TREND OF DAILY RAINFALL AND TEMPERATURE IN PENINSULAR MALAYSIA BASED ON GRIDDED DATA SET

*Chee-Loong Wong1,2, Zulkifli Yusop1,3 and Tarmizi Ismail1,3

1 Faculty of Civil Engineering, Universiti Teknologi Malaysia, 81310 Johor Bahru, Malaysia
2 Department of Irrigation and Drainage, Malaysia, Jalan Sultan Salahuddin Kuala Lumpur, Malaysia
3 Centre for Environmental Sustainability and Water Security, Research Institute for Sustainable Environment, Universiti Teknologi Malaysia, 81310 Johor Bahru, Malaysia

*Corresponding Author, Received: 8 May 2017, Revised: 24 Nov. 2017, Accepted: 20 Dec. 2017

ABSTRACT: A gridded data set with the size of 0.05 degree resolution (approximately 5.5km) which representing ground observations of daily rainfall and temperature of Peninsular Malaysia has been created over 1975-2006. The integration and processing of the variety of data sources and data assessment is also presented. The 32-year period of the daily gridded rainfall and temperature data set were assessed to see how the daily mean rainfall and temperature have changed over time and space. Northeast monsoon (NEM) contributes more rainfall over the country compare to southwest monsoon (SWM). The rainfall trend during NEM is found significantly increased at the 95% confidence level (7mm/season/year), meanwhile SWM rainfall does not pose any significant trends. Both NEM and SWM temperature trends show significant increasing trends at 95% confidence level at 0.32°C/decade and 0.31°C/decade, respectively over the 32-year period. A drastic increased of mean temperature (1.20°C) was found in Klang Valley over the 20-year period. The mean decadal temperature was found consistently decrease as it approached the northern, east coast and southern part of the country.

Keywords: Rainfall trend, temperature trend, gridded data set, Peninsular Malaysia

1. INTRODUCTION

An accurate and complete hydrometeorological data source plays an important role in large scale models to simulate hydrological land surface conditions, especially in related to climate and water resources studies. Though the nature of each individual model varies due to the applied physics, the basic requirement is to acquire a complete data sources in terms of space and time. In many flood modeling applications, the simulations are dependent upon the availability of station data and the distance of the nearest station [1]. Because of the limited observation coverage, the selected nearest station data could be a few hundred kilometres away from the research area. The spatial distribution of stations or stations density in observation network and erroneous data in temporal aspects are the key problems for obtaining good hydrometeorological data within the study area [2]. Besides, the differences in customary observation practice, e.g. for rainfall measurements between countries [3,4], or differences in measurement methods or report timing between variables also limits the research works, especially affecting the modelling simulations (calibrations and verifications).

A current trend in achieving better hydrological modelling results, is turning to remote sensing as a possible means for quantifying e.g. variable rainfall as an input in the hydrological models as explained by many researchers [5-8], particularly in areas with few rain gauges. However, few shortcomings of satellite data were identified e.g. short records of data, biases and discontinuities due to the disruptions of instruments or nature and frequent changes of algorithms causing lengthy calibration tasks [9,10]. Although the calibration methods are improved from time to time [10-12], gauge-based data remains as a primary data source for any hydrological and atmospheric models.

To overcome the spatial and temporal problems with the ground observations, gridded data sets become essential means for applications in hydrological research and climate studies. For examples, global and regional studies in hydrological cycles [1,13-16], water resources assessment in large river basin [15,17], hydrological extremes [18,19], flood forecasting [20,21], climate variability [22,23], climate change scenario projections [24,25] and high resolution climate model evaluation [26,27].

Motivated by the availability of long records and good quality of daily hydrometeorological gridded data set in Peninsular Malaysia [28], this study focuses on the trend assessment of the daily rainfall and temperature data set between 99.5°E to 104.5°E and 1°N to 7°N, with a 0.05° resolution. The study also covers the spatial and temporal variability analysis of rainfall and temperature in Peninsular Malaysia.
2. DATA COLLECTION

2.1 Data Sources and Quality Control

The daily rainfall and temperature observation data used in this study has been extended from the previous study by Wong [28] to 2006. The data sources are based on the ground observations data of Malaysian Department of Irrigation and Drainage (DID), the publicly accessible Global Summary of the Day (GSOD) archive compiled from reports transmitted over the Global Telecommunication System of the World Meteorological Organisation (WMO-GTS) at the National Climatic Data Center of the National Oceanic and Atmospheric Administration (NOAA/NCDC), the Global Energy and Water Balance Experiment (GEWEX) research programme Asian Monsoon Experiment (GAME) archive, and the Malaysian Meteorological Department (MMD). Table 1 shows the data sources for the period 1975-2006.

