Understanding the spatial distribution of eroded areas in the former rural homelands of South Africa: Comparative evidence from two new non-commercial multispectral sensors

Terrence Koena Sepuru and Timothy Dube

Abstract
In this study, we determine the most suitable multispectral sensor that can accurately detect and map eroded areas from other land cover types in Sekhukhune rural district, Limpopo Province, South Africa. Specifically, the study tested the ability of multi-date (wet and dry season) Landsat 8 OLI and Sentinel-2 MSI images in detecting and mapping eroded areas. The implementation was done, using a robust non-parametric classification ensemble: Discriminant Analysis (DA). Three sets of analysis were applied (Analysis 1: Spectral bands as independent dataset; Analysis 2: Spectral vegetation indices as independent and Analysis 3: Combined spectral bands and spectral vegetation indices). Overall classification accuracies ranging between 80% to 81.90% for MSI and 75.71%–80.95% for OLI were derived for the wet and dry season, respectively. The integration of spectral bands and spectral vegetation indices showed that Sentinel-2 (OA = 83, 81%), slightly performed better than Landsat 8, with 82, 86%. The use of bands and vegetation indices as independent dataset resulted in slightly weaker results for both sensors. Sentinel-2 MSI bands located in the NIR (0.785–0.900 μm), red edge (0.698–0.785 μm) and SWIR (1.565–2.280 μm) regions were selected as the most optimal for discriminating degraded soils from other land cover types. However, for Landsat 8 OLI, only the SWIR (1.560–2.300 μm), NIR (0.845–0.885 μm) region were selected as the best regions. Of the eighteen spectral vegetation indices computed, NDVI and SAVI and SAVI and Global Environmental Monitoring Index (GEMI) were ranked selected as the most suitable for detecting and mapping soil erosion. Additionally, SRTM DEM derived information illustrates that for both sensors eroded areas occur on sites that are 600 m and 900 m of altitude with similar trends observed in both dry and wet season maps. Findings of this work emphasize the importance of free and readily available new generation sensors in continuous landscape-scale soil erosion monitoring. Besides, such information can help to identify hotspots and potentially vulnerable areas, as well as aid in developing possible control and mitigation measures.

1. Introduction
Soil plays a vital role in many economies of the world, particularly in developing countries, such as South Africa, agriculture and forestry forms the backbone of the economy
A review on the outlook of Agriculture in South Africa by the Department of Agriculture Forestry and Fisheries (2010) shows that agriculture accounts for approximately 15.2% of the country’s Gross Domestic Product (GDP). Currently, the sector consisting of 82% (100 million hectares) of the South African land area whereas, in developed countries like Scotland, Europe, 79% of the land area is attributed to agriculture, accounting for 1.8% of the GDP and directly employing over 25,000 people (Scottish Environment Protection Agency, 2001). Despite their economic importance, soils in most developing countries are subject to continuous deterioration due to poor land management practices in place (Ighodaro et al., 2013), leading to severe soil degradation at a phenomenal rate (Pretorius, 1998; Garland et al., 2000; Le Roux et al., 2007). For example, in South Africa over 70% of the land is affected by soil erosion (Le Roux et al., 2007), with an estimated occurrence rate of 8–30 times faster than the rate of regeneration (Baade et al., 2012, Seutloali et al., 2017). Human activities like land clearing for farming, deforestation, overgrazing, or land abandonment coupled with climate change, accelerate the rate of land degradation (Alatorre and Beguería, 2009). In addition, the negative effects of land degradation are not limited to agriculture; they extend to other hydraulic structures, including reservoir sedimentation, which is associated with water treatment costs.

Table 1
LandSat 8 OLI and Sentinel-2 MSI spectral characteristics used in this study.

<table>
<thead>
<tr>
<th>Band#</th>
<th>LandSat 8 OLI spectral bands</th>
<th>Sentinel-2 multispectral bands</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>(Ultra Blue (coastal aerosol))</td>
<td>1 (coastal aerosol)</td>
</tr>
<tr>
<td>2</td>
<td>(Blue)</td>
<td>0.450–0.515</td>
</tr>
<tr>
<td>3</td>
<td>(Green)</td>
<td>0.525–0.660</td>
</tr>
<tr>
<td>4</td>
<td>(Red)</td>
<td>0.630–0.680</td>
</tr>
<tr>
<td>5</td>
<td>(NIR)</td>
<td>0.845–0.885</td>
</tr>
<tr>
<td>6</td>
<td>(SWIR)</td>
<td>1.560–1.660</td>
</tr>
<tr>
<td>7</td>
<td>(SWIR)</td>
<td>2.160–2.300</td>
</tr>
<tr>
<td>8</td>
<td>(Panchromatic)</td>
<td>0.500–0.660</td>
</tr>
<tr>
<td>9</td>
<td>(Cirrus)</td>
<td>1.30–1.38</td>
</tr>
<tr>
<td>10</td>
<td>(THRS)</td>
<td>10.60–11.19</td>
</tr>
<tr>
<td>11</td>
<td>(THRS)</td>
<td>11.50–12.031</td>
</tr>
</tbody>
</table>

*Bold represents the bands used in this study.*
This complicates water availability, which is already a problem, especially in the developing world of Africa. For example, South Africa alone losses approximately two billion rand annually including off-site costs for purification of water whose poor quality is caused by the siltation of dams in eroded surfaces rehabilitation (Department of Environmental Affairs and Tourism, 2006) whereas, other countries like Malawi and Kenya lost about US$2 billion and US$11 billion between 2001 and 2009 periods from land degradation, respectively (Kirui and Mirzabaev, 2015; Voortman et al., 2003; Nkonya et al., 2013). Additionally, Dube et al. (2017) indicated that pressure, due to these processes, has caused serious implications on rural economies; as most of them rely on agriculture as a source of living. Therefore, the need for reliable information on degraded areas and possible vulnerable areas has increased in order to understand the level of degradation and to come up with possible control measures. It is thus important to provide an accurate and
update, as well as detailed soil erosion maps for the former homelands of South Africa as baseline information for monitoring, rehabilitation and control purposes.

