Performance evaluation of Global Satellite Mapping of Precipitation (GSMaP) products over the Chaophraya River basin, Thailand

Watcharee Veerakachen¹, Mongkol Raksapatcharawong¹ and Shinta Seto²

¹Faculty of Engineering, Kasetsart University, Thailand

²Division of System Science, Nagasaki University, Japan

Abstract:

GSMaP_NRT (Near Real Time) is a viable tool to provide satellite-based precipitation data for further analysis. Its usefulness can be evident in the areas where continuous precipitation data is vital. This is why GSMaP_NRT performance has been evaluated globally. In this study, we evaluate its performance in terms of 1) rainfall detection capability based on Probability of Detection (POD) and False Alarm Ratio (FAR) and 2) estimation capability based on correlation coefficient and Root Mean Square Error (RMSE) over the Chaophraya River basin during 2009–2010. A non-realtime GSMaP_MVK (Moving Vector with Kalman filter) is also evaluated. Our results show that, at daily scale, both GSMaP_NRT/GSMaP_MVK performs well in rainy season (POD and FAR can reach 0.75/0.94 and 0.45/0.49, respectively) with acceptable RMSE of 14.64/14.23 mm. GSMaP_NRT tends to under-estimate whereas GSMaP_MVK slightly over-estimates the rain rates with correlation coefficient of 0.70 and 0.75. We conclude that GSMaP_NRT is considered good but not sufficient for near-realtime rainfall monitoring applications; whereas GSMaP_MVK is suitable for climate change studies.

KEYWORDS GSMaP; performance evaluation; rainfall detection; rainfall estimation

INTRODUCTION

Global climate change effects have been observed through occurrences of more frequent and more severe flood and drought cycles in Thailand during recent years. These events are considered as rain/no-rain triggered disasters. With precipitation data from reliable rainfall monitoring systems, inevitable losses and damages could be mitigated. Several near global coverage precipitation products using satellite observations have been developed worldwide such as Climate Prediction Center MORPHing technique – CMORPH (Joyce *et al.*, 2004), Precipitation Estimation from Remote Sensing Information using Artificial Neural Network – PERSIANN (Hsu *et al.*, 1997) and Global Satellite Mapping of Precipitation – GSMaP (Ushio *et al.*, 2009). For monitoring purpose, satellite precipitation products with high spatial and temporal resolution are of more interest. Most of them combine data from passive microwave (PM) and thermal infrared (TIR) sensors. Among these products is GSMaP which offers several versions of precipitation products for various applications. However, since rainfall

estimation models usually vary by climatic regions, a thorough accuracy assessment should be evaluated before applying these products for regional use. Although GSMaP_NRT products have been evaluated globally (Dinku *et al.*, 2010), the validation of GSMaP_NRT in Thailand has just started (Suwanprasert *et al.*, 2011); which showed that GSMaP_NRT overestimates the number of rainfall occurrences at daily time scale. However, for GSMaP_NRT to be effectively used for monitoring purpose, the validation of hourly rain rate is also necessary.

In this paper, two GSMaP precipitation products called GSMaP_NRT and GSMaP_MVK (the dataset v5.222 was used (gsmmap_mv5.222.20101101.0000.v5.222.dat.gz) which covers March 2000 to November 2010) are chosen for validation against rain gauge data at various time scales (hourly, daily, and monthly). The former is a near-realtime product for monitoring purpose whereas the latter is a more complicated, non-realtime product for climate change studies, e.g., drought and flood. The accuracy assessment was performed over a two-year period from 2009 to 2010.

STUDY AREA AND DATA SOURCES

Our study area is the Chaophraya River basin which is the most fertile paddy field area of Thailand. It is located between 13.5°–16.1°N in latitude and 99.5°–101°E in longitude with the area of 20,125 km². The basin is regarded as a principal waterway and is heavily affected by major floods and droughts. The rain gauge data used in this study were collected from 70 stations (locations shown in Figure 1) during the year 2009–2010 and were provided by Thailand Meteorological Department (TMD).

Among six GSMaP products, two versions of hourly GSMaP datasets were evaluated using rain gauge data. They use TIR derived motion vectors for propagating PM estimates in time and space using the Kalman filter approach, resulting in 0.1 degree spatial resolution and 1-hour temporal resolution. GSMaP_NRT, is the near-realtime version with 4-hour data latency that simply contains only the propagation process forward in time while GSMaP_MVK, the standard version, contains both forward and backward propagation processes to provide more accuracy. Their differences are summarized in Table I. Both datasets were selected and downloaded during the same period of rain gauge data.

