Monitoring Urban Expansion And Land Use/Land Cover Changes In Banadir, Somalia Using Google Earth Engine (GEE)

A M Hamud\textsuperscript{1}, H Z M Shafri\textsuperscript{1,\ast} and N S N Shaharum\textsuperscript{2}

\textsuperscript{1}Department of Civil Engineering, and Geospatial Information Science Research Centre (GISRC), Faculty of Engineering, Universiti Putra Malaysia, Serdang, Selangor, 43400, Malaysia
\textsuperscript{2}School of Engineering, Monash University Malaysia, Jalan Lagoon Selatan, 47500, Bandar Sunway, Selangor Darul Ehsan, Malaysia

helmi@upm.edu.my

Abstract. Land Use Land Cover (LULC) changes in the Banadir region are rapidly changing because of the increasing interaction of human activates with the environment as the population increases. However, there is no published evidence on this phenomenon. This study used multi-temporal Landsat images (1989, 2003, and 2018) to extract and evaluate LULC changes in Banadir Somalia. Based on Google Earth Engine (GEE) platform, two Machine Learning (ML) algorithms were used and compared, using supervised classification approaches via Support Vector Machine (SVM) and the Random Forest (RF) classifiers. This study shows that the average classification accuracy of SVM is showing the highest overall accuracy of 96.7\% (Kappa = 0.946), which is 1.7\% higher than that achieved by RF. The analysis of LULC changes in Banadir for the last three decades revealed that the urban area increased from 43.1 km\textsuperscript{2} in 1989 to 87.7 km\textsuperscript{2} in 2018, which means that the urban expanded or increased by two folds. Vegetation land cover dropped down from 276.5 km\textsuperscript{2} in 1989 to 96.42 km\textsuperscript{2} in 2018, which is a two-thirds decrease in vegetation in the last 30 years whereas bare surface area increased from 93.74 km\textsuperscript{2} in 1989 to 229.05 km\textsuperscript{2} in 2018, which is 1.96\% increase per year. This study concludes that Banadir has experienced a dramatic increase in the urban and bare surface and a decrease in vegetation for the past 30 years.

1. Introduction

Nowadays, LULC and urban expansion are worldwide challenges with environmental and socioeconomic consequences, because it may lead to different environmental challenges such as water quality degradation, air contamination, and loss in biodiversity [1]. LULC change; the primary driving force of worldwide ecological change is the backbone for the sustainable improvement issue. Urban expansion such as the movement of housing and commercial land to countryside places outside of region locations has extended been regarded as a sign of regional economic strength.

Urban expansion has increased the opportunities on several sides such as services, employment, and production which leads to maximizing rural-urban migration. Urbanization became a global socioeconomic challenge that had directed to the impact of LU change worldwide [2]. Africa is considered as the least urbanized, only 43\% of its population live in urban areas compared to Asia and Europe. Although it’s urban development and populations are well not documented [3]. Most of the African cities are facing a challenge related to urban growth such as unplanned and uncontrolled as well as an
informal settlement [4-7].

In Somalia, the last 30 years have been affected by civil war resulting migration of large-scale population and vegetation land change. Several remote sensing studies have revealed that war and conflict may lead to LULC changes [8]. Banadir region is among one of the rapid populations growing cities in East Africa with a population of 2.1 million in 2018, the average yearly growth of 3.52. Generally, LU and urban expansion in remote sensing require the examination of two registered satellite multispectral or aerial in the same geographical location obtained within the same geographic location amongst the two occasions regarded. Urban expansion monitoring is the method of studying the variations in the state of an object by remotely sensing it on various occasions. Expansion monitoring is the approach for figuring out the properties of LULC changes based on multi-temporal remote sensing data.

As far as we know, there are no studies done in urbanization and LULC change detection using cloud-based remote sensing platform Google Earth Engine (GEE) in Somalia. The objective of this paper is to investigate urban expansion and LULC Changes in the Banadir region by using GEE.

