WOODY VEGETATION COVER IN NAMIBIAN SAVANNAHS: A MODELLING APPROACH BASED ON REMOTE SENSING

With 7 figures, 1 table and 3 photos

HENDRIK WAGENSEIL and CYRUS SAMIMI

Keywords: Wood cover, PLS regression, Landsat ETM, SPOT VGT, Namibia
Gehölzdichte, PLS-Regression, Landsat ETM, SPOT VGT, Namibia

Zusammenfassung: Die Gehölzdichte in den Savannen Namibias: eine fernerkundungsgestützte Modellierung

Savannen sind durch eine Koexistenz von Gräsern und Gehölzen gekennzeichnet. Sie zählen heute zu den am dichtesten besiedelten Räumen der Tropen und unterliegen daher verstärkt anthropogenen Eingriffen. In den Trocken- und Dornsavannen des südlichen Afrikas, die vorwiegend weidewirtschaftlich genutzt werden, haben unangepasste Bestockungsraten vielerorts eine starke Verbuschung zur Folge, so dass sich die Anteile von Gräsern und Gehölzen verschieben und es zu einem Rückgang der Tragfähigkeit kommt. Ähnliches ist auch in Nationalparks zu beobachten, v.a. wenn durch Wildzäune die Wanderungsbewegungen der Tierpopulationen eingeschränkt werden.

Zur Erfassung derartiger Veränderungen der Vegetationsstruktur wird in dieser Studie ein Verfahren zur Modellierung der Gehölzdichte unter Verwendung grob aufgelöster, multitemporaler Satellitendaten vorgestellt. Zur Kalibrierung der Modellparameter sowie zur Ergebnisvalidierung dient eine Referenzkarte, die aus Felduntersuchungen sowie aus Landsat ETM-Daten mittlerer Auflösung erstellt wird. Der Modellansatz basiert auf der Annahme, dass sich Gräser und Gehölze unter den wechselfeuchten Klimabedingungen hinsichtlich ihrer Phänologie unterscheiden und sich diese Merkmale je nach Anteil beider Wuchsformen auf Bestandesebene und damit auch in Zeitreihen des NDVI manifestieren. Zur Quantifizierung der Unterschiede wird der NDVI-Verlauf über ein Polynom zweiten Grades beschrieben. Die Analyse der Modellparameter zeigt, dass insbesondere die Dauer der Trockenzeit sowie das trockenzeitliche NDVI-Minimum vom Anteil der Sträucher und Bäume abhängen und sich für eine Modellierung der Gehölzdichte eignen. Auf regionaler Maßstabsebene lassen sich somit räumliche Unterschiede in der Gehölzdichte mit Hilfe multitemporaler Satellitendaten erfassen.

Summary: Savannahs are characterized by a co-existence of grasses and trees. Nowadays, these ecosystems are considerably affected by human impact, as they are used for various types of agriculture and support a major part of the population within the tropics. E.g., it is widely accepted that in many regions of Southern Africa, livestock farming with inadequate stocking rates has facilitated bush encroachment and strongly alters vegetation structure and composition in these environments. Similar observations are reported from national parks, especially if fences confine the natural mobility of animals.

In this study, an approach is developed to model woody vegetation coverage from coarse resolution multitemporal NDVI data by making use of the different phenological characteristics of grasses and trees. A second order polynomial is used to describe NDVI curvature and to quantify phenological metrics. For model calibration and result validation, a reference map is initially generated based on extensive field work and medium resolution Landsat ETM data. The results suggest a close relationship between woody cover and dry season curvature and minimum, which are considered to be suitable phenological metrics for a modelling approach. Thus, spatial differences in woody vegetation cover were successfully measured on a regional scale by means of multitemporal NDVI data. The results clearly highlight the potentials of phenological parameters for mapping woody cover.

1 Background and objectives

Savannah ecosystems form a broad transition zone between the humid tropics and the subtropical deserts and are characterized by a co-occurrence of grasses and woody plants (HIGGINS et al. 2000). It is commonly agreed that the actual ratio of both growth forms results from a complex interaction of plant available moisture, plant available nutrients, fire and herbivory (e.g. SCHOLES a. WALKER 1993). Nowadays, land-use practices strongly influence savannah structure in many regions, as they are widely used for various types of agriculture (HIGGINS et al. 1999). In Namibia, and similarly in many regions of Southern Africa, the use of savannahs for livestock farming has facilitated bush encroachment (ROQUES et al. 2001; DE KLERK 2004, see Photo 1). Spatio-temporal variations in woody vegetation coverage in savannah ecosystems thus not only reflect actual site conditions and disturbances, but also indicate areas of human impact. However, these trends have not been quantified satisfactorily on regional scales. Hence, this study aims to develop an approach to estimate woody vegetation coverage from satellite imagery in order to enable an environmental monitoring on large scales.

Studying biophysical parameters by satellite remote sensing is widely performed by empirical modelling, i.e. a sample of ground-based observations is related to spectral image data of high spatial resolution by means of regression analysis. The relationship in turn is then used to provide spatially continuous estimates of the biophysical variables from the image data (e.g. KRAUS a. SAMIMI 2002; ARMSTON et al. 2004). However, applying these approaches on regional scales is hardly feasible due to high expenses for data acquisition and processing. For example, more than 40 Landsat ETM scenes are required to cover the entire Republic of Namibia.

