Quality of composites

Günther Haase, Swedish Meteorological and Hydrological Institute

Daniel Johnson, Swedish Meteorological and Hydrological Institute

OD12: Quality of composites

Günther Haase and Daniel Johnson

Swedish Meteorological and Hydrological Institute, Norrköping, Sweden

on behalf of EUMETNET OPERA

OPERA Working Document WD_XXX

May 15, 2018

1 Objective

The bid for EUMETNET OPERA 4 describes the OD12 work package (Quality of composites) as follows: *Monitoring of the quality of the composites. The bidders are requested to indicate how the quality is measured. Note that some monitoring tools are proposed to be developed already when completing OD1-OD8. It is advisible to co-operate with users of radar composite data, such as NWP and hydrological modelers. The EUMETNET H-SAF, for instance, has tools (CAVAL Work Package) which are useful to assess the quality of rainfall composites.*

2 Data

2.1 OPERA composites

OPERA produces three different composites (Matthews et al., 2012) which are updated every quarter of an hour (at minute 00, 15, 30, and 45) and issued ca. 15 minutes after data time:

- **Precipitation intensity:** each composite pixel is a weighted average of the lowest valid pixels of the contributing radars, weighted by the inverse of the beam altitude. Polar cells within a search radius of 2.5 km of the composite pixel are considered. Data measured below 200 m altitude are not used. Measured reflectivity values are converted to rainfall (mm/h) using the Marshall-Palmer equation.
- **Accumulated precipitation:** integral of the previous four 15-minute precipitation intensity products.
- **Maximum reflectivity:** Each composite pixel contains the maximum of all polar cell values of the contributing radars at that location.

The OPERA composites, covering large parts of Europe with a spatial resolution of 2×2 km^2 (Lambert Azimuthal Equal Area projection), are available in BUFR (since November 2010) and HDF5 (since January

2013). Data are corrected for topographical beam blockage using BALTRAD's *beamb* module and clutter-filtered using BALTRAD's *bropo* module and satellite imagery.

This study focuses only on the quality of the precipitation intensity composites. For comparison with a reference dataset (see section 2.2) hourly accumulations have been calculated. We do not use OPERA's accumulation product, as we want to keep track of all pixels in each single precipitation intensity composite. Data are aggregated to daily (0600 to 0600 UTC) precipitation sums (Fig 1 left).

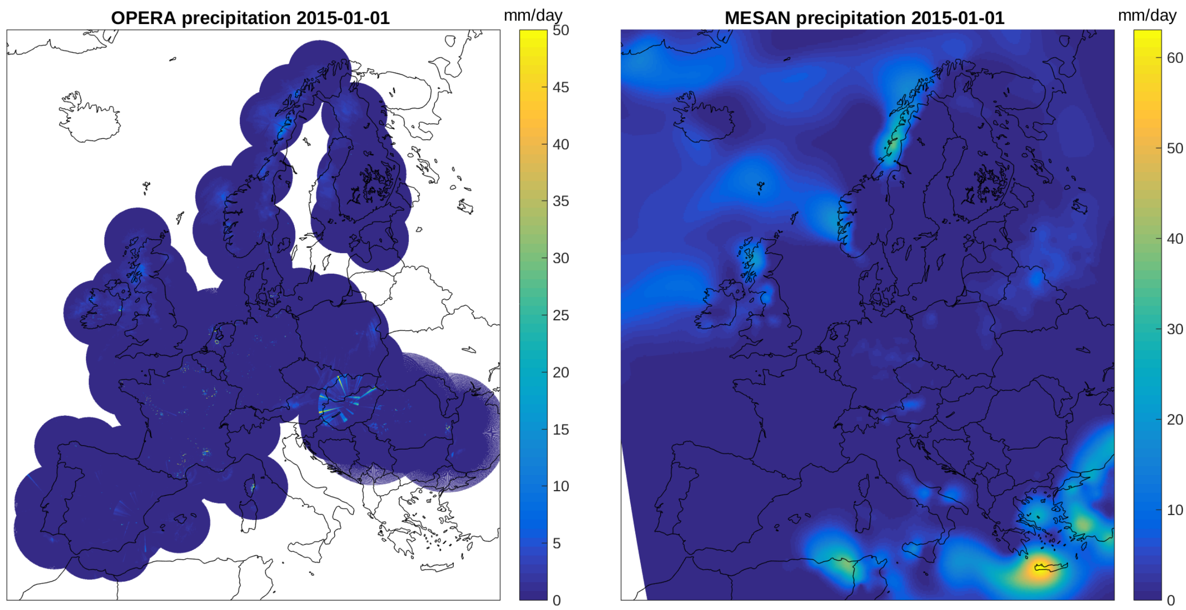


Figure 1: Accumulated precipitation 31 December 2014 0600 UTC – 1 January 2015 0600 UTC from OPERA precipitation intensity composites (left) and MESAN (right).

