ASSESSMENT OF MODIS LAND SURFACE TEMPERATURE DATA AGAINST GROUND METEOROLOGICAL DATA FOR THE WEST COAST OF PENINSULAR MALAYSIA

Nurul Iman Saiful Bahari (1), Farrah Melissa Muharam (2), Zed Zulkafli (3), Norida Mazlan (1)(2), Nor Azura Husin (4)

[1]Universiti Putra Malaysia, Institute of Tropical Agriculture and Food Security (ITAFoS), UPM Serdang, 43400, Malaysia
[2]Universiti Putra Malaysia, Faculty of Agriculture, UPM Serdang, 43400, Malaysia
[3] Universiti Putra Malaysia, Faculty of Engineering, UPM Serdang, 43400, Malaysia
[4] Universiti Putra Malaysia, Faculty of Computer Science and Information Technology, UPM Serdang, 43400, Malaysia
Email: nuruliman.sf@gmail.com; farrah.m@upm.edu.my; zeddiyana@upm.edu.my; noridamz@upm.edu.my; n_azura@upm.edu.my

KEYWORDS: Land surface temperature, air temperature, MODIS MOD11A1, tropical region.

ABSTRACT

Land surface temperature is one of the remote sensing products extracted from MODIS satellites. MODIS MOD11A1 provides land surface temperature data (T_s) on a daily basis at 1km spatial resolution. Although the T_s are freely available, the data suffers from missing values and unknown accuracy especially for tropical region such as Peninsular Malaysia. The weather in the tropics is mainly characterised with high rainfall intensity, high cloud cover and moderate temperature. The T_s for tropical regions is highly contaminated with cloud cover resulting in most part of the data being discarded even before they can be made ready for end-users. Therefore, this study was conducted with the objective to assess the accuracy of T_s in comparison to ground air temperature data (T_a) acquired from network of meteorological stations across west coast of Peninsular Malaysia. Daily T_a from 26 stations distributed were compared statistically against MODIS MOD11A1 T_s from 2012 to 2016. The MODIS-extracted T_s were compared with the pixel where the T_a from the ground meteorological stations were obtained. The performance of T_s in comparison to T_a analysis were quantified using measures such as root mean square error (RMSE) and correlation coefficient (r). The results illustrated were as follows: RMSE= 1.7 °C to 5.6 °C, and r = 0.1 to 0.6. The slightly high RMSE and low r values indicated that bias correction is necessary prior to the usage of MODIS MOD11A1 T_s.

1. INTRODUCTION

Air temperature data (T_a) is a useful information for various usages such as for monitoring of temperature trend (Tangang, Juneng and Ahmad 2007), global warming (Wai et al. 2005), crop yield modelling (Abraha and Savage 2008) and urban planning (Svensson and Eliasson 2002). T_a is a continuous, reliable and highly accurate weather measurements. Conventionally, T_a is automatically measured by a thermometer placed at a weather station, and transferred in real time to a computer or uploaded to a website. T_a is provided in a point form and therefore, only covers a limited range of area. Since T_a is only available for limited areas, T_a is unable to portray the heterogeneity of temperature over large areas. Hence, many studies have been carried out in attempt to interpolate T_a values. However, the inconsistent distance and poorly distributed location between each retrieval station affected the accuracy of the interpolated T_a (Wu and Li 2013). These limitations could be overcome with remotely sensed land surface temperature data that can provide estimated skin temperature data in a spatially wide area with high temporal resolution.

The launch of Moderate Resolution Imaging Spectroradiometer (MODIS) sensors on board of Terra satellite in 1999 has enabled remotely sensed day-time weather data to be retrieved. The MOD11A1 is one of the MODIS products and provides daily information about T_s at regional scale. T_s is measured from the radioactive skin temperature of mix elements of targeted earth's land surface calculated from solar radiation (Norman and Becker 1995). The elements of earth's land surface could vary from vegetated, agricultural, residential, forest, industrial and abandoned areas. T_s has been a great input in assisting many users to gain estimated temperature values for application of various fields such as hydrology (Parinussa et al. 2016), agriculture (Huang et al. 2015), ecology (Zhang et al. 2010), drought (Huang et al. 2017), evapotranspiration (Du et al. 2013), soil moisture estimation (Hosseini and Saradjian 2011), and more.

