Research article

Terrain-based Landscape Structure Classification in Relation to Remote Sensing Products and Soil Data for the Okavango Catchment

Jan Wehberg, Michael Bock, Thomas Weinzierl, Olaf Conrad, Jürgen Böhner, Marion Stellmes & Lars Landschreiber

Abstract: The landscape of the Okavango Catchment is structured much more diverse than is perceptible at first sight. From the mountainous Angolan highlands with altitudes of more than 1700 meters above sea level to less than 1000 meters at the delta and the surrounding Kalahari sands; accordingly the relief intensity decreases steadily, from high mountainous to undulating in the center parts and flat in the south. Finding a way to describe, analyze and outline this diverse landscape structure and to classify landscape units on the basis of available data is just as challenging as useful. This paper shows how to derive geomorphographic units (GMUs) as discrete terrain entities on the basis of a SRTM digital elevation model and how to valorize and verify these units. The GMUs semi-automatically computed with SAGA GIS are based on a set of local and regional continuous land surface variables. As a first result we obtained 17 GMU-classes, sorted into four main groups: major valley floors (MVF), tributary valley floors (TVF), sandvelds and longitudinal dunes (SLD), and slopes and summit areas (SSA), scattered over the catchment. To test the applicability, we correlated the GMUs with two different vegetation indices (NDVI, EVI), a MODIS landform classification and soil parameters. In conclusion, single GMU-classes respond reasonable in a distinct and specific way to the land cover and reproduce the physiogeographic settings of the Okavango catchment quite well. On the basis of such units, further delineations respectively mappings of, for example, vegetation or soil data is seen as the next step..

Abbreviations: ANOVA = Analysis of Variance; DEM = Digital Elevation Model; EOSDIS = Earth Observing System Data and Information System; GMUs = Gemorphographic Units; ISRIC = International Soil Reference and Information Centre; MODIS = Moderate-Resolution Imaging Spectroradiometer; MVF = Major Valley Floors; LSUs = Landscape Units; SAGA = System for Automated Geoscientific Analyses; SLD = Sandvelds and Longitudinal Dunes; SOTERSaf = Soil and Terrain Database for Southern Africa; SP = Subproject within TFO; SRTM = Shuttle Radar Topography Mission; SSA = Slopes and Summit Areas; TFO = The Future Okavango; TVF = Tributary Valley Floors; WRB = World Reference Base.

Classificação da estrutura da paisagem com base no terreno em relação aos produtos de sensoriamento remoto e dados do solo para a represa do Okavango

Resumo: A paisagem da represa do Okavango está estruturada muito mais diversamente do que é perceptível à primeira vista. Dos montananhosos planaltos de Angola com altitudes de mais de 1.700 metros acima do nível do mar, para menos de 1.000 metros no delta e nas areias circundantes do Kalahari; consequentemente, a intensidade do relevo diminui de forma constante, de altas montanhas a ondulações nas partes centrais e planas no sul. Encontrar uma maneira de descrever, analisar e delinear essa estrutura diversificada da paisagem e classificar as unidades de paisagem, com base em dados disponíveis é tão desafiador quanto útil . Este trabalho mostra como derivar unidades geomorfográficas (GMUs) como discretas individualidades do terreno com base em um modelo digital de elevação SRTM e como valorizar e verificar essas unidades. As GMUs semi-automaticamente computadas com SAGA GIS são baseadas em um conjunto de variáveis de superfície contínua de solos locais e regionais. Como primeiro resultado, obtivemos 17 classes de GMU, ordenadas em quatro grupos principais: grandes vales (MVF), vales afluentes (TVF), savanas e dunas longitudinais (SLD) e encostas e picos (SSA), espalhadas pela bacia hidrográfica. Para testar a aplicabilidade, correlacionamos as GMUs com dois índices diferentes de vegetação (NDVI e EVI), uma classificação de relevo MODIS e parâmetros do solo. Em conclusão, classes únicas de GMU respondem razoávelmente de forma distinta e específica à cobertura do solo e reproduzem muito bem as configurações fisio-geográficas da bacia do Okavango. Com base em tais unidades, mais delimitações como os mapeamentos de dados sobre o solo ou a vegetação, por exemplo, são vistos como o próximo passo.

Introduction

Semi-automated classification of landscape structure is of great interest, first and foremost in areas with sparse data availability. Against this background methods of deriving landform units by reproducing the relief of the earth in a classification framework are an established focus in automated Digital Elevation Model-based terrain analysis (Macmillan & Shary 2009). Distinct

landform units are important because it can be assumed that similar processes and conditions are specific for them, e.g. water distribution, material transport, plant productivity. Landform units can be related to scientific findings from field campaigns or similar, and thus help in their extrapolation of the knowledge into other regions. The concept of automated geomorphographic mapping as used here refers to the conceptual approach of Krug et al (2005), which represents a highly specialized terrain classification system. Other commonly used methods target a broader spectrum of different applications (Iwahashi & Pike 2007, Drăguţ & Eisank 2012). The applied GMU concept follows a hierarchical classification approach, which is mainly based on relative terrain position and slope gradient.

In order to assess how meaningful and significant the derived GMUs are with respect to other environmental variables, the GMUs have been correlated with MODIS based vegetation indices. Furthermore it is intended to get perception about plant productivity, allotment of landscape classes, and soil distribution in the catchment. The results serve as starting point for further investigations like soil mapping or for an adjustment of remote sensing products, like land cover classification. Two guiding questions are raised in this paper: First, is the delineation of the GMUs justifiable, meaningful and significant? Second, how is the catchment described by the GMUs with reference to landform classification, vegetation indices, and soil data?

The concept was chosen because its development precisely aims at the application that is among others required in the context of the TFO-project: identification of homogenous landform units to support other subprojects related to soil and vegetation science. Individual adaptation to the special landscape, however, is possible and even necessary to reflect the soil-forming processes as well as possible.

Methods and Material

In the following section the used data and methods are discussed. Some data is freely available whereas some has been collected or processed by other subprojects of the TFO project (SP03, SP09.1).
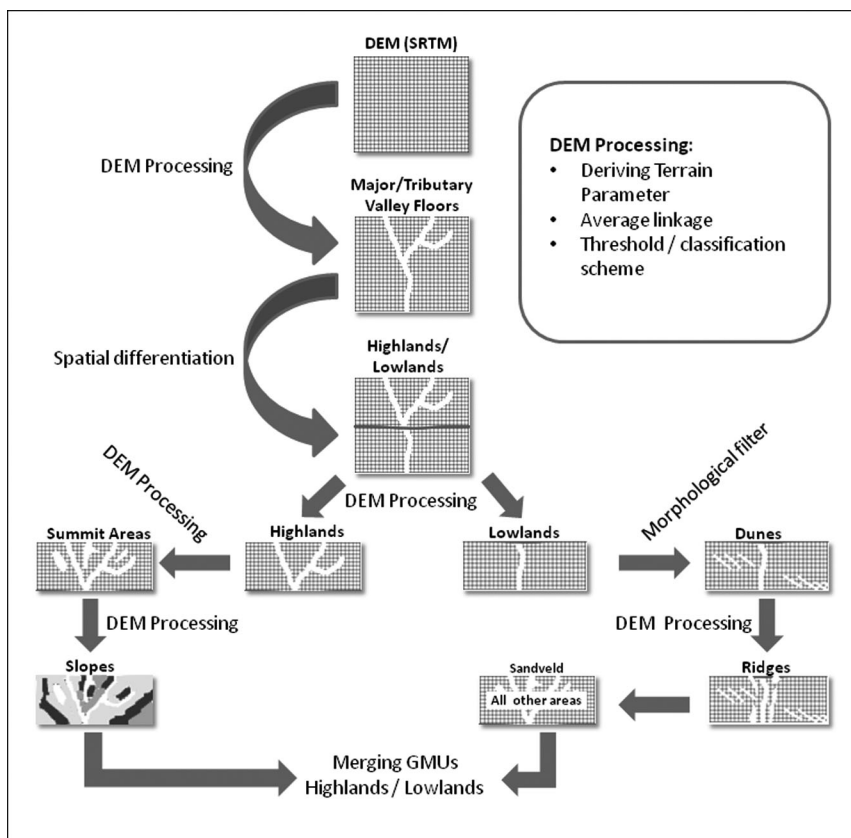

Fig. 1: Workflow for the GMU derivation.

Derivation of Geomorphographic Units

The derivation of geomorphographic units (GMUs) is based on a set of land surface variables, which are derived from a Digital Elevation Model (DEM, Figure 1). The elevation model used in this study is the established SRTM 3-arc-second DEM (Jarvis et al. 2008), which has a horizontal resolution of 90 by 90 meters. It was filtered with an extended version of the multidirectional Lee-Filter (Lee 1980) in order to reduce artificial noise and preserve important landscape elements at the same time. At first the filter creates bands of cells in all directions, e.g. with a length of 9 and a width of 3 cells, around every grid cell. Then the standard deviation (of elevation) is calculated for every band (direction). The new value (elevation without noise) for the grid cell is calculated from the mean elevation of the band with the lowest standard deviation and the original elevation of the cell (Köthe & Bock 2009). All steps of DEM analysis have been conducted with SAGA GIS. The delineation of discrete terrain units is based on a set of local and regional continuous land surface variables: SAGA wetness index, altitude above channel

network, slope, relative slope position, terrain classification index for lowlands (TCI$_{low}$, Bock et al. 2007).

