Summary

The article describes a number of performant quality control functions in order to assess the quality of precipitation measured by ground based stations. The overall system is detailed and provides a good understanding of different observation problems. The final outcome of the RainGaugeQC system is a quality index for each station.

General remarks

lines 67 and 73: There are numerous stations which are operated by former meteorologists or even active meteorologists on their private ground which provide data of a high quality. Often, operators of such stations are organised in amateur meteorological clubs. These clubs often publish data on the internet and togther with the conditions of the measurement locations, so that they are documented. Therefore, private stations do not guarantee a good data quality, but a considerable number of them is regularly monitored and their data of high quality. This needs to be verified for each private station, though.

> We are aware of this. However, in general, we cannot say anything with certainty about the fulfilment of WMO standards by private stations. Besides, the reliability of individual stations for very different reasons can change over time. Nevertheless, we have somewhat mitigated the meaning of this passage (line 75).

line 82: internationally, dual-sensor rain gauges are not the rule, even in national weather services. The WMO classifies station location (e.g. GIMO Guide of WMO (2021)) and has performed several gauge intercomparison exercises (Lanza / Vuerich, 2009).

> We are also aware that two-sensor stations are not standard. We have removed the sentence from lines 82-83, especially as it does not apply to non-professional stations.

Specific remarks

> Please note: all line numbers below refer to the original version of the article, not modified.

line 33: "... highly distorted" - depending on the operator, usually radar measurements are quantitatively less accurate, but not highly distorted (see e.g. WMO, 2024)

> This is a debatable issue. We have been working for many years on weather radar data quality control (RADVOL-QC system, references in the article) and in our opinion the radar data is "highly distorted". This is not visible on the data that comes to the users. However, we have removed the word "highly".

line 33: "Rain gauge measurements are still considered ..." - maybe you could be more specific, like: "In hydrology, rain gauge measurements are considered ..."

> Thanks - we have changed it.

line 42: it is correct that the authors have published relevant papers in this context. However, the reader would also appreciate more general publications, such as WMO BPG on radar data quality and Lanza / Vuerich (2009) on the WMO rain gauge intercomparison.

> Thanks for this reference - we have added it.

lines 110 and 115: radar data should only be incorporated after their quality control - else this is not state of the art (WMO, 2024).

Thanks also for this reference - we were not familiar with this document! We have added a relevant note to the text in line 110:

„but only after quality control (WMO-No. 1257, 2024)"

line 117: If you do not set a minimum threshold, your correlation risks to give you random results - this should be pointed out more clearly here.

We have added information on the threshold assumed:

"so a minimum precipitation threshold should be used to filter out data from such periods"

line 131: the adjustment to rain gauges modifies the radar time series temporally and thus bears the risk that a comparison to a rain gauge time series becomes more difficult. A time series with corrected radar data and without gauge adjustment might give better results, in particular if the gauge data to be analysed have been integrated into the adjustment.

We have modified the text:

"However, such measurements are not common, so remote sensing data such as radar observations, which are more widely available, can be used as a benchmark, but they require the QC to have been previously carried out."

lines 132 and 133: please give an indication of the length of the required time interval, e.g. one year.

de Voss (2019) used a 14-day interval (for 5-min data). We have added this information in line 133.

table 1: the column "type of rain gauges" would merit additional information: what is the minimum volume of the tipping bucket gauges? Are they unheated for the two first rows? Which measurement type is the third row? "heated" does not tell much ...

Minimum volume of the tipping bucket gauges applied at IMGW is 0.1 mm.

line 181: please comment on the accuracy of the daily gauges if there is rainfall at the time of the day change. How accurate can the daily precipitation amount be under such conditions? Manually operated gauges often have a time interval for the operator to read and check the rainfall amount which may be in the order of 15 minutes. Such information should be communicated, additionally to the formulation that their data "are believed to be more accurate".

This observation is confirmed by studies carried out on data from the IMGW rain gauge networks, e.g. Urban and Strug (2021). We have added this paper to the references and the text at line 182:

„…, which has been confirmed by extensive reliability analyses of different types of rain gauges at IMGW (Urban and Strug, 2021)."

line 189: Please comment on the range of 250 km. Please note that for hydrological quantitative use, distances beyond 120 km range are subject to higher uncertainties in the radar measurement due to the measurement height and measurement geometry. Which range is practically used for your applications?

