Effect of density of gauges on accuracy of merged GSMAP: case study of typhoon Morakot

G. OZAWA[1], **H. INOMATA**[2] **& K. FUKAMI**[3]

1 CTI Engineering Co., Ltd. CTI Fukuoka building, 2-4-12 Daimyo, Chuo-ku, Fukuoka-city 814-0041, Japan
g-ozawa@ctie.co.jp

2 National Institute for Land and Infrastructure Management, Ministry of Land, Infrastructure, Transport and Tourism, 1 Asahi, Tsukuba city, Ibaraki prefecture 305-0804, Japan

3 International Centre for Water Hazard and Risk Management under the auspices of UNESCO (UNESCO-ICHARM), Public Works Research Institute (PWRI), 1–6 Minamihara, Tsukuba city, Ibaraki prefecture 305-8516, Japan

Abstract Previous research has found that satellite-based rainfall data such as GSMaP_NRT in general tend to underestimate raingauge data, especially for heavy rainfall, which poses a problem in their use for applications. Thus, a correction method for satellite-based rainfall data to improve the accuracy is necessary for applications. In some ungauged basins, a few real time raingauges have been recently installed. In these basins a merging method for satellite-based rainfall and raingauge data can be applied, e.g. Inverse Distance Weighted interpolation (IDW), co-kriging. These merging methods are expected to improve the accuracy of satellite-based rainfall with raingauges. This paper revealed the relation between the density of raingauges and the accuracy of merged GSMaP by IDW. The result implies that a certain level of the density (one station/5000 km^2) of raingauges is necessary for applications using merged GSMaP.

Key words GSMaP; merging, IDW; ungauged basin; Morakot

INTRODUCTION

Satellite-based rainfall data are widely available for public use, and research has been in progress to apply this type of data to a broad range of areas including flood management in ungauged basins. In particular, a near-real time rainfall product called GSMaP_NRT (provided by Japan Aerospace Exploration Agency (JAXA), GSMaP) has drawn much attention because of its high spatial and temporal resolutions and short data latency.

However, previous research (Seto *et al.*, 2008) have found that satellite-based rainfall products in general tend to underestimate observed ground rainfall, especially for heavy rainfall, which poses a problem in their use for flood forecasting and warning. Thus, a correction method for satellite data to improve the accuracy is necessary to use them for flood runoff analysis.

In some ungauged basins, a few real time raingauge stations have been installed recently. In these basins, merging methods for satellite-based rainfall and raingauge data can be applied such as Inverse Distance Weighted interpolation (IDW), co-kriging. Merging methods combine the advantages of point data (e.g. raingauge) in terms of accuracy and advantages of area data (e.g. satellite-based rainfall data) in terms of wide-range data. These merging methods are expected to improve the accuracy of satellite-based rainfall with raingauge data.

This paper will describe the relation between the density of raingauges and the accuracy of merged satellite-based rainfall data.

STUDY BASIN AND DATA

Study basin

It is necessary that the target basin has a high-density of raingauge stations to validate the relation between the density of raingauges and accuracy of merged satellite-based rainfall data. In addition, it is necessary that the target basin is large enough to validate the accuracy of merging with number of density raingauges in a very large area (e.g. one raingauge in 10 000 km^2). As a result, the entire island of Taiwan (area: 36 180 km^2), shown in Fig. 1, was selected as the study area to investigate the accuracy improvement of the merging of satellite-based rainfall data with ground raingauge data. Typhoon Morakot in 2009 caused serious hazards such as floods and landslides in Taiwan, and the number of missing and dead people reached 698 and 59, respectively (Water Resources Agency, 2010).

The Water Resources Agency of the Ministry of Economic Affairs in Taiwan has installed and managed raingauge stations. At present, there are 321 stations (one station in every 112 km^2) in Taiwan. All of the stations are capable of rainfall observation with 1 h resolution. This study used raingauge data collected at those 321 stations. According to these raingauge data, the maximum hourly rainfall reached 136 mm/h, and the total maximum rainfall recorded was 2848.5 mm at an observation point from 16:00 h 6 August to 9:00 h 10 August 2009 (UTC).

