Title:	OPERA5 – OA2 report: Long Time Series of Radar Data
Prepared by:	L. Méri, M. Jurašek, Ľ. Okon, J. Kaňák Slovak Hydrometeorological Institute (SHMÚ)
Category:	Report
Abstract:	This report assesses the feasibility of collecting and providing a long-term time series of archived radar data predating the operational introduction of Odyssey in 2011. The study focused on evaluating the applicability of the OPERA Data Information Model (ODIM) and its HDF5 implementation for long-term climatological analyses, particularly when incorporating data from legacy radar systems. A key demonstration involved generating a 5 km CAPPI-level map of the average duration of reflectivity exceeding 40 dBZ, highlighting areas with potential flash-flood risk. The findings confirm that the ODIM-H5 format is well suited for storing and processing long-term radar volume data in a unified framework. Standardizing the data format enables efficient reprocessing across heterogeneous sources and supports the development of consistent processing chains for diverse radar datasets.

1. Motivations and Aims

The OA2 working package developed in the framework of the EUMETNET OPERA 5 project was intended to perform a feasibility study to collect and provide a long time series of archived data from before the operational introduction of Odyssey (2011). The data collected should be either single site data or re-analysis of composites. The study was focused on testing and evaluation of the usability of the OPERA Data Information Model and its hdf5 implementation in the case of long-term climatological studies involving different types of data sources – including some older radar systems.

Climatological processing requires reliable and precise input data, spatial homogeneity, temporal homogeneity and long-term availability among other. Radar derived climatological studies commonly rely on some already available derived products and composites (e.g. QPE, hail detection, storm cell tracks [Saltikoff et al. 2019] [Tabary et al. 2012], [Marra and Morin 2015], [Nisi et al. 2016]…), which does not satisfy these climatological criteria in many cases. QPE and hail detection is frequently not precise and reliable enough. The spatial homogeneity can be compromised by the widening radar beam, by the lowering of the sensitivity by the distance from the radar, by ground clutter and beam blockage, interferences or by meteorological conditions (bright band, data filtering by zero radial winds …). The temporal homogeneity can be affected by changing radar types and placement, scan strategies, measurement intervals, calibration methods, outages, processing algorithms. The long-term availability depends on archiving hardware and methods, data formats and processing software. The result of these influencing factors is that the archived derived products are often not well suitable for climatological studies.

The previously stated concerns imply, that there is an opportunity to use the "raw" radar volume data and reprocess them with the climatological criteria in mind. As the reprocessing involves different types of radar systems, there is a need for common data format. The OPERA Data Information Model implemented in hdf5 format (ODIM-H5) was chosen as a common radar volume data format, because it is offering a modern and relatively stable storage model since publishing the ODIM-H5 v2.0.1 [Michelson et al. 2010]. However, the usability of this format was to be tested on the case of radar data from before its publishing.

2. Short history of Slovak weather radars

Radar meteorology has a long history in Slovakia. In the 1960s, the output of the air traffic radars (precision approach radar, airport surveillance radar, provincial surveillance radar) was transmitted to the meteorological office in the next building via a closed video circuit. The first specialised meteorological radar was put into operation in 1972 in Malý Javorník (Malé Karpaty). Radar data were archived in paper form and most of the archive was lost in a flooded building.

The first automated meteorological radar MRL-5 was installed in eastern Slovakia at Kojšovká hoľa (Volovské vrchy) in the early 1990s. ARMS, the Automated Radar Measurement System developed at SHMU, was used and the whole territory of Slovakia was covered for the first time. In 1992, the Russian radar at Malý Javorník was replaced by the EEC DWSR92C Doppler weather radar. Archiving of data on small tape began. MRL-5 at Kojšovská hoľa was replaced in 2004 by Radtec RDR 250GC with ENIGMA signal processor and the use of CD, later DVD as archiving media started.

In 2010, SHMU purchased a new Tivoli archiving machine and started online archiving of radar data. In 2015, the renovation of the Slovak weather radar network started with the replacement of Malý Javorník and Kojšovská hoľa by new dual polarisation SELEX METEOR 735CDP radars. In 2016, two newly build radars were put to operation at Kubínska hoľa (Oravská Magura) and Špaňí laz (near Veľký Krtíš).

3. Radar Volume Data Used in the Study

After inspecting all the available archive data from the SHMÚ radar network data from 4 different radar systems were found, stored in 5 different file formats. The list of the used data is in the Table 1.

Radar System	Type	Data format	Available Period	Storage Media	Volume measurement frequency
MRL-5	Dual-band (X and S band), without Doppler	In-house developed binary files	1998 - 2004	Tapes, CDs	Every 30 min.
EEC DWSR 92C	Doppler C-band	EDGE proprietary binary files	1998 - 2015	Tapes, DVDs	Every 15, later 10 and 5 min.
Radtec RDR 250GC	Dual-Pol, Doppler, C-band	Muran (2 different versions: proprietary binary format and hdf5)	2005 - 2015	Tapes (on HPC), DVDs	Every 7,5, later 5 min.
Selex METEOR 735 CDP	Dual-Pol, Doppler, C-band	Rainbow XML with compressed binary data blobs	2015/16 - 2023	Tapes (on HPC)	Every 5 min.