Soil erosion, which is considered to be the most critical environmental problems leading to land degradation (Seutloali et al., 2017; Le Roux et al., 2007; Le Roux et al., 2008), has put a burden on national economies. Although great strides have been made in land degradation or soil erosion monitoring, the accuracy of the derived thematic maps remains questionable, especially in developing countries, as the availability of high-resolution data remains a challenge (Luleva et al., 2012). The Landsat series data have so far been the most commonly used remotely sensed data in soil erosion modeling and monitoring. For instance, Dube et al. (2017) assessed the potential of using freely available Landsat series in mapping degradation levels in Eastern Cape, South Africa. They found out that degraded areas can be detected from the Landsat series data. Additionally, the main benefit of Landsat sensor is the multi-temporal aspect (De Jong et al., 1999), although the medium spectral resolution of the scenes presents a limitation. In addition, Lo Curzio and Magliulo (2010), by means of Landsat data series assessed the spatial distribution of degraded areas, with a good spatial accuracy of 97.48% Overall Accuracy (OA) and limited cost in southern Italy. Moreover, Seutloali et al. (2017) mapped the severity of soil erosion, using the 30 m Landsat multispectral satellite data in the former South African homelands of Transkei and the results of the study have indicated that a variety of soil erosion levels (i.e. sheet, slight rills, deep rills, medium gullies to deep gullies) could be detected and mapped.

Even though Landsat data is viewed and appraised as the most reliable and appropriate dataset for soil erosion modeling, especially when compared with other multispectral sensors, such as MODIS or AVHRR which are provided at coarse spatial resolution, its applications remain largely restricted. This dataset’s associated slightly poor spatial resolution largely compromises its fullness in erosion-related studies, especially were farm or catchment level monitoring is required. For example, fine-scale erosional problems remain difficult to document from Landsat ETM + 7 which has since experienced scanline errors resulting in approximately 22% data loss. Most studies that have tested the utility of historical Landsat archival data concluded that the sensor has an outstanding performance in mapping soil erosion at a larger scale. However, the performance of Landsat ETM + 7 remains compromised for farm level monitoring, due to 22% data loss from the scanline errors.

The newly launched Landsat 8 unlike its predecessors, has been highly rated in most of its applications and these include biomass mapping (Yavaşlı, 2016; Zhang et al., 2017), land surface temperature mapping (Adeyeri et al., 2017; Avdan and Jovanovska, 2016), and invasive species mapping (Gavier-Pizarro et al., 2012; Wang et al., 2017). For instance, studies by Vågen et al. (2013) demonstrated that land degradation and soil health in Ethiopia at scales appropriate for management can be assessed, using Landsat-8 bands. Most of the studies associated the successful performance of the sensor to its improved sensing characteristics (Wulder et al., 2008). Landsat 8 OLI also occupies a unique spatial-temporal position in the sense that its bands can better detect and monitor human changes.
in land cover, whereas at the same time having an imaging footprint that is sufficiently large to enable wide-area applications (Wulder et al., 2012). Given the sensor’s recommended performance we, therefore, assume its improved sensing characteristics can aid in determining and mapping, as well as monitoring eroded areas in the former homelands of South Africa – a previous challenging task with broadband sensors like MODIS.

Similar to Landsat 8, Sentinel-2 MSI (Multi-Spectral Instrument) data available at 10 m spatial resolution is considered as one of the most possible solutions to most environmental related challenges in sub-Saharan African, due to its free availability, global footprint, high temporal resolution (± 5days), presence of new multiple bands, previously missing from the previous batch of broadband multispectral sensors, such as ASTER, Landsat 4, 5 and 7, MODIS etc. Studies that have applied sentinel 2 data, for instance in forest mapping, Korhonen et al. (2017) indicated that the sensor’s addition red edge (RE1, RE2, RE3) spectral bands have the capability to improve the accuracy of estimating key plant biophysical variables. They have also shown that the sensor provides the most unique and robust datasets required for understanding environmental problems. For example, Korhonen et al. (2017) have shown that the specific information content of a batch of the broadband multispectral Sentinel-2 sensor may be useful in the monitoring of canopy properties.

Testing the ability of this free-and-readily available remotely sensed datasets in soil erosion monitoring is, therefore critical. So far, the rich information contained in these sensors has not yet been fully exploited in mapping and monitoring eroded areas. This is primarily due to the fact that these sensors were launched recently. Among the different types of the readily-available multispectral remote sensing sensors, archive digital dataset with a wider swath-width (185 km Landsat 8 OLI; 290 km Sentinel-2 MSI) and a 16 (OLI) and 5 (MSI) day temporal resolution makes the two sensors to be perceived as the key primary data sources highly suitable for providing practical or operational regional or district level analysis of eroded areas. This study for the first time sought to assess soil erosion mapping abilities of two new non-commercial multispectral remote sensing data: Landsat 8 OLI and Sentinel-2 MSI in the Sekhukhune district, Limpopo Province of South Africa. To determine the most optimal bands and vegetation indices that can accurately detect and map soil erosion regardless of form. Further, the study wanted to determine whether the areas identified as degraded in the dry season could be as well be detected during the wet season or it was a dry season phenomenon associated with the non-cropping. The study also establishes if the observed soil erosion patterns are a function of elevation or land use.