Although remotely sensed T_s can provide high spatial and temporal data, it suffers from data uncertainty and should be assessed for accuracy. This is because the quality of the day-time T_s is highly dependent on the radiant temperature of the land surface elements and atmospheric condition of the particular day (Vancutsem et al. 2010). Many studies were carried out to determine the accuracy of daily T_s acquired from MODIS. Values from T_s and T_a were compared statistically to identify their relationship. Lu et al. (2010) conducted a comparison between day-time daily T_s with in situ data for arid area in Northwest China. The finding of the study disclosed that the relationship between both data varied accordingly to seasonality: winter, spring, summer, or autumn and weather conditions: rainy or dry with RMSE between 2.39ºC to 3.05 ºC. Furthermore, Da Silva et al. (2015) performed a study to identify the linear relationship between the accumulated monthly MODIS derived T_s in comparison to the accumulated monthly T_a, in continental and near sea regions of Portugal. It was discovered that both data were highly linearly related to each other with R^2 more than 0.98. In addition, Zhang et al. (2010) reported that the accuracy between MODIS derived T_s to maximum T_a was highly affected when there were partial cloud covered pixels. The MAE values were between 2.8 ºC to 4.1 ºC.

In the west coast of Peninsular Malaysia region, MODIS derived T_s are highly affected by cloud covers leading to most of the data were discarded before usage. Besides, the undetected partial cloud cover may result in lower T_s readings. Moreover, the interaction between solar radiations with particulate matter in the atmosphere and other atmospheric condition such as wind speed and precipitation may compromise the T_s readings. Thus, by considering all the factors contributing to the uncertainties of the MODIS LST product, it is necessary to perform a data assessment before further application. Unfortunately, no assessment has yet to be done to evaluate the MODIS LST product for Peninsular Malaysia scene, and therefore, this study aim to fill this gap.

2. DATA AND METHODOLOGY

2.1 Study Area

The Peninsular Malaysia, or also known as West Malaysia, is located at the south end of Asia continental plate (Figure 1). It is geographically located between 1°N to 7°N and 99°E to 105°E. Its weather is characterised as hot and humid tropics with temperature ranging between 23°C to 35°C, daily humidity levels exceeding 80%, and high rainfall capacity with high cloud cover throughout the year. There are two monsoonal winds characterising two rainy seasons known as the Southwest and Northeast monsoons that outset from June to September and November to March, respectively (Masseran and Razali 2016). Mountain ranges known as Titiwangsa mountain range acts as the backbone of Peninsular Malaysia separating between the eastern and western coastal plains. The western coastal plain is characterised by states and federal territories that are facing the Straits of Malacca. The west coast region is further categorised into northern region: Perlis, Kedah, Penang and Perak, central region: Selangor and federal territories of Kuala Lumpur and Putrajaya and southern region: Negeri Sembilan, Melaka and Johor.

