Development of a Localized Radar-Rain Gauge Co-Kriging QPE Scheme for Potential Use in Quality Control of Real-time Rainfall Data

H.Y. Yeung, C. Man\textsuperscript{1}, A. Seed\textsuperscript{2} & S.T. Chan

The Third WMO International Conference on Quantitative Precipitation Estimation and Quantitative Precipitation Forecasting and Hydrology, 18-22 October 2010, Nanjing, China

\textsuperscript{1} The Chinese University of Hong Kong
\textsuperscript{2} The Centre for Australian Weather and Climate Research
Development of a Localized Radar-Rain Gauge Co-Kriging QPE Scheme for Potential Use in Quality Control of Real-time Rainfall Data

Hon-yin YEUNG¹, Chun MAN², Alan SEED³ and Sai-tick CHAN¹
Hong Kong Observatory¹, Chinese University of Hong Kong², Centre for Australian Weather and Climate Research³

Lead author’s email address: hyyeung@hko.gov.hk
Presenting author's email address: stchan@hko.gov.hk

1. Introduction

Quality control (QC) is required before automatic rain gauge (RG) data, e.g. typical tipping-bucket RG, can be used quantitatively due to various types of systematic and random errors, caused by various factors including wind, wetting, evaporation, splashing, calibration, finite sampling, mechanical failure, funnel blockage, signal transmission interference, power failure, etc (Habib, 2001). Among the random errors, those arising from contamination during data transmission through radio telemetry systems and blockage of the rain gauge itself by external obstructions, like insects or fouling by birds, are crucial in heavy rain monitoring and assessment. While telemetry errors usually contaminates a single observation with an extreme value, a blocked gauge will return zero or unreasonably small values that fall well within the climatological range and are therefore difficult to detect. A QC procedure based on spatial consistency checking is proposed in this paper. Major techniques employed include a localized quantitative precipitation estimation (QPE) scheme based on the co-Kriging of radar and rain gauge data, as well as statistical analysis of the RG-QPE residuals in search for suitable QC criteria.

2. Data Sets

The two major data sets used in this study are (a) radar reflectivity at constant altitude and (b) 6-minute rainfall accumulation data from the networks of Hong Kong and Guangdong province of China. A total of 14 convective rainfall events occurred during 7 February – 1 June 2010 were selected for algorithm development and one event for performance assessment. The reflectivity data came from a mosaic of two S-band Doppler weather radars in Hong Kong updated every 6 minutes. The reflectivity signals were subjected to spurious filter and clutter filter control. To minimize the effect of anomalous propagation, the reflectivity data used for co-Kriging were sampled at 3 km above mean sea level. They are gridded onto a 480x480 rectangular array (about 1-km grid spacing), covering an area of 512x512 km² centred at Hong Kong. To make the co-Kriging calculation tractable, a 1-out-of-5 data thinning strategy was applied to the gridded reflectivity data, reducing the effective resolution to about 5.3 km. The rain gauge data is provided by two independent networks from Hong Kong and Guangdong, which together contain more than 1,200 real-time gauges. In this study, only the subset within the radar reflectivity domain was considered. At each update time, about 50 rain gauges would be available from Hong Kong but the actual availability from Guangdong could vary from 500 to 900 with a nominal number around 750. Roughly speaking, the average gauge separation over the two networks is about 10 km but the distribution is uneven with much higher density over the Pearl River Delta region, reducing the separation to about 5 km. Both the update schedule and rainfall accumulation periods are different in the two networks. In this study, the 6-minute accumulations with a 6-minute update cycle were chosen for combined analysis. To match these requirements, only about one third of the Hong Kong rain gauges could be used. Besides, the sensitivities of the rain gauges are also different. While the minimum rainfall amount to tip is 0.1 mm in Guangdong, that required by Hong Kong gauges is 0.5 mm.