2.2 MESAN (reference)

The MESAN system has been used at SMHI for many years to provide hourly mesoscale analyses over Sweden (Häggmark et al., 2000). Since 3 December 2014 a second domain has been setup covering large parts of Europe with a spatial resolution of 11 km. The analysis of accumulated surface precipitation (0600 to 0600 UTC) is based on an optimal interpolation technique using ECMWF model data (background) and rain gauge observations as input. For comparison with the OPERA composites the accumulations are interpolated to the radar grid (Fig 1 right). MESAN is considered as an independent reference to evaluate the quality of the OPERA composites, as no radar information has been used as input. However, the quality of MESAN's precipitation analyses has not been assessed yet.

3 Methods

Several factors can deteriorate the quality of radar data, like non-meteorological echoes (caused by ground and sea clutter, anomalous propagation, interferences etc.), beam blockage, or overshooting.

In this study the quality of the OPERA precipitation intensity composites has been evaluated in two ways: using only radar data (section 3.1) and using a reference dataset (section 3.2). Figure 2 gives an overview of the processing steps for both approaches. Overall, three years of data have been analyzed (2015–2017) limited by the availability of the MESAN analyses.

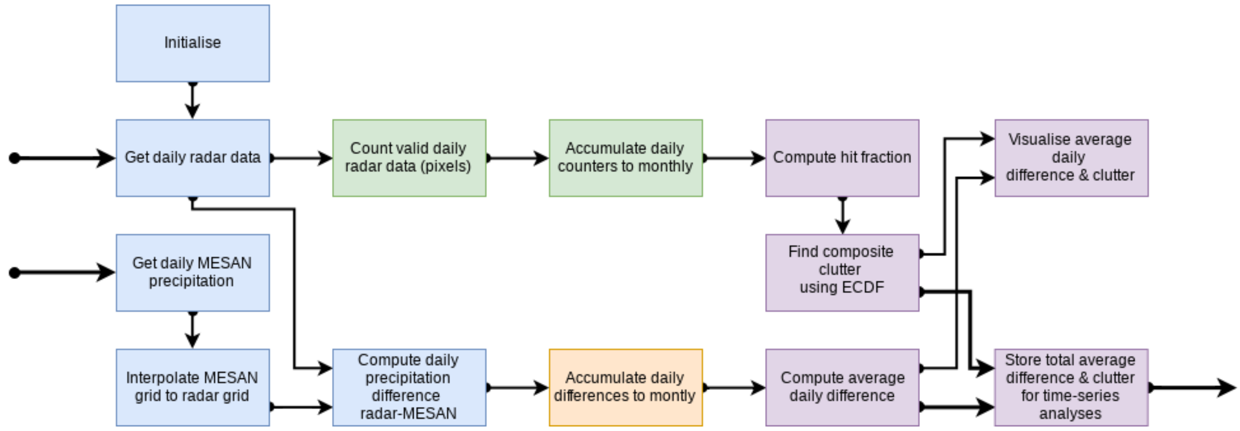


Figure 2: Monitoring the quality of OPERA composites based on hit fractions and precipitation intensity biases (OPERA–MESAN).

3.1 Monitoring the quality of composites without using reference data

The EUCOS Quality Monitoring Portal provides daily and monthly availability statistics on incoming radar data to Odyssey, processed radar data by Odyssey, and OPERA composites. If a radar has been used in a composite it should be listed in the `how/nodes` attribute of the corresponding ODIM_H5 file.

