Manchester Metropolitan University

Symeonakis, Elias and Korkofigkas, Antonis and Vamvoukakis, Giorgos and Stamou, Giorgos and Arnau-Rosalen, Eva (2020) Deep Learning Monitoring of Woody Vegetation Density in a South African Savannah Region. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives, 43. pp. 1645-1649. ISSN 1574-0846

https://e-space.mmu.ac.uk

DEEP LEARNING MONITORING OF WOODY VEGETATION DENSITY IN A SOUTH AFRICAN SAVANNAH REGION

E. Symeonakis [1]*, A. Korkofigkas [2], G. Vamvoukakis [2], G. Stamou [2], E. Arnau- Rosalén[1]

[1] Department of Natural Sciences, Manchester Metropolitan University, Manchester, M1 5GD, UK
[2] School of Electrical and Computer Engineering, National Technical University of Athens, Athens 15780, Greece
(E.Symeonakis@mmu.ac.uk)

Commission III, WG III/7

KEY WORDS: Fractional woody vegetation cover, bush encroachment monitoring, South Africa, Landsat, spatiotemporal metrics, deep learning, CNN, U-Net

ABSTRACT:

Bush encroachment in African savannahs has been identified as a land degradation process, mainly due to the detrimental effect it has on small pastoralist communities. Mapping and monitoring the extent covered by the woody component in savannahs has therefore become the focus of recent remote sensing-based studies. This is mainly due to the large spatial scale that the process of woody vegetation encroachment is related with and the fact that appropriate remote sensing data are now available free of charge. However, due to the nature of savannahs and the mixture of land cover types that commonly make up the signal of a single pixel, simply mapping the presence/absence of woody vegetation is somewhat limiting: it is more important to know whether an area is undergoing an increase in woody cover, ever if it is not the dominant cover type. More recent efforts have, therefore, focused in mapping the fraction of woody vegetation, which, clearly, is much more challenging. This paper proposes a methodological framework for mapping savannah woody vegetation and monitoring its evolution though time, based on very high-resolution data and multi-temporal medium-scale satellite imagery. We tested our approach in a South African savannah region, the Northwest Province (>104,000 km^2), 0.5m-pixel aerial photographs for sampling and validation and Landsat data.

1. INTRODUCTION

Savannah ecoregions are important ecosystems with high biodiversity. They provide a number of ecosystem services, e.g. grazing for pastoralist communities, or the supply of fuelwood, amongst others. Over the last years, savannahs have been under pressure from human activities, exacerbated by climate change, with dramatic shifts in savannah vegetation distribution and, consequently, alterations of their function. Bush encroachment, fuelwood overexploitation, increased carbon emissions, loss of biodiversity are processes that are often being flagged as of concern for most savannah ecoregions worldwide (Symeonakis et al. 2018). Therefore, monitoring their extent and composition is of importance and directly links to a number of UN Sustainable Development Goals (United Nations, 2015) and the target to achieve Land Degradation Neutrality (LDN) by 2030 (von Maltitz, et al., 2019).

Earth observation technologies are the only viable approach for achieving this due to the spatial coverage they provide, the ever-increasing access to open-source data archives and the computational and technological improvements. However, due to their structural properties and composition, traditional 'hard classification' mapping approaches are not helpful in successfully monitoring savannah condition (Higginbottom et al. 2018). Rather, the ability to assess the contribution of each of the main savannah vegetation components and the evolution of those through time is needed. Over recent years, a limited number of studies have addressed this issue with varying degrees of success (Ludwig, et al., 2019).

Here, we address the issue of accurately mapping the fractional cover of one of the main constituents of savannah ecoregions, woody vegetation, by testing a methodological framework

incorporating a deep learning approach in a southern African savannah context.

