Data quality in the BALTRAD+ Project

Jan Szturc[1], Daniel Michelson[2], Jarmo Koistinen[3], Günther Haase[2], Markus Peura[3],
Rashpal Gill[4], Martin B. Sørensen[4], Katarzyna Ośródka[1], Anna Jurczyk[1]

[1] Institute of Meteorology and Water Management, Warszawa, Poland
[2] Swedish Meteorological and Hydrological Institute, Norrköping, Sweden
[3] Finnish Meteorological Institute, Helsinki, Finland
[4] Danish Meteorological Institute, København, Denmark

(Dated: 30 May 2012)

1. Introduction

The BALTRAD/BALTRAD+ project „An advanced weather radar network for the Baltic Sea Region: BALTRAD" is carried out in the frame of Baltic Sea Region Programme 2007-2013 financed by the European Union. It is aiming at real-time exchange of weather radar data. Each partner in the developed decentralized network provides its polar data: radar reflectivity, Doppler data, and polarization moments where applicable. In a modern weather radar network processing of radar data alone is insufficient, so the most sophisticated correction algorithms are incorporated into the BALTRAD system (Ośródka et al., 2010).

Quality assurance (QA) in the BALTRAD+ covers all phases of weather radar process from scanning and signal processing, through volume raw data generation, to end user specific product generation. Two basic levels of QA are recognized: data corrections and data quality characterization. The developed algorithms for data corrections are to improve performance of radar production. The algorithms can be divided into categories based on analysis of scan geometry, spatial structure of reflectivity data, dual-polarization parameters, or external data (raingauge network, satellite, etc.).

2. Chain of data quality algorithms

The quality algorithms are designed in different ways in dependence on the type of data. In general, the following data types are distinguished (Holleman et al., 2006): volume of 3-dimensional radar reflectivity, Z_e, surface precipitation, R, and Doppler velocity, V. Corrections for wind volumes and products (e.g. for wind profiles to detect birds and insects) have not been included in QA at this stage of work.

2.1. Volume of 3-dimensional radar reflectivity Z_e data

The following order in processing of reflectivity volumes has been approved in BALTRAD+ project:

- Technical calibrations, such as instrumental accuracy, network power levels, antenna pointing angles, etc. (Gekat et al., 2003).
- Diagnosis of non-meteorological echoes (ground clutter, biological, interferences, etc.) and type of precipitation, especially hail (Peura, 2002; Ośródka et al., 2012).
- Correction of the effects of "Doppler snake" on reflectivity data, which are regions with radial velocity close to zero and thus often introducing reduced reflectivity due to high Doppler filtering.
- Analysis and correction of beam blockage (partial and total) due to terrain obstacles (Bech et al., 2007).
- Diagnosis of the probability of beam overshooting, which depends on distance to radar site and on the actual height of the vertical profile of reflectivity (VPR).
- Path attenuation in precipitation (Ośródka et al., 2012) and due to wet radar radome (Kurri and Huuskonen, 2008).
- Correction of intra-volume advection and fall velocity due to "time stamp". If possible each individual elevation scan should be independently time stamped, however it is common to only record the start time of the volume scan.
- Diagnosis of the strength of convection based on vertical profile of reflectivity (VPR), bright band recognition, or using external data, e.g. updraft effects visible on drop size distribution (DSD).
- Probabilistic diagnosis of the occurrence of overhanging precipitation in the lowest elevation, based on radar-derived VPR or NWP information (Pohjola and Koistinen, 2004).
- Diagnosis of precipitation type and particle size distribution (PSD) in the volume scan ("bin situ" – related to measured bins) (Park et al., 2009).

2.2. Quality algorithms for ground level conditions, especially for precipitation rate

After quality control of volumes, the next stage in weather radar data processing is generation of specific 2-D products. The most important products are related to surface precipitation field. At this stage the data should be quality controlled as well. The following chain of dedicated algorithms is recommended in BALTRAD+ project:

- Correction of the effects of the vertical profile of reflectivity (VPR) in reflectivity data (bright band correction and extrapolation from the lowest elevation to the ground).