Radars of POLRAD network measure reflectivity at a range of 250 km as standard, but these measurements are applied for different purposes. We are aware that for QPE these 250 km are far too far. On the other hand, all radar networks in Europe measure at ranges greater than 120 km (as far as we know, the minimum is 180 km).

In our case, we use data up to 215 km. This distance is the result of a compromise: (i) as short a range as possible, (ii) full coverage of the whole country with measurements. We didn't describe this in the article because radar measurements only play a supporting role in it. But since you find it useful, we have added the following in line 189:

“For the estimation of the precipitation field, data up to 215 km from the radar are used. This distance represents a balance between achieving the shortest possible range and ensuring complete coverage of the entire country with measurements.”

Figure 3: Why 215 km ranges? Can you please elaborate on this?

The explanation is as above.

line 210: please give more details on the satellite data used in the system! Which is the data source, how is it quantitatively transformed to precipitation? Which is the original resolution in time and space and how is it mapped to fit to the ground measured data?

We have not written more extensively about satellite precipitation because it is a side issue for this paper, although it is of course of interest. However, there is a reference in our paper to Jurczyk et al. (2020b), where this is described in detail.

Table 2: What do you understand with "gross errors"? Please explain more in detail or refer to the correspondent section in this paper

We have added sentences in the text in line 228:

“The GEC involves detecting when natural limits are exceeded, while the RC focuses on identifying when climate-based thresholds are surpassed at an individual gauge.”

Table 2: TCC - over which time interval does the comparison take place?

Details of the particular algorithms, including parameter values, are given in the relevant chapters. Time intervals in the TCC are given in line 276.

Table 2: SCC - please provide more details on the definition of outliers in this context!

We write about the SCC in chap. 3.5. However, there only the changes with respect to the original version (see Ośródka et al., 2022) are described. We have not repeated this basic information about the algorithms in the present work.

line 255: what would be a typical "specific time interval"? Please provide a range in minutes!

In Section 3.2 there is only a introduction to the directions in which the changes in RainGaugeQC have gone. All details are given in the relevant subsections of Section 3. The time interval in the TCC is 5 and 10 days - we write about this in line 276.

line 256: do you request a minimum amount of rainfall in radar and gauge?

Yes. We write about this in line 275 (0.025 mm).

line 256: When do you consider a correlation coefficient to be "good"? Please give more details in the assessment of the correlation quality!

We write about the relationship between the correlation coefficient and the assigned QI value in chapter 3.3 (formulae 3-5).

line 263: how do you carry out the unbiasing procedure? Please provide more details!

The BC (unbiasing) procedure is described in detail in section 3.4.

line 272: when do you consider a time series to be long? What is the minimum duration for this?

This is described in detail in section 3.3.

line 275: the amount of 0.025 mm is per which time period?

This 0.025-mm threshold is applied to 10-min totals. We write about this in line 276.

Regarding lines 255-275: After reading some of the Reviewer's previous comments on lines 255-275, we concluded that at the beginning of Section 3.2 of the article we did not clearly emphasise that here is only a preview of what has been changed, while the details will be developed in Sections 3.3 - 3.5. We have added the relevant information to the text in line 253:

„Here is a brief overview of the changes made to the RainGaugeQC algorithms, whereas detailed information can be found in Sections 3.3 to 3.5."

line 336: the formula implies that the QI value is reduced even for perfect data. Is this intended?

We agree: only for private stations we always reduce the quality by at least 0.1, which is due to their generally lower reliability. We have reason to assume that while professional stations meet WMO standards and are properly maintained and supervised, we are not so sure about private stations.

Formulas (9) and (11): it is unusual to work with a bias in this way - more often, a multiplicative approach is used (see chapter 3.3.5 of WMO, 2024). Your approach penalises a deviation of 5 mm equally for a rain gauge sum of 10 mm and of 50 mm where in the first case, this represents 50%, in the second one 10%.

Bias is applied multiplicatively (equation 14). Regarding the example of the 5 mm difference in rainfall of 10 and 50 mm: according to the similarity function (equation 7), these cases are similar to each other. Thus, if this example concerns professional/unprofessional rain gauges, then the QI is not reduced in both cases. However, if it applies to private rain gauges, then the bias is 1.5 and 1.1 respectively, i.e. they were in the range [0.2, 5.0] (equation 10), so in both cases QI = QI - 0.1.