2. Study area
The study area, Banadir region resides along the southeastern coast of Somalia in the Horn of Africa, a region that occupies the extreme northeast corner of the continent. It sits on coastal dunes in an extremely arid, desert environment Somali currently serves as the primary language of the region [9]. Other languages used less in the city are Arabic, Italian, and English. It is one of the biggest as well as the fastest-growing city in the country with around 91 km² coverage. Banadir region has 17 districts such as Bondhere, Daynil, Dharkenley, Hamar-Jajab, Hamar-Weyne, Heliwa, Hodan, Howl- Wadag, Karan, Shangani, Shibis, Waberi, Wadajir, Wardhigley, Kaxda, Abdiiaziz, and Yaqshid. It has a population of 2,425,000 residents [9].

![Banadir Map](image)

**Figure 1.** Banadir region capital of Somalia

3. Data and methods
In this study, satellite imageries of Landsat 5 (TM), Landsat 7 (ETM+) and Landsat 8 (OLI) with surface reflectance were used. Landsat data can improve the understandings of the land changes of the earth. This study applied machine learning algorithms, namely Random Forest (RF) and Support Vector Machine (SVM). SVM is a machine learning technique and binary classifier that works to identify the optimal hyperplane and promptly divides the data points into two classes. Furthermore, SVM is a powerful algorithm that can work well with a limited number of samples [10]. The RF algorithm is another type of machine learning technique used for image classification. It is a robust algorithm that works through building multiple decision trees and merges them to produce accurate output.
3.1. Method
The data processing stage has four parts. First, we selected Landsat data with cloud-free images from GEE archive, and then we used the following Landsat’s images: TM 1989, TM+ 2003 and OLI 2018. Second, pre-processed Landsat imageries available through GEE were used to assess LULC change across the Banadir region. Third, supervised classification using different algorithms, SVM and RF were used to evaluate the classification of LULC with overall accuracy assessment implemented. Finally, a change detection comparison between 1989, 2003, and 2018 was evaluated to monitor the LULC.

3.2. Data processing in GEE
GEE is the platform tool used to perform the processing and analyzing the satellite imagery during the last four decades. It includes two primary components that actually work with one another that are the GEE Explorer for data view and also the GEE API. Figure 2 shows the overall methodology for this study and a JavaScript API was used along with three different imageries of Landsat 5, Landsat 7, and Landsat 8 were obtained to perform the LULC.

![Flow chart of the methodology](image)

The GEE API can be used for both loading and visualizing massive satellite imagery as well as managing complex geostatistical and geospatial procedures on the imagery. Furthermore, GEE processes data through cloud computing and it provides a vast amount of dataset (e.g. Landsat, Sentinel, MODIS and others) and algorithms (e.g. SVM, RF, Naïve Bayes and others). A visual analysis of the classified pixels in all three different Landsat imageries was conducted in which a total of 30% of the total samples were used as the testing samples for evaluating the classified maps. Overall accuracy evaluation was conducted by using an accuracy assessment procedure in the GEE. The outputs generated within GEE were exported to ArcGIS for further analysis.

3.3. McNemar test
This study used two machine learning algorithms, namely SVM and RF to perform image classification. A statistical McNemar’s test is required to evaluate the significance of the difference between two classified maps. The test was carried out based on the 2 x 2 matrix table by comparing the pixels classified by Classifier 1 (C1) and Classifier 2 (C2) as shown in Table 1.
Table 1. 2 x 2 matrix table

| C2 | Correct | Wrong |
|----|---------|-------|
| CI |         |       |
| Correct | a | b |
| Wrong    | c | d |

The test is calculated based on the sample ratio [11] using the formula as follows:

\[ X^2 = \frac{(b - c)^2}{b + c} \]  

(1)

The test is considered to be statistically significant if \( p < 0.05 \), thus rejecting the null hypothesis [12].

4. Results

4.1. LULC classification accuracy evaluation

The overall accuracy and Kappa coefficient of LULC classification for two different machine learning algorithms, as shown in Table 2. In this study, the accuracy classification result of SVM was 96.7% (kappa = 0.946), which is 1.7% higher than that achieved by RF. This table also shows the accuracy and Kappa within a different time frame (1989, 2003, and 2018). In 1989, 2003, 2018 we achieved 94%, 96%, 95% when using RF compared to 96%, 98%, and 96% when using SVM respectively.