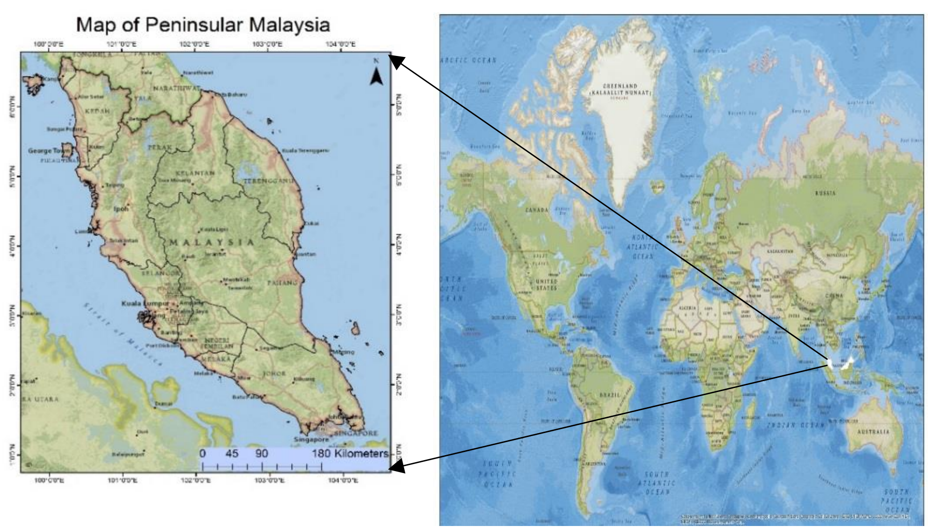

Figure 1: Map and location of Peninsular Malaysia (ESRI Basemap)

2.2 Data Collection

Ground measurement: T_a was obtained from 26 automatic weather stations (AWS) that belong to Department of Environment (DOE) for the period of 5 years between 2012 and 2016. The weather stations measure daily maximum, minimum and mean T_a. Maximum T_a indicate the highest temperature, while minimum T_a indicate the lowest temperature and mean T_a indicate the average temperature for a particular day. Nevertheless, only the maximum T_a was used in this study because of its compatibility with MOD11A1 which was observed during the day at about 10:30 a.m. local time. The AWS are set up 2m above ground and located at government facilities such as schools, police stations, hospitals and government offices. The AWS must be strategically located away from activities involving land use and land cover changes to ensure that there is no drastic changes in the temperature. Figure 2 shows the locations of the AWS in west coast of Peninsular Malaysia with information of average maximum T_a from 2012 to 2016. It is noticeable that there are 2 AWS located at Langkawi and Penang Islands. The AWS on the land are more concentrated around the city of Kuala Lumpur, while the others are scattered along the west coast.

MODIS LST Product: The MOD11A1 was obtained from https://search.earthdata.nasa.gov/search with 1km spatial resolution in a 1200 by 1200 km grid. The data was retrieved daily from MODIS sensors on board Terra satellites providing daytime T_s at about 10.30 a.m. local time. The T_s covering 2 tiles for Peninsular Malaysia was extracted from horizontal line 27 to 28 and vertical line 08 for years 2003 to 2016.

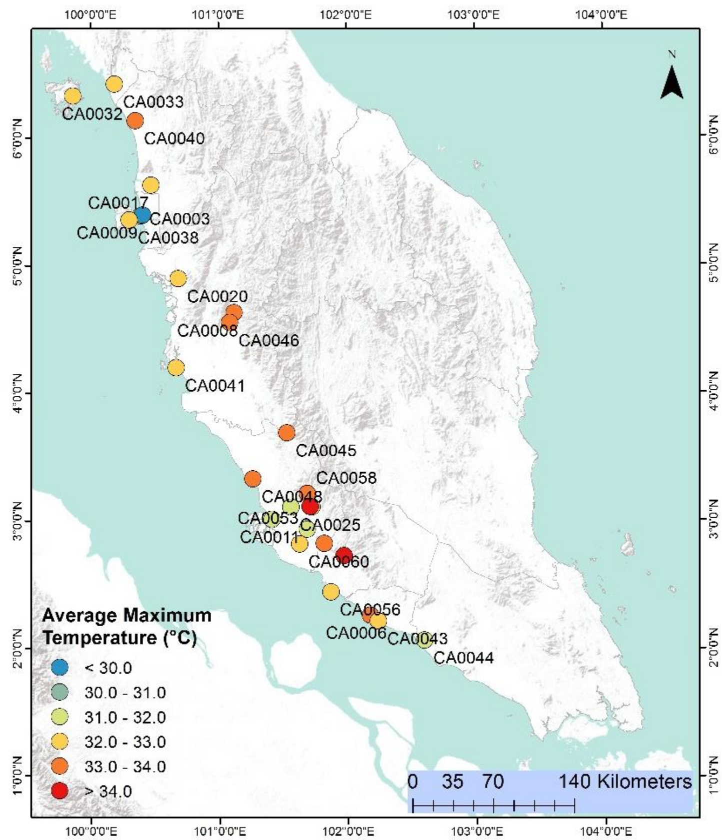

Figure 2: The average maximum of T_a for the AWS (ESRI Basemap)

