The development of a Kriging based Gauge and Radar merged product for real-time rainfall accumulation estimation

Sharon Jewell and Katie Norman

Met Office, FitzRoy Road, Exeter, UK

(Dated: 16th July 2014)

Sharon Jewell

1 Introduction

Weather Radar measurements allow high resolution, real-time estimates of precipitation at the surface to be made, which can be used by hydrological models for flood forecasting and warning applications. However, radar estimates are prone to errors due to the vertical profile of reflectivity, beam broadening, variations in the drop size distribution, attenuation and orographic enhancement. Conversely, rain-gauges provide an accurate but spatially sparse measurement of rainfall accumulations. Skilfully combining radar estimates with gauge measurements can produce a high resolution merged product to enable a better use of radar measurements in hydrological applications.

The success of a merging scheme depends on factors such as gauge network density, data quality, accumulation period and rainfall distribution. We present here our recent results produced during the development of a real-time gauge and radar merging scheme for England and Wales. Kriging has been used to combine 15 minute gauge and radar accumulations in near real-time and the effect of the choice of variogram and the gauge data quality on the merged product has been examined.

2 Kriging with External Drift (KED)

Kriging is a general term for a number of geostatistical techniques that can be used to interpolate the value of a random field at a location where the true value is unknown (Wackernagel (2010)). Kriging methods use a linear least-squares estimation algorithm to evaluate the field at a grid point by using observations at nearby locations. A key feature of the Kriging methods is that the positions of the observation points relative to each other are included through a variogram.

Recent work carried out at the Met Office has shown that Kriging with External Drift (KED) is the preferred method for generating a merged gauge-radar product in real time across England and Wales. This technique uses two independent but closely correlated datasets – in this case gauge measurements from a network of 1064 tipping bucket rain (TBR) gauges from Met Office and Environment Agency networks, and radar data on a 1 km grid. The output of the KED algorithm produces a map of rainfall accumulation across the entire domain with the resolution of the radar grid. For each pixel location the merged rainfall is calculated by adding up weighted contributions from the nearest neighbour gauges. The weighting factors for each gauge measurement depends on the distance between the gauge and the centre of the pixel and the covariance of the surrounding gauge measurements. In addition, three constraints are placed on the merged rainfall value as follows.

- The sum of the weighting factors must be equal to 1.
- The sum of the weighting factor multiplied by the radar estimate at each nearest-neighbour gauge location must equal the radar estimate at the pixel.
- The estimation covariance between the unknown true accumulation at a pixel location and the accumulation from nearest neighbours is minimized.

In this study two alternative variograms have been evaluated. The first is a parametric variogram generated from the gauge measurements for each 15 minute accumulation period by fitting a spherical form to the correlogram (Schiemann et. al 2011). The use of this variogram is denoted by KEDs and the advantage of using this variogram is that it is very quick to compute. Alternatively, the variogram can be generated non-parametrically from the 2D Fourier transform of the radar field in accordance with the Wiener-Khinchin theorem (Reif 1956, Velasco-Ferero et. al. 2009). This scheme is denoted by KEDn and has the advantage of being able to take in to account rainfall at all locations (rather than just gauge sites) but it takes a lot longer to compute.

3 Gauge Quality Control and Spatial Correlation

The rainfall data used for this study is from a dataset of measurements from 1064 TBR gauges sited in England and Wales between September 2010 and September 2011. The information available from each gauge is limited to the 15 minute rainfall accumulation and a “suspect” flag which is set to “1” if an error is suspected. However, no further diagnostics describing the nature of the error and its severity is available. To ensure that the number of gauges available for merging is maximised at the same time as minimising the propagation of gauge measurement errors into the merged product, a more refined gauge Quality Control (QC) process was developed. This process uses gauge time-of-tip information and rainfall

 sharon.jewell@metoffice.gov.uk

totals for the accumulation period to determine whether or not gauges flagged as "suspect" are providing erroneous measurements and therefore whether or not they can be retained in the dataset. Maintaining a high gauge density is particularly important for Kriging as the quality of merged accumulations decreases with decreasing gauge density (Jewell and Gaussiat (2014)).

