zef
Center for
Development Research
University of Bonn

ZEF-Discussion Papers on
Development Policy No. 218

Valerie Graw, Carsten Oldenburg, and Olena Dubovyk

Bush Encroachment Mapping for Africa: Multi-scale analysis with remote sensing and GIS

Bonn, July 2016

Valerie Graw, Carsten Oldenburg, and Olena Dubovyk, Bush Encroachment Mapping for Africa: Multi-scale analysis with remote sensing and GIS, ZEF – Discussion Papers on Development Policy No. 218, Center for Development Research, Bonn, July 2016, pp. 45.

ISSN: 1436-9931

Published by:
Zentrum für Entwicklungsforschung (ZEF)
Center for Development Research
Walter-Flex-Straße 3
D – 53113 Bonn
Germany
Phone: +49-228-73-1861
Fax: +49-228-73-1869
E-Mail: zef@uni-bonn.de
www.zef.de

The author[s]:
Valerie Graw, Center for Remote Sensing of Land Surfaces (ZFL), University of Bonn.
Contact: valerie.graw@uni-bonn.de
Carsten Oldenburg, Center for Remote Sensing of Land Surfaces (ZFL), University of Bonn.
Contact: carsten.oldenburg@uni-bonn.de
Olena Dubovyk, Center for Remote Sensing of Land Surfaces (ZFL), University of Bonn.
Contact: odubovyk@uni-bonn.de

Acknowledgements

We would like to thank the German Federal Ministry for Economic Cooperation and Development (BMZ) for funding this study as well as Johannes Schmidt and Yohannes Ayanu for providing field data from South Africa and Ethiopia for the development of the here presented approach.

Abstract

Bush encroachment (BE) describes a global problem severely affecting savanna ecosystems in Africa. Invasive species and woody vegetation spread out in areas where they are not naturally occurring and suppress endemic vegetation, mainly grasses. Livestock is directly affected by decreasing grasslands and inedible invasive species which are a result of the process of BE. For many small scale farmers in developing countries livestock represents a type of insurance particular in times of crop failure and droughts. Among that, BE is also becoming an increasing problem for crop production. Studies on the mapping of BE have so far only focused on smaller regions using high-resolution data and aerial photography. But they rarely provide information that goes beyond the local or national level. In our project, we aimed at a continental-wide assessment of BE. For this, we developed a process chain using a multi-scale approach to detect woody vegetation for the African continent. The resulted map was calibrated with field data provided by field surveys and experts in Southern and Eastern Africa. Supervised classification linked field data of woody vegetation, known as BE, to the respective pixel of multi-scale remote sensing data. The regression technique was based on random forests, a machine learning classification and regression approach programmed in R. Hotpots of woody vegetation were further overlaid with significant increasing Normalized Difference Vegetation Index (NDVI) trends which can refer to BE. Secondly, the probability of BE occurrence based on possible identified causes such as fire occurrence, mean annual precipitation rates, soil moisture, cattle density and CO_2 emissions was analyzed. By this, possible areas for BE occurrence based on their pre-conditions and risk factors were identified.

This approach includes multiple datasets derived from earth observation data to detect BE – a severe and ongoing global problem – at the continental level. Within the study´s duration of seven months, a method to upscale field data to a larger level could be developed. Nevertheless, improvement is needed to provide a reliable continental map on BE. Especially the integration of more field data will be needed which is currently under consideration. The identification of woody vegetation and the probability of its occurrence can help to prevent further ecosystem degradation. Moreover, sustainable land management strategies in these areas can be focused to support pastoralists and their livelihoods in rural areas.

Keywords: bush encroachment, remote sensing, multi-scale analysis, probability map, random forests, regression trees

Content

List of Figures vi

List of Tables vi

1. Introduction 1

2. Study Area 4

3. Data and Methods 6

    3.1 Remote Sensing Data 6

    3.2 Training Data – Field Data 7

4. The two-step approach: Analysis 8

    4.1 Woody Vegetation Map – A multi-scale analysis (1st Step) 9

        4.1.1 Data processing 9

        4.1.2 NDVI Trend analysis 12

        4.1.3 Results: Multi-Scale Analysis 13

    4.2 Bush encroachment/woody vegetation: Probability Map (2nd Step) 20

        4.2.1 Causes of Bush Encroachment 20

        4.2.2 Logistic regression modeling 28

        4.2.3 Results: Probability Mapping for Bush Encroachment 28

5. Discussion 34

6. Conclusion and Outlook 37

References 39

List of Figures

Figure 1 Land Cover Africa. 4
Figure 2: Two-Step approach for bush encroachment detection in Africa 8
Figure 3: Multi-scale Analysis 10
Figure 4: Process chain for the detection of bushes/woody vegetation in Africa 11
Figure 5: Prediction for woody vegetation probability in Africa 13
Figure 6: Woody Vegetation Map based on predicted data for 2007 and 2009 14
Figure 7: Overlap of NDVI increasing trends (trend analysis for: 1982-2013) and woody vegetation prediction 15
Figure 8: Example for R^2 and MSE with 500 regression trees for the South African up-scaling. 17
Figure 9: Landsat and MODIS Prediction Model for 2007 and 2009. 19
Figure 10: Predicted results versus validation data for the Ethiopian field dataset from 2013 19
Figure 11: Vertical resource partitioning by Walker and Noy-Meir 1982 21
Figure 12: Mean Annual Precipitation in Africa 22
Figure 13: Soil Moisture for 2010 24
Figure 14: Annual Sum Fire Occurrence 25
Figure 15: CO2 Emission Intensity per Pixel 26
Figure 16: Cattle Density of Africa 27
Figure 17: Distribution of woody vegetation and non-woody vegetation in percentage in the different land cover zones according to GlobCover 2009 30
Figure 18: ROC Curves for four possible models chosen via logistic regression modeling 32
Figure 19: Probability of bush encroachment based on the prediction for woody vegetation of 2007 and 2009 including three main impact variables: precipitation, soil moisture and cattle density 33

List of Tables

Table 1: Overview of remote sensing based field data 7
Table 2: Country Statistics on overlapping woody vegetation prediction and significant NDVI increase. 16
Table 3: R^2 and MSE for the Landsat and MODIS regression for 2007, 2009 and 2013 18
Table 4: Number of Points used for training and testing the prediction model. 18
Table 5: Input variables for logistic regression modeling. 29
Table 6: Logistic Regression Modeling for woody vegetation prediction. 30
Table 7: ROC Statistics. 31

1. Introduction

Bush encroachment (BE) is defined as the “suppression of palatable grasses and herbs by encroaching woody species” (Ward, 2005, 101). Grassland cover is decreasing as soils favor woody vegetation represented by shrubs and bushes over grasses due to several factors including livestock or climatic variables such as increasing rainfall. Besides “bush encroachment” the expansion of bushes is also known as “shrub expansion” (Naito & Cairns, 2011), “grassland-forest-transition” (Bai et al., 2004) or “woody vegetation cover change” (Fensham, Fairfax & Archer, 2005). The awareness of increasing bushes especially in Southern Africa started in the 1980s. During this time “bush encroachment” became a commonly used term (Welz, 2013). The expansion of woody vegetation cover over grasses is a global problem and highly diverse in terms of factors triggering this process but also with regard to occurrence across the globe. It represents a loss in soil fertility and thereby also an indicator of land degradation which is developing fast (MEA 2005; Dube, 2007).

The African continent is facing this problem especially in the Savanna ecosystems (Hudak & Wessman, 2001). It creates an imbalance of trees and grasses, while also describing a decrease in biodiversity as well as a decrease in the carrying capacity of an ecosystem (Kruger, 2002; de Klerk, 2004). The most affected ecosystems are savannas, often described as instable systems (Sankaran et al., 2005). But also desert and mesic grasslands and the Arctic tundra are facing BE (Naito & Cairns, 2011). With regard to decreasing grasslands also livestock and rural pastoralists depending on these areas are affected and therefore included in the ongoing discussion of this problem. While nearly half of the population in Africa lives in rural areas and around 70% of them are living in poverty natural resources are of high importance to sustain a living – especially with regard to crop production or livestock (IFAD, 2010; Nachtergaele et al., 2010). Livestock represents a type of insurance for the poor, particularly in times of crop failure or droughts (Speranza et al., 2008).

In South Africa BE already affects 10-20 million hectares of land (Ward, 2005). The co-existence of trees and grasses describes the characteristic of savanna ecosystems while BE refers to a meta-stability (Scholes & Archer, 1997). The increasing of woody vegetation and the decreasing of palatable grassland at the same time has multiple negative effects (Welz, 2013). Besides the suppression of grasslands and thereby food for herbivores it also impacts agricultural productivity and biodiversity in multiple ways (Ward, 2005; Kraaij & Ward, 2006; Wigley et al., 2009). For example the cape vulture, a bird species endemic in Southern Africa, is no longer breeding in Namibia as most open land is covered with shrub and tree cover due to BE (Bamford et al., 2007). As they need long flat areas to run and takeoff in the air this becomes more difficult with increasing woody vegetation (Welz, 2013). The thorny bushes moreover are not edible for most animals. In South Africa e.g. cheetahs are restricted in their hunting habits due to BE occurrence in their former hunting areas (Meik et al., 2002; Purchase et al., 2007).

South Africa and Namibia in particular are heavily faced by increasing woody vegetation in flat areas and grassland (Welz, 2013). Southern Africa is represented widely in studies on BE detection and management strategies to combat this process (Trollope, 1980; Ward, 2005; Kraaij & Ward, 2006; Wigley et al., 2009). Besides Africa also other continents are affected such as Australia, where foliage cover increased by around 11% within the last thirty years (Donohue et al., 2013), Northern America where especially mesquite – a *prosopis* species – is spreading (Grover & Musick, 1990; van Auken, 2000) and parts of Asia such as China (Peng et al., 2013). Among that a study by Naito and Cairns (2011) reported BE in the Arctic tundra.

BE is common in desert and mesic grasslands and Savannas but none of these ecosystems are clearly defined (Naito & Cairns, 2011). Similar difficulties arise when searching for causes triggering BE. While one factor can be triggering BE within one ecosystem it might hinder the process in another. As Savanna ecosystems are among the most vulnerable to the expansion of woody vegetation variables impacting this system with regard to BE were in focus. Besides increasing livestock factors like rainfall and soil moisture play a key role. Fire occurrence is important in Savanna ecosystems. Its absence is reported to trigger shrubland expansion (Scholes & Archer, 1997). Recent research promotes increasing CO_2 levels as cause for BE (Bond & Midgley, 2012).

Remote sensing is an outstanding tool to detect land cover change and its status of land over large areas. Earth observation data and techniques in general represent the only effective approach to map BE over large areas (Symeonakis & Higginbottom, 2014). A pixel – the smallest unit of a remote sensing raster dataset – contains the information for the user. Here, unique spectral characteristic can be used for classification and regression techniques. Identified land cover, the pixel value respectively, can then be linked to semantic information such as woody vegetation cover.

Many studies have already been carried out aiming at the mapping of BE at local and/or national scales in Africa. Here, focus regions and countries are Namibia (e.g. Meik et al., 2002; Wiegand et al., 2005; Symeonakis & Higginbottom, 2014), South Africa (e.g. Hudak & Weeman, 2001) and Botswana (e.g. Moleele et. al, 2002) in Southern Africa, and especially Ethiopia (e.g. Angassa & Oba, 2009; Ayanu et al., 2014) in Eastern Africa. But research on detection and mapping of BE so far has only been carried out on rather small scales. A regional or even continental mapping approach does not yet exist.

In this study a two-step approach is conducted. The first step includes the development of a method that provides the detection of woody vegetation using a machine learning classification approach that is trained with high resolution data in two different study sites. The BE map is calibrated with ground truth data provided by experts in Southern and Eastern Africa. By up-scaling location specific information to different levels of remote sensing imagery – 30 m resolution with Landsat images and 500 m resolution with Moderate Resolution Imaging Spectroradiomater (MODIS) data – a map shows the potential and actual areas of woody vegetation on the African continent This map represents a first result of possible up-scaling techniques for BE mapping. The second output provides a map on the probability of woody vegetation occurrence based on trigger factors of BE. Data on possible causal factors for BE were collected and furthermore modeled with woody vegetation prediction in Africa to identify areas at risk for BE. Remote sensing data and methods as well as analysis in a GIS build the main methods used. The approach is a first attempt to detect areas affected by BE but also to identify areas where BE is most likely to occur in the future based on logistic regression modeling. With further calibration it can support BE analysis in regions where further research and policy actions are needed.

2. Study Area

Land Cover of the African continent is highly diverse which results from different climatic and ecological preconditions. Figure 1 shows land cover in Africa which ranges from humid areas covered by rainforest in Central Africa to non-vegetated areas in the Sahara desert and marks the two field data locations. Grass- and shrubland can be observed in Savanna ecosystems, the Sahel zone and large patterns in Southern Africa, including Namibia and South Africa. Savannas are known for their tree-grass interaction and co-dominance which are widely discussed (Sankaran et al., 2005; Ward et al., 2012). About 20% of the earth´s land surface is covered by savanna ecosystems (Sankaran et al. 2005). They are located in tropical and subtropical regions with an alternation of wet and dry seasons (Scholes & Archer, 1997). The term savanna is not clearly defined by certain thresholds in temperature or rainfall. Although Ward et al. (2012) distinguished between an arid and humid Savanna type.

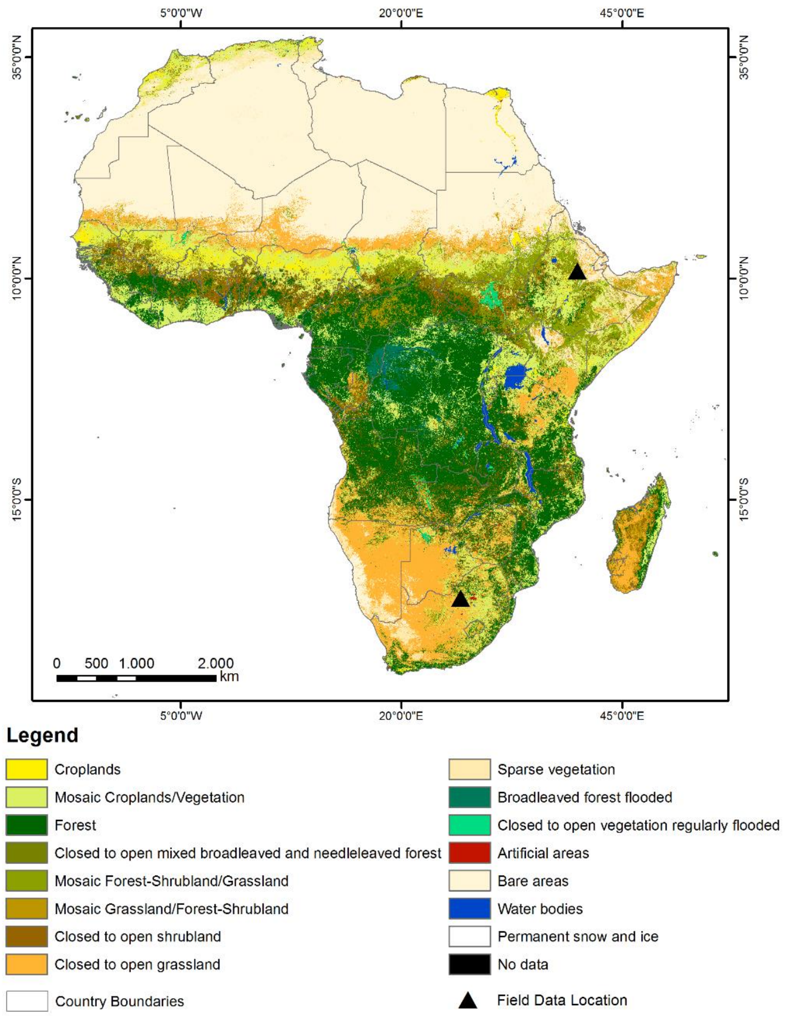

Figure 1: Land Cover Africa. Data: ESA Globcover Project, led by MEDIAS-France/Postel. Visualized and partly reclassified by authors. Data and description available via: http://due.esrin.esa.int/page_globcover.php, Arino et al. 2012

African savannas and grasslands are most commonly affected by BE (Ward, 2005; Kgosikoma & Mogotsi, 2013). In general Africa´s ecosystems are facing multiple disturbances such as wildfire and mega-herbivores (Bond, 2005). The woody vegetation which is increasing and suppresses grasses due to BE is mostly not edible or simply denied by livestock which defines the problem for pastoralists (Lesoli et al., 2013; Negasa et al., 2014).