2.Materials and methods
2.1 Description of the study area
The research was conducted in the Greater Sekhukhune District Municipality, Limpopo Province, South Africa. The area is located at the coordinates 24°23′27.52″S and 29°50′06.83″E and lies across the border of Mpumalanga and Limpopo province. The district comprises approximately 13 528 km² of geographical area, the majority of which is
rural. The district lies in the south-eastern part of the province and is comprised of five local municipalities: Elias Motsoaledi, Ephraim Mogale, Fetakgomo, Makhuduthamaga, and Tubatse. The district is situated in a semi-arid environment, with average annual rainfall ± 560 mm and temperatures showing a moderate fluctuation with average summer temperatures of ± 23 °C (Mpandeli et al., 2015; Stronkhorst et al., 2009). The topography of the area is generally.

Fig. 2. Average reflectance of eroded areas in relation to other land cover types derived using Landsat 8 OLI (b, dry season and c, wet season) Sentinel-2 MSI (a, dry season and d, wet season) (error bars indicate the level of separability).

Table 4
Classification accuracies derived using spectral dataset, spectral vegetation dataset as well as combined dataset.

<table>
<thead>
<tr>
<th></th>
<th>Landsat 8</th>
<th></th>
<th></th>
<th></th>
<th>Sentinel 2</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Analysis 1</td>
<td>Analysis 2</td>
<td>Analysis 3</td>
<td></td>
<td>Analysis 1</td>
<td>Analysis 2</td>
<td>Analysis 3</td>
<td></td>
</tr>
<tr>
<td>PA</td>
<td>UA</td>
<td>PA</td>
<td>UA</td>
<td>PA</td>
<td>UA</td>
<td>PA</td>
<td>UA</td>
<td>PA</td>
</tr>
<tr>
<td>Eroded areas</td>
<td>84.85</td>
<td>73.68</td>
<td>75.76</td>
<td>78.79</td>
<td>80.37</td>
<td>79.31</td>
<td>79.31</td>
<td>79.31</td>
</tr>
<tr>
<td>built-up</td>
<td>63.89</td>
<td>85.19</td>
<td>63.89</td>
<td>74.19</td>
<td>80.56</td>
<td>53.33</td>
<td>66.67</td>
<td>56.67</td>
</tr>
<tr>
<td>cultivated</td>
<td>76.37</td>
<td>96.30</td>
<td>76.47</td>
<td>86.67</td>
<td>73.51</td>
<td>83.33</td>
<td>81.08</td>
<td>85.00</td>
</tr>
<tr>
<td>Grass-shrubland</td>
<td>80.56</td>
<td>78.38</td>
<td>77.78</td>
<td>70.00</td>
<td>83.33</td>
<td>70.59</td>
<td>64.86</td>
<td>67.44</td>
</tr>
<tr>
<td>forest-woodland</td>
<td>82.35</td>
<td>66.67</td>
<td>82.35</td>
<td>66.87</td>
<td>83.24</td>
<td>85.00</td>
<td>77.27</td>
<td>85.59</td>
</tr>
<tr>
<td>water bodies</td>
<td>97.30</td>
<td>92.31</td>
<td>83.31</td>
<td>63.35</td>
<td>92.89</td>
<td>100.00</td>
<td>95.35</td>
<td>100.00</td>
</tr>
<tr>
<td>OA</td>
<td>80.95</td>
<td>75.71</td>
<td>82.86</td>
<td>91.89</td>
<td>94.44</td>
<td>80.00</td>
<td>81.90</td>
<td>83.81</td>
</tr>
<tr>
<td></td>
<td>Landsat 8</td>
<td></td>
<td></td>
<td></td>
<td>Sentinel 2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Analysis 1</td>
<td>Analysis 2</td>
<td>Analysis 3</td>
<td></td>
<td>Analysis 1</td>
<td>Analysis 2</td>
<td>Analysis 3</td>
<td></td>
</tr>
<tr>
<td>PA</td>
<td>UA</td>
<td>PA</td>
<td>UA</td>
<td>PA</td>
<td>PA</td>
<td>UA</td>
<td>PA</td>
<td>PA</td>
</tr>
<tr>
<td>Eroded areas</td>
<td>78.79</td>
<td>76.47</td>
<td>84.85</td>
<td>80.00</td>
<td>87.88</td>
<td>68.97</td>
<td>71.43</td>
<td>64.52</td>
</tr>
<tr>
<td>Build-up</td>
<td>75.00</td>
<td>81.82</td>
<td>75.00</td>
<td>62.79</td>
<td>88.89</td>
<td>63.33</td>
<td>61.29</td>
<td>63.33</td>
</tr>
<tr>
<td>Cultivated</td>
<td>32.35</td>
<td>84.62</td>
<td>50.00</td>
<td>80.95</td>
<td>84.71</td>
<td>69.44</td>
<td>86.21</td>
<td>63.89</td>
</tr>
<tr>
<td>Grass-shrubland</td>
<td>90.00</td>
<td>57.69</td>
<td>82.00</td>
<td>65.08</td>
<td>92.00</td>
<td>55.88</td>
<td>54.29</td>
<td>73.53</td>
</tr>
<tr>
<td>Woodland</td>
<td>65.00</td>
<td>76.47</td>
<td>75.00</td>
<td>83.33</td>
<td>75.00</td>
<td>82.50</td>
<td>71.74</td>
<td>80.00</td>
</tr>
<tr>
<td>Water bodies</td>
<td>89.19</td>
<td>94.29</td>
<td>59.46</td>
<td>73.33</td>
<td>89.19</td>
<td>97.56</td>
<td>97.56</td>
<td>87.89</td>
</tr>
<tr>
<td>OA</td>
<td>73.81</td>
<td>71.43</td>
<td>84.29</td>
<td>97.06</td>
<td>87.89</td>
<td>74.29</td>
<td>73.81</td>
<td>74.29</td>
</tr>
</tbody>
</table>