Table 1: List of the used radar volume data

The oldest data available were from the MRL-5 radar system located at the Kojsovska hola site. It was a dual-wavelength radar working at the X and S frequency bands without Doppler or dual-polarization processing capabilities. The data were stored on CDs and uploaded to our processing server. The file format was an in-house developed binary file – described in [Jurasek and Meri 2023]. Fortunately, the source-code of a processing and image generation software was found together with the volume files. This allowed us to develop a conversion tool from the in-house binary file format to the ODIM-H5 format. There were some minor problems detected in the conversion to ODIM-H5 format. The distinct scans were not well defined in the used binary in-house binary format. Because of the slow elevation speed, the transition between the elevation scans is smooth – it was hard to tell which elevation the given ray belongs when transitioning form one elevation angle to other. There were some extra sector scans (around 30 degrees wide in azimuth) at 10 degrees elevation angle (Fig. 1). Only the full or near full scans were added to the final ODIM-H5 files. The resulting volumes contain around 24 elevations every 30 minutes with data from the X and S channels (Fig. 2) from years 1998 to 2004.

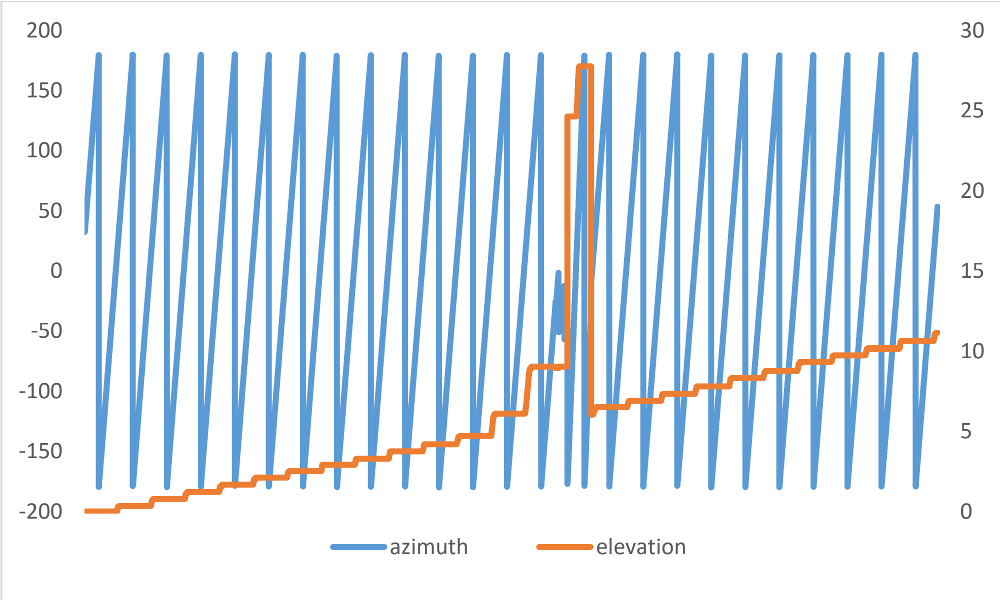

Fig. 1 : Scan strategy of the MRL-5 radar at the Kojsovska hola site.