According to different agro-ecological zones (AEZ) defined by different temperature regimes and rainfall patterns which favor different land uses also a variety of land use types take place across Africa. While areas in humid and sub-humid areas with sufficient rainfall favor crop cultivation grass- and shrubland areas are mainly suitable for livestock production and therefore pastoralists agricultural production is rather low.

Namibia and South Africa but also Botswana are often reported to be heavily affected by BE. With regard to Figure 1 these countries are characterized by large grass- and shrublands. Namibia e.g. has lost 26 Mio hectares of land due to BE (Welz, 2013). Nearly the whole country is affected but especially central parts of Namibia (Meik et al., 2002; Joubert et al., 2008). But also areas characterized here with “sparse vegetation” are known to be prone to BE including Northwestern Kenya and the northeastern parts of Ethiopia where one of our study sites was locate. The second study site was located in Northern Cape, South Africa. In both cases field data was available using aerial photos and high-resolution imagery. Both study sites are described in more detail in chapter 3.2.

3. Data and Methods

Field data from two study regions, Southern and Eastern Africa, was taken into account to develop a process chain for the mapping of BE in Africa. In Southern Africa the *acacia* type is commonly invading an area (Oldeland et al., 2010) while the invasive species *prosopis* (*prosopis juliflora*) was particularly reported in South Africa and Eastern Africa (Mwangi & Swallow, 2005; Zachariades, Hoffman & Roberts, 2011; Ayanu et al. 2014). For these two regions, from South Africa and Ethiopia, datasets were provided that contained classified woody vegetation based on known BE patterns.

In general, BE seems to not have as many negative stated effects in western Africa compared to e.g. Southern Africa. According to literature research only a very few studies which conduct research on BE in western Africa could be identified. An up-scaling approach therefore could help to get some more insights in woody vegetation cover and its dynamics in western Africa.

3.1 Remote Sensing Data

The geospatial location of the training sides led to the identification of related Landsat scenes. The scenes were chosen based on the time the field data were collected. By this the biggest overlap of ground truth information with remote sensing imagery was assured. If disturbances were noticed in the respective datasets, such as e.g. high cloud coverage, the next possible dataset was chosen. As Landsat 7 images are highly distorted data of Landsat 5 was integrated for the first up-scaling. Landsat was launched in 1972 and provides data with 30 m resolution. The sensor is able to provide an image for the exact same point every 16 days (USGS, 2013). Landsat 5, launched in 1984, comes with 4 spectral bands which allow multiple analysis and classification techniques.

MODIS multispectral data are used for the second and last level up-scaling. This data type is also used for covering whole Africa. As the surface reflectance data of MODIS with a spatial resolution of 250 m (MOD09GQ) only consist of two bands (Band 1 (RED) and Band 2 (NIR)) between 0.6 µm to 0.9 µm bandwidths, MODIS Surface Reflectance Data with 500 m resolution (MOD09GA) had to be included. The 500 m resolution MODIS data covers seven bands. In addition to channel 1 and 2 (RED and NIR) the bandwidths of the spectral channels 3 to 7 reach from 0.4µm to 2.1 µm.

3.2 Training Data – Field Data

Field data from two areas in Africa were available located in Eastern and Southern Africa. They served as a first input to develop a process chain for BE detection in Africa. Details about the location and data will be described in the following.

South Africa – Northern Cape Province, close to Kuruman

The data used for regression was based on aerial photography. Research on BE mapping was embedded in the project RCR of the University of Bonn and generated by Johannes Schmidt. The study areas were located in the Kalahari Bushveld near the City of Kuruman (Northern Cape). To classify bushes on six different farms RGB images (aerial photos with Red, Green and Blue bands) of 2007 with 0.5 m spatial resolution and panchromatic images of 2009 with 0.75 m resolution were taken into account. We used a binary classification of these images to implement a regression model with random forests to detect the possibility of BE in the Landsat and further on the MODIS image. Three different farms were included with in total 186,537 training pixels in the Landsat image for 2007 and 195,679 pixels for 2009. For MODIS 140,562 pixel were included for 2007 and 141,275 for 2009.

Ethiopia – Baadu case study site

Ayanu et al. (2014) mapped the invasive species *prosopis juliflora* in Ethiopia for different years (2000, 2005, 2010 and 2013). We used the latest dataset generated for the year 2013 for further classification and up-scaling. As this dataset has a 15 m resolution and is coarser than the field data used for South Africa, only the second upscale, from ASTER based data to MODIS, could be approached. This area is usually partly covered by clouds throughout the year which made it difficult to find completely cloud free images. Furthermore, several MODIS tiles cover the research area. Thus an image mosaic was created for the regression approach to upscale the classified dataset. In total 3,862 training pixels for MODIS were available.

Remote sensing data for up-scaling was always linked to the observation period of the field data. A detailed description of the data with respective dates is given in Table 1.

Table 1: Overview of remote sensing based field data

Year	Country	Study Site (Year)	Landsat (Path/Row)	Modis
2007	South Africa (1)	2007	Landsat 5 (173/079) 2007 178	MOD09A1 2007 169
2009	South Africa (2)	2009	Landsat 5 (173/079) 2009 151	MOD09A1 2009 153
2013	Ethiopia	2013	Landsat 8 (167/053) 2013 136	MOD09A1 2013 137

4. The two-step approach: Analysis

Detection of BE, woody vegetation respectively, on a continental scale was conducted with a two-step approach comprising (i) the development of a prediction map of woody vegetation in Africa overlapping with significant increasing vegetation trends and (ii) the generation of a probability map which shows areas where BE is likely occurring based on the prediction map out of Step 1 as resulted outcome and trigger factors of BE using logistic regression analysis. Within the first step a multi-scale process chain was developed in R. Here, field data on BE was used as training sample as they represented bushes as the outcome of BE processes. Based on this, an up-scaling process was subsequently conducted to the continental level. The processing included the application of a random forests regression model, a machine learning regression approach which represents an effective tool for prediction (Breiman, 2001). For this analysis two level of remote sensing data with different spatial resolution and extent were used for up-scaling: Landsat data for local level analysis and MODIS data for the continental coverage. The second step includes the generation of a BE probability map using logistic regression modeling. Possible causal factors were identified that might trigger BE according to literature research. Figure 2 gives an overview about input data, processing steps and the two outcomes.

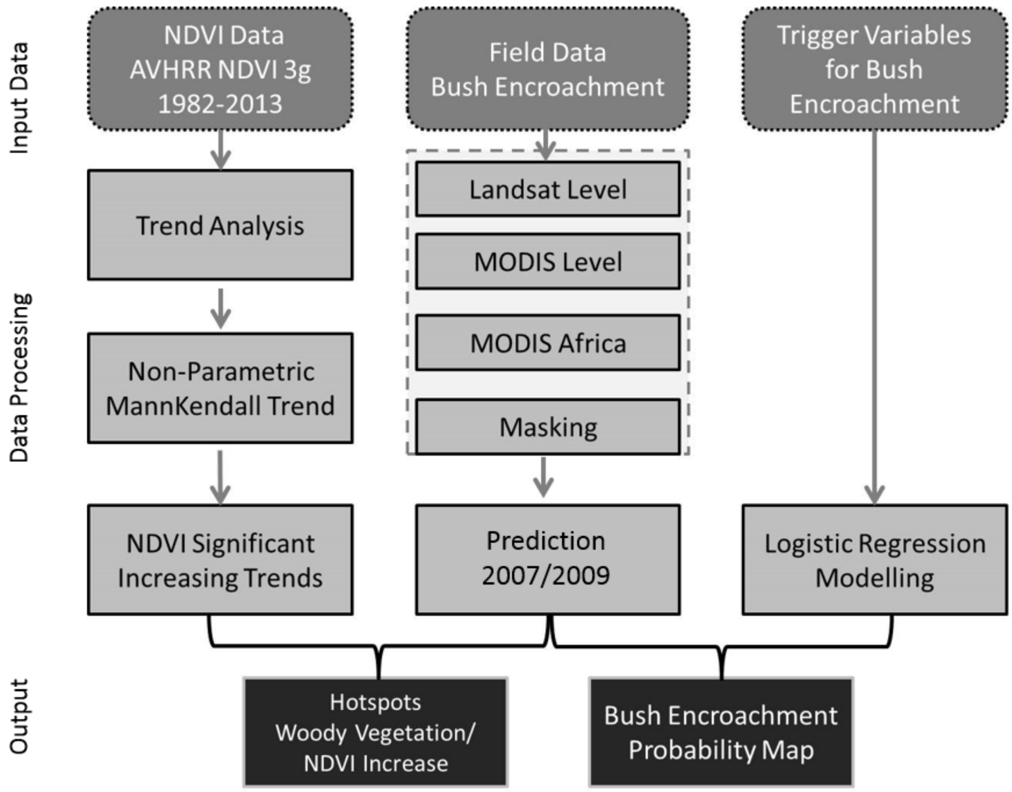

Figure 2: Two-Step approach for bush encroachment detection in Africa

4.1 Woody Vegetation Map – A multi-scale analysis (1st Step)

Already classified field data on BE was provided in two areas of Africa: Southern and Eastern Africa. Southern Africa is represented by field data from the Northern Cape in South Africa, where BE on different farms was mapped with aerial and panchromatic images. In East Africa field data could be provided for Ethiopia which focused on the detection of *prosopis juliflora*.

After detection of woody vegetation in the Landsat scenes for a first upscale and in the MODIS tiles for the second upscale the developed process chain was used to predict BE on the continental scale.

Multispectral data of MODIS (MODA9) as used for the second up-scaling was processed for whole Africa. All data used for the up-scaling matched the observation period of the field data. Field data was available from 2007 and 2009 for South Africa and for 2013 from Ethiopia. The developed process chains were applied to the respective datasets. Predictions based on field data could therefore be made for the three different years of the input datasets. Overlapping areas out of all predictions were identified in a GIS.

4.1.1 Data processing

Figure 3 shows the multi-scale analysis using one field dataset from South Africa. The classified dataset in (a) is based on aerial photos and shows one farm where a binary classification was conducted with "bushes" and "other". For the field data location, the respective Landsat scene (b) was identified which ideally covers the similar period of time when the aerial photos were taken. Only scenes with no or minimal cloud cover were taken into account to minimize disturbance and false classification within the image. Based on the area of the Landsat scene further on the respective MODIS tile was identified (c), ideally also within the same period of time and with minimal to no cloud cover. The "MODIS-level" was finally approached to map whole Africa using the developed processing chain (e).

The regression approach was used to derive the information of every pixel and afterwards link its spectral information and values to semantic information. Based on spectral characteristics a regression was conducted resulting in a binary classification: "bush" and "no bush". Based on this the percentage coverage of bushes for every Landsat pixel within the classification area was computed. The regression technique was based on random forests (RF) regression trees (RT), a machine learning regression approach (Breiman, 2001).

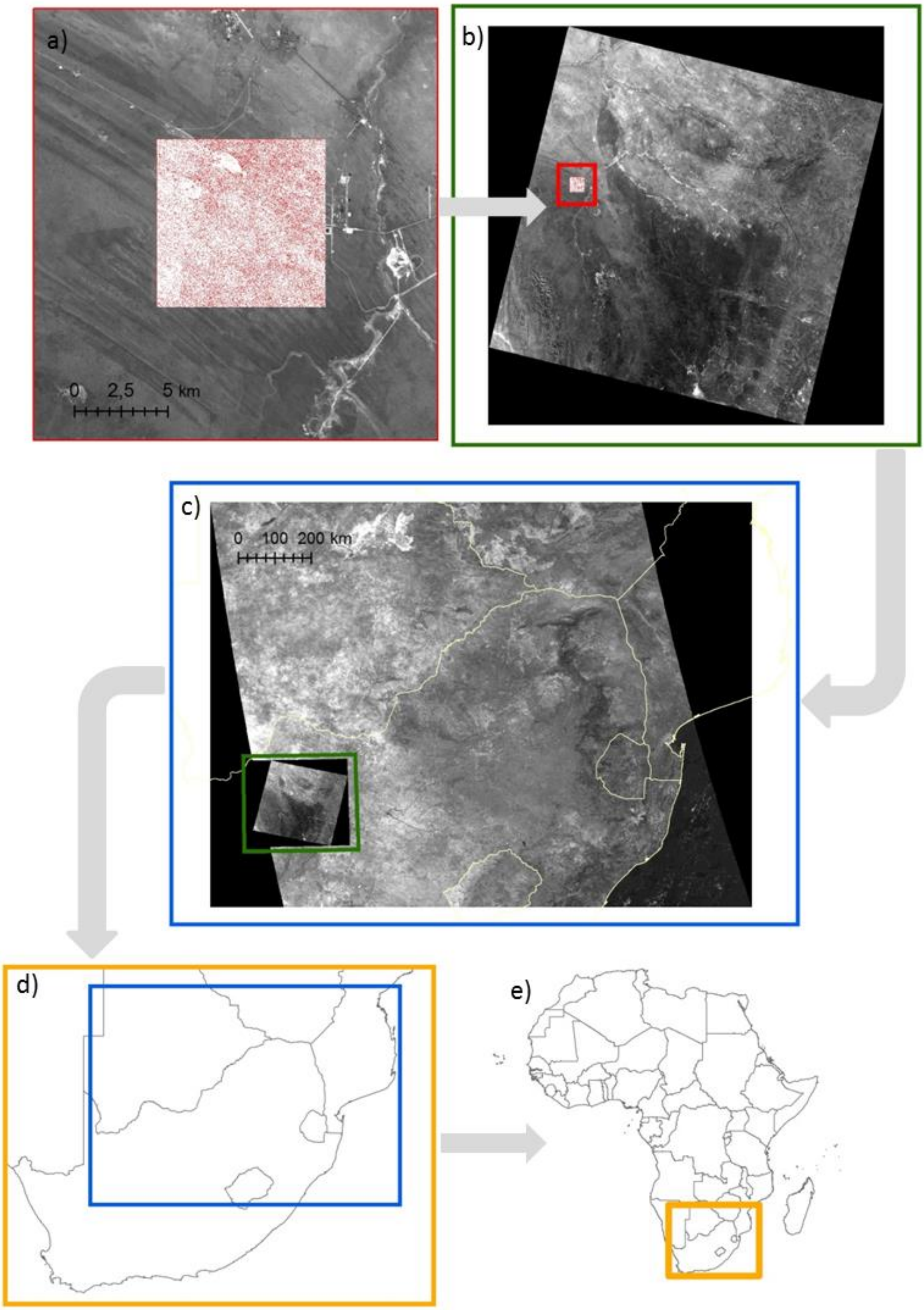

Figure 3: Multi-scale Analysis (Example South Africa). (a) Classified field data was linked to the respective pixel in the (b) Landsat image. The regression model then was applied to the Landsat scene and further to the second level using (c) MODIS tiles. The study area as represented in (c) is depicted in (d), its location within Africa in (e). Source: own draft.

Random forests regression trees were run to determine the degree of bush encroachment. The definition according to Breiman (2001) describes random forests as

*"a classifier consisting of a collection of tree-structured classifiers {h (**x**,_k), k = 1, . . .} where the {_k} are independent identically distributed random vectors and each tree casts a unit vote for the most popular class at input x" (2001: 6).*

Regression of remote sensing imagery integrating this technique has already been used to detect BE. Random forests were taken into account by Gessner et al. (2013) for their study which also included the use of multi-scale datasets. In their work, besides high-resolution imagery (Ikonos, Quickbird), Landsat data represents the second level before up-scaling to moderate-resolution imagery using the MODIS sensor. Two study sites in the savanna regions of Namibia were analyzed and showed that up-scaling to a higher level is possible and leads to good results if high-resolution data or precise ground truth information is used as baseline.

The advantages of applying RF and RT lie in the non-parametric nature, high regression accuracy and being capable of determining variable importance (Rodriguez-Galiano et al., 2012; Symeonakis & Higginbottom, 2014). According to Rodriguez-Galiano et al. (2012) RF is "highly accurate and robust to training data reduction and noise" (Symeonakis & Higginbottom, 2014: 30). These findings led to a preference in using random forests for this approach.

