Raingauge quality-control algorithms and the potential benefits for radar-based hydrological modelling

PHIL J. HOWARD, STEVEN J. COLE, ALICE J. ROBSON & ROBERT J. MOORE
Centre for Ecology & Hydrology, Wallingford, UK
philhw@ceh.ac.uk

Abstract Raingauges and weather radar are essential sources of rainfall information for hydrological modelling and forecasting. However, significant errors in raingauge time-series can drastically affect raingauge-only and combined radar-raingauge rainfall estimates. In turn, these errors can have a negative impact on hydrological model calibration, performance and failure diagnosis. This study considers the automated quality-control of 15-min rainfall totals obtained from 981 tipping-bucket raingauges across England and Wales. The Grid-to-Grid distributed hydrological model, now operated by the Flood Forecasting Centre in support of national flood warning, is used with gridded rainfall estimates to assess the utility of the raingauge quality-control procedures. Although a historical dataset is used here for demonstration and assessment purposes, the automated algorithms have been designed for implementation in real-time.

Key words quality control; raingauge errors; hydrological modelling; automated; flood forecasting; radar

INTRODUCTION

Good quality rainfall data are critical for accurate hydrological modelling. For operational application in real-time, for example in support of flood forecasting, there are competing demands on the timely availability of data and their quality. For raingauge data, operating agencies typically apply some form of manual quality-control, but this may be some time after data capture. To obtain improved real-time flood forecasts, much may be gained by applying automated quality-control algorithms to telemetry raingauge data.

Here we describe automated methods to improve the quality of raingauge data, such that a large number (981) of telemetry gauges may be included as live inputs to the Grid-to-Grid distributed hydrological model configured across England and Wales. Further to the requirement for real-time application, the size of the dataset alone makes the prospect of manual data quality-control daunting and impractically time consuming. It is not our intention to attempt to produce "perfect" rainfall time-series, but rather to robustly detect and effectively remove the most egregious faults in the raingauge network dataset.

Although time-of-tip records are available for many of the gauges in this study, they are not for all; hence the requirement for quality-control of a 15-minute interval time-series. This factor, combined with the spatial extent of the dataset and the long timescales considered here, leads to differences in our approach from those reported elsewhere (Jorgensen *et al.*, 1998; Steiner *et al.*, 1999; Stanzani *et al.*, 2000; González-Rouco *et al.*, 2001; Upton & Rahimi, 2003; Kondragunta & Shrestha, 2006).

The second part of this paper demonstrates the impact of the developed procedures on hydrological modelling, with examples from a historical study over the three years 2007–2009. Particular attention is given to the impact of the quality-control procedures when raingauge data are used in conjunction with radar rainfall data to obtain "best-estimates" of rainfall as input to the model.

QUALITY-CONTROL PROCEDURE

To enable real-time application and computational efficiency, all gauges are processed in parallel for each 15-min interval. A set of simple tests are performed on each gauge separately, before more involved comparisons to neighbours are made. Thus problems identifiable at the single gauge level should not adversely influence the quality-control of neighbours.

When a suspicious gauge value is detected, the value is flagged as missing and the details are logged. In the case of a historical study, past records are set to missing for the duration of the period over which the flag was set. For real-time application, it is thus necessary to run all checks at every time point, whereas for offline use, that requirement can be relaxed and long-term checks can be processed on a correspondingly less frequent basis (daily here).

In some operational situations, it may be desirable to substitute the suspicious records, for example by a Standard Average Annual Rainfall (SAAR)-weighted average of neighbouring values. However, this is neither necessary nor suitable for the present application where a gridded rainfall estimate calculated from functioning gauges already serves this purpose.

Single raingauge tests

For each raingauge, checks are made for exceedence of a threshold comfortably above historical maxima. First, each 15-minute value is checked against a 70 mm threshold, and then the accumulation over the last 24 h against a 350 mm threshold. Next, a set of checks which assess the likelihood of a partial blockage are performed. Complete blockages, where no tips at all are registered, can only be detected by comparison with neighbours.

