

# Spatial-Temporal Characterization of Atmospheric Aerosols via Airborne Spectral Imaging and Growing Hierarchical Self-Organizing Maps

John W. Makokha\*, Jared O. Odhiambo

Department of Science Technology and Engineering, Kibabii University, Bungoma, Kenya

Email: \*makokhajw@kibu.ac.ke

**How to cite this paper:** Makokha, J.W. and Odhiambo, J.O. (2018) Spatial-Temporal Characterization of Atmospheric Aerosols via Airborne Spectral Imaging and Growing Hierarchical Self-Organizing Maps. *Journal of Geoscience and Environment Protection*, 6, 101-110.

<https://doi.org/10.4236/gep.2018.66008>

**Received:** March 12, 2018

**Accepted:** June 22, 2018

**Published:** June 25, 2018

Copyright © 2018 by authors and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## Abstract

Neural network analysis based on Growing Hierarchical Self-Organizing Map (GHSOM) is used to examine Spatial-Temporal characteristics in Aerosol Optical Depth (AOD), Ångström Exponent (ÅE) and Precipitation Rate (PR) over selected East African sites from 2000 to 2014. The selected sites of study are Nairobi (1°S, 36°E), Mbita (0°S, 34°E), Mau Forest (0.0° - 0.6°S; 35.1°E - 35.7°E), Malindi (2°S, 40°E), Mount Kilimanjaro (3°S, 37°E) and Kampala (0°N, 32.1°E). GHSOM analysis reveals a marked spatial variability in AOD and ÅE that is associated to changing PR, urban heat islands, diffusion, direct emission, hygroscopic growth and their scavenging from the atmosphere specific to each site. Furthermore, spatial variability in AOD, ÅE and PR is distinct since each variable corresponds to a unique level of classification. On the other hand, GHSOM algorithm efficiently discriminated by means of clustering between AOD, ÅE and PR during Long and Short rain spells and dry spell over each variable emphasizing their temporal evolution. The utilization of GHSOM therefore confirms the fact that regional aerosol characteristics are highly variable be it spatially or temporally and as well modulated by PR received over each variable.

## Keywords

Aerosol Optical Depth, Ångström Exponent, Neural Network, Satellite Spectral Imaging, Precipitation Rate, East African Atmosphere

## 1. Introduction

Aerosols sources and their short lifetime (5 - 10 days) results into a spatial-temporal

heterogeneous field that makes aerosol characterization real challenge [1] [2]. Despite this, recent initiatives by organizations such as NASA among others have increasingly deployed a number of passive remote sensing platforms that provide systematic and accurate long-term measurements of aerosol characteristics over the globe. This initiative hasn't been reciprocated adequately by the science community since a large percentage of the data actually used is low, in part because of a lack of efficient and effective analysis tools. For example, less than 5% of all remotely sensed images are ever viewed by human eyes or actually used [3]. Therefore, the increasing quantity and type of data available for climate change research studies among them atmospheric aerosols require effective feature extraction methods such as self-organizing map (SOM) and the growing hierarchical self-organizing map (GHSOM). Additionally, accurate extraction of key features and characteristic patterns of variability from a large data set is vital to correctly monitor atmospheric processes and how they alter climate change [4].

Techniques for pattern detection *i.e.* clustering, classifying and feature extraction for multi-dimensional spectroscopic datasets are becoming increasingly important since the previous is growing in size and complexity. The SOM, an artificial neural network based on unsupervised learning, is an effective software tool of feature extraction [5] [6]. It provides a nonlinear cluster analysis, mapping high dimensional data onto a (usually) 2D output space while preserving the topological relationships of the input data. As a tool for pattern recognition and classification, the SOM analysis is in widespread use across a number of disciplines among the climate research [7] [8] [9].

Notwithstanding its wide applications, SOM analysis has inherent deficiencies. First, it utilizes static network architecture with respect to the number and arrangement of neural nodes that are predefined prior to the start of training. Second, hierarchical relations between the input data are difficult to be detected in the map display. These two issues have been addressed within a single framework of the GHSOM that is available [10] [11]. The GHSOM is composed of independent SOMs, which are allowed to grow in size during the training process until a quality criterion regarding data representation is met. This growth process is further continued to form a layered architecture such that hierarchical relations between input data are further detailed at lower layers of the neural network.