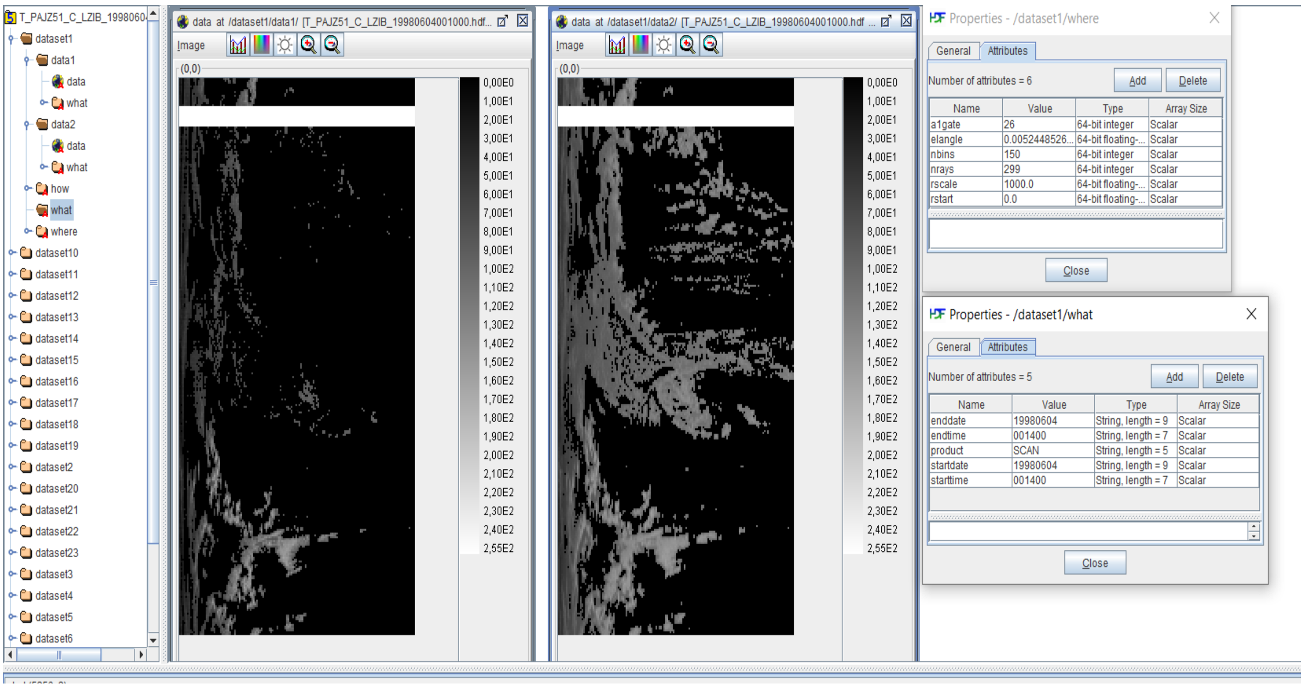

Fig. 2 : MRL-5 data from 04.05.1998 00:10 UTC in ODIM-H5 format. X-channel on the left, S-channel on the right. The white stripes are the rays depicted as no-data due to poorly defined scans.

The next radar system processed was the EEC DWSR 92C located at the Malý Javorník site. It was a C-band Doppler radar (without dual-pol.) working from 1998 to 2015. The archived data were stored first on magnetic tapes and later on DVDs in proprietary binary EDGE format. As data from this radar were sent to the Odyssey center from around 2012 an in-house developed converter to ODIM-H5 was already available. The measurement interval was changing during the lifetime of the radar from every 15 minutes to 10 and lately to every 5 minutes.

The third system was the Radtec RDR 250GC radar located at the Kojšovská hola site. It was a dual-pol. C-band Doppler radar working from 2005 to 2015. The volume data from this system were archived on DVDs and later on magnetic tapes on the archiving module of the HPC server of the SHMÚ. These data were stored in two different formats: first in proprietary binary Muran files, later in some custom Muran hdf5 files. In-house converters to ODIM-H5 were developed for both formats. The measurement interval of this system was every 7.5 minutes, later every 5 minutes.

The last system used in this study was the current radar network of the SHMÚ. It consists of four Selex METEOR 735 CDP dual-pol. C-band Doppler radars located at sites Malý Javorník (MJ), Kojšovská hola (KH), Kubínska hola (KB) and Španí laz (SL). They are in use from 2015 (Malý Javorník and Kojšovská hola) and from 2016 (Kubinská hola and Španí laz). The long-term storage of the volumes is done on magnetic tapes by the archivation module of the HPC server of the SHMÚ. The primary format of the volume files is Rainbow XML with data in binary blobs, but ODIM-H5 files are also produced by the Rainbow software. The measurement interval of this network is every 5 minutes.

4. Data processing

After all the radar volumes were converted to ODIM-H5 format the selection of an appropriate product was done, where these climatological concerns were taken in to account:

- resulting product needs to minimize effects of spatial inhomogeneities:
    - beam blockage
    - ground clutter
    - bright-band
    - sharp contours of the overlapping areas
    - higher or lower average values in the overlapping areas
- processing method needs to cope with the change from 2 to 4 available radar sites
- processing method needs to deal with temporal inhomogeneities
    - changing hardware, calibration and sensitivity
    - changing scan strategy from measuring every 30 min. up to every 5 min.

The most influential spatial factor in this case was the complex orography of Slovakia (Fig. 3) with ASL heights of the terrain between 100 and 2600 meters. Beam-blockage modelling at different ASL level was done to determine the lowest ASL level without significant beam-blockage. As the result the CAPPI-5km reflectivity (DBZH) composite was chosen as a product processed in this study. The 5 km level satisfies also the ground-clutter reduction criteria (ground clutter at this height is not present) and the bright-band reduction criteria (the 0°C isotherm below 5 km in the whole year in Slovakia).

There are sometimes sharp contours or higher/lower values perceived in the overlapping areas of radars when a long-time processing is done. This is probably due to the chosen compositing method, which is mainly using the maximum detected value in the given point. This leads to higher values in the overlapping areas as there is a chance to find a higher value as opposed to the areas where data is available only from one radar. To reduce this effect quality index (QI with values between 0.0 and 1.0) based compositing was used, where the QI is a weighting factor for the overlapping data. Three Qis were used: distance QI (decreases with the distance from the radar), beam-blockage QI (fraction of blocked energy), similarity QI (speckle, noise, ground clutter removal). The resulting QI used in the weighting was the product of these 3 QIs. The in-house developed qRad software [Meri et al. 2021] was used in the QI evaluation and compositing step.

The problem of the changing number of radar sites (and potential temporary outages of individual sites) the method of using "virtual" radars was developed. The method is based on computing the 5-year average difference for each month between the CAPPI-5km products from each radar-site. The 5-yaer average difference is mainly influenced by stable spatial factors – beam-blockage and relative height differences between the radars. These monthly average radar-to-radar differences were used to reconstruct the missing data by adding them to data from the available radars. The computation of the virtual value from the missing radar is described in the equation (1):

$$Z_j = \frac{\sum_{i=1}^{n}(Z_i+\Delta_{ij})\,QI_i}{\sum_{i=1}^{n} QI_i}, \qquad (1)$$

where Z_j is the computed virtual value of reflectivity, n is the number of available radars in the given point, Z_i is the reflectivity from the i-th available radar, Δ_{ij} is the average difference reflectivity in the CAPPI-5km level in the given point between missing and the i-th radar, and QI_i is the QI of the reflectivity measurement from the i-th radar. Of course the data can be reconstructed only in the overlapping areas. Example of average differences and the reconstructed CAPPI-5km field from a missing radar site is in the Fig. 4. The usage of the virtual radars when used for supplementing the two new radar sites to a composite with the old radar network with only 2 radars is depicted in Fig. 5.