3. Radar-Rain Gauge Co-Kriging

Borrowing ideas from geostatistics, the ordinary co-Kriging technique was explored for optimal combined analysis of rain-gauge measurements (as primary data) and reflectivity deduced rainfall (as subsidiary data). For rainfall events, especially those associated with intense convection, short-duration accumulations could vary significantly across space and haphazardly over time. We might therefore treat such quantities as random variables and apply ordinary co-Kriging analysis to combine them. Here, we did not assume any second order stationarity for the local mean and variance of these random variables. Instead, we assumed intrinsic stationarity and work directly with variograms. For background information on intrinsic stationarity, variograms, ordinary co-Kriging and the general solutions to the system of equations, interested readers are referred to the reference books by Wackernagel (1998) or Webster & Oliver (2001), as well as journal papers by Goovaerts (1998) and Phillips et al. (1997). For previous applications of the geostatistical technique to radar and rain-gauge data, reference could be made to Creutin et al. (1988), Schuumans et al. (2007) and Velasco-Forero et al. (2009). In the next two paragraphs, we summarized
the main points in our estimation for the rainfall variograms, which are vital to the calculation of individual data weights, and configuration of the localized co-Kriging analysis.

The variograms and cross-variograms are assumed to be dependent only on the distance separation $h$ between a pair of rain gauges or radar grid points and not on the location of the pair of points or the orientation of the two points relative to each other, i.e. isotropy. The validity of isotropy will be discussed in Section 7. During each rain gauge update, the number of possible gauge pairs varies significantly making individual variogram estimation difficult and unreliable in real-time. If we further assume that the spatial structure of a particular type of rainfall system changes slowly, which is often the case for widespread and persistent rainstorm situations, we could pool data pairs at different update times together to make variograms more representative. Even with such considerations, the resulting experimental variograms would still be far from smooth and not suitable for solving the systems of co-Kriging equations. In real-time situations, dynamic modelling of the variograms will offer the most accurate spatial description of individual rainfall events but its automation may not be robust in practice. Instead, a more practical strategy of pooling all 14 rainfall cases together to prepare a single set of average variograms was proposed here. Fig. 1(a)-(c) show the resulting three average variograms with their corresponding theoretical variograms (exponential model with nugget effect) annotated.

It is known that in Kriging or co-Kriging, data closer to an estimation point will have screening effect on more distant observations, thereby reducing the latter’s weights (Wackernagel, 1998). Moreover, the actual set of variables/equations has to be kept to a manageable number if the computation is to be completed within a data update cycle, which is 6 minutes for the networks of Guangdong and Hong Kong. From Fig. 1, the ranges of both gauge-gauge and radar-radar variograms are within 40 km. Although the average gauge-radar cross-variogram hinted for an unbounded variance, individual cross-variograms (not shown) were bounded with a range generally less than 60 km. We therefore considered a radius of 50 km for setting the size of a circular neighbourhood for performing co-Kriging analysis. With such localization, we found that the average number of enclosed rain gauges is 10 or more. At most locations, the maximum number of enclosed rain gauges are well above 20. We further checked in two cases that the co-Kriging QPE will become stabilized as the neighbourhood size was increased to about 30 km and beyond. The number of radar grid points within the neighbourhood is fixed to be about 270.

4. Characteristics of RG-QPE Residuals

To quantify the amount of spatial consistency or inconsistency of a rain gauge data with respect to the co-Kriging estimate, we calculated RG-QPE residual defined as:

