Comparison between different QPE based on: Microwave Links, Radar adjusted and Gauges.

Pier Paolo Alberoni[1], Anna Fornasiero[1], Giacomo. Roversi[2,3], Stefania Pasetti[2], Marco Folegani[2], Federico Porcù[3]

[1] *ARPAE Emilia Romagna, Italy*
[2] *MEEO S.r.l., Italy*
[3] *Department of Physics and Astronomy - University of Bologna, Italy*
(Dated: 22 June 2018)

1 Introduction

Currently the most used tools to estimate quantitative precipitation are rain gauges and weather radars, both with relative merits and defects. Rain gauges measurements are not reliable in case of intense precipitation events and have poor spatial representativeness due to high spatial and temporal variability of rainfall field. Weather radars monitor rainfall over large areas with high spatial resolution, but through indirect measurements and affected by various sources of error.

A complementary approach is represented by a technology, not yet been tested in Italy, that consists in the extrapolation of rainfall intensity data from cellular communication infrastructures (i.e. in the identification of rain-induced attenuation of microwave links), usually called Microwave link (MWL) networks or Commercial Microwave links (CML).

A validation activity of these different QPE – Quantitative Precipitation Estimates technologies was done over the Bologna province area (Italy) on the period February 20th - June 30th 2016 in the framework of the RainBO project LIFE15/CCA/IT/000035[1].

An accurate validation exercise was carried out by comparing microwave links QPE (MWL QPE) and radar adjusted QPEusing, as reference field, a Rain gauges measurements spatial analysis (ERG5, a modified Shepard interpolation).

The validation activity was developed in two steps: 1) a Qualitative analysis, mainly based on many daily accumulation maps comparisons; 2) a Quantitative analysis, carried out by using statistical indexes..

In the first phase, a visual inspection of the rainfall fields in the datasets has been done in order to understand the general behavior and to identify, if possible, any problems or drawbacks in the collected data set. A synoptic comparison was conducted by plotting together the rainfall estimates obtained from remote sensing tools (i.e. Microwave links and radar) with the measured reference field (i.e. rain gauges). In the second validation phase, the quantitative one, statistical indicators based on contingency table and continuous verification methods were calculated in order to estimate the reliability of the microwave link and the radar QPE methods.

Additionally some single event analysis was conducted in cases of false precipitation detection by MWL when compared to gauges, as in the case of abnormal attenuation detection by MWL due to the presence of snow or mixed-phase precipitation and a discussion of under sampling of precipitation in the convective season.

2 Validation datasets

The validation has been carried out comparing the rainfall obtained by Microwave links with other rainfall estimates available over the project domain. Further in detail, Microwave QPE have been compared with radar QPE, both raw and Gauge-corrected, rain gauges and the operational precipitation analysis (ERG5) made available by Arpae.

2.1 Microwave Links (MWL) data

Microwave data were collected as historical dataset, starting from February 2016 to June 2016, from two providers: Lepida and Vodafone. Lepida provides also real-time data, while Vodafone is evaluating the possibility to provide a short sample period of 2017, simulating the real-time data that can be provided in case a customer will finance an operative service.

The operative frequency of Vodafone links has a large range, for this reason the frequency window required by original Rainlink algorithm was extended from [12.5 - 40.5] GHz to [5.0 - 45.0] GHz in order not to miss any available information. The Lepida links work at a fixed frequency of 26 GHz.

The pathlength of the MWL varies fairly but most of these links are under 13 km and the average length is 5.6 km.

As the algorithm estimates the rainfall averaged along the link path, we expect an error not lower than the average length of the links. This interval can be even larger in case of less dense links, so we decided to increase the grid resolution from 1 km (the original one used by Wageningen University[2]) to 5 km, in order to filter any possible interpolation’s artificial

[1] https://www.rainbolife.eu/
[2] WUR - https://www.wur.nl/en/Persons/dr.ir.-A-Aart-Overeem.htm

effect. The interpolation grid adopted for the RainBO project is the one already in use by Arpae for ERG5 precipitation analysis.

Table 1:Microwaves link characteristics

	Lepida	**Vodafone**
Number of links	37	422
Coverage area	Bologna Apennines	Bologna metropolitan area
Historical data period	Starting from 28th March 2016 – till today	From 20th February 2016 to 30th June 2016
Temporal frequency	1 minute	15 minutes
Link average length	7.4 km	5.4 km
Frequency range	26 GHz	5-45 GHz
Time format in the original dataset	UTC	UTC+1
ACM – Attenuation Coefficient Measurement	Not used	Active for some links

2.2 Rain-gauge data

Rain gauge hourly data are provided by the Arpae Rirer (regional hydro-meteorological) network, established in 2001 by bringing together existing hydrological and meteorological station networks managed by different public bodies, operating all over the regional area. The network is composed by 498 stations with circa 900 sensors.

Real-time data are collected in the central data archive mainly by radio-link and are available for operational activities. Figure 1 shows the Rirer network and the location of the two radars which form the regional radar network.