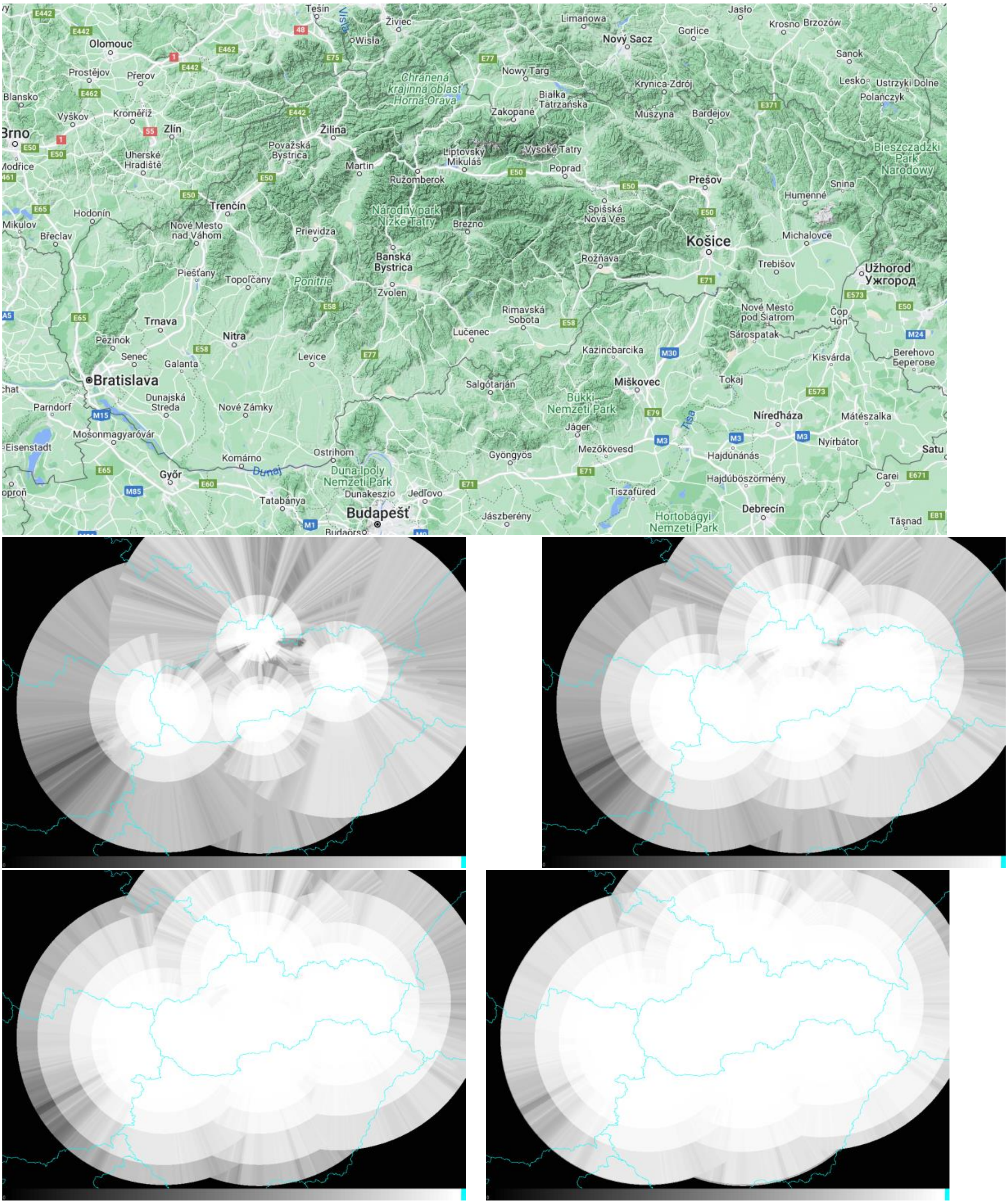

Fig. 3: Orography of Slovakia and the modelled beam-blockage from 4 radars at 2, 3, 4 and 5 km ASL (white – bot-blocked, gray – partially blocked).