Correspondence to: Watcharee Veerakachen, Faculty of Engineering, Kasetsart University, 50 Ngamwongwan Road, Ladyao Jatujak, Bangkok 10900 Thailand. E-mail: fengwrw@ku.ac.th ©2014, Japan Society of Hydrology and Water Resources.

Received 18 September, 2013
Accepted 28 December, 2013

METHODOLOGY

This study evaluates GSMaP performance in two aspects: rainfall detection capability based on rain/no rain events and rainfall estimate capability based on correlation coefficient and root mean square error (RMSE) of rain rates. Rainfall detection evaluation starts with defining four scenarios to express the relationship between rain/no rain events detected

by GSMaP_NRT and GSMaP_MVK and observed by rain gauges as shown in Table II. Consequently, performance indices which are accuracy (ACC), probability of detection (POD), and false alarm ratio (FAR) can be developed. Their expression and description are shown in Table III.

Rainfall estimate evaluation is based on a correlation and RMSE between rain rates estimated by GSMaP_NRT and GSMaP_MVK and observed by rain gauges located in

Figure 1. Thailand meteorological department rain gauge network consisting of 70 stations in the study area

Table I. Differences between GSMaP_NRT and GSMaP_MVK

Items	GSMaP_NRT	GSMaP_MVK
Input data availability	Input data which is available within 3 hours from observation	All available data
Input data: GEO Infrared radiometer	Globally-merged pixel-resolution data by JWA	Globally-merged full-resolution data by NOAA/CPC
Look-up table for the GSMaP microwave imager algorithm	Calculate once a day at around 3UTC and use nearest GANAL and MGDSST as inputs	Calculate once a day (00UTC) and use nearest to observation time
Look-up table for the GSMaP microwave sounder algorithm	N/A	Calculate once a day (00UTC)
The GSMaP microwave-IR combined algorithm	A simplified version of the algorithm (morphing and Kalman filtering by forward process)	A full version of the algorithm (morphing and Kalman filtering by forward and backward processes)

* This table is excerpted from GSMaP_MVK v5.222 User Guide (JAXA/EORC, 2012)

Table II. Scenarios description

		Observed by rain gauges	
		Rain	No Rain
Detected by GSMaP_NRT/GSMaP_MVK	Rain	Hit (<i>a</i>)	False alarm (<i>b</i>)
	No Rain	Miss (<i>c</i>)	Correct negative (<i>d</i>)

the same 0.1 deg grid interval (or pixel). Monthly rain scatter plot (1 point/gauge/month) is used to find the correlation coefficient between rain gauge data versus GSMaP_NRT and GSMaP_MVK products. At daily scale, there are two parts of analysis. Firstly, performance indices grouped by month are plotted to indicate how they exhibit over a number of years. Secondly, performance indices are divided into four cases: 2009, 2010, 2009_rain, and 2010_rain (the last two cases focus on the rainy season from May to October) to see how they exhibit during the rainy season compared to the whole year. At hourly scale, performance indices based on a similar four cases as daily scale are shown. These finer grain indices are substantial indicators for near-realtime applications.

RESULTS AND DISCUSSION

Figure 2 shows the monthly rain scatter plots of data from rain gauge versus GSMaP_NRT and GSMaP_MVK whose correlation coefficients are 0.70 and 0.75, respectively. From the zero-crossing trend line, GSMaP_MVK is more likely to detect rainfall than GSMaP_NRT. This is due to higher POD and FAR of GSMaP_MVK as will be discussed later.

Figure 3 (a)–(c) elaborates on the daily rainfall detection performance. We can see that both GSMaP_NRT and GSMaP_MVK share similar trends. In rainy season, more

rain events are observed and detected, hence, hits improve significantly whereas false alarms increase at lower rates. Table IV illustrates that GSMaP_MVK can detect more rain events which lead to higher hits and false alarms compared to GSMaP_NRT. This results in higher POD (>0.7 for GSMaP_NRT and >0.9 for GSMaP_MVK) and lower ACC (<0.55 for GSMaP_NRT and <0.5 for GSMaP_MVK). FAR also drops to nearly 0.3 for both due to improved hits. In general, GSMaP_MVK yields better rainfall detection performance confirming its effectiveness.

Figure 3 (d) shows the daily average RMSE which increases to more than 10 mm in rainy season for both GSMaP_NRT and GSMaP_MVK. This is because correct negatives turn into hits, false alarms, and misses as depicted in Table IV. Note that GSMaP_MVK achieves noticeably lower RMSE in most cases except during Sep–Oct 2010 possibly due to heavy rainfalls which caused flooding in October 2010.