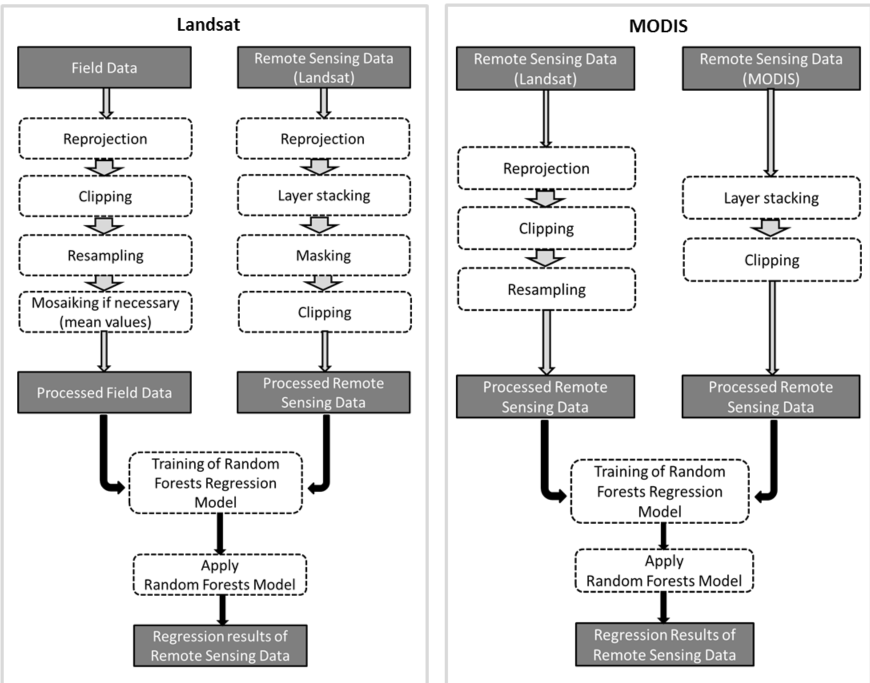

Figure 4: Left: Process chain for the detection of bushes/woody vegetation in Africa with the first upscale using Landsat data. Field data represents the training sites in South Africa and Ethiopia. Right: Process chain for the second up-scaling using the regression result of the Landsat data as input.

Figure 4 shows the process chain for both levels – Landsat and MODIS. Field data first had to be reprojected, clipped and resampled for further use. As some classified datasets in South Africa were overlapping mosaicking was applied by calculating the mean values of overlapping pixels. After reprojection the remote sensing data was stacked, masked and clipped to prepare them for further processing. The training sites, referring to the processed field data, trained the RF regression model for the first level of remote sensing data represented by Landsat imagery. The developed process chain is further applied to the next level applying the second level model to MODIS data (Figure 4, right) where the regression result of the Landsat scenes provide the training sites for further up-scaling.

The regression results of the remote sensing (RS) image – whether Landsat scene or MODIS tile – show the percentage area covered with woody vegetation per pixel.

4.1.2 NDVI Trend analysis

The Normalized Difference Vegetation Index (NDVI) provides information on the greenness of vegetation and thereby allows the observation of the status of land, especially when analyzing this index over time with time-series trend analysis (Nkonya et al., 2011).

Time series analysis of vegetation helps to detect certain trends of vegetation as e.g. observing land degradation processes via decreasing trends on the one hand. Increasing trends in vegetation on the other hand could not only refer to improved status of land, due to e.g. use of fertilizer or sustainable land management strategies, but at the same time describe BE and thereby a process of degradation as vegetation, here especially woody vegetation cover, is increasing while at the same time grasslands are suppressed (Bai, de Jong & van Lynden, 2010; Lesoli et al., 2013). Increasing vegetation trends over time are also reported to be related to BE in African savannas (Saha, Scanlon & D`Odorico, 2015, Mitchard & Flintrop, 2013). But in general, there is so far no reliable map which is dealing with the gain of woody vegetation cover instead of focusing only on forest loss (Mitchard & Flintrop, 2013).

Data on NDVI with values ranging from -1 to +1, is derived from remote sensing imagery using the red (RED) and near-infrared (NIR) band. The index is calculated by:

$$NDVI = \frac{NIR-RED}{NIR+RED}$$ (Huete et al., 2002; Jiang et al. 2008)

Woody vegetation predictions as conducted in chapter 4.1.1 were overlaid with NDVI increasing trends. NDVI panel data from 1982 to 2013 was observed using the GIMMS NOAA AVHRR NDVI3g dataset with 8 km spatial resolution (Pinzon, Brown & Tucker, 2007; Pinzon & Tucker, 2014) and a bi-weekly temporal resolution. Images for the year 1981 are available but as some images are missing to get the full year coverage the analysis started in 1982.

Pixel based trends were calculated in R based on annual sums. The Mann-Kendall non-parametric trend was conducted for the detection of significant trends. Pixel with no values and non-significant trends were masked out. Afterwards the dataset was reclassified to highlight only increasing significant trends. With reclassifying and overlaying using GIS techniques hotspots were consequently identified in which woody vegetation prediction and significant increasing NDVI trends overlapped.

4.1.3 Results: Multi-Scale Analysis

Among all datasets the prediction for 2007 showed the best results which are related to the input field data with the highest resolution of 0.5 m. The field data for Ethiopia with 15 m resolution was less accurate for the training of the classification compared to the two input datasets for South Africa. Woody vegetation cover on the African continent was therefore predicted much higher based when using the Ethiopian field dataset for the classification than in the other two regressions.

Figure 5 presents the prediction based on 2007 field data. Forested areas were masked according to the GlobCover 2009 classification as the forest areas obviously report high probability of woody vegetation cover.

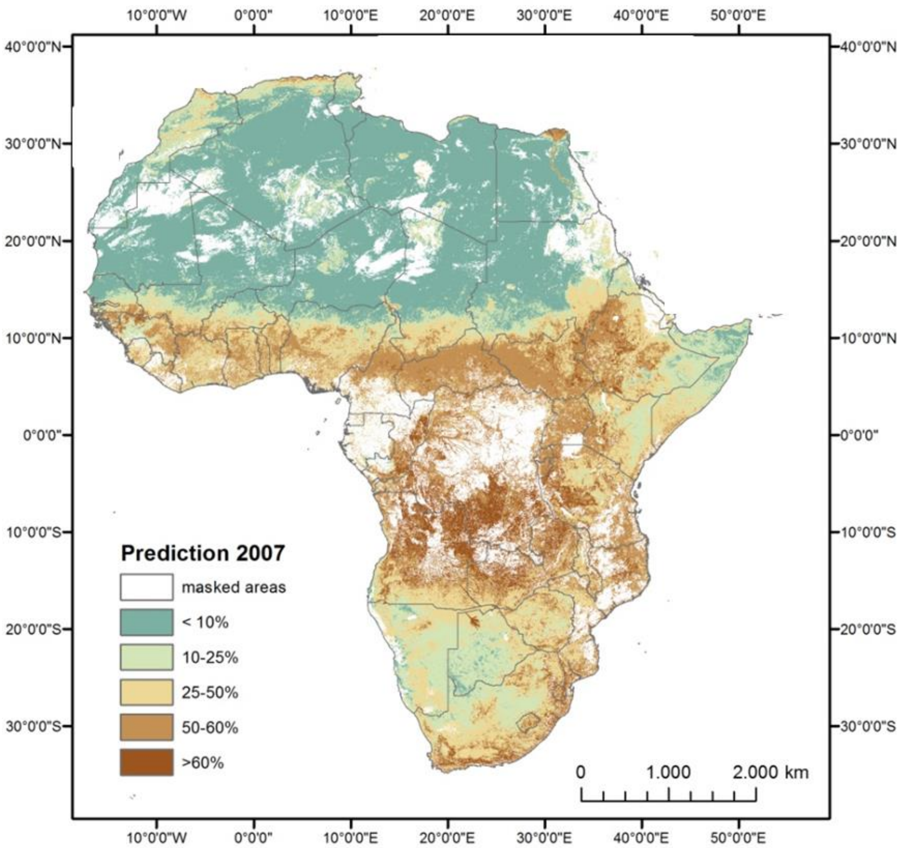

Figure 5: Prediction for woody vegetation probability in Africa based on 2007 field data from South Africa.

The predictions based on the 2007 and the 2009 field data were furthermore superimposed to identify those areas where both predictions match. We chose a threshold of 25 % which was the closest showing BE patterns in those areas where we knew that BE processes take place such as central Namibia, South Africa, parts of Botswana and Ethiopia. Brownish patterns show those areas overlapping in both predictions – 2007 and 2009 – while blue colored patterns refer to woody vegetation with more than 25 % for the 2007 prediction and purple patterns for the 2009 prediction respectively.

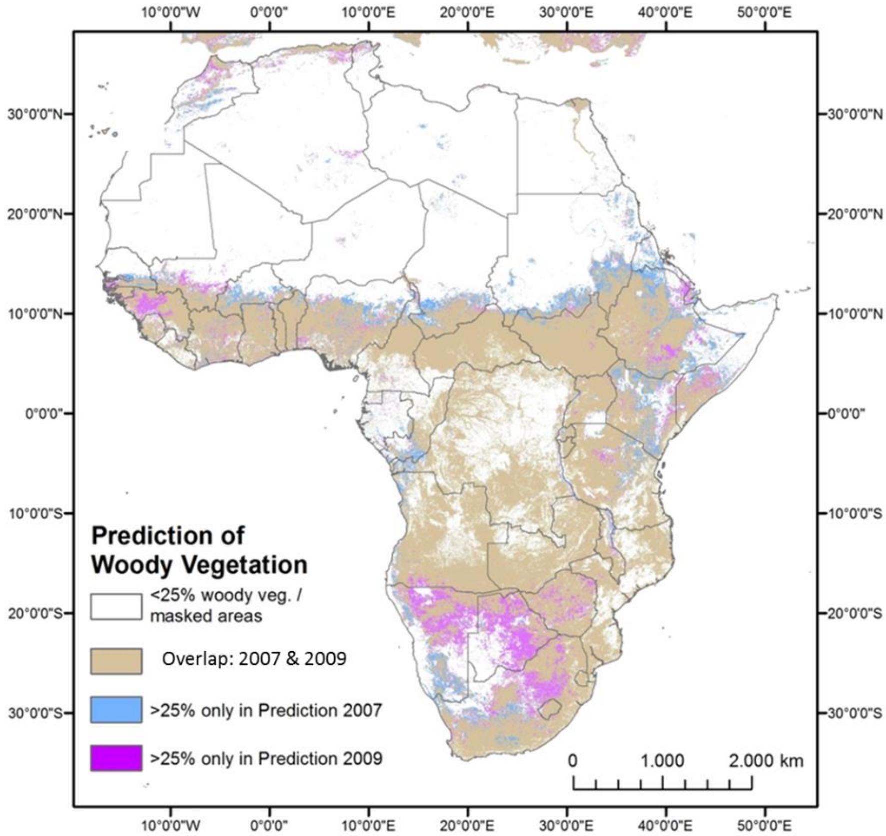

Figure 6: Woody Vegetation Map based on predicted data for 2007 and 2009. The threshold was set at 25% woody vegetation prediction for the given year based on the field data as training data for the classification.

As BE is a process we integrated the NDVI time series analysis to highlight those areas where besides the occurrence of woody vegetation significant increasing NDVI trends could be observed. Figure 7 shows the overlap of increasing NDVI trends based on the described method in chapter 4.1.1 and woody vegetation structures based on the prediction.

Southern Africa and parts of Eastern Africa, including also the study site in Ethiopia observing the spreading of *prosopis juliflora,* could be highlighted. *Prosopis juliflora* invention sites in the Baadu case study site were analyzed in the study of Ayanu et al. (2013). Based on Landsat and ASTER data classification was conducted to measure and detect the invasion of *prosopis juliflora* for several years from 2000 until 2013. The area of BE describing the invasion of *prosopis juliflora* is highlighted in yellow in the two smaller maps below the main map showing a close-up of Ethiopia and the Baadu case study area. Even though the Ethiopia field data is due to a much lower resolution and less promising results not included in the hotspot map (Figure 7), the area could be detected with the here documented approach by calculating increasing NDVI trends and woody vegetation cover based on BE training data for South Africa.

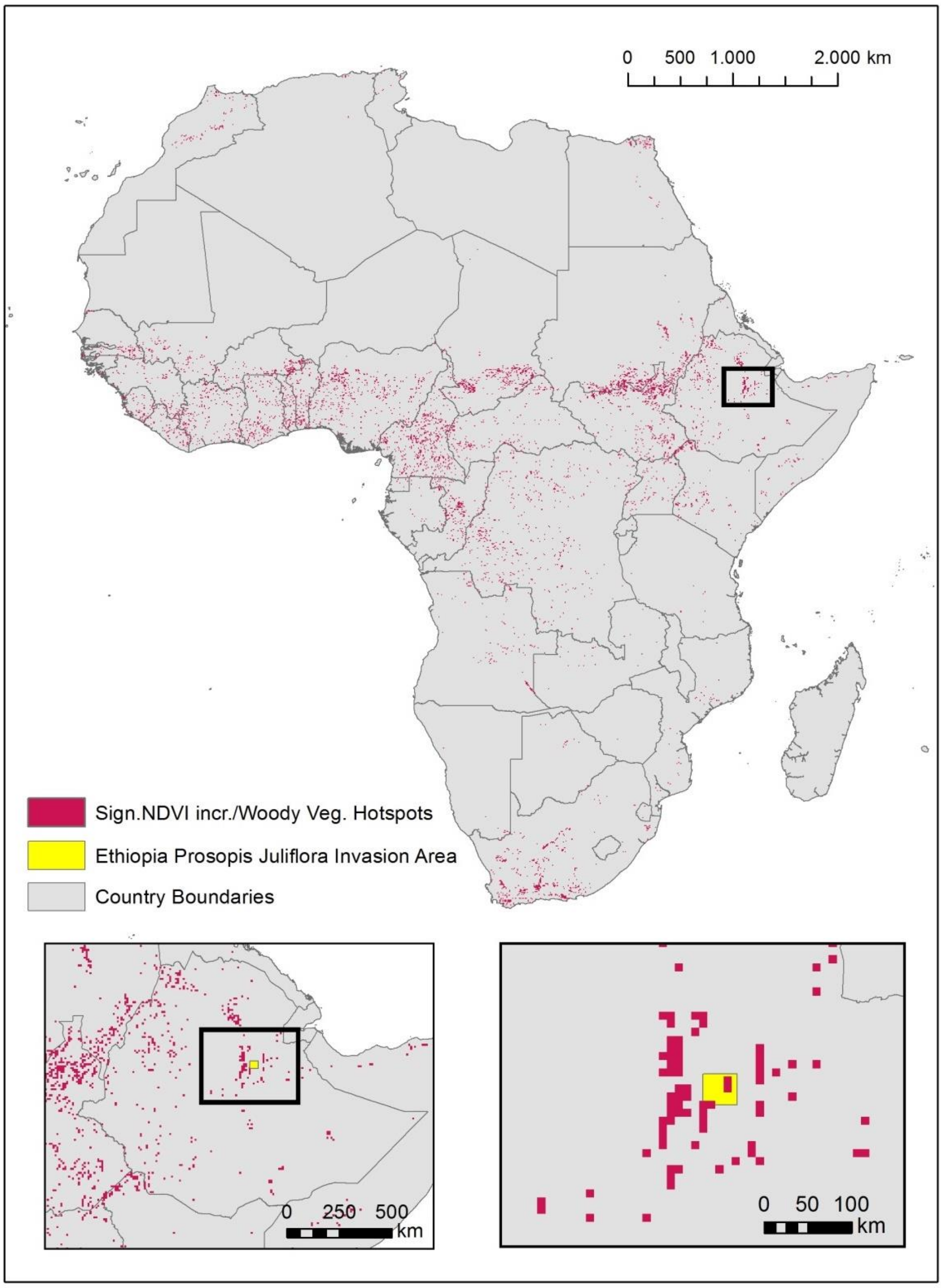

Figure 7: Overlap of NDVI increasing trends (trend analysis for: 1982-2013) and woody vegetation prediction.
NDVI data was derived from: http://ecocast.arc.nasa.gov/data/pub/gimms/3g.v0/

Nevertheless, according to the Hotspot-Map we could not identify areas in Namibia – a country which is severely affected by BE. Moreover, the approach detects BE in western Africa where we could not find any proof in the literature nor could verify these findings with experts in the area. Considering these results we can on the one hand identify areas affected by BE based on our method but at the same time other areas such as patterns in western Africa were inaccurate.

Table 2 lists the percentage area per country where possible BE is affecting parts of the country. Calculation is based on the overlap of NDVI significant increase and woody vegetation prediction.

Table 2: Country Statistics on overlapping woody vegetation prediction and significant NDVI increase.