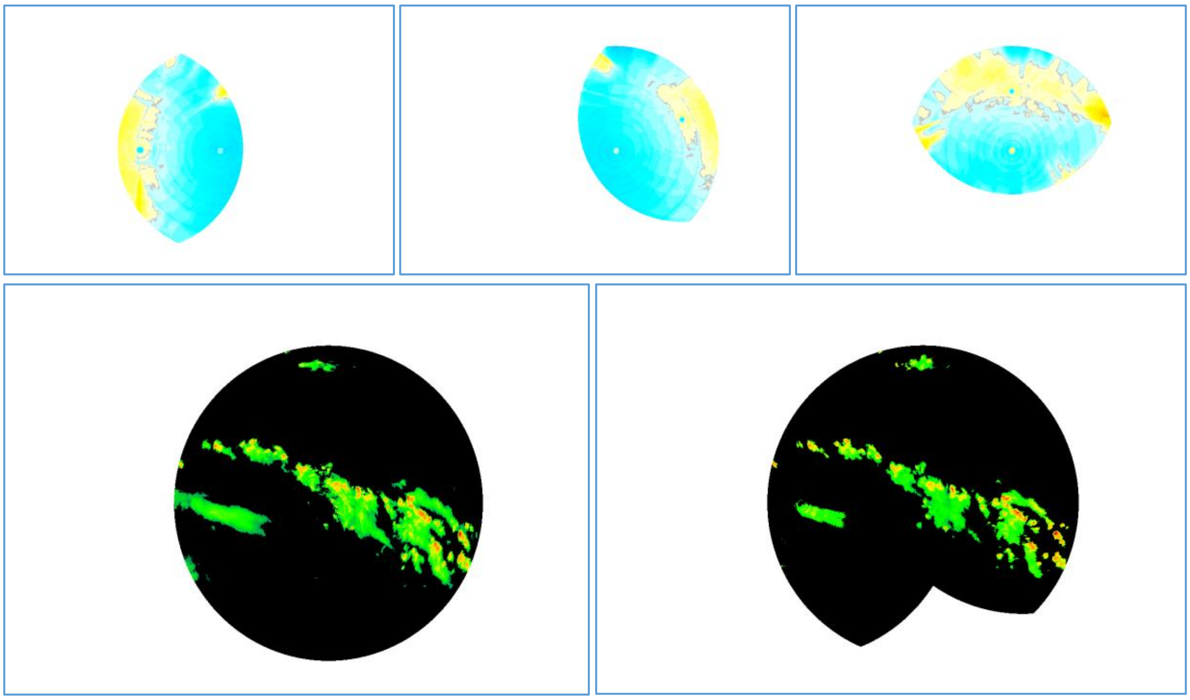

Fig. 4: Usage of average differences for reconstructing “virtual” radar products. Upper row, from left to right: average differences of the CAPPI-5km reflectivity between Malý Javornik vs. Španí laz, Kojšovská hola vs. Španí laz and Kubíska hola vs. Španí laz. Bottom row: the actual CAPPI-5km product from Španí laz (left) and the reconstructed CAPPI-5km product using only the 3 other radars and the average differences (right).

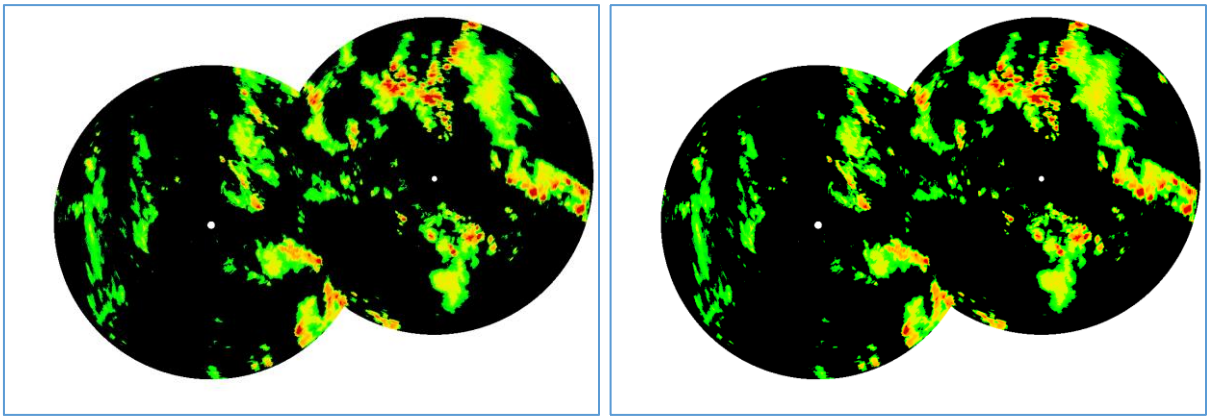

Fig. 5: Using virtual radars to add the reconstricted data from the two new radar sites (north and south center) to the origial network of two radars. On the left is the roiginal composite from two radars, on the right is the new composite using two original and two reconstructed radar products.

To eliminate the effect of changing hardware, calibration and sensitivity a high reflectivity threshold of 40 dBz was chosen, which is very likely to be detected by the radars. Reflectivity above 40 dBz at 5 km height is mainly related to severe storms. To reduce the effect of longer duration missing data, we decided not to compute the frequency of occurrence of the >40 dBz values, where a missing months-long interval can significantly alter the result. Instead we chose to compute the average duration of reflectivity >40 dBz in a given point in minutes. The resulting product should be an indicator of stationary storms occurrence. Areas with higher average duration of >40 dBz events are more prone to stationary storms which implies a higher risk of flash-floods in the given area.

The changing measurement interval (from every 15(30) minutes up to every 5 minutes) was treated by interpolating the successive products to 5 minute time-steps when the measurement interval was greater than 5 minutes. The 5 minute time step was supposed to be sufficient to study stationary storms, where one is interested for low speed cells with no significant displacement in 5 minutes. A simple motion vector algorithm was used, where the motion vector was determined by the displacement at the maximum correlation of the consecutive products. As the maximum interpolation time step were from only 2.5 to10 minutes, no significant displacement errors were expected for these small time intervals.