The terrain classification index for lowlands has recently been introduced by Bock et al. (2007) and plays a central role in the GMU derivation. It constitutes a combination of altitude above channel lines (AACL, inverted, log) and *SAGA wetness index* (SWI); both normalized (range 0 to 1) and integrated by:

$$TCI_{low} = \frac{2 \times AACL + SWI}{3}$$

The average linkage grouping which has been used for the GMU identification is an agglomerative algorithm in a connectivity based hierarchical clustering to combine neighboring cells into groups iteratively until a stop criterion is reached (Ester & Sander 2000). It starts with the single grid cells. In the first step, groups consist of maximal two grid cells ("groups" with only one cell may find a partner at a later stage). Thereafter groups are growing by joining with similar neighboring groups. The result is a grid with numerous small subareas which comprise cells with similar values. These will be arranged according to expert-

based threshold values, e.g. the bottom areas are defined by threshold values of the TCI_{low}. All other areas are defined corresponding to the steps mentioned above.

The Okavango Catchment consists of two major landscapes, a mountainous area in the north and flat plain areas in the south. To improve the performance of the terrain classification, both landscape types have been analyzed separately. The division was conducted semi-automatically by using a threshold of parameter surface roughness, measured on a coarse scale of 20km. The final course of the line was drawn manually. For better integration of subsequently derived terrain units, the subdivisions share an overlapping zone of 30km.

The classification resulted in the following units. (i) Four types of bottom areas (both lowlands and highlands, TCI_{low} < 0.29): class boundaries of TCI_{low} = 0.0, 0.089, 0.1, 0.185, 0.29, major or tributary valley (the lowest class of valley floors refers to the actual Okavango river bed). (ii) Ridges in sand fields (lowlands): threshold *SAGA wetness index* = 17. (iii) Dunes (Lowlands): identified in manually separated subareas by using a morphological filter which kept the minimum value in a radius of eight grid cells. This value was subtracted from the DEM. Areas with boundary values > 3 m were defined as dunes. The result was subsequently edited manually. (iv) Six types slope areas (Highlands): class boundaries refer the SOTER manual of procedure (van Engelen & Wen 1995, see Table 1). (v) Summit areas (Highlands): class boundary of 0.9 of relative slope position. (vi) Using generalization techniques, areas smaller than five raster cells (sink areas are excluded) were merged with neighboring areas.

With lengths of up to 100 km, longitudinal dune ridges are prominent features in the Northern Kalahari. Typical widths are about 2km and the crest-to-crest spacing between neighboring dunes ranges from 1 to 2.5km (Thomas et al. 2000). Except for some channel incisions the linear dunes are the only undulation in relief. The dune delineation is again based on relative height. A terrain base level is identified using a morphological erosion filter with a search radius in the dimension of the crest-to-crest distance and a relative height is calculated as difference to the original elevation. Contour lines derived from the relative height for several thresholds between 2

Table 1: Slope classes according to the SOTER manual of procedure (van Engelen & Wen 1995).

	Slope classes	Degree	Percent
F	flat	0 - 1.1	0 - 2
G	gently undulating	1.1 - 2.9	2 - 5
U	undulating	2.9 - 4.6	5 - 8
R	rolling	4.6 - 8.6	8 - 15
S	moderate steep	8.6 - 17.4	15 - 30
T	steep	above 17.4	above 30

and 5m have been visually examined regarding their suitability to delineate dunes. While 2m contours included many irrelevant features and often did not separate neighboring dunes, 5m contours appeared to be too restrictive. A best overall match has been found for a 4m threshold, which subsequently has been used for dune definition.

MODIS-Products

The MODIS vegetation indices products provide consistent information of the spatial and temporal behavior of vegetation conditions (Huete et al. 2002). We used the MOD13Q1 product which is distributed via the Earth Observing System Data and Information System (EOSDIS) of the National Aeronautics and Space Administration (NASA). This product includes two vegetation indices, which are (i) the Normalized Difference Vegetation Index (NDVI) and (ii) the Enhanced Vegetation Index (EVI). The indices are related to vegetation density and cover (Huete et al. 2002). The product consists of 16-day composites with a spatial resolution of 250 m and is available since April 2000.

Vegetation indices

In the present study we used both the NDVI and the EVI which are supposed to complement each other (Huete et al. 2002). The NDVI is an indicator for the assessment of productivity of photosynthetical active green vegetation.

The index is ratio-based and captures the spectral contrast between the red and near-infrared (NIR) reflectance (ρ) of the vegetation signal. It is defined as:

$$NDVI = \frac{\rho NIR - \rho RED}{\rho NIR + \rho RED}$$

ρRED :   reflectance red
ρNIR :   reflectance NIR

The values range from -1 and to 1 where water bodies attain values below 0, whereas most land surfaces are characterized by values between 0 and 1 with increasing values for increasing presence of photosynthetical active vegetation. Drawbacks of the NDVI are its saturation with high biomass values as well as the strong influence of the background signal on the NDVI on areas with low vegetation cover (Elvidge 1985).

The EVI was designed to improve the vegetation signal with enhanced sensitivity in high biomass regions by de-coupling the canopy background signal and minimizing atmospheric effects. Thus, an advantage of the EVI compared to the NDVI in low productivity areas is the lower interference with background signals generated by bare soil (Huete et al. 2002). The EVI responses sensitive to canopy characteristics such as leaf area index, canopy type, plant physiognomy and canopy structure (Gao et al. 2000; Huete et al. 2002). The EVI is defined as:

$$EVI = G \times \frac{(\rho NIR - \rho RED)}{((\rho NIR + C_1 \times \rho RED - C_2 \times \rho BLUE + L))}$$

ρRED: reflectance red
ρNIR: reflectance NIR
$\rho BLUE$:   reflectance blue
G: gain factor = 2.5
L: canopy background adjustment that addresses non-linear, differential NIR and red radiant transfer through a canopy; L =1
$C1, C2$:   coefficients of the aerosol resistance term, which uses the blue band to correct for aerosol influences in the red band; C1 =6, C2 = 7.5

For the analysis of the NDVI we used a composite covering a period at the end of March 2013 and for EVI the 16-day-composite covering a comparable period of April 2011. Time frames at the end of
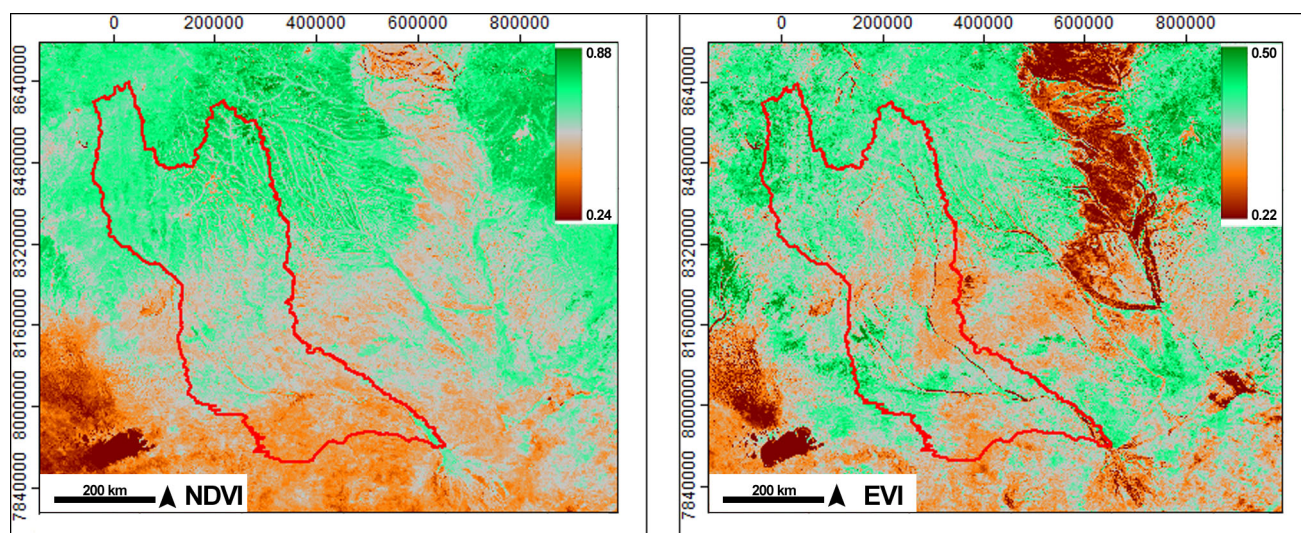
Fig. 2: MODIS-NDVI and -EVI in the Okavango catchment area. Note the high values in the mountainous regions and the low ones in the Kalahari, especially in the Etosha-Pan in the south west. Data Source: 16-day MODIS vegetation index data (MOD13Q1) provided by EOSDIS/NASA. The red polygon indicates the area covered by GMUs. Projection/Datum: UTM, WGS 1984, Zone 34S.

the rainy season had been considered because of a higher biomass at this time. Since it was not envisaged to promote a direct comparison of these vegetation indices, for the purpose of this paper it should be neglected that the scenes are from different years. Mean values for NDVI and EVI are extracted directly from SAGA GIS (Fig. 2).