2.3 Data Processing

Pre-processing of T_a: The AWS data was collected and transmitted automatically without human interference, and therefore, data screening process was necessary to evaluate the AWS data. The boxplot outlier analysis was conducted in order to identify extreme values in AWS data (Tukey 1977). The equation for lower and upper boundaries are given as in Equation 1 and 2.

$$\mathrm{lower}_{\mathrm{boundary}} = \mathrm{Q1} - 1.5(\mathrm{Q3} - \mathrm{Q1}) \tag{1}$$

$$\mathrm{upper}_{\mathrm{boundary}} = \mathrm{Q3} + 1.5(\mathrm{Q3} - \mathrm{Q1}) \tag{2}$$

Where Q1 and Q3 are equal to first and third quartile of the data.

Pre-processing of T_s: The downloaded MOD11A1 (h27v08 and h28v08) were projected and clipped to Peninsular Malaysia region. The absolute temperature value in Kelvin (K) were converted to degree Celsius (°C) as shown in Equation 3.

$$\mathrm{T_s} = 0.02\mathrm{T} - 273.15 \tag{3}$$

Where 0.02 is the scale factor, T is the absolute temperature in Kelvin and T_s is the land surface temperature in °C (Wan 2013). Cloud covered pixels were removed prior the end product released leading to missing data for most part of Peninsular Malaysia. Figure 3 illustrates the map of T_s distribution in Peninsular Malaysia for 23rd May 2017. The white patches indicate the missing data, mostly because of cloud cover. The T_s at the location of each AWS was then extracted. Further, the available pixels of T_s and T_a of the same locality were paired and fit for data analysis.