2. MATERIALS & METHODS

2.1 Study area

The study area is the Northwest Province (NWP) of South Africa. It covers an area of 104,882 km2 (Figure 1).
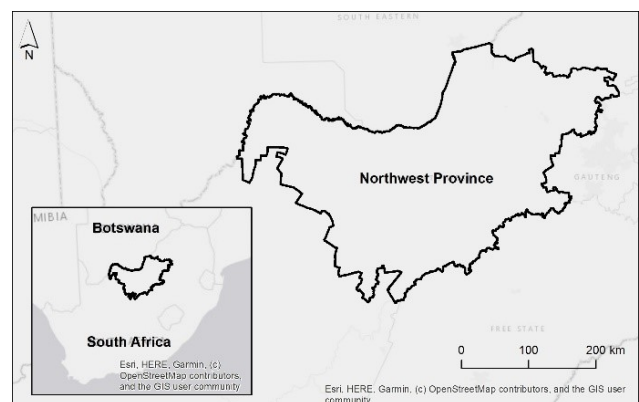

Figure 1. The study area of the Northwest Province, and its location within South Africa

Temperatures range from 17° to 31 °C in the summer and from 3° to 21 °C in the winter. Annual rainfall is ~360 mm (~14 in), with nearly all of it falling during the summer wet months, between October and April (Wikipedia, 2016). Around 70% of the Province falls within the Savannah Biome (Bushveld vegetation). The remainder falls within the Grassland Biome,

---

*   Corresponding author

which contains a variety of grasses typical of arid regions. Ten different vegetation types are found, mostly belonging to the thornveld, bushveld or savannah grassland categories (Walmsley and Walmsley, 2002). The vegetation variation follows the respective east-west variation in the climatic characteristics.

2.2 Datasets

Training and validation data came from the open-access 0.5m-pixel RGB aerial photography of the South African mapping agency, the National Geospatial Institute (NGI). The NGI makes the aerial photographs available as an ArGIS basemap dataset (ESRI, 2019). For our area of study, the photos were taken between 2009 and 2013.

We used all available Level 1 Landsat 5, 7 and 8 from the USGS archive from 1986 to 2019 with less than 80% cloud cover. This amounted to 16,456 images. The Landsat data were accessed through the data catalogue of Google Earth Engine (GEE; https://earthengine.google.com; Gorelick et al. 2017).

2.3 Methods

On the aerial photography data, we manually annotated 12048 points with classification in three classes, woody vegetation, non-woody vegetation, and non-vegetation. For 2525 of these points we estimated visually the woody cover percentage in 10% intervals (i.e. in 11 classes from 0% to 100%) for an area of 90m x 90m, corresponding to a 3x3 pixel area on the equivalent Landsat data.

We started our analysis from the epoch centred around 2010. We first trained a Random Forest Classifier on the 12048 aerial image data points, inputting the RGB colour and the Visible Vegetation Index (Liaw and Wiener, 2002) and outputting one of the three classes with an 88% accuracy. The model was then used to generate predictions for every point of a 180x180 pixel aerial image area centred around each of the 12048 points. The percentage of pixels classified as woody vegetation was used as an estimation of fractional woody cover of the equivalent 3x3 pixel (90m x 90m) Landsat area. The per-pixel predictions also constitute a 3-class semantic segmentation mask for the aerial image.

We then trained a deep learning image segmentation model (i.e. per pixel classifier) based on the U-Net Convolutional Neural Network architecture (Ronneberger et al., 2015), using the aerial images as input and the predicted masks as labels. The U-Net we employed was 3-layer deep with 32x32 resolution at the narrowest convolutional layer. It reached 93% accuracy on a held-out test dataset of 1200 images. The U-Net model was finally used to generate another set of 12048 fractional woody cover estimations.

Finally, we trained a woody coverage regression model. The input consisted of the Landsat bands and spatio-temporal variability metrics derived from the Landsat data in the five years between 2009 and 2013 of a 3x3 pixel Landsat area (Symeonakis et al., 2018). The Landsat metrics were calculated using Google Earth Engine. To find the best model architecture we performed a 5-fold cross validation grid search between various configurations of Random Forest Regressors, Gradient Boosted Regression Trees from the XGBoost (Chen et al., 2016) library and Multilayer Neural Networks.