$$\xi = 10\log \left( \frac{G}{K} \right) = 10\log G - 10\log K$$

in units of decibels (dB) and studied their frequency distributions. In the above formula, $G$ denotes the short-duration accumulation reported by a rain gauge and $K$ the corresponding radar-rain gauge co-Kriging estimate. We also studied the RG-QPE residual in linear scale, i.e. $D = G - K$, and call $D$ as “RG-QPE departure” for the sake of easy reference. To make $\xi$ well defined for all possible $G$ and $K$ values, an offset (0.08 mm) slightly less than the sensitivity of the Guangdong rain gauges (0.1 mm) was added when either of them was zero. We calculated both the residuals and departures for all the 14 rainfall events and compiled the resulting distributions as shown in Fig. 2 (a) and (b). It is noted that peaks show up in the positive tail of the departure distribution at 25.5 mm and beyond, signifying unreasonably large rain gauge values as compared to their reference QPE. Such peaks constitute one QC criterion (see Section 5 below). For false zeroes or unreasonably small rain gauge data, such kinds of abnormalities are rather subtle to be detected as they are typically buried well within the broad central peak of the distributions. To unearth the hidden mines, we need to throw away the uninterested samples, i.e. those with $K$ equal or close to zero, as well as those bearing a relatively small estimation error. The absolute RG-QPE departure standardized by the QPE estimation error $\sigma$, i.e. $\delta = \left| G - K \right|/\sigma$, was used to select the data samples of interest. When $K$ deviates significantly from a vanishing $G$, $\delta$ tends to be large especially when $\sigma$ is small. We plotted in Fig. 2(c) and (d) the reduced distributions obtained by throwing away samples with $\delta \leq 2$ and $\delta \leq 3$ respectively. As shown in Fig. 2(c), a peak started to emerge in the negative tail of the reduced distribution around -12 dB, signifying some unreasonably large negative residuals attributable to false zeroes or unreasonably small gauge values. The peak remained visible as the distribution was further reduced by retaining only those samples with $\delta > 3$ (Fig. 2(d) refers). This discovery constitutes another QC criterion as summarized in Section 5 below.

5. Rain Gauge Data QC

Based on the previous discussions and results mentioned in Section 3 and 4, we proposed a simple QC procedure for short-duration rainfall accumulation reported by automatic rain gauges as follows:
(i) pre-QC screening to retain only those data lower than a preset threshold value \( G \); (ii) perform spatial consistency check on data retained in step (i) — 
- calculate a reference QPE using localized radar-rain gauge co-Kriging; 
- calculate the RG-QPE residual, departure and standardized departure; and 
- compare against preset threshold values (denoting \( \xi_1 \), \( D_1 \) and \( \delta_1 \) respectively); (iii) assign QC flags “P” (for pass) or “R” (for reject) as follows: 
- if \( D \geq D_1 \), flag “R”; 
- else, if \( \delta \geq \delta_1 \) and \( \xi < \xi_1 \), flag “R”; 
- otherwise, flag “P”.

For the 6-minute rain gauge data from Guangdong and Hong Kong, the following set of threshold values was selected for testing: \( D_1 = 25 \) mm, \( \delta_1 = 2 \) and \( \xi_1 = -12 \) dB. Note that the results were relatively insensitive to the choice of \( \delta_1 \) and higher values may also be used. The pre-QC screening is necessary, otherwise unphysically large rainfall data may contaminate the co-Kriging QPE and ruin the subsequent QC decision. From the maximum instantaneous Jardi rainfall rate of 513 mm/h registered in Hong Kong in 1971, \( G \) was set to be 51.3 mm in 6 minutes. We assumed the same threshold value is applicable to other rain gauge locations in Guangdong.

6. Performance Assessment

As our localized co-Kriging scheme excludes observation data at the estimation point, the RG-QPE departures mentioned in Section 4 could also be used for cross validation purpose. To avoid faulty rain gauge data, we only counted cases with \( |D| \leq 10 \) mm (implying \( \delta \) generally above 20). From these, the root-mean-square error of the 6-minute co-Kriging QPE was estimated to be about 0.36 mm with a vanishing mean error. We noted from the scatter plot of \( K \) versus \( G \) (not shown) that the co-Kriging QPE tended to underestimate the 6-minute accumulations, which is not surprising as the most important piece of observation is purposely omitted. Despite this sub-optimal aspect, the co-Kriging QPE is still considered useful because the QC criteria as described in Section 5 above are not very sensitive to the absolute values of the QPE itself. As illustrated in Fig. 3, the co-Kriging rainfall map (a) captures the spatial details as shown in the radar picture (b) and reflects the rainfall intensity indicated by the rain gauge readings. For comparison, the rainfall distributions analyzed from pure radar reflectivity (stratiform-rain Marshall-Palmer relation) and rain gauge data (Barnes method) alone were also shown in Fig. 3 (c) and 3(d) respectively.

