Hydrometeor Classification for the DWD Weather Radar Network: First Verification Results

Jörg Steinert

Deutscher Wetterdienst, Frankfurter Str. 135, 63067 Offenbach am Main, Germany

(Dated: 18 July 2014)

1. Introduction

Although weather radar data have a high impact in national weather service environments, the significance is still increasing. Here, the wide horizontal coverage is used to get data in high spatial resolution for large areas. The idea is then to infer synoptic information from the collected data. Up to the present, the derivation of the precipitation rate was the main topic in this field. With the installation of dual-polarisation weather radars the possibility arises to investigate the hydrometeor type based on the additional measurements (e.g. Park et al. (2009) or Al-Sakka et al. (2013)). For the utilisation of the dual-polarisation technique, Deutscher Wetterdienst (DWD) started in 2010 with the replacement of the German C-band weather radar network. Beside this hardware renewal, the inhouse project "Radarmaßnahmen" was established for the development and implementation of a software framework called POLARA (POLArimetric Radar Algorithms, see Rathmann and Mott (2012)) for the integration of radar data related algorithms and methods. The aim of this paper is to give some information about the status of the hydrometeor classification at the DWD implemented in POLARA. To illustrate this, a brief description of the classification algorithm is presented in section 2 supported by an example weather case. In the following section 3 the verification of the hydrometeor classification will be addressed by the analysis of a two day weather case. After that the conclusion will give an overview of the results for this study.

2. Hydrometeor classification

The dual-polarisation radar delivers various additional parameters in comparison to the single-polarisation radar type. Especially the measurements of the reflectivity in horizontal polarisation (Zh), the differential reflectivity (ZDR), the co-polar correlation coefficient (ρ_{hv}) and the specific differential phase (KDP) are typical input data for the hydrometeor classification (cf. Straka et al. (2000)). Unfortunately, the radar information is not sufficient, as, for instance, light rain (with small spherical raindrops) and snowflakes give similar radar signatures. Consequently the inclusion of temperature information is appropriate to support the hydrometeor classification. For this, single level output from the numerical weather prediction model COSMO-DE was used. Selected were the height of the 0°C-isotherm (HZEROCL) and the height of the snow fall limit (SNOWLMT), see Baldauf et al. (2014). Whereby HZEROCL shall be used as an indicator for the upper limit of a melting layer and SNOWLMT used to discriminate between rain and snowfall.

The set of classes for the hydrometeor classification are a combination of user requests and technical distinguishable hydrometeor types. This concluded in classes for drizzle (DR), rain (RA), big drops (BD), wet snow (WS), dry snow (DS), ice crystals (IC), graupel (GR), rain or hail (RH) and hail (HA). Additionally, there was a demand for a separate detection of the melting layer (ML). For one thing the detected ML can support the hydrometeor classification and for another thing it is used in the quality assurance scheme for the radar data (Werner and Steinert, 2012).

The algorithm scheme for the hydrometeor classification is named "Hymec" and is basically a combination of the ML detection and the hydrometeor classification itself, see fig. 1 (right). The input radar data for Hymec pass a quality control scheme beforehand, see Werner (2014). Furthermore, there are two runs of Hymec, one before the attenuation correction of the reflectivity data and one thereafter, see fig. 1 (left). Here the first run of Hymec is done to aid the attenuation correction with initial hydrometeor classification products. More information to this can be found in Werner (2014). The classification results for external users or subsequent algorithms like quantitative precipitation estimation are taken out of the second run of Hymec.