A further gauge QC test was developed to quantify the "spatial correlation" of a gauge measurement (i.e. how well it relates to the surrounding rainfall). A gauge may legitimately appear spatially uncorrelated due to being located at the site of a highly localised convective event, in which case the measurement should be retained in the dataset; however, gauges erroneously giving a particularly high (or low) reading that has been missed by the raw data "suspect" flag can also be detected by the scheme and these are removed from the merging dataset. This process involves comparing the rainfall accumulation at the gauge location with the 10 nearest-neighbour gauges in order to identify any significant difference in rainfall within the region. In addition, the rainfall accumulation recorded by the radar at the gauge location is used to check that rainfall was detected in the local area during the accumulation period. This helps to determine whether a particularly high rainfall accumulation is a valid measurement, or if it is actually an error in the gauge measurement process (such as the TBR recording multiple tips).

4 Method

Two case study periods were used for evaluation: a 6-day period of predominantly stratiform events (1st – 6th October 2010), and another 6-day period of predominantly convective events (3rd – 8th August 2011). 15 minute gauge and radar accumulations were merged using the KEDn and KEDs merging schemes with different QC flag combinations as detailed below. In each case, one-third of the available gauges (354 in total) were removed from the merging dataset and set aside for use as cross-validation gauges to allow data quality and detection efficiency statistics to be calculated. In addition the merged rainfall accumulations were inspected visually to further assess the performance of the different algorithms.

4.1 Data quality assessment

The data-quality was assessed by using the Root Mean Square Error (RMSE) value for all cross-validation gauge readings above a rainfall accumulation threshold, *t.* The threshold ranged from 0.2 mm to 2.8 mm in steps of 0.2 mm (consistent with the gauge quantisation of the TBRs used).

4.2 Detection efficiency assessment

For the detection efficiency, binary statistics based on the number of hits (*gauge and merged values above threshold value*), misses (*gauge above threshold and merged product below*), false alarms (*gauge below and merged product above*) and correct rejections (*both below the threshold*) were calculated for the same thresholds as in 4.1. The skills used were:

(i) Critical Success Index (CSI):

This expression is an all-round assessment of skill with a maximum (and "perfect") score of 1. It is calculated for a number of different threshold values to allow the scheme performances at low, mid and high rainfall thresholds to be determined.

$$CSI = \frac{hits}{hits + misses + false\ alarms} \qquad (4.1)$$

(ii) Frequency Bias (fBIAS):

This expression determines whether a scheme has a tendency to overestimate the rainfall accumulation (fBIAS > 1) or underestimate (fBIAS < 1).

$$fBIAS = \frac{hits + false\ alarms}{hits + misses} \qquad (4.2)$$

5 Results

The impact of applying the refined gauge QC and spatial correlation tests to the gauge accumulations on the quality of the merged accumulation is shown below. The RMSE, Critical Success Index and Frequency Bias averaged over all cross-validation gauge accumulations during the two case study events are shown in Figure 1, though other events were tested during the project.

5.1 Influence of gauge QC on gauge availability

Of the non-zero gauge measurements recorded during each of the events, 15-20% had a "suspect" flag associated with them. If this number of gauge measurements were rejected without further investigation the average area per gauge would rise from 490 km^2 to 612 km^2 which would have a significant impact on the quality of the merged accumulation. However, the use of the refined QC and spatial check reduces the number of rejected gauges to around 2% producing an average area per gauge of 500 km^2.

5.2 Impact of the refined gauge QC process on the quality of KEDn merged 15 minute accumulations

The impact of selectively removing gauges that fail the refined QC and spatial correlation test is shown in figure 1 which illustrates the cross-validation statistics at all error-free cross-validation gauges. The results show that applying the refined QC improves the RMSE slightly in convective conditions and notably, when compared to the radar alone, has benefit where the raw QC has reduced the gauge density to a point where the RMSE is made worse. There is no overall impact on the CSI as a result of the QC but the merged product significantly improves the frequency bias, though the successive refinements to the gauge QC reduce the improvement to the frequency bias. This is because the refinement of the merging scheme is biased towards detecting suspiciously high rainfall amounts and removing these gauges from the merging scheme, hence the number of "false-alarms" reduces more significantly than the number of "misses".

