Mitigation & Identification for Local Aridity, Based of Vegetation Indices Combined with Spatial Statistics & Clustering K Means

Sri Yulianto Joko Praetyo, Kristoko Dwi Hartomo, Bistok Hasiholan Simanjuntak, Dian Widiyanto Candra
Satya Wacana Christian University

Email: sri.yulianto@uksw.edu, kristoko@uksw.edu, dian.chandra@uksw.edu

Abstract. This research aims to develop new methods of mapping aridity risk zones and their potential impacts on land fires using a combination of indices to identify land fires such as the CSI and NBR and the indices for NDVI and SAVI. The research location is in Gunung Merapi National Park (TNGM) and Gunung Merbabu National Park (TNGMb), in Central Java Province and Yogyakarta. The data used in this research is Landsat 8 OLI image year 2010 - 2018, DEM data from ASTER image in TNGM and TNGMb area using landsat 8 OLI image specification. The research was conducted in 3 stages such as pre-processing, image data extraction and post-processing. Global and Local Moran's Analysis on NDVI, SAVI, CSI and NBR vegetation indices data can be used as an indicator of aridity and potential land fires. The experiments show that the average is in class 4 including the moderate greenish classification. Moderate greenish is interpreted that the study area is overgrown with meadows, shrubs, barren, sandy, rocky areas and a low population of vegetation canopy that shows that the area is on the surrounding mountain peaks. The results of the analysis shows Positive Spatial Autocorrelation, the phenomenon of aridity has spatial connectivity between observed regions. Analysis of K Means on the high vegetation density conditions shows that the weight of the distance between the vegetation data to the centroid is shorter, therefore the data is concentrated on a region. In low vegetation density conditions, the weight of the distance between the data to the centroid is increasingly wide, therefore the data looks more widely distributed.

Keyword: aridity, satellite image, vegetation indices, spatial autocorrelation, Clustering K Mean
1. Introduction

Aridity is a disaster which is difficult to be detected and predicted accurately because its process takes place slowly and the impact is not directly recognized. It occurs in long term from months to years[1][2]. Aridity begins to be felt when agricultural land, rivers and springs dried up, degraded land and some areas experience land fires [3][4]. The severity of the aridity is affected by the rain duration, season patterns, ecological spatial patterns on the aridity affected areas, therefore the impact will vary from region to region [5][6][7]. The World Meteorological Organization (WMO), the current determination of aridity can be done using satellite images based on the calculation of vegetation indices which includes: Normalized Difference Vegetation Indices (NDVI), Normalized Difference Water Indices (NDWI), Enhanced Vegetation Indices (EVI), Vegetation Condition Indices (VCI), Temperature Condition Indices (TCI), and Vegetation Health Indices (VHI) [8][9][10][11][12]. NDVI represents the capability of photosynthesis, biomass, dominant plant species, biogeochemical processes at local and regional scale and a global model identified as greenish vegetation [13][14]. NDVI close to 1 indicates very densely packed vegetation, close to 0 indicating bare ground or very sparse vegetation [5]. A negative NDVI value indicates water bodies or urban areas. NDVI is used to monitor plant growth, land cover and biomass production capabilities [15][16]. The formula NDVI is:

\[
NDVI = \frac{\text{NIR} - \text{RED}}{\text{NIR} + \text{RED}}
\]

SAVI is vegetation indices that have been used to distinguish between post-fire areas and vegetated areas[17][18]. The L notation in SAVI equation is a constant notation that represents vegetation density with L = 0.25 if the vegetation density is high, L = 0.5 if the vegetation density is medium and L = 0.75 if the vegetation density is low [19][20]. The formula SAVI is:

\[
SAVI = (1 + L) \frac{\text{NIR} - \text{RED}}{\text{NIR} + \text{RED} + L}
\]

NBR is an indices used to distinguish the land that has been affected by fires or has not [21][22][23]. The formula NBR is:

\[
NBR = \frac{\text{NIR} - \text{SWIR}}{\text{NIR} + \text{SWIR}}
\]

EVI is effectively used to identify characteristics of areas which are not sensitive to NDVI such as distinguishing between water bodies with high moisture soils, biomass such as seasonal plant growth variations, vegetation variation and phenology, ranging from 1- to 1 but normal vegetation is at a value between 0.2 – 0.8 [11][10]. The formula EVI is:

\[
EVI = 2.5 \times \frac{\text{NIR} - \text{RED}}{\text{NIR} + \text{RED} + + L}
\]