It is important to note that both SOM and GHSOM techniques can be utilized through their respective MATLAB toolboxes that are available free online. The SOM MATLAB Toolbox (version 2.0) utilizes MATLAB structures, making it convenient to tailor the code for specific user needs and can be downloaded from a Website of the Helsinki University of Technology, Finland: <http://www.cis.hut.fi/projects/somtoolbox/>. On the other hand, the GHSOM Toolbox, developed jointly by the University of Aberdeen and Vienna University of Technology, can be downloaded at <http://www.ifs.tuwien.ac.at/~andi/ghsom/>.

The current study presents novel techniques for spatial-temporal aerosol characterization over East Africa for over a decade of monthly selected aerosol optical and microphysical properties *i.e.* Aerosol Optical Depth (AOD), Ångström Exponent ( $\text{\AA}$ ) and Mass Concentration (MC) from Moderate Resolution Imaging Spectrometer (MODIS) through SOM and GHSOM toolboxes in MATLAB. The two techniques *i.e.* SOM and GHSOM demonstrate their applicability in pattern detection, classification, feature extraction and temporal variations.

## 2. Methodology

### 2.1. Description of Study Area

The East Africa region covers diverse land forms comprising of glaciated mountains, Semi-Arid, Plateau and Coastal regions. Details and the map illustrating the study region and specifics on each site of study are as shown in **Figure 1**.



**Figure 1.** Study sites. Source: [12].

## 2.2. Growing Hierarchical Self-Organizing Maps (GHSOM)

The inherent deficiencies of SOM are well addressed by GHSOM through the use of an incrementally growing version of the SOM, which does not require the user to directly specify the size of the map beforehand and its enhanced ability to adapt to hierarchical structures in the data as illustrated in **Figure 2** [10] [11]. Prior to the training process, a “map” in layer 0 consisting of only one unit is created. This unit’s weight vector is initialized as the mean of all input vectors and its mean quantization error ( $MQE$ ) of unit  $i$  is computed as:

$$MQE_i = \frac{1}{|U_i|} \sum_{k \in U_i} \|x_k - m_i\|, \quad U_i = \{k | c_k = i\} \quad (1)$$

A new 2D array SOM is always created underneath layer 0 map so as to increase the map size so that all the spectroscopic data is well represented. A mean of all  $MQE_i$  normally defined as  $\langle MQE \rangle$  is compared to the  $MQE$  in the above layer  $\langle MQE \rangle_{above}$  and if the inequality in the following equation is fulfilled, then a new row or column of map units is inserted in the SOM.

$$\langle MQE \rangle > \tau_1 \langle MQE \rangle_{above} \quad (2)$$

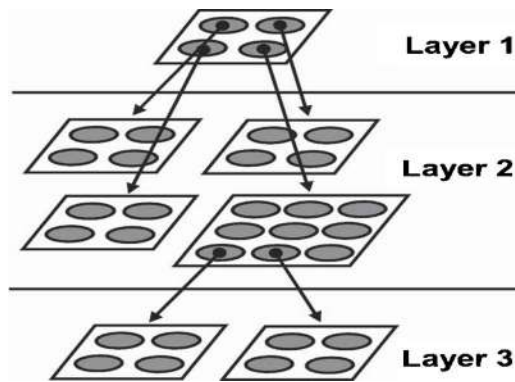
where  $\tau_1$  is the breadth controlling parameter. The GHSOM array of unit  $i$  with the largest  $MQE_i$  is normally the error unit. On the other hand, the unit with the largest distance with respect to the model vector is selected and a new row or column is inserted between the two. If the inequality in the above equation is not satisfied, then the next decision is whether to expand some units in the next hierarchical level or not. If the data mapped onto one single unit  $i$  still has a large variation *i.e.*

$$\langle MQE_i \rangle > \tau_2 \langle MQE \rangle_{above} \quad (3)$$

where  $\tau_2$  is the depth controlling parameter with the values of both  $\tau_1$  and  $\tau_2$  are chosen as  $1 > \tau_1 \gg \tau_2 > 0$ . Of significance to note, is the fact that smaller values of both  $\tau_1$  and  $\tau_2$  implies large SOM arrays and more layers GHSOM will have in the hierarchy respectively. The schematic of how GHSOM was implemented is as shown in **Figure 2**.

