AMTD amt-2015-370 – Reply to Referee #2

Jörg Burdanowitz, Christian Klepp, and Stephan Bakan

February 10, 2016

email address: joerg.burdanowitz@mpimet.mpg.de

In the following answer we proceed as follows. Text from the Referee #2 is shown in *italic*, our answer in **bold** and changes in the manuscript are highlighted in blue.

General Comments:

This manuscript presents a novel method for estimating automatically the precipitation phase from the disdrometer and auxiliary meteorological data in time scale of minutes over the ocean. The method is extension to earlier developed statistical model using now more variables in the regression analysis, and showing an improvement in accuracy reaching to 91 % for two phases and to 81.2 % including also mixed phase. The method is applied to an interesting and valuable data set of over four years of data from the Atlantic Ocean. And as stated in the manuscript, validation data sets for the satellite retrievals over ocean are rare and therefore this research is significant for ground validation e.g. with NASA GPM mission. The proposed algorithm can be applied to similar instruments and utilized as verification of climate models.
The scientific quality of the manuscript is good, the method is well explained, and the testing utilizing the four scores, accuracy, bias, misclassified and uncertain cases, are profoundly described and clarified with tables and figures. The results are compared to earlier studies in objective and quantitative manner, and explanations to found discrepancies are discussed. There are some shortages in the description of error sources of the measurement, which in my opinion should be addressed, and some statements or descriptions which require more clarification.

The details are shown below. The language is good, as far as, I as non-native can control it. The manuscript is mostly well-structured and easy to follow. My recommendation is to publish the manuscript, when the minor corrections are considered.

We would like to thank referee #2 for taking the time to review our manuscript. We appreciate the valuable comments and suggestions to improve our manuscript and the recommended literature. In the following we address all raised points one by one.

Specific Comments:

1. Reference on the page 13647 line 10: (Anagnostou et al. 1999). Is this proper reference here to present the different coastal and oceanic PSDs? The study is based on gauge and radar comparison, where in the study it is shown the different PSD shapes of different climatological conditions? The reference (Bumke and Seltmann, 2012) in the manuscript therefore reports no notable differences in DSDs in between continental and maritime areas.
We replaced the reference to Anagnostou (1998) by that of Kidd and Levizzani (2011). Furthermore, we added Bumke and Seltmann (2012) as reference. As further work, we plan to test their findings using OceanRAIN data.
"Furthermore, existing coastal and island-based precipitation measurements may not fully represent oceanic precipitation because the measured particle size distributions (PSDs), rates, and accumulations may differ from those measured over the open ocean (Kidd and Levizzani, 2011). However, Bumke and Seltmann (2012) found no difference between PSDs over coastal areas and open ocean."

2. Page 13647, line 22: extra-tropics, without dash.
Corrected.

3. Page 13648, line 14: "...perfectly agreed with observers log during the Lofoten Cyclones campaign in measuring snowfall events." This is strongly said, although the same is stated also in the reference. I would add the word "in detection of snowfall events", while measuring refers also to the quantitative analysis e.g. precipitation accumulation, and comparison in the reference in this respect is not satisfying. I would leave the word perfectly *out.*

We agree on the mismatch of "measuring" with respect to "observer's log" and thus replaced "measuring" by "detecting".

"For snow, a predecessor of the current ODM470 agreed with the observer's log during the Lofoten Cyclones campaign (LOFZY; Klepp et al., 2010) in detecting snowfall events."

4. Page 13652 line 19 - Page 13653 line 10: This section is taken straight from the earlier studies of some of the authors in the manuscript, and hence it is understandable that the results are adopted. And although the manuscript is not providing precipitation rate calculated with the equations (5) and (6), but only describing it as a possible predictor variable, I would like to see some discussion of the induced errors with the adopted assumptions. The selection of lump graupel m-D relation in Lempio et al. (2007) led to strong overestimation of precipitation rate in the comparison to Geonor/manual measurements in some of the study cases and this was stated in the Lempio et al. (2007) that the theoretical assumptions are not possibly valid for all winterly precipitation.

We agree that an assumption of lump graupel as an overall snow retrieval is far from perfect and introduces errors, thus we point this out clearer in the manuscript. However, it is important to note, that no unique snow retrieval exists because a disdrometer measures a cross-sectional area that does not directly relate to the required maximum dimension. Hence, it does not correspond to the liquid water equivalent or mass of the particle. The lump graupel assumption suits well because lump graupel is spherical and thus does not require a transfer function between cross-sectional area and maximum diameter. Snowflakes larger than 9 mm in diameter are rare compared to smaller snowflakes, thus this assumption makes sense. Individual snowfall events may surely lead to wrong estimates as they would for any other snow retrieval assumption, as well. We clarify this in the manuscript as follows.