All configurations were trained on 3 datasets: the manually annotated 2525 point dataset (A) and the two 12048 point

datasets estimated by the Random Forest Classifier and U-Net models above (datasets B and C). The model was then applied to 7 other epochs of the Landsat bands and metrics centred around 1988, 1993, 1998, 2003, 2008, 2013 and 2018.

3. RESULTS & DISCUSSION

The regression model results for the 3 datasets are depicted in Table 1. The best performing models were the ones trained using the masks from the U-Net model (dataset C).

	RF - A	RF - B	RF - C
MAE	0.2723	0.1847	0.1953
MSE	0.0955	0.0524	**0.0511**
	XG - A	XG - B	RF - C
MAE	0.2736	0.1888	0.1953
MSE	0.1055	0.05916	0.0581
	NN - A	NN - B	NN - C
MAE	0.2840	0.1786	**0.1693**
MSE	0.1629	0.0831	0.0873

Table 1. Mean Average Error and Mean Squared Error for the Random Forest (RF), Gradient Boost Tree (XG) and Neural Network (NN) regression models tested on the 3 datasets (A, B, and C)

Using the best performing regression model based on lowest Mean Squared Error (RF-C) we calculated the woody cover percentage for different areas and epochs. Figure 2 is the outcome for the entire study area, while Figure 3a is the same for an area on the border with Botswana, close to the town of Bray (Figure 3a). Figure 3b is the area around Bray as seen on the NGI 0.5m pixel aerial photos.

Figure 4 depicts the results for four out of the eight epochs of our study (1993, 2008, 2013 and 2018), as displaying all epochs would restrict the ability to visually assess the outcome. A zoomed in area in the west of the Province is also shown where the pattern of the increase until 2010 and the decrease in the last epoch is more evident.
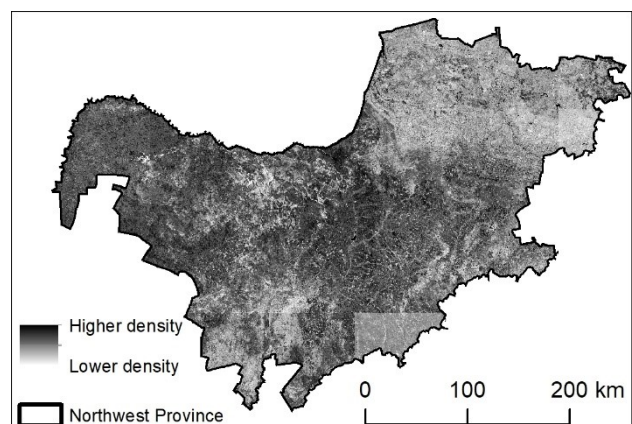

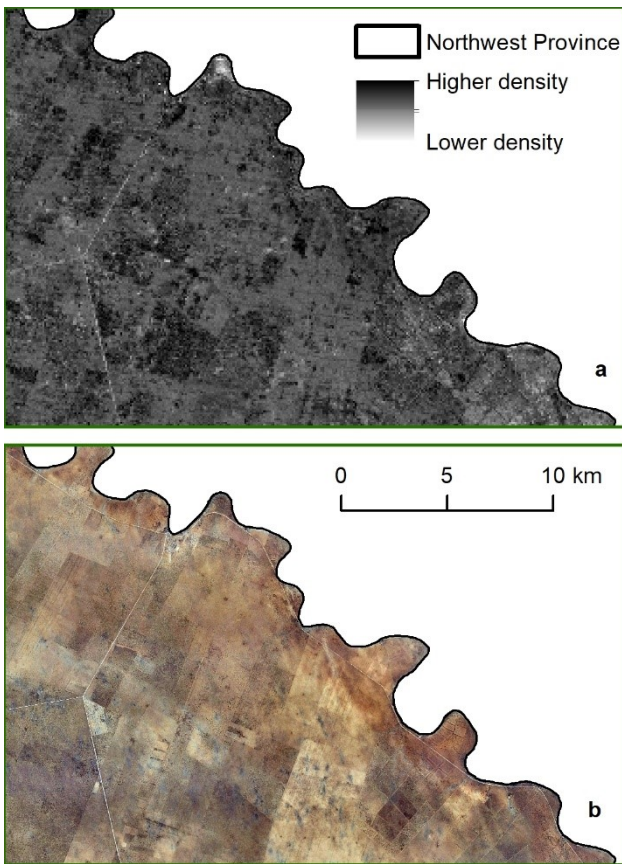Figure 2. Fractional woody cover as estimated for the year 2010 for the Northwest Province