Other aridity indicators are TCI, VCI and VHI, the vegetation indices used as indicator of aridity of vegetation in certain area, so these three indices are generally referred to as indices for aridity monitoring. The formula TCI, SAVI and VCI are:

\[
VCI = \frac{\text{NDVI}_{\min} - \text{NDVI}_{\max}}{\text{NDVI}_{\max} - \text{NDVI}_{\min}} \times 100
\]

\[
TCI = \frac{\text{LST}_{\max} - \text{LST}_{\min}}{\text{LST}_{\max} - \text{LST}_{\min}} \times 100
\]

\[
VHI = aVCI + (1 - a)TCI
\]

TCI is used to identify the occurrence of stress on vegetation caused by high temperatures and excess air humidity. VHI is used as health indicator of vegetation from a combination of TCI and VCI[16][24]. This study aims to develop new methods of mapping aridity risk zones and their potential impacts on land fires using a combination of indices to identify land fires, such as CSI and NBR and indices for the vegetation’s health such as NDVI, SAVI, TCI, VCI, and VHI.

2. Research Methods

The research location is in Gunung Merapi National Park (TNGM) and Gunung Merbabu National Park (TNGMb), covering an area of ± 6,410 hectares, located in Magelang, Boyolali and Klaten regencies of Central Java Province and Sleman regency of Province of Yogyakarta. The observation sites are in 234 villages, 17 sub-districts and overall are in 4 regencies. The location of the research in
position are 07°22' - 07°52' South Latitude, 110°15' - 110°37' East Longitude (TNGM) and 7°38' - 7°48' South Latitude and 110°32' - 110°48' East Longitude (TNGMb). The rainfall is min 902 and max 3627 mm/year. The data used in this research is Landsat 8 OLI image year 2010 - 2018, DEM data from ASTER image in TNGM and TNGMb area. The research was conducted in 3 stages such as initial data processing, image data extraction and post processing. The pre-processing of the image is divided into 3 procedures such as: Geometric correction, radiometric correction, and vector data correction. The image data extraction is done through 3 procedures such as: Extraction and making thematic map of vegetation indices such as CSI, NBR, NDVI, SAVI in all observation areas in 234 study area villages. Post processing analysis of vegetation indices data such as CSI, NBR, NDVI, SAVI using Spatial Statistic method and Clustering K-Mean method [25].

Figure 1. The observation location, there are 234 observation points in TNGM and TNGMb.

Moran's I analysis of vegetation indices data aims to determine spatially correlation patterns between observation points. Moran's I Global Equation is:

\[ I = \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} (x_i - \bar{x})(x_j - \bar{x})}{(\sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij}) \sum_{i=1}^{n} (x_i - \bar{x})^2} \]  

and \[ I_t = z_i \sum_j w_{ij} z_j \]

\( z_i \) is the value of \( z \) observation of each region \( i \), \( z_j \) is the deviation or standard deviation from the mean value, \( w_{ij} \) is the weight between observation region to \( i \) and to \( j \).

3. Result and Discussion

The identification and determination of aridity was done by using vegetation indices, namely NDVI and SAVI, and the identification of fire incident was determined using CSI and NBR vegetation indices. The regulation of classification NDVI values is, class 1 (-1 to -0.03) is land without vegetation, class 2 (-0.04 to 0.15) is very low vegetation, class 3 (0.16 to 0.25) is Low Vegetation, class 4 (0.26 to 0.35) is moderate vegetation, and class 5 (0.36 to 1.00) is High Vegetation[26]. NDVI, NBR, CSI and SAVI vegetation indices in the study area is in the observation period between year 2016 – 2018. Tables 1 and Figure 3 show that the average NDVI value in the study area is 0.5 that belongs to class 4, which is included in the moderate greenish classification. Moderate greenish (greenery) is interpreted that study areas are overgrown with meadows and shrubs as well as a low population of vegetable canopy. The maximum NDVI value is 2.0 which is interpreted that the study area is overgrown with vegetable canopy populations in good growth condition. This happens during the rainy season and only on conservation areas. The minimum value of NDVI in the study area is 0.11, which is represented by areas of moderate greenishness, barren, sandy and rocky areas, indicating areas around the peak of Mount Merapi and Mount Merbabu. This NDVI indices represents the condition of vegetation both in terms of quality such as the vegetation’s health condition and...
quantity such as the level of vegetation spread. The interference toward this reflectance is overcome by using the SAVI vegetation indices. The SAVI indices aims to reduce the effect of soil reflectance on indices results[11]. The average SAVI value is 0.38, the minimum is 0.42 and the maximum is 0.57. It is interpreted that there is moderate vegetation densities in the study area. The range of SAVI values is between -1 to 1, ranging from high vegetation canopy density to low vegetation canopy density. The average dNBR value is 0.01, at minimum 0.00 and maximum 0.18 which is interpreted as the ratio of the severity level of the region with relatively low fires, the study area experiences uncanopied vegetation growth and the probability of land fire in the study area is very low [18]. The average CSI value is 1.36, at minimum 1.55 and maximum 1.91 which is interpreted that in the study area there is no former land fires. Therefore, from all observations done in 2016 – 2018, there is no land fires detected by LANDSAT 8 sensors.