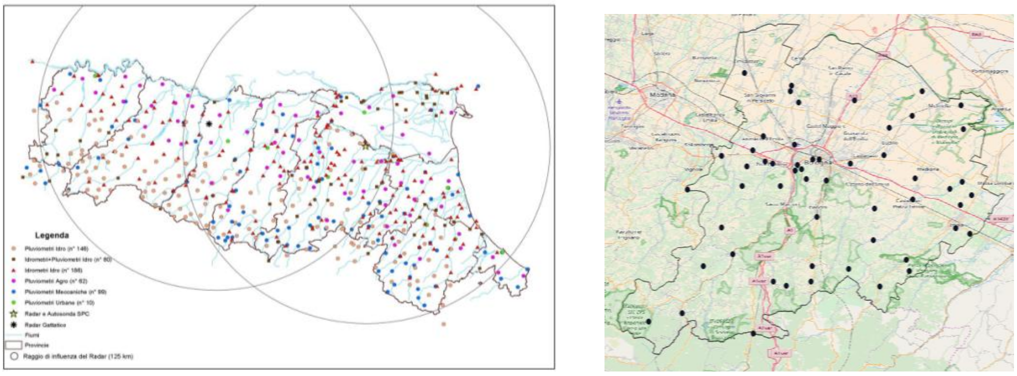

Figure 1: The Rirer network (left), Distribution of the Rirer gauges in the validation area (right)

2.3 Radar data

Radar dataset is based on hourly precipitation estimations obtained from the composit of the regional radar network managed by Arpae-SIMC. The regional network is composed by two C-Band systems, one located at San Pietro Capofiume (Bo) and the other in Gattatico (RE) (Figure 1). For each radar the precipitation rate at the ground is then interpolated from polar coordinates to a 256 X 256 Cartesian grid of 1 km x 1 km resolution, then single rain-rate estimates are merged to obtain a composite of both radar. Rain rates are obtained every 5 minutes and the final rain total over a 1-hour period is computed by an advective algorithm taking into account the movement of precipitating systems. The algorithm is based on the computation of maximum cross-correlation between consecutive maps, leading to the estimate of the displacement vector for each precipitation system. The precipitation field is then reconstructed every minute between the observed maps and finally cumulated over one hour. Finally radar QPE is adjusted with rain gauges data.

2.4 Erg5: rainfall analysis

The ERG5 gridded meteorological data set has been developed by Arpae-SIMC, in order to support agricultural services of Emilia-Romagna. ERG5 data are operationally produced since 2001, interpolating the hourly station data for the main meteorological variables (air temperature, relative humidity, precipitation, wind, solar irradiance) onto a 5x5 km grid covering the Emilia-Romagna region. Within the RainBO project, only ERG5 hourly cumulated precipitation data are considered. With respect to this parameter, ERG5 input data consists of all hourly cumulated precipitation data recorded by the Rirer network. All data are daily checked for their quality and undependable hourly data are not used as input of the analysis. The interpolation method used for hourly precipitation consists of a Shepard modified scheme using topographic distances instead of Cartesian distances. This allows the interpolation to take into account the influence of topography on precipitation, by making locations divided by orographic obstacles more distant than they would be if Cartesian distances were used. For project validation purpose only the erg5 grid point which are within the Bologna province has been used.

3 Qualitative analysis: daily maps

As stated above, daily rainfall amounts of MWL based estimation were visually compared with the others available estimates (radar, calibrated and uncalibrated, and ERG5), in order to get an idea on the quality of precipitation estimates retrieved by MWL. The ERG5 analysis was used as reference.

The daily rainfall pattern retrieved by MWL, as well as the hourly pattern, generally shows a reliable structure that sometimes reproduces the pattern features (rainfall peaks and gradients) observed by the others measuring tools. We state that the agreement between MWL estimates and the other ones is generally promising, even though some drawbacks still need to be addressed. These drawbacks cause a variable time matching of the different precipitation estimates. It is also important to highlight that the performance of the Microwave estimates seems to increase in the second half of the analysed period. The main reason lies, maybe, on the types of clouds producing precipitation. Stratiform precipitations, lasting for hours but mainly with moderate or light rain, were observed in the winter period. The spring period is mainly characterized by convective storm, even embedded in frontal structures, with showers and heavy rain. As the MWL estimate is based on the attenuation occurred along the link path and it is clear that attenuation is strongly related to rainfall intensity and that the signal is consequently stronger in the convective season. This behaviour is sometimes counter-balanced by the difficulty to detect scattered or isolated convective storm; we will further showcase some examples of isolated storms that not intersect, with its core, any microwave link in the project domain.