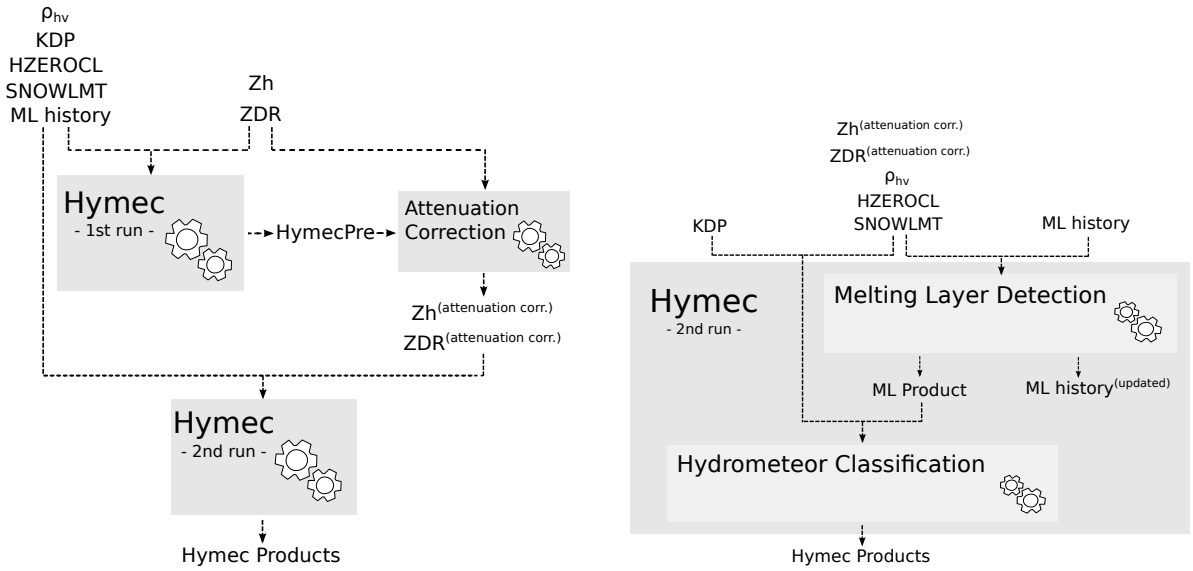

Figure 1: The structure of Hymec with the incoming and outgoing data and the related data paths. Left: The two runs of Hymec, Right: The structure inside Hymec at the example of the second Hymec run.

The implementation of the hydrometeor classification as well as the melting layer detection are based on a fuzzy-logic core. This kind of algorithm helps to handle signature overlaps for the hydrometeor classes. As initial version of the fuzzy-logic and the related membership functions the values of Park et al. (2009) are used by keeping in mind, that the values have to be adapted from S-Band to the operation at C-Band. Moreover, the ML detection includes a post-processing to underline the nature of the melting layer. In this step ML segments along the ray are combined and ML elements in neighbouring rays are considered. The outcome are more or less connected areas induced by melting particles, which appear as complete rings or at least ring segments in the PPI scans. An example of the hydrometeor classification is shown in fig. 2 on the right hand side together with the related attenuation corrected reflectivity on the left hand side. Additionally, Frech and Steinert (2014) give more information about the algorithm scheme and show the performance of Hymec for another weather case including a transition from rain into snowfall.

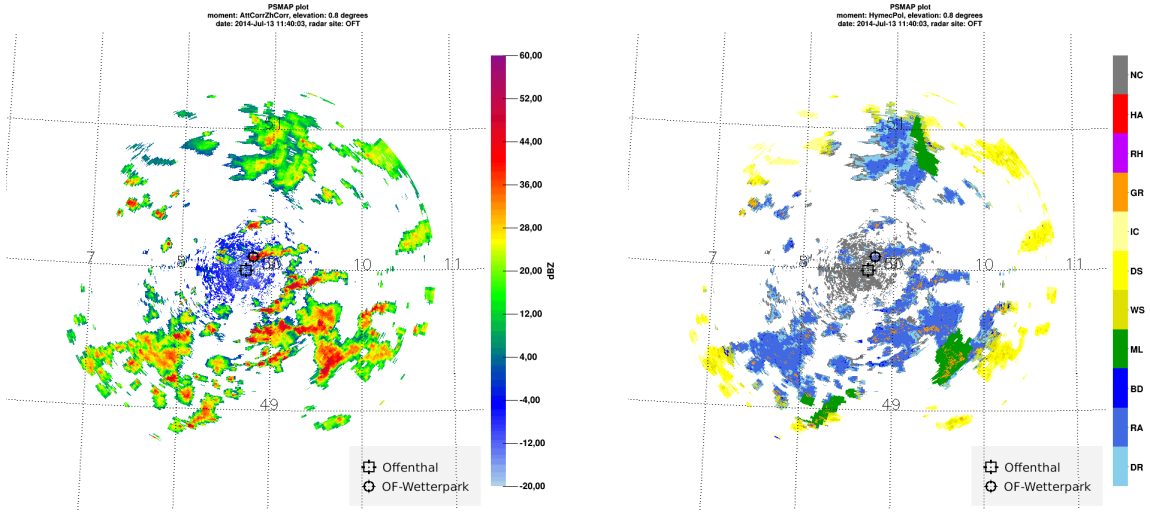

Figure 2: Example for the hydrometeor classification recorded at 11:40 UTC on 13.07.2014 from radar Offenthal (OFT). The data are taken from an operationally measured PPI scan with a variable low elevation (terrain following). Marked are the positions of the radar with a rectangle and the optical disdrometer (section 3) with a circle. Left: attenuation corrected reflectivity, Right: hydrometeor classification (description of the hydrometeor class abbreviations are given in the text of section 2 and NC means 'not classified').