Table 2. Accuracy evaluation of LULC classification results

|            | Random Forest | Support Vector Machine |
|------------|---------------|------------------------|
|            | Accuracy      | Kappa                  | Accuracy | Kappa |
| 1989       | 94%           | 0.922                  | 96%      | 0.924 |
| 2003       | 96%           | 0.953                  | 98%      | 0.962 |
| 2018       | 95%           | 0.931                  | 96%      | 0.953 |
| Average    | 95%           | 0.935                  | 96.7%    | 0.946 |

4.2. McNemar test result

Table 3. Matrix table for year 1989, 2003 and 2018.

|            | Support Vector Machine |
|------------|------------------------|
|            | Correct | Wrong |
| 1989       |         |       |
| Correct    | 83      | 0     |
| Wrong      | 3       | 8     |
| 2003       |         |       |
| Correct    | 114     | 3     |
| Wrong      | 1       | 1     |
| 2018       |         |       |
| Correct    | 110     | 2     |
| Wrong      | 0       | 0     |
McNemar’s test has been conducted to measure the significance of the SVM and RF algorithms in classifying the images, and the number of pixels tabulated in Table 3 was extracted from 1989, 2003 and 2018 classified maps respectively.

| Year | p-value |
|------|---------|
| 1989 | 0.08    |
| 2003 | 0.32    |
| 2018 | 0.16    |

The pixels in Table 3 were used to calculate the p-values as shown in Table 4. The p-values for 1989, 2003 and 2018 data are statistically not significant by producing p > 0.05. Thus, accepting the null hypothesis. Therefore, this test suggested that SVM and RF are not significantly different and both algorithms are suitable for accurate mapping of the study area.

4.3 LULC Changes in Banadir

Figures 3 and 4 show that the LULC change of Banadir from 1989 to 2003 and from 2003 to 2018 respectively. The urban area of Banadir increased from 43.1 km$^2$ in 1989 to 57.7 km$^2$ in 2003 and from 57.7 km$^2$ in 2003 to 87.7 km$^2$ in 2018. It means that in the first 15 years, the area has expanded approximately by 17% and in the second 15 years, Banadir urban area has grown up to 34%. The total expansion was 50%, which is a 2-fold increase in the last 30 years. However, the vegetation area decreased from 276.5 km$^2$ in 1989 to 135 km$^2$ in 2003 and a total of 96.42 km$^2$ of vegetation area was mapped in 2018. It shows a 51% decrease in that first 15 years and a 14% decrease in the next 15 years subsequently.

**Figure 3.** LULC Changes over the years

**Figure 4.** LULC Changes
Figure 4 shows an increment had occurred to built-up and bare surface from year 1989 to 2018. The vegetation loss from 275.42 km$^2$ to 135.32 km$^2$ (1989 – 2003) has contributed to the drastic increment of bare surface from 93.74 km$^2$ to 219.11 km$^2$. As shown in Table 5, the increment of built-up and bare surface gave an impact to vegetation loss over the years. The overall vegetation of the 30 years was two-thirds in decrease, as shown in Table 5. It may be due to increased droughts and charcoal business in Somalia. The bare surface has increased from 93.74 km$^2$ in 1989 to 219.1 km$^2$ in 2003 and 229.05 km$^2$ in 2018, which is 55%. Increase in the first 15 years and a 4% increase in the next 15 years. The overall bare surface had increased by 59% in the last 30 years.

| LULC       | 1989  | 2003  | 2018  |
|------------|-------|-------|-------|
|            | Area (km$^2$) | %    | Area (km$^2$) | %    | Area (km$^2$) | %    |
| Built-up   | 43.1442 | 10.46 | 57.726 | 14.00 | 87.7419 | 21.3 |
| Bare surface | 93.7404 | 22.73 | 219.1122 | 53.16 | 229.0545 | 55.43 |
| Vegetation | 275.4261 | 66.80 | 135.3231 | 32.83 | 96.4224 | 23.33 |
| Total      | 412.3107 | 100   | 412.3107 | 100   | 412.3107 | 100   |