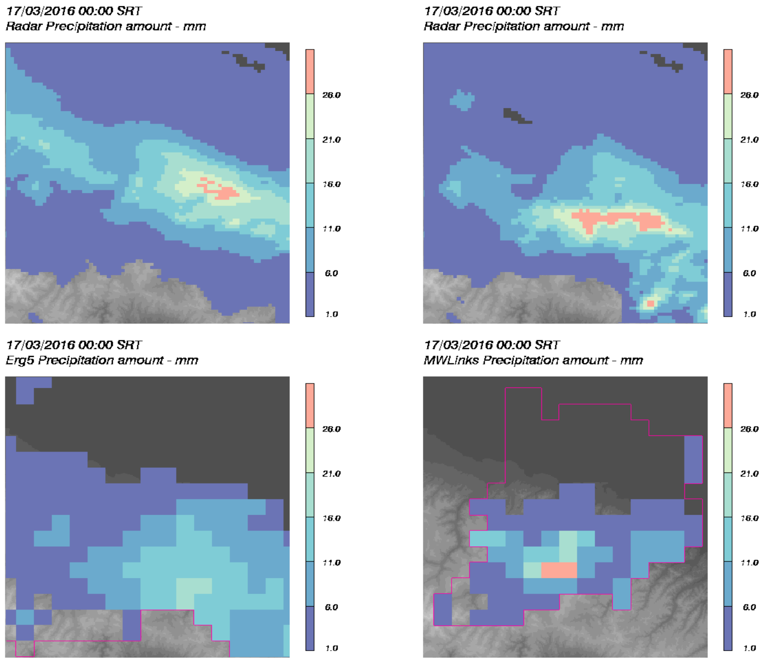

Figure 2: The Quantitative precipitation estimates accumulated from 16/03/2016 00 UTC to 17/03/2016 00 UTC. (Top Left) Radar QPE, (Top Right) Radar adjusted QPE; (Bottom Left) Erg5; (Bottom Right) Microwave Links QPE

On 16 March precipitation is mainly located over the Bologna hilly area, with a local maximum shown in Fig. 2. Both remote sensing methodologies identify correctly this feature, radar estimates spread light precipitation in the flat area north of Bologna while erg5 and MWL does not detect such rainfall.

Excellent results are achieved also in a more convective event such as the one displayed in Fig 3. On 11 May precipitation occurred with a well defined gradient in the West-East direction and some local maxima in the North-West and South-West part of the province. Microwave accumulation slightly underestimate the rainfall field while an overestimation is recorded in the non adjusted radar. Adjustment procedure well calibrates radar data as displayed in the top right panel of Fig. 3. Fine scale structure of daily amount is well detected in both remote sensing maps.

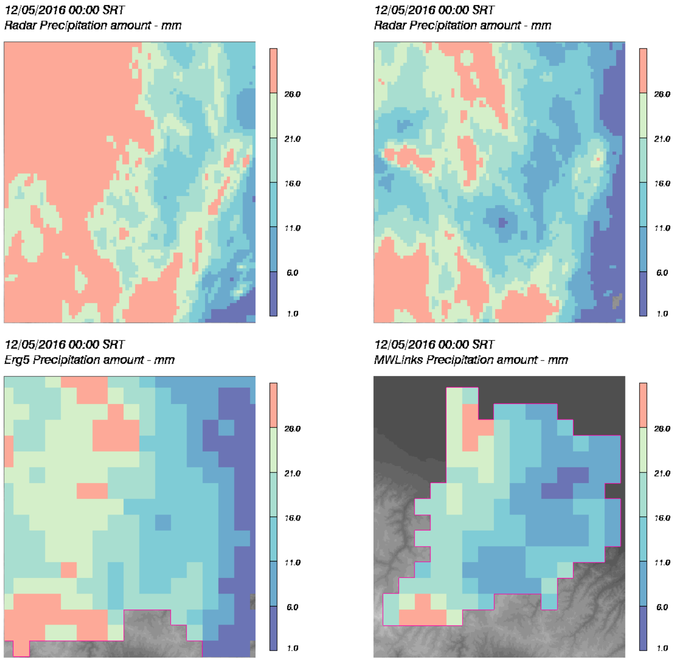

Figure 3: The Quantitative precipitation estimates accumulated from 11/05/2016 00 UTC to 12/05/2016 00 UTC UTC. (Top Left) Radar QPE, (Top Right) Radar adjusted QPE; (Bottom Left) Erg5; (Bottom Right) Microwave Links QPE

Another important result achieved through the qualitative analysis is the need of an enhanced quality check for the Microwave power data. In the punctual analysis carried out on the Budrio station (not reported here) we have found that the link is affected by an antenna structural deformation due to sun heating and/or wind pressure which caused a misalignment between antennas link and consequently a lack of power which is interpreted as an attenuation event. Malfunctioning effects are more frequent and stronger on this link, though we noticed that less severe effects are or could be present in other links. We evaluate that this problem affects roughly 10% of daily maps.

4 Quantitative analysis

Categorical statistics and continuous verification scores are calculated on time series of hourly precipitation distribution parameters (one for all areal mean) over the Bologna metropolitan area and estimated by radar, radar-adjusted (ADJ), Microwave links (MWL) and rain gauges, then compared with erg5 (used as reference field).