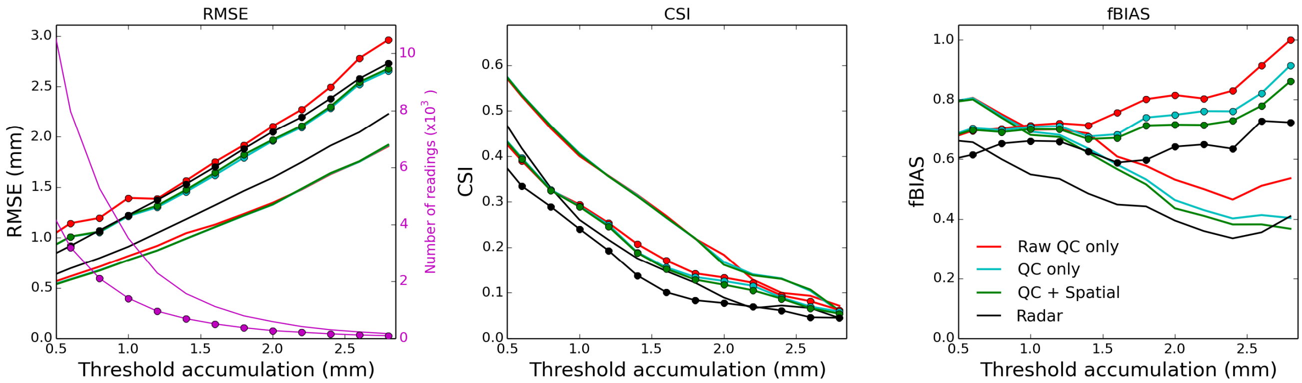

Figure1: Effect of gauge QC process on KEDn statistics at all valid cross-validation gauges as a function of rainfall threshold. The plain lines are for the stratiform case study and the circles are for the convective case study.

5.3 Effect of Variogram choice on the quality of the merged accumulation

The removal of gauges in the QC reduces the gauge density and therefore the number of measurements available for calculating the variogram as well as for the merging process itself. KEDs generally runs much faster than KEDn and has been identified as a possible alternative to KEDn. Figure 2 shows the results of comparing the two merged accumulations (from KEDs and KEDn) with the refined QC applied.

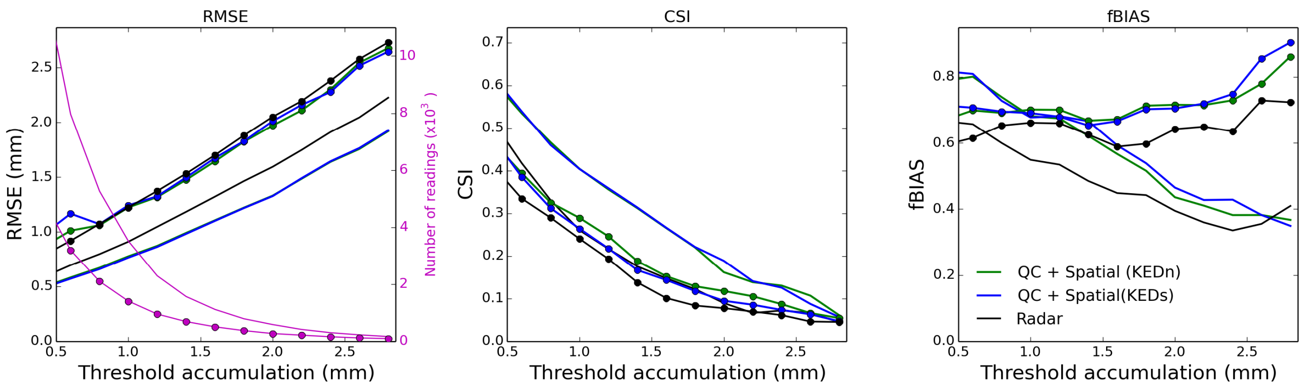

Figure 2: Effect of variogram choice on the quality of KEDn and KEDs at all valid cross-validation gauges, plotted as a function of rainfall threshold. The plain lines are for the stratiform case study and the circles are for the convective case study.