(\textit{PA} = \textit{Producer Accuracy}, \textit{UA} = \textit{User Accuracy}, and \textit{OA} = \textit{Overall Accuracy}. a) dry season b) wet season.)
2.2. Field data collection
The field survey was conducted from the 26th to 28th of June 2017, coinciding with the dates of remotely sensed data acquired for the study area. Data collection was done by recording coordinates at sub-meter accuracy using GPS device, to validate satellite remote sensing data. Eroded areas were identified during field surveys using random walks and google earth maps of the area. A similar approach was used in collecting data on other major land cover classes in the area, and these included built-up areas, cultivated areas, eroded areas forest-woodland, and grass-shrubland, vegetation cover and water bodies. Land cover classes were identified using visual observation. The vector maps of the study, courtesy of Sekhukhune District together with the aid of google earth were used to navigate to areas affected within the study area. During the field operation, a total of 300 (50 per class) points were recorded, using a Trimble GeoXH 6000 series handheld Global Position System (GPS) at sub-meter accuracy. These GPS points were used in extracting spectral data from the two satellite data sets. Furthermore, photographs of eroded areas and other land cover types were taken, using a handheld camera. During the collection of photographs, GPS coordinates were also recorded and these were used to verify the classified maps.

2.3. Remote sensing data acquisition and preprocessing
Landsat-8 and Sentinel-2 satellite images were used in mapping eroded areas across two seasons (dry and wet season) in the Greater Sekhukhune District Municipality (Table 1). Landsat-8 acquires global moderate-resolution measurements of the Earth’s surface in the visible, near-infrared, shortwave, and thermal infrared. The sensor was launched on the 11th of February 2013, with a combination of two push broom instruments: the Operational Land Imager (OLI) consisting of nine spectral bands (refer to Table 2) and (ii) the Thermal Infrared Sensor (TIRS) which encompasses thermal bands 10 and 11 at a 100 m spatial resolution. Sentinel-2 mission, launched on the 23rd of June 2015, is a land monitoring constellation of two satellites (Sentinel-2a and Sentinel-2b) providing global optical imagery with 13 spectral bands at a 5-day interval, using MSI (Multispectral Imager) instrument.

In this study, cloudless Landsat 8 OLI and Sentinel 2 MSI images were acquired respectively during the 1st of June 2017 and 31st July 2017 for the dry season and between 1st of December 2016 and 31st January 2017 accessed from the USGS Earth Resources Observation and Science (EROS) Centre archive (http://earthexplorer.usgs.gov/). Subsequently, the images were re-projected and mosaicked. Both images were atmospherically corrected, using the dark object subtraction (DOS1) model in QGIS version 2.1.8 software.

2.4. Digital elevation model data
The Shuttle Radar Topography Mission (SRTM)-derived digital elevation model (DEM) was used to generate information on the elevation of the area and this data was used to determine whether the occurrence eroded areas can be explained as a function elevation (slope). This study used SRTM DEM because of its higher spatial resolution (30 m/pixel) corresponding to that of the two sensors and accessibility. The DEM was converted to meters (m) and the elevation of the area under study ranged between 495 m and 2101 m.
DEM rugged ranging from hilly to mountainous with an average altitude of 494 m above sea level. Subsistence or smallholder agriculture accounts for 70% of the farming activities in the district, whilst the other 30% is commercial agriculture (Siambi et al., 2007). Preprocessing was done using ArcGIS tools 10.4.1 software.

![Study area map](http://repository.uwc.ac.za)

Fig. 1. Location of the study site derived, using Sentinel 2 bands.
Fig. 3. Overall classification accuracies for three analysis stages. a) dry season; b) wet season.

Table 5
Deviation of classification accuracies between Landsat 8 OLI and Sentinel-2 MSI

<table>
<thead>
<tr>
<th>Sensor</th>
<th>Parameter</th>
<th>Overall accuracy</th>
<th>Deviation in terms of accuracy (%)</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dry season</td>
<td>Landsat 8 OLI</td>
<td>80.95%</td>
<td></td>
<td>+5.24</td>
<td>-7.15</td>
<td>-1.91</td>
</tr>
<tr>
<td></td>
<td>Spectral indices (SI)</td>
<td>75.71%</td>
<td>-5.24</td>
<td>-</td>
<td>+7.15</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>BI + SI</td>
<td>82.86%</td>
<td>+1.91</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Spectral indices (Slh)</td>
<td>80.00%</td>
<td>-1.9</td>
<td>+3.81</td>
<td>-</td>
<td>-1.91</td>
</tr>
<tr>
<td></td>
<td>BI + SI</td>
<td>81.90%</td>
<td>-1.9</td>
<td>+1.91</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Spectral indices (SI)</td>
<td>83.81%</td>
<td></td>
<td>+3.81</td>
<td>-</td>
<td>-1.91</td>
</tr>
<tr>
<td>Wet season</td>
<td>Landsat 8 OLI</td>
<td>73.81%</td>
<td></td>
<td>-2.38</td>
<td>+10.48</td>
<td>+12.86</td>
</tr>
<tr>
<td></td>
<td>Spectral indices (Slh)</td>
<td>71.43%</td>
<td></td>
<td>-</td>
<td>+2.38</td>
<td>-12.86</td>
</tr>
<tr>
<td></td>
<td>BI + SI</td>
<td>84.29%</td>
<td>+10.48</td>
<td>+12.86</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Spectral indices (SI)</td>
<td>74.29%</td>
<td>+2.38</td>
<td>-</td>
<td>+12.86</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>BI + SI</td>
<td>73.81%</td>
<td>-0.48</td>
<td>+0.48</td>
<td>+</td>
<td>-0.48</td>
</tr>
</tbody>
</table>
2.5. Landsat 8 OLI and Sentinel-2 derived spectral data and vegetation indices