Alternatively, satellite data sources exist from moderate resolution sensors, e.g. NOAA AVHRR, SPOT VGT and TERRA-MODIS, which are especially designed for frequent imaging of large areas. Whereas their spatial resolution is commonly low (app. 1 x 1 km^2), a short revisiting period allows analysis of phenological patterns from these image data. This is commonly performed by using time series of the Normalized Difference Vegetation Index (NDVI), an image ratio derived from measurements of the red and near-infrared channels (see MATHER 1999), which is related to essential vegetation parameters such as green leaf biomass, leaf area index and percent green cover (e.g. CARLSON a. RIPLEY 1997; DIOUF a. LAMBIN 2001; SANNIER et al. 2002).

The objective of this study is to examine if differences in vegetation site phenology portray variations in woody coverage and can be used for mapping purposes in order to utilize the large area coverage of multitemporal image data. A modelling approach is developed using an averaged data set derived from six years of SPOT VEGETATION imagery. For result verification, an empirical model is initially developed from multispectral Landsat ETM+ data and field observations.

Photo 1: Bush encroachment with *Acacia tortillis*
Verbuschung mit *Acacia tortillis*

Photo 2: Test site with low woody coverage (0–15%)
Aufnahmefläche mit geringer Gehölzdichte (0–15%)

Photo 3: Test site with very high woody coverage (>60%)
Aufnahmefläche mit sehr hoher Gehölzdichte (>60%)

2 *Study area*

Field work was conducted in an extensive area of NW Namibia covering the Etosha National Park (ENP) in the east and extending to the southern part of Kunene region, also termed “Kaokoland”, in the west. The study area is delimited arbitrarily by the coverage of three Landsat ETM scenes (Fig. 1).

Climate within this area varies from semi-arid (west) to arid (east) with a rainy season in summer (November to April). Annual rainfall ranges from app. 450 mm in Etosha to less than 50 mm in the coastal desert on a long-term average (MENDELSOHN et al. 2002, Fig. 1). Vegetation dominantly consists of *Colophospermum mopane*, which is accompanied by *Terminalia prunioides*,