Country	ha	Country	ha
Algeria	175,00	Madagascar	0,00
Angola	2750,00	Malawi	0,00
Benin	3325,00	Mali	2650,00
Botswana	325,00	Mauritania	75,00
Burkina Faso	3275,00	Morocco	1925,00
Burundi	0,00	Mozambique	975,00
Cote d´Ivoire	5275,16	Namibia	325,00
Cameroon	12580,57	Niger	200,00
Central African Republic	4725,00	Nigeria	11475,26
Chad	7975,00	Rwanda	0,00
Congo	4150,00	Sao Tome and Principe	0,00
Congo DRC	11650,00	Senegal	1751,57
Djibouti	0,00	Seychelles	0,00
Egypt	1475,01	Sierra Leone	1431,02
Equatorial Guinea	150,00	Somalia	1425,00
Eritrea	125,00	South Africa	10850,00
Ethiopia	8275,00	South Sudan	12325,00
Gabon	975,00	Sudan	8750,00
Gambia	100,18	Swaziland	75,00
Ghana	4650,00	Tanzania	325,00
Guinea	2075,00	Togo	1800,00
Guinea-Bissau	256,49	Tunisia	275,00
Kenya	3225,00	Uganda	2125,00
Lesotho	0,00	Yemen	0,00
Liberia	2300,06	Zambia	575,00
Libya	32,59	Zimbabwe	0,00

As already derived from the visual interpretation of the Figure 7 countries such as South Sudan and Ethiopia in Eastern Africa, South Africa but also Cameroon, Nigeria and Cote d´Ivoire in western Africa show the highest numbers of overlapping NDVI increase and woody vegetation prediction.

Regressions were run with up to 500 trees. Figure 8 shows the model performance for R^2 and RMSE (Root Mean Square Error, or MSE (Mean Square Error)) with 500 regression trees generated with the input field data of 2009 on 0.75 m resolution for Landsat and MODIS. As seen in Figure 8 showing an example from 2009 R^2 improves with the number of trees while also the MSE decreases simultaneously.

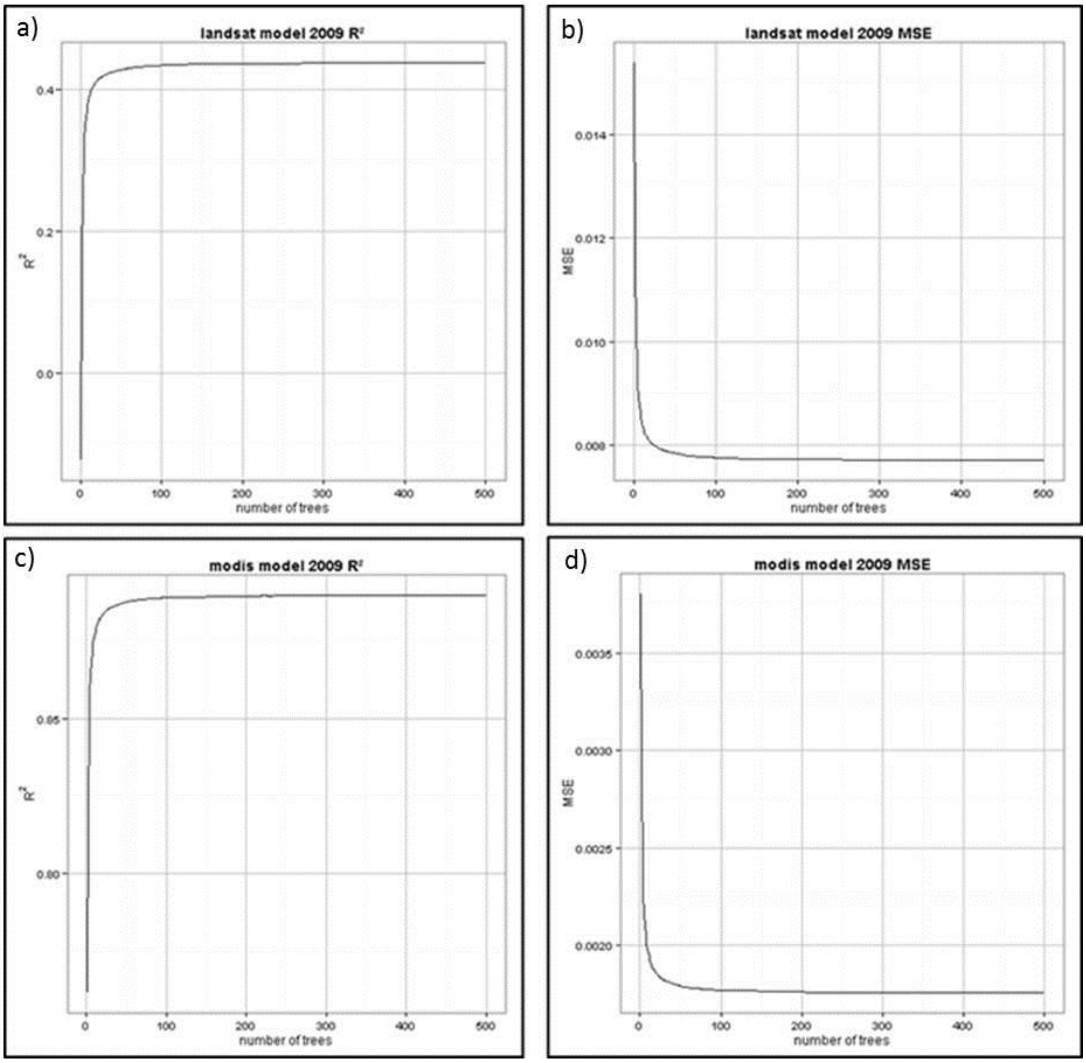

Figure 8: Example for R^2 and MSE with 500 regression trees for the South African up-scaling based on the year 2009: R^2 (a,c) and MSE (b,d) for Landsat (a,b) and MODIS (c,d) regression models for up to 500 regression trees.

For each field data and the respective year the R^2 and Mean Square Error (MSE) were calculated (see Table 3).

Table 3: R^2 and MSE for the Landsat and MODIS regression for 2007, 2009 and 2013

		2007	**2009**	**2013**
Landsat	R^2	0.34	0.44	-
	MSE	0.0048	0.0077	-
MODIS	R^2	0.94	0.89	0.647
	MSE	0.0011	0.0018	0.041

For South Africa first validation of the field data from 2007 conducted with 500 trees results in an R^2 of 0.34 and Mean Square Error (MSE) of 0.005. The up-scaling to MODIS data with 500 m resolution showed an R^2 of 0.94 and MSE of 0.001. Both levels were performing very well. Taking into account field data from 2009, again from South Africa, we could observe an MSE of 0.008 and an R^2 of 0.44 while for the MODIS regression the MSE was 0.002 and R^2 = 0.89.

In Ethiopia data on BE based on the spreading of *prosopis juliflora* was available for 2013. The same process chain as developed for South Africa was applied using the dataset of Ayanu et al. (2013). The first upscale to Landsat by using the initial dataset with 15 m resolution was not resulting in a clear correlation which was due to a coarser resolution of the dataset compared to the data available for South Africa. Therefore the data for Ethiopia was used as input for the second upscale to apply a regression to the MODIS tiles for 2013. Regression results for 2013 therefore are only listed for MODIS data. With regard to the coarser dataset we could also observe slightly weaker results.

According to the resolution of the remote sensing data regression results were performing better for the South African classification referring to higher R^2- and lower MSE-values in general (see also Table 3). To validate the model and assess its overall performance the field data samples were divided randomly into training samples (75%) and validation samples (25%). A description of the number of points that were used for training and testing can be found in Table 4.

Table 4: Number of Points used for training and testing the prediction model.

	# samples	**2007**	**2009**	**2013**
Landsat	training	139903	146760	-
	testing	46634	48919	-
MODIS	training	105422	105957	2897
	testing	35140	35318	965

The performance of the prediction model for each field dataset and the respective upscale is visualized in Figure 9. Here, the comparison of predicted versus validation data for Landsat (a, b) and MODIS (c,d) for the years 2007 and 2009 is visualized.

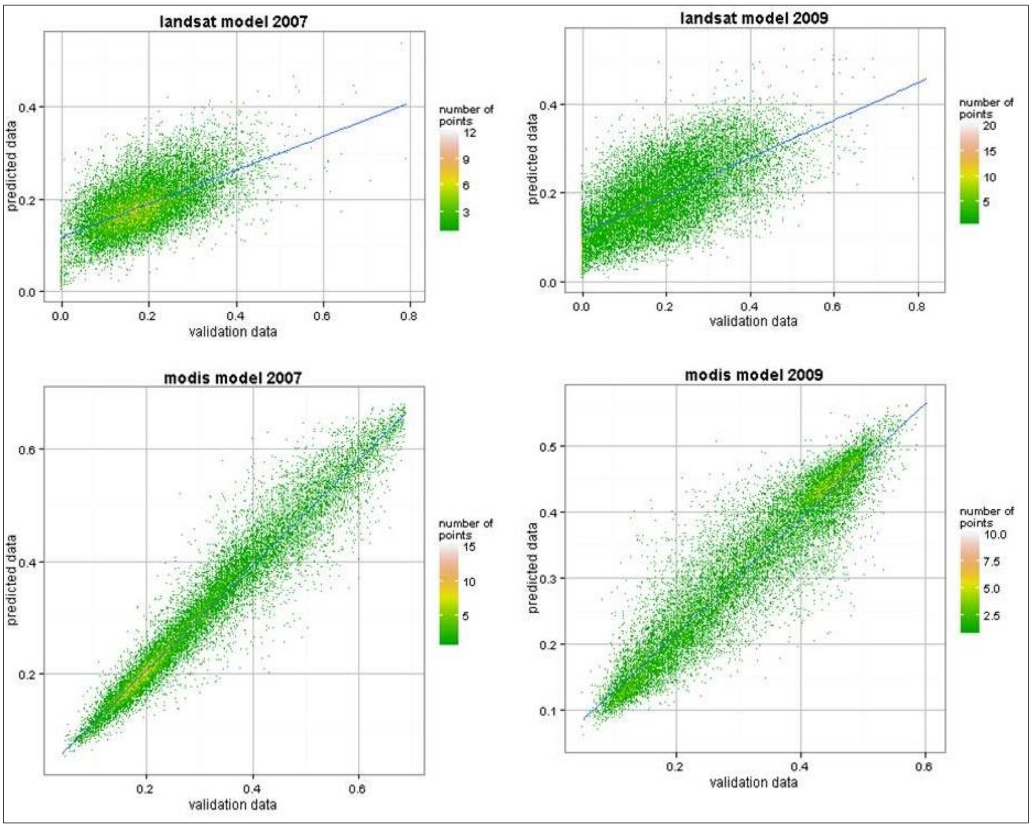

Figure 9: Landsat and MODIS Prediction Model for 2007 and 2009.

The upscale for Ethiopia could only be conducted for the second upscale. Comparing this figure with the prediction for 2007 and 2009 we can identify a much better performance for the datasets provided for South Africa in those two years. As seen in the regression results of the 2013 dataset based on Ethiopia field data we can recognize a less strong performance compared to the other datasets which is also due to less many field data pixel (Figure 10).

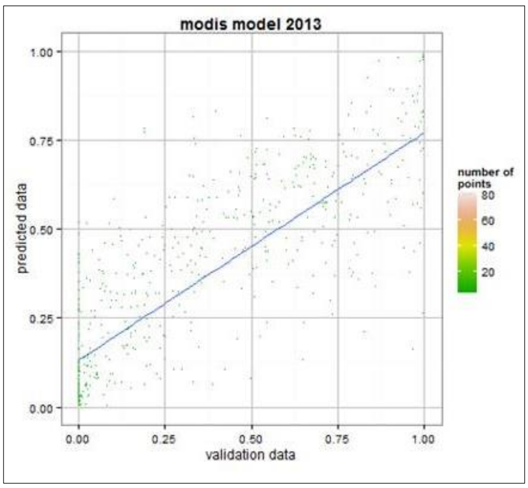

Figure 10: Predicted results versus validation data for the Ethiopian field dataset from 2013.

As areas that represent bare soil such as deserts, urban area or water bodies should be masked out of further analysis MODIS NDVI data (Product MOD13A1) with 500 m resolution was taken into account. All values with an NDVI index below 0.1 which represent the mentioned surfaces and do not refer to vegetative cover as such were masked in the final dataset.

4.2 Bush encroachment/woody vegetation: Probability Map (2nd Step)

The generation of the probability map for BE is based on driving factors that were reported to favor and trigger BE. Africa is a very diverse continent characterized by many different agro-ecological zones (AEZ) which also hold different land management practices. Some areas favor BE more than others. Based on literature research five main impact variables were identified that are considered to trigger BE. The woody vegetation map (see chapter 4.1) was used as a main input in the following analysis. We used logistic regression to identify and rank variables related to BE. The approach is based on Dubovyk et al. (2013).

4.2.1 Causes of Bush Encroachment

Four main causal factors were identified and will be discussed further:

- Precipitation
- Soil Moisture
- Livestock Pressure – Overgrazing
- Fire Occurrence
- Atmospheric CO_2

These indicators have been analyzed and identified in various studies (e.g. van Auken, 2000; Roques et al., 2001; Sankaran et al., 2005; Kulmatiski and Beard, 2013). Other factors that possibly have an impact on BE included soil type, overhunting or eradication of indigenous savanna animals, introduction of invasive species and mismanagement (Scholes & Archer, 1997; Sankaran et al. 2005; Hudak, 1999; Roques, Connor & Watkinson, 2001; Kgosikoma, Harvie & Mojeremane, 2012). The causes of BE have not yet been fully understood (Ward, 2005). The identification of drivers is still ongoing while the main drivers could more or less be named as seen above. Soil characteristics could not be identified as high impact variables for BE as studies observing the feedback loops of soil types and BE concluded that soil characteristics alone do not represent the main cause of BE (Ward, 2005; Wiegand et al., 2005). Land use management was not identified as a key variable as stated in a study by Wigley, Bond & Hoffman (2009). Moreover, it could be recognized that some indicators such as tree dependency on rainfall showed different results across savanna systems in South

America, Africa and Australia (Lehmann et al. 2009). Nevertheless the respective variables were discussed and analyzed for the African continent to test their influence on BE.

Invasive species could have been introduced to a certain area to improve land and productivity. This e.g. happened in Ethiopia, where *prosopis juliflora* was introduced in 1983 (Ayanu et al., 2014). Mesquite, another *prosopis* species, was on the other hand brought to South Africa to establish more fodder for livestock (Wise et al., 2012). In general, the intention when introducing a non-endemic species to an area is to improve a current ecosystem. But with the introduction of a new species which describes a kind of disturbance in an established ecosystem this system can get imbalanced and begins to change if preconditions favor this change. A study in South Africa observing three different land use practices in relation to BE could state that there is no causal relationship depending on the land management practices which could not state a clear impact on the process of BE (Wigley et al., 2009). The ongoing discussion on the causes of BE raises a difficulty to identify potential variables. Therefore the main factors were highlighted and investigated with regard to their impact on woody vegetation in Africa. A deeper insight in the four main variables will be given below.

i. Precipitation

The tree-grass co-dominance in savanna ecosystems was mainly explained by water availability as a main limiting factor (Ward, 2005). In this regard the "two-layer hypothesis" of Walter (1939) describing different levels of water availability for either grasses or woody vegetation became well known. Initially Walter (1939 and 1971) came up with this theory which was later simplified in Walker and Noy-Meir (1982) with the explanation of vertical niches (Ward et al., 2012). According to the hypothesis grasses do rely more on water or moisture which they can access in the topsoil while woody vegetation can additionally access subsoil water below the grass roots as shown in Figure 11 (Walker and Noy-Meir, 1982; Ward et al., 2012).

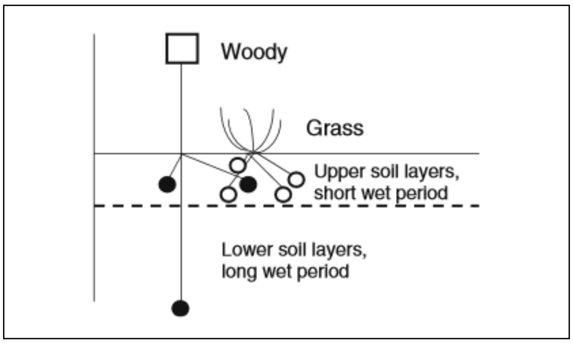

Figure 11: Vertical resource partitioning by Walker and Noy-Meir 1982, in (Ward et al., 2012)

Walter moreover stated that up to a mean annual precipitation (MAP) of 500 mm grasses dominate while above a MAP of 500 mm woody vegetation becomes the dominating vegetation (Walter 1939 in Ward et al., 2012). According to Sankaran et al. (2005) a threshold of 650 mm MAP can also describe the distinction of arid and humid but rather stable and unstable savanna systems which are shown in Figure 1. The threshold is not seen as a fixed cut-off line but rather as orientation. Below an MAP of 650mm savannas are considered to be stable as fire, herbivory and soil properties interact to reduce woody cover (Ward et al., 2012). Moreover below this threshold the system is constrained and increases also linearly with MAP (Sankaran et al., 2005).