Spectral reflectance values (Table 2) along with selected simple spectral band ratios were applied in this study. The choice of vegetation indices was based on their performance as demonstrated from previous studies (Vaidyanathan et al., 2002; Taruvinga, 2009; Singh et al., 2004) were retrieved from Landsat-8 OLI and Sentinel-2 data sets. The spectral reflectance was extracted from the images, using points collected during field surveys. The extraction was done, using the Hawths spatial analysis tool embedded in ArcGIS 10.4.1 software. Eighteen vegetation indices were computed (Table 2). The indices were selected based on their successful application in the classification and analysis of degraded surface mapping from the highlights made by previous studies and remotely sensed variables for validation (Seutloali et al., 2017; Kwanele and Njoya, 2017; King et al., 2005).

2.6. Statistical data analysis

One-way analysis of variance (ANOVA) was used to test whether there were significant differences between the mean reflectance of the six identified classes. Also, the variable ranking was performed to test the significant performance ($\alpha = 0.05$) of spectral indices. These six classes we tested for a significant difference ($\alpha = 0.05$) using reflected multivariate extracted values. The variable ranking was performed to identify the optimal spectral indices that can detect and discriminate eroded areas from other land cover types. To map eroded areas, three stages of analysis (presented in Table 3) were implemented based on the two sets of variables (bands and indices) derived from the two sensors.
2.7. Image classification and accuracy assessment

The Discriminant Analysis (DA) classification ensemble was used to discriminate eroded lands from other land cover types. DA is a non-parametric classification ensemble, which searches for a linear combination that best discriminates amongst land cover types (Sibanda et al., 2015). Sibanda et al. (2015) and Dube et al. (2017), indicated that this non-parametric classification, converts the reflectance data of land covers at each waveband into several components that account for the difference in reflectance amongst the land cover types. The DA algorithm assumes that the samples are random (Dube et al., 2017), which is the case with land management units samples that were used, hence appropriate for this kind of study. This algorithm provides cross-validated results with variable scores (Eigenvalues) that indicated the strength of a specific function in discriminating eroded surfaces under different land management units.

Eigenvectors also knew as variable scores were produced by DA and used to evaluate the relative contribution of each waveband and vegetation indices to the DA function that optimally discriminated eroded surfaces under different management practices. The DA algorithm applies the Box test (Chi-square asymptotic approximation), Box test (Fisher’s F asymptotic approximation), Mahalanobis distances, Wilks’s Lambda test (Rao’s approximation), and Kullback’s test to test whether within-class covariance matrices were
equal. These tests exhibited significant classification power (P < 0.05). The classification was conducted using XLSTAT for Microsoft Excel 2013 software and confusion matrices were derived. In each confusion matrix, the columns represented the test data, while the rows represented the classes to which each sample was allocated to by the DA classifier (Fig. 1).

To assess the classification accuracy of the results, quantity disagreement, and allocation disagreement was used following its best application from Sibanda et al. (2016) as a way of separating data into training tested data also recommended by Pontius and Millones (2011) as the successor of Kapa Statistic. Quantity disagreement is a sum of least perfect matches between the training (70%) and the testing (30%) reflectance datasets of each land management practice. Precisely, the quantity disagreement follows when the column total of a management practice class deviates from the row total of that class in a confusion matrix. To estimate the extent of the difference between Landsat 8 OLI data accuracy and that of Sentinel-2 MSI agreement between classification results and ground truth data was measured using the producer accuracy (PA), user accuracy (UA) and overall accuracy (OA) generated from the confusion matrices. These two parameters were used in accuracy assessment, as suggested by Pontius and Millones (2011).