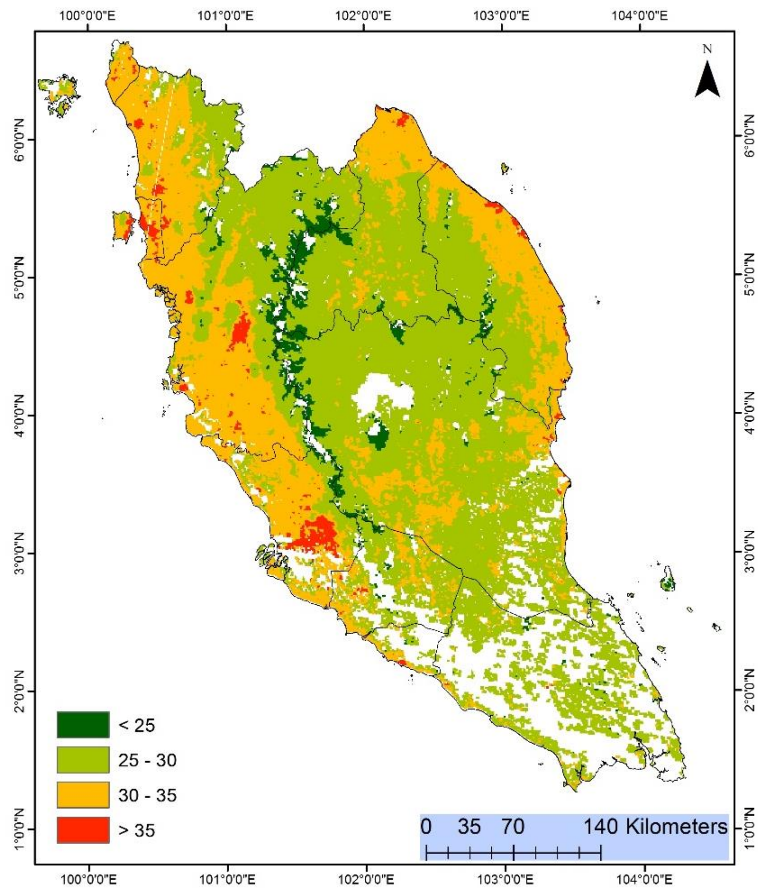

Figure 3: Map of T_s distribution in Peninsular Malaysia in degree Celsius for 23rd May 2007

2.4 Data Analysis

Evaluation Metrics: The T_s and T_a pixel-point pairs were compared to assess the variation of T_s from T_a using evaluation metrics such as root mean square error (RMSE), and correlation coefficient (r) (Benali et al. 2012). The equation for RMSE and r are as shown as Equation 4 and 5 respectively.

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(T_{si} - T_{ai})^2} \tag{4}$$

Where, $T_{si}= T_s$, $T_{ai}= T_a$ and N= sample size. Root mean square error (RMSE) is a measure to determine error between two variables comparing between simulated and observed value. A low RMSE indicates a good fit between two variables.

$$r = \frac{\sum_{i=1}^{N}(T_{si}-\overline{T}_s)(T_{ai}-\overline{T}_a)}{\sqrt{\sum_{i=1}^{N}(T_{si}-\overline{T}_s)\sum_{i=1}^{N}(T_{ai}-\overline{T}_a)^2}} \tag{5}$$

Where $\overline{T}_s$= average land surface temperature data and $\overline{T}_a$= average air temperature data. Correlation coefficient (r) measure strength and direction of relationship between both variables. The r value of 1 indicates a very strong positive correlation, 0 indicates no correlation and -1 indicates strong negative correlation.

3. RESULT AND DISCUSSION

Figure 4(a)-(b) illustrate the range of RMSE and r. Table 1 presents the performance for T_s in comparison to T_a in terms of evaluation metrics for all 26 stations. The RMSE were recorded between 1.7°C to 5.6°C with most of the RMSE scores range between 3°C to 4°C. The three highest RMSE values were recorded from Perai, Pulau Pinang (CA0003), USM, Pulau Pinang (CA0038) and Seberang Jaya, Pulau Pinang (CA0009) with RMSE scores of 5.6°C, 4.5°C and 4.4°C, respectively. Overall, the northern region depicted higher RMSE than the central and southern regions with associated average T_a of 32.0°C to 34.0°C, 31.0°C to 35.0°C, and 31.0°C to 33.0 °C, respectively.

The r values were reported to be very low with a range from 0.1 to 0.6 with majority of r range between 0.1 and 0.2. The lowest correlation between T_a and T_s with r = 0.1 was recorded from three stations namely, Presint 8, Putrajaya (CA0053), Port Dickson, Negeri Sembilan (CA0056) and Bukit Changgang, Selangor (CA0060). Nevertheless, the highest r was reported from Shah Alam, Selangor (CA0025) with r= 0.6.