MODIS Land Cover Classification
The GMUs were evaluated and analyzed by intersecting them with land cover information that was based on a MODIS EVI time series. The high temporal resolution of the MOD13Q1 enables the derivation of information describing the development of vegetation cover within the vegetation period. Phenology parameters were calculated for the period 2001 to 2012 using the software TIMESAT (Jönsson and Eklundh 2002). These parameters include for example start and end of the vegetation period as well as parameters that can be related to standing biomass and magnitude of seasonality.

Because different vegetation communities differ in their phenological characteristics, the aforementioned parameters can be employed to discriminate major land cover types. Based on the phenological information, major land cover strata within the Okavango Catchment were statistically distinguished by employing an unsupervised clustering algorithm (ISODATA) where pixels that are characterized by similar annual EVI

course will be summarized in one class. The nomenclature of the resulting 20 classes was developed based on extensive field work in the Okavango region and aims at identifying the dominant vegetation type that each class represents. This fact, as well as the spatial resolution of MODIS and the limited number of 20 classes, entail the potential shortcoming that the actual vegetation cover might locally differ from the nomenclature in the map (Stellmes et al. 2013).

Soil Data

For the comparison of soil data the WRB Reference Groups of the digital soil map by SOTERSaf at a 1:2 million scale were used (ISRIC 2001-2003). The compilation of the database was conducted according to the SOTER methodology, using existing information. Point source data were collected and provided by Subproject 03 of the TFO project at the Institute of Soil Science, University of Hamburg. 106 profiles of
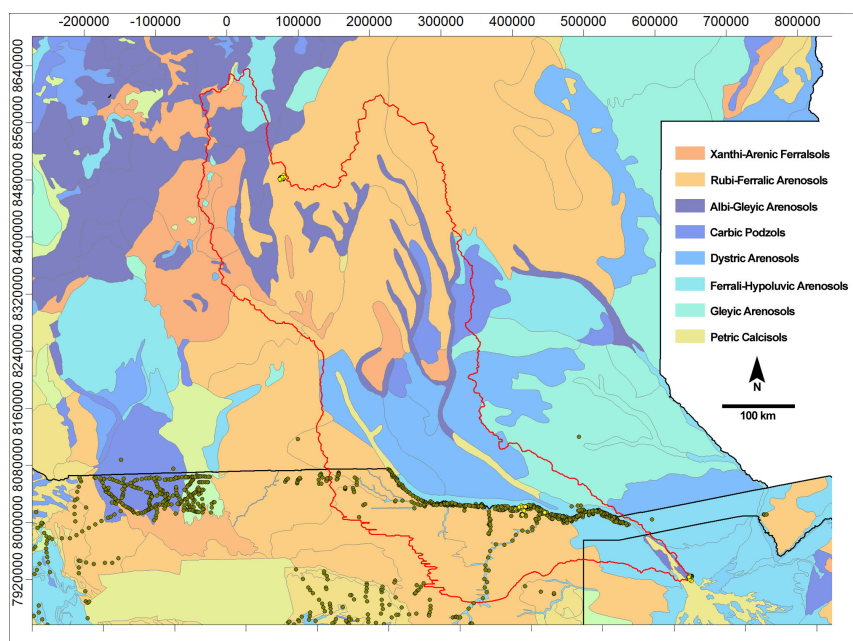

Fig. 3: WRB reference groups of the SOTERSaf data base and the profiles of SP03 (yellow dots) respectively the SOTERSaf data base (olive).

Table 2: Survey of the GMUs with attributes.

		GMU.Name	No. of Shapes	No. of Cells	Total Area [km²]	Mean Area [km2]	Slope Mean [°]
1		Major Valley Floor Level 1	36	207.148	1.858,65	51,63	0,48
2	MVF	Major Valley Floor Level 2	1.762	20.547	184,36	0,10	0,56
3		Major Valley Floor Level 3	1.151	158.328	1.420,61	1,23	0,68
4		Major Valley Floor Level 4	180	331.097	2.970,79	16,50	0,79
5		Tributary Valley Floor Level 1	237	3.602	32,32	0,14	0,69
6	TVF	Tributary Valley Floor Level 2	672	418.557	3.755,53	5,59	0,74
7		Tributary Valley Floor Level 3	16.036	316.796	2.842,47	0,18	1,42
8		Sandveld	3.857	6.965.897	62.501,93	16,2	0,43
9	SLD	Sandveld. other Ridges	3.437	1.812.633	16.263,96	4,73	1,35
10		Longitudinal Dunes	715	385.817	3.461,77	4,84	0,93
11		Slope Class F	31.269	2.464.321	22.111,27	0,71	0,66
12		Slope Class G	51.644	4.741.620	42.544,47	0,82	1,92
13		Slope Class U	54.165	1.646.464	14.773,00	0,27	3,51
14	SSA	Slope Class R	13.382	418.082	3.751,27	0,28	5,50l
15		Slope Class S	1.034	21.341	191,48	0,19	9,99
16		Slope Class T	19	147	1,32	0,07	18,15
17		Summit Areas	5.097	1.750.303	15.704,70	3,08	1,37
		Sum.Total	184.693	21.662.700	194.369,87	1,05	

the data base are classified according to WRB (for locations of soil samples see Fig. 3).

Statistics

For analysis of congruence the spatial extents of the GMUs were compared to the shared surface area respectively number of samples of the used data bases. The uneven distribution of samples within the catchment is a result of the very limited accessibility of the northern part, in part due to land mines. The SAGA module "Zonal Grid Statistics" was applied to create a contingency table of unique condition units (UCUs). These units are delineated from a zonal grid (i.e. GMUs) and optional categorical grids, e.g. landcover and soil. For the analysis

of the non-discrete respectively unclassified data sets, i.e. NDVI and EVI an ANOVA and a Tukey's Honestly Significant Difference Test (Abdi & Williams 2010) was done for all 184,693 subunits with SPSS (IBM Corp. 2012). All GIS work was done with SAGA (www.saga-gis.org) and ArcMap 10 (ESRI 2011).

Results

The delineation and analysis of the GMUs should be seen as the main result of this paper. Even though the results of the MODIS products and soil analysis contribute interesting facts, they are considered more as a tool for the actual purpose.

Geomorphographic Units (GMUs)

In total the generation of the GMUs resulted in 184.693 subunits (shapes) for 17 different GMU-classes in four main groups, covering an area of 194.370 square kilometers and an average area of subunits of slightly more than one square kilometer (Table 2). Regarding total area the sandveld is the largest unit whereas the steepest slope class (T) is the smallest (Fig. 4). An example for the GMUs in one of the core sites can be seen in Figure 5.