A cumulative hyetograph for a blocked gauge is shown by the black line in Fig. 1, characterised by a gently curved profile when compared with neighbouring gauges. To distinguish such patterns without reference to neighbours, we identify long decreasing sequences of recorded rain (>2.5 h duration) noting that shorter sequences may well be the tail-end of a rainfall event, on the basis that a blocked gauge tends to drain more quickly the fuller it is (Upton & Rahimi, 2003). Equality is counted as a decrease within otherwise-decreasing runs, to account for lack of knowledge about individual tip-times within an interval. Because more rain may fall while a gauge drains, we also allow sequences to count as decreasing, so long as the number of decreases less the number of increases is greater than half the total sequence length.

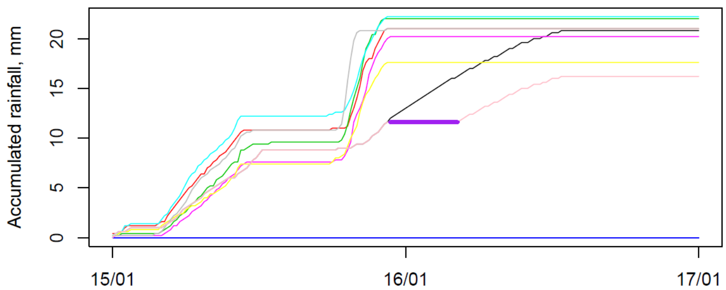

Fig. 1 Rainfall accumulations showing the effect of a partial blockage on the raingauge at Duston Mill (black line), from 15 to 16 January 2008. Also shown are its nearest seven neighbours (coloured lines), the quality-controlled accumulation (pink line) and the interval identified as suspicious by the tests (purple line). Note that records for the gauge at Litchborough (blue line) were missing for this period.

In the tail of such a blockage, there may be a sequence of 15-min time-steps with no recorded rain between two time-steps where single tips are recorded; knowing the tipping-bucket volume is important to identify a single tip. This is somewhat accounted for here by tracking the ongoing rate of tipping, with gaps between tips of up to 3 hours allowed. Further isolated tips are counted as independent events and may be detected later by multiple-gauge tests. Note that such small, sparse tips will have minimal impact on short-term hydrological modelling, though long-term water balances may be affected. Moreover, experience suggests that blockages can last for several weeks before being cleared: here, no attempt is made to maintain a gauge's blocked status and its responses to separate rainfall events are treated as independent.

We cannot expect to properly detect all problems. Figure 1 indicates the typical performance of the quality-control procedure: the blockage at Duston Mill is detected after the initial rainfall event has ended, but the blockage continues well beyond the detected range. Potential improvements to the current operation would be treating the gauge as missing over the entire event (or until what looks like a clearance occurs) and using a better classification of decreasing sequences.

Multiple-raingauge tests

Comparisons of each gauge's records against its neighbours' are performed once all single gauge tests have been completed for the current interval. Here the seven nearest gauges were used, seven being a compromise of proximity against inclusiveness suggested by Upton & Rahimi (2003). To avoid difficulties with particularly isolated gauges, a 50 km proximity threshold was applied. Also, a gauge is only compared against if it is deemed to have been functioning properly for at least two-thirds of the period under consideration.

Robust statistics (median and median absolute deviation) are used as the basis for detection of outliers, though other schemes could be considered (González-Rouco *et al.*, 2001). Use of these statistics somewhat mitigates the effect of multiple malfunctioning gauges in each other's proximity (Upton & Rahimi, 2003). Particular attention is given to detection of anomalously low values, since there is no simple and reliable method to distinguish intense localised rainfall from an error, other than the single-gauge tests already applied. This restriction could potentially be relaxed seasonally, as the likelihood of such events occurring reduces in winter.

Table 1 lists the quantities calculated for each gauge and the tests applied to distinguish outliers. For any quantity x, $[x]$ indicates the median of neighbours' values, x^+ the maximum, x^- the minimum and $\Delta(x)$ the median absolute deviation (MAD), that is $\left[\left|x-[x]\right|\right]$. τ_i is the tipping-bucket volume of gauge i. The thresholds are based on those found in (Upton & Rahimi, 2003).