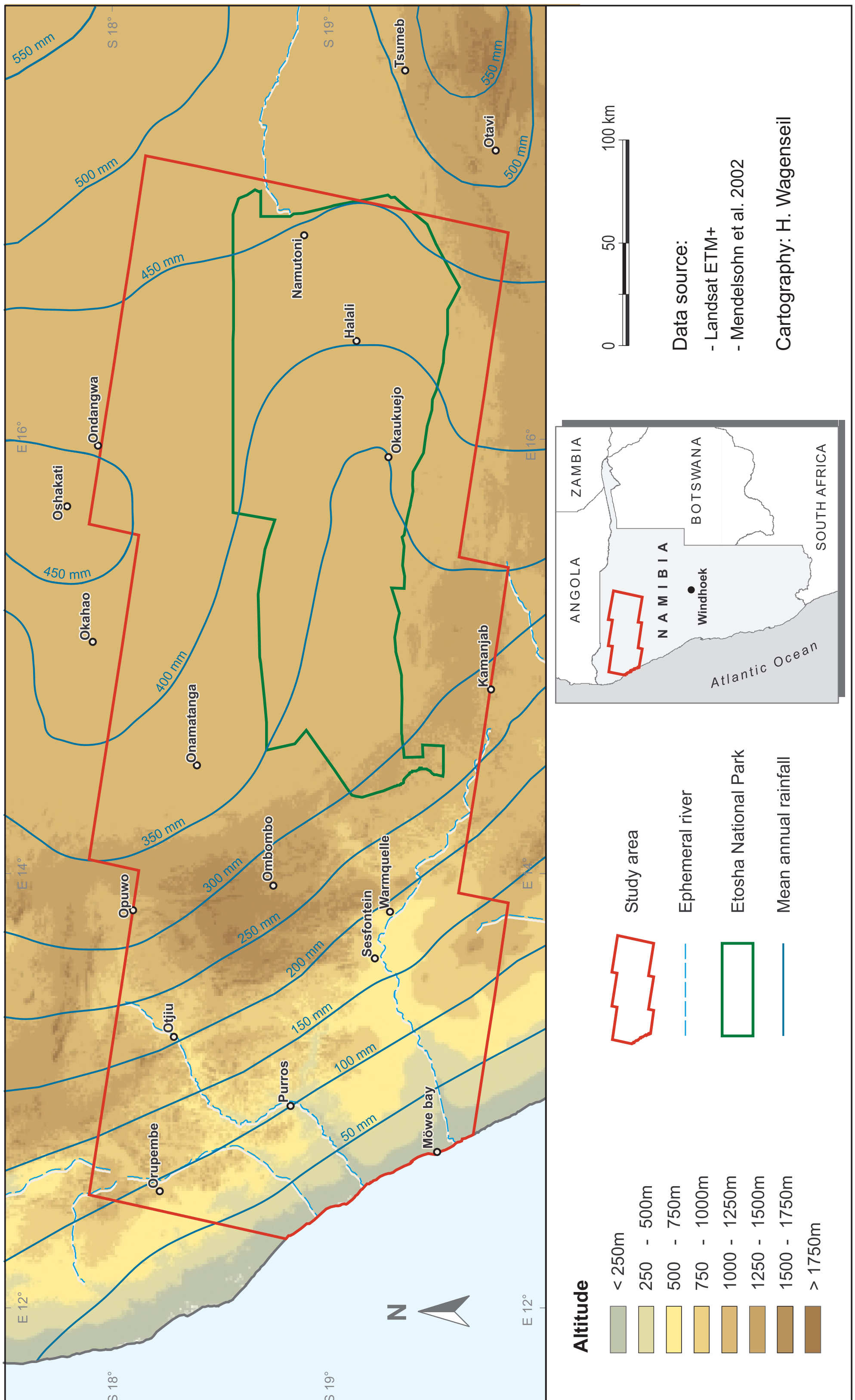

Fig. 1: Location of the study area
Lage des Arbeitsgebietes

Dichrostachys cinerea and different species of *Acacia*. In accordance with the rainfall gradient, vegetation structure changes from dense tree savannah to open shrub savannah, although substantial local variation can be observed (SANNIER et al. 2002; MENDELSOHN et al. 2002).

3 *Materials and methods*

3.1 Empirical modelling from Landsat ETM+

Three scenes of Landsat-7 ETM+ (path/row, acquisition date: 179/73, 15.06.2002; 180/73, 22.06.2002; 181/73, 29.06.2002) were acquired, providing observations of the study area at a spatial resolution of 30 x 30 m^2. An extensive geometric and radiometric pre-processing was conducted for the six solar reflective channels (1–5, 7), including calibration to at-satellite reflectance (HUANG et al. 2001), topographic correction (RIAÑO et al. 2003) and relative image normalization using pseudo-invariant features (FURBY a. CAMPBELL 2001). Finally, all scenes were combined to form one seamless image mosaic covering the entire region.

In order to relate the spectral image data to woody vegetation coverage, ground samples were taken throughout the study area at 102 locations. Each site spans an area of app. 100 x 100 m^2 thus corresponding to 3 x 3 Landsat pixels. Vegetation coverage of shrubs and trees was estimated from each plot and accurate site location within the image was ensured using a GPS. All samples were assigned to six coverage classes in order to reduce the error of estimation (Tab. 1, Photos 2 and 3).

Commonly, ordinary least squares regression (OLSR) is performed to determine a relationship between biophysical data and spectral image values (e.g. KRAUS a. SAMIMI 2002). However, premises of OLSR are commonly violated if multiple spectral variables are included due to multicollinearity of adjacent channels (MATHER 1999). Thus, within this study, partial least squares regression (PLSR, see TOBIAS 2006) was performed, allowing the inclusion of multiple channels or image transforms without addressing the multicollinearity problem in advance. PLSR combines features of principal component analysis and multilinear regression by extracting latent factors from a set of variables that account for most of the predictor variance while modelling the responses well (TOBIAS 2006). All six solar reflective channels as well as NDVI (see COHEN et al. 2003) and Tasseled Cap Greenness and Brightness (see KRAUS a. SAMIMI 2002) were included to predict woody vegetation cover classes. Leave-one-out cross-validation was performed to determine the number of factors that minimize the root-mean-square-error of prediction (RMSEP, Fig. 2a). Based on this decision criterion, the first six factors were chosen for the modelling approach. Final model quality was evaluated by plotting observations of woody coverage classes against PLSR predictions (Fig. 2b). In general there is a very good fit between model output and field data ($R^2 = 0.87$). The approach slightly tends to overestimate sites with low woody vegetation coverage while underestimating high density classes. Nevertheless, the cross-validation proves a high model quality and the empirical relationship between multispectral Landsat image data and woody vegetation density was thus used to estimate woody coverage for the study area (Fig. 3). The map output generally agrees with spatial patterns of woody vegetation observed during the field trips. Along the coast, woody plants are largely absent due to arid climate conditions. Coverage increases towards the east up to very high densities in the mountainous area between Warmquelle and Opuwo. The northwestern part of ENP is mainly formed by grass and shrub savannah with low woody coverage. Disregarding Cuvelai drainage in the North and Etosha pan, proportion of woody plants gradually increases towards the eastern part of the study area. This map is considered to be a reliable and valid reference, which can be used for calibration and result verification of the phenological modelling approach conducted in the next chapter.

Table 1: Coverage classes of woody vegetation (shrubs and trees)
Deckungsklassen der Gehölze (Büsche und Bäume)

Class	Coverage	Number of sites
1	0%	14
2	> 0% bis ≤15%	50
3	>15% bis ≤30%	7
4	>30% bis ≤45%	15
5	>45% bis ≤60%	11
6	>60%	1

3.2 Phenological modelling from SPOT VGT

In dry savannahs with distinct differences between rainy and dry season, plant growth is strictly water-limited and thus, vegetation phenology shows marked seasonal patterns. This even applies for the woody layer, which mainly consists of deciduous species in these environments (SCHULTZ 2000). Nevertheless, evi-

