STATISTICAL ANALYSIS OF ANNUAL RAINFALL PATTERNS IN PENINSULAR MALAYSIA USING TRMM ALGORITHM

Tam Tze Huey[a] and Ab Latif Ibrahim[b]

[a] Graduate Student, Department of Remote Sensing, Faculty of Geoinformation and Real Estate,
Universiti Teknologi Malaysia, 81310 UTM Johor Bahru, Johor, Malaysia;
Tel: +6(0) -17-6618991; Fax +6(0)-75566163
E-mail: thtam2@live.utm.my

[b]Associate Professor, Institute of Geospatial Science & Technology, Faculty of Geoinformation and Real Estate,
Universiti Teknologi Malaysia, 81310 UTM Johor Bahru, Johor, Malaysia;
Tel +6(0)-17-7570489; Fax +6(0)-75566163
E-mail: ablatif@utm.my

KEY WORDS: Rainfall, TRMM 3B43, Statistical Analysis, Rain Gauge

ABSTRACT: Water is one of the most important of natural resource for all living things. In certain areas, water tends to be very limited and there are situation where the availability of water is always deficient. This phenomena will altered by the effect of global warming. The main source of water is rainfall. Understanding of rainfall characteristics is very essential. The main objective of this study is to evaluate the reliability of Tropical Rainfall Measuring Mission (TRMM) 3B43 algorithm to study Peninsular Malaysia's rainfall distribution. Ten years of TRMM 3B43 data from 1998 to 2007 were validated by 24 rain gauge measurements. Four statistical parameters were used to determine the reliability of TRMM data. Results of this study show that TRMM data is overestimated and low accuracy. It might be due to the echoes error of Precipitation Radar (PR) on TRMM satellite. The TRMM data is still be considered reliable to analysis the annual Peninsular Malaysia rainfall pattern. Further research should be proceeded to analyze in-depth the results and enhance the reliability of TRMM data.

INTRODUCTION

Water is one of our most valuable natural resources and vital to all form of life (Sowry, 1976). Water also used for transportation, source of power, and serves many other useful purposes for domestic consumption, agriculture and industry. The main important source of water in any areas is rain. The amount or availability of water for various purposes is very much depending upon the amount of precipitation in that particular area. Excess or extended absence of rainfall will cause flooding and drought respectively. Adler *et al.*, (2000) stated that precipitation information is essential to understanding the hydrologic balance on a global scale and in understanding the complex interactions among the components within the hydrologic cycle.

Rainfall measured by rain gauge provides a fairly accurate rainfall measurement (Sinclair and Pegram, 2005). Rain gauge is an instrument that collects the rain by capturing a volume over a continuous or fixed time interval (Strangeways, 2007). The numbers of rain gauge are limited with less equipped in remote area and rainfall is high spatial and temporal variability. Therefore, rainfall data obtained by remote sensing technique is been widely used now because this technique was successfully in overcoming the major shortcoming of rain gauge measurement (Yilmaz *et al.*, 2005; Chanyatham and Kirtsaeng, 2011).

Methods of estimating rainfall by remote sensing technique can be classified into two groups: indirect and direct method (Carleton 1991; Barrett & Beaumont 1994). For the indirect method, cloud characteristics (cloud-top temperature) observed in visible (VIS) and infrared (IR) satellite imagery are used as indicators of the occurrence of precipitation. The direct method applied microwave (MW) techniques to obtain instantaneous rain rates. Surface-based remote sensing principal involves weather radar, though this is limited to relatively small areas and is mainly used for forecasting and research purposes (Simmers, 2005).

According to Jobard (2001), indirect or direct estimating rainfall has its own advantage. Visible and infrared technique is mostly appropriate for large and climatic scales. Microwave technique is suitable for instantaneous rain retrieval. Combined infrared and microwave technique performed better as estimate monthly rainfall averages. When estimate spatial averaged rainfall and time accumulated amounts, Jobard (2001) recommended using methods that combine multi-satellite source data, because they blend the physically-based gain information available from microwave measurement with geostationary-based infrared measurement to capture the space-time evolution of precipitating clouds (Jobard 2001). Therefore, TRMM 3B43 algorithm was chosen in this study.