## 2.3. Data

Level-3 MODIS gridded atmosphere monthly global product “MOD08\_M3” at spatial resolution of  $1^\circ \times 1^\circ$  [13], was used in the current study for spatio-temporal characterization of AOD (at 550 nm), Ångström Exponent (ÅE) (at 470 - 660 nm) and Precipitation Rate (PR) over selected sites of East Africa from 2000 to 2014. Based on **Figure 2**, the MODIS level 3 monthly data was rearranged in a 2D array with the rows and columns representing the temporal and spatial dimensions respectively. The row vector at each time step was used to update the weight of the SOM via an unsupervised learning algorithm. The outcome weight vectors of the SOM nodes are reshaped back into characteristic data patterns [4] [9]. Likewise, the same data arrangement was utilized in implementing the



**Figure 2.** Graphical representation of the hierarchical structure of the GHSOM where the four units in the first layer of SOM are expanded in the second layer and only two units of the second layer are further expanded in the third layer [9].

GHSOM based on **Figure 2**. GHSOM utilizes the hierarchical structure where the four units in the first layer of SOM are expanded in the second layer and only two units of the second layer are further expanded in the third layer and so on [9].

### 3. Results and Discussions

#### 3.1. Growing Hierarchical Self-Organizing Maps (GHSOM)

Aerosol Optical Depth (AOD), Angstrom Exponent ( $\text{\AA}E$ ) and Precipitation Rate (PR) data values consist of 62, 69 and 136 samples respectively with each having six variables denoting the spatial dimension. It was noted that GHSOM displayed a more detailed granularity in each of the variable datasets. For example, 34 - 38 and 41 - 43 clusters are displayed during GHSOM classification of AOD and  $\text{\AA}E$  respectively. Likewise, a total of 69 - 72 clusters are revealed during GHSOM classification of the PR dataset, the high number of clusters was attributed to the highest sample space as compared to the rest of the datasets used in this work. The darker and white colors in each variable space implies lower and high occurrence of the feature respectively in this case AOD,  $\text{\AA}E$  and PR. Classification of both AOD and  $\text{\AA}E$  using GHSOM was attributed to various factors among them aerosol transport, diffusion, direct emission, hygroscopic growth and their scavenging from the atmosphere.

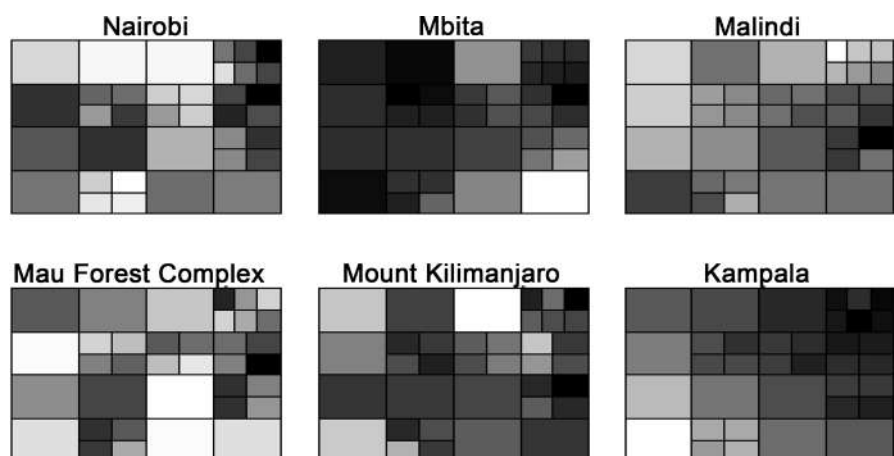
##### 3.1.1. GHSOM Analysis of Aerosol Optical Depth