3. Results

3.1. Discrimination of eroded areas from other land cover types

Fig. 2 shows the differences amongst averaged spectral bands values or curves of eroded areas and other land cover types for the wet and dry seasons. ANOVA results reveal significant differences (p < 0.05), which implies eroded areas can be discriminated from other land cover types.
It can be observed that sentinel-2 MSI could optimally discriminate eroded areas from other land cover types, using bands located in the NIR (0.785–0.900 um), and red edge (0.698–0.785 um) and SWIR (1.565–2.280 um) regions for both dry and wet season (Fig. 2b and d). Similarly, Landsat 8 OLI illustrated the ability of SWIR (1.560–2.300 um) region, followed by NIR (0.845–0.885 um) region to optimal in discriminate eroded areas eroded areas. However, the visible regions (from 0.433–0.578 um) from both sensors show close or inseparable reflectance curves, which implies weaker discrimination capabilities for both dry and wet season. But the visible (0.450–0.680 um) region of Landsat 8 still became worse than that of Sentinel-2. Moreover, it can be observed that when using Landsat 8 OLI data, eroded areas could be discriminated from the built-up and grass-shrublands in the SWIR region, when compared with the red and green region of the visible range on the electromagnetic (EM) spectrum. However, when using Sentinel 2 MSI data, more eroded areas were discriminated from the built-up and grass-shrublands in the red edge region as compared with the SWIR region in the wet season than dry season. When using Landsat 8 OLI data, the dry season shows that most of the eroded areas were discriminated from the woodland-forest in the SWIR region of the EM spectrum but there were some overlaps in the visible region and NIR of EM spectrum (Fig. 2). On the other hand, when using Sentinel 2 MSI, most of the woodland-forest unit was differentiated from the eroded areas.
in the SWIR, VRE, NIR and visible (red, green) regions of the EM spectrum except for the blue region of the spectrum (Fig. 2).

### 3.2 Spectral indices performance

Variable ranking (P < 0.05) was performed and of eighteen spectral indices used, results reveal that Landsat 8 derived NDVI and SAVI had the highest ranking in discriminating eroded areas from other land cover types. Using Sentinel-2, SAVI and GEMI had the highest ranking.

### 3.3 Image classification

#### 3.3.1 Analysis 1: soil erosion classification using spectral bands as an independent dataset

Table 4 illustrates overall classification accuracies derived from Landsat 8 OLI and Sentinel 2 as an independent data set for two seasons (dry and wet seasons). The results indicated that the spectral reflectance information of Landsat 8 OLI alone produced slightly good classification results, with an overall accuracy (OA) of 80.95% over Sentinel-2 in the dry season but Sentinel outperformed Landsat 8 with ± 0.48 OA in the wet season. The results showed the superiority of Landsat 8 over Sentinel-2 in the dry season by achieving an overall classification accuracy of ± 80 while Sentinel performed better in the wet season (Table 4). Both sensors yielded good user and producer accuracies of above 50% in all land cover types of the study area for both dry and wet season. In agreement to this, eroded areas had user accuracy (UA) of 79.31%, using Sentinel-2, higher than that of Landsat 8 (magnitude of 5.63%) in the dry season (Fig. 3). In the wet season, eroded areas specifically produced UA of 76% (Sentinel) and 66% using Landsat 8. Over the six classes, better soil erosion PA and UA were achieved from both sensors for the dry season data, whereas in the wet season, a UA of 96.30% using Landsat 8 OLI.
3.3.2. Analysis 2: soil erosion mapping classification using vegetation indices as an independent data

For the dry season, Landsat 8 OLI derived vegetation indices performed slightly less in detecting and mapping eroded areas with an OA of 75.71%. On the other hand, Sentinel-2 derived spectral indices had a considerably high OA of 81.90% (Table 4). For the wet season, derived vegetation indices performed slightly less in detecting and mapping eroded areas with an OA of less than 75.71% when compared with the dry season. When compared to classification results based on spectral bands (analysis I), the OA decreased by ± 5.24% for Landsat 8 OLI datasets and increased by ± 5.9% for Sentinel-2 MSI datasets in the wet season (Table 4). Furthermore, user and producer accuracies for Landsat 8 OLI also slightly decreased with most of the classes ranging from ± 2.78% to ± 18.92% whereas
Sentinel-2 maintained an increase of ± 2.58% to ± 14.70% in most of the classes (Table 5). The results attained using Landsat 8 OLI and Sentinel-2 derived vegetation indices separately produced slightly high classification accuracies when compared to the use of spectral bands as an independent dataset. Sentinel-2 achieved better accuracy results than Landsat 8 with user accuracy of 79.31% in validating eroded areas and whilst Landsat 8 OLI scored 78.13% of user accuracy, respectively. The difference with regards to the magnitude of performance using vegetation indices as a separate data set for detecting and mapping eroded areas was ± 6.19% between the two sensors. Landsat 8 OLI thus had slightly lower accuracies when compared to the 10-m Sentinel 2 data.

3.3.2 Analysis 3: soil erosion classification using a combination of spectral bands and spectral vegetation indices

As illustrated in Table 4 the combined dataset (i.e. spectral bands and vegetation indices) achieved high classification results, when compared to analysis 1 and analysis 2. For instance, in the dry season using the combined data set, Landsat 8 OLI and Sentinel-2 produced high overall accuracies of 82.86% and 83.81%, respectively. Landsat 8 OLI sensor performed slight less than the Sentinel-2 sensor. Eroded areas were classified as a
producer and user accuracy results of about 80% demonstrating an increase of more than 10% for both sensors, particularly in dry season. Furthermore, during the wet season eroded areas were classified with high UA and PA accuracies of ± 70% for both the sensors (Table 4), although less than that of the dry season. The results obtained from combined dataset significantly improved classification accuracies, although Sentinel-2 produced slightly good classification accuracy of 85% when compared Landsat 8. In general, both sensors yielded the best performance in classifying eroded areas based on combined spectral bands and spectral vegetation indices (Fig. 3).

### 3.3.3 Derived soil erosion maps

Fig. 4 shows the derived maps of the eroded areas. Overall it can be observed that both Landsat 8 and Sentinel-2 sensors have depicted a similar pattern in the distribution of eroded areas and the trend is identical across the wet and dry seasons. The derived maps indicate that the central part of Sekhukhune is more eroded when compared to other areas. However, comparatively, Landsat 8 (Fig. 4a and c) demonstrate high levels of erosion when compared to Sentinel 2 (Fig. 4b and d). For instance, Fig. 5 demonstrates the area in percentage of each classified land cover class in comparison of the two sensors. Landsat 8 showed that most classes are more than 5% when compared to Sentinel-2 in dry season while with more than 6% in a wet season (Fig. 5).