- Modification of reflectivity data due to hail for quantitative precipitation estimation (QPE).
- Correction due to falling velocity and horizontal advection to represent time stamp conditions at ground level (Lauri et al., 2012).
- Advection-based creation of virtual measurements at denser time intervals than the actual scanning interval, i.e. horizontal motion analysis of precipitation patterns and then interpolation the field applying it.
- Hydrometeor type analysis at ground level: vertical correction of the “bin situ” hydrometeor type applying data from automatic weather station (AWS) data, numerical weather prediction (NWP) models, or satellite observations.
- Dynamical conversion of radar reflectivity into the quantity of interest (rainfall intensity, snowfall intensity, visibility, etc.) – e.g. so called *R(Z)* relationship.
- Rain gauge adjustment of radar data (e.g. weighted multiple regression) and combination of the both data kinds (Todini, 2001).
- Nowcasting based on weather radar data extrapolation and more sophisticated techniques such as taking evolution of convective cells into account (Jurczyk et al., 2012).

Generally, in most cases both conventional and polarimetric algorithms are needed because particular quality algorithms significantly differ for these two technical solutions. It should be noted that the both algorithm lists above are optional, i.e. priorities and implementations will be performed according to local needs and resources – not necessarily all of the algorithms will be coded.

3. Description of algorithms included into the BALTRAD toolbox

The BALTRAD toolbox is a data processing and product generation framework that has been developed during the main-stage BALTRAD project (Henja et al., 2010). It is written in C with Python APIs; Java APIs are currently being added. The toolbox concept implies that common functionality is used to read, write, and manipulate data and metadata. We have achieved this by creating data objects that map closely to ODIM_H5 groups, attributes, and datasets (Michelson et al., 2011). The result is an efficient code base with a small memory footprint. Tools are added to the toolbox in any of the three ways described by Michelson et al. (2012). The most efficient way of integrating a new algorithm is to use the full features of the toolbox APIs when writing new code. Integrating existing code is done most efficiently by only using the common functionality for reading/writing ODIM data and then mapping toolbox objects to the code’s own objects. This strategy enables algorithm chains to be run efficiently in memory independently of who wrote the original code. Using the common I/O functionality also reduces the risk of data formatting incompatibilities and unnecessary duplication of work. More information is available on line at http://git.baltrad.eu/ .

The toolbox is being further elaborated during BALTRAD+. This section describes the data processing algorithms currently integrated into the toolbox, and algorithms that are being implemented into it.

3.1. Integrated quality assurance algorithms

The most uncomplicated tool implemented in the toolbox thus far is code contributed by KNMI for **monitoring of radar calibration stability** (Holleman et al., 2010) through the use of signals from the sun. This exemplifies the toolbox concept well: the original code is written in C, and the routines for reading input data were simply replaced with the common ones in the toolbox, so that this tool will read ODIM data. KNMI is not even a BALTRAD partner, showing how the community can benefit from its own contributions. There are two parts to the sun monitoring: 1) determination of antenna azimuthal and elevation angle accuracies, 2) absolute calibration levels based on external solar flux measurements. The latter monitoring procedure is not trivial, as it requires detailed knowledge of the radar’s calibration, and such information is simply not available for all radars. The former monitoring procedure is, in contrast, straightforward to apply provided all relevant metadata are available in ODIM v.2.1 files. This provides a harmonized means of monitoring the pointing accuracy of a single radar and a network of many radars.

Anomaly (non-precipitation echo) detection and removal is done with the bRopo package, which is a toolbox add-on. The original code (called ROPO) was developed by FMI (Peura, 2002) and it has been used operationally in real time for over 10 years there. We have integrated this tool using the toolbox APIs as a dependency instead of adding the ROPO code directly to the toolbox itself. This is a good way of integrating a set of existing tools and a significant amount of code. bRopo provides a set of image analysis based algorithms for identifying and removing speckle, clutter, biometeors (birds, insects), and echoes from external emitters. Output is similar to that illustrated in Fig. 1 below.