The effectiveness of the QC procedure was assessed via an in-depth study with the case of 22 May 2010. The major rainfall episodes occurred from 8:00 a.m. to 12:24 p.m. and the QC procedure was applied to all the 27,635 6-minute rain gauge reports available during the period. Among these reports, 22,197 registered 0.0 mm and the remaining 5,438 at 0.1 mm or above. Out of the 22,197 reported zeroes, the QC algorithm rejected 183 reports (about 0.8%). Eyeball counter-checking revealed that 55 of which could be retained as true zeroes as they were recorded by rain gauges located on the edge or just ahead of some approaching rain bands. Out of the 5,438 non-zero reports, the QC algorithm rejected 30 (about 0.6%), which were mainly 0.1 mm. Eyeball counter-checking retained 6 of them. Overall, the rejection ratio is about 0.8% (totaling 213 out of the total 27,635 data processed). Among these rejected data, 29% (61/213) were suspected to be wrongly rejected according to eyeball counter-checking. With respect to the entire set of data samples, the portion of wrong QC as suspected by human eyes amounted to about 0.2% (61/27,635). To further assess the false-zero detection rate of the QC algorithm in heavy rain situations, we conducted an objective test targeting to simulate situations where gauges covered by an intense rain band were clogged. Firstly, a total of 562 reports with \( K \geq 1 \) mm and \( G \geq 2 \) mm were selected. Besides ensuring a significant rainfall rate, this selection can also minimize the number of ambiguous situations such as having a rain gauge located on the edge of a rain band. Secondly, each such rain gauge report was replaced by a bogus zero and then the QC procedure was re-applied. A total of 514 successful rejections were resulted, implying a false-zero detection rate of about 91%. Examination on the unsuccessful detections showed that they were all marginal cases with both \( \delta \) and \( \xi \) close to their respective thresholds (i.e. \( \xi = 2 \) and \( \xi = -12 \) dB). That means higher detection rate could have been achieved if the thresholds were lowered (at the expense of a higher false alarm rate, of course).

It was noted that there was no extremely large faulty readings in this particular case. As such, we cannot assess the automatic detection rate for this kind of rain gauge error directly. Instead, we conducted a second objective test similar to the one for false zeroes. All the 22,069 true zero reports were used this time. To make allowance for estimation error, the bogus rain gauge data was set to be 25.5 mm in each turn. A total of 21,690 bogus data were successfully rejected, resulting in a detection rate of 98.3%. In the remaining cases, the corresponding QPE estimates from the co-Kriging analysis were all greater than 0.5 mm, resulting in the failure of the
QC test to reject the bogus faulty readings. For rainy situations, an objective test to assess the corresponding QC performance for erroneously large readings has yet to be devised and conducted.

7. Discussions

The problems identified during QC performance assessment were mainly of two types, namely underestimated reference QPE by co-Kriging and the detailed design of the QC algorithm. To address the QPE issue, radar-deduced rainfall accumulation may be improved by using a reflectivity-to-rain rate conversion formula more representative of the convective climatology of the region. Secondly, the reflectivity thinning strategy could be adjusted to allow higher density of radar grid points in order to capture highly localized intense convective cells. To this end, major optimization of the co-Kriging solver will be required as the total number of variables/equations will be increased non-linearly. Apart from the underestimated QPE trend, ambiguous situations as discussed in Section 6 also existed where a rain gauge is located on the edge of a rainy area. At first sight, this might seem to prompt for consideration of anisotropy. To address this issue, we studied the gauge-gauge variogram using the data at 11:30 am on 22 April 2010. It was a classic case of SW-NE oriented squall line moving mainly from west to east with a small southward component. The blue and red variograms in Fig. 1(d) corresponded to gauge pairs oriented NE-SW and NW-SE respectively. It is evident that the two variograms differ appreciably only in the farther range, say beyond 40 km with essentially the same correlation range. This justified our earlier assumption of isotropy for the co-Kriging neighbourhood. And this short-range isotropy is expected to hold in general due to the use of 6-minute accumulation, which only reflects the rainfall distribution over a very short time scale (a typical rain band, including squall line, may travel only a few kilometers and any anisotropy caused by advection along the direction of motion will not be apparent). The fast motion speed (of order 40 km/h) of the squall line in this case could also help smear out any anisotropy. Yet the “edge” ambiguity is believed to be a genuine issue requiring further treatment in the future.