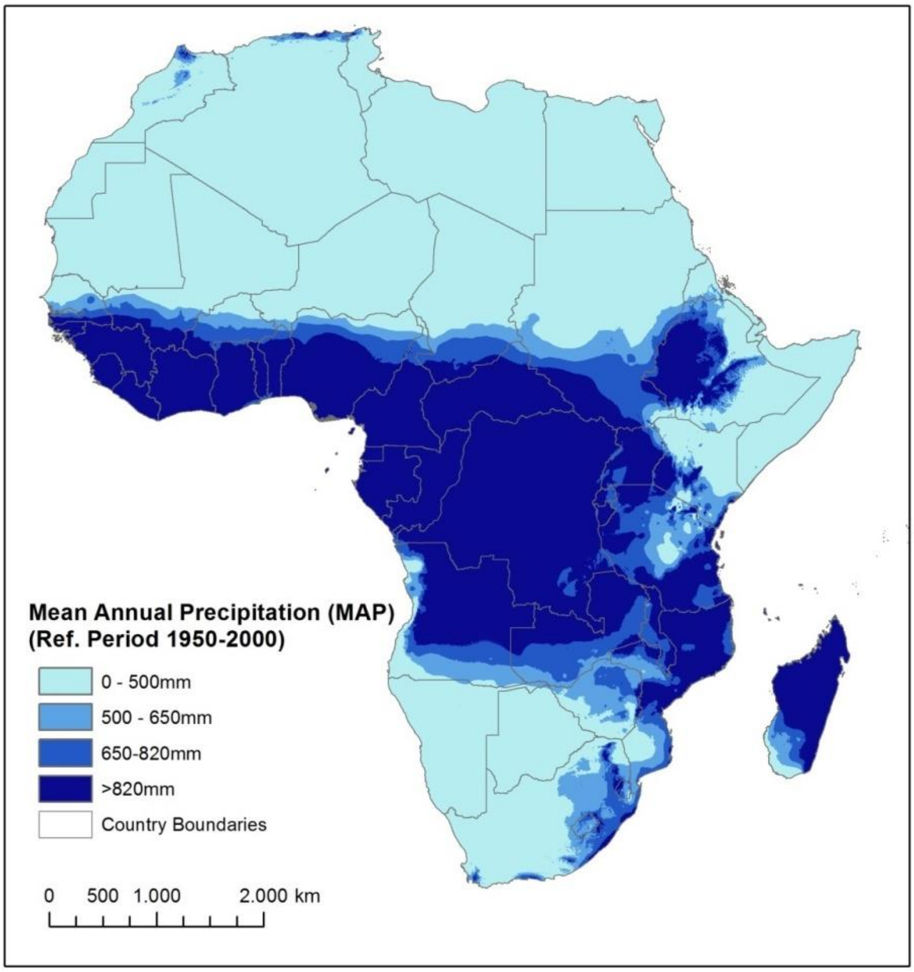

Figure 12: Mean Annual Precipitation in Africa based on the observation period from 1950-2010. Thresholds were chosen based on literature research linked to characteristics of savanna ecosystems and the stability of these systems. Source: WorldClim – Global Climate Data, Hijmana et al. 2005.

Figure 12 shows the MAP for Africa based on data from 1950-2010 (Hijmans et al., 2005). The color scheme describes the thresholds as mentioned by Walter (1939)(500 MAP) and Sankaran et al., (2005)(650 MAP) . Another threshold added is based on (Higgins et al., 2010) stating that in areas with MAP > 820 mm fire can only be confidently considered to affect savanna structure. But also grasslands in areas where these conditions are not all valid (Sankaran et al., 2005) are affected by BE. Looking at Namibia for example which is identified with a MAP below 500 mm according to the data it should not be affected that much by woody vegetation encroachment if Walter (1939) and even Sankaran et al. (2005) were right. This statement could also be valid for parts of South Africa. Based on these restrictions a probability analysis with possible impact factors is useful.

But in general this rainfall-dependent tree density relationship differs across savannas in South America, Africa and Australia (Lehmann et al., 2009).

ii. Soil Moisture

If precipitation rates change this also has an effect on soil moisture content and thereby can cause large-scale shifts in plant growth (Kulmatiski & Beard, 2013). Moreover there is a distinction between annual precipitation rates compared to precipitation intensity where last mentioned is an important aspect with regard to soil conditions as it could push soil water to deeper levels (Kulmatiski & Beard, 2013). If this happens the two-layer hypothesis of Walker and Noir-Meir (1982) becomes valid again. Moreover it is questionable which roles precipitation trends and precipitation variability play.

Soil moisture data (Essential Climate Variable (ECV) Surface soil Moisture (SM) (Product ECV SM 02.0) of ESA (Liu & Dorigo, 2012; Liu et al., 2011; Wagner et al., 2012) was furthermore included in the study for further analysis. The dataset is available for the time period 1978 to 2013, has a daily observation and a spatial resolution of 25km. Figure 13 Shows mean soil moisture calculated for the year 2010.

For the probability map annual data from 2000 to 2010 was analyzed to generate the mean over this period of time. Furthermore also trends were calculated to analyze their impact.

As identified in Figure 13 some areas contain no data. Here, areas with dense vegetation such as in tropical forest areas, with ice cover, strong topography, large fractional coverage of water, or extreme desert areas are masked out as it is not possible to gather meaningful soil moisture information in those areas[1]. As for further analysis these areas were masked out anyway as they are not relevant to BE as such these data gaps did not have an impact on the ongoing analysis.

[1] Information by ESA via: http://www.esa-soilmoisture-cci.org/node/136

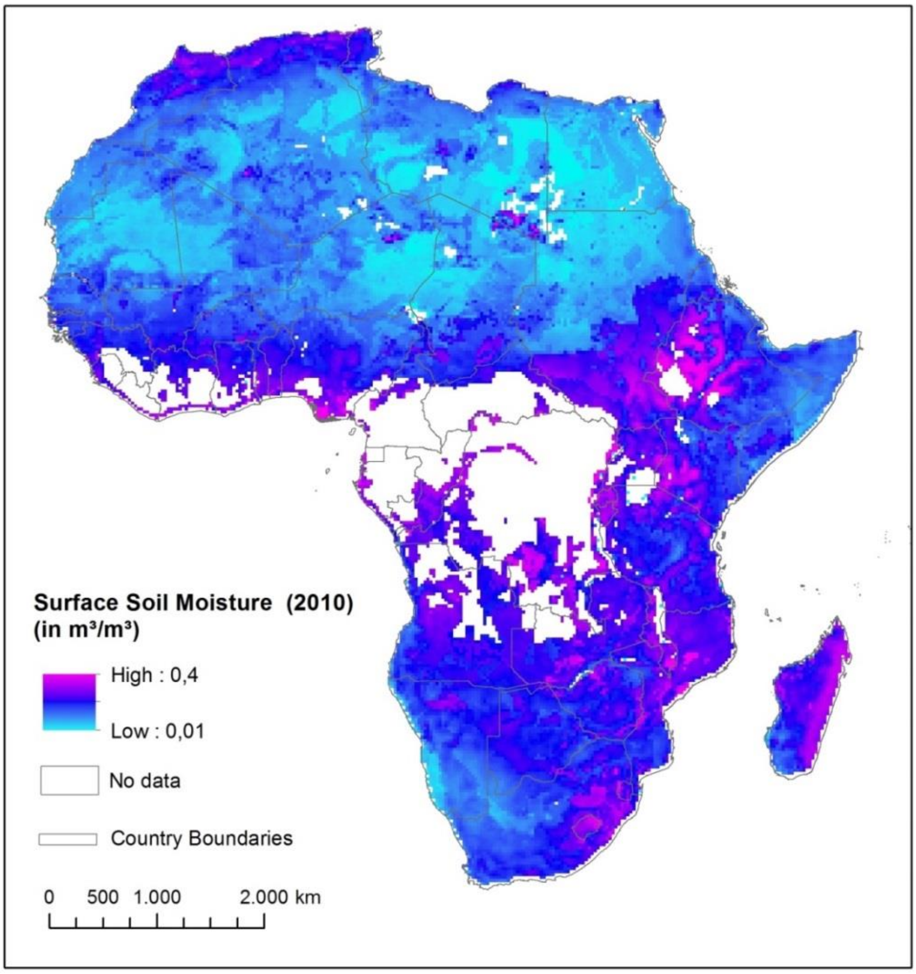

Figure 13: Soil Moisture for 2010 by calculating the mean annual surface soil moisture via zonal statistics. Data is based on ECV SM 02.0 by ESA (Liu & Dorigo, 2012; Liu et al., 2011; Wagner et al., 2012).

iii. Fire Occurrence

Fire does naturally occur in savanna systems and lead to a balanced ecosystem (Scholes and Archer, 1997; Williams et al., 1999; Roques et al., 2001; Van Langevelde et al., 2003; Govender et al., 2006). If natural disturbance factors such as fire, browsing or grazing are absent the change from open to closed bush savannas can happen very fast (Scholes and Archer, 1997). We can assume that in areas where an intrusion in a balanced ecosystem occurs the natural tree-grass co-dominance changes to an imbalanced system.

Fires are naturally occurring and help to maintain grasslands where preconditions are rather favoring woodlands or forest (Bond & Midgley, 2012). Larger trees are more fire resistant than grasses and regrow fast after a fire event (Bond & Midgley, 2012). According to Higgins et al (2010) fires affect savanna structure at MAP >820 mm. The suppression or prevention from fire is reported to trigger bush encroachment (Scholes & Archer, 1997; Lesoli et al. 2013; Kgosikoma & Mogotsi, 2013).

Data on fire occurrence was provided by NASA based on LANCE FIRMS MODIS Active Fire data. Five years from 2010 to 2014 were analyzed to find the mean fire occurrence among the region. The datasets provide location of the pixel with fire occurrence at a certain time during the year by a point-shapefile. Within the data files the date of fire occurrence is listed. Additionally a confidence measurement is provided which ranges between 0 and 100% (Giglio et al. 2003). Hotspots of fire occurrence are grouped based on their confidence value given in the dataset. According to Giglio (2010) high confidence of the occurrence of a hotspot is given at >80%. For the analysis all pixel with a confidence level of >80% confidence level were therefore taken into account as higher confidence levels can be used to reduce the number of false alarms (Giglio, 2010). With a 1km-grid that was produced for Africa the number of fires for each year that occurred within one grid was calculated with zonal statistics. The sum of fires for each year between 2010 and 2014 built the baseline for the mean annual sum of fire occurrence on the African continent among five years in the respective time period (Figure 14).

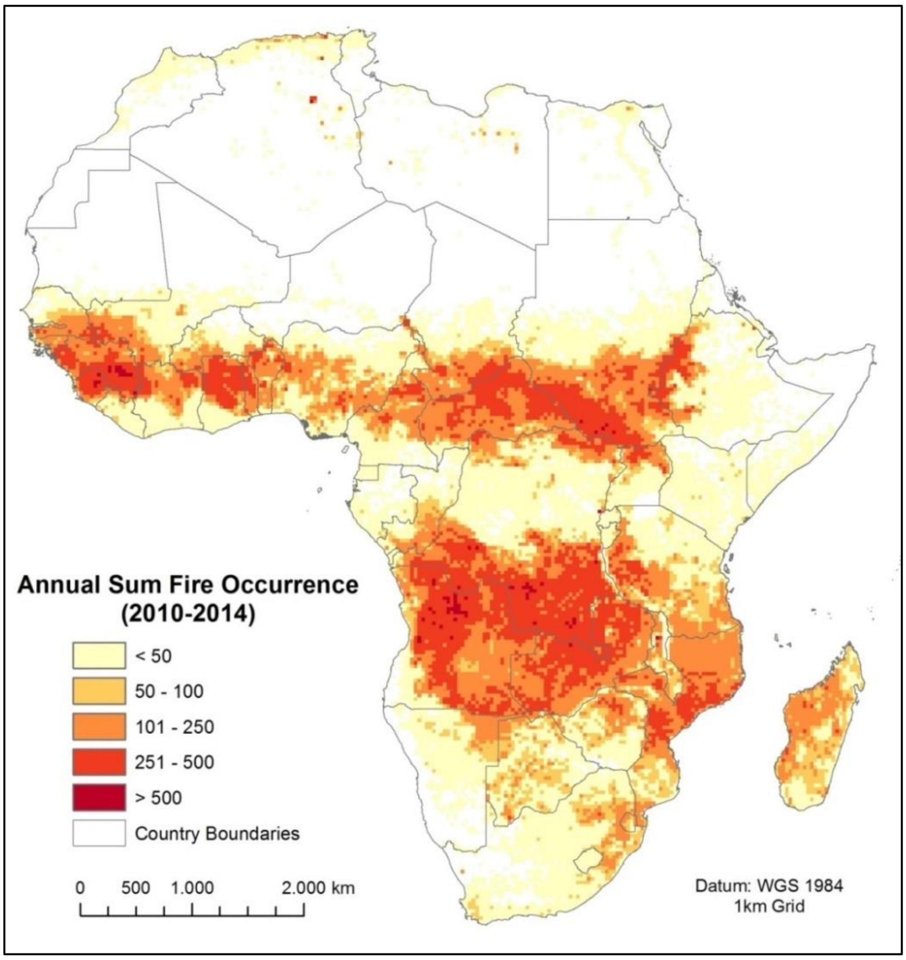

Figure 14: Annual Sum Fire Occurrence. Reference period: 2010-2014. The number of fires occurring each year with a confidence level of >80% were taken into account to generate a mean annual sum fire occurrence map for Africa. Data Source: LANCE FIRMS MODIS Active Fire data.

iv. Atmospheric CO_2

The impact of increasing atmospheric CO_2 on increasing woody vegetation and the suppression of grasslands is coming up more recently with regard to studies addressing this issue (Moleele et al., 2002; Bond et al., 2003; Ward, 2005; Buitenwerf et al., 2012). According to Bond and Midgley (2012) increasing CO_2 levels could favor trees over grasses. It can be assumed that because of increasing CO_2 levels trees grow faster and build their stems more rapidly and longer to be more fire prone. Moreover studies showed that in Savannas with higher CO_2 trees grew more than twice as fast as in a pre-industrial atmosphere (Kgope et al., 2010).

With regard to these information data on CO_2 emission from fossil fuel combustion was used as input variable for further analysis. Data was available via the Fossil Fuel Data Assimilation System (FFDAS) – Product FFDAS v2.0. The data can be downloaded with 0.1°/hourly spatial resolution and a monthly or annual temporal resolution.

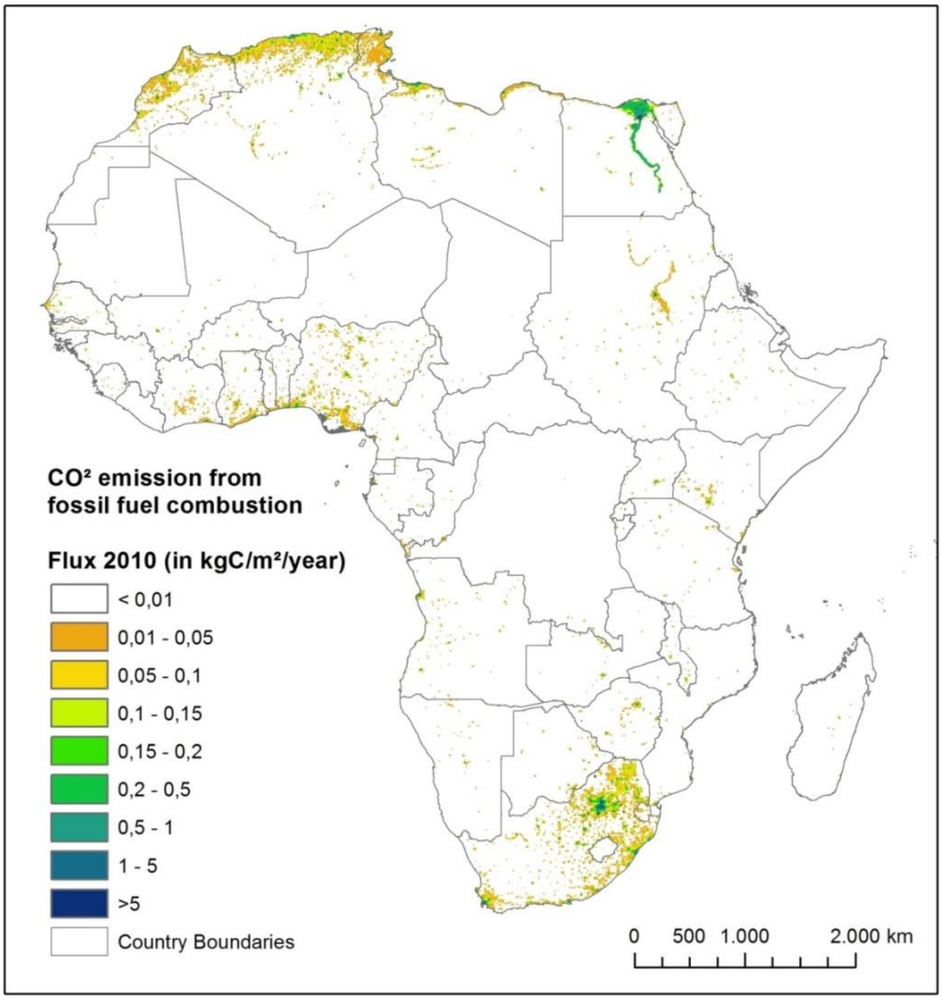

Figure 15: CO2 Emission Intensity per Pixel. Source: FFDAS 2010 via: http://hpcg.purdue.edu/FFDAS/index.php?page=media (Rayner et al., 2010)

Figure 15 shows the mean flux for 2010 expressing emissions as a product of areal population density, per capita economic activity, energy intensity of the economy, and carbon intensity of energy (Rayner et al., 2010).

v. Livestock Pressure – Overgrazing

Overgrazing is often mentioned to be one of the main cause for BE (Walter 1964, Scholes and Archer, 1997; Ward, 2005; Coetzee et al., 2008).

