Downscaling the Satellite Precipitation Data in Humid Tropical Region

Mohd. Rizaludin bin Mahmud

2015

Doctoral Dissertation

Tokyo Metropolitan University
FIGURES

<table>
<thead>
<tr>
<th>No.</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Peninsular Malaysia and its seasonal rainfall variation</td>
<td>8</td>
</tr>
<tr>
<td>2.2</td>
<td>Rain gauge distribution in Peninsula Malaysia</td>
<td>9</td>
</tr>
<tr>
<td>3.1</td>
<td>Ratio between 0.25º and 0.125º TRMM-rain gauge validation. (a)</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td>Northeast monsoon (NEM); (b) Inter-monsoon 1 (IM1).</td>
<td></td>
</tr>
<tr>
<td>3.2</td>
<td>Seasonal correlation of the TRMM 3B43 against ground rainfall at</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td>local climate region of Peninsular Malaysia. The line represents the</td>
<td></td>
</tr>
<tr>
<td></td>
<td>co-efficient of variation (COV) of areal rainfall surfaces</td>
<td></td>
</tr>
<tr>
<td>3.3</td>
<td>Temporal maps of seasonal correlation of the TRMM 3B43 against</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td>ground rainfall. (a) NEM : Northeast monsoon, (b) IM1 : Inter-monsoon</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1, (c) SWM : Southwest monsoon, (d) IM2 : Inter-monsoon 2</td>
<td></td>
</tr>
<tr>
<td>3.4</td>
<td>Average seasonal RMSE of the TRMM 3B43 against ground rainfall in</td>
<td>28</td>
</tr>
<tr>
<td></td>
<td>local climate regions and average rainfall from areal rainfall</td>
<td></td>
</tr>
<tr>
<td></td>
<td>surfaces for the period 1998–2010</td>
<td></td>
</tr>
<tr>
<td>3.5</td>
<td>Temporal maps of seasonal RMSE of the TRMM 3B43 against ground</td>
<td>28</td>
</tr>
<tr>
<td></td>
<td>rainfall. (a) NEM : Northeast monsoon, (b) IM1 : Inter-monsoon 1,</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(c) SWM : Southwest monsoon, (d) IM2 : Inter-monsoon 2</td>
<td></td>
</tr>
<tr>
<td>3.6</td>
<td>Average seasonal ratio values between TRMM rainfall estimates and</td>
<td>29</td>
</tr>
<tr>
<td></td>
<td>ground rainfall in local climate regions</td>
<td></td>
</tr>
<tr>
<td>3.7</td>
<td>Temporal maps of seasonal ratios of the TRMM 3B43 over ground</td>
<td>29</td>
</tr>
<tr>
<td></td>
<td>rainfall. (a) NEM : Northeast monsoon, (b) IM1 : Inter-monsoon 1,</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(c) SWM : Southwest monsoon, (d) IM2 : Inter-monsoon 2</td>
<td></td>
</tr>
<tr>
<td>3.8</td>
<td>Areas experiencing consistent TRMM underestimations (dashed polygons). The elevation map is derived from re-gridded Shuttle Radar Topography Mission (SRTM) data at 0.125 degree resolution</td>
<td>30</td>
</tr>
<tr>
<td>3.9</td>
<td>Average seasonal NSE of the TRMM 3B43 against ground rainfall in</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>local climate regions and average rainfall from areal rainfall surfaces for the period 1998–2010</td>
<td></td>
</tr>
</tbody>
</table>

...
4.1 Location of study site. (a) map of Southeast Asia (b) study area.
4.2 62 year-averaged monthly climate and rainfall of the study area (1949-2010). The record is taken from Kota Bharu station.
4.3 Methodology
4.4 TRMM pixels at the study site. Circles represent the rain gauges.
4.5 Scatter plot of bias ratio between PCA-against DA-based daily rainfall estimates.
4.6 Comparison between DA- and PCA-based rainfall estimates with different threshold. (a) bias ratio and (b) RMSE.
4.7 Relationship between the 3-hourly rain-rate and co-efficient of variation (COV).
5.1 Time-series between the ground areal rainfall, raw TRMM and HRC-downscale product from 2008-2010
5.2 Co-efficient of variation of the historical bias record from the satellite and ground areal rainfall
### TABLES

<table>
<thead>
<tr>
<th>No.</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>General information about the TRMM satellite.</td>
<td>7</td>
</tr>
<tr>
<td>3.1</td>
<td>Average annual scale statistical indicators for the direct transformed Tropical Rainfall Measuring Mission (TRMM) 3B43 data for average Peninsular Malaysia from 1998 to 2010</td>
<td>24</td>
</tr>
<tr>
<td>3.2</td>
<td>$T$-test results comparing the $0.250^\circ$ and $0.125^\circ$ validation results. (a) Correlation; (b) Root mean square error (RMSE); (c) Ratio and (d) Nash-Sutcliffe efficiency (NSE).</td>
<td>25</td>
</tr>
<tr>
<td>3.3</td>
<td>Cross-validation on the high resolution areal interpolated precipitation.</td>
<td>26</td>
</tr>
<tr>
<td>4.1</td>
<td>Comparison of ratio and root mean square error (RMSE) between two different daily rainfall estimates against interpolated areal rainfall. Negative sign on ratio indicates satellite underestimation condition.</td>
<td>49</td>
</tr>
<tr>
<td>4.2</td>
<td>Quantity of the rainy days that improved using the PCA-based approach.</td>
<td>49</td>
</tr>
<tr>
<td>4.3</td>
<td>Spatial correlation between TRMM satellite rainfall estimates from DA- and PCA-based approaches against areal rainfall from interpolated rain gauge measurement (RG).</td>
<td>49</td>
</tr>
<tr>
<td>5.1</td>
<td>Bias ratio comparison between raw TRMM data and HRC-downscale product.</td>
<td>64</td>
</tr>
<tr>
<td>5.2</td>
<td>Comparison between the HRC-downscale product and other satellite precipitation products</td>
<td>64</td>
</tr>
</tbody>
</table>
### APPENDIX

<table>
<thead>
<tr>
<th>No.</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Fig. A1-A3. Average monthly rainfall derived from TRMM 3B43 version 6 and 7 for Peninsular Malaysia from 1999 to 2001.</td>
<td>31</td>
</tr>
<tr>
<td></td>
<td>Table A1. Average monthly rainfall differences derived from TRMM 3B43 version 6 and 7 for Peninsular Malaysia from 1999 to 2001.</td>
<td>32</td>
</tr>
<tr>
<td></td>
<td>Table A2. T-test result of the TRMM 3B43 version 6 with version 7 for each corresponding monsoon season in every region.</td>
<td>33</td>
</tr>
<tr>
<td>2</td>
<td>Fig. A1. Hyetograph of the study area on November and December in 2003, 2005 and 2008.</td>
<td>55</td>
</tr>
</tbody>
</table>
ACKNOWLEDGMENT

I would like to express my utmost gratitude to everyone who contributed to this thesis either direct or indirectly. My foremost appreciation goes to the supervisors including Dr. Shinya Numata, Dr. Tetsuro Hosaka and Dr. Hiroshi Matsuyama of Graduate School of Urban Environmental Sciences, Tokyo Metropolitan University for their advice and supervision. Their enormous support in my research works has improved my capacity to become a sound researcher.

My special thanks to Prof. Dr. Mazlan Hashim and Geoscience & Digital Earth Centre, Technology University of Malaysia for his guide and supply of enormous hydrological and spatial data. My next appreciation is dedicated to the Dept. of Irrigation & Drainage, Malaysia, National Space Administration (NASA), and Japan Aerospace Exploration (JAXA) for the data access used in this study. I would like to thank Tokyo Metropolitan University and Tokyo Metropolitan Government for the award of Doctoral Fellowship of Asian Human Resources Program. Without their support, financially and socially, this thesis and its subsequent publications will not be materialized.

I want to thank my family, friends, colleague, acquaintances, and loved one for their support, love and companionship. Living in a foreign is a challenging and valuable experience. I will continue to learn, put the best hope and expectation, and living my life to fulfil my potential.

Thank you.
ABSTRACT

Rainfall played significant role in sustaining the nature ecotourism site in humid tropical region. The limitations of rain gauge measurement in remote, difficult access and risky areas require alternative support in rainfall measurement. Utilizing satellite precipitation data are useful option but faced with inherent constraints regarding its less defined accuracy and reliability for small-sized areas. This thesis conducted thorough investigation on one of the highest quality satellite precipitation data, namely Tropical Rainfall Measuring Mission (TRMM) and developing technique to improve their spatio-temporal accuracy. Using high resolution areal rainfall generated from dense rain gauge network over Peninsular Malaysia, TRMM indicated that their ability to represent temporal rainfall variation only good during the winter monsoon (Nov.-Apr.). The ability to depict local spatial rainfall variation was limited by the coarse grid size and differences of precipitation mechanism between satellite radar and rain gauge. Overcoming the large quantitative measurement errors that occurred during the heavy rainfall season (Nov.-Jan.) in sparse rain gauge areas can be alternatively conducted by estimating the rainfall using continuously correlated rain-rate images at daily scale. The coarse grid of TRMM can be refined from 0.25º to 0.06º using the monthly basis co-efficient derived from historical satellite-rain gauge error records. The TRMM satellite precipitation data can be useful in assisting the water related disaster management in the forest and related ecosystem and identifying suitable location and time to maximize the visitor’s satisfaction in water related ecotourism such as river, lake and streams. The output of this thesis contributed to the measures in maximizing the use of satellite precipitation data for ecohydrology applications in humid tropical region.
CHAPTER 1

GENERAL INTRODUCTION

Many potential ecotourism sites in humid tropical region are located in the tropical forest and related ecosystems (Menkhaus & Lober, 1996). This includes various types of tropical rainforest, mangrove, peat swamp, lake and rivers. Rainfall has been identified to be one of the crucial environmental variables that sustain those ecosystems. Rainfall play significant role as freshwater supplier for the plant to growth, medium of transportation for minerals from upper stream to downstream, and sole source of freshwater to lake and rivers (e.g, Pereira et al. 2014; Eklund et al. 1997). Excessive and shortage of rainfall might cause disturbances to tropical forest and its related ecosystems. During the heavy torrential rainfall, flood might occur. The flood has been recognized as a conspicuous threat on the tourism industry due to its impact on safety and physical effects on related ecosystems (Bernard & Cook, 2013). On the opposite case, a prolonged drought will cause rainfall shortage. The negative effect of drought on tropical forest was inevitable such as decreasing net productivity and water use (Virginia et al. 2001). In Southeast Asia, the seasonal Asian monsoon contributes significant role in characterizing the rainfall patterns (Dale, 1959; Fein & Stephens, 1988).

Prior to that, having rainfall information with sound spatial and temporal characteristic is required for monitoring and assessment. However, relying on conventional ground rain gauge observation to comply with that demand is difficult due to several inherent constraints. One of the limitations is the sparse or inadequate quantity of rain gauges. Often, this is caused by the natural condition of the thick, remote and difficult access to tropical rainforest and high operational cost. The next limitation is the inefficient rainfall management. The implication of this problem is the number of missing observation and difficulty in data retrieval. Using ground precipitation radar as a support is sensible. The radar operates by scanning the atmosphere and estimating the rain-rate based on the relationship between water
droplets and radar backscatter. They can provide near real-time information which suits the disaster mitigation and prevention, high accuracy estimation and has wide coverage (Joe, 1996). However, the drawbacks of the system are high infrastructure cost and data production, complicated data processing, and difficult data access, retrieval and archiving. As an alternative, the rain rate provided by the orbiting satellite precipitation radar that operated similarly with the ground radar can be a useful option.

The main advantages are that satellite precipitation data provide operational data supply, wide coverage, and open access data format. Enhanced spatial, temporal and measurement quality has increased the use of satellite precipitation data as source of rainfall information (Collischonn et al. 2008; Su et al. 2007; Adeyewa & Nakamura, 2003). One of the best and widely used products is the merged Tropical Rainfall Measuring Mission (TRMM) with other satellite data. They provided data at 0.25 degrees grid in every 3-hour temporal scale and covering 50 degrees latitude at both north and south.

However, despite the finest resolution, conspicuous limitation occurred in a small-sized region. This factor alone would cause insensitivity toward local rainfall pattern and measurement (Roongroj & Long, 2008; Mahmud, 2012; Behrangi et al. 2011). In addition, another contributor of the uncertainties is the precipitation estimation algorithm (Wolf et al. 2005). To anticipate this concern, a downscaling approaches in which the satellite precipitation data undergo post-processing to improve its quality (Yatagai et al. 2014; Tao & Barros, 2010; Cho & Choi, 2013). For a small-sized humid tropical region with significant seasonal influence on the rainfall variation, specific downscaling procedures that emphasized improvement on spatial and temporal aspect are strongly required. A sound downscaling algorithm is still rare in the region.
Despite several validations that have been conducted in humid tropical regions (Roongroj & Liu, 2008; Su et al. 2007; Collischonn et al. 2008; Semire et al. 2012), accuracy elements that need to be improved by small-sized regions are still not well justified. This is primarily due to the drawbacks of the validation method where coarse grid assessment or discrete rain gauges are employed and inadequacy of ground data is substantially used for evaluating the temporal accuracy. Identifying and understanding those accuracy limitations are important as reference for further downscaling activities. Prior to the concern, this study conducted a thorough examination of TRMM performance in a small-sized humid tropical region as a basis for the proper downscaling approach development. After that, a specific downscaling approach to improve the use of satellite precipitation data in small-sized humid tropical region is proposed.