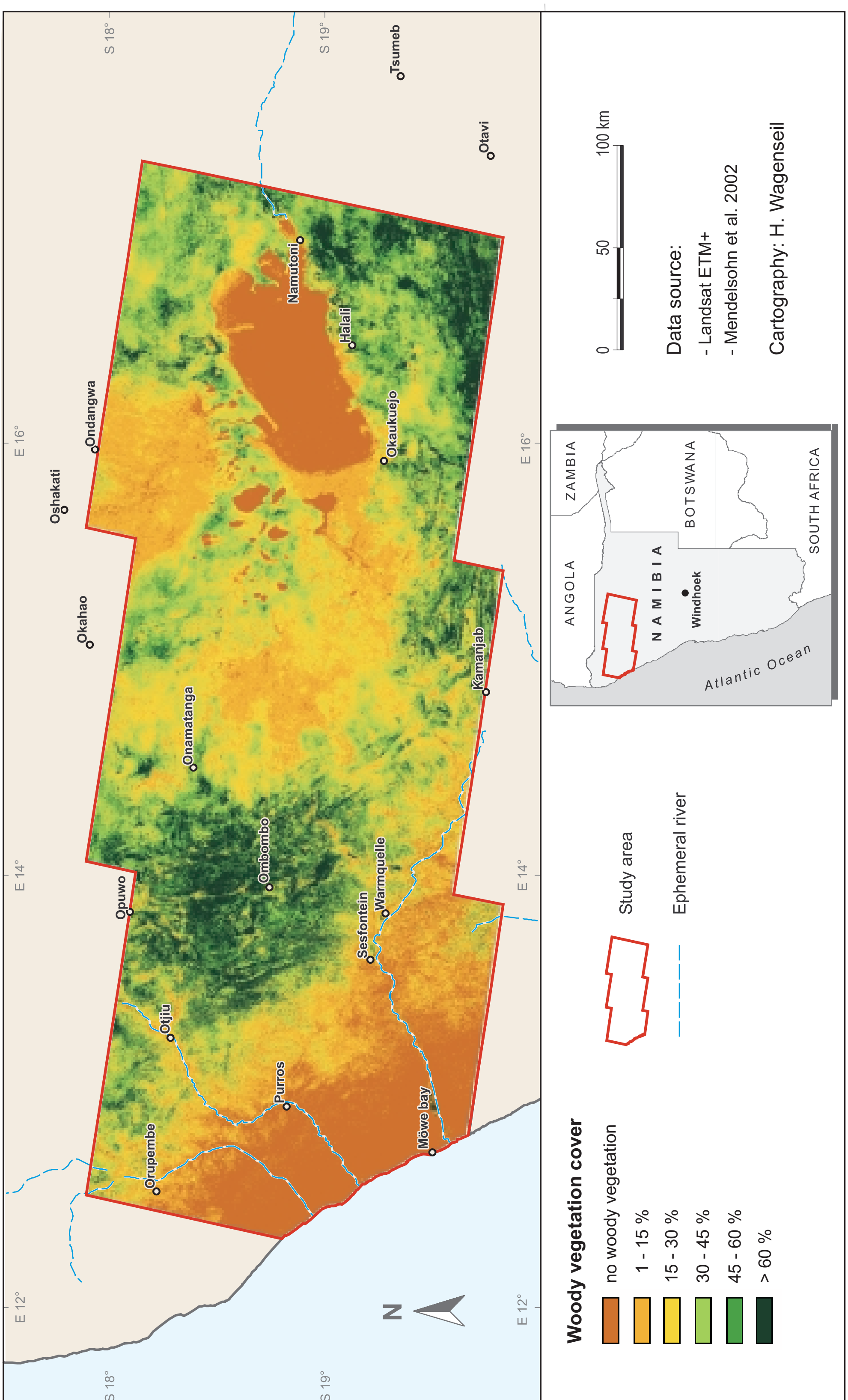

Fig. 3: Woody vegetation coverage estimated from Landsat ETM+
Deckungsgrad der Gehölze, abgeleitet aus Landsat ETM+

dence exists, that there are phenological differences between both life forms, grasses and woody plants, which are commonly attributed to different carbohydrate storage capacities (SCHOLES a. WALKER 1993, Fig. 4). Due to considerable reservoirs, trees and shrubs manage to achieve full leaf expansion already prior to the first summer rains, whereas grass storage capacity only supports the first green shoots. Additionally, woody plant foliage is maintained at least for several weeks after the last rains. Thus, notable differences in terms of growing season length can be expected between areas dominated by woody plants and those with grass prevalence.

A time series of NDVI derived from SPOT VGT image data was used to investigate, if these differences are also recognizable in multitemporal satellite image data and allow distinction of grassy and woody areas. Measurements from SPOT VGT are pre-processed geometrically and radiometrically prior to public data access. For this study, ten-day synthesis products (VGT S10, see VEGETATION PROGRAMME 2005) were used, consisting of ten-day maximum value composites (see HOLBEN 1986) of NDVI in a 1 x 1 km^2 spatial resolution. Temporal coverage comprises six years ranging from July 1998 to June 2004. However, the strong dependency of plant growth on water availability and the high spatio-temporal variation in actual rainfall within the study area (MENDELSOHN et al. 2002) produces significant NDVI intra- and inter-annual signal variation thus complicating data analysis of individual seasons (WAGENSEIL a. SAMIMI 2006). Hence, the six years of data were combined to produce NDVI data reflecting vegetation growth during an average season.

As anticipated, plant phenology captured by multitemporal NDVI varies significantly between sites with low and high woody coverage (Fig. 5). In case of few woody plants (Fig. 5a, Photo 2), the rainy season induces a short growing period mainly caused by greening of the herbaceous layer, followed by a long dry season with little green biomass and low NDVI respectively. Otherwise, the woody site (Fig. 5b, Photo 3) is characterized by expanded vegetation growth and correspondingly, a short dry period with higher NDVI values resulting from comparably high amounts of dry matter. In general, these findings agree with the observations made by SCHOLES and WALKER (1993) introduced above.