Data

In this study, GSMaP_NRT, a near-real time hourly product, is used as near-real time satellite-based rainfall data. GSMaP_NRT has been developed to deliver GSMaP_MVK, which is propagated both forward and backward in time between two successive images from satellite microwave radiometers, in real time. This product simulates rainfall-area movement based only on temporally forward-moving vectors provided by infrared radiometer observations; thus GSMaP NRT differs from GSMaP_MVK in terms of input datasets (e.g. in case of TRMM/TMI (sensor: TMI on satellite:TRMM), which is one of the passive microwave radiometers, the NASA/GSFC (NASA Goddard Space Flight Center) real time version is used in GSMaP_NRT, while the standard version is used in GSMaP_MVK, although the algorithms are the same for both). The product is available in approximately 4 hours after observations.

In addition, GSMaP_NRT has some error characteristics such as underestimation of orographic rainfall (Kubota *et al.*, 2009). In particular, the influence of the frequency of microwave radiometer observations on the GSMaP correction method due to low accuracy of extrapolation during no MWR observations has been indicated (Ozawa, 2010). To remove this uncertainty of GSMaP, only times of GSMaP data estimated by microwave radiometers are used in this paper.

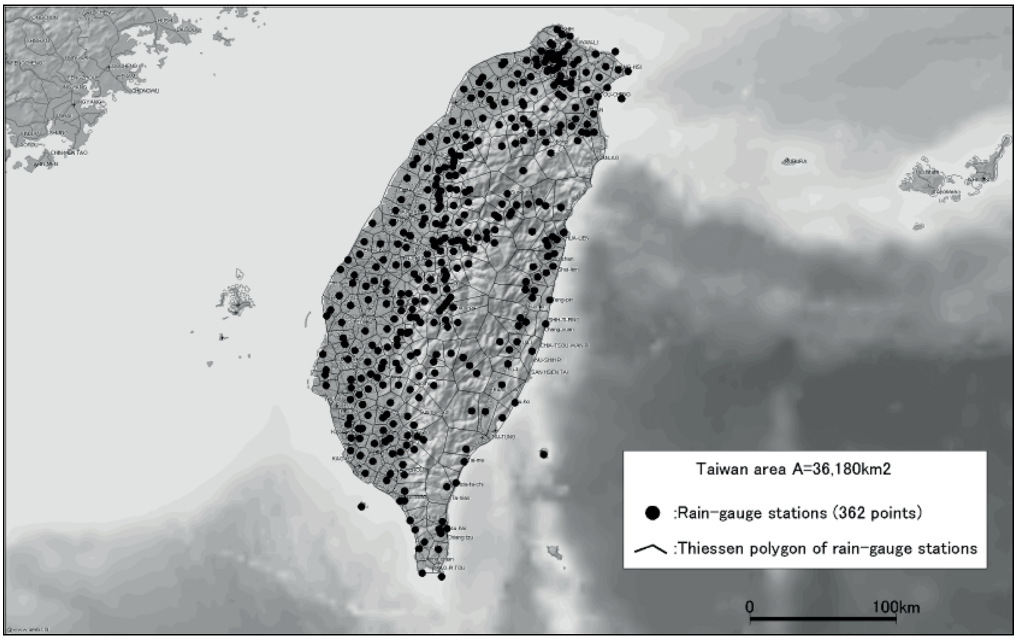

Fig. 1 Location of raingauge stations in Taiwan.

MERGING METHOD

Select of merging method

To combine raingauge and satellite-based rainfall data, it is necessary to select the merging method. In the past, several statistical merging schemes have been developed, such as conditional merging (Pegram, 2002) and Bayesian merging (Todini, 2001) and the Dynamic-window method

(Yamaguchi *et al.*, 1993). Co-kriging is one of the most popular multivariate spatial prediction methods, and it has been adapted as a merging method for satellite-based rainfall data with raingauges (e.g. Oke *et al.*, 2009). The Inverse Distance Weighted Method (IDW) is also adapted to merge a radar raingauge with a raingauge (Brandes, 1975).

The purpose of this paper is not to select the most accurate merging method, but to select one method which can be applied from ungauged basins to fully-gauged basins and conduct the sensitivity analysis of the relation between density of gauges and accuracy of merged satellite-based rainfall data. Thus, it does not matter what the merging method is if it can be applied from ungauged basins to fully-gauged basins.