Fig. 6 illustrates eroded areas as detected by the two sensors. Fig. 3 further shows the zoomed areas for clear visualization of the eroded areas. Extensive levels of eroded areas can be observed in Fig. 6(a)–(f). It can be observed that the majority of the land disturbances through erosion are concentrated along villages and agricultural fields. Fig. 7 as it also compares the two sensors; it can be perceived that sentinel-2 managed to detect eroded surfaces than Landsat 8 from the classified images. Moreover, it can be observed that the major disturbances in Sekhukhune are mainly related to croplands. Fig. 7(b) shows evidence of disturbed areas along agricultural farms. In addition, photographs were taken during field, observation indicates that some of the eroded surfaces are detected along villages (Fig. 8b and d). From the observation, it is noticed that some of the eroded surfaces were detected along mines as one anthropogenic activity (Figs. 7(d) and 8(c)) and also some of the open areas like abandoned croplands appear to experience severe erosion.

### 3.3.5. Relationship between eroded areas and elevation

Derived erosion layers were extracted and overlaid on the elevation map (Fig. 9). It can be observed that much of the area between 600 m and 900 m exhibited high levels of erosion when compared to low-lying and mountainous areas (Fig. 10). From the map, it can as well be noted that high proportions of eroded areas occurred in areas with an elevation between 600 and 1500m. Moreover, areas in the extremes (495 m–600 and 1800 m– 2101 m) areas show less rate of soil erosion (Fig. 10).
4. Discussion

The main essence of this study was to test the effectiveness of two new generation sensors in detecting and mapping the spatial distribution of eroded areas amongst other land cover types in the former homelands of Sekhukhune, South Africa. Accurate mapping of eroded areas provides a critical input dataset required for soil conservation strategies. Specifically, the study aimed at assessing soil erosion mapping abilities of two new non-commercial multispectral remote sensing data: Landsat 8 OLI and Sentinel-2, as well as determine the optimal bands and indices that can detect and map soil erosion in former homelands. Also, the study sought to find out if the variations in terms of soil erosion can be explained using variations in elevation.

Although the provision of remote sensing multispectral sensors provides an attractive alternative for mapping and monitoring eroded areas, one of their primary challenges is the inability to reduce the mapping error. Results of this study demonstrated the potential of the newly launched Sentinel-2 MSI in detecting and mapping eroded areas with overall accuracy results that are slightly higher than that of Landsat 8 OLI. Combined spectral vegetation indices and extracted Sentinel-2 MSI spectral reflectance information were used to accurately discriminate eroded surfaces from other land cover types and high
classification accuracies in terms of percentages (OA, PA, and UA) were observed as compared to Landsat 8 OLI. The same results were also observed when only extracted spectral information were used with Sentinel-2 MSI slightly outcompeting Landsat 8 OLI. However, when spectral vegetation indices were used, Landsat 8 OLI performed slightly better than Sentinel-2 MSI. Overall, Sentinel-2 MSI outperformed Landsat 8 OLI.

It can be observed that the use of the combined dataset improved the classification accuracies, were Sentinel-2 outperformed Landsat 8. The unique performance of the Sentinel 2 imagery can be attributed to the uniqueness of sensor design. For instance, Sentinel 2 is a pushbroom scanner, with numerous new and strategically positioned bands that provide unique information about the earth’s surface (Clark and Kilham, 2016; Guidici and Clark, 2017). For example, Sentinel-2’sNIR (0.785–0.900 um), red edge (0.698–0.785 um) and SWIR (1.565–2.280 um) region of the EM spectrum have been depicted as the most important bands providing separability windows for discriminating eroded surfaces from other land cover types. On the other hand, Landsat 8 OLI does not cover the red edge portion of the EM spectrum, hence slightly weaker performance. In that regard, the lack of information from the red edge region could also explain the unsatisfactory performance of Landsat 8 OLI in this study. The study by, Korhonen et al. (2017) has shown that the lack of red edge bands in most multispectral sensors downplays their potential in mapping environmental properties.

Fig. 9. Maps showing the relationship between elevation and eroded areas (a. dry & c. wet; b. dry & d. wet season for Landsat 8 and Sentinel 2 respectively).
Moreover, since accuracies increased in both sensors after combining spectral bands and vegetation indices, the results of the current study, therefore, clearly indicate the importance of combining vegetation indices with spectral bands in the discrimination of eroded areas from other land cover types. This combination agrees with the result from the study made by Sibanda et al. (2016) where they demonstrated that the high classification accuracies exhibited using vegetation indices and wavebands in spectrally discriminating grasses grown under different management practices. Similarly, Matongera et al. (2017) integrated the spectral bands and derived vegetation indices yielding the best overall classification accuracy (80.08% and 87.80% for Landsat 8 OLI and Worldview-2 respectively) in detection and mapping the spatial configuration of bracken fern weeds.