The average NDVI in the study area is 0.50 which is interpreted as a region with moderate and low greenishness, overgrown with grass and shrub vegetation. From the observed data of 18 periods, from 2016 to early 2018, the value of NDVI ≤ value 0,5 and close to 0,5 in the rainy season as can be seen in blue points (Figure 3). As an improvement in the accuracy of the NDVI analysis, a SAVI analysis was performed that eliminated the soil surface effect of LANDSAT sensor’s reflectance. The results show that the SAVI average is 0,38 which is interpreted as a region with medium to high aridity level. The SAVI values is on the range of ≥ 0,5 and it is relatively high at some observation points especially around the rainy season. This study shows that there is a high correlation between SAVI and NDVI, therefore the result of SAVI analysis and interpretation shows the same thing with NDVI such as there is no vegetation canopy in the study area and is overgrown with grass and shrubs vegetation type.

Table 1. Average, maximum and minimum values of vegetation indices in the study area.

| Vegetation Indices | Average | Min  | Max  |
|--------------------|---------|------|------|
| NDVI               | 0.50    | 0.11 | 2.09 |
| NBR                | 0.18    | 0.19 | 0.23 |
| dNBR               | 0.01    | 0.00 | 0.18 |
| CSI                | 1.36    | 1.55 | 1.91 |
| SAVI               | 0.38    | 0.42 | 0.57 |

SAVI is shown in Figure 3, on the orange colored points. SAVI is also influenced by local season patterns with an increased indices in the rainy season. The low vegetation does not have significant correlation to land fires, observing low CSI and NBR in the absence of former land fires (Figure 3). The visual analysis of aridity satellite imagery in the dry and rainy season in the study area shows significant difference in the greenish of the area, as shown in Figure 4. NBR and CSI are indices
which are sensitive to local seasonal patterns, evidenced by an increase in indices values during the rainy season. The spatial statistic analysis on vegetation indices data was aimed to observe the connectivity of aridity level between the observation areas and to identify aridity concentration pattern, outlier data and aridity potential distribution pattern in all study areas by applying indicator of vegetation indices such as NDVI, SAVI, CSI and NBR [19][21]. The spatial statistic analysis method used in this research is Moran's I and Getis Ord. The comparison of aridity pattern visualization of the study area of each vegetation indices and the results of Moran's indices analysis and Getis Ord indices analysis are presented in Table 2.

Table 2. Comparison of aridity pattern visualization of the study area of each vegetation indices

| Vegetation Indices | Global Moran's I | Local Moran's I | Interpretation |
|--------------------|------------------|-----------------|----------------|
|                    | Average Min Max  | Average Min Max |                |
| NDVI               | 0.618 0.709 0.479 | 0.544 -0.517 5.023 | Cluster        |
| SAVI               | 0.767 0.720 0.713 | 0.679 -0.887 6.891 | Cluster        |
| CSI                | 0.526 0.523 0.438 | 0.544 -0.517 5.023 | Cluster        |
| NBR                | 0.529 0.598 0.422 | 0.542 -0.517 4.771 | Cluster        |

The results of the analysis of the vegetation indices data indicate that all data meet the criteria of $I \geq E(I)$, therefore all the data is interpreted as clusters, or the data distribution is concentrated at a particular points. The concentrated data at a particular observation points indicates that not all observed areas are overgrown with vegetation. The area of concentrated vegetation based on Moran's I analysis is a moderate to high vegetation region based on the analysis of vegetation indices (SAVI and NDVI).
Figure 3. Analysis with the K-mean Clustering method to see the concentration of vegetation indices data during the rainy and dry seasons.