4.1 Categorical score verification

The scores are calculated using hourly mean over the area, either on the whole period and in rainy hours (i.e. where no precipitation has been detected; no rainy hour is defined as both hourly mean precipitation = 0 mm for Erg5 and hourly mean precipitation < 0.1 mm for the other estimates). This is done to remove the no rainy event, which is most frequent one (about 80%), from the dataset, focusing then the score computation on the rain occurrence. Categorical scores are computed for two different thresholds. The lower threshold distinguishes between Rain/no Rain (of course for the filtered distributions where no rainy hours has been removed this threshold distinguishes between Rain/VeryLightRain). The upper threshold focuses on the identification of heavy rain events. For the whole period and for the filtered subset the higher threshold was set to 2 mm. In the framework of the whole validation period (Table 2), a comprehensive reading of categorical statistics shows that at the end of the day both adjusted radar and MWL estimates show a comparable skill in the precipitation estimation. A deep analysis indicates that such result comes from two different behaviours: on the one end radar tends to overestimate precipitation (in terms of intensity and number of hours) while MWL tends to underestimate precipitation (in terms of intensity and number of hours).

In order to highlight the effect of different precipitation regimes on the categorical scores, we split the dataset into two sub-periods, the first one for the 'stratiform season' from February 1 to April 15 (Table 3) and the second one for the 'convective season' from April 16 to June 30 (Table 4). This splitting may be questionable, but we deem it very useful to explain the different estimation trends we found. Regarding the first period (Table 3) we duplicate the computation of categorical score removing a melting-layer event from the dataset: therefore we added an additional column to the previous table. For simplicity we refer to the first period as “winter” and to the second one as “convective”.

In these period we observe an opposite behaviour.

In winter time radar over-performs in the low threshold if compared to MWL. A general overestimation from radar is observed. Regarding MWL, we notice the effect of sampling the snow-melting also in terms of validation score as well. By lower threshold the 'cleaned dataset' performs worst results. One possible explanation is that by lower threshold, the hours not considered because of the misleading snow-melting effect, are interpreted as detected phenomena (in the contingency table) regardless the absolute hourly amount (i.e. both observation, erg5, and estimation, MWL, was above the threshold). Looking to higher threshold the rejection has a huge effect: FAR falls to 0, BIAS decreases from 1.53 to 0.79 and Threat Score raises from 0.41 to 0.67. With this threshold, data are wrongly registered as detected phenomena from MWL but not from Erg5, so we have removed a false alarm event. This analysis makes clear that we need to tackle such phenomena before to move to a real-time or near real-time use of MWL.

The performance of the MWL improves during the convective season and from low to high threshold. This is probably due to the better response of MWL to intense events. In the same time radar scores dramatically drops in convective season. Table 4 shows radar ADJ Threat Score falling around 0,5 from the 0,70 winter value. This worsening seems due to the growing of FAR (reflected also in BIAS), while POD has a positive score.

These results make clear the issue about the capability of rain gauges (and as consequence of erg5) and MWL network to sample small-scale phenomena such as thunderstorms captured by radar instead. Therefore radar shows a 'probably false' overestimation when compared to erg5 probably due to convective cores which are not detected by the rain gauges network. MWL, as rain gauges, under-samples convective cores and therefore the areal mean is lower.

Table 2: Categorical statistics scores on the filtered dataset (no rainy hours removed)

	Threshold	Radar	ADJ	Rain gauges	MWL
N. Samples		702	702	586	280
FAR	0,5	0,39	0,37	0,06	0,09
BIAS	0,5	1,57	1,53	0,996	0,67
POD	0,5	0,95	0,95	0,94	0,61
TS	0,5	0,59	0,61	0,89	0,57
FAR	2	0,58	0,39	0,12	0,08
BIAS	2	2,25	1,55	1,05	0,89
POD	2	0,95	0,95	0,93	0,82
TS	2	0,41	0,59	0,82	0,77

Table 3: Categorical statistics scores on the filtered dataset (no rainy hours removed). Period: 1 February- 15 April. Cleaned dataset means MWL discharging period 01:00 – 13:00 UTC of 03/03/2016 .

	Threshold	Radar	ADJ	MWL	'Cleaned ' MWL
N. Samples		347	351	101	87
FAR	0,5	0,27	0,26	0,02	0,03
BIAS	0,5	1,32	1,28	0,59	0,52
POD	0,5	0,96	0,95	0,58	0,5
TS	0,5	0,70	0,71	0,57	0,49
FAR	2	0,5	0,30	0,52	0,09
BIAS	2	2	1,44	1,53	0,79
POD	2	1	1	0,73	0,71
TS	2	0,5	0,70	0,41	0,67

Table 4: Categorical statistics scores on the filtered dataset (no rainy hours removed). Period: 16 April-30 June.