We use the reduction of QI due to Bias with caution due to some uncertainty also in the radar rainfall R (especially for the 10-min accumulation period), and even after adjustment it is not ground truth.

Figure 4: (a) please indicate the number of dry time spells for each of the months - for some statistics this plays an important role

To avoid the negative impact of dry time spells, we used a threshold of 0.5 mm for daily totals - days with lower rainfall were not included. The number of such days per month was as follows: Apr - 4, Jul - 7, Oct - 5, Jan - 5. We have added this information to line 415:

„Days with precipitation accumulation below 0.5 mm were not included in the calculations (in total 21 days in these four months)."

(b) how did you take into account the systematic bias of extreme values due to the interpolation of the gauges? Does the distance of each gauge play a role? How would the result have looked when comparing to radar data?

We are aware that interpolation of rain gauge data during convective precipitation can involve large errors depending on the distance from the nearest rain gauge. Therefore, professional rain gauges situated at manual gauge locations, which is a relatively common situation in the IMGW network, were not included in the statistics in order not to favour this category of data (lines 413-414). The average distance of professional and unprofessional rain gauges from manual ones was then similar.

line 422: how can you state that the reliability of both data sources is comparable, if your reference is biased? This is, also in the light of the important and illustrative discussion in this paragraph, a statement without foundation.

We have tried to minimise the bias associated with the interpolation of rain gauge data as we have written in response to the previous comment.

line 458: Your finding that gauge data in Junauary are the least reliable is at least surprising since point rainfall data in summer are less representative in space. It would therefore be beneficial to read a discussion on these findings, in particular considering the predominant rainfall types and their spatial variability. Is this influenced by low temperatures in winter?

We have added such an explanation to the text, line 460:

"The low percentage of QI = 1.0 in January for both data types is due to the methodology used to determine these values in the SCC algorithm (Section 3.5). It uses the spatial variability function (SVF), which quantifies the spatial variability of precipitation at each time step. The high variability of precipitation is associated with convective precipitation and the introduction of the SVF function is intended to prevent such precipitation from being treated too rigorously and decreasing QI of good measurements. However, convective precipitation is very rare in winter in Poland, hence the frequent reduction of QI for weak outliers."

line 521: is this a finding for this day only or are the non-professional gauges always biased towards higher values?

Table 3 shows that this is a general feature of non-professional gauges.

line 540: do you consider this high value to be an outlier or a true value? If a true value, please discuss the discrepancies that you can see in figure 7.

Fig. 7 shows a convective cell with a very small area and very high rainfall intensity. Precipitation from such cells is generally short-lived. The reasons for these discrepancies are probably: (1) the lack of professional rain gauges in the path of this cell, which resulted in a decrease of the multi-source estimate, as well as ineffective adjustment of radar data to the rain gauges, which also effected the multi-source estimate, (2) the radar measures instantaneous values of rainfall, so with a 5-min time step it may have missed the maximum rainfall.

We write about this in lines 524-528.

line 556: do you have an explanation for the underestimation? Was there snowfall?

This is explained in a modified sentence in lines 557-558:
"This is mainly due to the underestimation of weather radar, which has problems detecting precipitation from low clouds most commonly occurring in winter, usually as snow, at farther distances from the radar site."

Technical details

line 11: replace "10-min" by "10 min".

line 49: delete "often"

lines 76, 133: replace "np" by "e.g."

line 79: delete "very"

line 122: I suggest to replace "underestimated" by "underestimating the true rainfall"

line 141: replace "were" by "are"

lines 147 - 148: please provide the number of stations here or omit them from the two following bullet points in order to give a uniform information

line 292: please add "r" after "correlation coefficient" for clarity purposes!

We have included all of these above technical remarks in the article. Thank you for reviewing it so carefully!

References:

WMO (2024) Guide to Operational Weather Radar Best Practices (WMO-No. 1257). Volume VI: Weather Radar Data Processing. Provisional version at https://community.wmo.int/en/activity-areas/weather-radar-observations/best-practices-guidance

Lanza, L., Vuerich, E. (2009) The WMO Field Intercomparison of Rain Intensity Gauges. Atmospheric Research, Volume 94, Issue 4, December 2009, Pages 534-543.

WMO (2021) Guide to Instruments and Methods of Observation (WMO No. 8). https://community.wmo.int/en/activity-areas/imop/wmo-no_8.

Thank you for these references!