A storm event in a given point was defined as an event starting when the reflectivity value is greater than 40 dBz and ending when the reflectivity is lowering below the 40 dBz threshold. The maximum allowed time gap between >40 dBz values was 10 minutes (see Fig. 6).

Time	12:00	12:05	12:10	12:15	12:20	12:25	12:30	12:35	12:40	12:45	12:50	12:55	13:00
dBz	35	**41**	36	38	**42**	30	30	35	**45**	**47**	35	35	32

Event 1
duration = 20 min

Event 2
duration = 10 min

Fig. 6: Example of the storm event duration definition.

The resulting processing chain can be divided in 6 steps (Fig. 7):

1. ODIM-H5 volume generation – using the in-house developed converters and the already available ODIM-H5 volumes.
2. Generation of the single-site CAPPI-5km reflectivity products together with the corresponding QI fields.
3. Interpolation of the single-site products to a 5 minute interval by the motion vectors if necessary.
4. Reconstruction of the products from the missing radar sites by the “virtual” radars method, using the average differences.
5. Generation of the composite CAPPI-5km reflectivity product using the previously generated, interpolated and reconstructed single-site products and QI fields.
6. Computing the average duration of >40 dBz events from the composite CAPPI-5km maps.

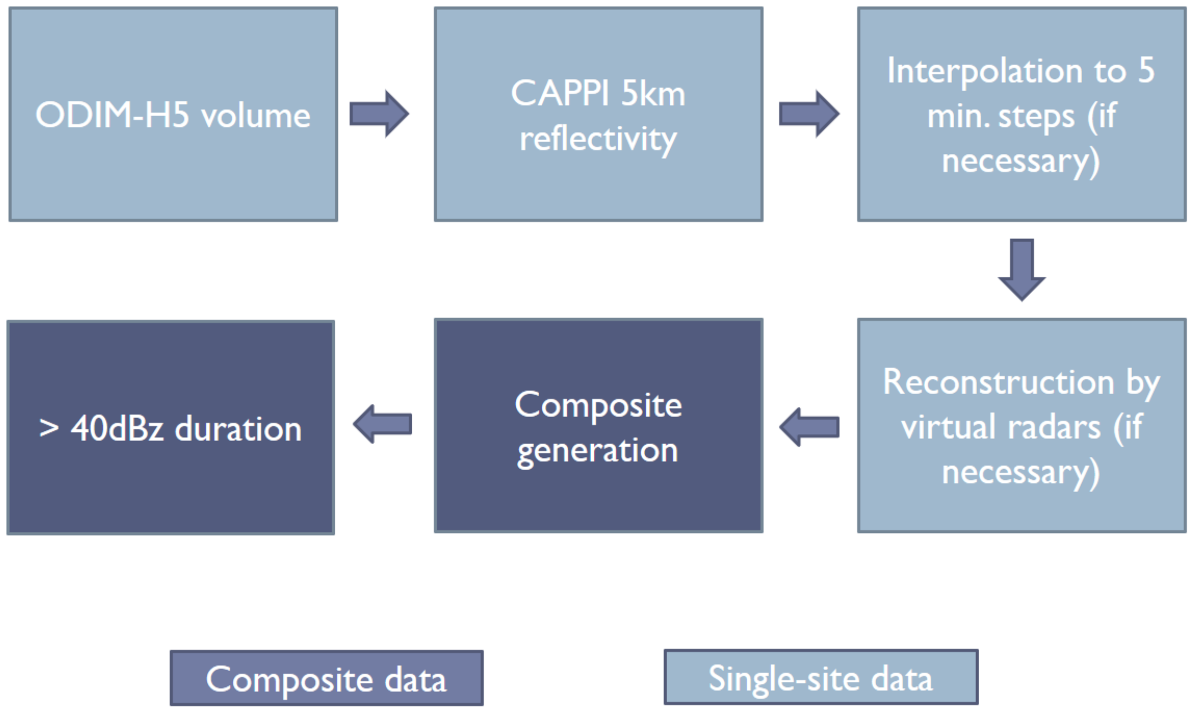

Fig. 7: The processing chain to generate the final results.

5. Results

A map of average duration of CAPPI 5 km reflectivity greater than 40 dBz in the period of 1.1.2007 to 31.12.2022 (16 years), characterizing intense rainfall and storm events was created (Fig. 8).