**Organization of the thesis**

In this thesis, the performance of the satellite precipitation data over small sized humid tropical region of Peninsular Malaysia is evaluated and improved through a process namely known as downscaling.

In the second chapter, the seasonal accuracy of the satellite precipitation data over Peninsular Malaysia was intensively examined. The intensive evaluation involved the use of the higher resolution of areal precipitation data and assessment in every local climate region during seasonal monsoon scale. The ability of the satellite precipitation data in depicting the local-scale seasonal temporal variation, spatial variation and measuring the actual rainfall were revisited.

The third chapter focuses on tackling the high bias occurred during the heavy seasonal tropical thunderstorms without the use of rain gauge data. The direct accumulation of three hourly rain rate measurements into daily scale introduced temporal uncertainties. Because anticipating the uncertainties is a challenging task in many sparse and ungauged basins, an alternative approach to use significantly correlated images via principal component to recalculate the daily rainfall estimation
was experimented.

The next chapter emphasized on an effort to refine the coarse spatial grid of the satellite precipitation data. The coarse 0.25 deg. resolution data are downscaled into 0.06 deg. resolution. By taking into consideration that the seasonal precipitation experienced low variance of changes, a high resolution scale co-efficient was derived from the previous 10 years satellite-ground bias ratio record.

In the last chapter, I summarize the opportunities, challenge and limitation of downscaling the satellite precipitation data in small sized humid tropical environment based on the preceding chapters. The importance of having a reliable precipitation measurement in assessing the ecohydrological aspect in the potential ecotourism spot in a tropical environment is discussed.
CHAPTER 2

STUDY SITE DESCRIPTION & RAINFALL DATA

Study Site Description

Peninsular Malaysia (98.7–104.5 E, 1.1–7.5 N) is located in the western part of Malaysia (Fig. 2.1). It has a population of 18 million and covers approximately 132,000km². The general land cover is agriculture (52%), forest (22%), and built up areas (26%) (MACRES & UTM, 2008). The climate is humid tropical one with temperatures ranging from 24 to 32°C and rainfall throughout the year, with a total annual rainfall of 1200mm. The rainfall distribution pattern over Peninsular Malaysia is strongly influenced by regional wind flows (Wong et al. 2009); therefore, it is important to describe it based on the seasonal monsoon flows. A combination of the Asian Monsoon flows and local topographic patterns classified the area into five local climate regions; (i) northwest; (ii) east; (iii) west; (iv) southwest and (v) highland (>400m a.s.l.) (Figure 2.2). Two distinct wet seasons run from November to February, during which the northeast monsoon (NEM) produces heavy rainfall in the east region; and from May until mid-September, when the southwest monsoon (SWM) affects some areas in the west and southwest regions (Figure 2.1). The northwest, west, and southwest regions experience two wet seasons per year, from mid-March until May (IM1) and from October until November (IM2); both are in the inter-monsoon period (Dale, 1959). These two subsequent inter-monsoon periods take place during the shifting of the primary NEM and SWM seasons and vice versa. Due to the change of wind direction and effect of local topography, substantial rainfall occurred.
Rain Gauge Data

A total of 984 rain gauges, covering the entire Peninsular Malaysia, were collected from the Malaysia Department of Irrigation and Drainage and used in this study (Figure 2.2). Vast amounts of rain gauge data are collected to provide high-resolution areal rainfall information on the ground at both spatial and temporal scales. The rain gauge measurement was conducted on a daily basis with 24 hours observation from 8.00 a.m. until 8.00 a.m. of the next day (GMT). The daily rainfall measurement was then summed for one month period to produce the monthly basis rainfall measurement. This data and its corresponding geographical coordinates was then exported into geographic information system (GIS) shapefile format. The rain gauge records were acquired as early of 1961 until 2011.

Tropical Rainfall Measuring Mission (TRMM) Satellite Data

The TRMM Multi-Satellite Precipitation Analysis (TMPA) data product which involved rain rate from multiple satellites and other data products were selected for this study. TRMM satellite was orbiting the earth at an altitude of 402km carrying three primary sensor; namely Precipitation radar (PR), TRMM microwave imager (TMI), and visible and infrared scanner (VIRS). The PR sensor is designed to provide detailed vertical distribution of radar reflectivity, related to the number and especially the size of precipitation inside systems. The TMI sensor is purposed to measure the vertical integrated ice and water path. Meanwhile, the VIRS will provide information on cloud top temperature and reflectance. By using the fundamental concept of precipitation and radar reflectivity, the rain rate is estimated. General description of the data product is summarized in Table 2.1 and detail information on the data including algorithms and other parameters can be referred to the TRMM Instruction Manuals (TRMM Precipitation Radar Team et al. 2005 & 2011).
Precipitation data products from the TRMM satellite, was used for the following reasons: (i) frequent and consistent data collection (daily scale up to monthly scale); (ii) good spatial resolution (0.25°); and (iii) public domain data access. The high spatial and temporal resolution of this data satisfies the requirement to be used as the primary input for hydrological modeling and spatial analysis. The data were downloaded via internet from the official websites of National Aeronautics and Space Administration (NASA) with the collaboration of Japanese Aerospace Exploration Agency (JAXA). The rainfall data can be accessed on the following web address; http://daac.gsfc.nasa.gov/data/datapool/TRMM/01_Data_Products/02_Gridded/index.html. The TMPA data for Peninsular Malaysia was identified and extracted using the corresponding global coordinates systems for Peninsular Malaysia.

Table 2.1  General information about the TRMM satellite.

<table>
<thead>
<tr>
<th>Satellite</th>
<th>Tropical Rainfall Measuring Mission</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Provider</td>
<td>National Aeronautics and Space Administration (NASA)</td>
</tr>
<tr>
<td>Data Format</td>
<td>HDF</td>
</tr>
<tr>
<td>Data Coverage</td>
<td>50° North and 50° South from the equatorial belt</td>
</tr>
<tr>
<td></td>
<td>0° East to 180° East and 0° West to 180° West</td>
</tr>
<tr>
<td>Data Grid Size</td>
<td>0.25 degrees</td>
</tr>
<tr>
<td>Information</td>
<td>Latitude (degree), Longitude (degree), Precipitation (scaled millimeters)</td>
</tr>
</tbody>
</table>
Figure 2.1 Peninsular Malaysia and its seasonal rainfall variation
Figure 2.2 Rain gauge distribution in Peninsula Malaysia
CHAPTER 3

ASSESSMENT OF EFFECTIVE SEASONAL DOWNSCALING OF TRMM PRECIPITATION DATA IN PENINSULAR MALAYSIA

INTRODUCTION

Precise seasonal monsoon rainfall measurement is critical for accurate hydrologic prediction, simulation, and assessment in humid tropical regions. Using Tropical Rainfall Measuring Mission (TRMM) satellite precipitation data as an alternative to conventional rain gauge measurement is one useful option (Wilby & Yu, 2013; Paiva et al. 2013). Despite its promising potential for many regions worldwide (Krakauer et al. 2013; Islam & Uyeda, 2007; Li et al. 2012; Franchito et al. 2009; Almazroui et al. 2011), its sensitivity for local-scale rainfall in a small region, particularly one located in Southeast Asia, is contentious due to inherent uncertainties. These uncertainties include the effect of upscaling the instantaneous rain rate to an effective temporal scale (Roongroj & Long, 2008), the insensitivity of the TRMM precipitation algorithms to low- and high-precipitation clouds (Varikoden et al. 2010; Prasetia et al. 2013), and the coarse grid size of the TRMM data for resolving local rainfall patterns (Mahmud, 2012). Prompt actions must be taken to mitigate these uncertainties in order to obtain improved rainfall estimates from the TRMM data that suit local scale applications. In general, the untreated TRMM data showed an increasingly biased result when forced to scale hydrology modeling down to the local level as the region size decreased from large (Collischonn et al. 2008) to medium (Behrangi et al. 2011) to small (Mahmud & Hashim, 2011).

An effective countermeasure to mitigate these uncertainties can be achieved through downscaling, which can be defined as the specific process that improves the sensitivity of satellite precipitation data to local rainfall properties. In the case of Peninsular Malaysia, improving the coarse resolution of the TRMM data and minimizing the quantitative seasonal error from the rainfall estimates are essential for
obtaining precise seasonal rainfall information due to its small size and the high rainfall excess. Having detailed information on TRMM uncertainties at a finer scale can be useful for developing and applying suitable downscale procedures for specific local areas. Since ideal downscaling is designed for a specific environmental niche (e.g., Park, 2013; Tao & Barros, 2010), having an intensive reference would reduce the laborious processing work and increase its efficiency.

Unfortunately, previous validations or related TRMM studies in this region were unable to provide the necessary information for seasonal downscaling. These include the inability to depict the spatial error distribution and correlation due to the use of discrete rain-gauge comparisons (Varikoden et al. 2010; Mahmud, 2012) and coarse-grid resolution (Roongroj & Long, 2008). In addition, the reported uncertainties were measured in instantaneous scale and expected to increase as the instantaneous rain rate is upscaled (Omotosho et al. 2013). Moreover, because TRMM data products and validation scope varied among previous studies, the generalization of these findings for specific local-scale application is difficult and inappropriate. This is due to differences in the processing scheme, spatial grid size, and temporal scale of the rainfall measurement (hourly, daily, monthly) of the data products. Hence, the uncertainties in the TRMM products were different even within similar regions (Adeyewa & Nakamura, 2003; Nicholson et al. 2003).

A suitable method to determine the detailed uncertainties for this region is through meticulous validation conducted using a high-resolution precipitation grid within the local climate scale. This is because the rainfall intensity varies between monsoon seasons and within a local rainfall region (Wong et al. 2009). The effects of local environmental factors on rainfall, including topography, prevailing local winds, and maritime effects (Cheang, 1980; Varikoden et al. 2011), are significant. Furthermore, from the perspective of catchment hydrology, the size of the effective catchment for water resources in this region is relatively small (Vegas-Vilarrubia et al. 1994). The grid-based assessment was able to report a comprehensive spatial-based uncertainties distribution that is useful for local scale application (Collischonn et al. 2008, Nicholson et al. 2003). Nonetheless, there are few intensive
spatial-based seasonal uncertainties reports for this region. Therefore, as a preliminary step for effective downscaling, a thorough validation is needed. With the launch of the TRMM successor, the Global Precipitation Mission (Hou et al. 2014), there is a bright prospect for the active use of satellite precipitation and this downscaling gap should be accomplished.

This paper validated the re-gridded TRMM precipitation data at a local climate scale in a small humid tropical region. The monthly rainfall estimates derived from the TRMM 3B43 was evaluated at the seasonal monsoon scale in the local climate region of Peninsular Malaysia using high-resolution areal precipitation data(0.125°). The areal rainfall was derived from a dense rain gauge network (n =984). Four relevant performance elements were evaluated: (i) the ability to depict temporal rainfall variation (measured using the correlated ground data); (ii) the quantitative error between the TRMM and ground rainfall (measured using the root mean square error (RMSE)); (iii) the ability to estimate the actual rainfall amount (measured using the ratio between TRMM and actual rainfall), and (iv) the relative ability to reproduce ground rain gauge observations (measured using the Nash-Sutcliffe efficiency)(Nash & Sutcliffe, 1970).

MATERIALS AND METHODS

Data

TRMM satellite precipitation data

Precipitation data of monthly product from the TRMM satellite, namely TRMM 3B43 Version 6 was selected. A total of 156 monthly rain rate images were downloaded, covering the13 year period from 1998-2010. I compared the TRMM 3B43 version 6 and the latest version 7 and found no significant difference between them (Appendix 1).
Rain gauge data from 1998-2010 were acquired. Detail of the rain gauge properties is provided in the Chapter 2.

**Methodology**

*Direct re-gridding of the TRMM monthly rainfall estimation*

The TRMM data for the experimental site were extracted from the global dataset and projected into a local coordinates system. The monthly rain rate estimates from the TRMM 3B43 data were converted into monthly rainfall estimates by multiplying the hours and the number of days in each month. These are the standard processing procedures as stated in the standard TRMM Spatial Data Information and Management Systems (TSDIS) for raw rain rate data version. The basic equation of this estimation is given below (Equation (1)). After the monthly rainfall estimates are computed, the initial grid of TRMM 3B43 was rescaled to a finer grid of 0.125°. The new rainfall estimate values are obtained using the nearest neighbor interpolation scheme. The nearest neighbor interpolation scheme is selected due to two reasons; (i) minimization of the effects of smoothing which may lead to generalization of the rainfall patterns; and (ii) because the size of each pixel of the TRMM precipitation data is constant, nearest neighbor interpolation is able to retain the original rainfall values when rescaled.

\[
R_{sat} = R_s \times D_m \times H_t
\]  

(Equation (1))

where \( R_{sat} \) is the monthly TRMM rainfall estimate, \( R_s \) is the hourly average rain rate in a particular month, \( D_m \) is total days per month, and \( H_t \) is the total hours per day which is set as a constant of 24.