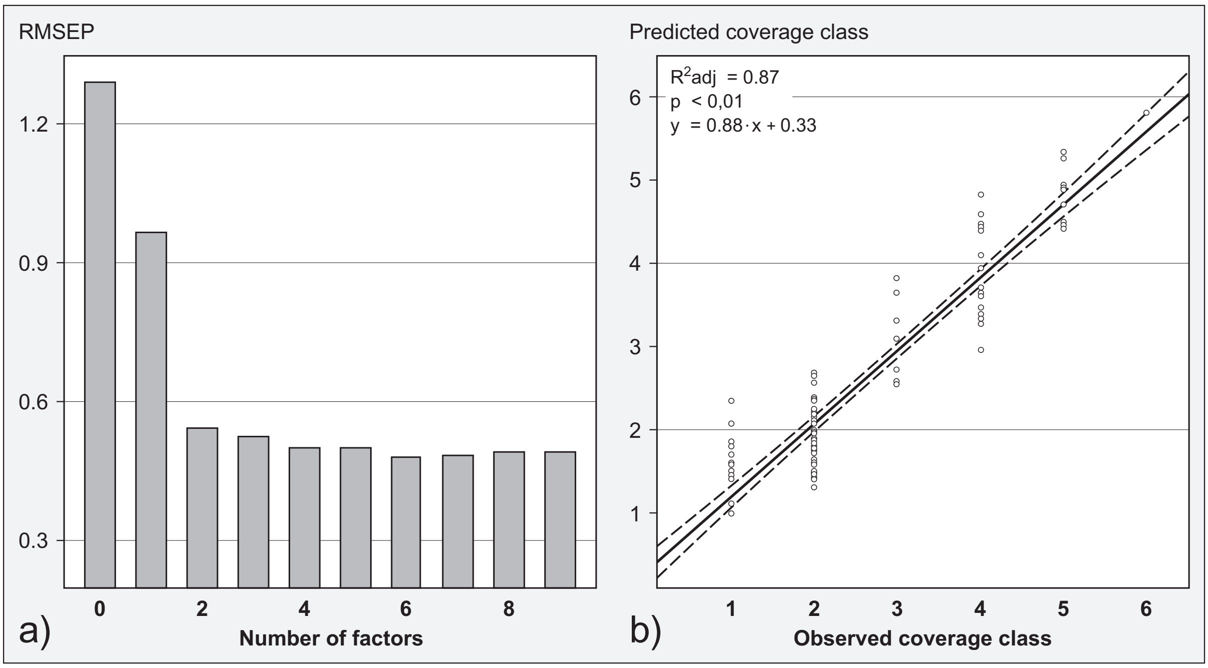

Fig. 2: a) Root-mean-square-error of prediction (RMSEP) in relation to the number of latent factors ($RMSEP_{Min}$ for n = 6); *b)* Woody coverage predicted by PLSR (n = 6) in relation to woody coverage observed in the field

a) Mittlerer quadratischer Vorhersagefehler in Abhängigkeit von der Anzahl latenter Faktoren ($RMSEP_{Min}$ bei n = 6); *b)* Vergleich zwischen modelliertem (n = 6) und beobachtetem Deckungsgrad der Gehölze

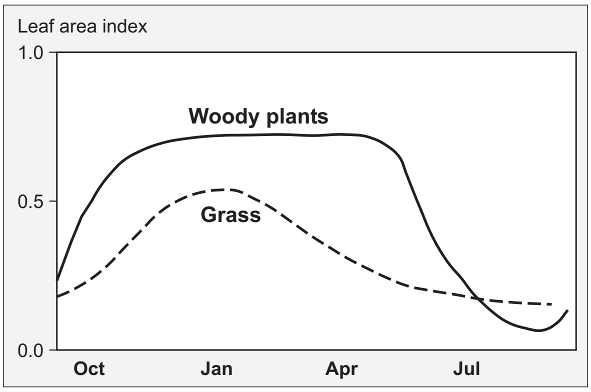

Fig. 4: Seasonal development of leaf area index for grasses and woody plants. Example from a savannah at Nylsvley/SA (Scholes a. Walker 1993, modified)

Jahresgang des Blattflächenindex für Gräser und Gehölze. Beispiel einer Savanne in Nylsvley/SA (Scholes a. Walker 1993, verändert)

In order to model NDVI curvature and to quantify these differences, a second order polynomial was fitted on a per-pixel basis (Fig. 5):

$$NDVI\ (t) = a + b * (t + c)^2 \quad (1)$$

With:

NDVI (t)	NDVI of a given pixel at time *t*
t	time of year
a, b, c	model parameters

Especially the polynomial parameters *a* and *b* are suitable to quantify the differences in NDVI phenology described previously. The principle is illustrated in figure 5. Whereas *a* represents the polynomial offset and is thus related to the seasonal minimum of NDVI, *b* describes the width of the parabolic function and is therefore considered to be an indicator of dry season length.