Figure 3. (a) Fractional woody cover as estimated for the year 2010 for an area near the town of Bray. (b) The area near Bray as seen on the NGI aerial photos
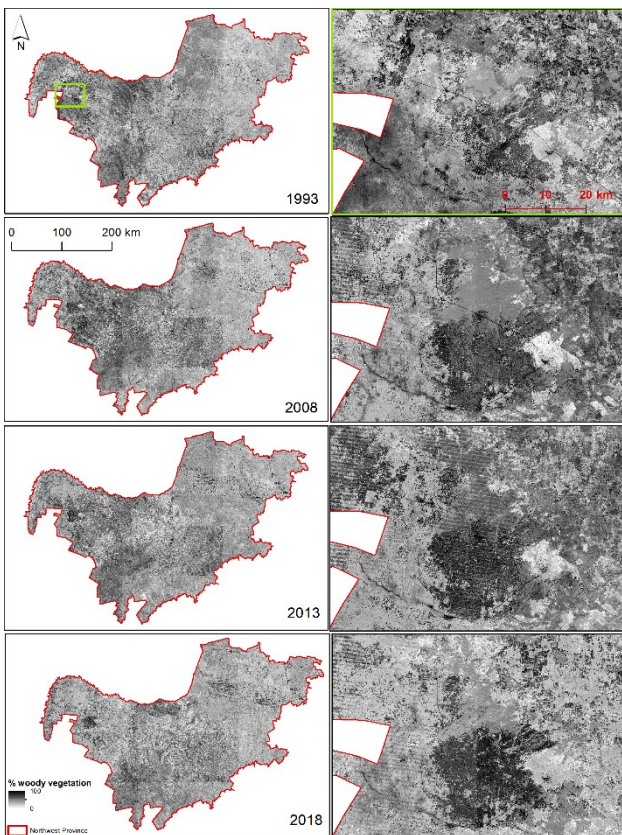

Figure 4. Fractional woody vegetation results for 1993, 2008, 2013 and 2018 (i.e. four out of the eight epochs of our study)

The multi-temporal results show an increasing trend in woody cover densities throughout most of the study area from 1988 to 2013. A decreasing trend is then observed in the last epoch (centred around 2018).

Figure 5 shows the overall trend in fractional woody vegetation cover over the 30 years of the period of study (1988-2018) for the entire Province (Figure 5a) and for an area on the border with Botswana, near the town of Tosca (Figure 5b).
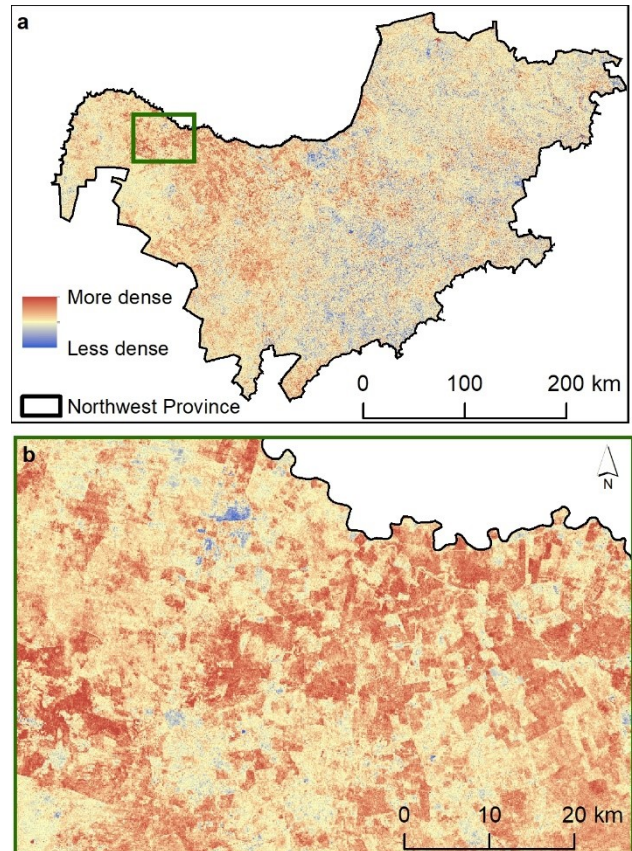