*Generation of gridded areal rainfall from rain gauge data*

Rainfall data from a total of 984 rain gauges were used to generate gridded areal precipitation surfaces at 0.125° resolution. That is a four-times areal
increment of the initial TRMM areal grid (0.25°). This rain gauge data was interpolated using the co-Kriging interpolation scheme, which is commonly used for creating areal rainfall (Mair& Fares, 2011). Cross-validation is conducted to determine the reliability of the interpolated areal rainfall surfaces. Two independent datasets were established, one used as a reference, the other used as a tester. The root mean square error (RMSE) and its mean percentage error (MAPE) are computed and then used as cross-validation metrics.

**Validation on TRMM rainfall against gridded areal rainfall**

Three statistical indicators were used in this study, (i) Pearson correlation (r); (ii) average ratio between TRMM estimates and rain gauge measurements; and (iii) root mean square error (RMSE). Linear regression was used to examine the strength of the linear relationship between TRMM rainfall estimates and rain gauge measurements. High correlation values indicated a high ability of TRMM to describe seasonal rainfall variation. The ratio between the TRMM precipitation estimates and the corresponding rain gauge measurement provides information on the capability of TRMM to describe the actual rainfall on the ground. Ratio values > 1 indicated an overestimation in the TRMM data, while those < 1 indicated an underestimation. A value of 1 indicated that TRMM estimated the actual rainfall perfectly. The RMSE was used to quantify the TRMM and rain gauge data in standard SI units (mm). In order to provide standardized TRMM-rain gauge validation across areas and seasons with different base rainfall rates, Nash-Sutcliff efficiency (NSE) was used. NSE value of 1 indicates TRMM agrees with rain gauge observations, while value of 0 means that TRMM is only as good as always predicting the mean observed precipitation value (Krakauer et al. 2013). Equations (2) to (5)are the standard equations used in this assessment for all indicators.

\[
r = \frac{n(\sum R_{sat})-(\sum R_{sat})(\sum R_{rg})}{\sqrt{\left[ n\sum R_{sat}^2-(\sum R_{sat})^2 \right] \left[ \sum R_{rg}^2-(\sum R_{rg})^2 \right]}}
\]

(2)

\[
t = \frac{\sum R_{sat}}{n \sum R_{rg}}
\]

(3)

\[
\]
\[ s_e = \sqrt{\frac{\sum (R_{sat} - R_{rg})^2}{n}} \]  

\[ NSE = 1 - \frac{\left( \frac{\langle R_{rg} - R_{sat} \rangle^2}{\left( \langle R_{rg} - R_{rg} \rangle \right)^2} \right)}{\left( \frac{\langle R_{rg} - R_{rg} \rangle^2}{\left( \langle R_{rg} - R_{rg} \rangle \right)^2} \right)} \]

where \( r \) is the Pearson’s correlation co-efficient, \( t \) is the ratio between TRMM estimates over rain gauge measurements, \( Se \) is the RMSE, \( R_{sat} \) is the monthly TRMM rainfall estimate, \( R_{rg} \) is the monthly rain gauge data, and \( n \) is the number of samples, \(<.> \) represent the averages.

All four statistical indicators were computed for every corresponding pixel of both rainfall datasets from 1998 to 2010. The spatial assessment was conducted at all five local climate regions: northwest, east, west, southwest, and high elevated terrain during four major monsoon seasons including NEM, IM1, SWM and IM2. Auxiliary information regarding seasonal average rainfall variation and the co-efficient of variation (COV) from all rain gauge data were calculated for further analysis of the TRMM correlation, the ratio, and RMSE. These two rainfall characteristics were selected due to inherent characteristics during each monsoon season.

**TRMM systematic bias and subgrid variation determination**

The validation is designed to determine the seasonal systematic error of TRMM at the fine grid of 0.125° resolution. Subsequently, the spatial downscale process may introduce subgrid variation effects. To anticipate this matter, a comparison between the validation results at initial resolution (0.25°) and downscale resolution (0.125°) is carried out. \( T \)-tests are employed to verify whether the results significantly vary at 95% confidence level. Significant differences will indicate significant effects of subgrid variation on the validation and \textit{vice versa}. 
RESULTS

TRMM seasonal correlation, ratio, and RMSE for peninsular scale

For the Peninsular Malaysia scale, the TRMM showed good spatial correlation with rain gauge data only during the NEM and IM1 (>0.65) (Table 3.1). The measurement errors were higher during the NEM and IM2. Overestimates of 70%–90% of actual rainfall measurements were seen in all seasons. Analysis of variance (ANOVA) was conducted with the null hypotheses that there were no significant differences in the statistical indicators within different monsoon seasons. The ANOVA result indicated that the seasonal correlation and RMSE were significantly different among monsoon seasons ($F_{0.05; 3,3312} = 846, F_{0.05; 3,3312} = 427$).

TRMM subgrid variation analysis

The comparison between the 0.25° and 0.125° grids (Table 3.2) have shown that there was not enough evidence to indicate a significant difference between 0.25° and 0.125° analyses (Table 3.2). However, there were inherent changes in the ratios (Table 3.2). The ratio for the 0.125° validation was increased during NEM (7%) and IM1 (11%) (Figure 3.3). During IM1, the ratio increased in all regions except the southwest. By contrast, during the NEM, only the northwest region was affected. Those affected areas experienced the lowest rainfall amount during that season. Based on the cumulative findings of the comparison, it clarified that direct re-gridding of the TRMM data had an imminent effect on the ability of the TRMM to depict spatial rainfall patterns during dry season but had no effect on the overall efficiency and other indicators. Therefore, the 0.125° validation presented later in this paper is assumed to be the TRMM systematic errors and other related sources (e.g., the post-rainfall estimation algorithm) except for the specific occasions in the spatial rainfall representation (ratio).
TRMM seasonal correlation in the local climate region

All the climate regions exhibited a similar trend in the seasonal correlation and significantly varied within seasons (Figure 3.4). The correlation tended to be high during NEM and IM1 but relatively low during the SWM and IM2; the correlation was highest during NEM (0.60–0.79) followed by IM1 (0.55–0.70), SWM (0.36–0.50) and IM2 (0.30–0.50). This indicates that the ability of TRMM to depict the seasonal rainfall variation during SWM and IM2 was relatively poor compared to that during the other two monsoon seasons. The correlation showed greater variation among regions during SWM and IM2 compared to NEM and IM1 (Figure 3.5a–d). The correlation during SWM was relatively good in the west and in parts of the southwest, but it was very low for a major part of the east coast (Figure 5c). During IM2, the upper region of Peninsular Malaysia, including the highlands, the upper part of the east coast, and the majority of the northwest had low correlations. A plausible reason for this seasonal influence on the TRMM correlation was caused by the seasonal co-efficient of variation (COV) trend of the ground rainfall (Figure 3.4). A higher COV was identified during NEM and IM1 than during SWM and IM2, and the seasonal difference was significant (F$_{0.05}$; 3, 3312 = 537).

TRMM seasonal quantitative error in local climate regions

The RMSE was higher during the wet seasons (NEM and IM2) (Figure 6) compared to the SWM and IM1 for all climate regions, except the northwest region during the NEM. The RMSE during the NEM and IM2 contributed the largest proportion of the annual error, approximately 56%–61% for all regions. During the wettest part of the NEM, the range of RMSE was 120–162mm, and it was mostly distributed in the east where rainfall was high (Figure 3.6). Consequently, temporal maps also showed that a high RMSE during the NEM (Figure 3.7a) occurred in the east including all the highlands in that region. During IM2, high RMSE values were concentrated mainly in the upper part of the Peninsula, which includes the northwest, upper part of the west, east, and northern highland regions (Figure 3.7d). No spatial distribution of RMSE during SWM and IM1 among regions was discernible (Figure
3.7b,c) with values of 80–100mm/month. Significant seasonal rainfall differences were identified from the ground rainfall (F0.05; 3, 3312 = 428). This indicates that occurrences of the large seasonal RMSE are associated with high rainfall during the NEM and IM (Figure 3.6). The cross-validation of the interpolated areal precipitation gridded product revealed that it increased linearly with the increment of seasonal rainfall (Table 3.3). The mean average error percentage ranged from 8% to 13% (13–22mm). This showed that the systematic error from the interpolation contributed a small portion of the uncertainty on the validation outcomes.

TRMM seasonal rainfall amount estimation in local climate regions

Positive average ratio indicated that the TRMM overestimated the actual rainfall in all climate regions during all monsoon seasons (Figure 3.8). The variation of ratios varied among the regions, particularly between northwest, southwest and west. The average bias of the TRMM data can be ranked from low to high as follows: southwest (30%–50%), highland (60%–70%), east (50%–100%), west (70%–100%) and northwest (120%–160%). In fact, the northwest area had a comparatively larger bias (>200%) than the other regions during all respective monsoon seasons. A distinct overestimate was indicated during IM1 in the east region. Temporal maps depicted that TRMM has largely overestimated the ground rainfall in the northwest region every monsoon season (Figure 3.9). In addition, there were areas that experienced consistent satellite underestimation (green colors)(Figure 6a–d). Further investigation into those areas using digital elevation models revealed that it was the underlying terrain from areas with high elevation (Figure 3.10). This signifies that the ability of TRMM to represent the actual rainfall pattern was region-specific rather than seasonal.
Figure 3.11 shows that the efficiency of TRMM varied within monsoon seasons. The average efficiency of TRMM was higher during the NEM (0.302) and IM1 (0.386) than during the SWM (−0.188) and IM2 (−0.191). During the wettest and driest seasons of the NEM and IM1, respectively, TRMM proved to be more efficient. The trend towards a stronger seasonal than regional influence indicated that the seasonal co-efficient of variation (COV) of ground rainfall has largely influenced the TRMM efficiency. The efficiency differences that occurred within regions during the NEM and IM1 were larger compared to SWM and IM2. This suggests that the spatial rainfall COV during the wettest and driest seasons is the driving factor that determines the efficiency of TRMM.

DISCUSSION

The seasonal scale validation of the TRMM precipitation data using a high-resolution areal precipitation grid for Peninsular Malaysia highlights the important characteristics of a small humid tropical basin in the equatorial region. In terms of temporal rainfall sensitivity, there were two major issues that needed to be investigated. The first was the weak correlation during the SWM and IM2, and the second was the large measurement error during the wet seasons of the NEM and IM2. From the perspective of an accurate representation of the seasonal spatial rainfall pattern, the TRMM sensitivity for each climate region must be improved, particularly in the northwest and west. These local scale seasonal uncertainties in Peninsular Malaysia were unique compared to the findings in a larger basin (Clarke et al. 2011) and were unidentified by previous studies (Varikoden et al. 2010). Downscaling the grid to a fine scale (0.125°) did not introduce drastic subgrid variation that affected the overall result, apart from the minimal impact on spatial variability during dry seasons. Since TRMM 3B43 monthly rainfall data product is not available in near real time (Liu et al. 2014) and the hydrological application had
to use a less calibrated data product, our findings provide useful information as a reference and reflect its limitation at a local basin scale.

The seasonal correlated behavior that could have a relationship with a coefficient of variation can be explained by the presence of low and heavy precipitation clouds. A similar condition also occurred in the previous daily basis validation (Varikoden et al. 2010), and the effect found in this study could be the result of the longer temporal scale (hourly and daily to monthly). Because TRMM has better sensitivity towards heavy rather than low precipitation clouds, a good correlation was obtained during NEM. The majority of the heaviest rainfall during NEM was primarily caused by the large-scale monsoon flows that contained heavy precipitation cloud (Cheang, 1980). However, during SWM and IM2, in which both heavy and low precipitation clouds occurred, the correlation decreased due to insensitivity towards a low precipitation cloud.

A large measurement error (RMSE) of TRMM during the wet seasons has frequently been identified in humid tropical regions (Franchito et al., 2009; Varikoden et al. 2010; Prasetya et al., 2013; Collischonn et al. 2008; Su et al. 2007). The major constituent was the effect of scaling up the instantaneous rain rate from an hourly to a monthly basis using the given scale factor. Other contributing factors were the temporal differences between TRMM and rain gauge measurements and the systematic error of interpolation. However, the effect of temporal differences was considered minimal, because the time mismatch was only 1h (~8.00–9.00am local time). In addition, the interpolation effect also showed a linear RMSE increment towards increasing rainfall in a relatively small proportion. The interesting finding is that the RMSE in this study was relatively lower than in previous studies conducted in Southeast Asia using a similar data product. Therefore, it is clarified that the RMSE in this region strongly depending on seasonal rainfall intensity, method to calculate the cumulative monthly rainfall and grid size used in the assessment.
The TRMM spatial specific-region behavior might be influenced by two factors; the coarse initial grid size of the TRMM (0.25°) and the mechanism of precipitation measurement. The coarse initial TRMM grid size made it less sensitive to local scale rainfall patterns and distribution. Apart from the systematic error of the TRMM, the subgrid variation showed at 0.125° that the TRMM data were unable to resolve the small-scale of the convective cloud distribution and its shorter temporal variation during the dry seasons. In addition, the insensitivity of the TRMM sensor algorithm towards those cloud characteristics can lead to miscalculation of the rain rate (Wolf et al. 2005). Another plausible factor is the difference between precipitation measurement mechanisms of the satellite and rain gauge. In principle, satellite precipitation radar is operated to estimate the rain rate in the atmosphere rather than by quantifying the actual rainfall on the ground (Iguchi et al. 2000). Consequently, the TRMM has limited sensitivity for areas with a complex spatial rainfall pattern due to environmental parameters including the orography effect, monsoon flow, and distance to the sea, such as in the east region.