Recent studies over the East African region show that aerosol characteristics are controlled directly by the local climate, *i.e.* Monsoonal precipitation [14]. Other modulators of aerosol characteristics include direct emissions *i.e.* through anthropogenic or natural activities over the entire East African atmosphere [12] [15] [16] [17] [18]. Based on the results obtained by the GHSOM algorithm, both temporal and spatial variability in AOD over each site is unique and majorly dependent on the seasonal variability. Monsoon precipitation accelerates wet scavenging of aerosols from the atmosphere hence the low AOD values during

wet season (the black clusters) over each of the six sites of study (variables) [19] [20]. It is of importance to note that the GHSOM algorithm efficiently discriminated, by means of clustering between AOD, during wet and dry seasons over each variable. The dark clusters correspond to long rain spells that are associated with enhanced scavenging of AOD hence, their low values while greyish clusters correspond to a less aerosol scavenging from the atmosphere due to low PR. Moreover, the white clusters reveal enhanced AOD values due to inefficient scavenging of atmospheric aerosols via dry deposition over each variable during dry season. The details of the GHSOM classification of AOD over the six variables are shown in **Figure 3**.

On the other hand, spatial variability in AOD is pronounced since each variable corresponds to a unique level of classification. This variability is not only dependent on seasons but also on aerosol transport, anthropogenic influence, diffusion, direct emission, hygroscopic growth and their scavenging from the atmosphere which explain the varied classification levels observed over each study site. Of significance to note is the fact that Mau Forest experiences low number of dark clusters but significantly more greyish and white clusters. This maybe as a result of continual biomass burning and forest clearance for agricultural use even after the process of land reclamation stated in 2008 [21]. Even though Nairobi experiences relatively higher precipitation rate ( $0.15 \pm 0.02$  mm/hr) (**Figure 5**), it has an enhanced number of greyish and white clusters that may be attributed to anthropogenic influences e.g. increasing populace, vehicular and industrial emissions and biomass and refuse burning [15].

Additionally, Mbita experiences low AOD values relatively occasioned by the Lake-land air mass exchange controlled precipitation rate ( $0.19 \pm 0.01$  mm/hr) (**Figure 5**) which enhances aerosol scavenging during the study period and hinders biomass burning activities [18] [19] [20]. Meanwhile, maritime conditions coupled with long distance transport of aerosols from the Arabian Peninsula desert via Monsoon winds [18] explain the observed high number of greyish and white clusters (high AOD values) over Malindi. The high AOD values are



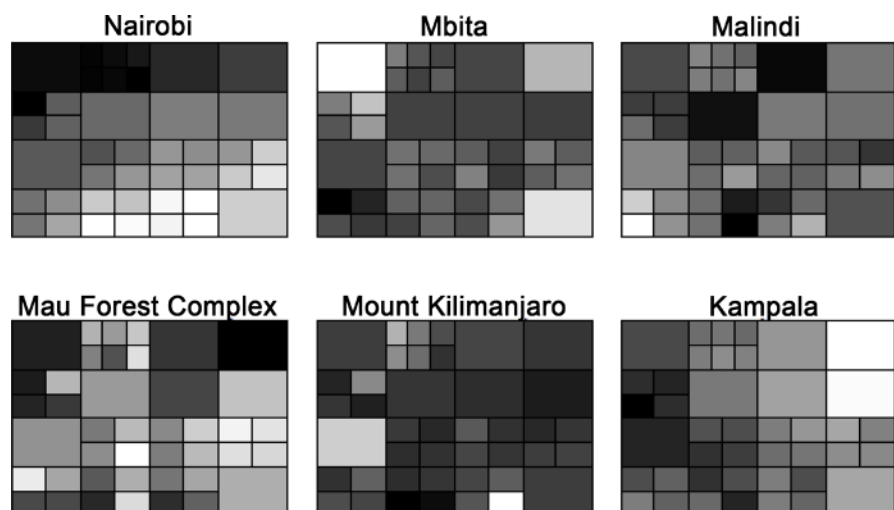
**Figure 3.** GHSOM classification of AOD over the six variables.

occasioned by significantly low precipitation rate ( $0.07 \pm 0.02$  mm/hr) observed in **Figure 5**. Likewise, Mount Kilimanjaro has been reclaimed in the recent past restraining the negative impacts of deforestation hence, explaining the dominance of greyish and dark clusters (low AOD values) during the study period [17]. The dominance of both greyish and white clusters for both AOD (**Figure 3**) and precipitation rate (**Figure 5**) implies high AOD and precipitation rate over Kampala during the study period. The observed high AOD values are associated to the high vehicular emissions from growing private motorized transport over the city [16].