Table V shows the daily statistics developed from occurrences in Table IV. The more complex algorithm of GSMaP_MVK has resulted in more hits and false alarms and becomes more evident during the rainy season. If we deliberate on each pair of scenarios, namely Hit/Miss and False alarm/Correct neg., we can deduce that GSMaP_MVK increases one and reduces the other at almost the same percentage. The first pair improves rainfall detection capability and therefore yields better POD. Nevertheless, ACC slightly decreases as the second pair accounts for more

Table III. Description of ACC, POD, and FAR

Index	Expression	Implication
ACC	$ACC = \frac{a + d}{a + b + c + d}$	The fraction of the correct detection to the number of overall events
POD	$POD = \frac{a}{a + c}$	The ratio of correct rain detection to the number of rain events observed
FAR	$FAR = \frac{b}{a + b}$	The fraction of rain detection that turn out to be wrong

Figure 2. Monthly rain scatter plots of rain gauge versus GSMaP_NRT (left) and GSMaP_MVK (right) indicating that GSMaP_NRT is underestimated whereas GSMaP_MVK is overestimated

Figure 3. Daily performance indices grouped by month. (a) ACC, (b) POD, (c) FAR, and (d) Average RMSE

Table IV. Percentage of scenarios based on daily rain

Scenarios	2009		2010		2009_rain		2010_rain	
	GSMaP_NRT	GSMaP_MVK	GSMaP_NRT	GSMaP_MVK	GSMaP_NRT	GSMaP_MVK	GSMaP_NRT	GSMaP_MVK
Hit	17.6%	22.4%	16.5%	22.4%	29.9%	38.0%	30.6%	38.5%
False alarm	17.9%	24.1%	16.1%	24.8%	24.6%	36.7%	25.9%	39.1%
Miss	7.5%	2.7%	6.3%	2.0%	11.6%	3.5%	10.2%	2.3%
Correct negative	57.1%	50.9%	61.0%	50.8%	33.9%	21.8%	33.3%	20.1%
Total	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

Table V. Performance indices based on daily rain

Indices	2009		2010		2009_rain		2010_rain	
	GSMaP_NRT	GSMaP_MVK	GSMaP_NRT	GSMaP_MVK	GSMaP_NRT	GSMaP_MVK	GSMaP_NRT	GSMaP_MVK
ACC	0.75	0.73	0.78	0.73	0.64	0.60	0.64	0.59
POD	0.70	0.89	0.72	0.92	0.72	0.92	0.75	0.94
FAR	0.50	0.52	0.50	0.53	0.45	0.49	0.46	0.50
RMSE (mm)	9.44	8.00	10.67	10.76	12.11	10.24	14.64	14.23

Table VI. Percentage of scenarios based on hourly rain

Scenarios	2009		2010		2009_rain		2010_rain	
	GSMaP_NRT	GSMaP_MVK	GSMaP_NRT	GSMaP_MVK	GSMaP_NRT	GSMaP_MVK	GSMaP_NRT	GSMaP_MVK
Hit	0.9%	2.2%	1.2%	2.6%	1.6%	3.8%	2.4%	4.6%
False alarm	5.3%	9.0%	4.9%	9.8%	8.5%	15.5%	8.8%	16.9%
Miss	2.1%	0.9%	1.9%	0.8%	3.7%	1.5%	3.5%	1.3%
Correct neg.	91.7%	87.9%	92.0%	86.8%	86.2%	79.2%	85.2%	77.2%
Total	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

Table VII. Performance indices based on hourly rain

Indices	2009		2010		2009_rain		2010_rain	
	GSMaP_NRT	GSMaP_MVK	GSMaP_NRT	GSMaP_MVK	GSMaP_NRT	GSMaP_MVK	GSMaP_NRT	GSMaP_MVK
ACC	0.93	0.90	0.93	0.89	0.88	0.83	0.88	0.82
POD	0.30	0.72	0.39	0.77	0.31	0.72	0.40	0.78
FAR	0.85	0.81	0.80	0.79	0.84	0.80	0.79	0.79
RMSE (mm)	1.38	1.27	1.51	1.58	1.77	1.63	2.05	2.09

Table VIII. Comparing performance indices of GSMaP_NRT in 2010

	Daily Scale			Hourly Scale	
	ACC	POD	FAR	RMSE	Correlation coefficient
This work (70 TMD rain gauge stations)	0.78	0.72	0.50	1.51	0.09
Suwanprasert <i>et al.</i> (2011) (17 TMD rain gauge stations with Inverse Distance Weighting (IDW) interpolation)	0.68	0.28	0.81	1.16	0.02

than 70% (40% in rainy season). As for FAR, this index improves accordingly to hits. In terms of rainfall estimation capability, GSMaP_MVK performs very well in keeping RMSE comparable to that of GSMaP_NRT considering its higher POD (especially during the rainy season in 2010 where heavy flooding occurred). This means GSMaP_MVK can better estimate high-variance seasonal rainfalls.