	Threshold	Radar	ADJ	MWL
N. Samples		355	351	179
FAR	0,5	0,50	0,48	0,13
BIAS	0,5	1,91	1,86	0,72
POD	0,5	0,95	0,96	0,63
TS	0,5	0,49	0,51	0,58
FAR	2	0,67	0,49	0,06
BIAS	2	2,61	1,70	0,78
POD	2	0,87	0,87	0,74
TS	2	0,32	0,48	0,71

4.2 Continuous score verification

Continuous verification is focused on the distribution of a continuous variable without performing any stratification as for categorical verification. Tables 5-7 report the mean error, the mean absolute error and the root mean square error for the whole period, the filtered rainy hours sub-set, and the two sub-periods: winter and convective. As already stated, overall behaviours (Table 5) are equivalent for adjusted radar and MWL, with a slightly underestimation (in hourly amount this time) for MWL and a corresponding overestimation for ADJ. Absolute errors and RMSE are very similar.

Table 5: Continuous scores on the filtered dataset (no rainy hours removed). Whole period.

	Radar	ADJ	Rain gauges	MWL
ME	0,506	0,282	0,017	-0,006
MAE	0,570	0,332	0,104	0,532
RMSE	0,908	0,503	0,163	1,064

If the two sub-periods (winter and summer) are considered, an opposite behaviour is registered between them.

In winter time (Table 6) radar ADJ get best scores respect to MWL, even considering the “cleaned” subset. Removing the snow-melting hours highlight the more pronounced underestimation where the mean error (ME) is -0.194 but also the greater (compared to radar ADJ) variability, as could be deducted from RMSE. We point out that MWL is much better than the raw radar estimate, this means that radar needs to be gauge-calibrated.

Table 6: Continuous verification scores on the filtered dataset (no rainy hours removed). Period: 1 February- 15 April. Cleaned dataset means MWL discharging period 01:00 – 13:00 UTC of 03/03/2016 .

	Radar	ADJ	MWL	'Cleaned MWL
ME	0,412	0,179	0,269	-0,194
MAE	0,496	0,241	0,841	0,461
RMSE	0,826	0,319	1,637	0,594

Convective season (Table 7) is characterized by a good estimation by MWL. An underestimation, for MWL, is still present, indeed the mean error is negative. We observe a worsening in the radar indexes. This depend to a greater variability in the rainfall pattern, due to the fine scale structure of precipitation field, which is better sampled by the radar.

Table 7: Continuous verification scores on the filtered dataset (no rainy hours removed). : 16 April-30 June.

	Radar	ADJ	MWL
ME	0,606	0,385	-0,160
MAE	0,642	0,423	0,358
RMSE	0,981	0,635	0,509

5 Single event based analysis

The statistical analysis carried out in the previous sections is not sufficient to present and analyse some data peculiarities needed to get a more complete understanding of the of the behaviours of the different precipitation estimates. Then we run an ad-hoc analysis on each event where different problems occur.

5.1 Detection of abnormal attenuation by MWL

On March 2nd a huge overestimation occurs on MWL estimate, as briefly previously introduced and further underlined in the quantitative analysis both with categorical and continuous scores.

The hourly precipitation analysis shows the passage of a meteorological system heading West, East. Precipitation is then detected and recorded in the datasets. Fig. 4 shows two representative hours. It is evident that MWL strongly overestimate precipitation in the Apennines area. Local hourly maximum reach 40 mm for MWL whereas the other estimate are below 5 mm (Fig. 5). We point out that MWL estimate is based on the measure of power attenuation along the link path. In this area links are normally placed on the top of hills and mountains, their path is free of any obstacle, and they can be located 10 or 100 meters off the ground. MWL detect a different hydrometeors in comparison with the ones collected by gauges (on the ground). the Radar, instead, will sample higher atmospheric volume, in order not to be contaminated by ground return.

Fig. 6 shows instantaneous temperature and hourly precipitation, for two Apennines stations (Sasso Marconi and Vergato). It is clear that the temperature close to the ground falls around 0°C during the precipitation event. We can infer that there was rain at ground, mixed-phase precipitation with the melting of snowflake along microwave link path and snow at higher levels where radar samples are taken. Microwave attenuation depends on the third power of the hydrometeors dimension: snowflake and aggregates have a huge horizontal dimension in the melting layer but, from an electromagnetic point of view, they act as a drop of water causing a big attenuation. This effect is similar to the well known "bright-band" in the radar meteorology.

Such phenomena occurred on March 3rd and other times, even if we are not able to demonstrate it. Similarities occur in other events as well, even the effect is not so strong.

A quality control is needed to tackle this problem in real-time operation.