There initially appears to be little difference between the KEDn and KEDs results. However, figure 3 shows the results when only the rainfall at spatially uncorrelated cross-validation gauge locations are considered. This allows an assessment of the ability of the merging algorithm to retain localised high rainfall accumulations at locations where a rain-gauge is not present. The RMSE results in particular show that the greatest difference between the KEDn and KEDs schemes arises during convective rainfall at the mid-range rainfall thresholds with the KEDn scheme performing similarly to radar (a desirable feature in this particular case) and the KEDs accumulations performing worse

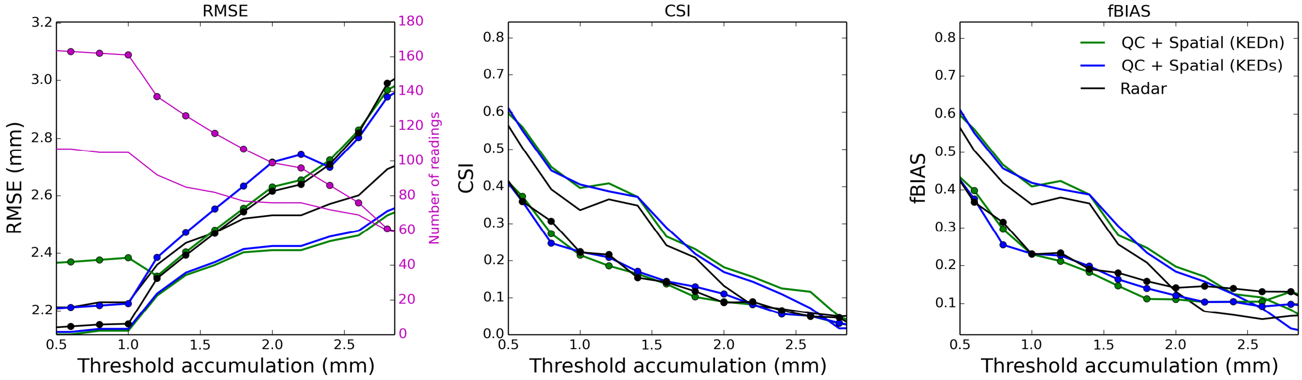

Figure 3: Effect of variogram choice on the quality of the merged accumulation, as validated at quality-controlled spatially uncorrelated cross-validation gauges, plotted as a function of accumulation threshold. The plain lines are for the stratiform case study and the circles are for the convective case study.

To realise the benefit of high-resolution QPEs from radar, the merging algorithms must preserve the spatial structure in this data. This can be evaluated by examining the rainfall accumulation images, examples of which can be seen in Figures 4 and 5. In the stratiform case (figure 5) the KEDs merged accumulation contains considerable artefacts due to gauge accumulations in the South-West, whereas the same area in the KEDn merged accumulation preserves more of the detail from the radar QPE. Therefore KEDn is likely to be more accurate in spatially inhomogeneous precipitation than the KEDs merged accumulation. This is also illustrated in Figure 5, particularly around the central region of the image where the fine detail of the radar image is lost in the KEDs merged product.

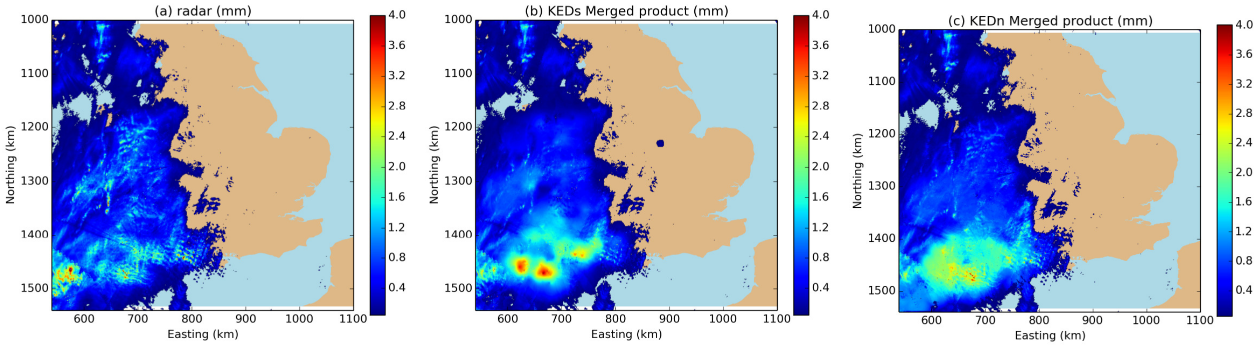

Figure 4: Examples of stratiform (a) radar QPE, and (b) KEDs merged and (c) KEDn merged 15 minute rainfall accumulations over England and Wales commencing at 04:15 on 3rd October 2010.