Beam blockage analysis and correction is performed using the beamb package, which is also a toolbox add-on because the package contains GTOPO30 topography tiles that are relatively large in size. This package has been developed by SMHI and it uses a geometric beam propagation model that oversamples the radar’s polar geometry in relation to the topography, as a means of resolving small-scale effects impacting on beam blockage. Currently, beamb assumes standard propagation conditions, but the code has been prepared to deal with non-standard conditions provided through NWP or radiosonde profiles. Beamb is configured to automatically generate blockage analyses based on the geometry of input data. These are stored and looked up as long as the geometry matches. These routines ensure that analyses will be available even when the scan strategy changes without advance notice. We have also prepared beamb to be able to generate results for any radar located in Europe and most of the middle-East, and Africa. The user can specify the radar’s half-power beamwidth when generating the beam-blockage analysis, and a threshold for accepted maximum blockage (e.g. 70%); data will be corrected up to that maximum, and above it data will be flagged with the “no-data” value. An example from a weather radar subject to extreme blockage is given in Fig. 1.

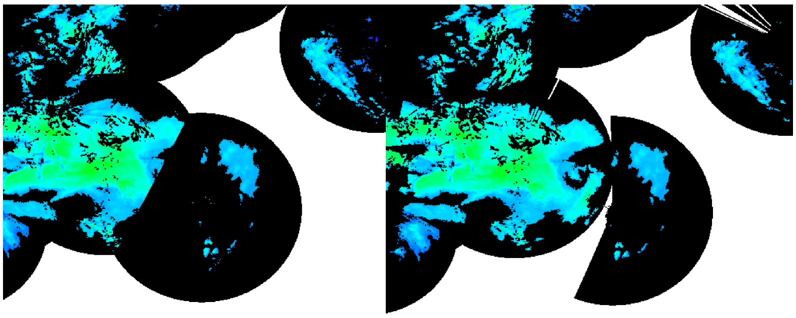

Figure 1. Detail from two experimental composites, one that hasn't been quality-controlled for beam blockage (left) and one that has been controlled (right). The corrected composite succeeds in capturing the precipitation that is blocked in the uncorrected composite.

An analysis of a radar's perceived maximum detection range given the current meteorological conditions can be expressed more simply as yielding quality information on the **probability of overshooting**. This is another example of code contributed from FMI (Koistinen and Hohti, 2010) that has been integrated directly into the toolbox. This information can be very valuable when characterizing data quality when precipitation systems are known to be shallow, e.g. during winter, but it should be emphasized that we don't use this information to correct or otherwise modify the original data. Probability of overshooting is a purely radar-based algorithm based on an elaborated analysis of the echo-top field (Fig. 2).

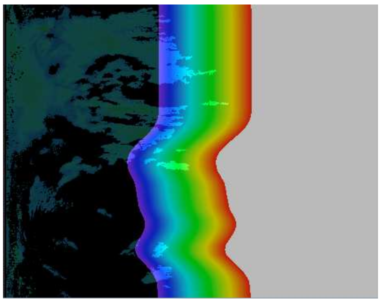

Figure 2. Reflectivity scan (rays clockwise top-to-bottom, bins left-to-right) from a radar with 460 km range, overlaid with quality indicator representing probability of overshooting.

Compositing is performed in the toolbox using a "slow" algorithm that transforms to the Cartesian grid directly from input polar volumes and scans. We call the algorithm slow because it is designed to be a reference that navigates all data as accurately as possible, and with no other optimizations that compromise accuracy. Currently implemented algorithms are nearest radar, lowest value to the Earth's surface, and a constrained maximum value. The generalized framework is prepared to accommodate other methods of compositing data. Although we call this algorithm slow, we have demonstrated how it can be successfully sped up through tiling to accommodate all European data within reasonable time constraints (Henja and Michelson, 2012). Explicit weighting of quality indicators/indices is not performed yet, although quality is implicitly addressed in other ways (Fig. 1). This work will be elaborated in BALTRAD+.