The results of this study show significant variations of the spatial distribution of eroded areas derived using the two sensors (Landsat 8 and Sentinel-2). For example, it can be observed from the results that the Sentinel-2 sensor with high spectral bands can depict eroded surfaces from other land cover types as when compared to the Landsat 8. Furthermore, spectral separability results indicated slight weaknesses of Landsat 8 when compared to Sentinel-2. However, this can be linked to the fact that images with low spectral reflectance have a challenge of mixed pixels. Although Landsat 8 demonstrated moderately poor quality (i.e. poor radiometric, spatial, spectral characteristics), particularly for detecting spatial occurrence of soil erosion, it holds a good record, especially large-scale mapping, as it is acknowledged by researcher on similar studies (Millington and Townshend, 1984; Whitlow, 1986; Vrieling, 2006; Zhou et al., 2008; Taruvinga, 2009; Seutloali et al., 2017; Dube et al., 2017). Based on the digital classification of Landsat thematic mapper and JERS-1 data, the study by Metternicht and Zinck (1998) detected and mapped different soil erosion feature in Bolivia and concluded that the synergy of Landsat TM provides a unique combination that allows more accurate identification of eroded areas. Furthermore, the study by Seutloali et al. (2017), in the former South African homelands of Transkei has indicated the effectiveness of utilizing Landsat data as a free and readily available multispectral remote sensing in mapping soil erosion levels. Even though Landsat 8 seems to yield comparatively good results than Sentinel-2 its spatial resolution makes it difficult to map eroded areas or vulnerable areas especially at plot or farm level due to mixed pixels.

The plausible classification results can also be attributed to Discriminant Analysis (DA) algorithm applied in mapping eroded areas, performed best results. The studies made by Sibanda et al. (2015) and Dube et al. (2017) indicated the potential of using this algorithm over other classification techniques. Unlike the traditional classification approaches, such as the maximum likelihood classification algorithms, Sibanda et al. (2015) have shown that the DA classification ensemble has the potential to spectrally detect and discriminate complex classes. Furthermore, Dube et al. (2017) stated that this classification ensemble is repeatable and simple to use and it is illustrated by its usefulness across a wide range of research areas, including natural resources management, electronics, finance, and accounting. The major hindrance of this algorithm, however, is that it requires data sets that are normally distributed of which according to Dube et al. (2017), is not usually the case.
The study further showed that soil erosion varies with a change in elevation. It can be observed that most the mapped eroded areas are found within areas that are not in slightly higher elevation. For example, the results showed that much of the eroded areas occurred in highly elevated areas i.e. between 600 m and 900 m when compared to low-lying or flat areas. Slightly elevated areas are likely to experience soil erosion due to runoff during the rainy season (Balaguer-Puig et al., 2017). Rain water has limited time to infiltrate into the soil as the areas will be a slope, hence more runoff and vice-versa.

![Graphs showing eroded areas in different elevation ranges.](http://repository.uwc.ac.za)

This observation is also confirmed by previous studies that have found out that highly elevated experience more runoff and consequently due to high gravitational force. For example, the study by Mondal et al. (2017) used open source DEMs of different resolution and ascertained their effects on soil erosion which mostly were detected in less steeped areas and reports that the DEMs gives better results with less uncertainty.
Future research studies should, therefore, focus on using Sentinel 2 MSI, as it provides a better alternative for mapping and monitoring at various scales given it’s the high resolution and other related sensing characteristics. Also, the free and readily available nature of the sensors makes it the most optimal solution for mapping soil erosion problems in sub-Saharan Africa, which is currently characterized with limited resources for accurate mapping of soil erosion for management and monitoring of environmental problems. In a nutshell from a land management side, the results of this study are vital and relevant to related stakeholders, i.e. environmental managers, soil scientists, and agriculturalists, as well as policymakers. The findings provide significant information on location and extent of the affected areas, and this will help in decision making, rehabilitation or remedial purposes.

5. Conclusions
The main aim of the study was to assess the effectiveness of Landsat 8 OLI and Sentinel-2 in mapping the spatial distribution of eroded areas in Sekhukhune district, Limpopo Province of South Africa. The findings of this work have shown that Sentinel 2 offers free, effective and time efficient of acquiring information on the spatial distribution of eroded areas. Sentinel 2 produced an overall classification accuracy of more than 80% whilst Landsat 8 with more than 75% of all tested analytical stages. The integration of Landsat 8 and Sentinel 2 derived raw spectral bands and vegetation indices significantly ($\alpha = 0.005$) improved the detection and mapping accuracies. The study further showed that soil erosion varies with a change in elevation. For example, much of the eroded areas were occur in elevated areas when compared to low-lying or flat surfaces. In summary, the findings of this study have shown that the new generation of readily available multispectral remote sensors together with discriminant analysis classification ensemble presents a potential for mapping and monitoring the spatial occurrence of eroded areas in resources constraints areas across different scales.

Acknowledgements
The Authors thank the National Aeronautics and Space Administration (NASA) through the United States Geological Survey (USGS) for the provision of Landsat 8 OLI and Sentinel-2 MSI images. Authors would also like to prolong their word of thanks to various people who helped with the fieldwork which includes Mr. Dhau I. and Mr. Thamaga K.H. They also thank the Department of Geography and Environmental studies in University of Limpopo and Risk and Vulnerability Science Centre (RVSC) for the Logistic support.
References


Dube, T., Mutanga, O., Sibanda, M., Seutoloali, K., Shoko, C., 2017. Use of Landsat series data to analyze the spatial and temporal variations of land degradation in a dispersive soil environment: a case of King Sabata Dalindyebo local municipality in the Eastern Cape Province, South Africa. Phys. Chem. Earth Parts A/B/C.


Siambi, M., Simalenga, T., Mpandeli, N.S., Liphadzi, K., Mailula, N., Mkhari, J., Ramugondo, R., 2007. Increased food security and income in the Limpopo Basin through integrated crop, water and soil fertility options and public-