The comparative efficiency of the TRMM vs. the ground areal rainfall showed that the seasonal effects were stronger than the regional effects. Prior to this, it is suggested that TRMM sensitivity towards high and low precipitation clouds that influenced the seasonal correlation was more influential at determining the efficiency of the TRMM than the coarse grid size and difference in precipitation measurement mechanisms between satellite and rain gauges. Improving the seasonal scale performance during the SWM and IM2 is critical to obtain high precision annual rainfall for our study site. This requires careful consideration and treatment of clouds with medium precipitation, which were common during this period.

Downscaling should be attempted to resolve the uncertainties found in this study. The large error propagated during the wet seasons of IM2 and NEM will lead to inaccurate rainfall estimates that will affect quantitative rainfall estimates. In addition, the low seasonal correlation during the SWM and IM1 will lead to inaccurate temporal monsoon variation prediction from TRMM data. Temporal downscaling is recommended to improve both correlation and measurement
differences. For the context of the Asian region, a procedure introduced by Yatagai et al. (2014) and Ryo et al. (2014) is scientifically sound. In terms of spatial rainfall variability, because the ability of TRMM to depict accurate spatial rainfall patterns varied within the local climate region, the use of spatial downscaling is recommended especially in the northwest and west regions.

However, there is a lack of spatial downscaling approaches adaptable to the humid tropics. Such a development would be the advances required in satellite precipitation data downscaling. The use of an alternative satellite data product, such as CMORPH or GSMaP, which are effective in high elevations (Dinku et al. 2007) as a substitute for TRMM or their integration, is a possible measure for deriving high precision precipitation. In short, downscaling satellite precipitation data to support local-scale hydrological applications is expected to be an active endeavor prior to the launch of the Global Precipitation Mission, the TRMM successor. Although this study used TRMM ver. 6, our findings provide useful, relevant information that can be applied to improve the subsequent versions of TRMM and its successor. The comparison between TRMM Version 7 datasets also showed no significant differences for this region (See Appendix 1).
CONCLUSIONS

This study validated the re-gridded TRMM 3B43 (0.125°) monthly rainfall estimates at seasonal monsoon scale at the local climate region of Peninsular Malaysia (1998–2010) using high resolution areal precipitation surfaces. The TRMM correlations varied seasonally and were high during the wettest period of the NEM (0.63–0.8) and its transitional period of IM1 (0.55–0.65) for all regions and decreased gradually over the remainder of the monsoon period. The ability of TRMM to estimate actual rainfall varied among local climatic regions and conditions were often overestimated; the extent of overestimation was 30%–50% in the southwest, 60%–70% in the highland region, 50%–100% in the east, 70%–100% in the west and 120%–160% in the northwest. The coarse grid size of the initial TRMM data has led to the insensitivity to local rainfall pattern that is strongly influenced by environmental factors and convective clouds. The quantitative errors (RMSE) were significantly related to the wet seasons in NEM and IM2. A large RMSE was identified in the east (120mm-NEM; 165mm-IM2), the highland region (130mm-NEM-140mm-IM2), and the northwest region (190mm-IM2) during both monsoon seasons. This study provides detailed spatial uncertainty patterns and variations of the direct spatial downscaling of TRMM which served as a critical reference for further downscaling activities of the satellite precipitation data.
Table 3.1 Average annual scale statistical indicators for the direct transformed Tropical Rainfall Measuring Mission (TRMM) 3B43 data for average Peninsular Malaysia from 1998 to 2010.

<table>
<thead>
<tr>
<th>Season</th>
<th>Correlation (r)</th>
<th>RMSE(mm)</th>
<th>Ratio</th>
<th>NSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>NEM</td>
<td>0.745</td>
<td>142</td>
<td>1.8</td>
<td>0.302</td>
</tr>
<tr>
<td>IM1</td>
<td>0.651</td>
<td>88</td>
<td>1.9</td>
<td>0.386</td>
</tr>
<tr>
<td>SWM</td>
<td>0.397</td>
<td>91</td>
<td>1.7</td>
<td>−0.188</td>
</tr>
<tr>
<td>IM2</td>
<td>0.406</td>
<td>123</td>
<td>1.7</td>
<td>−0.191</td>
</tr>
</tbody>
</table>
Table 3.2. *T*-test results comparing the $0.250^\circ$ and $0.125^\circ$ validation results. (a) Correlation; (b) Root mean square error (RMSE); (c) Ratio and (d) Nash-Sutcliffe efficiency (NSE).

(a)

<table>
<thead>
<tr>
<th>Monsoon Season</th>
<th>$0.250^\circ$ Mean (SD)</th>
<th>$0.125^\circ$ Mean (SD)</th>
<th><em>t</em>-test</th>
</tr>
</thead>
<tbody>
<tr>
<td>NEM</td>
<td>0.738 (0.118)</td>
<td>0.745 (0.122)</td>
<td><em>ns</em></td>
</tr>
<tr>
<td>IM1</td>
<td>0.654 (0.160)</td>
<td>0.651 (0.160)</td>
<td><em>ns</em></td>
</tr>
<tr>
<td>SWM</td>
<td>0.402 (0.177)</td>
<td>0.396 (0.173)</td>
<td><em>ns</em></td>
</tr>
<tr>
<td>IM2</td>
<td>0.408 (0.186)</td>
<td>0.406 (0.194)</td>
<td><em>ns</em></td>
</tr>
</tbody>
</table>

(b)

<table>
<thead>
<tr>
<th>Monsoon Season</th>
<th>$0.250^\circ$ Mean (SD)</th>
<th>$0.125^\circ$ Mean (SD)</th>
<th><em>t</em>-test</th>
</tr>
</thead>
<tbody>
<tr>
<td>NEM</td>
<td>141 (46)</td>
<td>142 (47)</td>
<td><em>ns</em></td>
</tr>
<tr>
<td>IM1</td>
<td>89 (27)</td>
<td>88 (26)</td>
<td><em>ns</em></td>
</tr>
<tr>
<td>SWM</td>
<td>91 (24)</td>
<td>91 (23)</td>
<td><em>ns</em></td>
</tr>
<tr>
<td>IM2</td>
<td>118 (40)</td>
<td>123 (43)</td>
<td><em>ns</em></td>
</tr>
</tbody>
</table>

(c)

<table>
<thead>
<tr>
<th>Monsoon Season</th>
<th>$0.250^\circ$ Mean (SD)</th>
<th>$0.125^\circ$ Mean (SD)</th>
<th><em>t</em>-test</th>
</tr>
</thead>
<tbody>
<tr>
<td>NEM</td>
<td>1.71 (1.59)</td>
<td>1.78 (1.00)</td>
<td><em>ns</em></td>
</tr>
<tr>
<td>IM1</td>
<td>1.67 (1.35)</td>
<td>1.89 (1.50)</td>
<td><em>ns</em></td>
</tr>
<tr>
<td>SWM</td>
<td>1.68 (0.77)</td>
<td>1.68 (0.67)</td>
<td><em>ns</em></td>
</tr>
<tr>
<td>IM2</td>
<td>1.64 (0.97)</td>
<td>1.66 (1.25)</td>
<td><em>ns</em></td>
</tr>
</tbody>
</table>

(d)

<table>
<thead>
<tr>
<th>Monsoon Season</th>
<th>$0.250^\circ$ Mean (SD)</th>
<th>$0.125^\circ$ Mean (SD)</th>
<th><em>t</em>-test</th>
</tr>
</thead>
<tbody>
<tr>
<td>NEM</td>
<td>0.313 (0.213)</td>
<td>0.302 (0.220)</td>
<td><em>ns</em></td>
</tr>
<tr>
<td>IM1</td>
<td>0.402 (0.161)</td>
<td>0.386 (0.113)</td>
<td><em>ns</em></td>
</tr>
<tr>
<td>SWM</td>
<td>-0.184 (0.429)</td>
<td>-0.188 (0.433)</td>
<td><em>ns</em></td>
</tr>
<tr>
<td>IM2</td>
<td>-0.171 (0.510)</td>
<td>-0.191 (0.522)</td>
<td><em>ns</em></td>
</tr>
</tbody>
</table>

Note. *M* = Mean; *SD* = Standard deviation; *ns* = not significant; $p<0.05$. 
Table 3.3. Cross-validation on the high resolution areal interpolated precipitation.

<table>
<thead>
<tr>
<th>Monsoon Season</th>
<th>Cross-Validation Metrics</th>
<th>Average Ground Rainfall (mm)</th>
<th>Percentage upon the TRMM-RG RMSE(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RMSE (mm)</td>
<td>MAPE (%)</td>
<td></td>
</tr>
<tr>
<td>NEM</td>
<td>22</td>
<td>13</td>
<td>198</td>
</tr>
<tr>
<td>IM1</td>
<td>15</td>
<td>10</td>
<td>112</td>
</tr>
<tr>
<td>SWM</td>
<td>13</td>
<td>8</td>
<td>118</td>
</tr>
<tr>
<td>IM2</td>
<td>20</td>
<td>11</td>
<td>171</td>
</tr>
</tbody>
</table>

Figure 3.1. Ratio between 0.25° and 0.125° TRMM-rain gauge validation. (a) Northeast monsoon (NEM); (b) Inter-monsoon 1 (IM1).
Figure 3.2. Seasonal correlation of the TRMM 3B43 against ground rainfall at local climate region of Peninsular Malaysia. The line represents the co-efficient of variation (COV) of areal rainfall surfaces.

Figure 3.3. Temporal maps of seasonal correlation of the TRMM 3B43 against ground rainfall. (a) NEM : Northeast monsoon, (b) IM1 : Inter-monsoon 1, (c) SWM : Southwest monsoon, (d) IM2 : Inter-monsoon 2.
Figure 3.4. Average seasonal RMSE of the TRMM 3B43 against ground rainfall in local climate regions and average rainfall from areal rainfall surfaces for the period 1998–2010.

Figure 3.5. Temporal maps of seasonal RMSE of the TRMM 3B43 against ground rainfall. (a) NEM : Northeast monsoon, (b) IM1 : Inter-monsoon 1, (c) SWM : Southwest monsoon, (d) IM2 : Inter-monsoon 2.
Figure 3.6. Average seasonal ratio values between TRMM rainfall estimates and ground rainfall in local climate regions.

Figure 3.7. Temporal maps of seasonal ratios of the TRMM 3B43 over ground rainfall. (a) NEM : Northeast monsoon, (b) IM1 : Inter-monsoon 1, (c) SWM : Southwest monsoon, (d) IM2 : Inter-monsoon 2.
Figure 3.8. Areas experiencing consistent TRMM underestimations (dashed polygons). The elevation map is derived from re-gridded Shuttle Radar Topography Mission (SRTM) data at 0.125 degree resolution.

Figure 3.9. Average seasonal NSE of the TRMM 3B43 against ground rainfall in local climate regions and average rainfall from areal rainfall surfaces for the period 1998–2010.
APPENDIX 1


(Year : 1999)

<table>
<thead>
<tr>
<th>Rainfall (mm)</th>
<th>TRMM Version 6</th>
<th>TRMM Version 7</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fig. A1
Fig. A2

(Year: 2000)

**Rainfall (mm)**

- TRMM Version 6
- TRMM Version 7

---

**Fig. A2**

---

32
(Year: 2001)

**Rainfall (mm)**

- TRMM Version 6
- TRMM Version 7

---

**Monsoon seasons**

Fig. A3
Table A1. Average monthly rainfall differences derived from TRMM 3B43 version 6 and 7 for Peninsular Malaysia from 1999 to 2001.

<table>
<thead>
<tr>
<th>Monsoon seasons</th>
<th>Monthly difference (%)</th>
<th>1999</th>
<th>2000</th>
<th>2001</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>NEM</td>
<td>10</td>
<td>8</td>
<td>6</td>
<td></td>
<td>8</td>
</tr>
<tr>
<td>IM1</td>
<td>14</td>
<td>12</td>
<td>11</td>
<td></td>
<td>12</td>
</tr>
<tr>
<td>SWM</td>
<td>7</td>
<td>7</td>
<td>10</td>
<td></td>
<td>8</td>
</tr>
<tr>
<td>IM2</td>
<td>5</td>
<td>6</td>
<td>11</td>
<td></td>
<td>7</td>
</tr>
</tbody>
</table>

Table A2. T-test result of the TRMM 3B43 version 6 with version 7 for each corresponding monsoon season in every region.