Table 1 Multiple-raingauge quality-control tests.

Quantity calculated for gauge i	Criteria identifying fault
ν, the last day's median value	$\nu < \min(\nu^- - \tau_i, \frac{1}{2}\nu^-)$ and $[\nu] - \nu > 10\Delta(\nu)$.
$\bar{\nu}$, the last day's maximum value	$\bar{\nu} < \min(\bar{\nu}^- - \tau_i, \frac{1}{2}\bar{\nu}^-)$ and $[\bar{\nu}] - \bar{\nu} > 10\Delta(\bar{\nu})$ or
	$\bar{\nu} > \max(\bar{\nu}^+ + 96\tau_i, 2\bar{\nu}^+)$ and $\bar{\nu} - [\bar{\nu}] > 10\Delta(\bar{\nu})$.
D, the total over the last 96 days	$D < \frac{1}{2}D^-$ and $[D] - D > 5\Delta(D)$, or
	$D > 2D^+$ and $D - [D] > 5\Delta(D)$.
χ, the mean of the cross-correlation of gauge *i* with its neighbours, for the last 96 days' totals	$\chi < \frac{1}{2}\chi^-$, $[\chi] > \frac{1}{2}$ and $[\chi] - \chi > 10\Delta(\chi)$.

A period of around 3 months was found to be suitable for studying long-term behaviour. Many detected faults persisted for durations significant on such a timescale (and longer in some cases). In practice, a lag of up to one day was allowed when determining the correlation between each pair of gauges and the largest value taken, to avoid problems with events/faults at the end of a day.

APPLICATION TO HYDROLOGICAL MODEL

The distributed hydrological model used to assess the impact of the quality-control procedure is the Grid-to-Grid (G2G) model (Bell *et al.*, 2007), which is in operational use across England and Wales by the Flood Forecasting Centre. This employs gridded rainfall, calculated from networks of raingauges and weather radars, to provide estimates of river flow on a 1 km scale across the model domain. Raingauge data, mostly in time-of-tip form, were pre-processed into consistent series of 15-min totals across the 3-year study period. Radar data were in the form of 5-min quality-controlled rain-rates.

Configuration and estimation of gridded rainfall inputs to the G2G model, using a multiquadric surface-fitting approach, is discussed in Cole & Moore (2008). That approach is applied here to the rainfall values from the network of 981 raingauges at each 15-minute time-step to obtain raingauge-only gridded rainfalls. Similar surface-fitting applied to a ratio of the gauge to the coincident radar value at each raingauge location is used to obtain raingauge-adjusted radar gridded rainfall.

Results of an offline case study

The spatial distribution of raingauge and radar locations is displayed in the left panel of Fig. 2. The centre and right panels give an overview of the long-term impact of the raingauge quality-control procedure across the whole of England and Wales. Several excessively high (white areas) and low (dark blue spots) rainfall accumulations were successfully removed during quality-control. Note the adverse effect that the sparseness of the raingauge network has over northwest Wales, and Anglesey in particular.