Figure 5. Fractional woody vegetation change from 1988 to 2018, over (a) the study area; (b) over an area on the border with Botswana

Results showing an increasing trend corroborate some of our earlier findings in the region (Higginbottom et al. 2018; Symeonakis et al. 2014) and are in line with the scientific consensus that, due to climate warming and the subsequent effect of carbon fertilisation, the increasing trend in woody densification and encroachment might be continuous in space and time (Eldridge et al., 2011; Ward, 2005).

Areas that exhibit a decreasing trend appear to be in disagreement with this notion. However, it must be noted that a number of bush encroachment control measures are in place in the Province, with varying degrees of success. Such measures range from reactive management approaches (including manual, mechanical control, and chemical control) to a combination of proactive land management practices (such as grazing control, fire and post-burn management (Dreber, et al., 2019; Harmse, etal., 2016; Turpie et al., 2019).

In this preliminary study, we used a limited amount of annotated data, occupying only a small fraction of the study area. Future work will address this issue by employing the machine learning technique of active learning (Settles, 2009): after selecting a set

of images in which the U-Net model has the lowest average prediction confidence, we will involve a human annotator to fine-tune the U-Net model on new annotations, repeating the process until no significant improvement in accuracy between iterations is achieved.

Finally, a planned field visit later in the year will focus on validating the latest results with in-situ data.

4. CONCLUSION

Long-term monitoring of woody savannah cover is needed to enhance the understanding of broad-scale changes in woody vegetation and the possible relationship between such changes and ecosystem resilience or degradation. The Landsat archive has been the workhorse for characterizing land cover using optical data given its long-standing archival imagery and high spatial resolution. We employed a deep learning approach to spatiotemporal metrics calculated with GEE from thousands of Landsat multi-temporal imagery to map the evolution of the fractional woody vegetation in an area of South Africa. We conclude that our approach is beneficial compared to previous attempts (e.g. using machine learning techniques) and should therefore be preferred for such monitoring endeavours in dryland environments.

REFERENCES

Bastin, J.-F.; Berrahmouni, N.; Grainger, A.; Maniatis, D.; Mollicone, D.; Moore, R.; Patriarca, C.; Picard, N.; Sparrow, B.; Abraham, E.M.; et al. The extent of forest in dryland biomes. Science 2017, 356, 635–638.

Chen T, Guestrin C. Xgboost: A scalable tree boosting system. InProceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining 2016 Aug 13 (pp. 785-794).

Dreber, N.; van Rooyen, S.E.; Kellner, K. One savanna, many shapes: How bush control affects the woody layer in the southern Kalahari, *South African Journal of Botany*, 2019, 125, 511-520.

Eldridge, D.J.; Bowker, M.A.; Maestre, F.T.; Roger, E.' Reynolds, J.F.; Whitford, W.G. Impacts of shrub encroachment on ecosystem structure and functioning: towards a global synthesis. *Ecology Letters*, 2011, 14(7), 709-722.

ESRI, 2019. The South African National Geospatial Institute (NGI) 50cm Colour Imagery https://www.arcgis.com/home/item.html?id=9d01fa9041264cb2 83c353a5a613c81e available at https://webgis.esri-southafrica.com/arcgis/rest/services/NGI/RSA_NGI_AERIAL/I mageServer (4 May 2020)

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D.; Moore, R. Google Earth Engine: planetary-scale geospatial analysis for everyone. Remote Sens. Environ. 2017, 202, 18–27.