Table VII shows the hourly statistics developed from occurrences in Table VI. GSMaP_MVK still behaves similarly to the daily statistics, especially its POD which is almost double that of GSMaP_NRT. For applications requiring near-realtime monitoring such as floods and landslides, GSMaP_MVK is considered a good candidate when its data can be available in realtime whereas GSMaP_NRT alone is not sufficient for this purpose.

Table VIII compares our results with Suwanprasert *et al.* (2011). Evidently, using a different number of rain gauges can produce quite different results. Using Inverse Distance Weighting (IDW) interpolation with only 17 gauge stations may be insufficient to represent rainfall observation over the basin. On one hand, rain events that could have been observed with rain gauges were incorrectly interpolated as no-rain events—hits were turned into false alarms. This explains lower ACC and POD, and much higher FAR (even in the rainy season). The extremely low coefficient implies that GSMaP_NRT is not suitable for hourly rainfall estimation (i.e., insufficient for near-realtime applications).

CONCLUSIONS

This work evaluates performance of GSMaP_NRT and GSMaP_MVK using rain gauge data over the Chaophraya River basin at various temporal scales. From monthly rain scatter plot, GSMaP_NRT is under-estimated, but GSMaP_MVK is slightly over-estimated with correlation coefficient of 0.70 and 0.75, respectively. From daily rain statistics, both GSMaP_NRT and GSMaP_MVK perform better in terms of POD and FAR during the rainy season. RMSE is higher as a result of less correct negatives. However, GSMaP_MVK outperforms GSMaP_NRT in

terms of POD and RMSE because of its complex algorithm. This advantage becomes more evident at hourly rain statistics where POD almost doubles (up to 0.78). Although GSMaP_MVK performance is suitable for near real-time applications such as flood monitoring, its computation latency is prohibitive. Alternatively, GSMaP_NRT can be used but is insufficient for this purpose.

ACKNOWLEDGMENTS

This study was supported by the Science and Technology Research Partnership for Sustainable Development (JST-JICA) under the Integrated Study on Hydro-Meteorological Prediction and Adaptation to Climate Change in Thailand (IMPAC-T) project. The authors thank Thailand Meteorology Department (TMD) for their kind provision of rain gauge data.

REFERENCES

- Dinku T, Ruiz F, Connor JS, Ceccato P. 2010. Validation and intercomparison of satellite rainfall estimates over Colombia. *Journal of Applied Meteorology and Climatology* **49**: 1004–1014. DOI: 10.1175/2009JAMC2260.1.
- Hsu KL, Gao X, Sorooshian S, Gupta HV. 1997. Precipitation estimation from remotely sensed information using artificial neural networks. *Journal of Applied Meteorology and Climatology* **36**: 1176–1190. DOI: 10.1175/1520-0450(1997)036<1176:PEFRSI>2.0.CO;2.
- Japan Aerospace Exploration Agency/Earth Observation Research Center (JAXA/EORC). 2012. User's Guide For Global Satellite Mapping of Precipitation Microwave-IR Combined Product (GSMaP_MVK) Version 5.222, Japan, 7 pp.
- Joyce RJ, Janowiak JE, Arkin PA, Xie P. 2004. CMORPH: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. *Journal of Hydrometeorology* **5**: 487–503. DOI: 10.1175/1525-7541(2004)005<0487:CAMPGP>2.0.CO;2.
- Suwanprasert K, Seto S, Kaewrueng S. 2011. Validation of GSMaP

rainfall estimates with local rainfall gauge in Thailand. *Proceedings of the 9th International Symposium on Southeast Asian Water Environment December 1–3, 2011 Bangkok, Thailand*. 65–72.

Ushio T, Sasashige K, Kubota T, Shige S, Okamoto K, Aonashi K, Inoue T, Takahashi N, Iguchi T, Kachi M, Oki R, Morimoto

T. 2009. A Kalman filter approach to the Global Satellite Mapping of Precipitation (GSMaP) from combined passive microwave and infrared radiometric data. *Journal of the Meteorological Society of Japan* **87A**: 137–151. DOI: 10.2151/jmsj.87A.137.