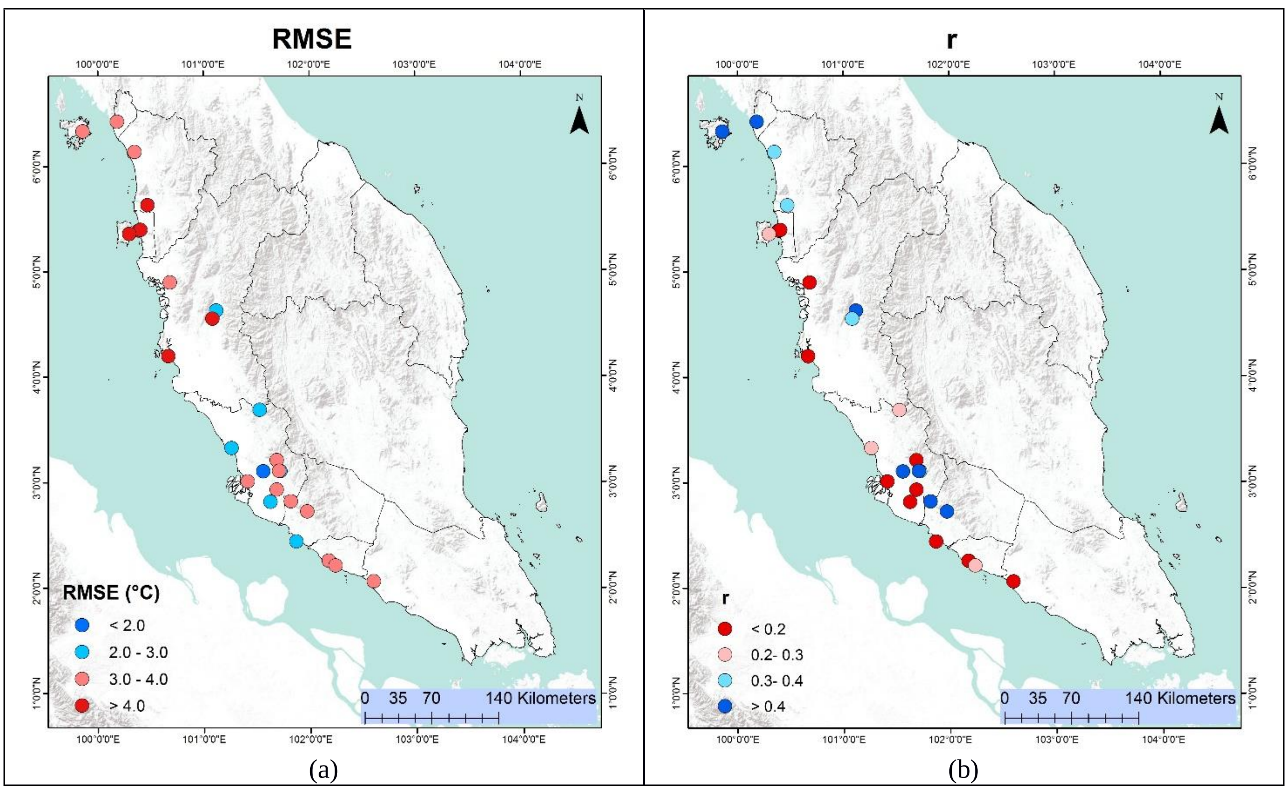

Figure 4: The performance of evaluation metrics for comparison between T_s as simulated data and T_a as observed data. (a) RMSE and (b) r. (ESRI Basemap)

Table 1: The performance of evaluation metrics for comparison between T_s as simulated data and T_a as observed data

No	Site State	Station ID	RMSE (°C)	r
1	Perai, Pulau Pinang	CA0003	5.6	0.2
2	Bukit Rambai, Melaka	CA0006	3.3	0.2
3	Ipoh, Perak	CA0008	2.9	0.5
4	Seberang Jaya, Pulau Pinang	CA0009	4.4	0.2
5	Nilai, Negeri Sembilan	CA0010	3.4	0.5
6	Klang, Selangor	CA0011	3.8	0.2
7	Petaling Jaya, Selangor	CA0016	3.1	0.5
8	Sungai Petani, Kedah	CA0017	4.3	0.4
9	Taiping, Perak	CA0020	3.6	0.2
10	Shah Alam, Selangor	CA0025	1.7	0.6
11	Langkawi, Kedah	CA0032	3.6	0.5
12	Kangar, Perlis	CA0033	3.3	0.5
13	USM, Pulau Pinang	CA0038	4.5	0.3
14	Alor Setar, Kedah	CA0040	3.2	0.4
15	Seri Manjung, Perak	CA0041	4.1	0.2
16	Bandaraya Melaka, Melaka	CA0043	4.0	0.3
17	Muar, Johor	CA0044	3.8	0.2
18	Tanjung Malim, Perak	CA0045	2.7	0.3
19	Pegoh, Perak	CA0046	4.2	0.4
20	Seremban, Negeri Sembilan	CA0047	3.4	0.5
21	Kuala Selangor, Selangor	CA0048	2.8	0.3
22	Presint 8, Putrajaya	CA0053	3.9	0.1
23	Cheras, Kuala Lumpur	CA0054	2.4	0.4
24	Port Dickson, Negeri Sembilan	CA0056	2.5	0.1
25	Batu Muda, Kuala Lumpur	CA0058	4.0	0.2
26	Bukit Changgang, Selangor	CA0060	2.9	0.1