Especially in arid savannas overgrazing can be a more important causal factor than CO_2 effects (Bond and Midgley, 2012).

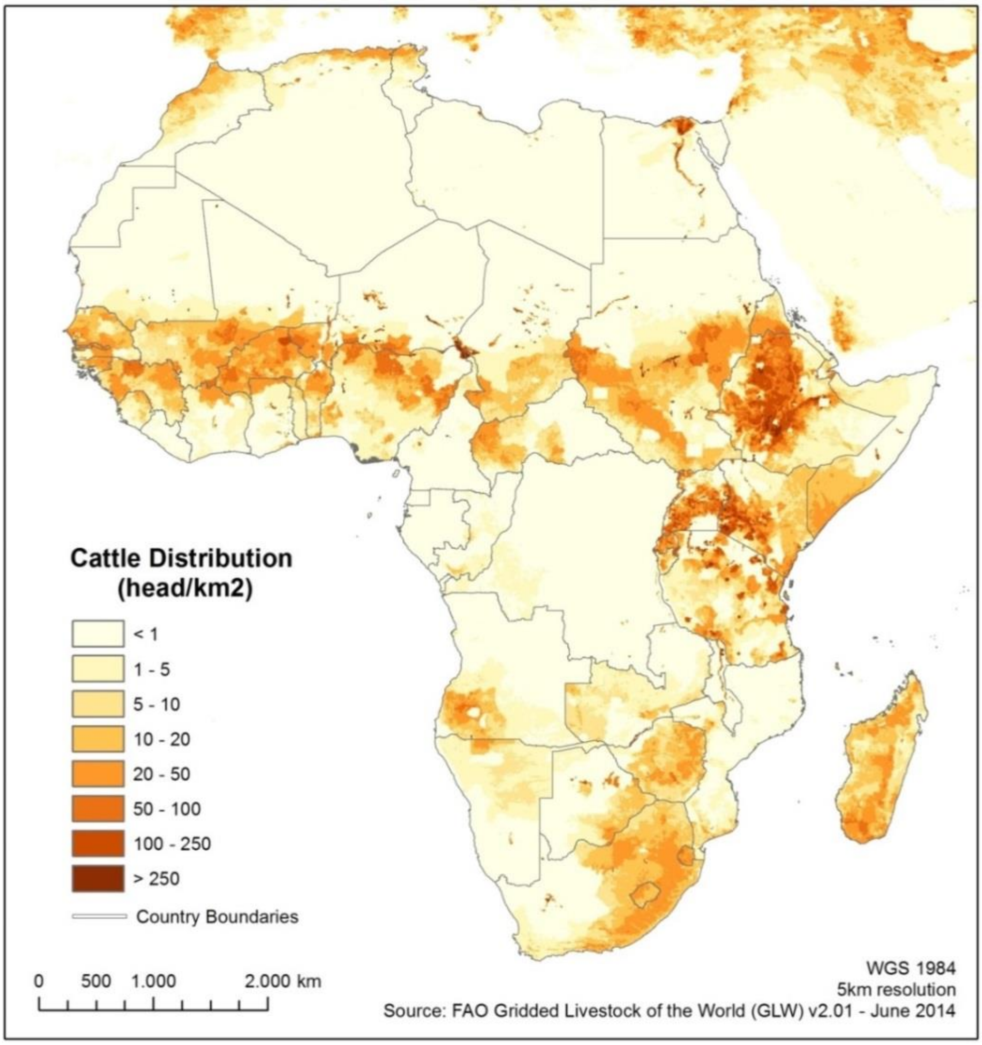

Figure 16: Cattle Density of Africa. Gridded Livestock of the World (GLW) v2.01 – June 2014 with reference period of 2010, FAO.

A study in Botswana conducted by Moleele et al. (2002) came to the conclusion that besides soil moisture, nutrient concentration and low fire frequency high cattle density is leading to increasing woody vegetation which thereby specifies livestock pressure and overgrazing as input variables for the probability analysis. Some studies deny this link but as it is reported to be an impact factor for BE in most of the literature it was a possible dataset to include in the analysis. Data as shown in Figure 16 is based on “Gridded Livestock of the world” (GLW) by FAO. Different livestock densities, including cattle density, are provided with 1 km resolution. Moreover the updated dataset from 2014 with the reference period from 2010 was chosen. As data for the whole African continent was needed a global or at least regional dataset needed to be found which could be provided with the here described dataset of FAO. Nevertheless it has to be mentioned that livestock in Africa is rather disperse which is not included in the general description of “livestock density”. Still, it was taken into account for logistic regression modeling to analyze the impact of the provided dataset and compare it with other possible indicators.

4.2.2 Logistic regression modeling

All datasets were resampled to the same pixel size. A 10-km grid was created in which the respective datasets were calculated via zonal statistics. The analysis was conducted in ArcGIS and QGIS. Depending on the dataset different approaches were applied. For soil moisture, precipitation, cattle density and CO_2 the mean value within each grid was taken into account. For fire occurrence the sum of fires that occurred within the 10km by 10km grid within the respective period of time was calculated. Finally, the generated information on woody vegetation prediction was also calculated via zonal statistics. For logistic regression modeling a binary variable was created based on the predicted woody vegetation layer. All pixels with a probability of more than 25% woody vegetation were reclassified as 1, all others as 0. For the 10 km by 10 km the majority rule via zonal statistics was approached.

Information on land cover was integrated using data by GlobCover 2009 which includes in total 23 land cover classes. Analysis on woody vegetation cover was therefore also distinguished by land cover classes.

4.2.3 Results: Probability Mapping for Bush Encroachment

The probability map was generated based on statistical and geospatial analysis with input data of the five main impact variables discussed in chapter 4.2.1. The approach is based on Dubovyk et al. (2013) using spatial logistic regression modeling.

Logistic regression modeling was applied to quantify the impact of all trigger variables for BE and calculate probabilities of woody vegetation occurrence in Africa. Table 5 gives an overview about the input variables for the model.

Table 5: Input variables for logistic regression modeling.

Variable	Description	Source	Nature of Variable
		I Dependent	
Prediction	Woody Vegetation Predictions based on 2007 and 2009 field data	Up-scaling approach based on random forests (see chapter 4.2)	Binary
		II Independent	
Precipitation (MAP)	Mean Annual Precipitation (observation period 1950-2000)	WorldClim – Global Climate Hijmans et al. 2005	Continuous
Soil Moisture	Soil Moisture Data (observation period 2010)	ESA Product ECV SM 02.0	Continuous
Fire Occurrence	Mean Fire Occurrence (observation period 2000-2010)	LANCE FIRMS MODIS Active Fire data	Continuous
CO_2 Mean	Mean CO_2 flux (observation period 1999-2009)	Fossil Fuel Data Assimilation System (FFDAS) – Product FFDAS v2.0	Continuous
Cattle Density	Cattle density (2010)	FAO Gridded Livestock of the World (GLW) v2.01 – June 2014 (reference period 2010)	Continuous
Land Cover	Global Land Cover (GlobCover2009)	Arino et al. 2012	Additional Dataset

The nature of the dependent variable, represented by the overlay of the two predictions for 2007 and 2009, was binary. Among the two classes 0 describes “no woody vegetation” referring to a prediction of less than 25% woody vegetation cover; and 1 defines “woody vegetation” cover with more than 25% prediction of woody vegetation cover. The logistic regression model was based on Hosmer and Lemeshow (2000), as also used in Dubovy et a. (2013):

$$\text{P(y) = 1/1 + exp} - P(y) = \frac{1}{1} + \exp - \left(\beta_0 + \sum_{i=1}^{n} \beta_i x_i\right)$$

P(y) describes the probability of the dependent variable *(y)* being woody vegetation linked to BE *(1)* given the independent factors $x_1, \dots, x_n$. Variance inflation factors (VIF) are calculated to avoid multi-collinearity of the variables. As no variable crossed a VIF greater than 2 all variables could be taken into account. In total 258,378 observation points were included.

In addition to the explanatory variables land cover information based on Globcover 2009 (Arino et al. 2012) was integrated. By this, also individual behavior within the land cover classes could be analyzed. Classes such as forests were dropped during the analysis as woody vegetation cover obviously is occurring with high probabilities and not referring to BE as such. In total 23 different land cover classes were integrated, including forest and bare areas, to get a first overview on distribution of woody vegetation in the different classes. Logistic regression modeling was conducted with STATA.

For some insights in the distribution of woody and non-woody vegetation Figure 17 gives the percentage of combined larger classes with regard to woody vegetation cover. Shrublands and forests can be identified to be mainly covered with high probabilities of woody vegetation.

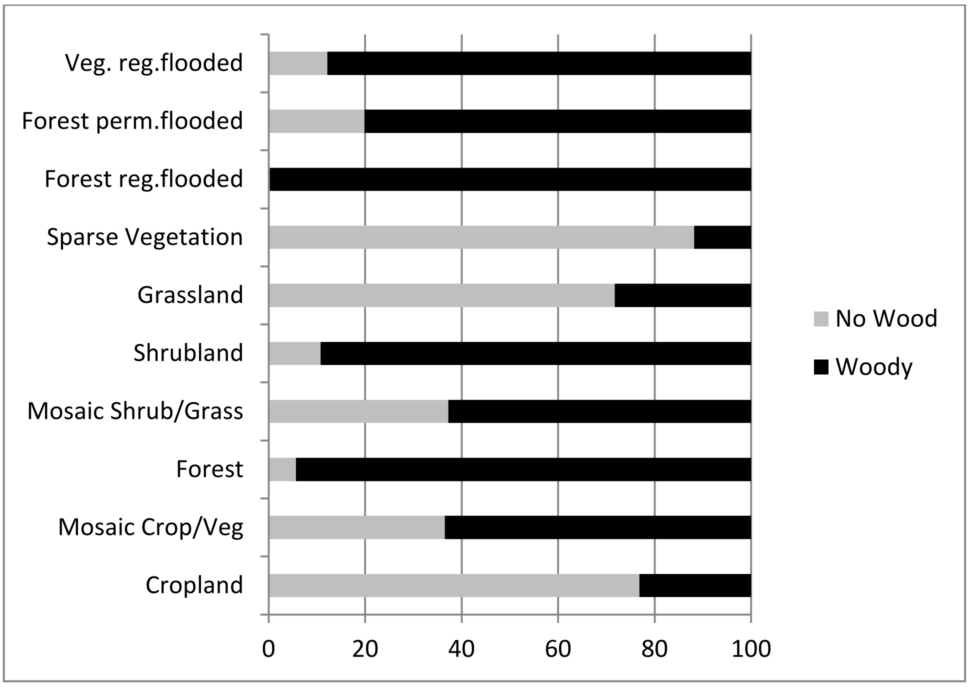

Figure 17: Distribution of woody vegetation and non-woody vegetation in percentage in the different land cover zones according to GlobCover 2009. The ratio among all pixels within a land cover class was taken into account.

All impact variables were standardized as their value range was not equally distributed and to furthermore eliminate false alarms. Logistic regression explaining woody vegetation occurrence with regard to the five main trigger factors showed significance for all variables (see Table 6).

Table 6: Logistic Regression Modeling for woody vegetation prediction.

	Odds Ratio	Std. Err.	P>z		[95% Conf. Intervall]	
Cattle Density	1.03	0.01	0.000	***	1.02	1.04
Precipitation (MAP)	6.71	0.06	0.000	***	6.60	6.82
Soil Moisture	1.52	0.01	0.000	***	1.51	1.54
Fire Occurrence	2.41	0.03	0.000	***	2.36	2.46
CO^2 Emissions	1.04	0.01	0.002	**	1.02	1.07

As land cover plays a role with regard to BE, especially in grasslands and bushlands, logistic regression modeling integrated analysis of the behavior of all variables in different land cover classes.

Although, all variables showed high significance, three variables – Precipitation, Soil Moisture and Cattle Density – on the one hand showed higher effect size and thereby higher impact on the model than Fire Occurrence and CO_2 emissions on the other hand. Focusing on the main impact variables in a regression and dropping the two variables on fire occurrence and CO_2 emissions while linking the information to land cover still resulted in a slightly similar explanation of the variance of the input variable with an R^2 of 0.61(dropping Fire Occurrence and CO_2 emissions) compared to 0.67 (including all five possible trigger variables).

Based on logistic regression the predicted values were calculated. Statistics and ROC were calculated for both models as described above: the three main impact variables Precipitation, Soil Moisture and Livestock Density and a model including all five variables. Both models were calculated in dependency of land cover and without. As forest areas were dropped during the analysis, as they obviously show high probabilities of woody vegetation, these classes were masked.

With a receiver operating characteristic (ROC) the different logistic regression models were assessed for their accuracy in predicting woody vegetation that could be related to BE. The ROC is used to assess the quality of the model and compare different model scenarios (Pontius & Schneider, 2001). The two different scenarios as described above were taken into account. One includes (i) all five impact variables and (ii) only the three variables with the highest impact, both considering the effect of land cover. The second scenario was not linked to land cover information and tested the general effect of the two different sets of variables on woody vegetation patterns.

According to the ROC statistics (Table 7) and the ROC curves (Figure 18) the model including all 5 variables in dependency to land cover is reporting the best results (LC5var).It reports an accuracy of around 93 %. When only three variables were included accuracy is only dropping marginal by 0.01 %. If land cover information is neglected logistic regression modeling is still able to predict 82 % or 83 % respectively of woody vegetation cover linked to bush encroachment in a 10 by 10 km grid.

Table 7: ROC Statistics.