In order to examine usability of parabolic minimum and width for modelling woody coverage, both parameters were linearly related to the reference compiled from Landsat imagery. To adjust the different spatial resolutions of both datasets, the mean cover class for each 1 x 1 km^2 VGT pixel was initially computed from the 30 x 30 m^2 ETM map. Results are provided in figures 6a and 6b suggesting a close relationship (R^2 = 0.75) between minimum NDVI (parameter *a*) and woody coverage (Fig. 6a), whereas parabolic width (parameter *b*) shows greater scatter and correlation is less pronounced though acceptable with R^2 = 0.62 (Fig. 6b). However, both variables are considered to contribute significant information and were combined within a multilinear regression model to estimate woody coverage (R^2 = 0.76, Fig. 7).

4 Results and discussion

Due to the coarser spatial resolution, the phenological model output displays a more generalized result with less detailed structures. Nevertheless, spatial patterns of woody density from both approaches show a high correspondence. Woody vegetation is commonly

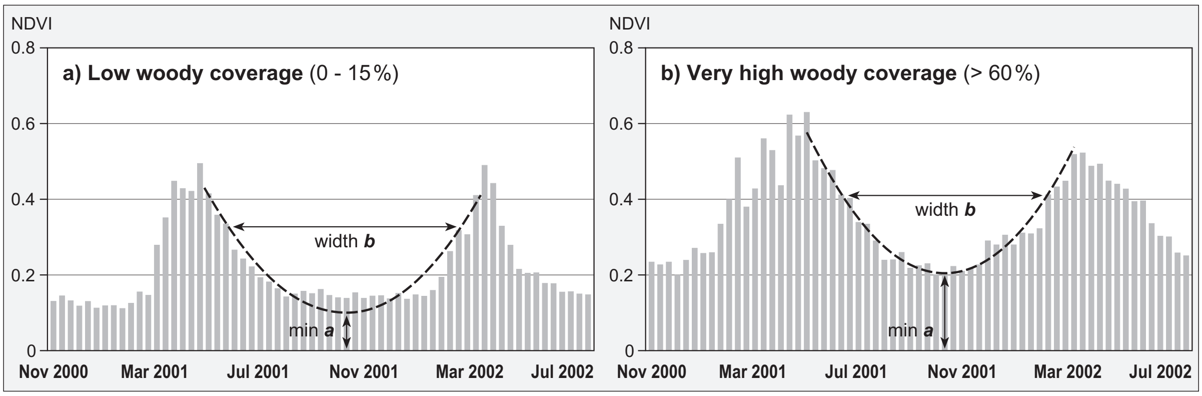

Fig. 5: NDVI trajectories (bars) for areas with low woody coverage (a) and high woody coverage (b) and corresponding polynomial curve fit

NDVI Profile (Säulen) für Gebiete mit niedriger (a) und hoher (b) Gehölzdichte mit den zugehörigen Polynomen