<table>
<thead>
<tr>
<th>Data sources</th>
<th>Variable</th>
<th>Duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>DID</td>
<td>Rainfall</td>
<td>1975-2006</td>
</tr>
<tr>
<td>MMD</td>
<td>Rainfall</td>
<td>2002-2006</td>
</tr>
<tr>
<td>GAME*</td>
<td>Temperature</td>
<td>1997-2002</td>
</tr>
<tr>
<td>WMO-GTS**</td>
<td>Temperature</td>
<td>1975-2006</td>
</tr>
</tbody>
</table>

* from GEWEX Asian Monsoon Experiment (GAME)
** retrieved from NOAA/NCDC

Fig. 1 shows the interpolation domain and distribution of the stations. For the purpose of interpolation, the station data includes several WMO-GTS stations in Thailand and Sumatra, which are located just outside of the grid as shown in Fig. 1.

The data quality control procedure follow that described by Wong [28]. Three levels of screening and filtering of data were conducted in order to obtain the most reliable and consistent data sources. The details were highlighted in gridded data set development by Wong [28]. The long-term MMD statistics become reference for the quality monitoring purpose. The quality control of WMO-GTS data was based on the quality flags in the data sources files. There are no additional quality checks on MMD and GAME data, assuming the data were checked by MMD.

Fig. 2 presents the number of available daily observations from the individual data sources and their combination for the period from 1975 to 2006. The total number of rainfall observations is relatively stable, except the lack of WMO-GTS observations during short periods. The combination of three datasets has an average of 122 daily observations with a maximum of 167 for 2001 and 2003. Meanwhile, the average numbers of daily observations are 32 for temperature. The temperature records show a very small number of days were the number of observations significantly drops.

2.2 Terrain Data

Static data, i.e. terrain characteristics are constant over time. It is impossible to develop or operate large-scale modeling without such data. Elevation data have been derived from the Shuttle Radar Topography (STRM 30) 1 km version 2 data set [29], and were averaged into the output grid resolution.
The dynamic terrain data are associated with the land cover properties in the study area. Land cover type data were obtained from MODIS MOD12Q1 annual V004 1 km data products for the period 2001-2004, using the International Geosphere-Biosphere Programme (IGBP) biome classification. The land cover data are used to obtain the effective momentum roughness length for wind speed interpolation by aggregating the roughness length values of all cover types within a grid box using a cover type fraction-weighted averaging scheme [30]. The relations between land cover types and roughness length are based on [31] and [32].

3. METHODOLOGY

3.1 Data Interpolation

The interpolation scheme is based on an adaption of Shepard’s [33] angular distance weighting (ADW) procedure covers 32 years (1975 – 2006) period. Details description of the data set and gridding procedure are given in Wong [28]. In order to account for orography, interpolation of temperature are carried out on sea level equivalent values and then converted to grid box values using the derived DEM. Temperature is related to elevation, the temperature lapse rate formulations applied for the U.S. standard atmosphere [34] is used. There is no additional processing is carried out for rainfall, which is interpolated directly from point observation values. Since very few stations are available at higher altitude, it is not possible to identity clear orography effect on rainfall especially in daily time series.

The interpolation method requires an understanding of the spatial correlation structure of the station data as explained by Wong [28]. The inter-station correlations were investigated to determine the distances over which observed climate variables are related. This allows defining the distance weighting. The Pearson correlation coefficients between all pairs of stations for the variables are interpolated and plotting against distance. The spatial relationships between stations vary depending on the density of station coverage. In order to obtain daily spatial correlation coefficients for all pairs of rainfall and temperature stations, they were binned into separation distance intervals. In each bin, the mean and standard deviation of the correlation coefficient was calculated and an exponential curve was fitted through the mean distance-correlation points. The correlation decay distance (CDD) then equals to the distance where the mean correlation coefficient drops to 1/e.