<table>
<thead>
<tr>
<th>Monsoon season/ Years</th>
<th>1999</th>
<th>2000</th>
<th>2001</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>E</td>
<td>W</td>
</tr>
<tr>
<td>NEM</td>
<td>ns</td>
<td>ns</td>
<td>ns</td>
</tr>
<tr>
<td>IM1</td>
<td>ns</td>
<td>ns</td>
<td>ns</td>
</tr>
<tr>
<td>SWM</td>
<td>ns</td>
<td>ns</td>
<td>ns</td>
</tr>
<tr>
<td>IM2</td>
<td>ns</td>
<td>ns</td>
<td>ns</td>
</tr>
</tbody>
</table>

* ns: no significant difference; T-test was conducted under 95% confidence level.
CHAPTER 4

TEMPORAL DOWNSCALING OF TRMM RAIN-RATE IMAGES USING PRINCIPAL COMPONENT ANALYSIS DURING HEAVY TROPICAL THUNDERSTORMS SEASON

INTRODUCTION

Heavy thunderstorms are significant phenomena that drive hydrological dynamics in humid tropical catchments. Accurate rainfall measurements during these events are critical for water resources applications, disaster mitigation strategies, and hydrological assessments (Thompson, 1999). Using the Tropical Rainfall Measuring Mission (TRMM) satellite precipitation 3-hourly rain rate data as an alternative support to rain gauge measurements is practical because of their wide spatial coverage, grid-based information, and high temporal resolution. Unfortunately, a series of validation and application studies in respective regions indicated that TRMM data produce large uncertainties during heavy rainfall events. One of the conspicuous uncertainties sources of uncertainty is the large measurement bias, particularly rainfall overestimation (Mahmud et al. 2015; Prasetia et al. 2014; Su et al. 2007; Collischonn et al. 2008; Varikoden et al. 2010; Roongroj & Long, 2008).

A plausible reason of the bias is the use of an inappropriate technique for estimating daily rainfall from 3-hourly rain-rate satellite images. A common technique used for this is to sum every scaled up 3-hourly rainfall measurement for an entire 24 hour period. This direct accumulation (DA) approach has two temporal deficiencies. The first is the effect of scaling up the 3-hourly average instantaneous rain-rates to 3-hourly rainfall. In the DA approach the rain rate within is assumed to be constant over a 3-hour period. That claim however is contentious as the literature showed that rainfall in humid tropics can vary diurnally during the peak monsoon period (Oki & Musiakke, 1994; Wang & Lin-Ho, 2001). The second deficiency is the use of different mechanism for measuring the precipitation. The general principle of precipitation radar
is estimating the rain-rate in the atmosphere based on the backscatter signal; meanwhile a rain gauge is measuring rainfall on the ground (Iguchi et al. 2000; Berne & Krajewski, 2013). The rainfall pattern on the ground varies dynamically due to significant influence from local environmental factors, namely prevalent wind and topography (Cheang, 1980; Varikoden et al. 2011). Hence, the rain-rate information obtained by the rain gauge and by the satellite sensor can be mismatched. As a consequence, direct summation of the scaled up 3-hourly rainfall into daily rainfall can lead to large miscalculation. Many existing temporal downscaling approaches that minimize temporal error requires support from rain gauge data (e.g., Yatagai et al. 2014; Ryo et al. 2014, Tao & Barros, 2010). Overcoming that limitation is difficult especially when sources of ground reference data, particularly rain gauges, are sparse, inadequate or unavailable. Efforts to anticipate with this conflict are rare, and subjected to be further investigation.

Tackling temporal deficiencies without the presence of rain gauges is a huge challenge. A better and more effective technique that can mitigate temporal deficiencies produced using the DA approach is encouraged. Plausible methods would include correlating coherent rain-rate information or appropriately recalculating the total daily rain-rate. Both measures can be implemented by using a principal components analysis (PCA). The significant information from multi-dimensional remotely-sensed images are compressed into fewer components, and a new pixel value is calculated based on their corresponding linear relationship with the effective component (Weng, 2011; Mather & Koch, 2010). By adapting this simple concept to our conflict, a similar spatial pattern of rain-rate from the series of 3-hourly TRMM rain-rate images can be accumulated into significant components, and the new adjusted rain-rate can be recomputed based on the correlation towards the respective component.

Pioneer effort on modification of the PCA original concept of projecting significant information from multiple digital images into minimum effective components for assisting information extraction and reducing processing works can be found in Jensen (1996). Consequently, PCA has proven an effective tool in reducing noise in data and in compressing related spectral information from multiple remote sensors for various
purposes including the atmospheric and related sciences (Demsar et al. 2013). The PCA approach is commonly applied for qualitative analyses such as classification or features identification, where significant information on specific features was compressed from multiple spectral bands (e.g., Rogge et al. 2014). The notion of applying PCA in the estimation of daily rainfall estimates using TRMM rain-rate imageries is scientifically sound but the potential is yet to be examined.

I investigated the utility of using PCA transformation to produce higher accuracy of daily rainfall estimates from TRMM images during heavy tropical thunderstorms. Our hypothesis in this study was that PCA is able to compress the significant temporal rain-rate data from eight continuous 3-hourly rain-rate datasets and recompute the rain-rate estimates based on the correlation value with the significant axis. To validate the effectiveness of PCA, the daily rainfall estimates from PCA transformation was compared with those from the common procedure of direct summation of continuous 3-hourly rain-rate data using actual rainfall measurement. The experiment was carried out during heavy thunderstorm seasons in three different years in a small-sized basin at the humid tropical region of peninsular Malaysia. The result of this study provides useful solutions for high-accuracy estimation of daily rainfall using TRMM rain-rate images in areas with sparse or insufficient rain gauge data during heavy tropical thunderstorms.

**STUDY SITE DESCRIPTION**

Kelantan basin is located at the upper east coast of Peninsular Malaysia, a west part of Malaysia (Fig. 4.1). The climate is humid tropical one with an average temperature between 25 to 28°C throughout the year with annual rainfall of 2709mm. Figure 4.2 showed the climate record for this basin for 60 year period. It covered almost 965 km square in acreage and experienced heavy rainfall in Asian monsoon season especially from Northeast Monsoon flows which primarily occurred from October to January. It carried moist air, strong prevalent wind and heavy thunderstorm (Cheang, 1980). This period is considered as the wettest period throughout the year. Highest monthly rainfall in this season is 1058mm and it is recorded in Kota Bharu station in November 2008 (Fig. 4.1b). Rainfall during Northeast Monsoon contributes
approximately 50% of the total annual rainfall. In fact, water-related disasters especially flood and drought are common in this monsoon season (e.g., DID, 2005, 2009). The land cover proportion consists of tropical forest (51%), plantation and agriculture (39%) and others (urban, peat swamp, water bodies, etc) (10%)(MACRES & UTM, 2008). Consequently, precise determination of rainfall data is crucial to support socio-economic activities, ecological needs, and sustains human lives.

MATERIALS & METHOD

There are three main phases in this study. The first phase is the preparation of the TRMM rain-rate data. The second phase is the estimation of daily rainfall using PCA approach and the third phase involved the comparison between the PCA and direct accumulation approach. Flow of the methodology is showed in figure 4.3.

Phase 1- TRMM satellite rain rate data sources & preparation

The TRMM 3B42 Version 7 data product provides 3-hourly rain-rate data over the global scale (50°N ranges to 50°S) at the finest resolution of 0.25 degrees. TRMM satellite is a joint venture project of JAXA (Japan Aerospace and Exploration), NASDA (National Space Development Agency - Japan) and NASA (National Aeronautics and Space Administration - U.S.A.) which provides primary rain-rate products from their three major sensors TMI (TRMM Microwave Imager), PR (Precipitation Radar) and VIS (Visible Infrared Scanner) over the cloud. Detailed information about TRMM rainfall algorithm and concept can be found in the TRMM Version 7 manual by TRMM Precipitation Radar Team et al. (2011) and Huffman et al. (2007). TRMM data were obtained from the public domain server at http://daac.gsfc.nasa.gov/data/datapool/TRMM/01_Data_Products/02_Gridded/index.html. Once the data is obtained, specific area of interest is extracted. Geographical coordinates were used to identify the study area. The next step of the pre-processing is transferring the rain-rate images which were in global geographical coordinates to the local coordinate system. The rain-rate images, initially in the format of WGS 84 (World Geodetic System) coordinate system, were transformed using transformation co-efficient
to MRSO (Malaya Rectified Skew Orthomorphic) projection system. Image processing software was used to carry out the process. Detail of the transformation process (co-efficient, concept, etc) can be found in updated report by Malaysian Survey Dept. (2009). The transformation is necessary in order to validate the results against the rain gauges which were mostly having local coordinates. This is because validating the PCA-transformed daily rainfall estimates is conducted against the rain gauges which have local projection format. In addition, another reason is to make the output of this study to be compatible for the local hydrology and environmental water management systems.

8 datasets of 3-hourly rain-rate imageries from 0900 GMT until 0900 GMT of the next day were assumed to be the total rain rate in a single day. The reason for the hour selection (0900-0900) is to synchronize the satellite-based daily rainfall estimates with the actual rain gauge measurement standard which is conducted from 0800 until 0800 of the next day. Because of the constraint where TRMM data was supplied at specific 3-hourly gap starting from 0000 GMT and perfect simultaneous sampling was impossible, the hour selection was opted to minimize the time-lag between TRMM and rain gauge data. The common approach of estimating daily rainfall from 3-hourly rain rate data by accumulating 3-hourly rainfall for the entire 24 hour period is conducted (Fig. 4.4). Throughout this chapter, it is referred as direct accumulation approach (DA). In this study, the introduced PCA-based rainfall performance was relatively evaluated against the DA-based rainfall.

For this preliminary experiment, 91 rainy days with various intensities during the seasonal heavy monsoon thunderstorm period from three different years, 2003, 2005 and 2008 were selected. These days were sampled during November and December. Equivalent number of samples for each rainfall intensity category was acquired. Based on the rainfall distribution from the six months of heavy thunderstorm seasons, the rainfall is classified as low (<13mm)(1st quartile), medium (13-29mm, 2nd quartile) and high (>30mm, 3rd quartile). The daily rainfall average of the selected basins for two consecutive months (Nov. and Dec.) of the heavy thunderstorm seasons in year 2003, 2005 and 2008 is included in Appendix 2.
Phase 2- Principal component analysis for effective daily rainfall estimates from TRMM rain-rate imageries.

Principal components were computed for 8 datasets of 3-hourly rain-rate data from TRMM satellite covering 24 hours daily interval from 0900 to 0900 of the next day. The original concept of applying PCA to compress information from multi spectral satellite remote sensing imageries introduced by Jensen (1996) was adapted in this study with slight modification. Below is the summary of the fundamental steps of application of PCA transformation for effective daily rainfall estimates;

i. PCA computes the eigenvalues and eigenvectors for all the 8 TRMM 3-hourly rain rate images which occur in a single day and creates 8 PCs corresponding to the similar number of input imageries.

ii. The second step is determining the effective PC. Effective PC means having large variance which in this case is referred to the rain rate information. The preferred condition is where the eigenvalue exceeds 1. Because the total eigenvalue is similar with number of 3-hourly rain rate images, having a PC with eigenvalue more than 1 means that more than 3-hourly information were compressed in a PC. This condition meets the hypothesis where significant correlated temporal rain-rate images can be compressed into minimal component and get the proportion recalculated through eigenvector. However, to evaluate the robustness of this approach, the threshold for the PC selection is set based on three condition, eigenvalue > 1, > 0.5 and >2.

iii. The condition set in this study is that the eigenvalues for the PC must exceed 1. This condition was chosen based on common practice found in PCA applications since there is less specific information regarding the standard when dealing with remotely-sensed imageries. Most successful attempts have PC eigenvalue standard from 1.5 to 5.4 (Estornell et al. 2013; Munyati 2004; Jensen, 1996)

iv. The next step is to recalculate the total rain rate for each PC. This is done by multiplying the original rain-rate value and corresponding eigenvector.
The positive and negative sign was neglected since it only represents the linear relationship trend against the respective axis. The new rain rate was the sum of all multiplication of eigenvector and original datasets. Equation 1 described the calculation process.

The final step is to compute the total daily rainfall. Once the total rain-rate for every selected PC is obtained, it has to be multiplied by 3 since the original nature of this rain rate is 3-hourly average. The total daily rainfall is acquired by summing the products from every PC. The computation is expressed in equation 2.

\[
PC_k = \sum_{i=1}^{n} (E_{ki} \times P_i) \quad Eqn. 1
\]

\[
PCA \text{ daily rainfall estimates} = \sum_{i=k}^{n} (PC_i \times 3) \quad Eqn. 2
\]

where, \(n\) is number of rain-rate images, \(PC\) is principal components on total rain-rate, \(E\) is eigenvectors, \(P\) is original rain-rate.