### 3.1.2. GHSOM Analysis of Angstrom Exponent

Spectral curvature of aerosol extinction plays an important role when calculating the ÅE with only two wavelengths. ÅE calculated from longer wavelength pairs ( $\lambda = 670, 870$  nm) are sensitive to the fine mode fraction of aerosols but not the fine mode effective radius; conversely, shorter wavelength pairs ( $\lambda = 380, 440$  nm) are sensitive to the fine mode effective radius but not the fine mode fraction [22]. Hence it is important to consider the wavelength pair used to calculate the ÅE when making qualitative assessments about the corresponding aerosol size distributions. The present study presents spatio-temporal characterization of ÅE at 470 - 660 nm wavelength range via GHSOM algorithm. The wavelength range used captures more information on ÅE capability to act as a qualitative indicator of aerosol particle size distribution over the study sites. The results are presented in **Figure 4**.

As noted earlier, enhanced clarity in the ÅE classification was observed over each variable as indicated in **Figure 4**. Similarly, we note three clear clusters in ÅE *i.e.* black, greyish and white clusters over each variable. The black, greyish and white clusters imply low, medium and high ÅE values experienced over each site as indicated in **Figure 4**. Based on **Figure 4**, Nairobi ÅE variability spans the range  $0.94 - 1.68 \pm 0.06$  during the study period. Black clusters correspond to



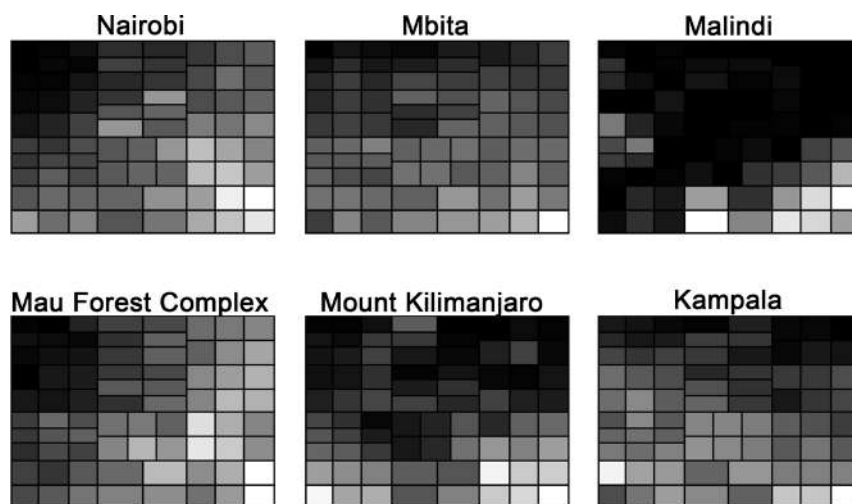
**Figure 4.** GHSOM classification of ÅE over the six variables.

low ÅE values which imply the dominance of aerosol particles in the 660 nm wavelength. On the contrary, the white clusters are indicative of high ÅE values of up to  $1.68 \pm 0.06$  which corresponds to the dominance of aerosol particles in the 470 nm wavelength over Nairobi. The greyish clusters over the site are indicative of the proper mixing of aerosol particles which is attributed to vehicular and industrial emissions and biomass and refuse burning over the site [15]. GHSOM reveals a unique ÅE classification over Mbita *i.e.* limited black and white clusters which imply that there are limited periods of time when the site experienced extreme low and high ÅE values. In general, the site is dominated by the greyish clusters ( $0.98 - 1.64 \pm 0.05$ ) which suggests proper mixing of aerosol particles from biomass burning over the site during the study period [18] [19] [20].