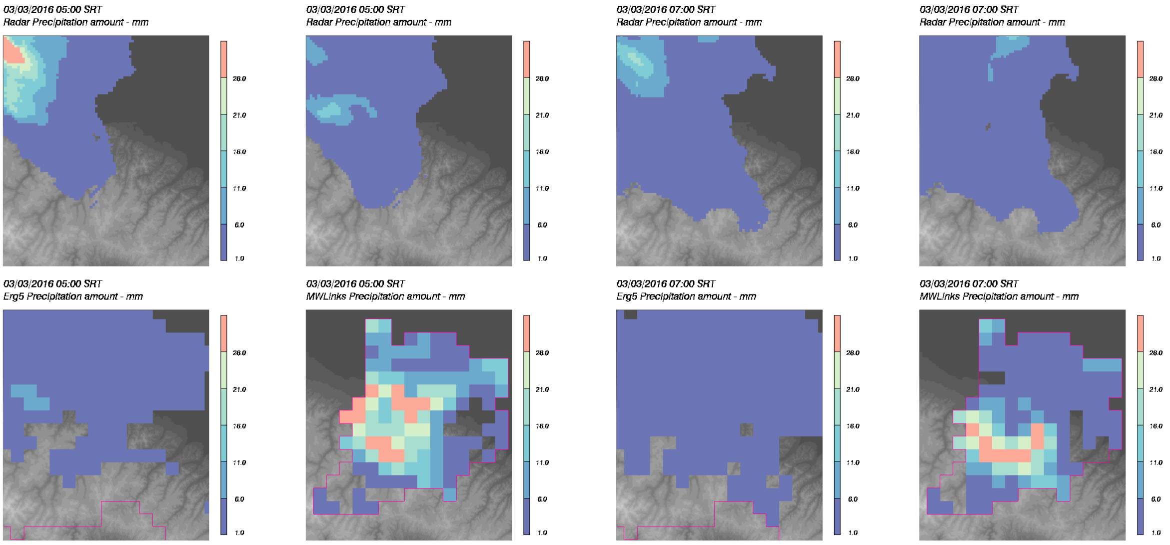

Figure 4: Quantitative precipitation estimates accumulated at 05_00 UTC (Left) and 07:00 UTC (Right) on 10/03/2016. Each group is composed by four panels : (Top Left) Radar, (Top Right) Radar adjusted; (Bottom Left) Erg5 ; (Bottom Right) MWL.

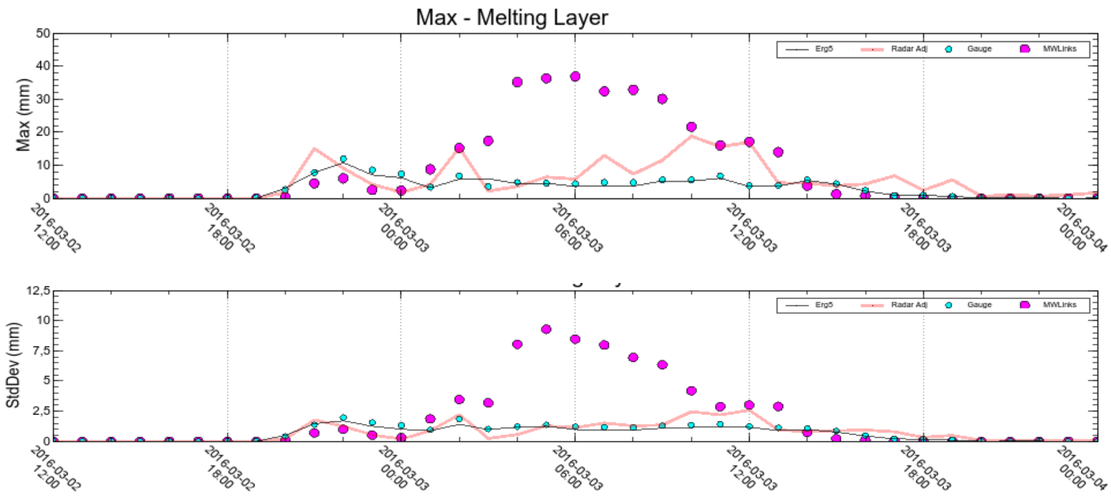

Figure 5: Time series of Hourly maxima (Top Panel) and Standard Deviation (Bottom Panel) from 02/03/2016 12:00 UTC to 04/03/2016 00 UTC

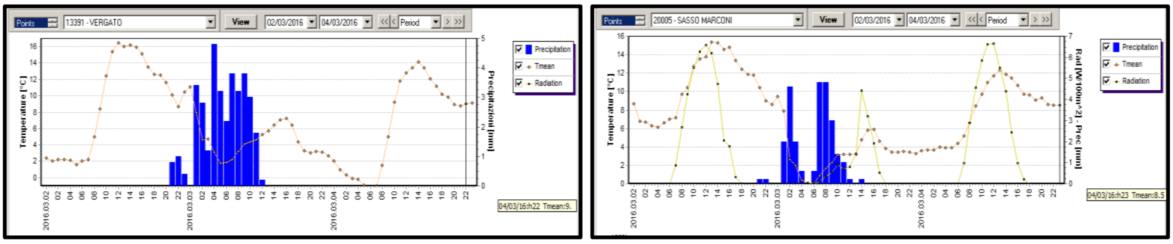

Figure 6: Time series of temperature and hourly precipitation: Sasso Marconi (Left) and Vergato (Right) from 02/03/2016 00:00 UTC to 05/03/2016 00 UTC

5.2 Undersampling of precipitation in convective season

Scattered precipitation is a challenge for every gauge network. We have pointed out in the quantitative analysis section that radar score drop down in the second period considered. On figure 7 the time series of hourly maximum has been plotted for some sub-periods of the convective season to highlight this problem.