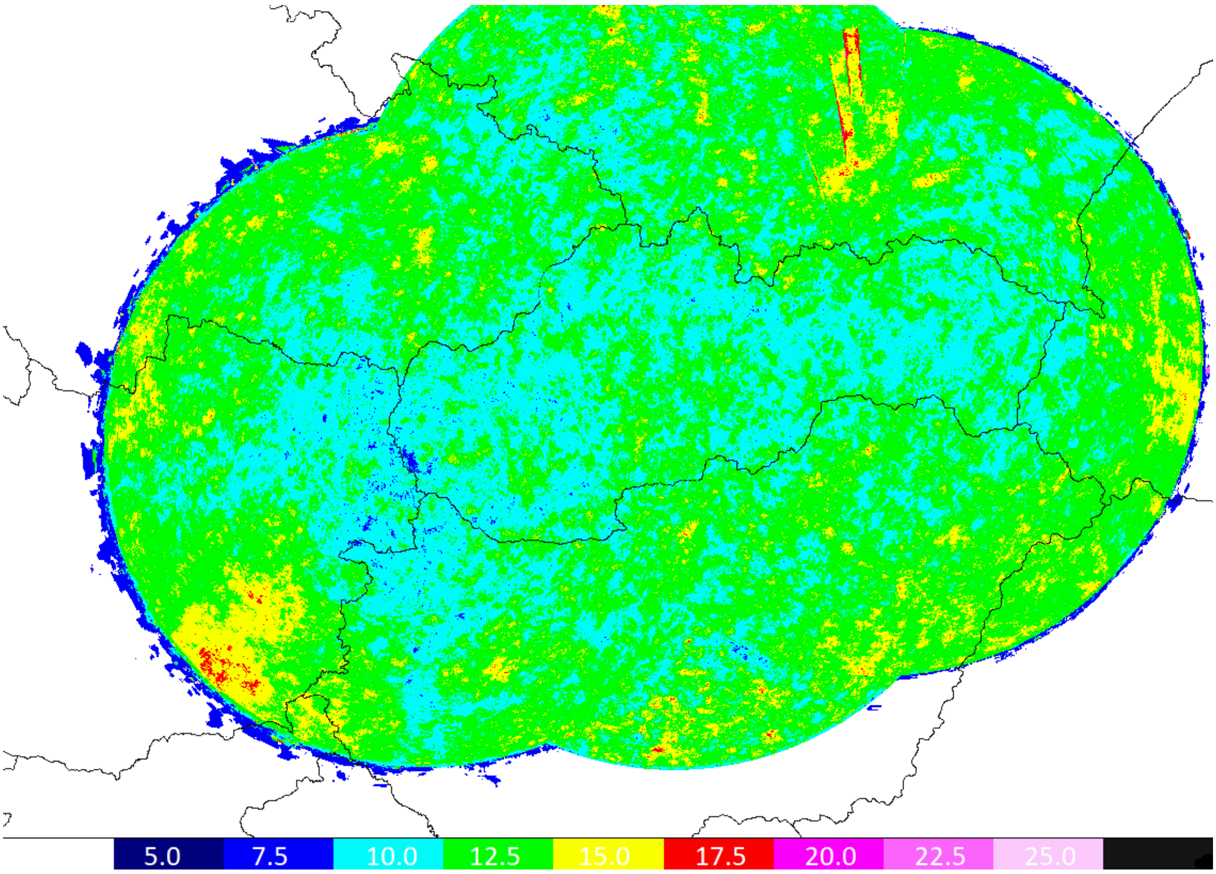

Fig. 8: Average duration of CAPPI 5 km reflectivity >40 dBz in minutes.

Several effects were found on the resulting map (see the annotated map in Fig. 9). There are generally low differences in the duration of >40 dBz events at CAPPI-5km in Slovakia. The values vary between 7.5 and 15.0 minutes, with maximum in the Štiavnicke vrchy and Belianske Tatry regions and minimum in the western part of the country in the downwind region of the Alps. Realistically higher values are in the region of the southern slopes of the Alps.

Somewhat unrealistic effect is the higher duration on average towards the edges of the radar coverage. This can be a result of the widening radar beam, which hit also significant parts below the 5 km level, where higher reflectivity is expected on average. This effect is not detected in Slovakia because of the good coverage and the compositing method used. Another unrealistic effect are the two rays with high values in Poland. This are probably some residual interferences detected by the Kojšovská hola site. The lower values around the coverage are a result of changing the maximum detectable range during the processed period.

In general, there are no evident effects of ground clutter, beam-blockage or bright-band and there are no contours or higher values on the overlapping areas.