In order to provide a measure of the expected accuracy, a jackknife cross-validation analysis has been carried out by repeated deletion and replacement of individual stations followed by a comparison of the grid box values with those of the deleted station. Overall daily error statistics for the entire area are computed in terms of mean absolute error (MAE) and bias. The results for temperature fields showed MAE is 0.4 °C with a bias of 0.1 °C. For rainfall a comparison at jackknifed grid box level resulted in a MAE of 5.2 mm/d and a bias of −0.1 mm/d (−1.9%). These results show that interpolation on a daily basis is significantly more error-prone that for the long-term average. On short time scales, the largely convective rainfall exhibits a large amount of spatial variability over short distances. In some cases, multiple rainfall stations located in the same grid box and that are averaged by the interpolation scheme show distinct differences in daily rainfall.

3.2 Trend detection

The annual and monsoon time series of rainfall of the climatic regions were analyzed using the Mann-Kendall nonparametric trend detection test [35,36]. The Mann-Kendall method has been widely used and tested as an effective method to evaluate a presence of a statistically significant trend in climatological and hydrological time series [17,37-39]. In the trend test, the null hypothesis \( H_0 \) is that there is no trend in the series (the data are independent and identically distributed). The alternative hypothesis \( H_1 \) is that a trend exists in the data. The Kendall \( S \)-statistic is obtained from comparison between all possible \((x, y)\) pairs of data and is given by

\[
S = \sum_{i=1}^{n} \sum_{j=i+1}^{n} \text{sgn}(x_i - x_j) \text{sgn}(y_i - y_j) \tag{1}
\]

where, the sign function is defined (for any variable \( u \)) as

\[
\text{sgn}(u) = \begin{cases} 
-1 & \text{if } u < 0 \\
0 & \text{if } u = 0 \\
+1 & \text{if } u > 0
\end{cases}
\]

Under the null hypothesis, the statistic \( S \) is approximately normally distributed with zero mean and variance are given by

\[
\text{Var}(S) = \frac{n(n-1)(2n+5) - \sum t_i(i-1)(2i+5)}{18} \tag{3}
\]

where \( t_i \) is the number of ties of extent \( i \) in either the \( x \) or \( y \) data. The summation term in the numerator is used only if the data series contains tied values. The standardized test statistic \( Z \) is obtained as

\[
Z = \begin{cases} 
\frac{S + 1}{\sqrt{\text{Var}(S)}} & \text{if } S < 0 \\
0 & \text{if } S = 0 \\
\frac{S - 1}{\sqrt{\text{Var}(S)}} & \text{if } S > 0
\end{cases}
\]

67
The test statistic $Z$ is used to measure the significance of a trend. In a two sided test, $H_0$ should be accepted if $|Z|$ is greater than $Z_{0.025}$, where $\alpha$ represents the chosen significance level (e.g. 5\% with $Z_{0.025} = 1.96$) then the null hypothesis is rejected implying that the trend is significant. A positive $Z$ value indicates an upward trend, whereas a negative $Z$ value indicates a downward trend. If a significant trend is present, the average rate of increase or decrease can be obtained from the slope of a simple linear regression.

4. DATA SET EVALUATION AND ANALYSIS

4.1 Rainfall

An overview of rainfall distribution of Peninsular Malaysia as a whole is presented. The gridded data set at 0.05 degree resolution was used to analyze the monthly rainfall distribution in Peninsular Malaysia. The mean monthly areal rainfall data for entire Peninsular Malaysia is shown in Fig. 3.

![Fig. 3 Box and whisker plot of monthly areal mean rainfall (1975-2006) in Peninsular Malaysia. The asterisk denotes the mean value, the solid line is the median, the height of the box is the difference between the third and first quartiles (IQR). Any data observation which lies 1.5 IQR lower than the first quartile or 1.5 IQR higher than the third quartile can considered an outlier in the statistical sense, indicated by open circles.](image)

Fig. 3 Box and whisker plot of monthly areal mean rainfall (1975-2006) in Peninsular Malaysia. The asterisk denotes the mean value, the solid line is the median, the height of the box is the difference between the third and first quartiles (IQR). Any data observation which lies 1.5 IQR lower than the first quartile or 1.5 IQR higher than the third quartile can considered an outlier in the statistical sense, indicated by open circles.