MODIS land cover classification

The MODIS landform classification (conducted by TFO-Subproject 09.1, Schneibel et al. 2013) shows a clear
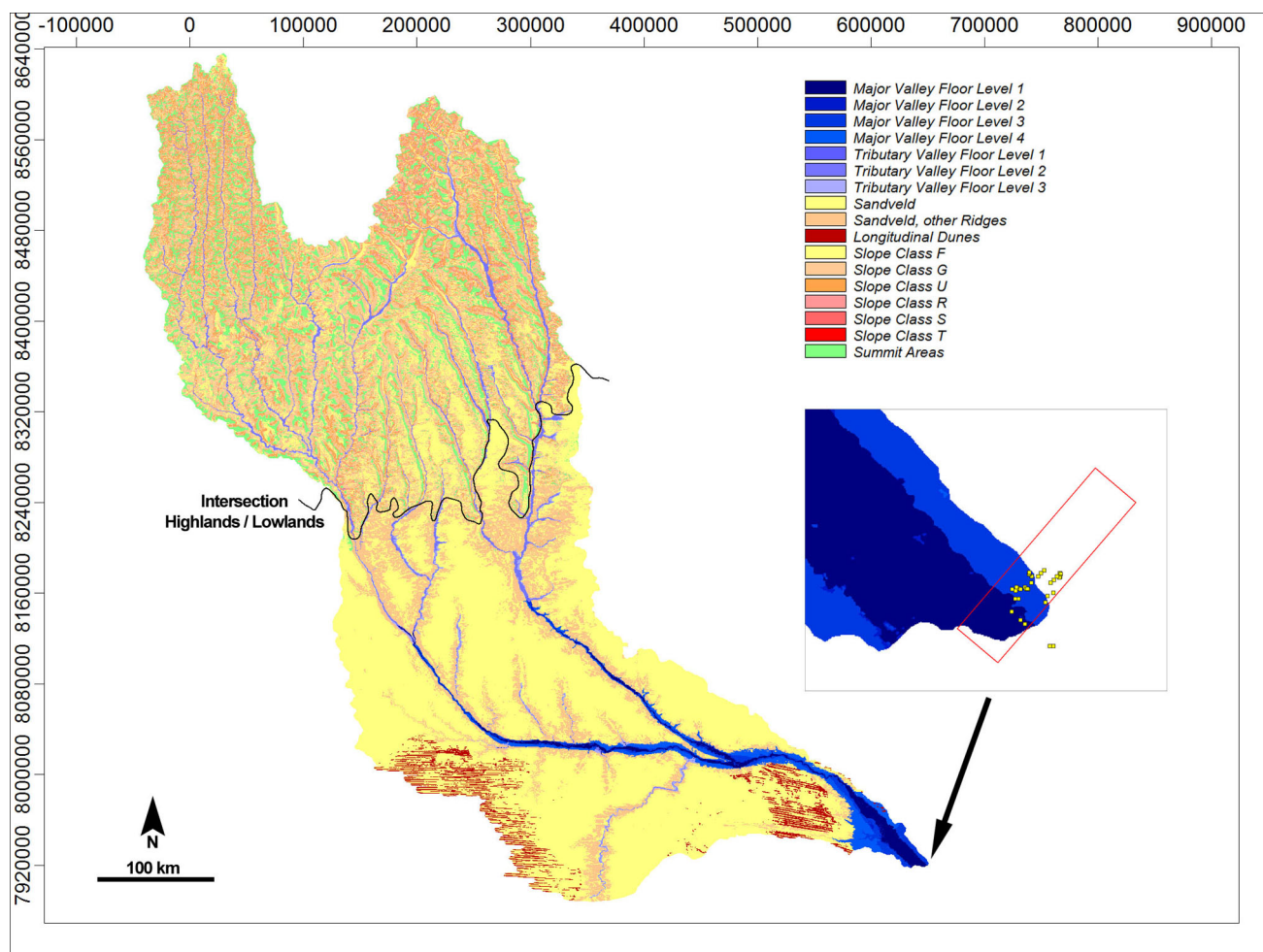
Fig. 4: Map of the geomorphographic units within the Okavango catchment. In the south, i.e. in the delta area, the delineation of the GMU is limited by the low relief intensity (small map, core site Seronga and profiles of SP03).

distribution over the GMU (Fig. 6). As expected, flooded areas are concentrated in the major valley floors and the tributaries. Settlements and sparse shrubland, fields or urban areas is also only to be found in the major valley floors. Regarding forest types there is a clear distinction, e.g. between the Miombo forest types at slopes and summit areas, the mixed Burkea Woodlands (with *Terminalia sericea*) at upper valley floors, tributary valley floors level 1 and SLD, and mixed woodlands (including *Colophospermum mopane*) only occuring at upper MVF. The thornbush savannah is concentrated at valley floors and SLD while wet grasslands and peatlands are abundant in tributary valley floors level 2 and 3. Shrublands are, apart from longitudinal dunes, especially found at
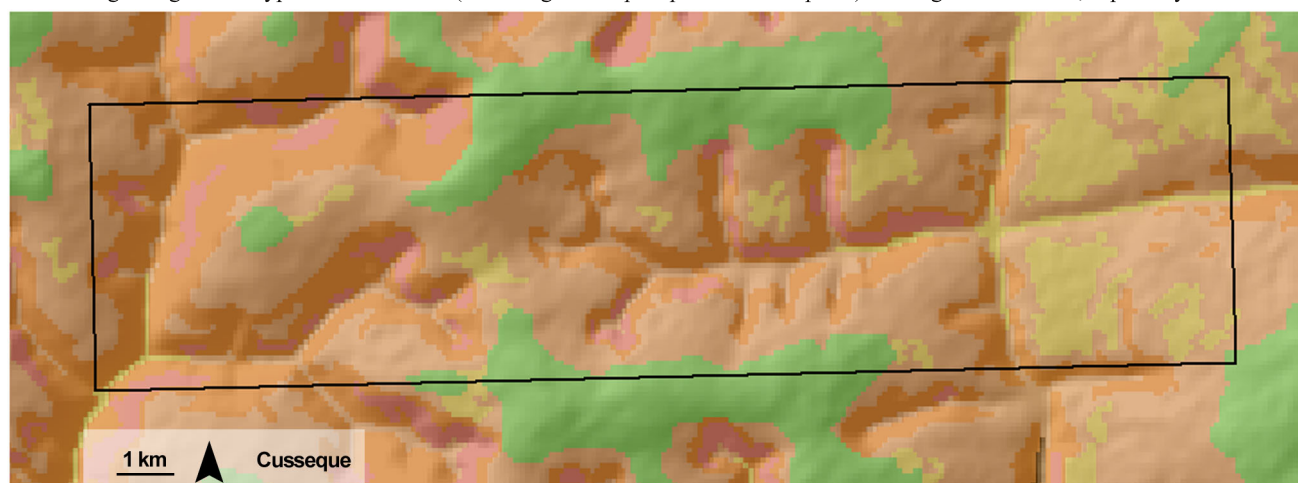

Fig. 5: GMUs in the core site Cusseque. Summit areas are green, a piece of tributary valley floor is visible in the southeast (brown). All other classes are slopes with different inclinations.
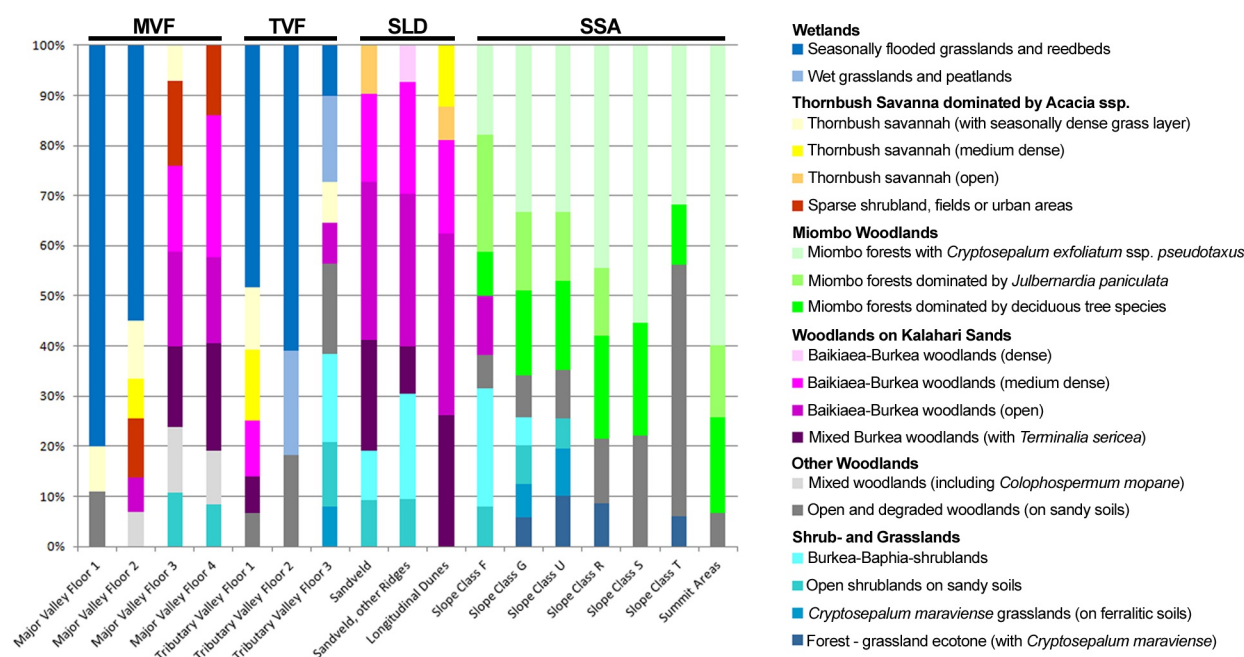
Fig. 6: Distribution of MODIS landforms for each GMU. Only classes with a share of more than five percent per GMU.

SLD, TVF3 and lower SSA. Woodlands on Kalahari sands (Baikiaea-Burkea woodlands) have their main area of distribution at SLD. The class *Parinari capensis* grasslands (on humid sands) (see Stellmes et al. 2013) was sorted out because it never reaches a share of more than five percent in the GMUs. A sample application for the GMUs compared to the MODIS land cover classification can be seen for the core site Cusseque in Figure 7.

GMUs versus MODIS Vegetation Indices NDVI and EVI

Overall, the Normalized Differenced Vegetation Index NDVI, has its highest values in the northern parts of the catchment and in the panhandle (Fig. 2).