Hansen, M.C.; Egorov, A.; Roy, D.P.; Potapov, P.; Ju, J.; Turubanova, S.; Kommareddy, I.; Loveland, T.R. Continuous fields of land cover for the conterminous United States using Landsat data: First results from the Web-Enabled Landsat Data (WELD) project. Remote Sens. Lett. 2011, 2, 279–288.

Harmse, C.J.; Kellner, K.; Dreber, N. Restoring productive rangelands: A comparative assessment of selective and non-selective chemical bush control in a semi-arid Kalahari savanna. *Journal of Arid Environments*, 2016, 135, 39-49.

Higginbottom, T.; Symeonakis, E.; Meyer, H.; van der Linden, S. Mapping Woody Cover in Semi-arid Savannahs using Multi-seasonal Composites from Landsat Data. ISPRS J. Photogramm. Remote Sens. 2018, 139, 88-102. https://doi.org/10.1016/j.isprsjprs.2018.02.010

Kennedy, R.E.; Andréfouët, S.; Cohen, W.B.; Gómez, C.; Griffiths, P.; Hais, M.; Healey, S.P.; Helmer, E.H.; Hostert, P.; Lyons, M.B. Bringing an ecological view of change to Landsat-based remote sensing. Front. Ecol. Environ. 2014, 12, 339–346. [Green Version]

Knorn, J.; Rabe, A.; Radeloff, V.C.; Kuemmerle, T.; Kozak, J.; Hostert, P. Land cover mapping of large areas using chain classification of neighboring Landsat satellite images. *Remote Sensing of Environment*, 2009, 113, 957–964.

Kuemmerle, T.; Erb, K.; Meyfroidt, P.; Müller, D.; Verburg, P.H.; Estel, S.; Haberl, H.; Hostert, P.; Jepsen, M.R.; Kastner, T.; et al. Challenges and opportunities in mapping land use intensity globally. *Current Opinion in Environmental Sustainability*, 2013, 5, 484–493.

Lehmann, E.A.; Wallace, J.F.; Caccetta, P.A.; Furby, S.L.; Zdunic, K. Forest cover trends from time series Landsat data for the Australian continent. *International Journal of Applied Earth Observation and Geoinformation*, 2013, 21, 453–462.

Liaw A, Wiener M. Classification and regression by randomForest. R news. 2002 Dec 3;2(3):18-22.

Ludwig, M.; Morgenthal, T.; Detsch, F.; Higginbottom, T.P.; Valdes, M.L.; Nauß, T.; Meyer, H. Machine learning and multi-sensor based modelling of woody vegetation in the Molopo Area, South Africa. *Remote Sensing of Environment*, 2019, 222, 195-203.

Maestre, F.T.; Eldridge, D.J.; Soliveres, S.; Kéfi, S.; Delgado-Baquerizo, M.; Bowker, M.A.; García-Palacios, P.; Gaitán, J.; Gallardo, A.; Lázaro, R. Structure and functioning of dryland ecosystems in a changing world. *Annual Review of Ecology, Evolution, and Systematics*, 2016, 47, 215–237.

Mathieu, R.; Naidoo, L.; Cho, M.A.; Leblon, B.; Main, R.; Wessels, K.; Asner, G.P.; Buckley, J.; Van Aardt, J.; Erasmus, B.F.N.; et al. Toward structural assessment of semi-arid African savannahs and woodlands: The potential of multitemporal polarimetric RADARSAT-2 fine beam images. *Remote Sensing of Environment*, 2013, 138, 215–231.

Millennium Ecosystem Assessment. Ecosystems and human well-being: Synthesis; World Resources Institute: Washington, DC, USA, 2005.

Naidoo, L.; Mathieu, R.; Main, R.; Wessels, K.; Asner, G.P. L-band Synthetic Aperture Radar imagery performs better than optical datasets at retrieving woody fractional cover in deciduous, dry savannahs. *International Journal of Applied Earth Observation and Geoinformation*, 2016, 52, 54–64.