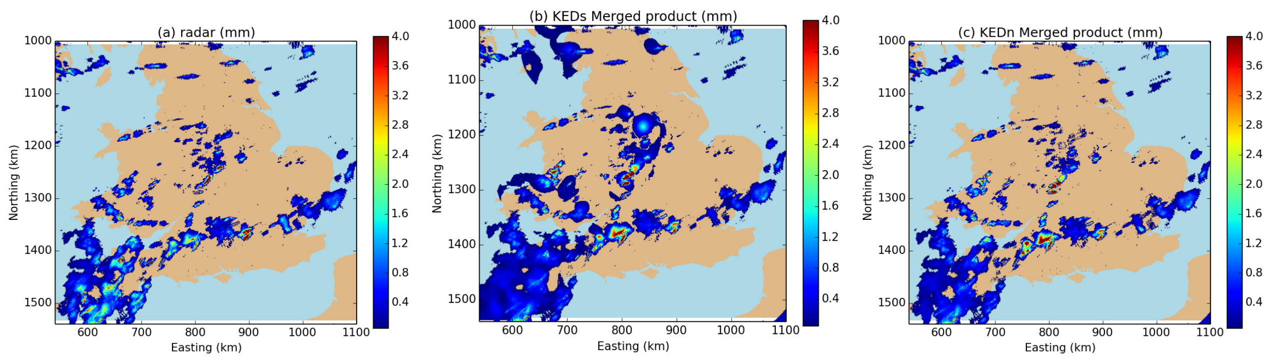

Figure 5: Examples of convective (a) radar QPE, and (b) KEDs merged and (c) KEDn merged 15 minute rainfall accumulations over England and Wales commencing at 15:30 on 7th August 2011.

6 Conclusions

This work has demonstrated that the use of a refined method for quality controlling gauge data, to improve the number of gauges available for merging, results in an improvement in the quality of the merged data. In addition, the use of a spatial correlation test can provide an additional gauge QC check for erroneous gauges, allowing gauges which are located at a site of a localised heavy rainfall event that is isolated from the neighbouring gauges to be distinguished from an incorrect gauge reading e.g. due to double-tips. The benefit of this process is two-fold; firstly, the number of gauges maintained in the merging scheme is not reduced unnecessarily and secondly, the fine-detail within the rainfall field is maintained. From examining the effect of the gauge QC process on the RMSE scores, at high rainfall thresholds the error drops by around 0.05 mm in stratiform conditions and 0.3 mm in convective conditions.

Comparing the cross validation statistics for the merged accumulations with parametric and non-parametric variograms shows that there is little difference in the quantitative accuracy of the merged accumulation when compared to gauge accumulations. However, closer examination shows that the non-parametric variogram retains a greater degree of spatial variability from the radar QPE in the merged rainfall accumulation. When the RMSE scores for the merged product are compared to the radar-only results, at the highest rainfall thresholds, the error drops by 0.3 mm in stratiform conditions and 0.1 mm during convective rainfall. These results again highlight the benefits of merging gauge and radar data to improve the quality of QPEs.

Acknowledgement

The authors would like to acknowledge joint funding for this project from the Met Office and the Environment Agency. The gauge QC and spatial correlation tests were developed by Shona Bryce and Selena Georgiou of the Met Office.

References

Jewell, S. and Gaussiat, N. (2014) A review of rain-gauge-radar merging schemes for flood forecasting applications, *Submitted to the Quarterly Journal of the Royal Meteorological Society*

Reif, F. (1956) Fundamentals of Statistical and Thermal Physics, McGraw-Hill

Schiemann, R., R. Erdin, M. Willi, C. Frei, M. Berenguer, and D. Sempere-Torres, 2011: Geostatistical radar-raingauge combination with nonparametric correlograms: methodological considerations and application in Switzerland. Hydrology and Earth System Sciences, 15, 22.

Velasco-Forero, C.A., Sempere-Torres, D., Cassiraga, E.F. and Gomez-Hernandez, J.J. (2009). A non-parametric automatic blending methodology to estimate rainfall fields from rain gauge and radar data. *Advances in Water Resources*, **32** pp986-1002

Wackernagel, H. (2010) Multivariate geostatistics: an introduction with applications, Springer, 3rd completely revised edition