3.2. Quality control algorithms being integrated

RADVOL-QC

This set of quality assurance algorithms developed at IMGW is dedicated to non-polarimetric radars (Ośródka et al., 2012). Its specific is that it runs in two channels: data quality correction and characterization. The corrected radar data are 3-D radar reflectivity volumes. The quality information for QPE purposes is described by means of quality index (QI) from range (0, 1) (Einfalt et al., 2010).

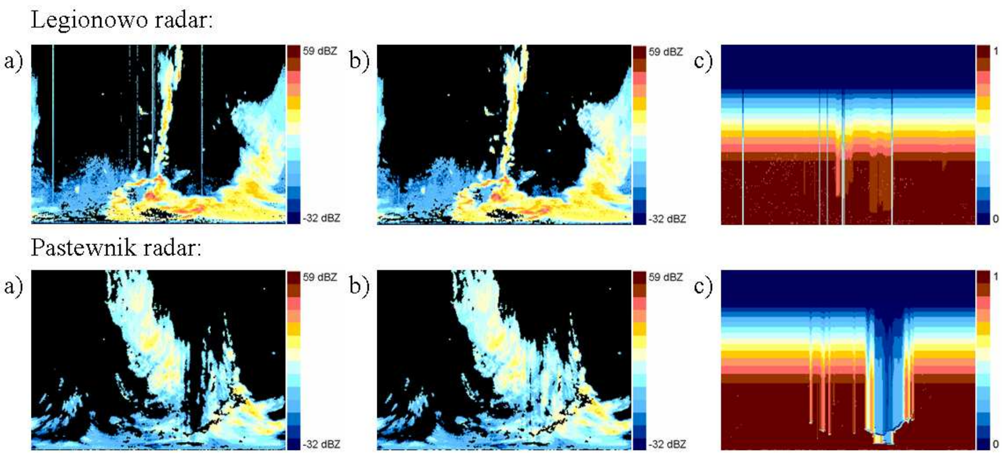

Fig. 3. Example of performance of all correction algorithms for the elevation (0.5°): a) raw data (in dBZ); b) corrected data; c) total quality index QI. Radar range is 250 km. The panels represent range (y-axis) vs. azimuth (x-axis) displays.

The following quality correction algorithms are included into the system at present stage of work: (i) BLOCK. This algorithm applies to a few errors in radar measurements related to especially ground clutter removal and partial and total beam blockage. It is based on digital terrain map and analysing radar beam path. (ii) SPIKE. Using this algorithms' set allows to remove different types of spurious echoes called wide, narrow and high spikes. Generally, they are geometrically-shaped non-meteorological echoes characterized by their specific spatial structure that clearly differs from precipitation field pattern: the shape of such echo is similar to a spike along the whole or large part of a single or a few neighbouring radar beams. The spikes often result from Sun radiation or external interferences. (iii) SPECK. This kind of contamination in radar observation is caused by noise in radar signal. The specks are isolated radar gates with or without echo which are removed after analysis of the given pixel's neighbourhood. (iv) ATT. The base for the algorithm is attenuation coefficient estimated for average propagation conditions. Attenuation in precipitation is difficult to determine because of iterative character of used algorithms, so certain thresholds are included.

Quality index is calculated taking into consideration not only aforementioned quality factors but the additional ones such as: SYS – describing impact of radar technical parameters on data quality, and BROAD – calculating broadening of radar beam with the distance from the radar site.

The quality algorithms were tested on radar data from Polish weather radar network POLRAD (Fig. 3) and then operationally implemented.

Rack

Rack is a standalone executable software developed at FMI. Its most relevant module in this context is AnDRe which provides anomaly detection and removal. Detection is based on non-polarimetric dBZ data. The detectors are similar to those of ROPO, but re-implemented in C++. The detection and removal are separate processes which means that processing by Rack can be combined with other software providing similar functionality. In addition, Rack provides compositing algorithms with various compositing rules, including quality-weighted compositing. An example of a resulting quality field is shown in Fig. 4. The program is explained in further detail by Peura (2012).