The Tropical Rainfall Measuring Mission (TRMM), jointly sponsored by the National Aeronautics and Space Administration (NASA) of the United States and the Japanese Aerospace Exploration Agency (JAEA, formally National Space Development Agency (NASDA) of Japan), is the first coordinated international satellite mission to study tropical and sub-tropical rain systems. TRMM satellite is variation from a low-inclination orbit combining a suite of sensors to overcome many of the limitations of remote sensors previously used for such measurements from space. The TRMM provides visible, infrared, and microwave observations of tropical and subtropical rain systems, as well as lighting and cloud and radiation measurements. The satellite observations area complemented by ground radar and rain gauge measurements to validate satellite rain estimation algorithms (Simpson, 1988).

TRMM satellite flies in a low inclination (35 degree), non-sun-synchronous, and highly precessing orbit, which allows it to fly over each position on the Earth' surface at a different local time each day so as to achieve the TRMM objectives. The TRMM satellite has five sensors on board. The five sensors include Precipitation Radar (PR), TRMM Microwave Imager (TMI), Visible Infrared Scanner (VIRS), Lighting Imaging Sensors (LIS) and Cloud's and Earth's Radiant Energy System (CERES). Lightning Imaging Sensor (LIS) and a Clouds and Earth's Radiant Energy System (CERES) are carried on the TRMM satellite. The LIS is a calibrated optical sensor operating at 0.7774 µm and observes distribution and variability of lightning. The horizontal resolution of LIS at nadir is 5 km and the swath width is 590 km. The CERES is a visible/ infrared sensor which measures emitted and reflected radiative energy from the surface of the Earth and the atmosphere and its constituents. The TRMM CERES operates at 0.3 to 5.0 µm in the shortwave range and 8.0 to 12.0 µm in the longwave range. LIS and CERES data are available from the NASA Global Hydrology Resource Center (http://ghrc.msfc.nasa.gov) and the NASA Langley Research Center (http://asd-www.larc.nasa.gov/ceres/ASDceres.html) respectively. Table 1 shows TRMM sensors performance characteristics and Figure 1 shows TRMM sensors.

Table 1: TRMM sensors performance characteristics.

	Precipitation Radar	TRMM Microwave Imager	Visible and Infrared Scanner
Frequency/Wavelength	Vertical Polarization: 13.8 GHz	Dual Polarization: 10.65, 19.35, 37, 85.5 GHz; Vertical: 21 GHz	0.63, 1.6, 3.75, 10.8 & 12 µm
Scanning Mode	Cross Track	Conical	Cross Track
Ground Resolution	4.3 km (*5.0 km) at nadir	4.4 km (*5.1 km) at 85.5 GHz	2.2 km (*2.4 km) at nadir
Swath Width	215 km (*247 km)	760 km (*878 km)	720 km (*833 km)
Science Applications	3-D rainfall distribution over both land and oceans, and latent heat release into the atmosphere	Surface rain rate type, distribution, and structure; other atmospheric and oceanic parameters	Cloud parameters, fire, pollution

Note: The TRMM satellite was boosted from 350 to 402 km altitude in August 2001.

* Numbers in parentheses represent post boost values

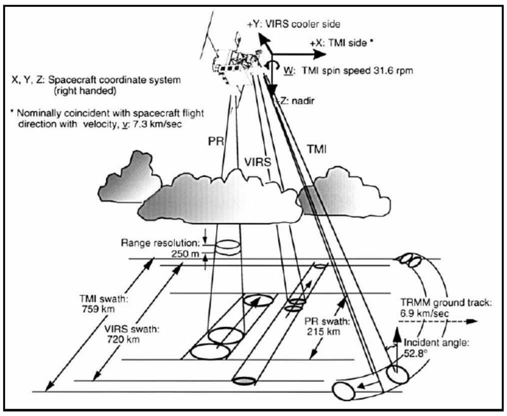

Figure 1: TRMM sensor.