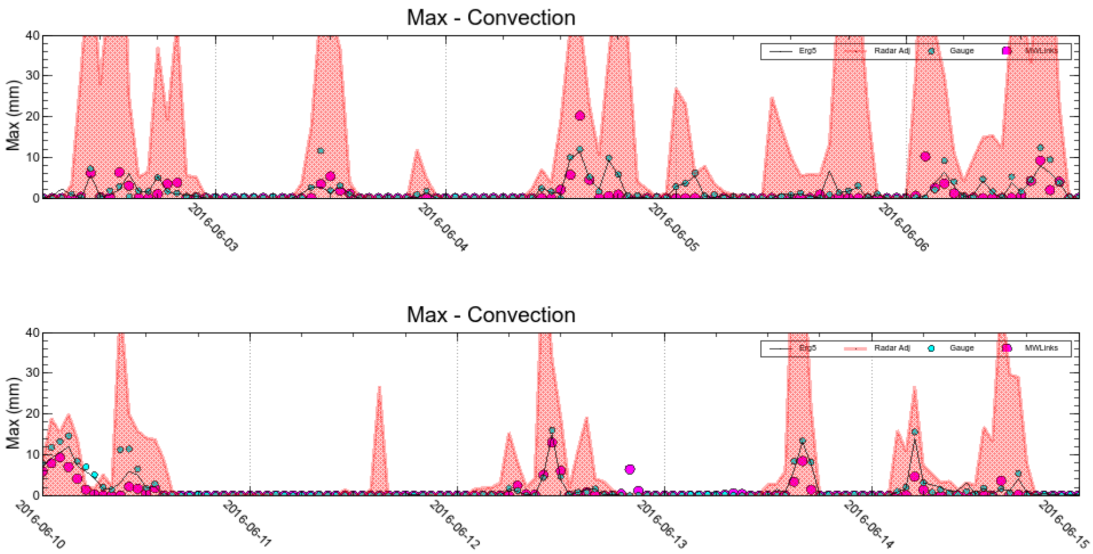

Figure 7: Time series of hourly maximum (Top panel) from 02/06/2016 00 UTC to 07/06/2016 00 UTC, (Bottom panel) from 10/06/2016 00 UTC to 15/06/2016 00 UTC.

In most of events plotted radar catches a signal stronger than in QPE datasets. Few examples of hourly patterns are shown in Fig. 8 and 9.

On May 14 at 20:00 UTC radar detects in the North of Bologna a convective storm, embedded in widespread precipitation pattern (Fig. 8). The storm marginally rides over some gauges and the analysed erg5 field does not detect such structure. MWL shows a local maximum in the same location where radar catch the storm. A closer look to link data shows similar problems in gauge and MWL data. Indeed no link cross the storm's core path. We point out that a merge of MWL data from different providers can increase the sampling capability and at the end precipitation retrieving.

On June 12t at 14:00 UTC a small storm rides over the Apennines slopes (Fig. 9). As in the previous case, gauge network recorded only the widespread precipitation in which storm is embedded and even MWL does not register the storm but only the surrounding precipitation.