Rack supports reading and writing HDF5 files applying OPERA information model (ODIM). Rack is able to convert its data structures to native HDF5 structures in memory, which may be one general means for integrating Rack in other systems.

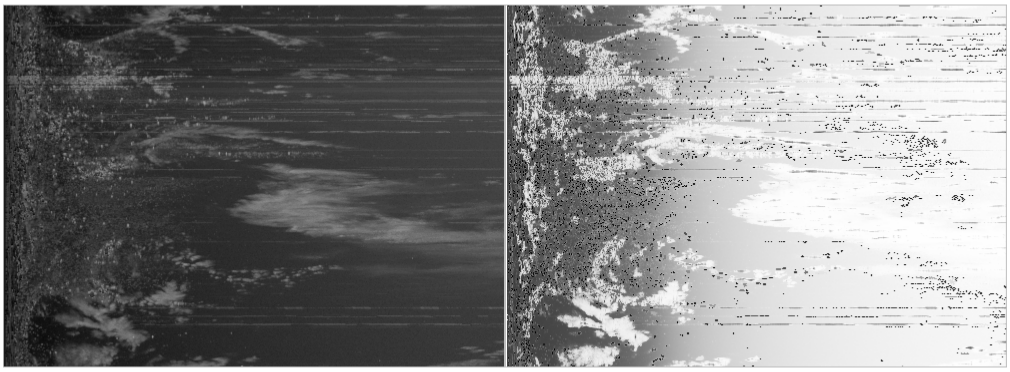

Fig. 4. Original sweep (on the left) and quality field (on the right) based on biometeor, emitter, ship, and speckle detections computed by AnDRe module of Rack.

HMC (Hydrometeor Classifier)

A hydrometeor classifier (HMC) using dual polarization C-band data has been developed with partial funding from the BALTRAD project (Gill et. al, 2012). The HMC also includes algorithms to correct for rain attenuations for reflectivity (Z) and differential reflectivity (Z_{DR}). The classification scheme is based on fuzzy logic and the membership functions are represented by 1-dimensional beta functions.

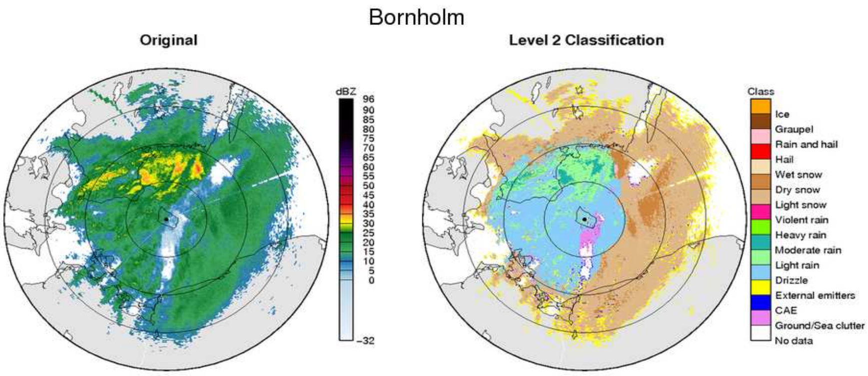

Fig. 5. The original radar image (on the left) and its corresponding level 2 hydrometeor classification (on the right).

In the current version, the algorithm can undertake the so-called level 1 and level 2 classifications. In the level 1 classification a radar echo is classified into one of four simple classes: precipitation, clutter, clean air echoes, and electrical signals from external emitters. Similarly, in the case of level 2 classification a radar echo is classified into one of 11 classes; ground clutter, sea clutter, external emitters, clean air echoes, light rain, moderate rain, heavy rain, violent rain, light snow, moderate to heavy snow, and hail/rain mixture. In the level 2 classification the melting layer heights from the numerical weather prediction model are used to aid the classification. Melting layer determination algorithm using the dual polarization parameters alone has also been developed as part of the HMC. This algorithm is under going evaluations before its use in the HMC scheme (Fig. 5).