STUDY AREA

Malaysia located in Southeast Asia. There are two main lands that separated by South China Sea where Peninsular Malaysia to the west and East Malaysia to the east. Peninsular Malaysia is located south of Thailand, north of Singapore and east of the Indonesian island of Sumatra. East Malaysia is located on the island of Borneo and shares borders with Brunei and Indonesia. Figure 2 show the location of Malaysia. Malaysia' climate is categorized as equatorial, therefore, being hot and humid throughout the year with generally light winds. The average rainfall is 2500 mm and the average temperature is 27 °C a year. The country is therefore rich in water resources as compared to the other regions of the world.

Malaysia is governed by two monsoon seasons, i.e. Northeast Monsoon and Southwest Monsoon. The Northeast Monsoon started from early November to March, originating from China and the north Pacific, brings heavy rainfall to the east coast states of the Peninsular Malaysia. The consequence is frequently causes those area have widespread floods. It also causes the wettest season in East Malaysia (Sabah and Sarawak). The Southwest Monsoon (from the deserts of Australia) from late of May and ends in September, and is dried period for the whole country. April and October form transitions between the two monsoons called as inter-monsoon and bring heavy rainfall.

The average annual water resources on a total land mass of 330,000 km^2 amount to 990 billion m^3. Out of which, 360 billion m3, or 36% returns to the atmosphere as evapotranspiration, 566 billion m^3, or 57% appear as surface runoff and the remaining 64 billion m^3, or 7% go to the recharge of groundwater. Of the total 566 billion m^3 of surface runoff, 147 billion m^3 are found in Peninsular Malaysia, 113 billion m^3 in Sabah and 306 billion m^3 in Sarawak.

Figure 2: Geography of Malaysia.

DATA AND METHODS

Two main types of data were used in this study, i.e. rain gauge data and satellite estimated data. Total of 24 rain gauges data in monthly format from 1998 to 2007 were provided by Malaysian Meteorological Department (MMD). These data were used for validating of TRMM data by using statistical approaches.

TRMM and Other Data Precipitation Product (TRMM Product 3B43) version 6 data were used in this study and this data are freely accessed at http://disc2.nascom.nasa.gov/Giovanni/tovas/. Besides TRMM 3B43, this website provide other data, e.g. GPCP, TRMM 3B42 and etc.

The purpose of Algorithm 3B-43 is to produce the "Tropical Rainfall Measuring Mission (TRMM) and Other Data" best-estimate precipitation rate and root-mean-square (RMS) precipitation-error estimates. These gridded estimates are on a calendar month temporal resolution and a 0.25-degree by 0.25-degree spatial resolution global band extending from 50 degrees South to 50 degrees North latitude. Algorithm 3B-43 is executed once per calendar month to produce the single, best-estimate precipitation rate and RMS precipitation-error estimate field (3B-43) by combining the 3-hourly merged high-quality/IR estimates with the monthly accumulated Climate Assessment and Monitoring System (CAMS) or Global Precipitation Climatology Centre (GPCC) rain gauge analysis (3A-45). The 3-hourly merged high quality/IR estimates are summed for the calendar month, and then the rain gauge data are used to apply a large-scale bias adjustment to the 3B-42 estimates over land. The monthly gauge-adjusted merged estimate is then combined directly with the rain gauge estimates using inverse error variance weighting (Tropical Rainfall Measuring Mission, Global Space Flight Center).

The downloaded TRMM 3B43 data was saved in ASCII file, which store 3 important types of data, i.e. coordinate (X and Y) and accumulated rainfall data (mm). This ASCII file can read by Microsoft Excel. It became well distributed point format after imported into ArcGIS 9.3 software. Interpolation technique under spatial analyst toolbox in ArcGIS is responsible for conversion of point to raster. Kriging technique was adopted for the interpolating point data to spatial data because this technique provides more reliable results than other interpolation methods, and the standard error can be stated (Haberlandt, 2006; Philips *et al.*, 1992).