Malindi shows the least ÅE values of range  $0.18 - 0.60 \pm 0.06$  as compared to all the other study sites in the region. The low values confirm the fact that the site experiences maritime conditions accompanied by sea salt and sea spray aerosol particles plus long distance transport of aerosols from the Arabian Peninsula desert via Monsoon winds. Continual biomass burning and forest clearance for agricultural use over Mau Forest Complex [21], enhance the region's aerosol particles whose ÅE values span in the range ( $0.46 - 0.89 \pm 0.07$ ). The low ÅE values infer the dominance of aerosol particles in the 670 nm wavelength. Aerosol particles from high energy use and emissions associated with the growth of private motorized transport over Kampala [16], dominate the 470 nm wavelength, this may explain the higher span range of  $1.30 - 1.83 \pm 0.06$  in the ÅE values observed over the site as compared to the rest of the study sites.

### 3.1.3. GHSOM Analysis of Precipitation Rate

As noted earlier, darker and white colors infers lower and high manifestation PR over each study site. It's clear that Malindi experiences the lowest PR as compared to the rest of the region. Additionally, Mbita-Kampala and Nairobi-Mau Forest Complex PR are correlated during the study period as shown **Figure 5**.



**Figure 5.** GHSOM classification of precipitation rate over the six variables.

## 4. Conclusion

MODIS Terra monthly AOD and ÅE level 3 data from 2000 to 2014 are used to spatio-temporal characterization of AOD, ÅE and PR over selected study sites of the East African region using GHSOM algorithm. It is possible to use the neural network techniques in studying spatial-temporal characteristics over the region with enhanced efficiency. The GHSOM algorithm classification of both AOD and ÅE is attributed to various factors among them aerosol transport, diffusion, direct emission, hygroscopic growth and their scavenging from the atmosphere. The East African region experiences diverse and highly variable aerosol characteristics as revealed by GHSOM.

## Acknowledgements

This work was supported by the National Council for Science and Technology Grant funded by the Government of Kenya (NCST/ST & I/RCD/4<sup>TH</sup> call PhD/201). The authors wish to thank the NASA Goddard Earth Science Distributed Active Archive for MODIS Level 3, TRMM rainfall data which served as a complement to the meteorological data from the Kenya Meteorological Department.

## References

- [1] Solomon, S. (2007) Climate Change 2007: The Physical Science Basis: Working Group I Contribution to the Fourth Assessment Report of the IPCC (Vol. 4). Cambridge University Press, Cambridge.
- [2] Stocker, T. (2014) Climate Change 2013: The Physical Science Basis: Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge.
- [3] Petrou, M. (2004) Preface. *Pattern Recognition Letters*, **25**, 1459. <https://doi.org/10.1016/j.patrec.2004.05.021>
- [4] Liu, Y. and Weisberg, R.H. (2011) A Review of Self-Organizing Map Applications in Meteorology and Oceanography. In: Mwasiagi, J.I., Ed., *Self Organizing Maps - Applications and Novel Algorithm Design*, InTech. <http://www.intechopen.com/books/selforganizing-maps-applications-and-novel-algorithm-design/a-review-of-self-organizing-map-applications-inmeteorology-and-oceanography>
- [5] Kohonen, T. (1982) Self-Organized Information of Topologically Correct Features Maps. *Biological Cybernetics*, **43**, 59-69. <https://doi.org/10.1007/BF00337288>
- [6] Kohonen, T. (2001) Self-Organizing Maps. Springer-Verlag, New York, Berlin, Heidelberg. <https://doi.org/10.1007/978-3-642-56927-2>
- [7] Oja, M., Kaski, S. and Kohonen, T. (2003) Bibliography of Self-Organizing Map (SOM) Papers: 1998-2001 Addendum. *Neural Computing Surveys*, **3**, 1-156.
- [8] Hong, Y., Hsu, K., Sorooshian, S. and Gao, X. (2004) Precipitation Estimation from Remotely Sensed Imagery Using an Artificial Neural Network Cloud Classification System. *Journal of Applied Meteorology*, **43**, 1834-1853. <https://doi.org/10.1175/JAM2173.1>
- [9] Liu, Y., Weisberg, R.H. and He, R. (2006) Sea Surface Temperature Patterns on the