A simple approach to measure the quality of radar composites is to count how often each pixel is illuminated within a month (`NODATA` pixels are rejected). This number is then normalized with the expected number of composites per month. Note that a pixel can be covered by one or several radars. In Fig 3 only pixels exceeding a precipitation intensity threshold of 0.0 mm/h (left) and 0.1 mm/h (ca. 7 dBZ) (right), respectively, have been considered. Artefacts like interferences and ground clutter are best visible in Fig 3 (left) (e.g. over Spain and France). The high values along the Norwegian coast might be caused by persistent precipitation together with ground/sea clutter. Appendix A includes monthly hit fraction maps for 2015–2017.

In an earlier study Norin (2015) analyzed NORDRAD composites and scans from individual Swedish radars with respect to clutter. He developed an algorithm which identifies clutter automatically without the need for external data. It is based on hit fractions per pixel (f)

$$f = \frac{\text{Number of pixels greater than a defined precipitation intensity threshold}}{\text{Number of pixels unequal } \texttt{NODATA}}$$

and can be computed with the following steps:

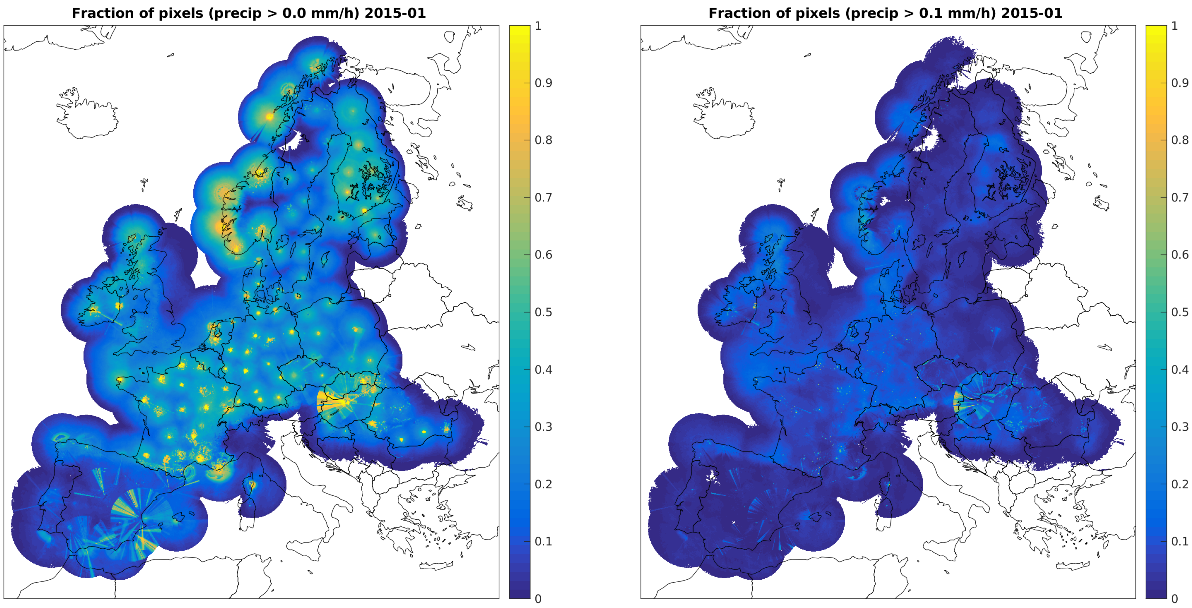
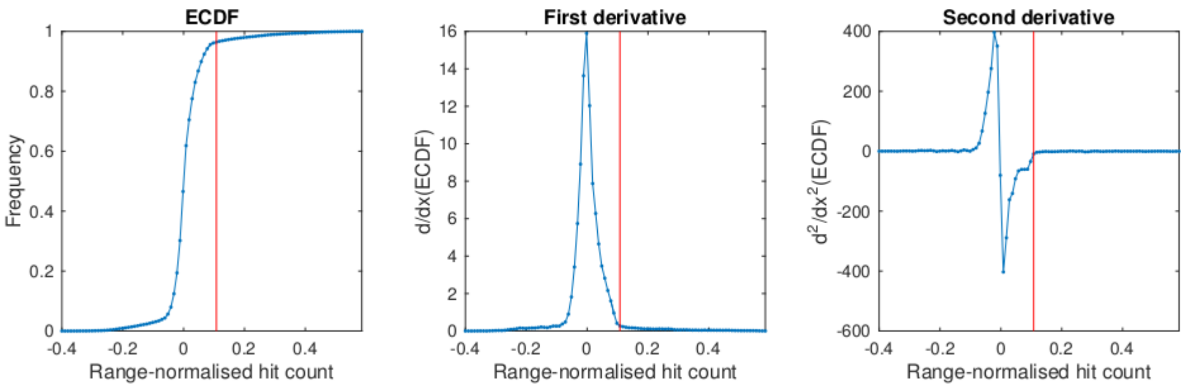


Figure 3: Fraction of pixels with precipitation intensities greater than 0.0 mm/h (left) and 0.1 mm/h (right) derived from all available OPERA precipitation intensity composites (January 2015).