Regarding the GMUs the highest values for the NDVI occur in the summit areas and slope classes G to T (Table 3). The middle section comprises all tributary valley floors, flat slopes, major valley floors 1, and sandveld/other ridges. Major valley floors 2, 3 and 4, sandveld and longitudinal dunes show the lowest NDVI values, which means a difference to the EVI. Nevertheless, both vegetation indices classify the slopes and summits as the most productive areas.

Table 4 shows the significance levels of a Tukey's Honestly Significant Difference Test. GMU 1 and 16 have a great share of non-significant variances (see discussion). The tributary valley floor level 3, the slopes (except 16 = slope class T) and the summits are significant in most of the cases. Slopes

and Summit Areas (SSA) differ within the main class and compared to the other main classes regarding NDVI mean values. In general, the highest values for the EVI are found in the northern parts of the catchment and in the vicinity of the main river bed (Fig. 2). At the GMU-level the highest mean values for the EVI are found at the summit areas and steeper slope classes, except slope class T (Table 4). Longitudinal dunes, sandvelds, flat slopes and the upper valley floors (level 3 each and major valley floor 4) are located in the middle section of the values. Valley floor level 1 and 2 (both tributary and major) show the lowest EVI values (Table 4). The Table indicates that most of the variances of the GMUs are distinct from each other. Again, GMU 16, the steepest slope class (T) has the lowest
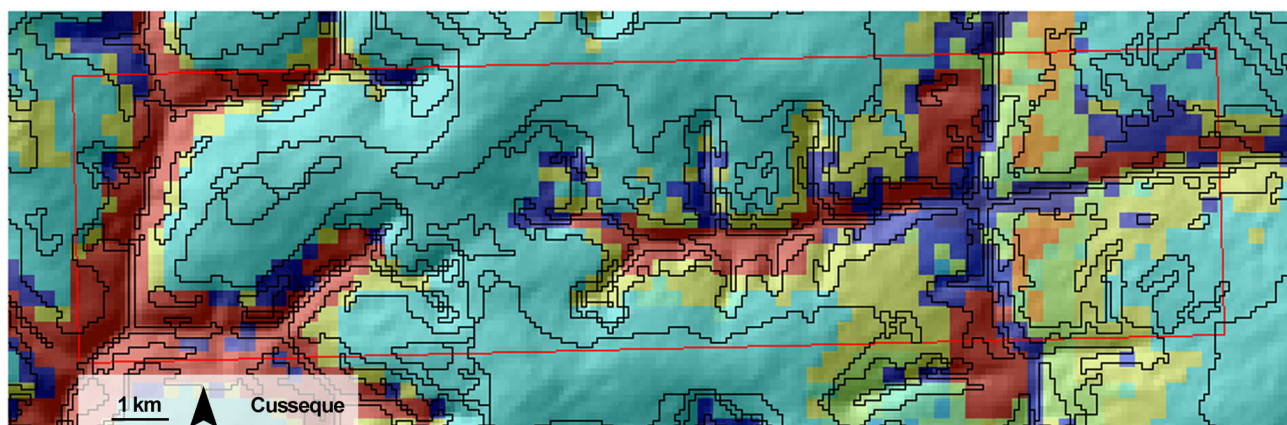

Fig. 7: MODIS land cover classification in the Core Site Cusseque. The vegetation cover consists predominantely of dense and medium dense Miombo forest (cyan) and grassland on ferraltic soils (red). The GMU delineation is shown black outlined.

Table 3: Mean values for vegetation indices NDVI and EVI in the GMUs.

	GMU / Name	Mean NDVI	Mean EVI
MVF	Major Valley Floor Level 1	0.56	0.24
	Major Valley Floor Level 2	0.51	0.26
	Major Valley Floor Level 3	0.47	0.35
	Major Valley Floor Level 4	0.47	0.37
TVF	Tributary Valley Floor Level 1	0.54	0.27
	Tributary Valley Floor Level 2	0.61	0.32
	Tributary Valley Floor Level 3	0.57	0.35
SLD	Sandveld	0.50	0.36
	Sandveld, other ridges	0.54	0.36
	Longitudinal Dunes	0.48	0.38
SSA	Slope Class F	0.60	0.36
	Slope Class G	0.66	0.39
	Slope Class U	0.67	0.40
	Slope Class R	0.68	0.40
	Slope Class S	0.70	0.39
	Slope Class T	0.71	0.38
	Summit Areas	0.71	0.40

share of significant differences. This is likely due to the very small number of shapes within the GMU.

GMU versus Soil Type/WRB Reference Groups

The WRB reference groups of the Harmonized World Soil Database (FAO/IIASA/ISRIC/ISSCAS/JRC 2012) are likewise suitable for an analysis in the scope of this paper, a comparison with the GMUs. Again, a clear distinction between the GMUs is visible (Fig. 8). Fluvisols are restricted to valley floors (one exception is the GMU Sandveld/other ridges). Perennial swamps are only to be found, as one can expect, in the major valley floor 1. Arenosols are more or less evenly distributed over all GMU with a higher percentage of Rubi-Ferralic Arenosols in the sandvelds, upper tributary valley floors, slopes and summits. Ferralsols are abundant in the upper valley floors (level 2 and 3), the slopes (especially class T) and the summit areas.

For further analyses of soils, 631 soil profiles of the data base of the Ministry of Environment and Tourism (unpublished) and 106 profiles of SP03 with a

Table 4: Analysis of variance of NDVI mean values (above right half) and EVI mean values (bottom left half) for every single subunit (Tukey's Test, Honestly Significant Difference): Significance level α (confidence interval 95%) of multiple comparisons. Significant values are displayed in bold.

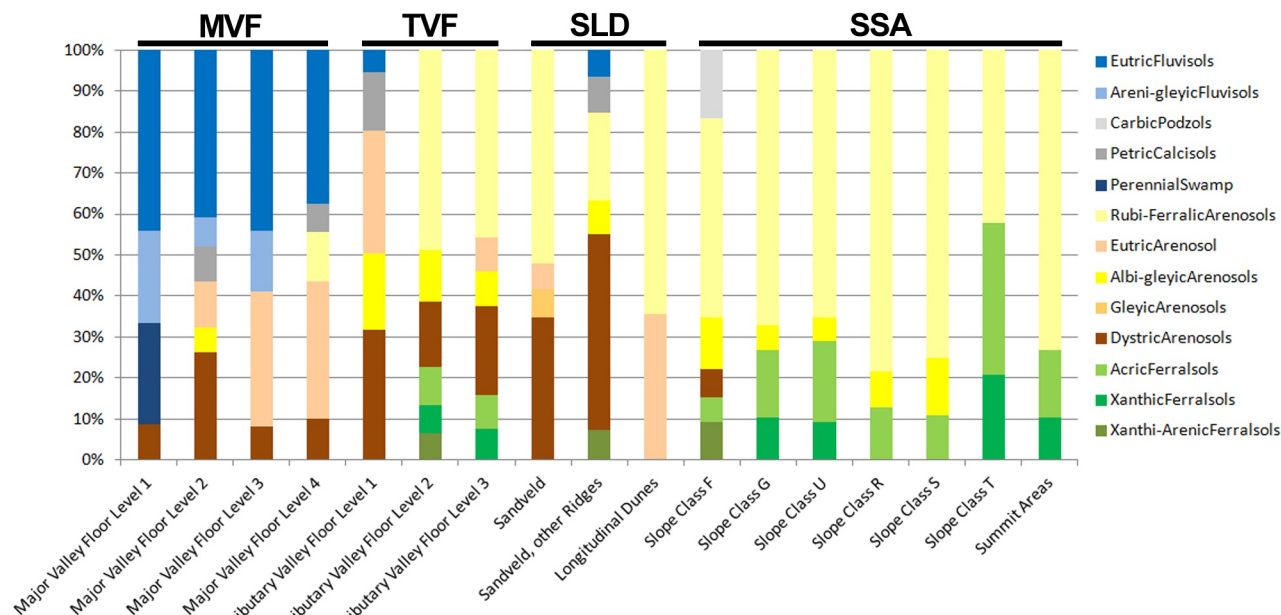|	NDVI →	MVF				TVF			SLD			SSA						
EVI ↓	GMU	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
MVF	1		0.974	0.986	0.922	1	1	0.998	1	1	0.617	0.592	0.268	0.075	**0.039**	**0.003**	0.54	**0.008**
	2	**0.028**		1	1	0.998	**0.11**	0	0.352	0.167	0.096	0	0	0	0	0	**0.01**	0
	3	0	0		1	1	0.183	0	0.535	0.276	**0.006**	0	0	0	0	0	**0.013**	0
	4	0	0	0		0.976	0.089	0	0.377	0.243	0.994	0	0	0	0	0	**0.006**	0
TVF	5	0.112	1	0	0		1	**0.013**	1	1	0.434	0	0	0	0	0	0.09	0
	6	0	0	0	0.997	0		0	0.952	0.986	0	0	0	0	0	0	0.099	0
	7	0	0	0	0	0	0		0	0	0	0	0	0	0	0	0.796	0
SLD	8	0	0	0	0	0	0	0.682		1	0	0	0	0	0	0	**0.035**	0
	9	0	0	0	0	0	0	0	0		0	0	0	0	0	0	**0.039**	0
	10	0	0	0	0	0	0	0	0	0		0	0	0	0	0	**0.001**	0
SSA	11	0	0	0	0	0	0	0	0	0	0.983		0	0	0	0	0.994	0
	12	0	0	0	0	0	0	0	0	0	1	0		0	0	0	1	0
	13	0	0	0	0	0	0	0	0	0	0	0	0		**0.005**	0	1	0
	14	0	0	0	0	0	0	0	0	0	0	0	0	0.998		**0.003**	1	0
	15	0	0	0	0	0	0	0	0	0	0.174	0	0	0.417	0.222		1	0.98
	16	0	0	0	0.322	0	0.595	1	1	1	0.996	1	0.998	0.606	0.566	0.855		1
	17	0	0	0	0	0	0	0	0	0	0.803	0	0	0	0	0.907	0.952	