"For snow, the measured cross-sectional area differs from the required maximum dimension of the particle due to the non-spherical shape of snowflakes. This difference requires applying a transfer function. However, Lempio et al. (2007) found that the product of particle terminal fall speed and particle mass (liquid water equivalent) as a function of cross-sectional area is in the same order of magnitude for various frozen precipitation particle types. Hence, no transfer function between cross-sectional area and maximum diameter is required when using a spherical lump graupel assumption.

The lump graupel assumption works well for frozen precipitation particles between 0.4 and 9 mm in diameter, whereas particles exceeding 9 mm in diameter rarely occur. Nevertheless, events with large particles introduce larger errors to the estimate in the same way as the retrieval quality may largely differ for individual snowfall events. Overall, no unique snowfall retrieval can be derived using optical disdrometers without recording the individual particle shape. Compared to a Geonor gauge, the optical disdrometer agreed well in most cases and overestimated a few light snowfall cases during the 1999/2000 winter period at Uppsala (Lempio et al., 2007)."
In terms of precipitation rate uncertainty, the rain retrieval performs accurately as supported in several direct comparisons (Eigenbrodt company test site, shipboard intercomparisons with a ship rain gauge using side collectors, and SPICE extreme rainfall events in Boulder, CO). However, mixed-phase and snow that both use the snow retrieval hold higher and yet unknown uncertainties that certainly vary from case to case. To quantify this uncertainty, the ODM470 was part of the global WMO International Solid Precipitation InterComparison Experiment (SPICE), which took place at Marshall field site in Boulder (CO, USA) during the winters 2013/14–2014/15. The results are currently under investigation by an international science team and results are expected in late 2016.

In Klepp et al. (2010) I think this has been overlooked by stating that the lump graupel is the most frequently occurring precipitation type over the cold-season Norwegian Sea. The data set given in the reference is in my opinion statistically too small for such a statement, and therefore the m-D relation cannot be applied generally as the manuscript implies. Battaglia et al. (2010) considers the error sources of Parsivel instrument in measuring the snow and many of the problems are applicable to ODM470 as well.
We agree that the formulation is misleading. Even though lump graupel was the most frequently observed type of precipitation during the Lofoten Cyclones campaign (LOFZY) it does not necessarily mean that this remains true at all times. To clarify this we constrained the statement as follows.
"Klepp et al. (2010) observed lump graupel being the most frequently occurring precipitation type over the cold-season Norwegian Sea during the LOFZY campaign."
Thanks for suggesting Battaglia et al. (2010) as a reference, which we added to the manuscript.

Battaglia et al. (2010) discuss several sources of error for a snow-measuring PARSIVEL whereof those for particle shape and orientation, margin effects, and coinciding particles also apply to the ODM470. However, the PARSIVEL is more sensitive to influences by wind speed and wind direction on the falling precipitation particles because the PARSIVEL has a fixed non-pivotable horizontal optical sensing area.

5. Page 13653, section 2.2.: How the time periods of manual observations (every 3 h) are considered in the comparison to the automatic observations (1 minute resolution). Please clarify.

Thank you for pointing out that our description how the three hourly observations relate to 1-minute time steps was not clear enough. The simplest approach would be to apply the PP inferred from a 3 hourly observed ww code to all following 1 minute time steps. In the manuscript we describe how we used ancillary data to better bridge the gap between two 3 hourly observational time steps. However, if the ancillary data does not help to confidently identify the PP we simply use the PP inferred from the ship observation. We added these information to the manuscript.

"Ancillary data means in particular atmospheric variables measured onboard the ship including the ODM470, such as air temperature, humidity, and precipitation rate. This ancillary data is available at much higher resolution of 1 minute compared to the 3 hourly observations. Initially, we assign the PP derived from the ww code directly to every single minute of precipitation that follows a 3 hourly observation as a first-guess information. If available, air temperature as one of the ancillary data serves to possibly correct this first guess PP."

6. Page 13656, line 23: Does the maximum particle diameter mean maximum of the observed particle diameters of 1 minute? Please clarify.

Yes, the maximum particle diameter means the maximum particle size measured within one minute. Accordingly, we added the information to clarify this.

"Instead of D, we use the 99th percentile of D, D_{99}, which is a measure for the maximum particle diameter measured within one minute but excluding erroneously large particles possibly caused by particle coincidences, drip-off drops or other artifacts."

7. Page 13658, line 11 and line 22: Clarification of the terms rain disagreement and snow disagreement, is it defined in respect to manual ob-

servations or model. This would improve also the bias score definition.
We see the ambiguity of our definition and added the following example to clarify this.
"For instance, rain disagreement means that the statistical model predicts rain that disagrees with the manual PP reference data indicating snow."
For the bias: "Accordingly, a bias score of $b < 1$ represents an overprediction of snow by the model, whereas $b > 1$ represents an overprediction of rain by the model."