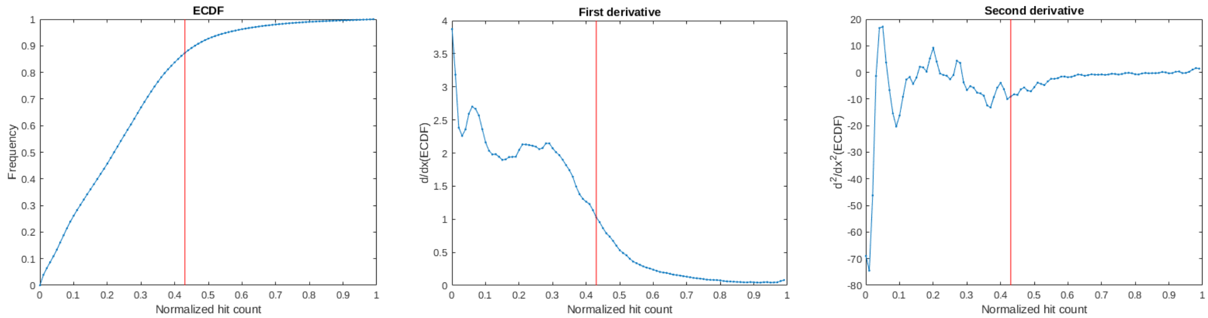
1. Remove pixels with a high f (0.99 in this study) as they are are most likely clutter-contaminated. This is mainly to reduce computational costs.
2. Normalize hit fractions by their azimuthal average for each range bin (applies only to scans).
3. Calculate the empirical cumulative distribution function (ECDF). Figure 4a (left) shows an example for radar Kiruna (Sweden).
4. Calculate the first and second derivative of the ECDF (Fig 4a center and right, respectively).
5. Find where the last (local) minimum of the second derivative is below a certain threshold (-10 in this study).
6. Find where the second derivative, after its last (local) minimum, exceeds the above mentioned threshold. The corresponding hit fraction is called the critical hit fraction. Pixels with hit fractions greater than the critical hit fraction are referred to as clutter.

We have applied the algorithm to OPERA precipitation intensity composites without adjustments. Figure 4b shows an example for January 2015. It is still possible to estimate critical hit fractions although the ECDF and its derivatives are less smooth than for individual scans (Fig 4a). This is probably due to the heterogeneous OPERA network including more than 160 radars from 25 countries and the heterogeneity of the precipitation climatology throughout Europe.

Figure 5 shows pixels in the OPERA precipitation intensity composite, where the fraction of pixels with precipitation intensities greater than 0.0 mm/h (left) and 0.1 mm/h (right) is greater than the corresponding critical hit fractions derived from the ECDF analysis (January 2015). The majority of the detected clutter disappears if the composite pixels are thresholded by 0.1 mm/h (Fig 5 right). Obviously the method fails in areas with persistent precipitation and/or low data availability. Erroneous flagged areas are e.g. visible over Iceland (July 2015 and December 2016), Norway (November and December 2015, December 2016, January, November, and December 2017), and France (March, April, and June 2015). Appendix A includes monthly



(a) 0.5° elevation angle scans for radar Kiruna (June 2014). Figures from Norin (2015).



(b) OPERA precipitation intensity composites (January 2015).

Figure 4: The empirical cumulative distribution function (ECDF) (left) together with its first (center) and second derivative (right). The red line indicates the critical hit fraction.

clutter maps for 2015–2017. The results indicate that complementary methods for clutter identification are needed.