Poulter, B.; Frank, D.; Ciais, P.; Myneni, R.B.; Andela, N.; Bi, J.; Broquet, G.; Canadell, J.G.; Chevallier, F.; Liu, Y.Y. Contribution of semi-arid ecosystems to interannual variability of the global carbon cycle. *Nature,* 2014, 509, 600–603.

Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. InInternational Conference on Medical image computing and computer-assisted intervention 2015 Oct 5 (pp. 234-241). Springer, Cham.

Schneibel, A.; Frantz, D.; Röder, A.; Stellmes, M.; Fischer, K.; Hill, J. Using Annual Landsat Time Series for the Detection of Dry Forest Degradation Processes in South-Central Angola. Remote Sens. 2017, 9, 905.

Scholes, R.J.; Walker, B.H. An African Savanna: Synthesis of the Nylsvley Study; Cambridge University Press: Cambridge, UK, 1993.

Senf, C.; Leitão, P.J.; Pflugmacher, D.; van der Linden, S.; Hostert, P. Mapping land cover in complex Mediterranean landscapes using Landsat: Improved classification accuracies from integrating multi-seasonal and synthetic imagery. *Remote Sensing of Environment*, 2015, 156, 527–536.

Settles B. Active learning literature survey. University of Wisconsin-Madison Department of Computer Sciences; 2009.

Symeonakis, E.; Higginbottom, T.P.; Bush Encroachment Monitoring Using Multi-Temporal Landsat Data and Random Forests. *International Archives of Photogrammetry and Remote Sensing Spat. Inf. Sci.*, 2014, XL-2, 40, 29–35. https://doi.org/10.5194/isprsarchives-XL-2-29-2014

Symeonakis, E.; Higginbottom, T.P.; Petroulaki, K.; Rabe, A. Optimisation of savannah land cover characterisation with optical and SAR data. *Remote Sensing*, 10(4), 499. https://doi.org/10.3390/rs10040499

Turpie, J.; Botha, P.; Coldrey, K.; Forsythe, K.; Knowles, T.; Letley, G.; Mills, A. Towards a Policy on Indigenous Bush Encroachment in South Africa, Anchor Environmental Consultants, Cape Town, 2019

UNCCD. Global Land Outlook; UNCCD: Paris, France, 2017; p. 338.

United Nations. Transforming our world: The 2030 Agenda for Sustainable Development. Resolution Adopted by the General Assembly on 25 September 2015. In A/RES/70/1; United Nations: New York, NY, USA, 2015; Available online: https://www.un.org/ga/search/view_doc.asp?symbol=A/RES/70/1&Lang=E (4 May 2020).

Von Maltitz, G.P.; Gambizo, J.; Kellner, K.; Rambau, T.; Lindeque, L.; Kgope, B. Experiences from the South African land degradation neutrality target setting process. Environmental Science & Policy, 2019, 101, 54-62.

Walmsley, D. and Walmsley, J., 2002. North West Province State of the Environment Report – Overview 2002. Available online at: www.nwpg.org.za/soer (31 March 2016)

Ward, D. Do we understand the causes of bush encroachment in African savannas? *African Journal of Range and Forage Science*, 2005, 22(2), pp.101-105.

Wessels, K.; Prince, S.; Malherbe, J.; Small, J.; Frost, P.; Van Zyl, D. Can human-induced land degradation be distinguished from the effects of rainfall variability? A case study in South Africa. *Journal of Arid Environments*, 2007, 68, 271–297.

Wessels, K.J.; van den Bergh, F.; Scholes, R.J. Limits to detectability of land degradation by trend analysis of vegetation index data. *Remote Sensing of Environment*, 2012, 125, 10–22.

Wulder, M.A.; Masek, J.G.; Cohen, W.B.; Loveland, T.R.; Woodcock, C.E. Opening the archive: How free data has enabled the science and monitoring promise of Landsat. *Remote Sensing of Environment*, 2012, 122, 2–10.