Fig. 8: GMUs in relation to the WRB Reference Groups of the HWSD Soil data base.

preliminary classification according to the WRB reference groups were taken into account (Fig. 9). Because of the uncertainties regarding the classification we refrain from mixing the data bases. The classification of "old floodplains" as Fluvisols is obsolete; whilst the Fluvisols are occurring rather in the recent floodplains [Gröngröft, personal statement]. Because of the low number of samples the results are quite uncertain and hardly comparable.

Nevertheless, some trends are visible (Fig. 9). The data of the Ministry of Environment and Tourism Namibia shows the highest share of Ferralic Arenosols in the sandvelds, whilst in the valley floors Haplic Calcisols and Arenic Fluvisols are much more important. For the data of SP03 it can be stated that Luvisols are exclusively found in the valley floors; the share of Arenosols are highest in the SLD-group; Plinthisols, Ali/Acrisols, and Histosols are only to be found in the SSA-group. The valley floors and the SLD-group are much more homogeneous regarding the WRB reference groups than the SSA-group is.

GMUs versus Soil Texture

Figure 10 shows an analysis of soil types of the "Soil property maps of Africa at 1 km" of the International Soil Reference and Information Centre (ISRIC 2013). Again, some trends are reasonable. For example the sandy clays are restricted to the MVF. On the other hand the results are not adequate, e.g. the lack of sand in the longitudinal dunes (Petersen 2008), or the relatively little share of sand in the sandveld.

An analysis of soil types of the SP03-data was conducted, comprising 163 topsoils samples (after excluding plots with no data or plots outside of the GMU coverage in Seronga, which leaves 163 samples, cf. Figure 11). Silty clay can be found mainly in TVF. Loam is more or less evenly distributed over MVF, TVF, and SLD. Sandy loam and loamy sand occurs in all four groups without focus. Sand has its main area of distribution in the SLD. The TVF are much more diverse regarding the soil texture than the other groups, except for very small shares in the MVF.

Table 5 shows the mean values of the ratio of clay and silt to sand, as well as the minimum and maximum values. A high ratio stands for a high sand content. Table 6 shows the analysis of variance for all mean values. Only 6 comparisons are significant.

Discussion

In general, DEM based terrain analysis and especially the segmentation of morphological units provides suitable baseline information for analyzing the landscape structure with respect to vegetation, soil, landforms etc. Especially in areas with poor data availability, an auxiliary method like a semi-automated terrain-based landscape structure classification delivers useful approaches for further investigations; on the other hand this procedure is very challenging.

The GMU approach delivers, even if not simple in the implementation, very useful landscape units in a reproducible manner, which can be applied in various ways; even if the GMU approach has to be adapted respectively revised for the application in other areas. Overall, analyzes done in this paper show that
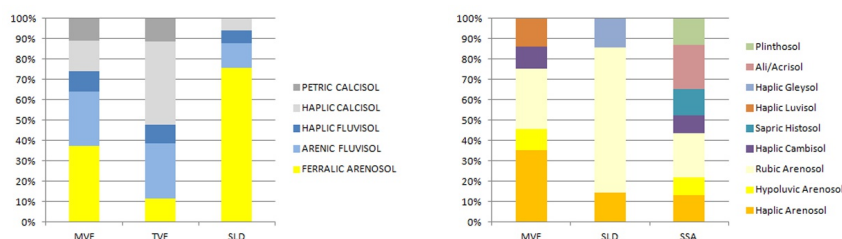

Fig. 9: WRB reference groups of the Ministry of Environment and Tourism data base (631 Profiles, to the left). Right: 106 profiles with WRB reference group of the SP03, Institute of Soil Science, Hamburg University are considered, shown are these with a share of more than five percent
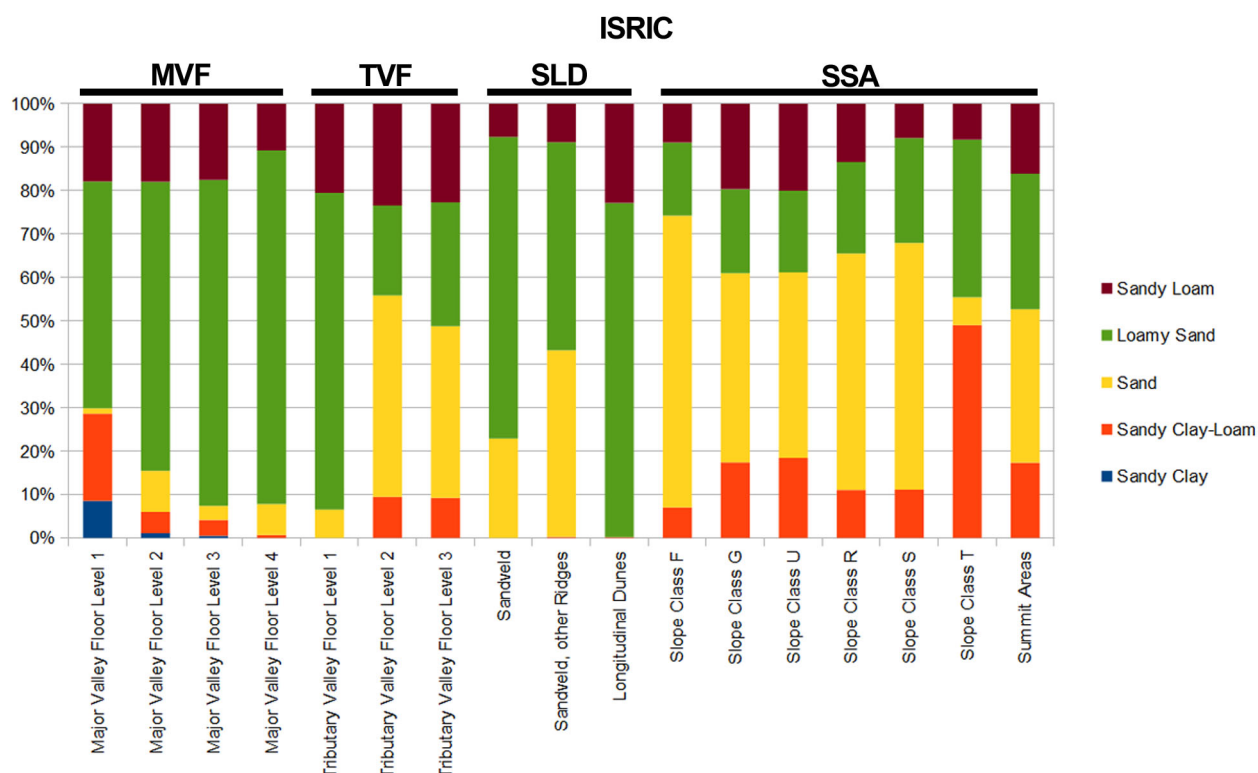
ISRIC

Fig. 10: ISRIC, Top Soils, Percentage of Soil Types within the GMUs.

there are distinct determinations between the GMUs and that every single of the 17 GMUs stands for specific settings/characteristics of the landscape.

Despite the fact that the GMUs are based only on DEM and expert knowledge at large scale, the results are presentable even at local scale. In Figure 12 the GMUs are compared to the Landscape Units (LSUs) of SP03. The LSUs were

Table 5: Ratio of clay and silt to sand of 163 topsoils of SP03 data base.