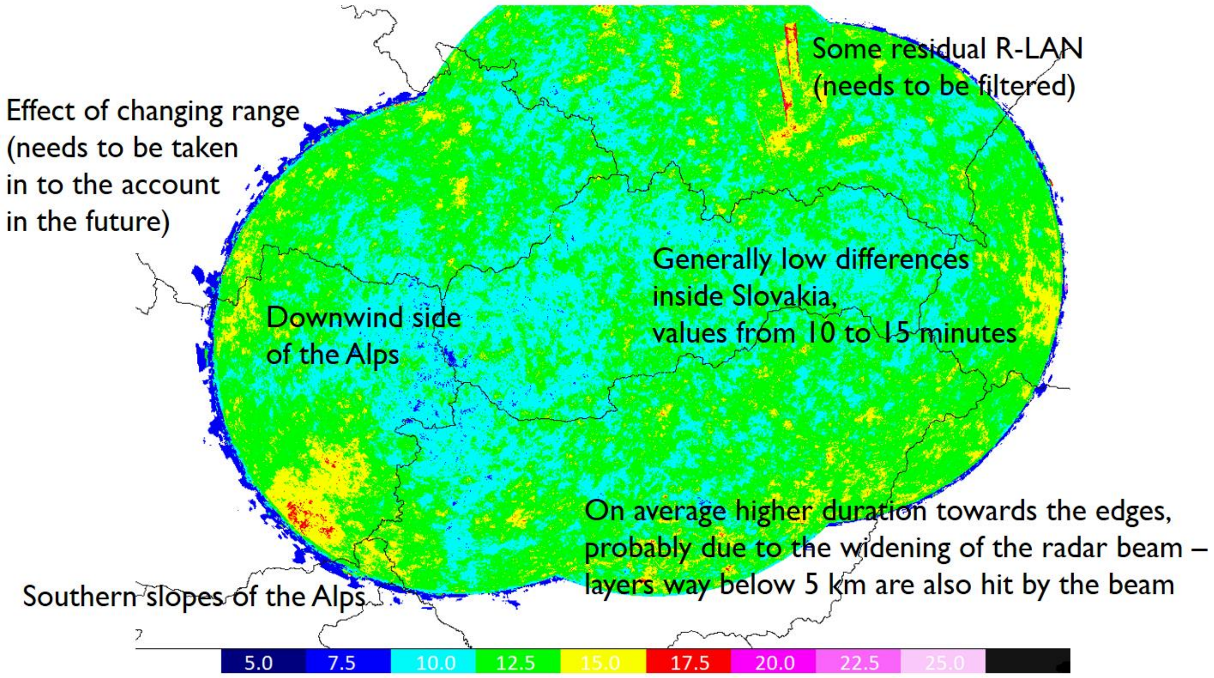

Fig. 9: The final map annotated with the main effects on the results.

6. Conclusion

A large scale of radar volume data conversion to ODIM-H5 format was done in the framework of the OA2 working package. Some minor challenges were only found in the conversion of the MRL-5 data due to the poorly defined scans in the original storage format. Overall the ODIM-H5 information model and data format is well suited to store and process long-term radar volume data in a unified way. The unification of the data format significantly simplifies the reprocessing of archive volume data and allows a wide range of reprocessing actions. It allows to use a common and uniform processing chain for radar data from different sources. The in-house built conversion software and the acquired skills can be useful also for other OPERA members.

From the climatological point of view several spatial and temporal homogeneity concerns needs to be taken in to account when radar data are reprocessed. An upper atmospheric parameter (above 4 or 5 km) needs to be considered in a heterogeneous topography region like Slovakia to mitigate the undesired spatial inhomogeneity effects. Several methods needs to be used to mitigate the temporal inhomogeneity issues – reconstructing "virtual" radar data, interpolation to a finer time-resolution.

This feasibility study demonstrated an approach to process long-term radar data by creating a map of average duration of >40 dBz reflectivity in the CAPPI-5km level to depict potentially hazardous areas from the flash-flood sensitivity point of view.

Some minor issues were detected in the resulting map, which needs to be taken in to account in the future works. The MRL-5 data were not used in the final product, because a reliable ground-clutter filtering algorithm needs to be implemented for this type of radar (no Doppler information available).

Results were published at EMS 2023 [Méri et al. 2023].

7. Reference

Jurašek, M., Méri, L.,2023:ET2023a_OPERA_OA2_Longtime_series.pptx, https://agora.fmi.fi/display/OPERA/ET2023a+Teleconference?preview=%2F120062663%2F120063536%2FET2023a_OPERA_OA2_Longtime_series.pptx

Marra, F., and E. Morin, 2015: Use of radar QPE for the derivation of intensity–duration–frequency curves in a range of climatic regimes. J. Hydrol., 531, 427–440, https://doi.org/10.1016/j.jhydrol.2015.08.064.

Méri, L.; Gaál, L.; Bartok, J.; Gažák, M.; Gera, M.; Jurašek, M.; Kelemen, M. Improved Radar Composites and Enhanced Value of Meteorological Radar Data Using Different Quality Indices. Sustainability 2021, 13, 5285. https://doi.org/10.3390/su13095285

Méri, L., Jurašek, M., Okon, Ľ., and Kaňák, J.: Long time radar data series processing using OPERA ODIM, EMS Annual Meeting 2023, Bratislava, Slovakia, 4–8 Sep 2023, EMS2023-337, https://doi.org/10.5194/ems2023-337, 2023.

Michelson, D. B., Lewandowski, R., Szewczykowski, M., Beekhuis, H., 2010: EUMETNET OPERA weather radar information model for implementation with the HDF5 file format, Version 2.0.1, OPERA Working Document WD_2008_03, https://www.eumetnet.eu/wp-content/uploads/2019/05/OPERA-ODIM_H5-v2.01.pdf

Nisi, L., O. Martius, A. Hering, M. Kunz, and U. Germann, 2016: Spatial and temporal distribution of hailstorms in the Alpine region: A long-term, high resolution, radar-based analysis. Quart. J. Roy. Meteor. Soc., 142, 1590–1604, https://doi.org/10.1002/qj.2771.

Saltikoff, E., and Coauthors, 2019: An Overview of Using Weather Radar for Climatological Studies: Successes, Challenges, and Potential. Bull. Amer. Meteor. Soc., 100, 1739–1752, https://doi.org/10.1175/BAMS-D-18-0166.1.

Tabary, P., P. Dupuy, G. L'Henaff, C. Gueguen, L. Moulin, O. Laurantin, C. Merlier, and J.-M. Soubeyroux, 2012: A 10-year (1997–2006) reanalysis of quantitative precipitation estimation over France: Methodology and first results. IAHS Publ., 351, 255–260.