The highest mean monthly rainfall of 311 mm is observed in November, equivalent to 13\% of the mean annual rainfall. The lowest mean monthly rainfall of 113 mm occurs in February, which contributes about 5\% to the mean annual rainfall. It is noted that during the NEM from November to March both the maximum and minimum monthly rainfall are observed. A relative high rainfall variation is observed during the NEM than the Southwest Monsoon (SWM), which occurs from May to September. The result agrees with the findings by Moten [40], where maximum rainfall is observed near the end of the year during the Northeast Monsoon (NEM). A secondary maximum is found during the intermonsoon months (April or May).

![Fig. 4 Peninsular Malaysia rainfall distribution (1975-2006).](image)

Fig. 4 Peninsular Malaysia rainfall distribution (1975-2006).

Fig. 4 shows the distribution of annual average rainfall in Peninsular Malaysia. It is noted that the rainfall in the east coast region of Peninsular Malaysia is mostly influenced by the NEM, particularly during November or December [41,42]. The NEM influence was even reported to have more than 70\% rainfall contribution over most parts of the Peninsular [40]. However, there is only 44\% of annual rainfall throughout 32-year analysis occurred during the NEM. In this study, both monsoons contribute 81\% of the total annual rainfall over the country as shown in Table 2. There is 37\% of annual rainfall which occurred during SWM period. During the NEM, the dry northeasterly wind becomes wet during the passage over the South China Sea. The interaction between with the land along the east coast area creates deep convection clouds and rainfall [43].

The inland region received a relatively smaller amount of mean annual rainfall, i.e. 1950 mm/year as shown in Fig. 4. The reduction of rainfall amount in the inland is due to Titiwangsa mountain range forms the backbone of the Peninsula, from southern Thailand running approximately south-southeast over a distance of 480 km and separating the eastern part from the western part [41]. It appears to block the westward progression of the climatic system and therefore inhibits excessive rainfall over the inland areas [44]. The rainfall produced in this region is mainly due to local convection caused by intense heating of the land surface [45].
Table 2: Monsoons rainfall contributions in Peninsular Malaysia.

<table>
<thead>
<tr>
<th>Rainfall</th>
<th>NEM (Nov-Mar)</th>
<th>SWM (May-Sep)</th>
<th>Total Monsoons</th>
</tr>
</thead>
<tbody>
<tr>
<td>Annual Rainfall (mm)</td>
<td>mm</td>
<td>%</td>
<td>mm</td>
</tr>
<tr>
<td>2343</td>
<td>1034</td>
<td>44</td>
<td>864</td>
</tr>
</tbody>
</table>

4.2 Rainfall Trends

The annual and monsoons rainfall trends of Peninsular Malaysia are presented in Fig. 5. The mean annual rainfall shows minor increase trend (5 mm/year) over 32-year period, however it does not pose any significant increasing trends at the 95% confidence level under Mann-Kendall analysis.

For the mean NEM rainfall in the Peninsula, increasing trends of rainfall are significantly found at the 95% confidence level (7mm/season/year). It was probably due to the accumulation of increasing trend of other months though they are not significant at 95% confidence level. During SWM monsoon, Peninsular Malaysia shows −1.7 mm/season/year of rainfall declining but it does not pose significant trends at the 95% confidence level. Generally, the trends of mean annual and monsoons rainfall are probably governed by the topographic characteristics and local or regional conditions that have an effect on the interannual rainfall variability, which is superimposed on the large-scale weather conditions.

4.3 Temperature Trends

Fig. 6 shows the annual mean temperature trend of Peninsular Malaysia for 1975-2006. Significant increasing temperature trend at the 95% confidence level was found at about 0.32°C/decade for annual temperature in Peninsular Malaysia. This warming trend is comparable to those reported by Tangang et al. [46] where their results showed that the warming trend of between 2.7–4.0 °C/100 years. The warming trend at 0.30°C/decade agrees with the simulated temperature increase range by MMD [47] and NAHRIM [23] where the country will be warmer by about 2°C in the next 50 years. Both NEM and SWM temperature trends show significant increasing temperature trends at the 95% confidence level at 0.32°C/decade and 0.31°C/decade, respectively over the 32-year period (as shown in Fig. 6(b) and 6(c)).