3. Verification of the hydrometeor classification

For the analysis of the hydrometeor type of a specific region, the determination by utilisation of remote sensing data is just the first step. Moreover, the accuracy of the estimated hydrometeor class is important to know for the developer and especially the user. For the verification, the application of data from devices which deliver time series point measurements are an appropriate tool. As device a so called laser precipitation monitor (LPM) was used. This optical disdrometer measures the

particle spectra and derives beside the precipitation intensities the hydrometeor type of the particles, too. The main calculated parameters are available for an integration time of 1 and 5 minutes. In fig. 3 (left) an example weather case is shown with the data from an LPM located at the Wetterpark in Offenbach/Main (height of the station: 118 m ASL). Starting from the DWD radar Offenthal the direction to this LPM is 23.9° (referred to north) and the distance is 12.8 km, see markings in fig. 2. For the coordinates of the LPM a time series was then generated out of the radar PPI scans. This time series is displayed in fig. 3 (right) and formed the base for the comparison with the LPM information. From the radar data just the measurements of the terrain following sweep (in direction to the disdrometer the elevation is 1°) with 1° beam width and a range resolution of 250 m was used. The bottom of the radar ray has a mean value of 363 m ASL and so the height difference between the disdrometer and the radar measurement is about 250 m. Especially for low melting layers this height difference is important.

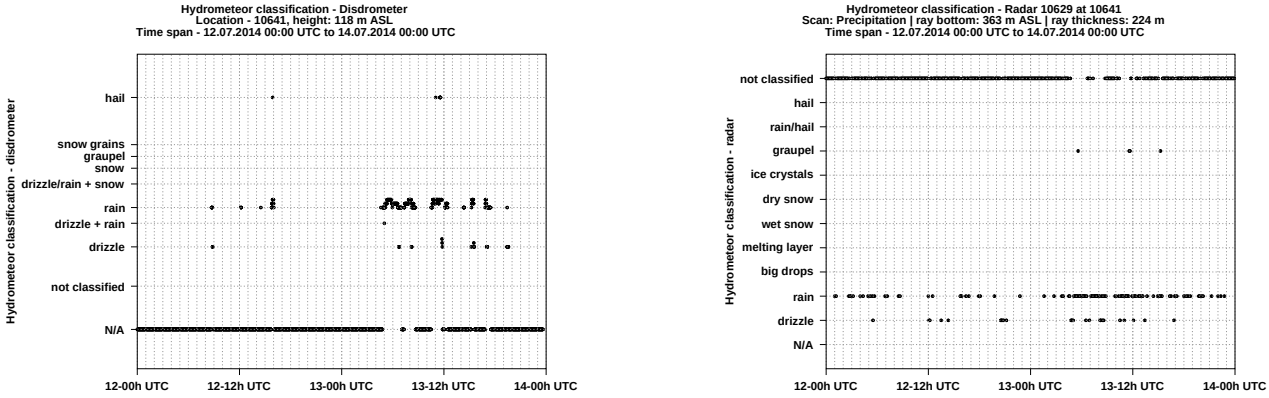

Figure 3: Time series of the hydrometeor classification at the postion of the disdrometer (Wetterpark Offenbach) for time span between 0:00 UTC on 12.07.2014 and 0:00 UTC on 14.07.2014. Left: results from disdrometer, Right: hydrometeor classes based on data recorded with radar OFT.

The presented case in fig. 3 covers a time span of two days, in that some convective cells pass the disdrometer. As an example, both systems estimate around 11:40 UTC on 13.07.2014, the occurrence of graupel or hail, see figs. 2 and 3. Furthermore, the radar data shows a temporal overestimation of rain events in contrast to the disdrometer data. An explanation might be that some of the convective rain cells are small in their horizontal dimensions and therefore didn't pass the disdrometer directly. The study of longer time spans shall deliver more information on such effects.