- West Florida Shelf Using the Growing Hierarchical Self-Organizing Maps. *Journal of Atmospheric and Oceanic Technology*, **23**, 325-328.  
<https://doi.org/10.1175/JTECH1848.1>
- [10] Dittenbach, M., Rauber, A. and Merkl, D. (2002) Uncovering the Hierarchical Structure in Data Using the Growing Hierarchical Self-Organizing Map. *Neurocomputing*, **48**, 199-216. [https://doi.org/10.1016/S0925-2312\(01\)00655-5](https://doi.org/10.1016/S0925-2312(01)00655-5)
- [11] Pampalk, E., Widmer, G. and Chan, A. (2004) A New Approach to Hierarchical Clustering and Structuring of Data with Self-Organizing Maps. *Intelligent Data Analysis*, **8**, 131-149.
- [12] Makokha, J.W., Odhiambo, J.O. and Godfrey, J.S. (2017) Trend Analysis of Aerosol Optical Depth and Angstrom Exponent Anomaly over East Africa. *Atmospheric and Climate Sciences*, **7**, 588-603. <https://doi.org/10.4236/acs.2017.74043>
- [13] Ichoku, C., Kaufman, Y.J., Remmer, L.A. and Levy, R. (2004) Global Aerosol Remote Sensing from MODIS. *Advances in Space Research*, **34**, 820-827.  
<https://doi.org/10.1016/j.asr.2003.07.071>
- [14] De Graaf, M., Tilstra, L.G., Aben, I. and Stammes, P. (2010) Satellite Observations of the Seasonal Cycles of Absorbing Aerosols in Africa Related to the Monsoon Rainfall, 1995-2008. *Atmospheric Environment*, **44**, 1274-1283.  
<https://doi.org/10.1016/j.atmosenv.2009.12.038>
- [15] van Vliet, E.D.S. and Kinney, P.L. (2007) Impacts of Roadway Emissions on Urban Particulate Matter Concentrations in Sub-Saharan Africa: New Evidence from Nairobi, Kenya. *Environmental Research Letters*, **2**, 045028.  
<https://doi.org/10.1088/1748-9326/2/4/045028>
- [16] Mabasi, T. (2009) Assessing the Impacts, Vulnerability, Mitigation, and Adaptation to Climate Change in Kampala City. Fifth Urban Research Symposium. Kampala.
- [17] Fairman, J.G., Nair, U.S., Christopher, S.A. and Mölg, T. (2011) Land Use Change Impacts on Regional Climate over Kilimanjaro. *Journal of Geophysical Research: Atmospheres*, **116**, D03110.
- [18] Makokha, J.W. and Angeyo, H.K. (2013) Investigation of Radiative Characteristics of the Kenyan Atmosphere due to Aerosols Using Sun Spectrophotometry Measurements and the COART Model. *Aerosol and Air Quality Research*, **13**, 201-208.
- [19] Ngaina, J.K. and Mutai, B.K. (2013) Observational Evidence of Climate Change on Extreme Events over East Africa. *Global Meteorology*, **2**, 6-12.
- [20] Ngaina, J.N., Mutai, B.K., Ininda, J.M. and Muthama, J.N. (2014) Monitoring Spatial-Temporal Variability of Aerosol over Kenya. *Ethiopian Journal of Environmental Studies and Management*, **7**, 244-252.
- [21] National Environmental Management Authority, Kenya (NEMA) (2013) Mau at a Glance. NEMA Report.
- [22] Eck, T., Holben, B.N., Ward, D.E., Dubovik, O., Reid, J.S., Smirnov, A., Mukelabai, M.M., Hsu, N.C., O'Neill, N.T. and Slutsker, I. (2001) Characterization of the Optical Properties of Biomass Burning Aerosols in Zambia during the 1997 ZIBBEE Field Campaign. *Journal of Geophysical Research* **106**, 3425-3448.  
<https://doi.org/10.1029/2000JD900555>