In merging methods, the Dynamic-window method is based on using basins with a high-density of raingauge stations, such as in Japan. This means the method can not be adapted to ungauged basins. The co-kriging method estimates rainfall distribution based on the spatial relativity between raingauges and satellite-based rainfall, i.e. this method can not be applied in ungauged basins since there are few raingauges in the basin and it is difficult to find enough spatial relativity to merge. However, IDW does not have a limitation for merging since it does not depend on spatial relativity and the number of raingauge stations. Therefore IDW is selected as a merging method in this paper and validates the relation between accuracy of merged satellite-based rainfall data with raingauge data and density of raingauges.

In the IDW method, the weight at each grid is derived from equation (1) and then the correction coefficient is derived from equation (2) by using a correction coefficient at raingauge point (ratio of raingauge to GSMaP).

$$W(r_i) = \exp(-r_i^2 / A) \tag{1}$$

$$f = \sum_{i=1}^{N} W(r_i) \cdot f_{0i} / \sum_{i=1}^{N} W(r_i) \tag{2}$$

r_i is distance from grid to raingauge point; A is coefficient in terms of distance; f_{0i} is correction coefficient at raingauge point (R_{obs}/R_{sat}); N is the number of raingauges; f is the correction coefficient at each grid point calculated by IDW; R_{sat} is rainfall intensity of GSMaP mm/h; and R_{obs} is rainfall intensity of raingauge mm/h.

Methodology

Target times are set before, on and after the peak rainfall and zero rainfall (total: 4 h) in one heavy rainfall event, typhoon Morakot in Taiwan from (1) to (4), as shown in Fig. 2. In GSMaP, these 4 h estimated by using microwave radiometer observations are selected since the accuracy of GSMaP estimated by not using microwave radiometer observations is lower (Ozawa *et al.*, 2011). Since the merged rainfall intensity is influenced by the location of raingauges, drastically in the case of a low density of raingauges, three cases are prepared in the case of density lower than one station in 5143 km^2 and the number of total validation cases are 22 as follows: 1/167, 1/228, 1/400, 1/800, 1/1000, 1/1573, 1/2010, 1/3000, 1/5143, 1/12 060, 1/18 090 and 1/36 180 km^2.

The R statistical package gstat (Pebesma, 2004) available at http://cran.r-project.org/web/packages/gstat/ is used in this study to implement the IDW algorithm.

Criteria for accuracy analysis

In this paper, the following two comparisons were conducted to reveal the relation between density of raingauge stations and accuracy of merged GSMaP: (a) Comparison of basin average rainfall of original GSMaP and merged GSMaP; (b) Comparison of basin average rainfall of raingauge and merged GSMaP.

In comparison (a), the effect of accuracy improvement of merging method can be revealed. In comparison (b), the advantages and disadvantages of merging method can be revealed. These validations are evaluated by using the accuracy criteria below:

(1) Relative error rate of merged GSMaP [%]:

$$E_{merged} = (R_{merged} - R_{obs}) / R_{obs} \times 100$$

(2) Decrease ratio of relative error rate of merged GSMaP comparing with GSMaP or IDW_gauges:

$$DR_{GSMaP} = (E_{GSMaP} - E_{merged}) / E_{GSMaP} \times 100,$$

$$DR_{IDW\_gauges} = (E_{IDW\_gauges} - E_{merged}) / E_{IDW\_gauges} \times 100$$

(3) CE (Coefficient of efficiency) of merged GSMaP or IDW-gauges:

$$CE_{merged} = 1 - \sum_{i=1}^{N} (R_{merged} - R_{obs})^2 / \sum_{i=1}^{N} (\overline{R}_{obs} - R_{obs})^2$$

$$CE_{IDW\_gauges} = 1 - \sum_{i=1}^{N} (R_{IDW-gauges} - R_{obs})^2 / \sum_{i=1}^{N} (\overline{R}_{obs} - R_{obs})^2$$

Where, R_{merged}: rainfall intensity of merged GSMaP mm/h; R_{obs}: rainfall intensity of raingauge mm/h; R_{IDW_gauges}: rainfall intensity of IDW interpolated rain-gauge mm/h; $\overline{R}_{obs}$: basin-average rainfall intensity of raingauge mm/h; N: number of gauge stations

RESULTS

Comparison of basin average rainfall of original GSMaP and merged GSMaP

Figure 3 shows the relation between density of raingauge for merging and decrease ratio of relative error rate of merged GSMaP comparing with raw GSMaP by plotted all times (4 h). Accuracy of merged GSMaP has a higher accuracy in higher density of raingauge than one raingauge station in 5000 km^2 (1/5000 km^2). This result shows that raingauge data higher density than 1/5000 km^2 is necessary for accuracy improvement for merging method on GSMaP. Otherwise, the accuracy of merged GSMaP is inclined to be lower accuracy than the original GSMaP.