Ratio Clay+Silt / Sand	MVF	TVF	SLD	SSA
Mean	5.62	1.60	7.93	6.93
Min	0.03	0.05	0.23	1.74
Max	12.47	4.00	12.47	12.47
No. Profiles	70	16	35	42

preliminary derived by using a 300 x 300 m² point grid as base for a first visual stratification of vegetation cover and soil color. The subsequent soil survey by SP03 served as ground truthing for the final delineation. Even though the LSUs are generated in a very different way with other specific intensions, at local level the major features are well reproduced. On
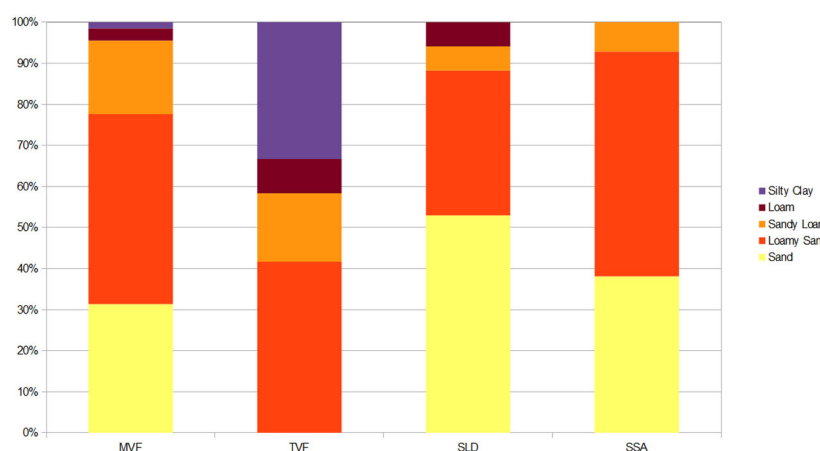
the other hand the limitations of the GMUs are obvious. Small structures like the dry riverbed are recognizable but are only in parts classified correctly. In this case the ephemeral river is part of the sandveld which leads to the fact that the "Gleysol" (Fig. 12) of the Fumbe basin appears as part of it.

In general very small structures, like riverbeds in the mountainous areas, could not be included by the GMU-algorithm, i.e. the spatial resolution of the SRTM is not sufficient. Very flat areas with low

relief intensity are a limitation for the GMU approach as well. Therefore, the Delta is not covered by the GMUs. That means that some of the soil profiles of SP03 are outside of the GMU coverage. Another problem with the delineation of the GMUs is the separation of highlands and lowlands at an early stage of the workflow. Indisputably there was no alternative to this way of proceeding to merge the gently undulated lowlands with the mountainous areas of Angola. Without this segmentation, either the structures in

Fig. 11: GMUs versus soil texture of 163 topsoil samples of SP03 (percentage per GMU main group).
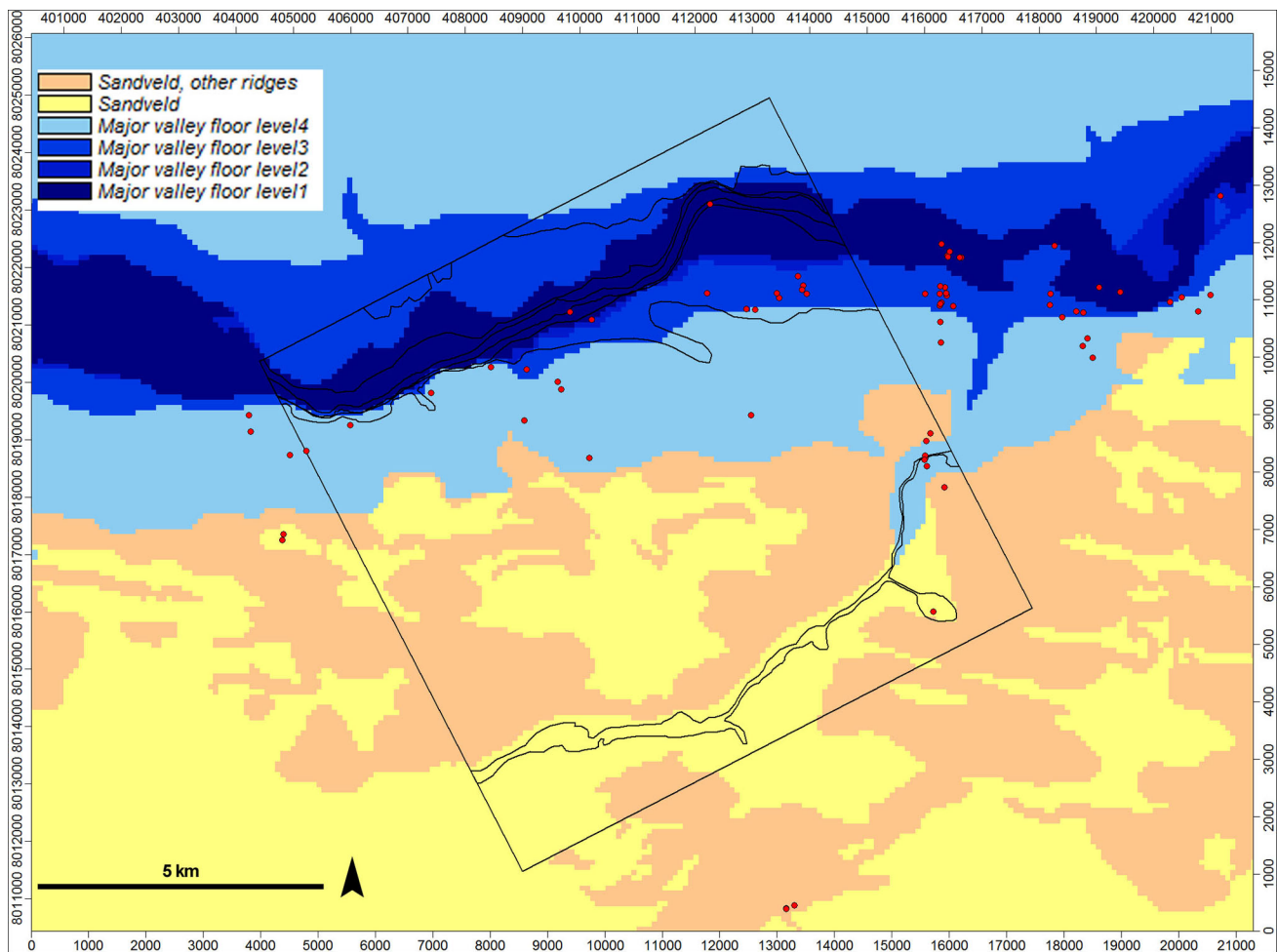
Fig. 12: GMUs (background, colored) compared to LSUs (outlined, black) of SP03. Soil Profiles of SP03 shown in red.

the north or these in the south would not be outlined in a sufficient way. However, this segmentation leads to the fact that some of the GMUs are entirely restricted to the northern or to the southern parts. If, for example, the summit areas or steep slopes show the highest EVI values it is also due to the circumstances that these areas are only to be found in the high productive, more or less rain fed areas in Angola.

NDVI and EVI are suitable for different purposes and thus cannot be compared directly. Moreover, variations can be caused by time of the "shot" (NDVI is from 2013.03.22 and EVI from a comparable period of the year 2011). However, in this paper the vegetation indices were not discussed exhaustive. Productive areas are situated in the slopes and summits, which is mainly due to the fact that these areas are located in the northern parts (see above) of the catchment and does not mean that these areas are in general more productive than the riversides. For the purpose of this paper it is sufficient that the Anova/Tukey HSD-Test shows mainly significant

Table 6: Analysis of variance of mean values for the clay+silt/sand-ratio for the subunits (Tukey's Test, Honestly Significant Difference): Significance level (confidence interval 95%) of multiple comparisons. Significant values are displayed in bold.

		MVF		TVF		SLD		SSA		
		3	**4**	**5**	**6**	**9**	**10**	**11**	**12**	**13**
MVF	**1**	0.913	0.998	0.469	0.264	0.941	0.971	1.000	1.000	0.604
	3		0.214	0.874	0.891	0.115	0.097	1.000	0.614	**0.032**
	4			0.184	**0.019**	1.000	1.000	0.993	1.000	0.917
TVF	**5**				1.000	0.104	0.121	0.795	0.336	**0.037**
	6					**0.010**	**0.008**	0.869	0.082	**0.003**
SLD	**9**						1.000	0.934	0.962	0.999
	10							0.964	0.984	0.985
SSA	**11**								1.000	0.660
	12									0.634

Table 7: Survey of the GMU main groups with selected parameters.

