E-GVAP. At the moment only 4 stations data are supplied, while we have 25 more. Now Estonian and Swedish geodetic services are conducting tests on processing the GNSS raw data acquired by Estonian Land Board. Also, ESTEA is highly interested in the usage of the third parties' meteorological observations, which are made by some Estonian schools and universities, the State transport department, Tallinn's water utility, the Centre of Environmental Studies and others. This could help to densify the observation network, which is important for the high-resolution domain. As a matter of research, my supervisor at Tallinn University of Technology and I are interested in assimilating the cloud base height measured by ceilometers. I am going to contact Siebren de Haan, who is developing the method, and test it with the Estonian ceilometers' data.

Acknowledgements

I acknowledge the ACCORD consortium for giving the opportunity to visit the Met Norway. I thank the host Roger Randriamampianina for organizing the working conditions and supervising me through the process. I also would like to thank Per Dahlgren, Roel Stappers, Yurii Batrak and Máté Mile for comprehensive help.