Figure 6 shows the monthly critical hit fractions and total clutter fractions for pixels with precipitation intensities greater than 0.0 mm/h and 0.1 mm/h, respectively. By thresholding precipitation intensities by 0.1 mm/h, the critical hit fraction decreases. The same is valid for the total clutter fraction. It is clearly visible that the total clutter fraction for the 0.0 mm/h thresholded composites drops by more than 10% at the beginning of 2016 (after a temporary minimum in December 2015 related to a high critical hit fraction), which coincides with the implementation of a new clutter removal and beam blockage correction in Odyssey. It increased slightly in 2017. The total clutter fraction for the 0.1 mm/h thresholded composites is quite low and stable over time. This means that a majority of clutter has corresponding precipitation intensities of less than or equal to 0.1 mm/h.

In this section we presented an algorithm which identifies clutter automatically without the need for external data. It has proven to be a useful tool for site analysis but has difficulties in assessing an inhomogeneous network. However, major changes in data processing can be detected. Instead of running the ECDF analysis over the whole OPERA domain, it could be applied on more homogeneous subdomains. In a more conservative approach the critical hit fraction is assumed to be constant over time.

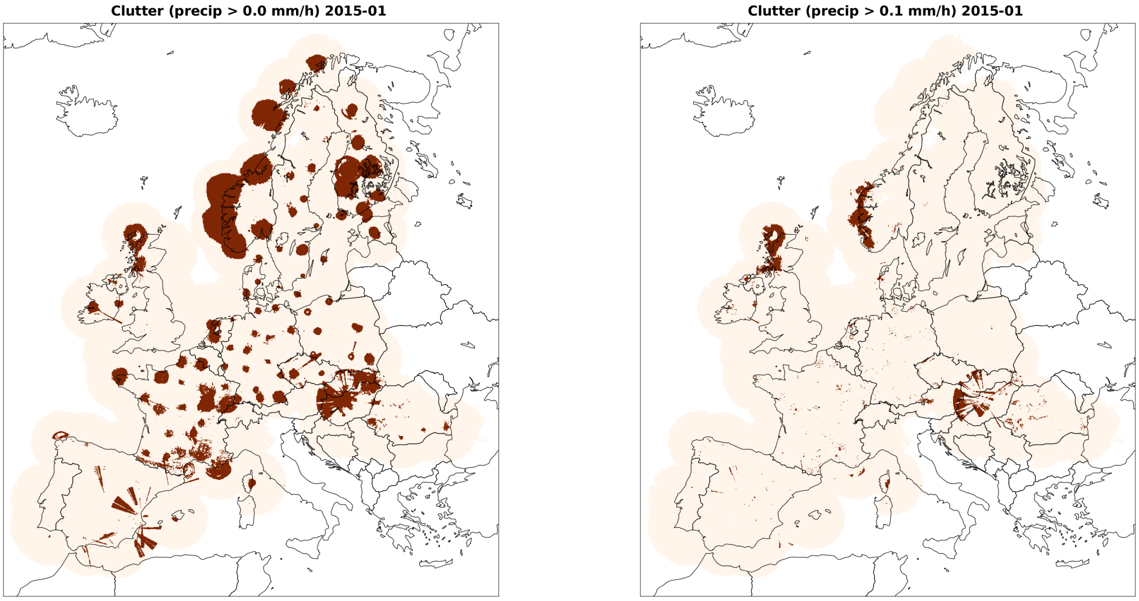


Figure 5: Pixels, where the fraction of pixels with precipitation intensities greater than 0.0 mm/h (left) and 0.1 mm/h (right) is greater than the corresponding critical hit fractions (43% and 20%, respectively), are marked in dark red. Pixels, where the fraction of pixels is less than or equal to the critical hit fractions are marked in light red (January 2015).