The spatial rainfall distribution was overlaid on the rain gauge stations in vector format (point) in order to acquire mean precipitation estimates based on the coordinate of rain gauge stations. Comparison between rain gauge data and the mean precipitation estimates was carried out by using 4 statistical formulas with the help of Microsoft Excel.

The 4 types of statistical parameters include, coefficient correlation, bias or mean error, mean absolute error and root mean square error. The correlation coefficient (*r*) is one of the most common and most useful statistics. A correlation is a single number that describes the degree of relationship between two variables.

The correlation coefficient is defined as follows:

$$r^2 = \frac{Cov(SD_i, GD_i)}{\sigma_{SD}\sigma_{GD}}$$

The mean error (ME), also called additive bias indicative the average direction of the deviation from observed values, but may not reflect the magnitude of the error. It measures the average error of a number of observations found by taking the mean of the taking the mean value of the positive and negative errors without regard to sign.

The mean error is defined as follows:

The mean absolute error (MAE) measures the average magnitude of the errors in a set of estimated value, without considering their direction. It measures accuracy for continuous variables.

The mean absolute error is defined as follows:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|SD_i - GD_i| \qquad 3$$

The root mean square error (RMSE) is a quadratic scoring rule which measures the average magnitude of the error. Compared to the MAE, the RMSE gives greater weight to large errors than to small errors in the average.

The root mean square is defined as follows:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(SD_i - GD_i)^2} \qquad 4$$

In all the previous equations, n is the total number of stations, SD is the sensor data – TRMM Microwave Radiometer Surface Rainfall Retrievals – and GD is the rainfall gauge data, and $\overline{SD}$, $\overline{GD}$ are the average of the sensor and gauge data.

RESULTS AND ANALYSIS

Comparison between mean precipitation estimates and rain gauge data was carried out. The mean precipitation estimates given similar tendency in term of statistic results. Table 1 showed that the mean precipitation estimates is overestimated. The MAE's value relatively large, therefore, the accuracy of either both data is quite low. Based on correlation coefficient, some rain gauge stations have good correlation with the mean precipitation estimates.

There are reasons can be explained the large deviation between the rain gauge data and the mean precipitation estimates. Rain gauge data, basically, subjected to two errors, i.e., systematic error and random error. The systematic errors include losses due to wind, wetting, evaporation, and splashing (Habib *et al.*, 2008). Another source of local TB gauge errors is due to dynamic calibration effects. TB gauges are typically calibrated by the manufacturer at a fixed rainfall rate that is typically low (few mm per h); however, such static calibration does not guarantee the conformance of the tipping-bucket size with the nominal calibration volume at high rainfall rates. Due to the measurement principle of the TB gauges, some of the incident rainfall is missed between successive tips of the bucket, which results in an underestimation of actual rainfall volumes and intensities. This underestimation is negligible at small rainfall rates, but grows nonlinearly with the increase of rainfall rates (Niemczynowicz 1986; Humphrey *et al.*, 1997; Luyckx and Berlamont 2001).

Besides these systematic errors, TB measurements are also associated with local random errors (Fankhauser 1997; Yu *et al.*, 1997; Nystuen and Proni 1996). These errors are mainly related to the discrete sampling mechanism of the TB gauge and are caused by uncertainties in defining start and end of rain event, partial filling of the bucket and instabilities in the water inflow into the collecting funnel. Ciach (2003) and Habib *et al.*, (2001) showed that such random errors have significant magnitudes, mainly at small rainfall intensities and short time scales.

Error also arises in the estimation of rainfall using satellite data. Echoes of scatters produced by active precipitation radar (PR) on the TRMM satellite include rain echo and surface echo (Nirala and Cracknell, 1998; Strangeways, 2007). Nirala and Cracknel, (1998) studied that the surface echoes are much stronger than rain echoes. This causing a problem to discriminate the rain echoes from the surface echoes. In the other word, rain echoes received by satellite contain surface echoes. Therefore, surface echoes give rise to another influence on the quality of satellite data.