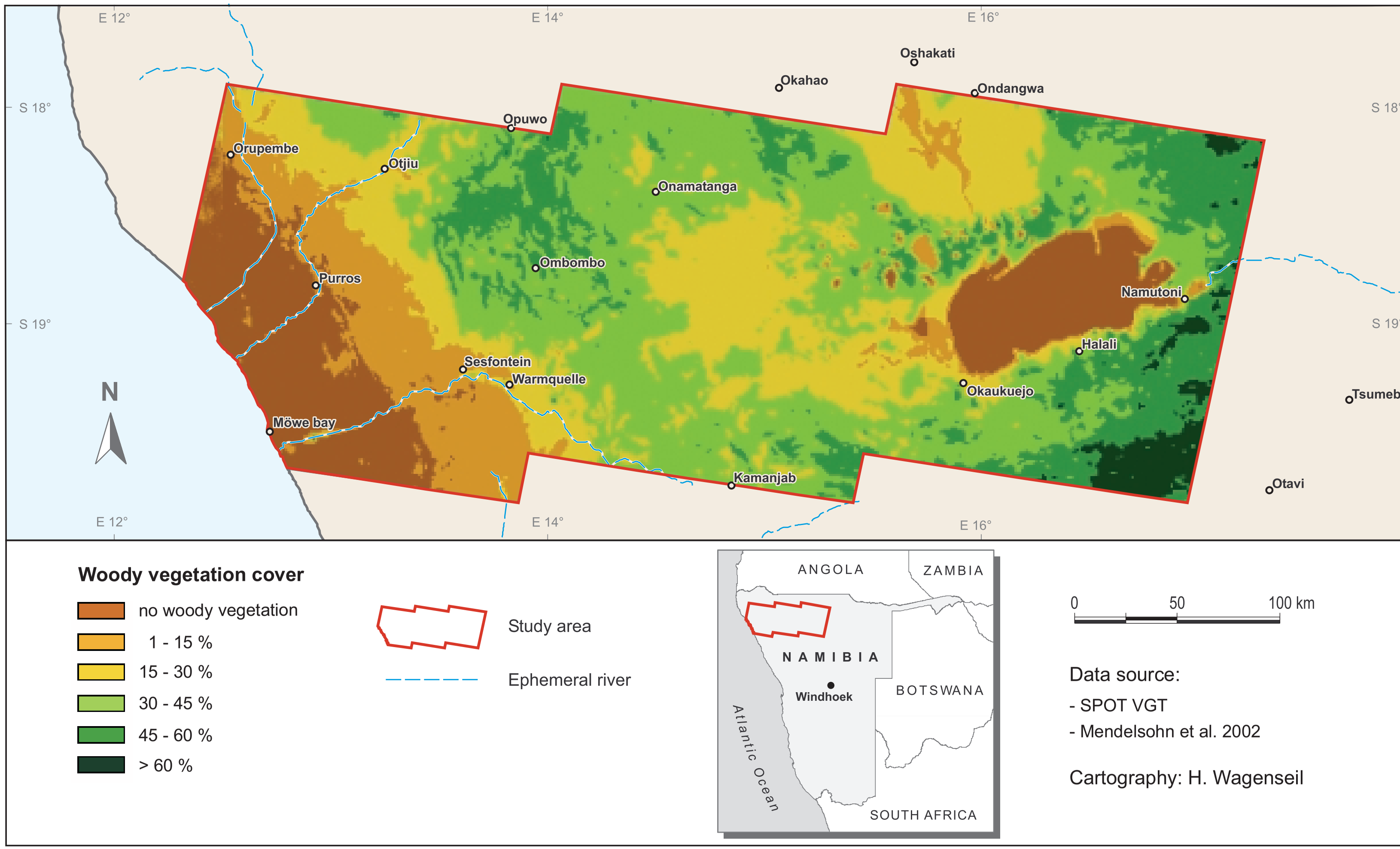

Fig. 7: Woody vegetation coverage estimated from multitemporal SPOT VGT NDVI
Deckungsgrad der Gehölze, abgeleitet aus SPOT VGT NDVI-Zeitreihen

absent in the coastal region made up by the Namib desert as well as inside the Etosha pan, a saline desert in the eastern part. Areas with low woody coverage (0–15% and 15–30%) form a transition zone in the west between the desert regions and the mountainous savannahs south of Opuwo, whose spatial dimension and percentage coverage are slightly underestimated by the phenological approach. Sparse woody vegetation also characterizes the central part of the image, whereas NDVI phenology suggests higher percentage coverage compared to the empirical findings. This observation also applies for the Cuvelai river system north of the Etosha pan. Areas of medium and high woody coverage are displayed in both maps for the northeast and southeast though again slight differences in absolute percentages occur.

Despite the good correspondence in terms of spatial patterns, deviations in estimated percentage cover require a short discussion about possible causes.

First, vegetation phenology and thus, NDVI derived metrics describing certain phenological features not only vary with vegetation composition but also with climatic conditions. As noted earlier, mean annual rainfall decreases within the study area from app. 450 mm in the east to 50 mm and less in the west. Similarly, rainy season length is shortened from four/five month to two and even less (Mendelsohn et al. 2002) meaning that the possible duration of the growing period is also reduced. Therefore, woody vegetation might be overestimated in the eastern part of the study area due to a comparably long rainy season, whereas coverage within the mountainous savannahs in the west is underestimated, resulting from an extended dry season. Additionally, a dense herbaceous layer appears to provoke overestimations in areas of low woody cover. This can be noticed in the central part of the region, which is made up by grass savannah with only little woody vegetation (Photo 2).

5 Conclusion and perspective

The approach presented in this paper clearly highlights that phenological differences between grasses and woody plants are portrayed in time series of NDVI and were successfully used for estimating spatial differences in woody coverage. Even if spatial resolution of multitemporal data sources is commonly low, their coverage allows application of these approaches on a regional scale. The principal findings of this study can be applied to archived NOAA AVHRR data for evaluation and quantitative analysis of temporal vegetation changes such as bush encroachment or deforestation in savannah ecosystems with deciduous woody plants.

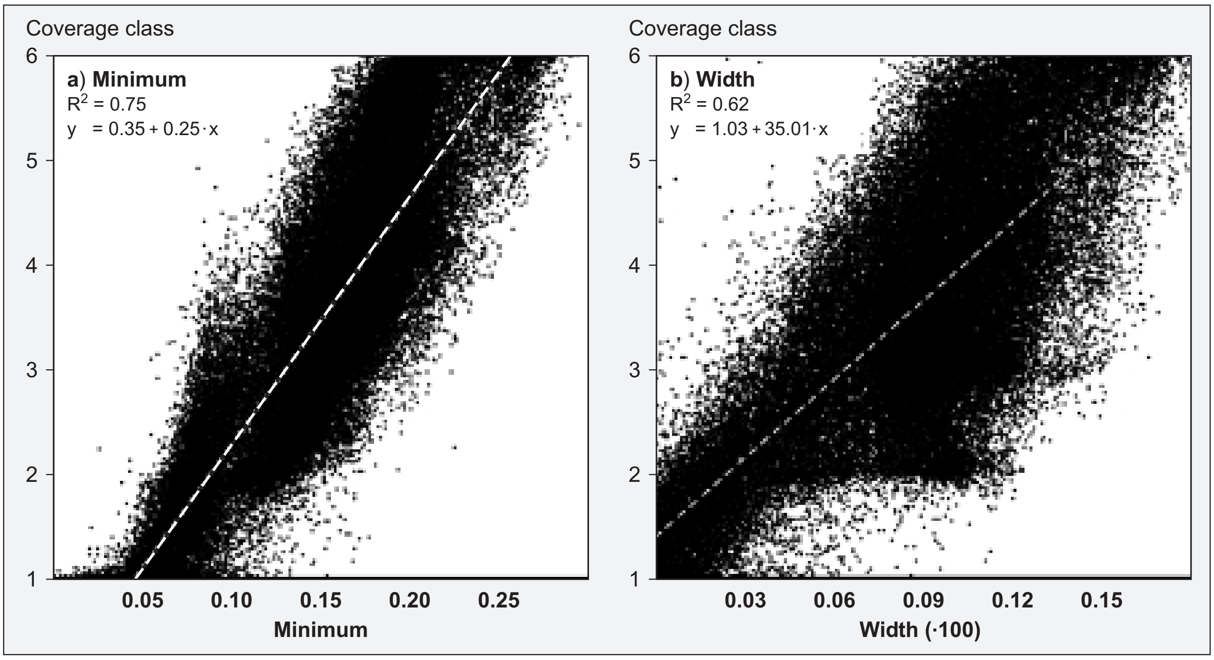

Fig. 6: Relationships between woody coverage and minimum (a) and width (b) of the parabolic curve fit

