General comments

The aim of the paper is clear. It is validation of satellite OPEMW precipitation product with ground-based weather radar and rain gauges network in Italy. The study is interesting both from the scientific as well as the practical perspective mostly because of the spatial and the temporal scales of it. Validation is made for the territory of all Italy and based on annual datasets derived from OPEMW products, 20 weather radars and more than 3000 rain gauges. Considerable size of database and its processing burden shall be appreciated.

Specific comments

1) Results of OPEMW sri products comparisons with RGN sri and RNC sri are discussed mainly from the statistical perspective. It means that visible scatter of points for sri >7-10mmh^{-1} in Fig. 7-9 is explained by the low number of cases (as seen in Fig. 6). It is possible good but not complete explanation of results. Most probably the discussion of results should be enhanced by considering the mechanism of precipitation. Maybe PEMW algorithm works better for stratiform precipitation than for convective rainfalls which somehow explains better performance for small sri values bins up to about 7mmh^{-1} and for winter periods when convection is rarely observed (see Fig. 8&9)? Maybe also PEMW algorithm works better for stratiform precipitation than for orographic rainfalls which to some extent explains the location of red areas in Fig. 10&11 over Alps region? Diagnosis of the scale of precipitation type influence on PEMW algorithm performance is probably hard to be done, especially for the whole database. So consider this comment mainly as suggestion, possibly for future research. However maybe it is possible to test this hypothesis at least for a sample of data.
2) The spatial-temporal analysis of the results for OPEMW and ground-reference (either RGN or RNC) sri products showed an increase in mean absolute difference over the Alps and along the northern Apennines during winter (see Fig. 10&11). This was related to the presence of snow on the ground, which is a well-known source of uncertainty for passive microwave. This is also probably not the complete explanation. Most probably the performance of rain gauges networks in mountainous regions of Italy is strongly affected by the form of precipitation (snow instead rainfall) and air temperature. It could be expected that the most gauges used in Italy are still tipping buckets. For this type of gauges direct snow precipitation measurements are impossible and melting of snow is necessary at their orifices (heating of snow and ice leads to evaporation losses). This could be the source of substantial measuring errors. Another question is if the density of rain gauges networks over the Alps and along the northern Apennines is the same as in other parts of Italy. There is no information about the gauges density distribution over Italy. Maybe "*larger errors affecting the rain gauges deployed in Sicily*" could be explained by instrumental differences (older type, less precise gauges) and/or sparser gauges density? Finally some focus could be given on the uncertainty of radar measurements of snow and snow&rain precipitations over mountainous regions.
3) The use of adjective: "significant" in numerous parts of manuscript should be reconsidered. Especially the phrase "significant difference" suggests that some statistical hypothesis was tested and for example rejected at some probability level. See page 4295 or 4292:

*Here we do not notice **significant difference** (visible diference?) between the four seasons, all of which tend to confirm the results in Fig. 7.*

4) It is proposed to add min-max. or standard deviation (error bars) for OPEWM sri values in Fig. 7-9. It should be done at least for Fig. 7 to display the ranges of OPEWM sri values estimated for the following bins of RGN and RNC sri products. In order to keep Fig. 7 visibility it could be composed of four plots derived separately for RGN and RNC sri products and red/blue markers.
5) The summary and conclusions paragraph should be rewritten. At current form it seems to be only the study summary. Some sentences seems to be almost copied from the body of former chapters.

Technical corrections

Page 4286:

The data set considered here covers one full year (July 2011–June 2012); data were processed for searching space-time 5 colocation, ensuring data quality, and finally computing statistical scores for quantitative performance evaluation.

Sentence is too long and its meaning is not clear.

Page 4287 – this sentence is not clear.

*Similarly, FOVs with less than 10 rain gauges are discarded, and only those for which more than 95% of the rain gauge (*gauges???*) detects either rain or not rain are considered.*

Page 4288:

….the contingency table reports the hits, misses, correct negatives, and false alarms of OPEMW rain detection (0/1 for rainy/non-rainy) with respect to RNC and RGN sri.

Sentence is not clear. Please explain the meaning of *correct negatives.* Are correct negatives the same as null (N) events – see Appendix A.

Page 4290:

Use of "–" suggests the range of values, which is not the case, as for example:

The perfect value for HHS and ETS is 1.0 while here these get to HHS = 0.42–0.45 and ETS = 0.27–0.29 (the first number being with respect to RGN while the second to RNC).

Page 4290:

This demonstrates the increasing OPEMW detection skills as the rainfall becomes more and more significant.

Maybe it is better to write: *…the rainfall becomes more intense* or *…rainfall intensity increases*

Page 4293:

*which is a well-know (*well-known?*) source of uncertainty for passive microwave estimates of rainfall.*

Page 4296:

…and it is the ratio of the estimated to observed rain areas, thus indicating whether there is a tendency to over or underestimate…

Please, change *estimated* to detected to be in agreement with the rest of paragraph.