Phase 3 - Validating the effectiveness of PCA-derived daily rainfall estimates

The purpose of this validation process is to determine the effectiveness of PCA transformation to produce precise daily rainfall estimates compared to the normal procedures of estimating daily rainfall by directly accumulating every 3-hour rain-rate for 24 hour period. Two statistical indicators were used in this verification, first is the average ratio between the satellite rainfall estimates over the actual rainfall measurement, and the second is the root mean square error (RMSE). Interpolated areal rainfall from rain gauge data was used as reference data. Dense rain gauge network data were then interpolated using the co-Kriging interpolation scheme which has been one of the commonly used approaches in generating effective areal rainfall (Mair & Fares, 2011). Total of 78 rain gauge records collected from Department of Irrigation and Drainage Malaysia were used as inputs.
Calculating bias ratio provides information on the ability of TRMM on describing the actual rainfall on the ground. The bias ratio percentage value of 0% represents perfect condition of satellite rainfall estimates with actual rainfall. Meanwhile, positive percentage of bias ratio value indicates the condition of satellite overestimation while negative means underestimation. The RMSE was used to quantitatively measure the average difference between both satellite and ground measurements. The ratio and RMSE were calculated for each corresponding pixel. Equation 3 and 4 described the calculation process of the ratio and RMSE respectively. In addition, number of rainy days which successfully improved in term of bias and RMSE were also identified. This evaluation intends to examine the overall consistency of PCA performance over large amount of samples.

\[ \text{Bias ratio} = \frac{1}{n} \sum \frac{R_{sat}}{R_{rg}} \times 100 \]  \hspace{2cm} \text{Eqn. 3}

\[ \text{RMSE} = \sqrt{\frac{\sum (R_{sat} - R_{rg})^2}{n}} \]  \hspace{2cm} \text{Eqn. 4}

where, \( n \) is number of pixels, \( R_{sat} \) is satellite derived rainfall, \( R_{rg} \) is areal interpolated rainfall from rain gauge data.

To further examine the PCA effectiveness, related variables associated with the heavy tropical thunderstorm were calculated using the raw TRMM rain-rate images, namely co-efficient of variation (COV) and rainfall intensity. A linear relationship among those two variables was also investigated as it may influence the PCA performance. Regarding the impact of PCA on altering the actual spatial rainfall patterns on the ground, spatial correlation between DA and PCA-based rainfall against the areal rainfall is computed. Any significant changes between those two rainfall products indicated PCA effects on spatial rainfall properties.
RESULTS

Effectiveness of PCA-based approach in high precision daily rainfall estimation

Comparison of principal component analysis (PCA) and direct accumulation (DA) approaches to rainfall estimation

PCA-based rainfall estimates showed varied performance in reducing bias and improving RMSE compared with DA-based rainfall estimates at different threshold and rainfall intensity (Table 4.1). On average, the PCA results using threshold 1.0 and 2.0 showed positive improvement in reducing bias ratio and RMSE. However, only PCA with threshold 1.0 produced more consistent performance in reducing bias ratio on each rainfall intensity classes (55-67%). PCA results using threshold 2.0 resulted increased bias ratio in medium rainfall intensity (~52%). Nevertheless, in RMSE reduction, the improvement indicated by PCA with threshold 2.0 was relatively higher (30%) compared to the threshold 1.0 (22%). PCA with threshold 0.5 showed the worst performance where no improvement was shown either in bias ratio and RMSE.

Statistics on successful improved days by PCA approach

The number of days that bias ratio and RMSE were successfully improved using the PCA technique varied by threshold and were ranked as the following: threshold 1.0, 2.0 and 0.5 (Table 4.2). It was found that 71% and 56% of the total samples were improved in bias and RMSE reduction, respectively. The percentage of successful treated days by PCA (threshold 1.0) was declining as the rainfall intensity increases. It was found that despite the better performance of PCA with 2.0 threshold in reducing the RMSE, the number of samples involved were lower compared to PCA with 1.0. This informed us that the PCA results with threshold 1.0 have positive effects consistently on many rainy days compared to the other threshold results.
PCA threshold effects on bias and RMSE reduction

Different threshold setting produced varied results to the bias ratio and RMSE reduction. The degree of positive improvement was only found when the PCA threshold was 1 and higher (Fig. 4.5). Using low threshold (0.5) would resulted negative improvement. Because employing lower threshold led to inclusion of many PCs, overestimation occurred due to the redundant rain-rate recalculation. Employing higher threshold (>2) was more effective in the condition of high bias ratio. This was clearly demonstrated during the low rainfall intensity condition where large bias ratio cases dominate (Fig. 4.6a). However, it produced negative impact during the low and medium negative bias (satellite underestimate). The excessive discard of the PCs components which occurred when using higher threshold, has caused large underestimation. This effect has largely influenced the PCA performance with threshold 2.0 during medium rainfall intensity (13-29mm) where bias ratio increment was indicated. Meanwhile, the initial 1.0 threshold produced consistent bias ratio reduction ability throughout every rainfall intensity condition particularly in medium and heavy ones.

RMSE increased with increasing rainfall intensity. PCA with threshold 1.0 and higher were able to reduce the RMSE during all rainfall intensity (Fig. 4.6b). In several large error occasions that occurred in medium and high rainfall intensity, the reduction can reached up to 50%. Using 2.0 threshold produced higher degree of RMSE reduction compared to 1.0. This was primarily due to the smaller number of PCs accounted. However, the error reduction capacity was declining as the error amount increased. Applying lower threshold (<1) has resulted an error increment in every rainfall intensity condition.

Another plausible factor that can explain performance of PCA instead of threshold variations is the spatial variability characteristics from low to high rainfall intensity condition. Further analysis on the rainfall intensity and co-efficient of variation (COV) which represents the spatial variability of the rain rate images indicated moderate non-linear relationship (Fig. 4.7). As low rainfall produced high heterogeneity (high COV), the orthogonal and uncorrelated rain rate images were effectively
identified. Thus, this lead to the explanation where discarding this PCs resulted larger impact on both bias ratio and RMSE reduction compared to the recalculation of the correlated rain-rate based on their corresponding eigenvectors. However, due to the moderate and non-linear relationship between COV and rainfall intensity, the direct relationship between COV to PCA transformation is difficult to be established.

Qualitative analysis of PCA-based rainfall compared to the ground areal rainfall

PCA approach was found to have less impact on the rainfall patterns. The computed spatial correlation between DA and PCA-based rainfall estimates at different threshold showed no improvement (Table 3). Small improvement was indicated at low and medium rainfall. Negative effect was found in higher rainfall cases. ANOVA test indicated that there is no significant difference between those results ($F_{0.05; 3, 360} = 0.015$). This means that PCA adjustment did not alter the rainy days where spatial correlation was high neither improving the spatial quality of low correlated rainfall images.

DISCUSSION

This study had demonstrated that PCA transformation can contribute to increased accuracy of daily rainfall estimates compared with the DA approach when using series of 3-hourly TRMM rain-rate images during heavy tropical thunderstorms season. PCA concept of using only the compressed significant rain-rate images in effective PCs play larger role in improving the bias ratio and RMSE compared to recalculate new rain-rates based on their contributions. However, selecting appropriate threshold was pivotal in determining the positive outcomes. The PCA approach with threshold 1.0 produced positive consistent performance for different rainfall intensity and involving larger number of rainy days compared to the other higher threshold. Choosing low threshold would lead to dramatic rainfall overestimation and introduced larger error. On the other hand, using higher threshold would lead to over rainfall reduction due to excessive discarded PCs. This condition commenced large satellite
underestimation if the pre-adjusted bias ratio is low (positive) or already in underestimation condition, which yield negative improvement.

Despite the PCA declining performance as the rainfall intensity increases, with appropriate threshold selection it would able to improve the accuracy of the satellite rainfall estimates during medium and heavy rainfall condition which contributed approximately 88% of the rainfall during the heavy rainfall thunderstorm season. In addition, the promising performance showed by the PCA during low rainfall indicated that the proposed technique would contribute to enhancing the analysis on drought phenomena. The uncertainties on satellite underestimation which primarily occurred during lower rainfall intensity condition was unable to be effectively treated by using PCA approach. This leaves gap for alternative technique to overcome the drawbacks. Instead, higher bias ratio during low rainfall intensity found in this study was rarely reported because many validation studies prominently highlight the false alarm ratio during heavy rainfall condition (e.g., Roongroj & Long et al, 2008).

The results of PCA-based rainfall estimates were comparable in accuracy to the general TRMM results obtained by previous studies performing reliable hydrological simulations in tropical basins. It was found that efficient hydrological modeling can be achieved from TRMM daily rainfall estimates, with average quantitative RMSE of ±90mm and +10-12% bias (Su et al. 2007; Collischonn et al. 2008). Our effective PCA-adjusted rainfall estimates (threshold 1.0) had lower average of RMSE (±30mm) but relatively having larger bias range (+8-24%) despite a 62% reduction compared with the DA approach. This study suggests that the PCA approach can be used as input to produce more precise hydrological simulation in tropical basins that lack or have only sparse rain gauge networks.

Whether the PCA-based rainfall estimate was able to improve the TRMM quantitative accuracy locally and at climatically identical sites cannot be determined by relevant comparisons with other studies for several reasons. First, there is the issue of differences in temporal scale. Validation conducted at a similar site on peninsular Malaysia by Semire et al. (2012) and in a nearby region of southern Thailand by
Roongroj and Long (2008) implied the need for annual and monthly scale assessments. In another case study, the TRMM data products used were also different (Omotosho et al. 2013). The only relevant reference in the literature discussed the reported accuracy by Varikoden et al. (2010). However, based on the visual interpretation from the correlation graph, showed that the RMSE obtained in this study was comparable to theirs (±35mm).

The application of PCA procedures introduced in this study provided an alternative solution to obtaining reliable daily rainfall estimates from TRMM rain-rate data during heavy tropical thunderstorms season in a region with sparse rain gauge and related reference data. Hence, other temporal downscaling technique required surrogate information and complex computation process (Tao & Barros, 2010; Ryo et al. 2014; Lidard et al. 2007). The output of PCA in this study also revealed useful spatio-temporal information, which temporal rain-rate played significant role in signifying the total daily rainfall estimates. Despite the simplicity and no rain gauge usage in the processing, the PCA approach has high potential to be applied as an automated technique to obtain increased accuracy of TRMM daily rainfall estimates especially in humid tropical regions with sparse rain gauges for appropriate calibration.

Future research should employ the use of the higher temporal scale rain gauge measurement (i.e. hourly). It would accurately verify whether the associated correlated rain-rate images eventually constitute to actual measured rainfall variation on the ground. Thus, actual rainfall temporal variation can confirm the assumption made by the PCA approach. In addition, using the updated version of TRMM version 7 is recommended as the processing method was improvised. With the launch of TRMM’s successor, Global Precipitation Mission (GPM), prospect are bright for the future use of satellite precipitation data and therefore advances in the data treatment are significant. Bridging the gap between space-based global scale measurements into appropriate local-scale information is paramount in developing an operational framework using space-based integrated systems in environmental hydrology (Pan et al. 2008).
CONCLUSION

This chapter presented the application of PCA transformation for high-accuracy estimation of daily rainfall using temporal 3-hourly TRMM rain-rate images during heavy tropical thunderstorms season in the small-sized catchment. Preliminary analysis showed that PCA transformation with specified threshold of 1.0 outperformed the common approach of direct accumulation (DA) in producing accurate daily rainfall estimates. The PCA transformation successfully reduced bias ratio and measurement differences at averages of 62% and 22%, respectively. The best average accuracy of PCA-based daily rainfall estimates during heavy thunderstorm (>30mm) that largely contributed to the total rainfall during the wettest season is bias+24% and standard error of ±30mm. The ability of PCA transformation to correlate significant rain rate and recalculate the temporal proportion was influenced by the threshold selection and rainfall intensity at both spatial and temporal scales. The effect toward spatial qualitative rainfall element was minimal, and meaningful rainfall information was retained. The output of this study is useful in minimizing large bias during heavy thunderstorms season using common DA approach to obtain high precision TRMM daily rainfall estimates in areas of sparse and inefficient rain gauge management.
Table 4.1 Comparison of ratio and root mean square error (RMSE) between two different daily rainfall estimates against interpolated areal rainfall. Negative sign on ratio indicates satellite underestimation condition.

<table>
<thead>
<tr>
<th>Rainfall intensity (mm/d)</th>
<th>Total no. of sample days</th>
<th>Ratio (%)</th>
<th>Improvement (%)</th>
<th>RMSE (mm)</th>
<th>Improvement (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>PCA DA Threshold</td>
<td>PCA Threshold</td>
<td>PCA DA Threshold</td>
<td>PCA Threshold</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.0 0.5 2.0</td>
<td>1.0 0.5 2.0</td>
<td>1.0 0.5 2.0</td>
<td>1.0 0.5 2.0</td>
</tr>
<tr>
<td>3-12</td>
<td>25</td>
<td>67 -90 81</td>
<td>16 10 20</td>
<td>8</td>
<td>38 -24 50</td>
</tr>
<tr>
<td>12-29</td>
<td>36</td>
<td>65 -291 -52</td>
<td>27 23 36</td>
<td>20</td>
<td>15 -32 26</td>
</tr>
<tr>
<td>&gt;30</td>
<td>30</td>
<td>55 -72 47</td>
<td>64 56 75</td>
<td>55</td>
<td>13 -18 14</td>
</tr>
</tbody>
</table>

Table 4.2 Quantity of the rainy days that improved using the PCA-based approach.

<table>
<thead>
<tr>
<th>Rainfall intensity (mm/day)</th>
<th>Total days</th>
<th>Bias ratio</th>
<th>RMSE</th>
<th>PCA threshold</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>PCA threshold</td>
<td></td>
<td>1.0 0.5 2.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.0 0.5 2.0</td>
</tr>
<tr>
<td>3-12</td>
<td>25</td>
<td>20</td>
<td>4</td>
<td>16</td>
</tr>
<tr>
<td>12-29</td>
<td>36</td>
<td>27</td>
<td>5</td>
<td>28</td>
</tr>
<tr>
<td>&gt;30</td>
<td>30</td>
<td>17</td>
<td>8</td>
<td>14</td>
</tr>
</tbody>
</table>

Table 4.3 Spatial correlation between TRMM satellite rainfall estimates from DA- and PCA-based approaches against areal rainfall from interpolated rain gauge measurement (RG).