Zusammenhang zwischen Gehölzdichte und Parabelminimum (a) bzw. Parabelbreite (b)

Acknowledgements

This research project was funded by the Deutsche Forschungsgemeinschaft (DFG) under contract SA 775/5–1, the "Frau Dorothea und Dr. Dr. Richard Zantner-Busch-Stiftung" and the German Academic Exchange Service (DAAD). The authors thank Johan Le Roux, Ministry of Environment and Tourism in Namibia, for support during the field work and Dr. Henning Schröder for many useful comments on data evaluation.

References

ARMSTON, J.; DANAHER, T. a. COLLETT, L. (2004): A regression approach for mapping woody foliage projective cover in Queensland with Landsat data. In: Proceedings of the 12th Australasian Remote Sensing and Photogrammetry Conference. Freemantle, 1–13.

CARLSON, T. a. RIPLEY, D. (1997): On the relation between NDVI, fractional vegetation cover, and leaf area index. In: Remote Sensing of Environment 62, 241–252.

COHEN, W.; MAIERSPERGER, T.; GOWER, S. a. TURNER, D. (2003): An improved strategy for regression of biophysical variables and Landsat ETM+ data. In: Remote Sensing of Environment 84, 561–571.

DE KLERK, J. (2004): Bush encroachment in Namibia. Report on phase 1 of the bush encroachment research, monitoring and management project. Windhoek.

DIOUF, A. a. LAMBIN, E. (2001): Monitoring land-cover changes in semi-arid regions: Remote sensing data and field observations in the Ferlo, Senegal. In: Journal of Arid Environments 48, 129–148.

FURBY, S. a. CAMPBELL, N. (2001): Calibrating images from different dates to 'like-value' digital counts. In: Remote Sensing of Environment 77, 186–196.

HIGGINS, S.; SHACKLETON, C. a. ROBINSON, E. (1999): Changes in woody community structure and composition under constrasting landuse systems in a semi-arid savanna, South Africa. In: Journal of Biogeography 26, 619–627.

HIGGINS, S.; BOND, W. a. TROLLOPE, W. (2000): Fire, resprouting and variability: a recipe for grass – tree coexistence in savanna. In: Journal of Ecology 88, 213–229.

HOLBEN, B. (1986): Characteristics of maximum-value composite images from temporal AVHRR data. In: International Journal of Remote Sensing 7, 1417–1434.

HUANG, C.; YANG, L.; HOMER, C.; WYLIE, B.; VOGELMANN, J. a. DEFELICE, T. (2001): At-satellite reflectance: A first order normalization of Landsat ETM+ images. http://landcover.usgs.gov/pdf/huang2 (Accessed: 15.07.2003).

KRAUS, T. a. SAMIMI, C. (2002): Biomass estimation for land use management and fire management using Landsat-TM and -ETM+. In: Erdkunde 56, 130–143.

MATHER, P. (1999): Computer processing of remotely-sensed images. Chichester.

MENDELSOHN, J.; JARVIS, A.; ROBERTS, C. a. ROBERTSON, T. (2002): Atlas of Namibia. Cape Town.

RIAÑO, D.; CHUVIECO, E.; SALAS, J. a. AGUADO, I. (2003): Assessment of different topographic corrections in Landsat-TM data for mapping vegetation types. In: IEEE Transactions on Geoscience and Remote Sensing 41, 1056–1061.

ROQUES, K.; O'CONNOR, T. a. WATKINSON, A. (2001): Dynamics of shrub encroachment in an African savanna: relative influences of fire, herbivory, rainfall and density dependence. In: Journal of Applied Ecology 38, 268–280.

SANNIER, C.; TAYLOR, J. a. DU PLESSIS, W. (2002): Real-time monitoring of vegetation biomass with NOAA-AVHRR in Etosha National Park, Namibia, for fire risk assessment. In: International Journal of Remote Sensing 23, 71–89.

SCHOLES, R. a. WALKER, B. (1993): An African savanna: synthesis of the Nylsvley study. Cambridge.

SCHULTZ, J. (2000): Handbuch der Ökozonen. Stuttgart.

TOBIAS, R. (2006): An introduction to partial least squares regression. http://support.sas.com/rnd/app/papers/pls.pdf (Accessed: 15.03.2006).

VEGETATION PROGRAMME (2005): VEGETATION programme web page. http://www.spot-vegetation.com (Accessed: 17.01.2005).

WAGENSEIL, H. a. SAMIMI, C. (2006): Assessing spatio-temporal variations in plant phenology using Fourier analysis on NDVI time series: results from a dry savannah environment in Namibia. In: International Journal of Remote Sensing 27 (16), 3455–3471.