Apart from the qualitative analysis, the calculation of verification measures give a direct imagination of the performance of the hydrometeor classification. As a first step for this verification, contingency tables were built from the two independent time series data sets. At the beginning a 2x2 matrix version (table 1) of the contingency table was set up for the interpretation of single hydrometeor types. In a later stage the contingency table shall be expanded to the direct analysis of more classes.

Table 1: Generic version of the contingency table to compare the results of the disdrometer (LPM) and the radar derived result for a selected hydrometeor class.

	LPM: class observed	LPM: class not observed
Radar: class estimated	a	b
Radar: class not estimated	c	d

With the four entries in table 1 various statistical quantities can be calculated. Very common ones are the probability of detection (POD) (eq. 3.1) and the false alarm ratio (FAR) (eq. 3.2). The POD is also named as hit rate and gives the number of hits in relation to all events considered to designated hydrometeor class by the disdrometer.

$$POD = \frac{a}{a+c} \tag{3.1}$$

$$FAR = \frac{b}{a+b} \tag{3.2}$$

Furthermore the so called probability of false detection POFD (e.g. Barnes et al. (2009)) can be given with eq. 3.3 and shows the number of false alarms in relation to all events which are not designated to the specific hydrometeor class.

$$POFD = \frac{b}{b+d} \tag{3.3}$$

Therewith POFD can be seen as opponent to the POD. To analyse the contingency table with one verification measure the extremal dependence index (EDI) was chosen, see Ferro and Stephenson (2011). In eq. 3.4 the notation of Mariani et al. (2014) is shown.

$$EDI = \frac{\ln(POFD) - \ln(POD)}{\ln(POFD) + \ln(POD)} \tag{3.4}$$

Whereas the POD, FAR and POFD varies between 0 and 1, EDI gives values from -1 to 1.

As to be seen in fig. 3 the list of hydrometeor classes differ between the disdrometer and the radar derived results. Because of this, more or less similar hydrometeor types were combined to class groups to realise the comparison. For this study the number of class groups is set to a minimum to just give an impression of the envisaged verification. The four class groups are rain, snow, hail and no information (N/A). The included classes from the disdrometer and the radar point of view are summarised in table 2. At this point it has to be stated, that the number of class groups has to be increased for further analysis, because otherwise the number of detected hydrometeor classes can't be verified.

Table 2: The assignment of the hydrometeor classes of the disdrometer and radar results (see fig. 3) to the class groups for verification.

class group	disdrometer hydrometeor classes	Hymec hydrometeor classes
rain	drizzle, drizzle + rain, rain	drizzle, rain, big drops
snow	drizzle/rain + snow, snow, snow grains	melting layer, wet snow, dry snow, ice crystals
hail	graupel, hail	graupel, rain/hail, hail
N/A	not classififed, no information	not classified, no information

Now, table 3 gives the contingency tables from time series data presented in fig. 3 with the utilisation of the class groups from table 2. For the designated class group, the values in the 'not obs.' (not observed) columns are related to events that are connected to the other three class groups. The same apply to the 'not estimated' row. Additionally, the estimates of POD, FAR, POFD and EDI are presented for each class group at the end of the table.

Table 3: Contingency tables with absolute values related to the time series data in fig. 3 and the class groups in table 2. Moreover the verification measures for each class group are put at the bottom.

	disdrometer results (ground truth)							
	rain		snow		hail		N/A	
	obs.	not obs.	obs.	not obs.	obs.	not obs.	obs.	not obs.
Hymec: class estimated	64	14	0	0	1	3	475	19
Hymec: class not estimated	22	476	0	576	0	572	14	68
POD	0.74		N/A		1		0.97	
FAR	0.18		N/A		0.75		0.04	
POFD	0.03		0		0.01		0.22	
EDI	0.85		N/A		1		0.96	

The first look at the contingency tables in table 3 shows that the number of N/A values is with 489 (85% of the events) very high. There are no snow events and just 1 true hail event. The remaining 86 events (64 estimated + 22 not estimated by Hymec) are classified by the disdrometer as rain. Because of the convective nature of the two day weather case this distribution seems reasonable.

Further analysis on the single classes delivers a good result for snow, because no false alarms were generated from Hymec. The hail class seems to be well detected by just looking at the POD, POFD and EDI values. But the FAR is with a value of 0.75 high, though just related to 3 false alarms and 1 hit. This shine out of the FAR has to be discussed because the high value of the false negatives (572) is not considered. Rain as the main precipitation class group in this weather case gives ordinary verification measures. Because snow and hail are not overestimated, the assumption is, that the misses for rain refers to not available (N/A) events. An improvement in this area shall be regarded.