In Fig. 4, the accuracy of merged GSMaP in each density of raingauges are plotted as a different time (before the peak, on the peak, after the peak and almost no rainfall) and accuracy of original GSMaP are shown as horizontal lines. Comparing each accuracy for same time data, the merged GSMaP has a higher accuracy than original GSMaP in most cases of higher density than 1/5000 km^2, as mentioned before. However, in some cases, especially in Fig. 4(d) of almost no rainfall, merged GSMaP has a lower accuracy than original even in cases merged with higher-

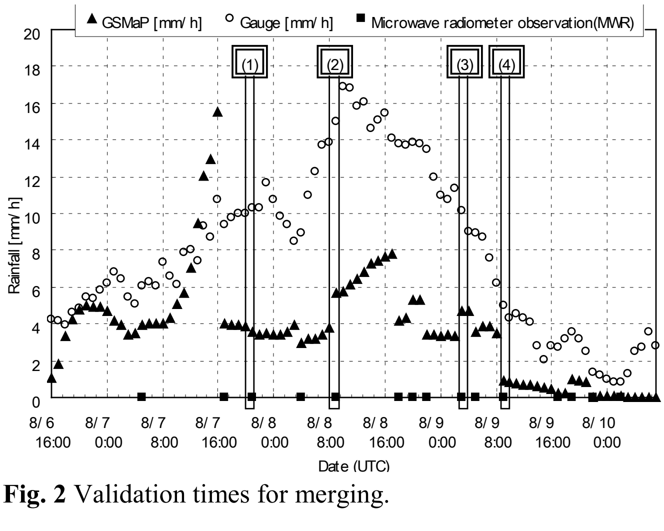

Fig. 2 Validation times for merging.

density gauges. This lower accuracy comes from overestimation of merged GSMaP. This reveals that the ratio of GSMaP to raingauge becomes several hundred when GSMaP is very small and as a result the overestimated ratio multiplied to surrounding GSMaP grids causes the overestimation in merged GSMaP. Therefore, there many overestimations can be found in cases of light rainfall like Fig.4(d), almost no rainfall. This shows the necessity of setting the upper limit (threshold) for correction coefficient.

Comparison of merged GSMaP and raingauge data

If the accuracy of merged GSMaP is lower than the accuracy of raingauge rainfall interpolated by using the same gauges used for merging, there is no advantage to using the merging method.

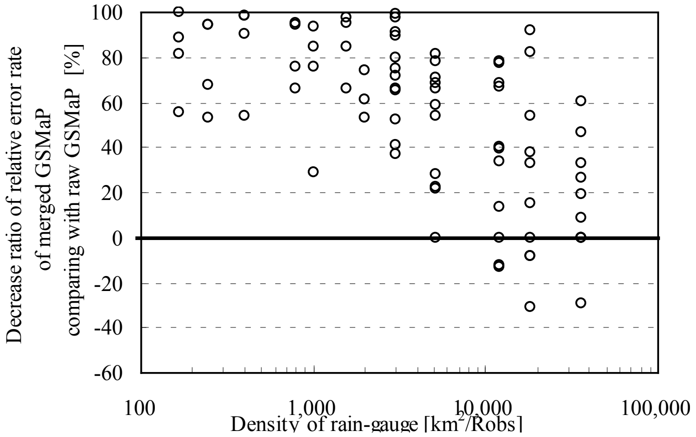

Fig. 3 Relation between density of raingauge and decrease ratio of relative error rate of merged GSMaP comparing with raw GSMaP.

