

STSM Report: Algorithm operationalisation in BALTRAD: upgrade and back processing

Liesbeth Verlinden

9 - 13 January 2017

Abstract

This report gives an overview of my STSM at the Swedish Meteorological and Hydrological Institute from 9 till 13 January 2017. The original planning is given by the titles of each section, a summary of the actions taken around that topic are then written underneath.

This STSM is linked to the STSM of Matúš Tejiščák, who focused on the technical aspects, so for completeness please see also his STSM report and the documentation he wrote.

Contents

1	Merging data files	2
1.1	Merging scans	2
1.2	Merging volume files	3
1.3	Additional issues	3
1.4	Discussion	4
2	Upgrading vol2bird	4
3	Testing the production pipeline from Baltrad to the server in the Netherlands	4
4	Back processing the data of migration season of autumn 2016	4

1 Merging data files

We started the STSM by checking the raw data available on the Baltrad server and how the merging of the polar volume (pvol) files was done so far.

It is good to point out that countries send in their radar data in different formats: some send in polar volumes, while other countries send in scans, some send in all available quantities (DBZ, TH, VRAD, and so on), while others send only the quantities DBZ and VRAD in the same scan or polar volume, while still others send in DBZ and VRAD in separate scans or polar volumes. An overview of how the data arrives in Baltrad per country is given in Table 1.

Country	Pvol/scan	Quantities	Merged?
BE	pvol	DBZ, VRAD	x
CH	scan	DBZ	x
CZ	pvol	DBZ, VRAD	✓
DE	scan	all	x
DK	pvol	all	x
ES	pvol	DBZ	x
FI	scan	DBZ VRAD	✓
FR	scan	DBZ VRAD	x
HR	pvol	DBZ VRAD	✓
NL	pvol	DBZ VRAD	✓
NO	pvol	DBZ, VRAD	x
PL	pvol	DBZ, VRAD	x
PT	pvol	DBZ	x
SE	scan	DBZ, VRAD	✓
SI	pvol	DBZ, VRAD	✓
SK	pvol	DBZ, VRAD	✓
UK	scan	DBZ, VRAD	x

Table 1: *Overview of countries we are allowed to use the data from, how they send in data and if that was merged correctly on the Baltrad server upon start of the STSM. Quantities separated with a comma are sent in in separate polar volumes/scans, quantities given without a comma are sent in in the same polar volume/scan.*

On the Baltrad server scans are merged into polar volumes and polar volumes that contain DBZ and VRAD separately are also merged together. We found that there were issues with both merging processes.

1.1 Merging scans

The process of merging scans is done by two scripts. We found that on the Baltrad server they outputted incorrect data: sometimes wrong quantities were

	DBZ	VRAD
number of rays	720	360
length of gate (m)	250	125

Table 2: Overview of differences between Norwegian scans containing DBZ and VRAD.

merged, or dropped, or the wrong files were processed. For details about the failure, see the documentation at the end of Matúš’ STSM report. When the code was run on Matúš’ computer, it worked fine.

1.2 Merging volume files

Although Anders Henja had written code to merge the volume files that contain DBZ and VRAD separately, we found out that the code actually never was placed on the Baltrad server, but that merging was done by a script put there by other users. This script did merge the volume files, though not correctly, e.g. sometimes scans of the *same* quantity but taken at different time stamps were merged together.

We tested the merging code written by Anders on Matúš’ computer. Apart from some small adjustments needed, the code worked fine.

1.3 Additional issues

We also found that some countries send in irregular data with the DBZ and VRAD scans having mismatches in:

- resolution (number of rays/bins),
- elevation angle,
- time stamps.

Mismatch in resolution Countries sending in data with different resolutions are Norway and the UK. E.g. Norwegian DBZ scans have 720 rays per scan, while the VRAD scans only have 360 rays. Additionally for Norway we found that the number of gates per ray differs as well. DBZ rays have gates of 250m, while VRAD rays have gates of 125m. An overview is given in Table 2. We did not yet check if the UK data also has different gates per ray.

Mismatch in elevation angle and time The minimum difference in elevation angle between scans containing different quantities was 0.1° , though this could increase to 0.7° (Belgium).

Mismatches in timing ranged from almost 4 minutes (Poland) till 10 minutes (Norway).

1.4 Discussion

We inquired at Phil Stepanian if a mismatch in resolution of the number of rays could be solved via scaling down. According to him linear downsampling in azimuth and range of DBZH to match VRADH will work and is a perfectly acceptable technique. A proof of concept for this is written (see Matúš' STSM report for details), though unfortunately the time scope of the STSM did not allow to finish it.

To decide whether scans with different time stamp and elevation angle can be merged and still give valuable data requires expert knowledge. During this STSM, we did not have time to look further into this.

Matúš has written a wrapper code, adapted to the programs and dependencies available on the Baltrad server, that both merges scans and polar volumes and produces vol2bird profiles. This should enhance the performance of the production pipeline and facilitate the implementation of any possible further changes to the merging code, vol2bird and the production pipeline on Baltrad.

At time of writing this report is the wrapper code finished and sent to Baltrad, where it is awaiting testing and implementation.

2 Upgrading vol2bird

We found that the git repository of vol2bird was connected to a side branch on github and not the master branch, which had not been updated since June 2016. We changed the repository to the master branch and updated vol2bird.

3 Testing the production pipeline from Baltrad to the server in the Netherlands

The week before the STSM Günter had already provided us a new password to access the ftp server, which had solved the production pipeline problem. Checking of the downloaded files on the beehub-server during the STSM showed that these were copied correctly. However, given the issues with some of the raw data and merging we identified (see Section 1 on page 2), not all of the newly downloaded vol2bird output files on the server are correct.

4 Back processing the data of the migration season of autumn 2016

The raw data of autumn 2016 is stored on a supercomputer in Linköping. During the STSM Günther started to copy the data, around 2TB, to the Baltrad server. Due to access restrictions to the supercomputer this copying could only be done during working hours. As a consequence, by the end of the week only 2 months of data were copied on the Baltrad server and not yet available to us. Therefore,

Matúš set up a secure pipeline so that Günther could transfer the raw data to our own computers. The back processing of the data therefore will be done after this STSM.