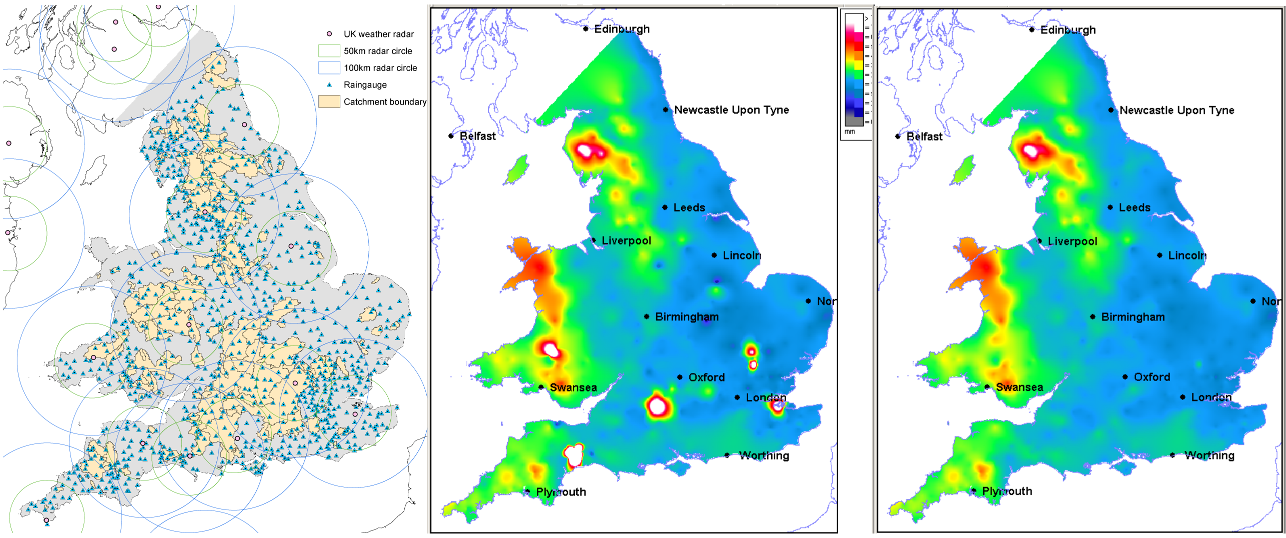

Fig. 2 Distribution across England and Wales of raingauges, weather radar and catchments used in case study (left). Rainfall accumulations for the 100 days up to 1 June 2009, for the multiquadric surface derived from 15 min totals of raw raingauge data (centre), and from quality-controlled data (right).

Summary statistics indicating the performance of G2G when using the different rainfalls as input are presented in Table 2. A three-month period from January to March 2008 was chosen to assess G2G performance over the set of 155 catchments whose boundaries are displayed in the left panel of Fig. 2. R^2 Efficiency and % bias (positive indicates overestimation) performance statistics, as a mean over all flow gauges, indicate the overall improvement obtained through raingauge quality-control. However, the signal is rather weakened by the averaging, since the majority of raingauges and time intervals experience no quality-control: 1.6% of values were identified as suspicious. For this period and set of catchments, the radar appears to significantly overestimate rainfall and use of either the raingauge-only or raingauge-adjusted radar rainfall estimates gives much improved G2G performance. Note that the G2G model has been calibrated using the quality-controlled raingauge-only rainfall estimates, as being the most consistent in time and space.

The benefit to hydrological modelling of raingauge quality-control is demonstrated much more clearly at the level of individual events. Figure 3 shows the accumulation over 15 days of rainfall in the vicinity of Blackburn, based on gauge-adjusted radar estimates with and without quality-control applied to the raingauges. There is clearly a severe fault with one nearby raingauge, which has recorded 30 times more rainfall than the average of its neighbours.

Table 2 Performance statistics for the Grid-to-Grid model over the three months from January to March 2008, using different rainfall estimates as input.

Gridded rainfall estimate	Mean R^2 Efficiency (Mean % Bias)	
	No raingauge quality-control	With raingauge quality-control
Raingauge-only	0.66 (–7.3)	0.68 (–5.9)
Raingauge-adjusted radar	0.63 (–3.2)	0.64 (–1.9)
Radar-only	0.47 (11)	N/A