		Area	Std. Err.	Correctly Classified
Land Cover Dependent	3 variables (MAP, SM, Cattle Dens)	0.93	0.03	0.88
	all 5 impact variables	0.93	0.03	0.89
Non-Land Cover Dependent	3 variables (MAP, SM, Cattle Dens)	0.84	0.05	0.86
	all 5 impact variables	0.82	0.06	0.87

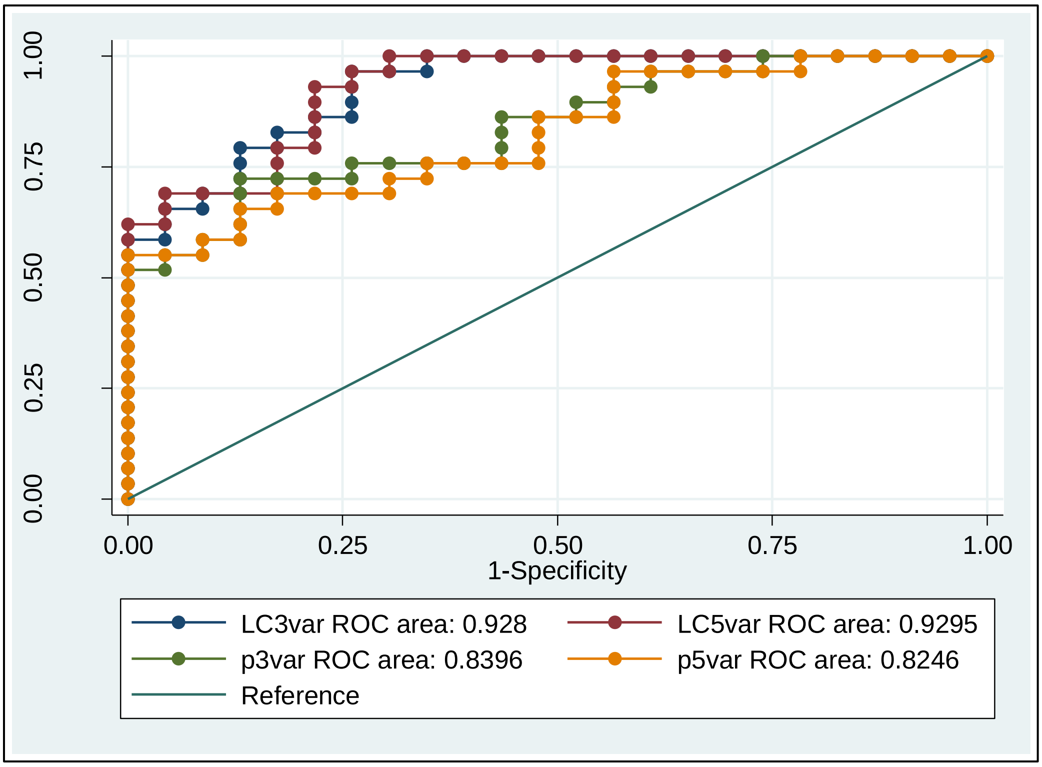

Figure 18: ROC Curves for four possible models chosen via logistic regression modeling and analysis of impact variables. LC3var describes the model including the three main impact variables (Precipitation, Soil Moisture and Cattle Density) in dependency of land cover, LC5var includes all five main impact variables discussed in this section. P3var and p5var furthermore refer to the models including all five impact variables (p5var) and only three main impact variables (p3var) without being linked to land cover information.

Even though the results for the model including all five variables represented the best but model validation did not show high discrepancy. Figure 19 represents the probability map resulting out of the logistic regression modeling including the three variables with the highest impact considering land cover information. Compared with a second map which we calculated for all five impact variables we could observe that both maps showed high probability (dark brownish) of occurring woody vegetation triggered by the included variables in Eastern Africa. Here, especially Ethiopia and parts of Kenya and Uganda as well as South Sudan were highlighted. But also parts of western Africa show high probability. It was also interesting that dropping the two variables with less impact – fire occurrence and CO_2 emissions – did not result in a big shift regarding the probability patterns in most parts of Africa. Slightly different patterns again could be found in Easter Africa, especially Ethiopia and South Sudan.

But we could also observed that those areas were BE is reported to be a severe problem were not depicted with high probabilities. This is e.g. valid for Namibia and parts of South Africa. Moreover, high probabilities are depicted in western Africa where BE is not stated as a problem as such. Regarding the map this statement cannot be confirmed.

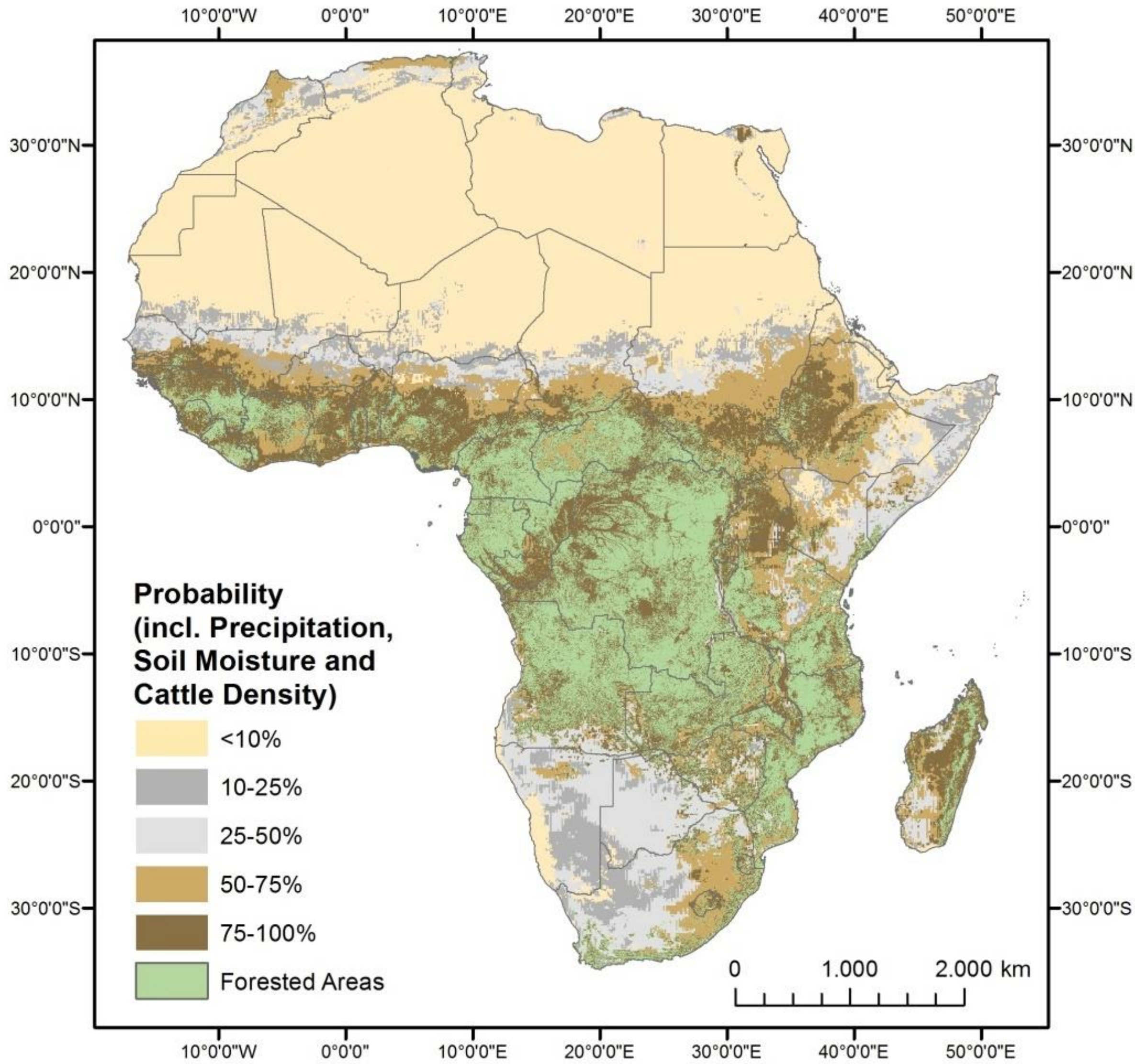

Figure 19: Probability of bush encroachment based on the prediction for woody vegetation of 2007 and 2009 including three main impact variables: precipitation, soil moisture and cattle density under consideration of land cover.

5. Discussion

BE mapping so far has only been carried out on local scales with a few examples representing national approaches. The objective of this study was to develop a process-chain which helps to upscale field data on BE detection to a much larger extent. Here, based on field data a process chain was developed to detect and predict woody vegetation and BE on the African continent. Field data was available for two regions in Africa: two study sites in South Africa and one study site in Ethiopia. Both training datasets represent rather small regions. Nevertheless we tested to upscale this data to a larger area. Due to the large spatial extent, this study covers multispectral datasets in different resolutions. A trend analysis of NDVI was integrated to get closer to the analysis of BE as a process happening over time. Significant increasing NDVI trends and the prediction of woody vegetation based on the results derived from the regression modeling of field data helped to identify those areas where woody vegetation cover is possibly linked to BE. NDVI data was furthermore used to mask urban or water covered areas.

Determined by the input field data, the first upscale was conducted with Landsat imagery of 30 m resolution and the second upscale based on MODIS imagery with 500 m resolution. The input field data for South Africa came with a higher resolution than the one of Ethiopia. Field data from South Africa had a resolution of 0.5 m (2007) and 0.75 m (2009). The dataset for Ethiopia was only available with a 15 m-resolution as it was based on ASTER imagery and thereby provided a much lower resolution compared to the two study sites in South Africa. Here, the first up-scaling to Landsat imagery resulted in weak regression results and therefore was upscaled to MODIS imagery with 500 m resolution only. Even though regression results, referring to a relatively high R^2 (0.65) and a low RMSE (0.04), were performing more than acceptable the output dataset showed significant differences compared to the predicted datasets based on the high resolution field data from South Africa. Integrating this data for further analysis was thereby less promising. The datasets were comparable but probabilities for woody vegetation were much higher in the prediction for 2013, generated with input data from Ethiopia, which is due to a finer resolution of the South Africa dataset that resulted in the predictions for 2007 and 2009. But we could observe equally distributed patterns in all predictions. The higher the resolution of the input dataset the more predictions were comparable with land cover classes in the respective regions with regard to the percentage of woody vegetation cover. According to these findings, the 2007 prediction relying on the input field data with the highest resolution reported the best results. We can agree with Gessner et al. (2013) stating that high-resolution data for accurate up-scaling with regression tree modeling is necessary. Although, the study area of Gessner et al. (2013) only covered small regions in Namibia including two training sites both within Savanna ecosystems, these conclusions are also valid when up-scaling to a much larger extent which could be confirmed by our study.

NDVI trend analysis was conducted to detect significant increasing and decreasing vegetation trends with non-parametric Mann Kendall trend analysis during the observation period from 1982 to 2013. Increasing vegetation trends not only refer to e.g. the use of innovative technologies meaning the application of fertilizer or improved irrigation techniques that result in increasing vegetation cover and/or density. These trends can also be related to BE processes. Therefore an overlay was processed in a GIS to identify areas where significant increasing vegetation trends overlap with woody vegetation detection according to the predictions for 2007 and 2009. Even though the dataset of Ethiopia was not included in the hotspot analysis it could be shown that the Baadu case study site analyzed by Ayanu et al. (2013) which showed invasion processes by *prosopis juliflora* – a BE species – was located in one of the detected hotspots. Besides that we found many detected hotspots in western Africa where according to literature and experts the occurrence of BE is rather low. Also did we observe a lack of hotspots in Namibia and South Africa where we expected them. We assume that due to an increase of more high resolution training data these results might improve. But in general we would agree that a combination of vegetation development and detection of woody vegetation represent a good method to get more information about the presence of BE considering the inclusion of more good training data.

There is still ongoing discussion about the factors triggering the process of BE. Different variables are still approached to detect if they favor the invasion of bushes by suppressing grasslands. Five main variables were identified to explain predicted woody vegetation cover based on BE study sites. These were: precipitation (mean annual precipitation (MAP)), soil moisture, fire occurrence, cattle density and CO_2 emissions. Logistic regression modeling showed significance for all five variables. Nevertheless two variables, fire occurrence and CO_2 emissions, resulted in a lower impact for the model. Therefore, two predictions including land cover information were calculated and mapped. Results only had slight changes but in all cases highlighted parts of the Savanna regions in western Africa and especially Eastern Africa with Ethiopia and southern Sudan. It has to be mentioned that those indicators are chosen according to the literature and there is no agreement on their impact but still indetermination if they trigger BE or not and to what extent. The accuracy of the chosen models was analyzed with ROC statistics. All models showed a high accuracy reporting correctly classified pixel for the prediction of around 90 %. The model with the best results included all five impact variables including also land cover information and the individual behavior of these variables within each class. To improve this study further analysis might also search for other impact variables that might play a role. But of course data availability on a large scale as the African continent might be a problem.

The combination of different approaches helps to get a better understanding of BE processes at continental scale. Nevertheless, high resolution data is necessary to train the model and finally get the most accurate results. This was already reported in studies before and could be shown at continental level. Although, there is still doubt whether certain indicators actually favor BE, logistic regression modeling could help to prove the significance of all impact variables including mean annual precipitation, soil moisture, cattle density, fire occurrence and CO_2 emissions. There might be individual cases on the global scale where certain indicators have a bigger impact than others. This approach helped to get a regional view of the problem and detect possible BE patterns as well as regions that are vulnerable to BE due to their preconditions.

Further approaches should take into consideration using high-resolution images input datasets to increase the accuracy of the results. In our study, impact variables were chosen by literature review. With a geodatabase for Africa and woody vegetation hotspot modeling as shown in this study also indicators that might be related to BE but are not yet under consideration could provide more detailed information and especially insight in the process of BE. Input of more high resolution field data from different parts of Africa moreover would be useful as species might differ in their presence in certain areas. Western Africa so far has only a very few studies focusing the problem of BE. According to our studies there are certain patterns in western Africa where BE might be a problem as also preconditions would favor BE.

6. Conclusion and Outlook

Bush encroachment is a global problem and often refers in literature to a process of environmental degradation. In Africa pastoralists are directly affected by the invasion of woody vegetation and the simultaneous suppression of grasses due to the increasing of unpalatable species for their livestock. Therefore, analysis on the extent of BE and the variables that influence this process was approached in this study with a multi-scale approach using remote sensing and GIS. Two main outputs were developed: prediction of woody vegetation cover and a probability map for BE, both for the African continent. The combination of remote sensing and GIS here proved to be a suitable tool to get information about processes occurring on a larger scale to go beyond the local analysis. By detecting hotspots of BE and regions that could be affected due to their preconditions, which favor impact factors triggering BE, this study contributed to research to protect pastoral environments and rural areas in Africa and thereby support sustainable land use. Continental and even global mapping approaches can only give a more generalized view on local processes. Nevertheless, they support the detection of areas where further research is necessary via employing detailed, local analysis and for guidance of the field campaigns.

Our study showed that up-scaling from field data to the continental level can be conducted with a process chain based on the random forests regression trees. The resolution of the input dataset determined the reliability of the results. The woody vegetation cover linked to BE was predicted for whole Africa using mainly two field BE datasets from two different areas in South Africa and one area in Ethiopia. The BE map based on the input dataset with the highest resolution – 0.5 m – lead to the best results (with R^2 = 0.87 for up-scaling to the MODIS-level) taking into account the performance of predicted versus validation data. A dataset from Ethiopia with a resolution of 15 m per pixel e.g. lead to less promising (with R^2 = 0.65 for the same upscale).

Multiple indications can be taken into account to refer to BE processes. Besides impact or trigger variables that are linked to BE according to multiple studies conducted in Africa also NDVI trend analysis could provide further information. While decreasing NDVI trends are mainly linked to degradation processes increasing trends can state the process of BE; BE describes the invasion and increasing woody vegetation not naturally occurring in the area. The overlay of significant increasing NDVI trends with prediction of woody vegetation cover derived from field data could detect areas where BE patterns are reported. A future work should be conducted on integration of more field data in the developed approach to further validate the produced BE map for African continent.

Logistic regression modeling was conducted including five variables that were reported to impact BE according to literature research. Dependent variable was based on the woody vegetation prediction. The explanatory variables included mean annual precipitation, soil moisture, cattle density, fire occurrence and CO_2 emissions. All variables showed significance even though variables such as precipitation and soil moisture had a much higher impact than the other variables.

This study aimed at the development of a multi-scale approach to detect BE on the continental level. Training data for this study was only available for three study sites in Southern and Eastern Africa which could not represent the African continent sufficiently. The here presented approach could highlight some areas where BE is actually occurring and stated as a problem but neglected others and presented false results in regions where BE is not reported to be occurring. Africa is highly diverse with regard to its many agro-ecological zones. For further analysis besides the integration of more training samples represented by e.g. Google Earth Imagery and RGB-classification techniques, we favor a more ecosystem-driven approach which additionally includes local drivers of bush encroachment instead of focusing on general impact factors.

Our study showed that analysis on different scales with the use of different methods leads to promising results to detect and analyze a complex problem such as BE. But still further research is needed which includes the integration of more high resolution data and a local analysis of impact variables to not generalize individual BE species and get more reliable results.

References

Arino, O., Ramos, P.,Jose, J. Kalogirou, V., Bontemps, S. Defourny, P., Van Bogaert, E., 2012. Global Land Cover Map for 2009 (GlobCover 2009). European Space Agency (ESA) & Université catolique de Louvain (UCL).

Angassa A, Oba G. 2008. Effects of management and time on mechanisms of bush encroachment in southern Ethiopia. Afr. J. Eco., 46, 186-196.

Ayanu, Y., Jentsch, A., Müller-Mahn, D., Rettberg, S., Romankiewicz, C., Koellner, T., 2014. Ecosystem engineer unleashed: Prosopis juliflora threatening ecosystem services? Regional Environmental Change 15, 155–167.

Bai, Y., Broersma, K., Thompson, D. and T. J. Ross, 2004. Landscape-level dynamics of grassland-forest transitions in British Columbia. Journal of Range Mangement. 57 (1), 66-75.