To sum up, the verification of this two day weather case makes a good agreement between radar-based Hymec results and the disdrometer measurements. Nevertheless there is some work to do to improve the hydrometeor classification for example by further adjusting of the membership functions. Referring to the hail event around 11:40 UTC on 13.07.2014, the author observed graupel through the windows of his flat in Offenbach. For this time stamp Hymec says graupel and the disdrometer estimates hydrometeor class hail.

4. Conclusions

The implemented hydrometeor classification together with the melting layer detection gives the hydrometeor classes in the spatial resolution of the radar data. For the verification of the estimated hydrometeor classes the temporal availability of the observations by the disdrometer is very important to have a good cross-section for various weather events. Related to the geographical conditions some hydrometeor types are rare in some season. For example, in winter the precipitation of snow is typical, whereas hail is rather unusual. The consequence is, that a verification, which wants to consider all hydrometeor types, has to be installed for longer periods. The utilisation of optical disdrometers builts the base for the verification and the measures presented in section 3 give a hint of the performance of the hydrometeor classification algorithm. The idea is now to expand the verification for further positions where such disdrometers are located. By the looking at longer temporal distances and different seasons this verification shall give measures which will help to improve the hydrometeor classification.

Acknowledgement

As the hydrometeor classification is not applicable without high quality radar data, the preceding quality assurance is very important. Because of this I would like to acknowledge Manuel Werner (DWD) for providing the quality controlled and attenuation corrected input data.

References

H. Al-Sakka, A.-A. Boumahmoud, B. Fradon, S. J. Frasier, and P. Tabary, "A new fuzzy logic hydrometeor classification scheme applied to the French X-, C-, and S-band polarimetric radars," *J. Appl. Meteor. Climatol.*, vol. 52, pp. 2328–2344, 2013.

M. Baldauf, J. Förstner, S. Klink, T. Reinhardt, C. Schraff, A. Seifert, and K. Stephan, "Kurze Beschreibung des Lokal-Modells Kürzestfrist COSMO-DE (LMK) und seiner Datenbanken auf dem Datenserver des DWD," Deutscher Wetterdienst, Tech. Rep., April 2014.

L. R. Barnes, D. M. Schultz, E. C. Gruntfest, M. H. Hayden, and C. C. Benight, "Corrigendum: False alarm rate or false alarm ratio?" *Weather and Forecasting*, vol. 24, no. 5, pp. 1452–1454, 2009.

C. A. T. Ferro and D. B. Stephenson, "Extremal dependence indices: Improved verification measures for deterministic forecasts of rare binary events," *Weather and Forecasting*, vol. 26, no. 5, pp. 699–713, 2011.

M. Frech and J. Steinert, "Polarimetric radar observations during an orographic rain event - performance of a hydrometeor classification scheme," *accepted for publication, Hydrology and Earth System Sciences Discussions*, 2014.

S. Mariani, M. Casaioli, A. Lanciani, S. Flavoni, and C. Accadia, "QPF performance of the updated SIMM forecasting system using reforecasts," *Meteorological Applications*, 2014.

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

N. Rathmann and M. Mott, "Effective radar algorithm software development at the DWD," in *7th Europ. Conf. On Radar in Meteor. and Hydrol.*, no. NET316, 2012. [Online]. Available: http://www.meteo.fr/cic/meetings/2012/ERAD/extended_abs/NET_316_ext_abs.pdf

J. M. Straka, D. S. Zrnić, and A. V. Ryzhkov, "Bulk hydrometeor classification and quantification using polarimetric radar data: Synthesis of relations," *Journal of Applied Meteorology*, vol. 39, no. 8, pp. 1341–1372, 2000.

M. Werner, "A new radar data post-processing quality control workflow for the DWD weather radar network," in *8th Europ. Conf. On Radar in Meteor. and Hydrol.*, no. 079, 2014.

M. Werner and J. Steinert, "New quality assurance algorithms for the DWD polarimetric C-band weather radar network," in *7th Europ. Conf. On Radar in Meteor. and Hydrol.*, no. NET403, 2012. [Online]. Available: http://www.meteo.fr/cic/meetings/2012/ERAD/extended_abs/NET_403_ext_abs.pdf