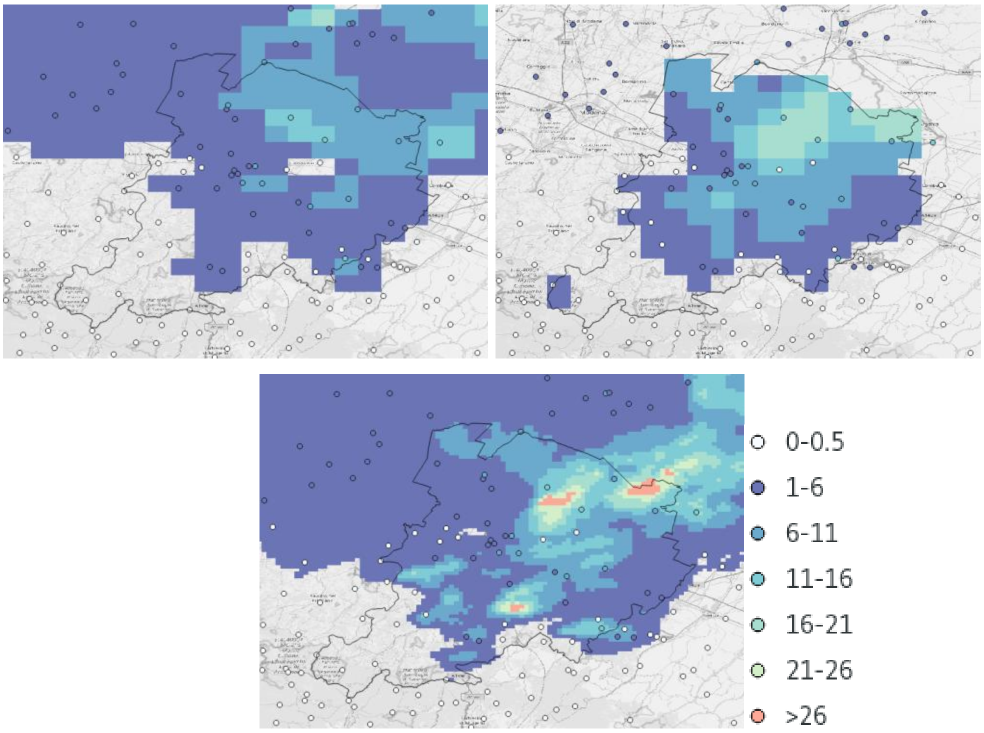

Figure 8: 14/05/2016 20:00 UTC hourly precipitation from different datasets, rain gauge data are overplotted in each panels: (Top Left) Erg5; (Top Right) MWL; Bottom Left) Radar Adj.

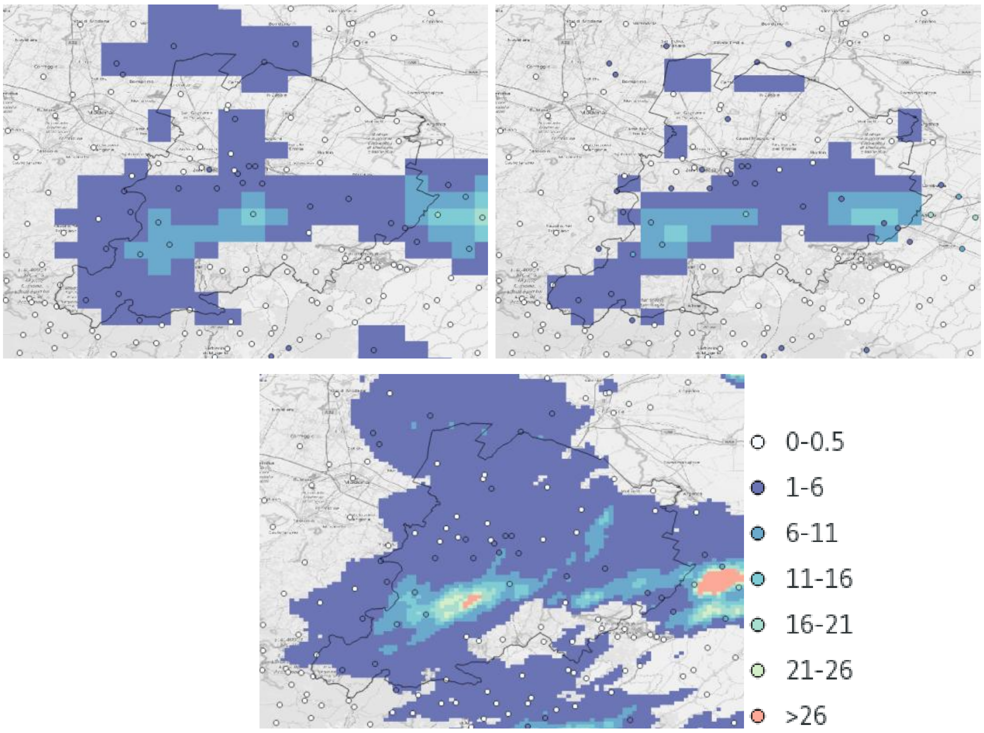

Figure 9: 12/06/2016 14:00 UTC hourly precipitation from different datasets, rain gauge data are overplotted in each panels: (Top Left) Erg5; (Top Right) MWL; Bottom Left) Radar Adj.

6 Conclusions

The agreement between MWL QPE and the other estimates is quite promising even if it requires a tuning activity in order to make the technology used in real-time or near real-time. The qualitative analysis shows that the performance of the MWL estimate seems to increase in the second half of the analyzed period, mainly characterized by convective storms, where showers and heavy rain play a predominant role even if embedded in frontal structures.

This behavior is confirmed by the quantitative analysis done by using statistical indicators: the scores of rainfall estimates by microwave links data get better during Summer and Spring seasons.

The validation has pointed out that a Microwave-based QPE method slightly underestimates precipitation occurrence both in spatial averages and punctual amounts, while Adjusted radar has complementary features.

The low spatial coverage of the microwave link network is an important issue to be addressed in the future, in collaboration with data providers, in order to increase the quality of the retrieved field.

Additional quality control procedures have to be developed during any validation follow-up activities in order to reduce the impact of false signals on the estimated rain fields. The automatic identification of false signals has to be tackled in particular in the following two scenarios: antennas structural deformations due to solar radiation or strong winds and bright-bands detection due to snow-melting layers crossing the microwave link path.

Acknowledgement

This work has been done within the framework of the RainBO project LIFE15/CCA/IT/000035