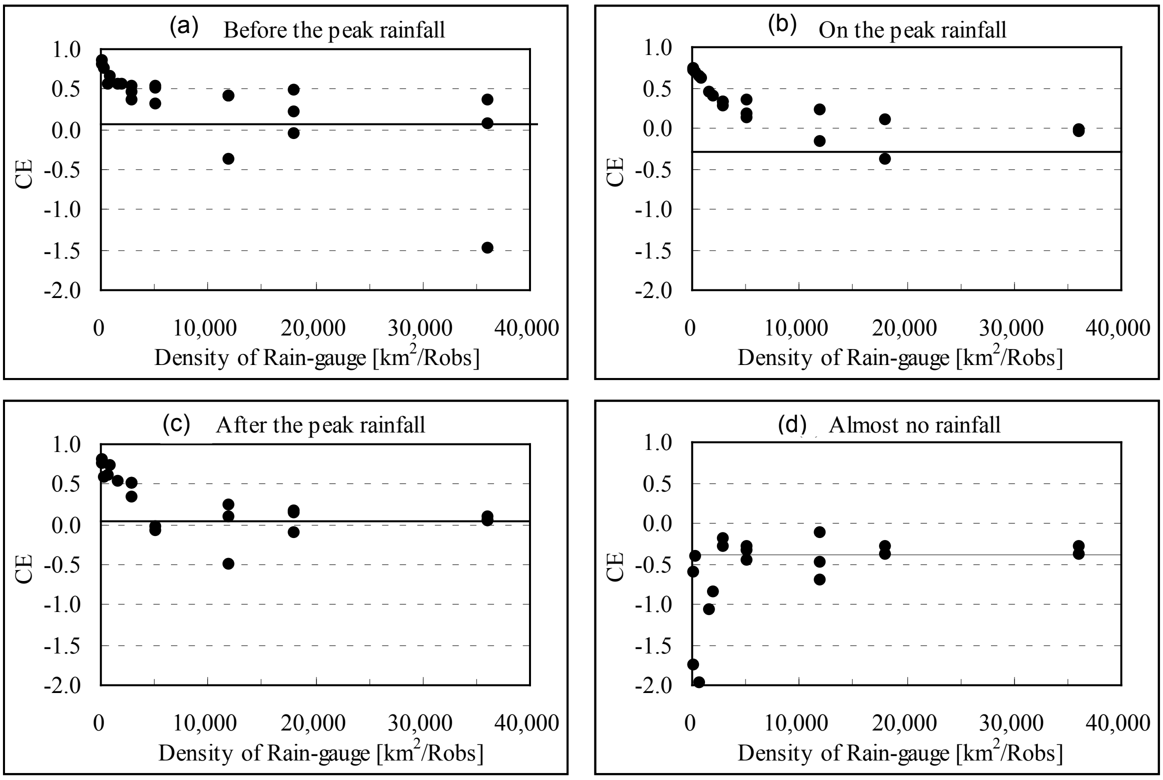

Fig. 4 Relation between density of raingauge and accuracy criteria (CE) at different times.

Therefore, the accuracy of merged GSMaP is compared with the accuracy of rainfall distribution interpolated (by IDW) by using the same gauges used for merging in this section. In addition, the following validation is conducted except for the rainfall time in Fig. 4(d) almost no rainfall, since the correction coefficient becomes extremely high in light rainfall, as mentioned previously.

Accuracy of merged GSMaP and accuracy of IDW interpolated raingauge data (IDW_gauge) are plotted based on the accuracy criterion CE in Fig. 5. The accuracy of merged GSMaP (average CE of merged GSMaP) is lower than the one of IDW_gauge (average CE of IDW_gauge) in higher density cases (up to 1/5000 km^2). However, merged GSMaP has a higher accuracy in all times in lower density than 1/10 000km^2. This is because the target rainfall event is typhoon, whose scale of rainfall distribution is tens of thousands km^2 and then the merging method can use GSMaP, which can represent macro-scale rainfall distribution. Thus, merged GSMaP utilizing lower density raingauge than 1/10 000 to 20 000km^2 has a higher accuracy than IDW interpolated raingauge rainfall distribution.

It is necessary to note that this study needs to increase the number of validations of other rainfall types and other raingauge locations since this paper is based on the case study on only one typhoon event and the number of lower density cases is limited (three cases).