4.4 Temperature Spatial Variation

To analyze the spatial variation of mean temperatures of Peninsular Malaysia, the 32-year of mean temperature gridded data was divided into three decades, i.e. 1975-1984, 1985-1994 and 1995-2006. The difference between the earlier and later decade was calculated to explore the spatial variation of mean monthly temperatures of Peninsular Malaysia as shown in Fig. 7.
Fig. 7: Decadal temperature (a) differences between baseline (1975-1984) and intermediate decade (1985-1994) (b) differences between intermediate (1985-1994) and most recent decade (1995-2006) of Peninsular Malaysia.

The initial decade (1975-1984) was used as the baseline to compare with the intermediate decade (1985-1994) as shown in Fig. 7(a). It is noted that the inland region (especially Klang Valley) encountered higher degree of temperature between 0.35°C/decade to 0.5°C/decade. It dispersed towards northern and southern part of Peninsular Malaysia. The most northern part of the country has minima of differences mean temperature (less than 0.05°C/decade) in terms of decade-to-decade comparison. Meanwhile, the southern of the country was at the range of 0.3°C/decade - 0.35°C/decade changes over 10 years period. Fig. 7(b) shows the difference of mean temperature between the intermediate decade (1985-1994) and the most recent decade (1995-2006). A drastic increased of mean temperature (more than 0.6°C/decade) was noted in Klang Valley and west coast of Peninsular Malaysia.

The mean temperature differences consistently decrease as it approached the northern, east coast and southern part of the country. The northern and southern part shows the approximately 0.35°C/decade or lesser of mean temperature differences. Majority of country area has an increased of mean temperature in the range of 0.35°C/decade between 0.6°C/decade over the latest decade-to-decade comparison. The mean temperature of Klang Valley areas was approximately 1.2°C warmer in 20-year period. This may be caused by the urbanization and industrialization in the country [48]. For other parts of the country, the mean temperature is generally on significant increasing trend as discussed earlier. The reason for the rise of warming rates in the recent decade is the higher frequency of El Niño events [46,47].

5. CONCLUSION

A high resolution, 0.05 degree gridded daily hydrometeorological data set was used to assess the rainfall and temperature variability in Peninsular Malaysia. The rainfall distribution of Peninsular Malaysia is generally influenced by two monsoons season, i.e. NEM and SWM. NEM has higher impact over the Study region compared to SWM. As most of the important rainfall trends may not be identified at the coarse time resolutions. A finer temporal scale is needed to investigate the rainfall trends in the future study.

Significant increasing trends (0.32°C/decade) at the 95% confidence level have been found in the mean temperature for the whole country. It was in line with the finding by other researchers [23,46,47]. The mean temperature of the country is generally spatially on increasing mode over the 32-year period. Klang Valley and west coast of Peninsular Malaysia encounter higher increase of mean temperature (the range of 0.95°C to 1.2°C) between 1975-1984 and 1994-2006. In contrast, the northern and southern parts of the country were on slow increasing pace (less than 0.35°C/decade) compared to other regions.

Future studies are likely to include the delineation of climatic regions using statistical analysis over the daily gridded data set in Peninsular Malaysia. It will be useful to further explore the regional characteristics of rainfall and temperature distribution through trend analysis and spatial rainfall and temperature variability analysis on different temporal resolution for different climatic regions in the country. It not only provides better understanding over the spatial and temporal climatology characteristics of Peninsular Malaysia, but also serve as the future fundamental for flood/hydrological research development in the country.

6. ACKNOWLEDGEMENTS

The authors thank the Department of Irrigation and Drainage Malaysia (DID), Malaysia Meteorological Department (MMD), and NOAA/NCDC for providing the data.

7. REFERENCES


[26] Masson, V.; Champeaux, J.-L.; Chauvin, F.; Meriguet, C.; Lacaze, R. A global database of