As for the design of the QC algorithm, one possible improvement may be achieved by considering also longer period accumulations, e.g. 30-minute or 60-minute, valid at the same time of the 6-minute rainfall. These information could be even more useful for identifying false-zero type of errors as they are likely to persist in time in general. Except at the 30- and 00-minute of the hour, currently there was no direct report of such longer-period accumulations from the Guangdong network during each update. If required, such accumulations have to be composed from the “raw” 6-minute data, a process in which data availability will introduce another QC problem. Another issue revealed in assessing the QC performance for large errors is that the simple check of \( D \geq D_c \) is insufficient for discriminating rain gauge reports that are not particularly large but are nevertheless inconsistent with the amount of rainfall seen from nearby gauges or radar reflectivity values. Skill similar to those explained for detecting false zeroes may be explored in the future.

8. Concluding Remarks

A localized version of co-Kriging analysis was developed specifically for QPE using rain gauge and radar reflectivity data taken in a subtropical environment of southern China, where highly uneven convective rainfall distributions are common in spring and summer. Such co-Kriging QPE provides a kind of “proxy” for measuring spatial consistency and a method for rain gauge data QC. Potential application of such technique in quality control over real-time rain gauge rainfall with short accumulation periods was explored. Preliminary validation findings based on selected cases in Guangdong and Hong Kong showed that the proposed QC scheme performed satisfactorily with high error detection rate. Under the existing radar data thinning strategy, the current co-Kriging solver requires a few minutes to complete a QC cycle for all gauges on a commodity Intel PC installed with Core i7 CPU at 2.67 GHz and Intel Math Kernel Library. But as discussed in Section 7, denser radar grid points is deemed necessary and a major optimization exercise will be required for enhanced operational performance. The limited coverage of Hong Kong radars confined the present study to a range of about 256 km, leaving the quality of about one quarter of the 1,200 gauges unchecked. Combined use with Guangdong radar data will be pursued to overcome the coverage limitation, as well as mitigating the range issue with reflectivity. In the future, the QC algorithm will be enhanced to cater also for the 5-minute accumulations provided by the other two thirds of rain gauges in Hong Kong. Moreover, 60-minute accumulations will also be explored for possible further enhancement to the current algorithm.

Acknowledgements

The authors wish to acknowledge and thank the Guangdong Meteorological Bureau for providing the rain gauge data in Guangdong used in this study. Thanks also go to Dr. Carlos A. Velasco-Forero for his critical review on the preliminary research results from this study.
References


---

Fig. 1 Average variograms and cross-variogram generated from 14 cases selected during 7 Feb – 1 Jun 2010. (a) gauge-gauge variogram; (b) radar-radar variogram; and (c) gauge-radar cross-variogram. The respective theoretical variograms are annotated and shown as a red curve in each plots. Plot (d) is the gauge-gauge variogram generated using the case data of 11:30 am on 22 April 2010 for studying anisotropy. The blue and red variograms in (d) corresponded to gauge pairs oriented NE-SW and NW-SE respectively.
Fig. 2 Frequency distributions of RG-QPE residuals: (a) all data samples; (b) all data samples but in linear scale (departures); (c) data samples with $\delta > 2$ only; (d) data samples with $\delta > 3$. 
Fig. 3 Comparing QPEs (60-minute accumulations) and radar reflectivity — example of 22 April 2010 ending at 11:00 am: (a) grid-based radar-rain gauge co-Kriging; (b) 3-km radar reflectivity at 11:00 am, 22 April 2010; (c) reflectivity-deduced rainfall using $Z = 200 \cdot R^{1.6}$; (d) Barnes analysis using rain gauge data only. Note that in the maximum rainfall category analyzed in (a) is 40-50 mm (near the left edge) whereas it is only 20-30 mm in the other two analyses in (c) and (d). Actual rain gauge data (numbers overlaid on the maps) matched closely with rainfall isohyets in (a).