<table>
<thead>
<tr>
<th>Rainfall range (mm/day)</th>
<th>DA PCA threshold</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1.0 0.5 2.0</td>
</tr>
<tr>
<td>3-12</td>
<td>0.151 0.159 0.154 0.115</td>
</tr>
<tr>
<td>12-29</td>
<td>0.247 0.248 0.240 0.266</td>
</tr>
<tr>
<td>&gt;30</td>
<td>0.206 0.199 0.202 0.189</td>
</tr>
</tbody>
</table>
Figure 4.1. Location of study site. (a) map of Southeast Asia (b) study area.

Figure 4.2. 62 year-averaged monthly climate and rainfall of the study area (1949-2010). The record is taken from Kota Bharu station.
Figure 4.3. Methodology
Figure 4.4. TRMM pixels at the study site. Circles represent the rain gauges.

Figure 4.5. Scatter plot of bias ratio between PCA-against DA-based daily rainfall estimates.
Figure 4.6. Comparison between DA- and PCA-based rainfall estimates with different threshold. (a) bias ratio and (b) RMSE. *The horizontal line in (a) is representing the perfect ratio (1.0). Ratio values that higher and lower than 1 were considered as satellite overestimate and underestimate, respectively.
Figure 4.7. Relationship between the 3-hourly rain-rate and co-efficient of variation (COV).

$R^2 = 0.6849$
APPENDIX 2

Figure A1. Hyetograph of the study area on November and December in 2003, 2005 and 2008.
CHAPTER 5

SPATIAL DOWNSCALING OF SATELLITE PRECIPITATION DATA USING FINE SCALE CO-EFFICIENT DERIVED FROM HIGH RESOLUTION AREAL PRECIPITATION ANALYSIS

INTRODUCTION

Utilizing satellite precipitation as an alternative to support the rain gauge measurement is considered as a useful option for acquiring effective spatio-temporal rainfall measurement. However, the application of the satellite precipitation data for small-sized region is often hindered by their coarse grid size (Mahmud et al. 2015; Behrangi et al. 2011). Most of the respective sites were located at Southeast Asia, which comprises many small sized peninsula and island. An effort to improve the grid known as spatial downscaling has become the contemporary focus for scientist in respective fields. An appropriate spatial downscaling algorithm for humid tropical environment is rarely reported despite the enormous amount of rainfall received. In addition, many related equatorial regions are developing countries whereas exposed to the ineffectiveness of rainfall monitoring activities due to the economical, technical and social constraints (Musiake, 2003).

At present, advances in spatial downscaling were centralized by using the higher spatial resolution of related environmental parameter as a rainfall predictor for the improved grid precipitation data. The key of applying such approach is determined by the strength of the relationship between the rainfall and its explanatory variables at site specific. Often, multiple regression-based relationship between rainfall, vegetation and elevation were used (Chen et al. 2014; Cho and Choi, 2013; Park, 2013; Immerzeel et al. 2009; Jian et al. 2013; and Shi et al. 2015). Employing such technique in the dominant humid tropical region is typically less
suitable due to the weak degree of relationship compared to temperate environment. In addition, there are few literatures emphasized on that particular relationship. Another dominant factor should be considered to be proxy variables or assumption.

Literature cited that rainfall distribution in the tropics is closely associated with water vapor (Muller et al. 2009). Asian monsoon season plays significant role in characterizing the local rainfall variation in many parts in the tropical region of Southeast Asia (Dale, 1959; Cheang, 1980). There is evidence indicating that despite the rainfall pattern changes at both regional and global scales proved to be prominent in recent decades, the local seasonal rainfall is remained low in variance (Wong et al. 2008; Aldrian et al. 2007; Fein & Stephens, 1988). This provides opportunities where fixed seasonal rainfall can be used as a proxy to downscale the satellite precipitation. It is possible with regards that the seasonal rainfall co-efficient is introduced at a finer resolution than the initial satellite resolution. This assumption is yet to be investigated.

Prior to that, this study hypothesized that in the low seasonal rainfall variance region the historical satellite and ground precipitation bias records can be used as a calibration co-efficient for future dataset. In the condition where the specific pixel basis co-efficient value was available at finer resolution grid than the initial satellite resolution, it is plausible to produce the new satellite precipitation with improved grid resolution. To test this hypothesis, Peninsular Malaysia is selected as an experimental site. The co-efficient of variation (COV) seasonal precipitation was reported to be low (Wong et al. 2008). An improved high resolution co-efficient (HRC) is expected to be derived using the ratio records between high resolution ground (0.06 deg.) and satellite rainfall from 1998-2007. A total of 980 rain gauges network was used to create a higher resolution areal rainfall (0.06 deg.). A new satellite precipitation data with similar resolution (0.06 deg.) is created using the HRC from 2008-2010 and evaluated.
MATERIALS & METHOD

Study Site

Peninsular Malaysia is selected in this study. Detail can be found in the Chapter 2.

Data

Satellite Precipitation & Rain Gauge Data

13 years satellite precipitation data from 1998-2010 provided by Tropical Rainfall Measuring Mission (TRMM) were used. Detail pre-processing can be found in Chapter 2. For the rain gauge data, similar monthly basis rainfall was used in Chapter 2. In this study, the rain gauge & satellite data were divided into two separate datasets. One is for the derivation of the co-efficient (1998-2007) and second for accuracy validation (2008-2010).

Downscaling TRMM data using high resolution co-efficient

Phase 1 - Preparation of satellite and ground precipitation data

A monthly basis of areal precipitation at 0.06 deg. resolution was created from the dense rain gauge network by using the kriging interpolation scheme. The satellite data is being re-gridded from the original 0.25 deg. resolution into similar resolution of 0.06 deg.

Phase 2 – High resolution co-efficient (HRC) derivation

The high resolution co-efficient (HRC) is derived through two fundamental steps. First, the re-gridded satellite precipitation is divided with the ground precipitation. This process is conducted at monthly basis using the data from 1998-
2007. Prior to this process, each pixel will have a unique monthly basis ratio. The second step is calculating the average bias ratio for each pixel. The outcomes of this process were 12 images with specific pixel basis co-efficient. The HRC value of 1.0 indicating the perfect condition between satellite and rain gauge measurement and thus, require no modification. Meanwhile, HRC value that more than 1.0 indicating overestimate and vice versa. This output is also known as high resolution co-efficient (HRC) later in this text. Equation 1 showed the HRC calculation.

\[
HRC_{(i,j)} = \frac{1}{n} \sum_n \frac{Sat_{(i,j)}}{Rg_{(i,j)}}
\]

where \( HRC \) is the high resolution co-efficient, \( Sat \) is satellite precipitation data, \( Rg \) is the areal ground precipitation data.

**Phase 3 – Downscaling the TRMM satellite data using HRC**

To produce the improved satellite precipitation using the HRC, an independent TRMM rainfall data from 2008-2010 were used. Each pixel value of the re-gridded raw TRMM data (0.06°) was divided with the corresponding HRC for each pixel. The modified value by the HRC will produced the accurate proportion of rainfall at 0.06°. Equation 2 showed the equation of the downscale process.

\[
DSat_{(i,j)} = \frac{RSat_{(i,j)}}{HRC_{(i,j)}}
\]

Eqn. 2
where $DSat$ is the downscaled satellite precipitation $(0.06\degree)$, $RSat$ is the raw satellite precipitation data $(0.06\degree)$, $HRC$ is the downscale co-efficient $(0.06\degree)$.

**Phase 4 – Accuracy validation**

To verify the performance of the HRC, the bias ratio and root mean square error (RMSE) between the HRC-based product and areal precipitation from 2008-2010 is calculated. In addition, the co-efficient of variation (COV) for the historical bias records is computed. This is purposely to determine whether the bias records were developed under low seasonal variance or not.

**RESULTS**

*HRC-based downscale product performance*

The three years average showed that the HRC based precipitation product with lower bias ratio compared to the raw $(0.25\text{ deg.})$ data product at all respective climatic region (Table 5.1). The bias ratio reduction capacity varied from 53 to 77%. The impact of HRC on relatively lower bias ratio (West & Highland) resulted in changes of the ratio trend from satellite overestimate to underestimate. It was considered as a positive change because the adjusted ratio is lower than the raw data. This result indicated that the use of HRC is able to produce high resolution areal precipitation with smaller bias ratio.

The multiple time-series graphs showed that there was a small declination in the correlation and reduced error between the satellite and ground rainfall measurement after the use of HRC in all climate regions except west (Figure 5.1). In the west region (Figure 5.1b), the HRC product has underestimated the ground rainfall measurement regardless of monsoon seasons. This indicated that the ability of HRC to produce good quality of high resolution areal rainfall is limited in the west...
region. From the common correlation and error trend found in every region, it is found that the smaller sized grid transformation leads to error minimization but minor decreases the correlation. Plausible reason of this condition was that the generation of smaller grid size areal rainfall introduced higher spatial rainfall variation as the heterogeneity of the pixel values has increased.

Historical bias ratio records co-efficient of variation & HRC performance

The co-efficient of variation (COV) of the ratio between satellite and ground rainfall ranged from 52% (southwest) to 77% (northwest) (Figure 2). The COV score is the highest (~113%) during the dry season at all regions (Feb). This informed us that the satellite precipitation has limitation to represent the local rainfall variation during dry seasons. Large and frequent outliers were identified indicating that downscaling the grid size induced higher rainfall value variation. Based on this evidences, this study clarified that despite the high COV of average ratio records between satellite and ground measurement, the value to a certain extent capable to calculate the rainfall value for the newly created finer resolution grid (0.06 deg.). This finding also informed us that the effectiveness of HRC is less related with the COV. Instead, the present large overestimates from the raw TRMM data (0.25°) were simply minimized through the use of quite robust HRC value.

Comparison between HRC-downscale product and other satellite precipitation product

The TRMM-HRC product did produce lower bias ratio compared with other coarser grid resolution data products at most local climate regions (Table 5.2). TRMM downscale performance is comparable to the high resolution GsMAP precipitation product (0.1 deg.) except in the west region. The coarser resolution data were unable to depict the local spatial rainfall in the northwest and west region. This
indicates that the improved grid of HRC data product is able to portray the spatial rainfall variation and measurement more consistently in every climate region.

**DISCUSSION**

The use of HRC produced higher resolution (0.06 deg.) of satellite precipitation product is superior to their initial grid size (0.25 deg.) with improved spatial representation of local rainfall variation. The trade-off of the process is a small declination of depicting the temporal local rainfall variation. The HRC results were found to be less affected by the low COV of previous historical bias ratio records. Because the bias ratio from the raw TRMM data was high (average = 88%) the HRC robustly minimized those biases. Instead, the impact of HRC in low bias region caused changes in bias ratio trend from overestimate to underestimate. The use of HRC, however, did not result in negative impact to the TRMM data and is reliable to be used as reference for spatial downscaling. Comparatively, the performance of HRC product is better than other satellite precipitation products at the study site.

There were three plausible reasons for the high COV. First is due to the process of spatial downscaling where number of pixels increased. When more pixels were generated and bias is calculated intensively by per pixel basis, larger variability between satellite and ground precipitation possibly occurred (Tao & Barros, 2010). The second factor was the short temporal sample (10 years). Many related Asian monsoon studies encouraged longer temporal sample (>30 years)(e.g., Wong et al. 2008). The last factor is the monsoon seasonal shift and changes due to regional climate change because of El-Nino Southern Oscillation (ENSO)(Hendon, 2003; Loo et al. 2014). There were high magnitude of El-Nino observed in 2002 and La-Nina from 1999-2000 which coincides with our sampling period. This informed us that careful selection and addition of the sampling years and period would be recommended for future works.
The HRC procedures which require simple calculation and process make it suitable for operational application in the case of previous high resolution bias ratio can be acquired. In addition, its independency from using ground reference is recommended technique for a region where rain gauge information is difficult to obtain due to ineffective management. Other spatial downscaling approach which uses advanced mathematics (e.g. Tao & Barros, 2010; Yatagai et al. 2014) is more complex and requires present rain gauge information. Furthermore, the downscale grid size (0.25 deg.) was coarser that achieved in this study. The high resolution precipitation product generated in this study would overcome uncertainties for small-sized Southeast Asia humid tropical region (Mahmud et al. 2015; Roongroj & Long, 2008; Prasetia et al. 2013).

CONCLUSION

This chapter tested the hypothesis whether the high resolution of historical bias records in the low variance seasonal monsoon rainfall can be used as co-efficient to spatially downscale the TRMM satellite precipitation data. The use of high resolution co-efficient (HRC) was able to produce improved TRMM satellite precipitation data from 0.25 to 0.06 deg. with lower bias ratio. The trade-off of the process is small declination of correlation. The HRC however is produced under the high variance seasonal condition (63%), thus indicating weak relationship between HRC and the output. It is identified that the initial TRMM data possess high bias. The variability increased when more pixels were created in the downscaling process. The HRC robustly minimized those errors but no negative impacts on the quality were indicated. The simplistic and effective procedure is applicable to spatially downscale the satellite precipitation in low variance of seasonal monsoon.
Table 5.1 Bias ratio comparison between raw TRMM data and HRC-downscale product.