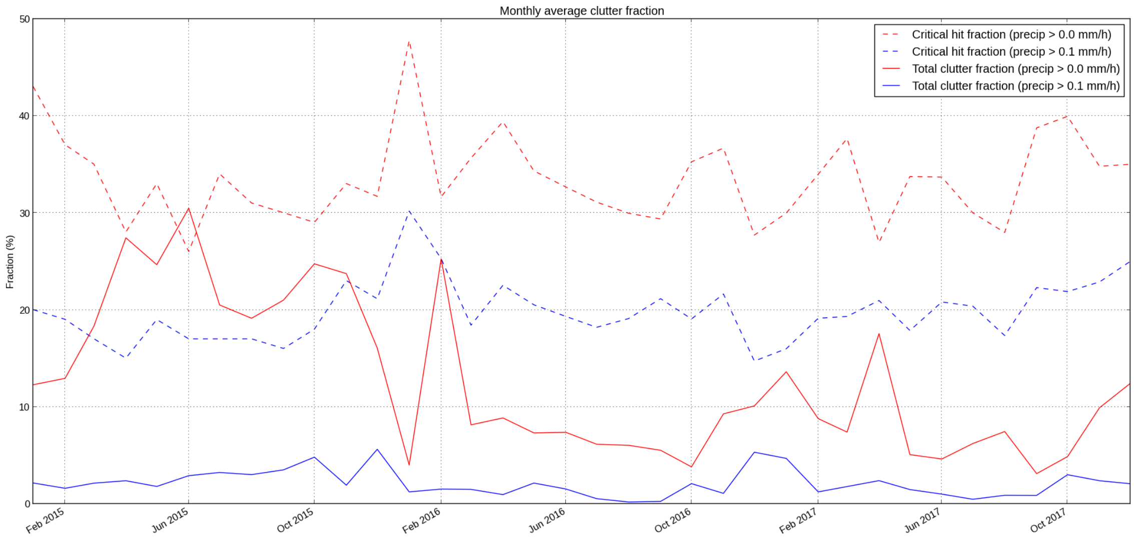


Figure 6: Monthly critical hit fraction (dashed lines) and total clutter fraction (solid lines) for pixels with precipitation intensities greater than 0.0 mm/h and 0.1 mm/h, respectively.

3.2 Monitoring the quality of composites using reference data

Another approach to monitor the quality of the OPERA composites is to use independent reference data, e.g MESAN (section 2.2).

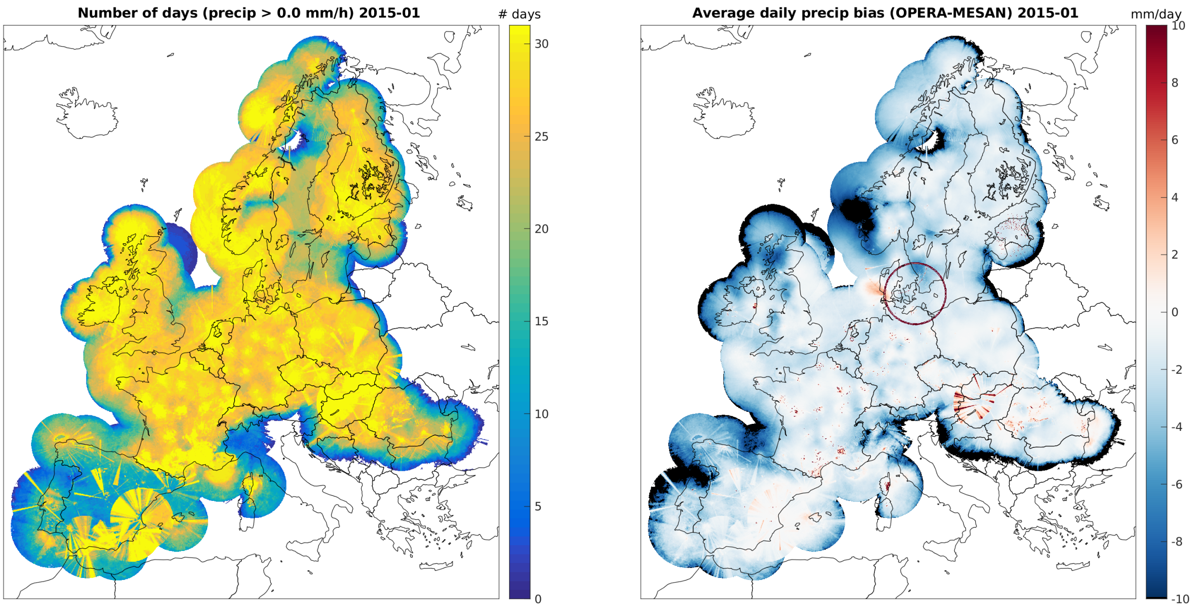


Figure 7: Number of days in January 2015, where at least one composite contains a pixel with a precipitation intensity greater than 0.0 mm/h (left). Daily precipitation intensity bias (OPERA–MESAN) normalized by the number of days, where at least one composite contains a pixel with a precipitation intensity greater than 0.0 mm/h (right).