According to Adeyewa and Nakamura (2003) and Franchito *et al.*, (2009), for considerably large number of grid points, the percentage standard error, or percentage standard deviation, reduces to relative rmse or % rmse. The latter has, therefore, been used to evaluate the reliability of TRMM algorithm in the different zones under consideration and their geographical and seasonal variations. When the rmse of an estimates algorithm is less than 50% of measured rainfall amount, such estimates algorithm is considered to be reliable in relative term. In contrast, when the rmse was equal to or greater than 50% of the magnitude of the reference rainfall, the estimates algorithm was considered unreliable for the region and for the particular season.

Annual rainfall received by 24 rain gauges, usually nearly or above 2000 mm. Fifty percent of the measured rainfall is approximately 1000 mm. Value of rmse in Table 1 mostly is below 1000 mm except rain gauge station at Temerloh and Batu Embun. RMSE value of Temerloh and Batu Embun were 1073.05mm and 929.784mm respectively, which is greater than and nearly equal to 50% of the measured rainfall.

Table 1: Statistical parameters for annually (the statistical analysis was calculated for the period of 1998 to 2007).

Station Name	r	BIAS [mm]	MAE [mm]	RMSE [mm]
Chuping	0.259	296.467	367.30	440.536
Alor Setar	-0.226	166.804	363.96	527.481
Butterowrth	0.907	126.826	203.77	219.048
Ipoh	0.789	-208.830	317.78	379.397
Cameron Highlands	0.406	-328.960	372.44	483.079
Subang	0.923	111.646	148.21	353.056
Petaling Jaya	0.832	-302.960	334.66	369.073
KLIA Sepang	0.416	546.153	546.15	631.852
Malacca	0.498	325.631	335.75	403.729
Batu Pahat	0.838	615.948	615.95	647.786
Senai	0.752	235.581	321.53	367.481
Kluang	0.921	708.483	708.48	733.856
Mersing	0.648	-235.980	388.20	464.499
Muadzam Shah	0.856	560.683	560.68	621.771
Temerloh	0.604	1013.970	1013.97	1073.05
Kuantan	0.876	-270.870	284.55	332.992
Batu Embun	0.783	907.400	907.40	929.784
Kuala Terengganu Airport	0.823	-26.707	198.49	288.672
Kota Bahru	0.934	20.571	178.54	212.249
Pulau Langkawi	0.261	-371.910	470.22	544.993
Bayan Lepas	0.844	-21.686	204.60	244.49
Kuala Krai	0.355	548.821	555.52	707.343
Sitiawan	0.583	455.665	455.67	511.043
Lubok Merbau	0.118	629.610	629.61	701.294

CONCLUSION

The following conclusions can be drawn based on the findings of the study. This study was a very direct and simple statistical analysis for comparing the 11-years mean precipitation estimates and rain gauge data. The main objective of this study is to evaluate the reliability of TRMM 3B43 algorithm applied in this country. Statistical results showed that TRMM 3B43 algorithm was reliable to be used.

In this study, some assumptions was explained the errors of rain gauge and also for the satellite rainfall estimation. Therefore, the discrepancy of both rainfall estimation methods is very large. In the future, study of the occurrence of errors is becoming very important in order to improve the satellite rainfall estimation and rain gauge observation.

REFERENCES

Adeyewa, Z. D. and K. Nakamura, (2003). Validation of TRMM Radar Rainfall Data over Major Climatic Regions in Africa. *Journal of Applied Meteorology and Climatology*. 42(2), 331-347.

Adler. R. F., Huffman, G. J., Bolvin, D. T., Curtis, S., and Nelkin, E. J. (2000). Tropical Rainfall Distributions Determined Using TRMM Combined with other Satellite and Rain Gauge Information. *Journal of Applied Meteorology*. 39, 2007-2023. doi: http://dx.doi.org/10.1175/1520-0426(2003)20<752:LREITB>2.0.CO;2

Chanyatham, T. and Kirtsaeng, S. (2011). Comparison and Analysis of Remote Sensing-based and Ground-based Precipitation Data Over India. *Chiang Mai Journal of Science*. 38(4), 541-550.