The spatial Autocorrelation analysis used in this study aims to see the clustering based on similarity of vegetation indices values in each observed area, measured using Moran's I function. As an effort to improve the accuracy of the system, an analysis using Clustering K-Mean method is done. It is a method used to classify data based on the weight of the distance between data observed by centroid. The K-Mean Clustering method approach in this study shows that the spatial spectrum groupings are more accurate and form consistent vegetation indices patterns for each data. Figure 4. shows the analysis of Clustering K-Mean vegetation indices in two seasons such as rainy and dry season. This analysis shows different patterns of different vegetation indices concentrations depending on the weight of the distance of the data with the centroid. In high vegetation density conditions, the weight of the distance between the vegetation data with the centroid is shorter, so that the data is seen to be concentrated on the region. In low vegetation density conditions, the weight of the distance between the data with the centroid is wider, so that the data looks more widely distributed. However, it appears that vegetation data groups are formed in smaller sizes with shorter spacing. This can be clearly seen by observing NDVI and SAVI in Figure 4. The NDVI and SAVI indices show the efficiency of photosynthesis process and potential productivity in all areas, which is marked in the image as a vegetation group. The dominance of the green colored data group shows the dominance of grids that characterizes as vegetation. The CSI and NBR indices show different patterns to the NDVI and SAVI. The dominance of the brown and red colored data groups indicates the dominance of grids that characterizes as non-vegetation surface area or low vegetation growth area and indicates the occurrence of aridity.

4. Conclusion
The Spatial Autocorrelation analysis, particularly Global Moran's and Local Moran's on NDVI, SAVI, CSI and NBR vegetation indices data can be used as an indicator of aridity and potential land fires. The experiments using LANDSAT 8 OLI Imagery in TNGM and TNGMb with 234 villages as the observation areas indicates that the average is in class 4, which is included in moderate greenish classification. Moderate greenish classification (greenery) is interpreted that the study area is overgrown with grasslands, shrubs, barren, sandy, rocky areas and a low population of vegetation canopy, showing areas around the mountain peaks. Some areas in the moderate greenish classification have forest ecosystem and field ecosystem types. The field ecosystem is an artificial ecosystem adjacent to the villages and is dominated by crops such as cassava, red chili, corn. The analysis results using Global Moran's and Local Moran's indicates Positive Spatial Autocorrelation, means that aridity phenomena has spatial connectivity between observed regions. The Clustering K Means Analysis on the high vegetation density conditions, shows that the weight of the distance between the vegetation data to the centroid is shorter, therefore the data is seen to be concentrated on a region. In low vegetation density conditions, therefore the weight of the distance between the data to the centroid is wider, so the data is seen to be more widely distributed. The dominance of the green colored data group indicates the dominance of the grids that marks as vegetation. The dominance of the brown and red data groups indicates the dominance of the grids that marks as a non-vegetation surface area or low vegetation growth and indicates the occurrence of aridity.

Acknowledgment
This scientific work is based on the research results conducted on the financing of Research Grand PUPT The Ministries Of Research, Technology, And Higher Education Republic of Indonesia 2018/2020.

References
[1] S. M. Vicente-Serrano, “Evaluating the impact of drought using remote sensing in a Mediterranean, Semi-arid Region,” Nat. Hazards, vol. 40, no. 1, pp. 173–208, 2007.

[2] N. P. Robinson et al., “A dynamic landsat derived normalized difference vegetation index (NDVI) product for the conterminous United States,” Remote Sens., vol. 9, no. 8, 2017.

[3] B. C. Gao, “NDWI - A normalized difference water index for remote sensing of vegetation liquid water from space,” Remote Sens. Environ., vol. 58, no. 3, pp. 257–266, 1996.

[4] M. K. Abuzar et al., “Drought risk assessment using GIS and remote sensing: A case study of District Khushab, Pakistan,” 15th Int. Conf. Environ. Sci. Technol., no. September, 2017.

[5] F. Caparrini and F. Manzella, “Hydrometeorological and vegetation indices for the drought monitoring system in Tuscany region, Italy,” Adv. Geosci., vol. 17, pp. 105–110, 2008.

[6] S. Wang, C. Huang, L. Zhang, Y. Lin, Y. Cen, and T. Wu, “Monitoring and assessing the 2012 drought in the great plains: Analyzing satellite-retrieved solar-induced chlorophyll fluorescence, drought indices, and gross primary production,” Remote Sens., vol. 8, no. 2, 2016.

[7] A. Zargar, R. Sadiq, B. Naser, and F. I. Khan, “A review of drought indices,” Environ. Rev., vol. 19, no. NA, pp. 333–349, 2011.