Figure 7 (left) shows the number of days in January 2015, where at least one composite contains a pixel with a precipitation intensity greater than 0.0 mm/h. This dataset is then used to normalize the daily precipitation intensity difference (OPERA–MESAN). Figure 7 (right) shows the corresponding bias for January 2015. Appendix A includes monthly bias maps for 2015–2017.

In general, weather radars underestimate precipitation at the borders of the network. This is especially evident at Northern latitudes but also in the South in winter. On the other hand radars tend to overestimate precipitation nearby possibly due to ground clutter, e.g. in France. The bias depends on the national network and the season. There is a tendency of persistent overestimation of radar precipitation intensities in southeastern Europe, especially in summer.

The circle around radar Stevns (Denmark) is an artefact caused by an error in the radar processing software; it is only visible in Fig 7 (right). The high precipitation intensities appear probably only in a couple of composites per month and are therefore not visible in Figs 3, 5, and 7 (left). The circle vanished in September 2017. Another circle around radar Sindal (Denmark) disappeared in May 2017.

The interferences around radar Budapest (Hungary) are clearly visible in Figs 3, 5, and 7 (right), as they are strong and persistent. On the other hand, the interferences in Spain are quite weak but persistent. They are visible in both countries 2015–2017. The interference at radar Luleå (Sweden) is visible 2015–2017.

The sea clutter south of radar Collobrières (France) in Figs 3 (left), 5 (left), and 7 (left) is quite weak but persistent. There is also persistent precipitation together with ground/sea clutter along the Norwegian coast, which is much stronger.

Anomalous propagation (ducting) of radar Luleå (Sweden) caused high reflectivities at the Finnish west coast (August 2015).

Something strange happens with radar Fljotsdalsheidi (Iceland) April–December 2017. The artefacts at radar Keflavik (Iceland) and radar Riga (Latvia) are probably related to low data availability.

The beam blockage correction for radar Bømlo seems to overcompensate precipitation intensities (2017).

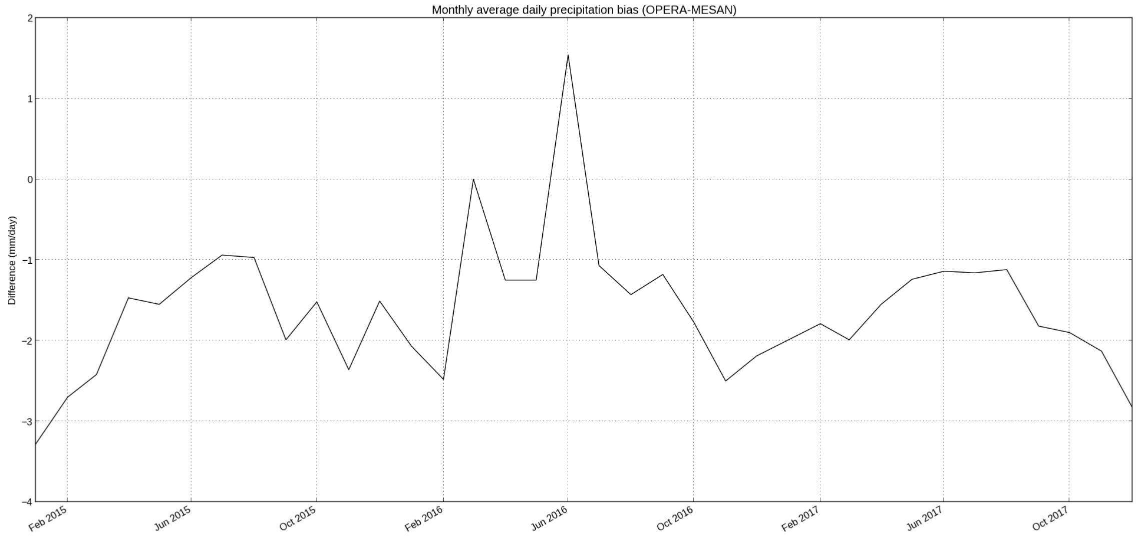


Figure 8: Monthly averaged daily precipitation intensity bias (OPERA–MESAN) over the OPERA composite domain.