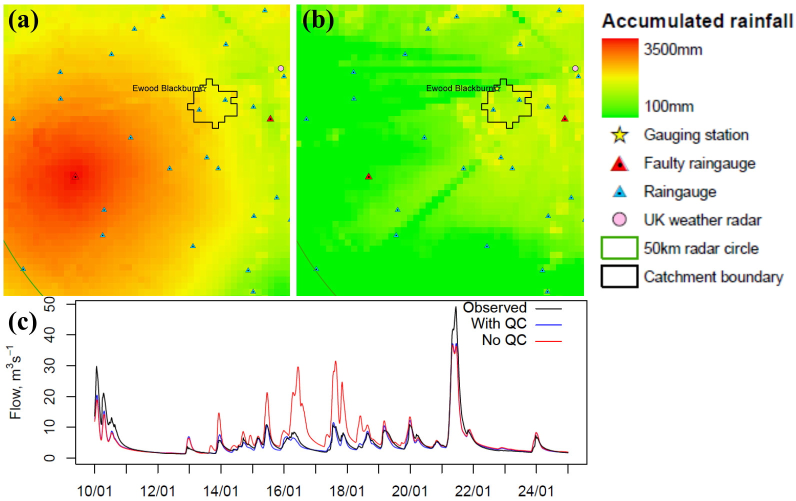

Fig. 3 Accumulated rainfall in the vicinity of Blackburn, over 15 days from 10 to 25 January 2008, using raingauge-adjusted radar: (a) without quality-control and (b) with quality-control of raingauge data. (c) Comparison of Grid-to-Grid modelled flow for the River Darwen at Ewood Blackburn using (a) as input (red line) and (b) as input (blue line).

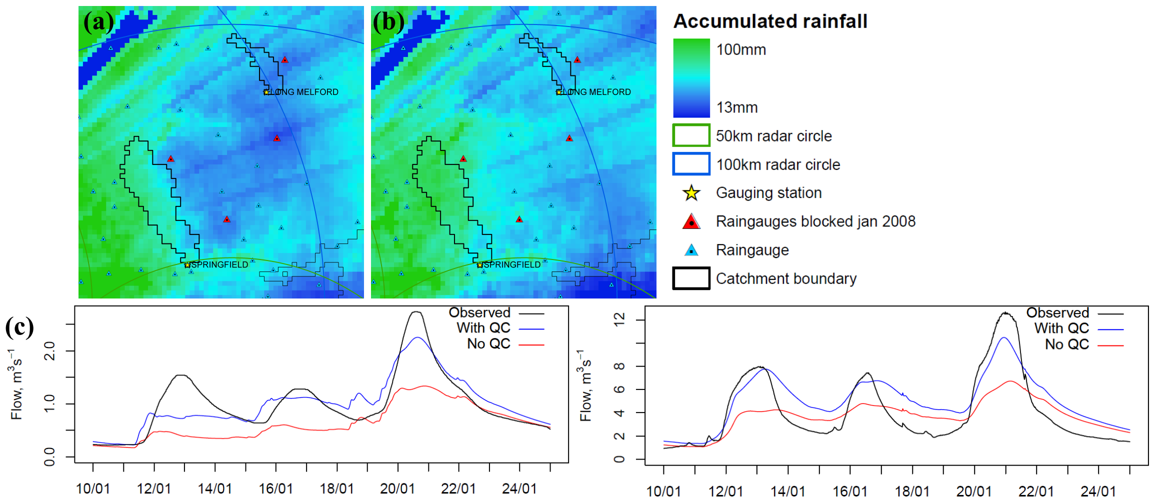

Fig. 4 Accumulated rainfall in the vicinity of Sudbury, over 15 days from 10 to 25 January 2008, using raingauge-adjusted radar: (a) without quality-control and (b) with quality-control of raingauge data. (c) Comparison of Grid-to-Grid modelled flow Chad Brook at Long Melford (left) and for the Chelmer at Springfield (right) using (a) as input (red line) and (b) as input (blue line).

Due to the large multiquadric offset parameter used in the gauge-adjustment process (Cole & Moore, 2008), this fault produces a large bias even over raingauges that are functioning well. Figure 3 shows observed and modelled flow for the River Darwen at Ewood Blackburn (drained area

39 km^2) for the same period. When gauge-adjusted radar rainfall is used as input to G2G, quality-control leads to an increase in the R^2 Efficiency of the resulting modelled flows, from 0.42 to 0.90.