The high variability in relationship between T_s and T_a is expected because in nature they are both measuring different nature of temperature. T_s is a result of radiative temperature over a large area of a ground surface with heterogeneous land cover and constantly changing depending on the atmosphere condition and other solar interaction, while T_a is a temperature measurement of a localized area in point form of a homogeneous area (Da Silva et al. 2015). Although T_s and T_a are different measurement of temperature in nature, Benali et al. (2012) and Fu et al. (2011) have reported that there was a linear relationship between these two temperatures. Nevertheless, the differences between these two temperatures can be contributed by mixed land and cloud covers that were not eliminated during preprocessing of the MOD11A1 product. This is in line with a finding from a study done by Langer et al. (2010), who found that pixels from T_s product that have cloud cover in it tended to measure T_s at 5 ºC to 10 ºC lower than the actual T_s value. Therefore, there is a need to calibrate the extreme values in T_s before usage.

4. CONCLUSION

An assessment has been carried out to compare MODIS derived T_s with ground measured T_a for west coast of Peninsular Malaysia scene for 2003 until 2016. It is found that T_s suffered from missing data especially along the mountain ranges due to the cloud cover. The evaluation metrics suggested that T_s is highly variable from T_a, possibly due the effect of undetected cloud associated with land cover and other atmospheric condition. This signifies the important to assess the MODIS derived T_s prior to the usage for Peninsular Malaysia scene and conduct bias correction prior to the use of MODIS T_s.

ACKNOWLEDGEMENTS

This work was financed by Higher Institution Center of Excellence (HICoE) research grant (HICoE-ITAFoS/2017/FC3, vote no. 6369105) by the Ministry of Higher Education (MoHE). The authors would like to acknowledge the air temperature data providers, namely Department of Environment, Ministry of Energy, Science, Technology, Environment and Climate Change. We thank the National Aeronautics and Space and Space Administration for providing the freely downloadable MODIS LST products.

REFERENCES

Abraha, M. G., and Savage, M. J., 2008. Comparison of estimates of daily solar radiation from air temperature range for application in crop simulations. Agricultural and Forest Meteorology, 148(3), pp.401-416.

Benali, A., Carvalho, A. C., Nunes, J. P., Carvalhais, N., and Santos, A., 2012. Estimating air surface temperature in Portugal using MODIS LST data. Remote Sensing of Environment, 124, pp. 108–121.

Da Silva, J. M., Damásio, C. V., Sousa, A. M., Bugalho, L., Pessanha, L., and Quaresma, P., 2015. Agriculture pest and disease risk maps considering MSG satellite data and land surface temperature. International Journal of Applied Earth Observation and Geoinformation, 38, pp. 40-50.

Du, J., Song, K., Wang, Z., Zhang, B., and Liu, D., 2013. Evapotranspiration estimation based on MODIS products and surface energy balance algorithms for land (SEBAL) model in Sanjiang Plain, Northeast China. Chinese geographical science, 23(1), pp. 73-91.

Duan, S. B., Li, Z. L., Li, H., Göttsche, F. M., Wu, H., Zhao, W., Leng, P., Zhang, X. and Coll, C., 2019. Validation of Collection 6 MODIS land surface temperature product using in situ measurements. Remote sensing of environment, 225, pp.16-29.