Figure 8 shows a time series of the monthly averaged daily precipitation intensity difference (OPERA–MESAN) for the entire OPERA composite area.

The bias is mostly negative, indicating that OPERA underestimates precipitation intensity compared to MESAN. There are two peaks with a significantly higher bias, in March and June 2016. These are mainly caused by radar Ängelholm (Sweden) and radar Riga (Latvia), respectively.

The bias has its maxima in summer and its minima in winter, because beam overshooting is more common in winter.

The introduction of the new clutter removal and beam blockage correction at the end of 2015 had no impact on the bias (Fig 8), probably because the clutter echoes were too weak (see clutter maps for November 2015 in appendix A).

In this section we analyzed the bias between OPERA and MESAN. The monthly average of the daily precipitation difference clearly reveals a seasonal pattern. Additionally, the distribution of the bias through time should be visualized (e.g. by percentiles). As already suggested in section 3.1 a regional analysis of the bias might give more insight. Finally, it is recommended that erroneous incoming data to Odyssey should be flagged using the `how/malfunc` attribute and that Odyssey applies this flag when generating composites.

4 Conclusions and recommendations

In this study the quality of OPERA composites has been evaluated in two ways: using only radar data (section 3.1) and using a reference dataset (section 3.2). Both methods complement each other.

Here are some major conclusions and recommendations based on this study:

- The presented clutter analysis has proven to be a useful tool for site analysis but it has difficulties in assessing an inhomogeneous network. Obviously the method fails in areas with persistent precipitation and/or low data availability. However, major changes in the data processing can be detected. A temporal invariant (regional) critical hit fraction might solve this dilemma.
- It would be interesting to further investigate the latitudinal (regional) dependency of the precipitation intensity bias. In addition to MESAN there are other reference datasets available to evaluate the quality of the OPERA composites, e.g. a European daily high-resolution gridded dataset of surface temperature and precipitation[1] (Haylock et al., 2008) or the ERA5 dataset provided by ECMWF[2].
- The monthly hit fraction, clutter, and bias maps together with the corresponding time series plots presented in this study are useful tools to monitor the quality of the OPERA composites and should therefore be produced on a regular basis. They allow to identify anomalies (e.g. ground clutter, interferences, spatial discontinuities, and unrealistic patterns) and to evaluate changes in the Odyssey processing chain.
- The `QIND` information in the (ODIM_H5) composites should also be evaluated. More quality fields might need to be added to the composites.
- Finally, it is recommended that erroneous incoming data to Odyssey should be flagged using the `how/malfunc` attribute and that Odyssey applies this flag when generating composites.

[1] `http://www.ecad.eu/download/ensembles/ensembles.php`
[2] `https://software.ecmwf.int/wiki/display/CKB/ERA5+data+documentation`

5 Acknowledgments

The authors would like to thank David Dufton (SMHI) for the fruitful discussions and Susanna Hopsch (SMHI) for the MESAN support. Thanks also to Laurent Delobbe (RMI) for his very constructive review.

References

L. Häggmark, K.-I. Ivarsson, S. Gollvik, and P.-O. Olofsson. Mesan, an operational mesoscale analysis system. *Tellus*, 52A:2–20, 2000.

M. R. Haylock, N. Hofstra, A. M. G. Klein Tank, E. J. Klok, P. D. Jones, and M. New. A European daily high-resolution gridded data set of surface temperature and precipitation for 1950–2006. *Journal of Geophysical Research: Atmospheres*, 113(D20):n/a–n/a, 2008. ISSN 2156-2202. doi: 10.1029/2008JD010201. URL `http://dx.doi.org/10.1029/2008JD010201`. D20119.

S. Matthews, P. Dupuy, R. Scovell, A. Kergomard, B. Urban, A. Huuskonen, A. Smith, and N. Gaussiat. EUMETNET OPERA Radar Data Center: Providing operational, homogeneous European radar rainfall composites. In *Proc. Weather Radar and Hydrology Symp., Exeter, United Kingdom, IAHS, 9–14.*, 2012.

L. Norin. Kvalitetsövervakning av radarprodukter. Technical report, Swedish Meteorological and Hydrological Institute, 2015.

A Quality analysis

The complete report including the appendix is available at SMHI. Please contact Günther Haase (gunther.haase@smhi.se).