Ciach, G. J. (2003). Local random errors in tipping-bucket rain gauge measurements. *Journal of Atmospheric Ocean Technology*, 20(5), 752–759.

Fankhauser, R. (1997). Measurement properties of tipping bucket rain gauges and their influence on urban runoff simulation. *Water Science and Technology*, 36(8–9), 7–12.

Franchito, S. H., Brahmananda, R., Vasques, A. C., Santo, C. M., and Conforte, J. C. (2009). Validation of TRMM PR Monthly rainfall over Brazil. *Journal of Geophysical Research.* 114: D02105, DOI:10.1029/2007JD009580.

Haberlandt, U. (2006). Geostatistical Interpolation of Hourly Precipitation from Rain Gauge and Radar or a Large-Scale Extreme Rainfall Event. *Journal of Hydrology*, doi:10.1016/j.jhydrol.2006.06.028.

Habib, E., Krajewski, W. F., and Kruger, A. (2001). Sampling errors of tipping-bucket rain gauge measurements. *Journal of Hydrology Engineering*, 6(2), 159– 166.

Habib, E. H., Meselhe, E. A., and Aduvala, A. V. (2008). Effect of Local Errors of Tipping-Bucket Rain Gauges on Rainfall-Runoff Simulations. *Journal of Hydrologic Engineering*. 13(6), 488-496, DOI: 10.1061/(ASCE)1084-0699(2008)13:6(488).

Humphrey, M. D., Istok, J. D., Lee, J. Y., Hevesi, J. A., and Flint, A. L. (1997). A new method for automated dynamic calibration of tipping bucket rain gauges. *Journal of Atmopsheric and Oceanic Technology*, 14(6), 1513–1519.

Luyckx, G., and Berlamont, J., (2001). Simplified method to correct rainfall measurements from tipping bucket rain gauges. *Proceeding Urban Drainage Modeling*, ASCE, 767–776.

Niemczynowicz, J. (1986). The dynamic calibration of tipping-bucket raingauges. *Nordic Hydrology*, 17(3), 203–214.

Nirala, M. L., and Cracknell, A. P. (1998). Rainfall Estimation Using TRMM Satellite Data. 1998 IEEE International Geoscience and Remote Sensing Symposium Proceedings: Volume IV. 6-10. July Sheraton Seattle, Seattle, WA, USA,135-137.

Nystuen, J. A., and Proni, J. R. (1996). A comparison of automatic rain gauges. *Journal of Atmospheric and Oceanic Technology.* 13(1), 62–73. doi: http://dx.doi.org/10.1175/1520-0426(1996)013<0062:ACOARG>2.0.CO;2

Philips, D. L., Dolph, J., and Marks, D. (1992). A Comparison of Geostatistical Prcedures for Spatial Analysis of Precipitation in Mountainous terrain. Agricultural and Forest *Meteorology*. 58, 119-141.

Sinclair, S., and Pegram, G. (2005). Combining radar dan rain gauge rainfall estimates using conditional merging. *Atmospheric Sciences Letters*. 6(1), 19-22.

Strangeways, I. *Precipitation: theory, measurement and distribution*. Cambridge: Cambridge University Press. 2007.

Tropical Rainfall Measuring Mission, Global Space Flight Center. *Algorithm 3B43 – TRMM and Other Data Precipitation*. Retrieved January 12, 2010, from http://trmm.gsfc.nasa.gov/3b43.html.

Yilmaz, K. K., Hogue, T. S., Hsu, K. L., Sorooshian, S., Gupta, H. V., and Wagener, T. 2005. Intercomparison of Rain Gauge, Radar, and Satellite-Based Precipitation Estimates with Emphasis on Hydrologic Forecasting. *American Meteorological Society*. 6, 497-517.

Yu, B., Ciesiolka, C. A. A., Rose, C. W., and Coughlan, K. J. (1997). A note on sampling errors in the rainfall and runoff data collected using tipping bucket technology. *Transactions of the ASAE*, 40(5), 1305–1309.