5. Discussion

The built-up area was developed due to the rapid growth of the settlement area after the war [13] and causes the built-up area to expand gradually, for instance for the first 15 years, it has been expanded by 17% and in the second 15 years, it has doubled to 34% as shown in Table 5. There are some new districts emerged due to the increased built-up areas such as Kaxda, Ceelasha, and Daynile. Figure 5 shows the factors that could lead to climate change resulting from the reduction of vegetation cover and the growing number of the urban and bare surface areas. A study in Somalia that was carried out by [14] has successfully mapped degraded areas occurring from 1982 to 2009 into several rates: strong, moderate and light. One of the types of degradation is the reduction of vegetation cover that had caused habitat loss and riverbank erosion leading to wildlife extinctions and flood [15]. Furthermore, the loss of vegetation cover was found to be a major issue in Banadir in which this study has resulted in vegetation reduction from 275.42 km$^2$ to 96.42 km$^2$ and increment of the bare surface from 93.74 km$^2$ to 229.05 km$^2$ from 1989 to 2018 (as shown in Table 5).

![Figure 5](image_url)

**Figure 5.** Example of some of the factors that contribute to changes to LULC in the Banadir region, a) land degradation, b) urbanization and c) bare surface.

The environmental impacts such as climate change and drought have contributed to the high number of the bare surface accretion as shown in Figure 5(c). Without proper monitoring, this could lead to an
increase in temperature and dry out the water resources which will affect agricultural productivity and the green vegetations will have minimum water supply to survive [16]. Furthermore, the increment of built-up and bare surface areas has contributed to a high rate of degraded areas. Figures 5(b) show the urbanization areas in Daarul Salam City and Kaxda area. One of the main reasons is due to the new settlement areas needed to accommodate more than 500,000 people that have moved into Banadir as recorded by [16] from 2016 to 2019.

Hussein (2017) highlighted five factors as the drivers of drought in Somalia which are population growth, urbanization, deforestation, soil erosion and climate change. Furthermore, Somalia has experienced more than 55% of vegetation-cover loss for 40 years (1977 – 2016) in which resulted in global warming and emissions of CO2 [17]. The unplanned urbanization growth had caused economic and environmental impacts such as land degradation as shown in Figures 5(a). Besides drought, the vegetation loss occurred was due to human activities such as charcoal production, loss of soil fertility, poor cultivation practices of productive land and conflict over natural resources as stated by the [18]. The challenges caused by rapid unplanned urbanizations in Banadir region are an insufficient number of health facilities, inadequate human and financial resources, limited coordination between line ministries, regional and district authorities, implementing health partners, and uncontrolled private (commercial) medical sector. These aforementioned challenges had triggered various issues such as malnutrition, limited access to safe water, poor sanitation, the lack of preventive public health interventions and if no necessary actions are taken, these issues will get worse and become a life-threatening illness.

6. Conclusion
This study has presented the utilisation of SVM and RF algorithms via GEE platform in mapping and performing change detection analysis. The results showed that SVM algorithm produced better outputs with 96%, 98%, and 96% and the RF algorithm with 94%, 96%, and 95% of accuracies for 1989, 2003, and 2018 respectively. Both algorithms performed rather similarly (according to McNemar Test) and thus both can be used for monitoring the changes in Banadir region. Furthermore, the utilisation of machine learning algorithms to classify Landsat data taken from different years in GEE platform have shown that the GEE platform is a great tool for analyzing different data simultaneously as well as generating LULC maps over a large area in a short time. Furthermore, the versatility of the coding platform available in GEE platform was able to produce quality and cloud-free images. Thus, this study has successfully mapped and performed change detection analysis in Banadir region using the GEE platform via RF and SVM machine learning algorithms. In addition, the environmental impacts occurred in Banadir region were assessed through the LULC changes analysis outputs in which the increment of the population and urbanizations are the main contributions to land degradation and the increasing bare surface area. Hence, the information will be useful to the government agencies for future planning especially on mitigating the activities involved in Banadir region.

References
[1] Maimaitijiang M, Ghulam A, Sandoval JO, Maimaitiyiming M (2015) Drivers of land cover and land use changes in St. Louis metropolitan area over the past 40 years characterized by remote sensing and census population data. Int J of Applied Earth Obs and Geoinfo 35:161-174. https://doi.org/10.1016/j.jag.2014.08.020.
[2] Chikowore T, Willems E (2017) Identifying the changes in the quality of life of Southern African Development Community (SADC) migrants in South Africa from 2001 to 2011. South African Geo J 99:86-112. https://doi.org/10.1080/03736245.2016.1208577.
[3] UNDESA (2011) World Urbanisation Prospects: The 2011 Revision. New York: Population Division of the Department of Economic and Social Affairs of the United Nations Secretariat.
[4] Magidi J, Ahmed F (2018) Assessing urban sprawl using remote sensing and landscape metrics: A case study of City of Tshwane, South Africa (1984–2015). The Egyptian J of Remote Sens and
Space Sci. https://doi.org/10.1016/j.ejrs.2018.07.003.