One of the by-product of the HMC algorithm has been that it can be used to remove the false echoes in, amongst others, the original radar reflectivity product, Z_{HH}. The latter product has been much appreciated by the end users, such as the operational meteorologists, as it significantly improves the radar data quality.

In the future, as part of BALTRAD+ project, further improvements to the algorithm are planned such as fine tuning the membership functions for hail. As hail is observed in very small regions occupying few pixels, it has been a challenge to extract these cells in the radar data whilst ensuring they are not contaminated by other hydrometeor classes.

De-aliasing

The radial velocity of scattering particles is determined from their observed phase difference between successive radar pulses. There is a maximum velocity that can be determined unambiguously. This maximum velocity is called the Nyquist velocity and depends on the pulse repetition frequency and the wavelength of the radar. Velocities higher than the unambiguous velocity will be folded back into the fundamental velocity interval. This process is called aliasing. Velocity aliasing can usually be identified in radar images by detecting abrupt velocity changes between neighboring measurements.

The de-aliasing algorithm developed at SMHI is a robust tool based on a linear wind model and designed to eliminate multiple folding (Haase and Landelius, 2004). It consists of four steps: (i) Wind observations are mapped from polar to Cartesian space. (ii) A set of test functions with different wind velocities and directions is created and mapped to Cartesian space. (iii) The test function closest to the observation is determined. (iv) The correct Nyquist interval is determined and the observed wind is de-aliased.

Note that the vertical wind component is neglected in this approach. To improve the quality of the algorithm, low wind velocities need to be filtered beforehand as strong ground clutter may not be suppressed completely. Moreover, de-aliasing should not be performed on sparely and/or unevenly distributed data in polar space. Weather radar wind profiles might be less affected by aliasing errors as they are mostly derived from higher elevation angles with a higher Nyquist velocity. Unlike many other concepts, the proposed de-aliasing algorithm does not depend on wind information from a nearby sounding (e.g., radiosonde or wind profiler) or from a numerical weather prediction model. Fig. 6 shows how the new de-aliasing algorithm improves the observed radial wind for the Swedish weather radar in Vara.

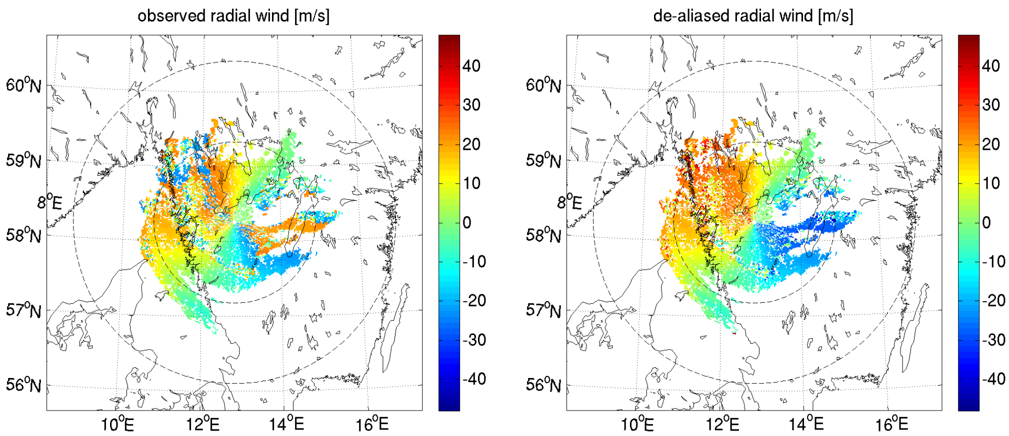

Fig. 6. Observed and de-aliased radial wind. The dashed lines indicate the 120-km and 240-km-range ring, respectively.

3.3. Examples of quality assurance algorithms to be defined

At present many QA algorithms listed in Section 2 are still in “brainstorming” phase. Some of them have been developed more so that we have at least a first guess solution draft available. Those are described tentatively in this subsection.