Esri., 2019. "Topographic" [basemap]. Scale Not Given. "National Geographic World Map". http://www.arcgis.com/home/webmap/viewer.html?useExisting=1&layers=30e5fe3149c34df1ba922e6f5bbf808f. (August 5, 2019).

Fu, G., Shen, Z., Zhang, X., Shi, P., Zhang, Y. and Wu, J., 2011. Estimating air temperature of an alpine meadow on the Northern Tibetan Plateau using MODIS land surface temperature. Acta Ecologica Sinica, 31(1), pp.8-13.

Hosseini, M., and Saradjian, M. R., 2011. Multi index-based soil moisture estimation using MODIS images. International journal of remote sensing, 32(21), pp. 6799-6809.

Huang, R., Zhang, C., Huang, J., Zhu, D., Wang, L., and Liu, J., 2015. Mapping of daily mean air temperature in agricultural regions using daytime and night time land surface temperatures derived from TERRA and AQUA MODIS data. Remote Sensing, 7(7), pp. 8728-8756.

Huang, D., Liu, J. F., Wang, W. Z., and Peng, A. B., 2017. The evolution analysis of seasonal drought in the upper and middle reaches of Huai River basin based on two different types of drought index. In: IOP Conference Series: Earth and Environmental Science, 59(1), pp. 012042.

Langer, M., Westermann, S., and Boike, J., 2010. Spatial and temporal variations of summer surface temperatures of wet polygonal tundra in Siberia-implications for MODIS LST based permafrost monitoring. Remote Sensing of Environment, 114(9), pp. 2059-2069.

Lu, L., Zhang, T., Wang, T., and Zhou, X., 2018. Evaluation of Collection-6 MODIS land surface temperature product using multi-year ground measurements in an arid area of Northwest China. Remote sensing, 10(11), pp. 1852.

Masseran, N., and Razali, A. M., 2016. Modelling the wind direction behaviours during the monsoon seasons in Peninsular Malaysia. Renewable and Sustainable Energy Reviews, 56, pp. 1419-1430.

Norman, J. M., and Becker, F., 1995. Terminology in thermal infrared remote sensing of natural surfaces. Agricultural and Forest Meteorology, 77(3-4), pp. 153-166.

Parinussa, R., Lakshmi, V., Johnson, F., and Sharma, A., 2016. Comparing and combining remotely sensed land surface temperature products for improved hydrological applications. Remote Sensing, 8(2), pp. 162.

Svensson, M. K., and Eliasson, I., 2002. Diurnal air temperatures in built-up areas in relation to urban planning. Landscape and urban planning, 61(1), pp. 37-54.

Tangang, F. T., Juneng, L. and Ahmad, S., 2007. Trend and inter-annual variability of temperature in Malaysia: 1961–2002. Theoretical and Applied Climatology, 89(3-4), pp.127-141.

Tukey, J. W., 1977. Exploratory data analysis. Addison-Wesley. Reading, PA.

Vancutsem, C., Ceccato, P., Dinku, T., and Connor, S. J., 2010. Evaluation of MODIS land surface temperature data to estimate air temperature in different ecosystems over Africa. Remote Sensing of Environment, 114(2), pp.449-465.

Wai, N. M., Camerlengo, A., Khairi, A., and Wahab, A., 2005. A study of global warming in Malaysia. Jurnal Teknologi, 42(F), Jun. 2005, pp. 1–10.

Wan, Z., 2013. Collection-6 MODIS land surface temperature products users' guide. ERI, University of California, Santa Barbara.

Wu, T., and Li, Y., 2013. Spatial interpolation of temperature in the United States using residual kriging. Applied Geography, 44, pp. 112-120.

Zhang, Y., Yang, S., Ouyang, W., Zeng, H., and Cai, M., 2010. Applying multi-source remote sensing data on estimating ecological water requirement of grassland in ungauged region. In: Procedia Environmental Sciences, 2, pp. 953-963.