[5] Fenta AA, Yasuda H, Haregeweyn N, Belay AS, Hadush Z, Gebremedhin MA, Mekonnen G (2017) The dynamics of urban expansion and land use/land cover changes using remote sensing and spatial metrics: the case of Mekelle City of northern Ethiopia. Int J of Remote Sens 38:4107-4129. https://doi.org/10.1080/01431161.2017.1317936.

[6] Kukkonen MO, Muhammad MJ, Käyhkö N, Luoto M (2018) Urban expansion in Zanzibar City, Tanzania: Analyzing quantity, spatial patterns and effects of alternative planning approaches. LU Policy 71:554-565. https://doi.org/10.1016/j.landusepol.2017.11.007.

[7] Abudu D, Echima RA, Andogah G (2018) Spatial assessment of urban sprawl in Arua Municipality, Uganda. The Egyptian J of Remote Sens and Space Sci. https://doi.org/10.1016/j.ejrs.2018.01.008.

[8] Al-doski J, Mansor SB, Shafri HZM (2013) War Impacts Studies Using Remote Sensing. IOSR J. Appl. Geol. Geophys 1:11-15.

[9] Unruh JD (2015) Mogadishu, Somalia. In Department of geography, Indiana University, Bloomington.

[10] Shafri HZM, Ramle FSH (2009). A Comparison of Support Vector Machine and Decision Tree Classifications Using Satellite Data of Langkawi Island. Information Technology Journal, 8(1):64-70.

[11] De Leeuw J, Jia H, Yang L, Liu X, Schmidt K, Skidmore AK (2006) Comparing accuracy assessments to infer superiority of image classification methods. Int J of Remote Sens 27:223-232. https://doi.org/10.1080/01431160500275762.

[12] Duro DC, Franklin SE, Dubé MG (2012) A comparison of pixel-based and object-based image analysis with selected machine learning algorithms for the classification of agricultural landscapes using SPOT-5 HRG imagery. Remote sensing of environment 118:259-272. http://dx.doi.org/10.1016/j.rse.2011.11.020.

[13] Abdulle A, Tan AA, Pradhan B, Abdullahi S (2016) Temporal assessment on land use land cover of Somalia after the effect of the civil war using remote sensing. In IOP Conference Series: Earth and Env Sci 37:1-012063. https://doi.org/10.1088/1755-1315/37/1/012063.

[14] Omuto CT, Balint Z, Alim MS (2014). A framework for national assessment of land degradation in the drylands: a case study of Somalia. Land degradation & development 25(2):105-119. https://doi.org/10.1002/ldr.1151.

[15] OCHA (2018, April 26) Somalia: Flash Floods. Reliefweb. Available online: https://reliefweb.int/flash/somalia-ff-2018-00041-som.

[16] UNDP (2018) Somalia Drought Impact & Needs Assessment. Volume 1 Synthesis Report. Available online: http://documents.vsemirnyjbank.org/curated/ra/901031516986381462/pdf/122991-v1-GSURR-Somalia-DINA-Report-Volume-I-180116-Digital.pdf.

[17] Hussein SMSSH (2017) Understanding the Drivers of Drought in Somalia: Environmental Degradation as a Drought Determinant. Available online: https://media.africaportal.org/documents/Understanding_the_Drivers_of_Drought_in_Somalia.pdf

[18] State Minister of Environment (2015) Somalia’s intended nationally determined contributions (INDCs). Available online: https://www4.unfccc.int/sites/ndcstaging/PublishedDocuments/Somalia%20First/Somalia%27s%20INDCs.pdf

Acknowledgements
The authors would like to acknowledge the supports from Universiti Putra Malaysia (UPM) and comments from the anonymous reviewers in improving this paper.