The largest sampling bias (difference between a radar measurement aloft or “bin situ” and conditions at ground level) at long distances is, on average, introduced by the negative vertical reflectivity gradient and partial beam overshooting in the snowfall part of precipitation, i.e. the vertical profile of reflectivity (VPR). The magnitude of this underestimating in radar reflectivity factor is typically 2-15 dB at ranges of 50-250 km from radar. A real time VPR correction method of the measured dBZ has been implemented at FMI to represent effective radar reflectivity factor (Z_e) in snowfall and in rain at ground level (Koistinen et al., 2003). The method aims to be invariant to the occurrence of rain, snow and melting in a composite volume of multiple radars. It also corrects for the overestimation due to interception with the bright band and underestimation due to the effects of partial beam overshooting and negative vertical reflectivity gradients. Prior to any application each measured VPR must pass automatic pattern recognition algorithms diagnosing precipitation type at ground, existence and altitude of the bright band, overhanging precipitation, evaporation, residual clutter and clear air echo (insects or birds) in the profiles. Each measured VPR is attached with a quality weight expressing the quality of the profile for the purposes of QPE. The correction term on dB scale is a time-space variable, obtained as a weighted average of an ensemble of VPR corrections derived individually at each grid point of a multi radar composite of lowest elevation angle scans. The time ensemble of corrections applies VPRs collected at each radar during the 6 hours preceding each volume scan. The time ensemble applies constantly also parameterized climatological VPRs, diagnosed and fitted with the real time height of the freezing level obtained from a numerical weather prediction model. The spatial ensemble of VPRs is obtained from the radars neighboring the one whose measurements are corrected. In absence of calibration errors the resulting ground level estimate of dBZ is a seamless continuous field in a radar network.

At FMI hydrometeor phase analysis at ground level is coupled with the VPR correction described above. The output of the VPR scheme is ground level estimate of dBZ either in dry snowfall or in rainfall – the effects of melting are eliminated (as far as we can). Thus a proper analysis is needed at ground level to find in real time the water phase of hydrometeors in order to perform proper $R(Z)$ or $S(Z_e)$ conversion (S denotes snowfall intensity in units of melted water, i.e. mm/h). The analysis is based on a two-parameter model (temperature and relative humidity at the height of 2 m above ground) as is explained in Saltikoff et. al. (2010). In areas where melting snow falls at ground we blend both estimates R and S applying local weights at each grid point which are proportional to the probability that the phase is rain (R) or dry snow (S).

Especially in cases of strong winds and snowfall the combined effect of slow falling velocity of snow particles with high horizontal wind speed introduce the phenomenon that particles measured at long ranges with radar (typically at altitudes of 1-4 km even with the lowest elevation angle) are drifted dozens or sometimes more than hundred kilometers before they land on the ground. The practical implication is that the real ground level pattern of precipitation intensity at the same time stamp moment as that in the low elevation PPI scan can differ significantly from the latter measurement. Thus even the qualitative precipitation pattern obtained from radar can be erroneous compared to the real one at ground level. As Lauri et al. (2012) have shown this bias can be corrected. However, real time calculation requires large computational resources as the correction applies derivation of 4-dimensional particle trajectories using 4-D wind fields from NWP output.

Acknowledgements

The paper was worked out in the frame of the BALTRAD (2009-2012) and BALTRAD+ (2012-2014) Projects „An advanced weather radar network for the Baltic Sea Region: BALTRAD" (Baltic Sea Region Programme 2007-2013), part-financed by the European Union (ERDF and ENPI).

References

Bech J., Gjertsen U., Haase G., 2007: Modelling weather radar beam propagation and topographical blockage at northern high latitudes. *Q. J. R. Meteorol. Soc*., **133**, 1191–1204.

Einfalt T., Szturc J., Ośródka K., 2010: The quality index for radar precipitation data – a tower of Babel*? Atmos. Sci. Let*., **11**, 139–144.