A more subtle effect is produced by raingauges that are failing to record or underestimate the proper number of tips for some reason, such as full or partial blockage. Note that blockages may tend to cluster due to their environmental and seasonal nature. Figure 4 shows accumulations from gauge-adjusted radar rainfall, over the same period as Fig. 3, but for a much drier area near Sudbury. Several raingauges in the vicinity were detected as blocked and/or completely failing to register rainfall during this period. This leads to an overly low estimate of rainfall over the area when quality-control is not applied. The observed and modelled flow for two nearby river gauging stations is plotted in Fig. 4 and the improvement in modelling due to quality-control is again clear. For Chad Brook at Long Melford (drained area 47 km^2) the R^2 Efficiency climbs from 0.07 to 0.74 and for the River Chelmer at Springfield (drained area 190 km^2) it increases from 0.49 to 0.67.

CONCLUSIONS

It has been shown that raingauge-radar merging schemes can be vulnerable to raingauge errors. The automated quality-control routines presented here can successfully correct some of the worst of the faulty raingauge readings, leading to significantly improved performance for a distributed hydrological model that uses their derived rainfall estimates as input. There are clear benefits for historical archives of raingauge data and their use in forecast model calibration.

In real-time operation, the benefits of the quality-control algorithms presented here may be reduced, since it is no use identifying the previous month's readings as suspicious when only the most recent records may impact on flood forecasts. A refined procedure would have to focus more on detecting anomalous individual tips, perhaps via statistical comparison with historical records (González-Rouco *et al.*, 2001). There remains a problem with the spatial nature of the raingauge network since rainfall, or its absence, at one gauge may look suspicious, but 15 minutes later it becomes apparent that the storm responsible has moved to or from neighbouring gauge locations. Ideally, the algorithms would draw on radar data to obtain better criteria for detecting raingauge errors (Steiner *et al.*, 1999; Stanzani *et al.*, 2000).

Acknowledgements The Environment Agency and Met Office are thanked for making available the data used in this work. Science budget funding from the Centre for Ecology & Hydrology supported the research.

REFERENCES

Bell, V. A., Kay, A. L., Jones, R. G., Moore, R. J. & Reynard, N. S. (2009) Use of soil data in a grid-based hydrological model to estimate spatial variation in changing flood risk across the UK. *J. Hydrol.* 377, 335–350.

Cole, S. J. & Moore, R. J. (2008) Hydrological modelling using raingauge- and radar-based estimators of areal rainfall. *J. Hydrol.* 358(3–4), 159–181.

González-Rouco, J. F., Jiménez, J. L., Quesada, V. & Valero, F. (2001). Quality control and homogeneity of precipitation data in the southwest of Europe. *J. Climate* 14(5), 964–978.

Jorgensen, H., Rosenorn, S., Madsen, H. & Mikkelsen, P. (1998). Quality control of rain data used for urban runoff systems. *Water Sci. Tech.* 37(11), 113–120.

Kondragunta, C. R. & Shrestha, K. (2006). Automated real-time operational rain gauge quality-control tools in NWS hydrologic operations. In: *Proc. 20th AMS Conference on Hydrology*, P2.4, American Meteorological Society.

Moore, R. J., Cole, S. J., Bell, V. A. & Jones, D. A. (2006) Issues in flood forecasting: ungauged basins, extreme floods and uncertainty. In: *Frontiers in Flood Research* (ed. by I. Tchiguirinskaia, K. N. N. Thein & P. Hubert), 103–122. 8th Kovacs Colloquium, UNESCO, Paris, June/July 2006. IAHS Publ. 305. IAHS Press, Wallingford, UK.

Stanzani, R., Alberoni, P., Nanni, S., Mulazzani, C. & Pasquali, A. (2000). Raingauge and C-band radar monthly rainfall comparison in the Po plain area. *Phys. Chem. Earth (B)* 25(10–12), 981–984.

Steiner, M., Smith, J. A., Burges, S. J., Alonso, C. V. & Darden, R. W. (1999). Effect of bias adjustment and rain gauge data quality control on radar rainfall estimation. *Water Resour. Res.* 35(8), 2487.

Upton, G. J. G. & Rahimi, A. R. (2003). On-line detection of errors in tipping-bucket raingauges. *J. Hydrol.* 278(1–4), 197–212.