Bai, Z.G.; de Jong, R. van Lynden, S. 2010. An Update of GLADA - Global Assessment of Land Degradation and Improvement. GLADA Report Update. ISRIC Report 2010/8. Available via: http://library.wur.nl/WebQuery/file/isric/fulltext/isricu_t4e524398_001.pdf

Bamford, A.J., Diekmann, M., Monadjem, A., Mendelsohn, J., 2007. Ranging behaviour of Cape Vultures Gyps coprotheres from an endangered population in Namibia. Bird Conservation International. 17, 331–339.

Bond, W.J., 2005. Large parts of the world are brown or black: A different view on the "Green World"hypothesis. Journal of Vegetation Science. 16, 261–266.

Bond, W.J., Midgley, G.F., 2012. Carbon dioxide and the uneasy interactions of trees and savannah grasses. Philosophical Transactions of The Royal Society. 367, 601–612. doi:10.1098/rstb.2011.0182

Bond, W.J., Midgley, G.F., Woodward, F.I., 2003. The importance of low atmospheric CO2 and fire in promoting the spread of grasslands and savannas. Global Change Biology. 9, 973–982.

Breiman, L. 2001. Random Forest. Machine Learning 45, 5-32.

Buitenwerf, R., Bond, W.J., Stevens, N., Trollope, W.S.W., 2012. Increased tree densities in South African savannas: >50 years of data suggests CO2 as a driver. Global Change Biology 18, 675–684.

Coetzee, B.W.T., Tincani, L., Wodu, Z., Mwasi, S.M., 2008. Overgrazing and bush encroachment by Tarchonanthus camphoratus in a semi-arid savanna. African Journal of Ecology 46, 449–451.

Donohue, R.J., Roderick, M.L., McVicar, T.R., Farquhar, G.D., 2013. Impact of CO2 fertilization on maximum foliage cover across the globe's warm, arid environments. Geophysical Research Letters. 40, 3031–3035.

de Klerk, J.N., 2004. Bush Encroachment in Namibia: Report on phase 1 of the bush -encroachment Research, Monitoring and Management Project. John Meinert -Printing, Windhoek.

Dube, O.P., 2007. Fire Weather and Land Degradation. In: Sivakuma, M. & Ndiang´ui, Ndegwa (eds): Climate and Land Degradation. Springer Science and Business Media. Heidelberg.

Dubovyk, O., Menz, G., Conrad, C., Kan, E., Machwitz, M., Khamzina, A., 2013. Spatio-temporal analyses of cropland degradation in the irrigated lowlands of Uzbekistan using remote-sensing and logistic regression modeling. Environmental Monitoring and Assessment, 185(6), 4775–4790.

FAO, 2011. The state of the world's land and water resources for food and agriculture (SOLAW) – Managing systems at risk. Food and Agriculture Organization of the United Nations, Rome and Earthscan, London.

FAO, 2013. Tackling Climate Change through livestock. A Global Assessment of Emissions and mitigation opportunities. Rome, Italy. Available via: http://www.fao.org/3/8d293990-ea82-5cc7-83c6-8c6f461627de/i3437e.pdf

Fensham, R. J., R. J. Fairfax & S. Archer, 2005. Rainfall, land use and woody vegetation cover change in semi-arid Australian savanna. Journal of Ecology 93:596–606.

Gessner,, U., Machwitz, M., Conrad, C. & S. Dech, 2013. Estimating the fractional cover of growth forms and bare surface savannas. A multi-resolution approach based on regression tree ensemble. Remote Sensing of Environment. Vol. 129, 90-102.

Giglio, L., Descloitres, J., Justice, C.O. & Y.J. Kaufman, 2003. An Enhanced Contextual Fire Detection Algorithm for MODIS. Remote Sensing of Environment. Vol. 87, 273-282.

Govender, N., Trollope, W.S.W., Van Wilgen, B.W., 2006. The effect of fire season, fire frequency, rainfall and management on fire intensity in savanna vegetation in South Africa. J. Appl. Ecol. 43, 748–758.

Grover, H.D., Musick, H.B., 1990. Shrubland encroachment in southern New Mexico, U.S.A.: An analysis of desertification processes in the American southwest. Climate Change 17, 305–330.

Higgins, S.I., Scheiter, S., Sankaran, M., 2010. The stability of African savannas: insights from the indirect estimation of the parameters of a dynamic model. Ecology 91, 1682–1692.

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G., Jarvis, A., 2005. Very high resolution interpolated climate surfaces for global land areas. International Journal Climatology, 25, 1965–1978.

Hudak, A.T., 1999. Rangeland Mismanagement in South Africa: Failure to Apply Ecological Knowledge. Human Ecology. Vol. 27, 55-78.

Hudak, A.T., Wessman, C.A., 2001. Textural analysis of high resolution imagery to quantify bush encroachment in Madikwe Game Reserve, South Africa, 1955-1996. International Journal of Remote Sensing 22, 2731–2740.

Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X., L.G. Ferreira, 2002. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. Remote Sensing of Environment, 83(1–2), 195–213.

IFAD, 2010. Rural Poverty Report. New realities, new challenges: new opportunities for tomorrow´s generation. Quintly.

Ifejika Speranza, C., Kiteme, B., Wiesmann, U., 2008. Droughts and famines: The underlying factors and the causal links among agro-pastoral households in semi-arid Makueni district, Kenya. Global Environmental Change 18, 220–233.

Jiang, Z., Huete, A. R., Didan, K., Miura, T., 2008. Development of a two-band enhanced vegetation index without a blue band. Remote Sensing of Environment, 112(10), 3833–3845.

Joubert, D.F., Rothauge, A., Smit, G.N., 2008. A conceptual model of vegetation dynamics in the semiarid Highland savanna of Namibia, with particular reference to bush thickening by Acacia mellifera. Journal of Arid Environments 72, 2201–2210.

Kgosikoma, O., Mojeremane, W., B.A. Harvie, 2013. Grazing management systems and their effects on savanna ecosystem dynamics. A review. Journal of Ecology and the Natural Environment. Vol. 5(6), 88-94. June 2013. Kgope, B.S., Bond, W.J., Midgley, G.F., 2010. Growth responses of African savanna trees implicate atmospheric [CO2] as a driver of past and current changes in savanna tree cover. Austral Ecology 35, 451–463.

Kraaij, T., Ward, D., 2006. Effects of rain, nitrogen, fire and grazing on tree recruitment and early survival in bush-encroached savanna, South Africa. Plant Ecology 186, 235–246.

Kruger, B., 2002. Bush encroachment Research, Monitoring and Management Project - (BERMMP): Bio-physical and Socio-economic impact of Bush Encroachment in the - Communal areas of Namibia. Desert Research Foundation of Namibia, Windhoek

Kulmatiski, A., Beard, K.H., 2013. Woody plant encroachment facilitated by increased precipitation intensity. Nature Climate Change 3, 833–837.

Lehmann, C.E.R., Ratnam, J., L.B. Hutley, 2009. Which of these continents is not like the other? Comparisons of tropical savanna systems: key questions and challenges. New Phytol. 181, 508–511.

Lesoli, M.S., Gxasheka, M., Solomon, T.B., B. Moyo, 2013. Integrated Plant Invasion and Bush Encroachment Management on Southern African Rangelands. In: Price, A.J. and J.A.Kelton (eds.): Herbicides - Current Research and Case Studies in Use. Chapter 11, 259-313.

Liu, Y. Y., W. A. Dorigo, et al., 2012. "Trend-preserving blending of passive and active microwave soil moisture retrievals." Remote Sensing of Environment 123: 280-297.

Liu, Y. Y., Parinussa, R. M., Dorigo, W. A., De Jeu, R. A. M., Wagner, W., van Dijk, A. I. J. M., McCabe, M. F., Evans, J. P., 2011. Developing an improved soil moisture dataset by blending passive and active microwave satellite-based retrievals. Hydrology and Earth System Sciences, 15, 425-436

MEA, 2005. Millennium Ecosystem Assessment, Ecosystems and Human Well-being: Synthesis. Island Press, Washington, D.C. S.

Meik, J.M., Jeo, R.M., Mendelson III, J.R., Jenks, K.E., 2002. Effects of bush encroachment on an assemblage of diurnal lizard species in central Namibia. Biol. Conserv. 106, 29–36.

Mitchard, E. T. A., & Flintrop, C. M., 2013. Woody encroachment and forest degradation in sub-Saharan Africa’s woodlands and savannas 1982–2006. Philosophical Transactions of the Royal Society B: Biological Sciences, 368(1625), 20120406.

Moleele, N.M., Ringrose, S., Matheson, W., Vanderpost, C., 2002. More woody plants? the status of bush encroachment in Botswana’s grazing areas. Journal of Environmental Management. 64, 3–11.

Mwangi, E., Swallow, B., 2008. Prosopis juliflora invasion and Rural livelihoods in the Lake Baringo area of Kenya. Conservation and Society 6 (2), 130-140

Nachtergaele, F., Petri, M., Biancalani, R., Van Lynden, G., & Van Velthuizen, H., 2010. Global Land Degradation Information System (GLADIS). Beta Version. An Information Database for Land Degradation Assessment at Global Level. Land Degradation Assessment in Drylands. FAO.

Naito, A. T., Cairns, D. M., 2011. Patterns and processes of global shrub expansion. Progress in Physical Geography 35:423–442

Negasa, B., Eba, B., Tuffa, S., Bayissa, B., Doyo, J., Husen, D. 2014. Control of bush encroachment in Borana zone of southern Ethiopia: effects of different control techniques on rangeland vegetation and tick populations. Pastoralism: Research, Policy and Practice. Vol. 4(18).

Nkonya, E., N. Gerber, P. Baumgartner, J. von Braun, A. de Pinto, V. Graw, E. Kato, J. Kloos, Walter, T., 2011. The Economics of Land Degradation. Toward an Integrated Global Assessment. Peter Lang, Frankfurt am Main, Berlin, Bern, Bruxelles, New York, Oxford, Wien, Development Economics and Policy - Volume 66*. Further Informati*on:

Oldeland, J. , Dorigo, W., Wesuls, D., Jürgens N., 2010. Mapping Bush Encroachment Species by Seasonal Differences in Hyperspectral Imagery. Remote Sensing. Vol 2(6), 1416-1438.

Peng, H.-Y., Li, X.-Y., Li, G.-Y., Zhang, Z.-H., Zhang, S.-Y., Li, L., Zhao, G.-Q., Jiang, Z.-Y., Ma, Y.-J., 2013. Shrub encroachment with increasing anthropogenic disturbance in the semiarid Inner Mongolian grasslands of China. CATENA 109, 39–48.

Pinzon J.E., Brown M.E., Tucker, C.J., 2007. Global Inventory Modeling and Mapping Studies (GIMMS) Satellite Drift Corrected and NOAA-16 incorporated Normalized Difference Vegetation Index (NDVI), Monthly 1981-2006.

Pinzon J.E. and C.J. Tucker, 2014. A Non-Stationary 1981-2012 AVHRR NDVI3g Time Series. Remote Sensing, 6, 6929-6960.

Pontius Jr., R.G., Schneider, L.C., 2001. Land-cover change model validation by an ROC method for the Ipswich watershed, Massachusetts, USA. Agriculture, Ecosystems and Environment. Vol 85, 239-248.

Purchase, G., Marker, L., Marnewick, K., Klein, R., Williams, S., 2007. Regional Assessment of the Status, Distribution and Conservation Needs of Cheetahs in Southern Africa.

Rayner, P. J.; Raupach, M. R.; Paget, M.; Peylin, P., Koffi, E.,2010. A new global gridded data set of CO^2 emissions from fossil fuel combustion: Methodology and evaluation. In: Journal of Geophysical Research Vol. 115.

Rodrigues-Galiano, V.F., Ghimire, B., Rogan, J., Chica-Olmo, M., Rigol-Sanchez, J.P., 2012. An assessment of the effectiveness of a random forest classifier for land-cover classification. ISPRS Journal of Photogrammetry and Remote Sensing, Vol. 67, 93-104.

Roques, K. G., O’Connor, T. G., Watkinson, A. R., 2001. Dynamics of shrub encroachment in an African savanna: relative influences of fire, herbivory, rainfall and density dependence. J. Appl. Ecol. 38, 268–280.

Saha, M.V., Scanlon, T.M., D`Odorico, P. 2015. Examining the linkage between shrub encroachment and recent greening in water-limited southern Africa. Ecosphere 6 (9): 156.

Sankaran, M., Hanan, N.P., Scholes, R.J., Ratnam, J., Augustine, D.J., Cade, B.S., Gignoux, J., Higgins, S.I., Le Roux, X., Ludwig, F., Ardo, J., Banyikwa, F., Bronn, A., Bucini, G., Caylor, K.K., Coughenour, M.B., Diouf, A., Ekaya, W., Feral, C.J., February, E.C., Frost, P.G.H., Hiernaux, P., Hrabar, H., Metzger, K.L., Prins, H.H.T., Ringrose, S., Sea, W., Tews, J., Worden, J., Zambatis, N., 2005. Determinants of woody cover in African savannas. Nature 438, 846–849.

Scholes, R.J., Archer, S.R., 1997. Tree-Grass Interactions in Savannas. Annu. Rev. Ecol. Syst. 28, 517–544.

Symeonakis, E., Higginbottom, T.P., 2014. Bush encroachment monitoring using multi-temporal Landsat data and random forests. ISPRS - The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences Volume XL-2, 2014 ISPRS Technical Commission II Symposium, 6 – 8 October 2014, Toronto, Canada

Trollope, W.S.W., 1980. Controlling bush encroachment with fire in the savanna areas of South Africa. Proc. Annu. Congr. Grassl. Soc. South. Afr. 15, 173–177.

van Auken, O.W., 2000. Shrub Invasions of North American Semiarid Grasslands. Annu. Rev. Ecol. Syst. 31, 197–215.

Van Langevelde, F., Van De Vijver, C.A.D.M., Kumar, L., Van De Koppel, J., De Ridder, N., Van Andel, J., Skidmore, A.K., Hearne, J.W., Stroosnijder, L., Bond, W.J., Prins, H.H.T., Rietkerk, M., 2003. Effects of fire and herbivory on the stability of savanna ecosystems. Ecology 84, 337–350.

Wagner, W., W. Dorigo, R. de Jeu, D. Fernandez, J. Benveniste, E. Haas, Ertl, M., 2012. Fusion of active and passive microwave observations to create an Essential Climate Variable data record on soil moisture. ISPRS Annals of the Photogrammetry, Remote Sensing and

Spatial Information Sciences (ISPRS Annals), Volume I-7, XXII ISPRS Congress, Melbourne, Australia, 25 August-1 September 2012, 315-321

Walker B.H., Noy-Meir, I., 1982. Aspects of the stability and resilience of savanna ecosystems. In: Huntley BJ and B.H. Walker (eds.). Ecology of tropical savannas. Berlin: Springer, 556–590.

Walter, H., 1939. Grasland, Savanne und Busch in der arideren Teile Afrikas in ihrer ökologischen Bedingtheit. Jahrbücher für Wissenschaftliche Botanik. Band 87

Walter, H., 1981. Ecology of tropical and subtropical vegetation. Edinburgh. Oliver & Boyd.

Ward, D., 2005. Do we understand the causes of bush encroachment in African savannas? African Journal of Range and Forage Science. 22, 101–105.

Ward, D., Wiegand, K., Getzin, S., 2012. Walter's two-layer hypothesis revisited: back to the roots! Oecologia 172, 617–630.

Welz, A., 2013. The Surprising Role of CO_2 Changes on the African Savanna. Environment 360. Available via: http://e360.yale.edu/feature/the_surprising_role_of_co2_in_changes_on_the_african_savanna/2663/

Wiegand, K., Ward, D., Saltz, D., 2005. Multi-scale patterns and bush encroachment in an arid savanna with a shallow soil layer. J. Veg. Sci. 16, 311–320.

Wigley, B.J., Bond, W.J., Hoffman, M.T., 2009. Bush encroachment under three contrasting land-use practices in a mesic South African savanna. African Journal of Ecology 47, 62–70.

Williams, R.J., Cook, G.D., Gill, A.M., Moore, P.H.R., 1999. Fire regime, fire intensity and tree survival in a tropical savanna in northern Australia. Australian Journal of Ecology 24, 50–59.

Wise, R.M., van Wilgen, B.W., Le Maitre, D.C., 2012. Costs, benefits and management options for an invasive alien tree species: The case of mesquite in the Northern Cape, South Africa. Journal of Arid Environments, 84, 80–90.

Zachariades, C., Hoffmann, J., Roberts, A., 2011. Biological control of mesquite (prosopis species)(fabaceae) in South Africa. African Entomology. 19 (SP), 402–415