[8] M. R. Amri et al., Risiko bencana indonesia. Jakarta, 2016.

[9] F. Zambrano, M. Lillo-Saavedra, K. Verbist, and O. Lagos, “Sixteen years of agricultural drought assessment of the biobio region in chile using a 250 m resolution vegetation condition index (VCI),” Remote Sens., vol. 8, no. 6, pp. 1–20, 2016.

[10] WMO and GWP, Handbook of drought indicators and indices, no. 1173. 2016.

[11] V. Vani, “Comparative Study Of Ndvi And Savi Vegetation Indices In Anantapur District Semi-Arid Areas,” vol. 8, no. 4, pp. 559–566, 2017.

[12] H. T. Tran, J. B. Campbell, T. D. Tran, and H. T. Tran, “Monitoring drought vulnerability using multispectral indices observed from sequential remote sensing (Case Study: Tuy Phong, Binh Thuan, Vietnam),” GIScience Remote Sens., vol. 54, no. 2, pp. 167–184, 2017.

[13] E. E. Small, C. J. Roesler, and K. M. Larson, “Vegetation Response to the 2012 – 2014 California Drought from GPS and Optical Measurements,” pp. 1–16, 2018.

[14] Z. Mašková, F. Zemek, and J. Květ, “Normalized difference vegetation index (NDVI) in the management of mountain meadows,” Boreal Environ. Res., vol. 13, no. 5, pp. 417–432, 2008.

[15] C. S. Gherghina, C. Maftei, and C. Filip, “Assessment of Multi-spectral Vegetation Indices using Remote Sensing and Grid Computing,” Int. J. Comput., vol. 5, no. 4, 2011.

[16] F. Ghaleb, M. Mario, and A. N. Sandra, “Regional Landsat-Based Drought Monitoring from 1982 to 2014,” pp. 563–577, 2015.

[17] B. Basso, D. Cammarano, and P. De Vita, “Remotely sensed vegetation indices: Theory and applications for crop management,” Ital. J. Agrometeorol., vol. 53, no. 1, pp. 36–53, 2004.

[18] H. Tonbul, T. Kavzoglu, and S. Kaya, “Assessment of fire severity and post-fire regeneration based on topographical features using multi-temporal Landsat imagery: A case study in Mersin, Turkey,” Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. - ISPRS Arch., vol. 41, no. July, pp. 763–769, 2016.

[19] S. S. Panda, D. P. Ames, and S. Panigrahi, “Application of vegetation indices for agricultural crop yield prediction using neural network techniques,” Remote Sens., vol. 2, no. 3, pp. 673–696, 2010.

[20] J. L. Hatfield and J. H. Prueger, “Value of using different vegetative indices to quantify agricultural crop characteristics at different growth stages under varying management practices,” Remote Sens., vol. 2, no. 2, pp. 562–578, 2010.

[21] L. Scheipers, B. Haest, S. Veraverbeke, T. Spanhove, J. Vanden Borre, and R. Goossens, “Burned area detection and burn severity assessment of a heathland fire in belgium using airborne imaging spectroscopy (APEX),” Remote Sens., vol. 6, no. 3, pp. 1803–1826, 2014.

[22] X. Yang, S. Zhao, X. Qin, N. Zhao, and L. Liang, “Mapping of urban surface water bodies from sentinel-2 MSI imagery at 10 m resolution via NDWI-based image sharpening,” Remote Sens., vol. 9, no. 6, pp. 1–19, 2017.
[23] F. Nioti, F. Xystrakis, N. Koutsias, and P. Dimopoulos, “A remote sensing and GIS approach to study the long-term vegetation recovery of a fire-affected pine forest in southern Greece,” *Remote Sens.*, vol. 7, no. 6, pp. 7712–7731, 2015.

[24] A. L. F., R. Hidayat, and Haris, “Comparison between remote-sensing-based drought indices in East Java This,” *IOP Conf. Ser. Earth Environ. Sci. 54*, vol. 9, no. 54, 2017.

[25] P. R. Hakim, A. Rahman, and E. Rachim, “Model Koreksi Geometri Sistematik Data Imager Pushbroom Menggunakan Metode Proyeksi Kolinear,” *J. Teknol. Dirgant.*, vol. 10, no. 2, pp. 121–132, 2012.

[26] Z. Hasan, *P 12/Menhut-II/2012 ttg Perubahan kedua atas Permenhut no P 32/Menhut-II/2009/ tentang tata cara penyusunan RTKRHL DAS*. 2012.