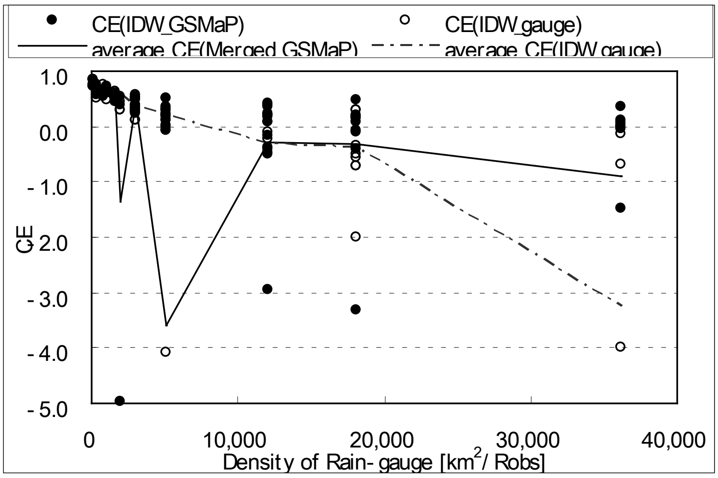

Fig. 5 Accuracy comparison of merged GSMaP and IDW interpolated raingauge.

SUMMARY

As a result of the experiment in the typhoon in Taiwan, the following conclusions are drawn in the comparison of raw GSMaP and merged GSMaP with raingauge by IDW:

(a) Merged GSMaP has a higher accuracy than raw GSMaP in higher density than one raingauge station in 5000 km^2. This result implies that a certain level of density (one station/5000 km^2) of raingauge stations is necessary in flood forecasting systems using merged GSMaP.

(b) In some cases of light rainfall period, merged GSMaP has a lower accuracy than raw GSMaP since the correction coefficient calculated as GSMaP/raingauge is overestimated in very light rainfall of GSMaP. This result shows the necessity of threshold for correction coefficient.

In the comparison of Merged GSMaP and raingauge data interpolated by IDW:

(c) Merged GSMaP does not have higher accuracy than raingauge data interpolated by IDW in higher density than 1/5000 km^2. However, in most cases of lower-density more than 1/10 000 km^2, merged GSMaP has a higher accuracy. This is because the target rainfall type in this paper is typhoon whose scale is several thousand km^2. While the rainfall distribution can be represented by high density raingauge data in higher density cases, the distribution can not be represented by low density raingauge data in lower density cases. Then the merged method can improve the accuracy in lower density cases since the rainfall distribution of GSMaP can represent the macro-scale rain distribution.

Since the layout of raingauge stations are limited (only a few cases) in this paper, it is necessary to note that all cases should be experimented in the future. In addition, sensitivity analysis should be conducted not only on a typhoon event but also on other rainfall types such as frontal rainfall and convective rainfall.

REFERENCES

Brandes, E. A. (1975) Optimizing rainfall estimates with the aid of radar. *J. Met.* 14, 1339–1345.

Kubota, T. *et al.* (2009) Verification of high-resolution satellite-based rainfall estimates around Japan using a gauge-calibrated ground–radar dataset. *J. Met. Soc. Japan* 87A, 203–222.

Ozawa, G. *et al.* (2011) Applicability of GSMAP correction method to typhoon "Morakot" in Taiwan. *Ann. J. Hydraul. Engng* 55, S445–450.

Oke, A. M. C. *et al.* (2009) The use of TRMM satellite data as a predictor in the spatial interpolation of daily precipitation over Australia. In: *18th World IMACS / MODSIM Congress* (Cairns, Australia).

Pebesma, E. J. (2004) Multivariable geostatistics in S: the gstat package. *Computers & Geosciences* 30, 683–691.

Pegram, G. G. S. (2002) Spatial interpolation and mapping of rainfall: 2. Optimal integration of rain gauge, radar and satellite-derived data in the production of daily rainfall maps. In: *Progress Report to the Water Research Commission*, for the Period April 2001 to March 2002 on contract K5/1153.

Seto, S. *et al.* (2008) Simulations of flood detection all over Japan by using high-resolution satellite precipitation maps. *Ann. J. Hydraul. Engng* 52, 355–360 (in Japanese).

Todini, E. (2001) A Bayesian technique for conditioning radar precipitation estimates to raingauge measurements. *Hydrol. and Earth System Sci.* 5(2), 187–199.

Water Resources Agency, Ministry of Economic Affairs in Taiwan (2010) Challenge to flood disaster. Available: http://gweb.wra.gov.tw/Opac/mdl_wra/dfile.aspx?Dfile=ebook\s2010.pdf (in Chinese).

Yamaguchi, T. *et al.* (1993) Accuracy Improvement of Radar Raingauges. *Memoirs of River Information Research* 1, 91–103. (in Japanese).