Gekat F., Meischner P., Friedrich K., Hagen M., Koistinen J., Michelson D.B., Huuskonen A., 2003: The state of weather radar operations, networks and products. In: *Weather radar. Principles and advanced applications*, P. Meischner (Ed.), Springer-Verlag, Berlin – Heidelberg, 1–51.

Gill R.S., Soeresen M.B., Boevith T., Koistinen J., Peura M., Michelson D., Cremonini R., 2012: BALTRAD dual polarization hydrometeor classifier. ERAD 2012 (this volume).

Haase G., Landelius T., 2004: Dealiasing of Doppler radar velocities using a torus mapping. *J. Atmos. Oceanic Techn.*, **21**, 1566–1573.

Henja A., Szewczykowski M., Ernes S., Michelson D., 2010: The BALTRAD technical platform. *Proceedings of ERAD 2010*.

Henja A., Michelson D., 2012: Improving the quality of European weather radar composites with the BALTRAD toolbox. ERAD 2012 (this volume).

Holleman I., Michelson D., Galli G., Germann U., Peura M., 2006: Quality information for radars and radar data. OPERA workpackage 1.2 (document: OPERA_2005_19).

Holleman I., Huuskonen A., Kurri M., Beekhuis H., 2010: Operational monitoring of weather radar receiving chain using the Sun. *J. Atmos. Oceanic Technol.*, **27**, 159–166.

Jurczyk A., Szturc J., Ośródka K., 2012: Object-oriented nowcasting of convective cells based on multi-source data. Manuscript.

Koistinen J., Michelson D.B., Hohti H., Peura M., 2003: Operational measurement of precipitation in cold climates. In: *Weather radar. Principles and advanced applications*, P. Meischner (Ed.), Springer-Verlag, Berlin – Heidelberg, 78–114.

Koistinen J., Hohti H., 2010: Operational diagnosis of precipitation detection range. *Proceedings of ERAD 2010* (short abstract).

Kurri M., Huuskonen A., 2008: Measurements of the transmission loss of a radome at different rain intensities. *J. Atmos. Oceanic Techn*., **25**, 1590–1599.

Lauri T., Koistinen J., Moisseev D., 2012: Advection-based adjustment of radar measurements. *Mon. Weather Rev.*, **140**, 1014–1022.

Michelson D.B., Lewandowski R., Szewczykowski M., Beekhuis H., 2011: EUMETNET OPERA weather radar information model for implementation with the HDF5 file format. OPERA Working Document WD_2008_03, Version 2.1.

Michelson D., Koistinen J., Peltonen T., Szturc J., Rasmussen M.R., 2012: Advanced weather radar networking with BALTRAD+. ERAD 2012 (this volume).

Ośródka K., Szturc J., Jurczyk A., Michelson D., Haase G., Peura M., 2010: Data quality in the BALTRAD processing chain. *Proceedings of ERAD 2010. Advances in Radar Applications*, 6–10 September, Sibiu, Romania, 355–361.

Ośródka K., Szturc J., Jurczyk A., 2012: Chain of data quality algorithms for 3-D single-polarization radar reflectivity (RADVOL-QC system). *Meteorol. Appl.*, **18** (Accepted).

Park H., Ryzhkov A.V., Zrnić D.S., Kim, K.-E., 2009: The hydrometeor classification algorithm for the polarimetric WSR-88D: Description and application to an MCS. *Weather Forecast.*, **24**, 730–748.

Peura M., 2002: Computer vision methods for anomaly removal. *Proceedings of ERAD 2002*, 312–317.

Peura M., 2012: Rack – a program for anomaly detection, product generation, and compositing. ERAD 2012 (this volume).

Pohjola P., Koistinen J., 2004: Identification and elimination of overhanging precipitation. *Proceedings of ERAD 2004,* 91–93.

Saltikoff E., Huuskonen A., Hohti H., Koistinen J., Järvinen H., 2010: Quality assurance in the FMI Doppler Weather Radar Network. *Boreal Env. Res*., **15**, 579–594.