<table>
<thead>
<tr>
<th>Region</th>
<th>Raw</th>
<th>HRC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Northwest</td>
<td>3.14</td>
<td>1.61</td>
</tr>
<tr>
<td>East</td>
<td>1.88</td>
<td>1.20</td>
</tr>
<tr>
<td>West</td>
<td>1.34</td>
<td>0.84</td>
</tr>
<tr>
<td>Southwest</td>
<td>1.78</td>
<td>1.32</td>
</tr>
<tr>
<td>Highland</td>
<td>1.27</td>
<td>0.91</td>
</tr>
</tbody>
</table>

*1.0 is perfect ratio, >1.0 is satellite overestimate, <1.0 is satellite underestimate.

Table 5.2 Comparison between the HRC-downscale product and other satellite precipitation products.

<table>
<thead>
<tr>
<th>Satellite Precipitation Products</th>
<th>Grid size (deg.)</th>
<th>Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>N</td>
</tr>
<tr>
<td>TRMM V7 - HRC</td>
<td>0.06</td>
<td>2.2</td>
</tr>
<tr>
<td>GsMAP</td>
<td>0.10</td>
<td>1.5</td>
</tr>
<tr>
<td>PERSIANN</td>
<td>0.25</td>
<td>15.0</td>
</tr>
<tr>
<td>CMORPH</td>
<td>0.25</td>
<td>10.6</td>
</tr>
<tr>
<td>CPC</td>
<td>0.50</td>
<td>21.0</td>
</tr>
<tr>
<td>GPCP</td>
<td>1.00</td>
<td>27.0</td>
</tr>
<tr>
<td>CMAP</td>
<td>2.50</td>
<td>15.0</td>
</tr>
</tbody>
</table>
Figure 5.1 Time-series between the ground areal rainfall, raw TRMM and HRC-downscale product from 2008-2010.
Figure 5.2 Co-efficient of variation of the historical bias record from the satellite and ground areal rainfall
CHAPTER 6

GENERAL DISCUSSION

The accuracy of Tropical Rainfall Measuring Mission (TRMM) over Peninsular Malaysia

The assessment using high resolution areal rainfall revealed three major findings. I found that the ability of TRMM to depict the seasonal rainfall significantly varied with monsoon seasons. Good correlations were found during very wet (Nov.-Feb.) and dry (Mar.-Apr.) monsoon seasons. Second, the high measurement differences directly associated with increasing rainfall. For the first and second condition, it was expected to be related with the low sensitivity of TRMM toward the medium precipitated cloud suggested by other studies conducted in the nearby region. Third, the TRMM ability to represent the spatial pattern was found to be varied by local climate region. The terrain variation and different precipitation mechanism between satellite and rain gauge were the main explanatory factors for the case. Overall effectiveness of the TRMM was relatively higher during very wet and dry monsoon seasons.

Based on my findings, it is assumed that employing TRMM for assessing the ecohydrological condition in humid tropical region of Peninsular Malaysia was depending on the size of the site, location and monsoon seasons. The smaller grid assessment has provided useful information on the limitation of TRMM in specific regions and monsoon seasons over Peninsular Malaysia. Assessing very wet and dry monsoon seasons were plausible. The proposed downscaling technique is plausible to be implemented in other humid tropical region with similar hydrometeorological characteristics. An improved TRMM product with smaller grid would increase the opportunity to examine smaller scale site with better accuracy. The monthly and annual rainfall analysis can be conducted with good precision. However, the daily assessment was still doubted due to its high variability.
Opportunity for downscaling the TRMM in humid tropical region

Direct re-gridded of the TRMM coarse pixels was not sufficient to achieve good seasonal spatio-temporal accuracy. A seasonal downscaling during summer monsoon (May-October) season is required. The large measurement error that occurred during heavy monsoon seasons is necessary to improve the quality of TRMM rainfall estimates. Improving the coarse grid of the TRMM may lead to increase sensitivity toward low and medium precipitated cloud that is dominant during the medium wet monsoon seasons in Peninsular Malaysia. Due to lack of literature regarding suitable spatial downscale approach in humid tropical region, it remains a challenging gap due to availability of high resolution proxy data.

Temporal downscaling during heavy tropical monsoon thunderstorm season using Principal Component Analysis

Correlating the significant rain rate into single axis and computing its value based on the correlation strength prove to be reliable in producing higher precision of rainfall estimates. The performance of the TRMM however, is influenced by the threshold selection. Threshold values higher than 1.0 were suitable for the low rainfall intensity condition. Meanwhile, the threshold value equal to 1.0 was preferred for medium and high rainfall intensity. The PCA approach did not affect the spatial quality of the rainfall images. The proposed approach is practically useful for no or sparse rain gauge condition since the procedure requires no rain gauge data.
Using seasonal co-efficient to spatially downscale the TRMM in humid tropical region

Because of the unsuitability of vegetation and elevation as proxy downscale variables in humid tropical region, the potential of consistent seasonality as predictor variable was investigated. Bias records between satellite and areal rainfall that were generated at high resolution (0.06 deg.) was capable to downscale the TRMM data from 0.25 deg. to 0.06 deg. The downscale product has a higher capacity to represent the local rainfall variation. However, there is no convincing evidence to support that the effectiveness of downscale product was due to the low-variance of seasonal bias records between satellite and ground rainfall. It is suspected that the high resolution co-efficient was robustly minimized the presently large bias possessed of the raw TRMM data.

Implications of TRMM satellite precipitation data in ecotourism management

The output of this study has significant implication in sustaining the tourism sector in humid tropical region directly and indirectly.

Directly, the high accuracy of satellite precipitation data during the heavy rainfall season would provide support for effective disaster management of the ecotourism locations. The improved resolution of rainfall data (~40km²) enabled thorough rainfall assessment at sub-basin scale of tropical forest reserve and national parks, particularly the small-size types (<500km²) which was limited by the initial rainfall data (~625km²). Primarily, the vulnerability of related ecotourism sites to water related disaster, namely flood and landslide can be assessed effectively using efficient rainfall data. Many tropical forests and the related ecosystems commonly have poor on site rainfall measurement due to remote location, high infrastructure cost of ground instrument and difficult access. In Peninsular Malaysia, seasonal flood during the peak of northeast monsoon (Nov.-Dec.) (Chan, 1997) is very common and has affected the ecotourism sites, including the forest, upper stream water fall, downstream river, and river banks in the east coast (The Star online press, 2014).
Heavy and prolonged rainfall has been identified as a major factor that triggers landslides in several tropical sites (APDC, 2006; Larsen & Simon, 1993). Hiking and forest tracking were among ecotourism activities that might be threatened by those risks. This disaster would negatively affect the ecotourism sites condition as well as the visitor’s safety.

In consequences, prevention and mitigation measures for those tourism sites can be improved. For example, by using the previous rainfall information, responsible authorities can acknowledge tourist with potential risk areas related to flood areas. This is useful for them in planning their trip to minimize risk of unprecedented casualties. In addition, the appropriate trail and necessary precautions to hikers in potential disaster hotspot also can be provided. For mitigation, a site specific plan for evacuation and disaster prevention measures can be improved to increase safety to tourist. Safety and risk perception is an important variable influencing the tourist travel patterns (Kellens et al. 2011). Having a good framework for disaster management is essential in minimizing the risk as well as keeping healthy perception to tourist.

Another contribution of efficient rainfall maps is to assist the related tourism authorities to maximize the visitor satisfaction in rainfall or water related ecotourism attractions. Tourism attraction like tropical waterfall, river, and lake were sustained and influenced by seasonal rainfall. Because the satellite precipitation data have good performance during heavy and dry seasons, the integrated rainfall data and other related maps would help to identify suitable period to the related tourism activities. As a result, proper visiting schedule and plan can be prepared for the prospective ecotourism locations. For instance, water sports including river rafting and kayaking were not suitable during the dry season due to low water level. Related activity such as fishing was also indirectly influenced by seasonal rainfall. For example, relationship between fish assemblages and seasonal rainfall in tropical site has been discovered (e.g., Barletta et al. 2008). Furthermore, other associated activities such as wildlife sighting may be difficult and limited if heavy torrential rainfall occurred. In nature-based tourism, tourist tends to show a high desire in experiencing culture,
nature and wilderness (Kalterborn et al. 2011). Maintaining good tourist satisfaction toward the nature would ensure the longevity of the tourism sector.

The indirect implication of the improved quality of satellite precipitation is it can support the comprehensive assessment of tropical vector diseases namely dengue and malaria in the ecotourism sites. The spreading tropical vector diseases due to seasonal rainfall variation have been reported in several tropical forest environmental settings (Wiwannitkit, 2004). Because tourism involved human traveling from short to large scale distance, consideration of the risk associated with health and disease is significant.

The satellite precipitation data are useful as rainfall measurement support in sparse or scarce rain gauge in a tropical ecotourism location. It directly contributed to the assessment of flood and landslide disaster threatening both the ecosystem and tourist. Thorough assessment of site specific condition would be helpful in assisting the disaster prevention and mitigation activities. The condition of ecotourism sites related to rainfall also can be determined to maximize the visitor satisfaction. Indirectly, intensive rainfall data can be used to examine communal tropical disease threatening the health in ecotourism sites.
SUMMARY

CHAPTER 1

The importance of rainfall in sustaining the tropical rainforest and related ecosystems with ecotourism site is stressed. Satellite precipitation data can be a useful alternative support to conventional rain gauge measurement. However, its accuracy in humid tropical region is hindered by several factors, including coarse grid size and measurement accuracy. The study conducted a thorough assessment using high resolution areal rainfall and technique development in anticipating both temporal and spatial drawbacks of widely used Tropical Rainfall Measuring Mission (TRMM) in humid tropical region.

CHAPTER 2

Peninsular Malaysia is selected for all experiments in this thesis. High density rain gauge network was acquired. Tropical Rainfall Measuring Mission data were used due to its advance of spatio-temporal resolution and sensor properties.

CHAPTER 3

An intensive seasonal scale validation of satellite precipitation data from Tropical Rainfall Measuring Mission (TRMM) was conducted in a small sized humid tropical environment. The improved validation accounted two important characteristics; local climate regions and fine grid resolution (0.125 deg.) of ground reference rainfall from the dense rain gauge network. The ability of TRMM to depict seasonal variation and quantification of exact rainfall on the ground in standard unit (RMSE) significantly varied between monsoon seasons. The seasonal correlation was good only during very wet and dry monsoon seasons. Measurement differences were significantly high only during wet seasons. The capability of TRMM to depict local rainfall pattern varied regionally and were dominated with satellite overestimation at an average of 60%. The usage of high resolution areal precipitation had imminent
effect only on the ability of the TRMM to depict spatial rainfall patterns during the dry season because local regional rainfall patterns were more heterogeneous. The overall efficiency of TRMM was only found to be good during very wet and dry season.

CHAPTER 4

A large measurement differences typically occurred during heavy tropical thunderstorm. Anticipating the conflict is possible using the ground reference rain gauge data, but difficult if the rain gauge is absent. A thorough investigation indicated that the severe overestimated case was caused by the conventional daily rainfall estimate procedure and different method of rain rate estimation between the rain gauge and satellite. The conventional direct accumulation of three hourly rainfall data exposed to excessive use of rain rate data used during heavy rainfall condition. The different method of rain rate estimation might cause the misrepresented actual rain event on the ground because satellite rainfall detection is depending on the relationship between the attenuation and rain rate, not rainfall on the ground. An alternative approach to anticipate this limitation without using rain gauge is proposed. I hypothesized that the correlated rain rate images can be compressed into fewer components and their corresponding rain rate is re-computed can minimize the large bias. Principal component analysis was employed and the result agreed with the hypothesis. However, it was influenced by the rainfall intensity and eigenvalue thresholds used in selecting the compressed components.
CHAPTER 5

An effort to refine the coarse grid of Tropical Rainfall Measuring Mission (TRMM) satellite precipitation data from 0.25 deg. to 0.06 deg. using the seasonal monsoon co-efficient was conducted. Based on the known literature that the co-efficient of variation of seasonal rainfall is low, it is assumed that the seasonal satellite-ground bias is also low and able to be used as a co-efficient to correct the future dataset. Despite the positive result of the HRC-downscale product, with a low bias ratio compared to the raw TRMM, there is less evidence to relate to low variance of monsoon seasonality. It was found that the HRC robustly minimized the large present error of TRMM data. The HRC-downscale product however is comparable and relatively better than other coarse grid satellite precipitation.

CHAPTER 6

The general findings from each respective chapter were discussed. It includes the detail spatio-temporal accuracy of TRMM data at local climate regions, opportunity for downscaling, tackling the large bias during heavy monsoon thunderstorm season using principal component analysis, and producing higher resolution grid of TRMM data using spatial downscaling approach based on seasonal co-efficient of variation. The implications of TRMM satellite precipitation data to ecotourism sites and activities in the humid tropical region were discussed.
REFERENCES


Mahmud, M.R. (2012). Run-off modeling and mapping using rainfall and evapotranspiration estimates from remote sensing satellite data in Peninsular Malaysia;