GMU Upper Level	Distribution/Characteristics (Landform, Productivity, Soil)
Major Valley Floors MVF	• Small in total area (only 3% of the catchment) but partly with the largest continuous domain (Okavango river bed). Permanently flat. • Primarily persistent in the southern part • Flooded areas (often, very often) to great extent, settlements/ smallholder agriculture/open Savanna (only group), Mopane woodlands (only supergroup), Terminalia woodlands, some dense Miombo forest in class 1 • Low values for NDVI (around 0.5), partly very low values for EVI [because of water surface areas in 1 and 2] (lower than 0.26 in 1 and 2, similar to sandvelds and dunes in 3 and 4. • perennial swamps, Fluvisols, Eutric Arenosols, Distric Arenosols, Petric Calcisols • Soil texture: Sand, loamy sand, sandy loam
Tributary Valley Floors TVF	• Small in area and quite dispersed. Often flat. • Primarily persistent in the northern part • Flooded areas (often, very often), grassland on sandy soils (only supergroup), open Savanna • Slightly higher NDVI values than major valley floors and sandveld and dunes (nearly 0.6), increasing values for EVI from level 1 to 3, in total low productivity (around 0.3). • Mostly various Arenosols (Eutric, Rubi-Ferralic), some Ferralsols, Petric Calcisols in level 1, Albi-Gleyic Arenosols • Soil texture: Loamy sand and silty clay
Sandvelds and Longitudinal Dunes SLD	• Second largest supergroup with a share of 42% of the catchment, relatively large mean areas. Not continuously flat. • Primarily persistent in the southern parts • Medium dense woodland, Terminalia woodlands, grassland • Low values for NDVI (around 0.5). EVI values very homogenous from 0.36 to 0.38. • Dominant WRB Reference Group: Dystric Arenosols (highest share), Eutric and Rubi-Ferralic Arenosols • Soil texture: Sand
Slopes and Summit Areas SSA	• Large Area (more than half of the catchment), but very disperse (85% of the number of shapes), very small mean areas. From flat to steep. • Steep slopes and summits primarily persistent in the northern part • Miombo forest (dense/medium to dense, only supergroup), Brachystegia Miombo, grassland agriculture (slope class G, U, R), some grassland with Terminalia (but no Terminalia woodlands) • Highest NDVI values (almost 0.7), highest values for EVI (around 0.39). • Rubi-Ferralic Arenosols very common, Acric Ferralsols, Xanthic Ferralsols • Soil texture: Loamy sand

differences between the GMUs.

Some GMUs are very small and covered only by a few shapes. GMU Nr. 16, the steepest slope class (T) has only 19 shapes and an area of only 1.32 square kilometers which makes it the smallest GMU. This is reflected by the analysis of variance of NDVI mean values where it has the lowest share of significant differences. It should be taken in consideration to allocate this GMU to the slope class S. The major valley floors level 1 are small in numbers as well, but this GMU has a by far wider area, because the wide and coherent areas of the panhandle and the Okavango river bed are part of this GMU.

Because of the specific workflow the land surface variables for the GMUs are, in some cases, not directly comparable to each other. For example the values of the SAGA wetness index SWI for the sandvelds and longitudinal dunes are higher than these for the tributary valley floors.

The units are coherent and reproduce the physical settings well. According to the GMUs the landscape of the Okavango catchment can be described in short as shown in Table 7.

Conclusion

The GMU concept is a promising tool for landscape analysis and mapping purposes. It provides added value especially in areas with poor data coverage as it can provide information about geomorphological settings. We think it will prove useful for extrapolation of point data such as soil properties or even support validation purposes of remote sensing products in different contexts. The approach may be further tested through a transfer to other regions and enhanced by automating more processes. The concept performs best in mountainous regions where relief

intensity is high. In flat areas and areas with small entities (like the longitudinal dunes) a post-processing is required. Thus, the approach is constrained by the quality of the digital elevation model that serves as input.

Acknowledgements

References

Abdi, H., Williams, L.J. (2010): Tukey's Honestly Significant Difference (HSD) Test. In Salkind, N. [Ed.]; Encyclopedia of Research Design. Thousand Oaks, CA: Sage.

Bock, M., Böhner, J., Conrad, O., Köthe, R., Ringeler, A. (2007): Methods for creating Functional Soil Databases and applying Digital Soil Mapping with SAGA GIS. - In: Hengl, T. et al. [Eds.] Status and prospect of soil information in south-eastern Europe: soil databases, projects and applications. p.149–163, Scientific and Technical Research series, Office for Official Publications of the European Communities; Luxemburg.

Drăguţ, L., Eisank, C. (2012): Automated object-based classification of topography from SRTM data. – Geomorphology **141**: 21–33.

Elvidge, C.D., Lyon, R.J.P. (1985): Influence of rock soil spectral variation on the assessment of green biomass. – Remote Sensing of Environment **17**: 265–279. CrossRef

ESRI 2011. ArcGIS Desktop: Release 10. Redlands, CA: Environmental Systems Research Institute.

Ester, M., Sander, J. (2000): Knowledge Discovery in Databases. VERLAG Berlin.

FAO/IIASA/ISRIC/ISSCAS/JRC, 2012. Harmonized World Soil Database (version 1.2). FAO, Rome, Italy and IIASA, Laxenburg, Austria.

Gao, X., Huete, A.R., Ni, W., Miura, T. (2000): Optical – biophysical relationships of vegetation spectra without background contamination. – Remote Sensing of Environment **74**: 609–620. CrossRef

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

IBM Corp. Released 2012. IBM SPSS Statistics for Windows, Version 21.0. Armonk, NY: IBM Corp.

ISRIC – World Soil Information (2001-2003): Development of a soil and terrain database for Southern Africa (SOTERSAF). URL: www.isric.org/projects/soter-southern-africa-sotersaf

ISRIC – World Soil Information (2013): Soil property maps of Africa at 1 km. Available for download at www.isric.org.

IUSS Working Group WRB (2006): World reference base for soil resources 2006. World Soil Resources Reports No. 103. FAO, Rome

Iwahashi, J., Pike, R.J. (2007): Automated classifications of topography from DEMs by an unsupervised nested-means algorithm and a three-part geometric signature. – Geomorphology **86**: 409–440. CrossRef

Jarvis, A., Reuter, H.I., Nelson, A., Guevara, E. (2008): Hole-filled SRTM for the globe Version 4, available from the CGIAR-CSI SRTM 90m Database URL: http://srtm.csi.cgiar.org

Jönsson, P., Eklundh, L. (2002): Seasonality extraction by function fitting to time-series of satellite sensor data. – IEEE Transactions on Geoscience and Remote Sensing **40(8)**: 1824–1832. CrossRef

Köthe, R., Bock, M. (2009): Preprocessing of Digital Elevation Models - derived from Laser Scanning and Radar Interferometry - for Terrain Analysis in Geosciences. – Proceedings of Geomorphometry, Zurich, Switzerland, 31 August - 2 September: 155–161; Zürich. URL: http://www.geomorphometry.org/system/files/koethe2009geomorphometry.pdf

Krug, D., Hagedorn, J., Adler, G.H., Bock, M., Gehrt, E., Köthe, R., Stegger, U. (2005): Die morphographische Karte der Bundesrepublik Deutschland im Maßstab 1:1.000.000 (GMK1000). Mitt. DBG **107(2)**: 737–XXX.

Lee, J.-S. (1980): Digital Image Enhancement and Noise Filtering by Use of Local Statistics. – IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume PAMI-2, Number 2, pages 165–168, March 1980. CrossRef

McCloy, K.R. (2006): Resource Management Information Systems: Remote Sensing, GIS and Modeling 2nd edition. Taylor and Francis Group.

MacMillan, R.A., Shary, P.A. (2009): Landforms and Landform Elements in Geomorphometry. In: Hengl, T., Reuter, H.I.: Geomorphometry. Concepts, Software, Applications. VERLAG: Amsterdam. CrossRef

Petersen, A. (2008): Pedodiversity of southern African Drylands. – Hamburger Bodenkundliche Arbeiten **61**: 1–374.

SAGA GIS System for Automated Geoscientific Analyses. URL: http://www.saga-gis.org

Schneibel, A., Stellmes, M., Frantz, D., Finckh, M., Revermann, R. (2013): Cusseque - Earth Observation. – Biodiversity & Ecology **5**: 55–57. CrossRef

Stellmes, M., Frantz, D., Finckh, M., Revermann, R. (2013): The fire regime within the Okavango Catchment as seen by MODIS. – Biodiversity & Ecology **5**: 351–362. CrossRef

Thomas, D.S.G., O'Connor, P.W., Bateman, M.D., Shaw, P.A., Stokes, S., Nash, D.J. (2000): Dune activity as a record of late Quaternary aridity in the Northern Kalahari: new evidence from northern Namibia interpreted in the context of regional arid and humid chronologies. – Palaeogeography, Palaeoclimatology, Palaeoecology **156(3)**: 243–259. CrossRef

van Engelen, V.W.P., Wen, T.T. (1995): Global and National Soil and Terrain Digital Databases (SOTER). Procedures Manual (revised ed.). UNEP, Nairobi; ISSS, Vienna; FAO, Rome; ISRIC, Wageningen. 125 p.

Affiliations

Jan Wehberg* (wehberg@geowiss.uni-hamburg.de), Michael Bock (bock@geowiss.uni-hamburg.de), Thomas Weinzierl (thomas.weinzierl@uni-hamburg.de), Olaf Conrad (conrad@geowiss.uni-hamburg.de), Jürgen Böhner (boehner@geowiss.uni-hamburg.de)
Department of Geography
University of Hamburg
Bundesstraße 55
22222 Hamburg, GERMANY

Marion Stellmes (stellmes@uni-trier.de)
Department of Environmental Remote Sensing and Geoinformatics, Faculty of Regional and Environmental Sciences
University of Trier
Behringstrasse 21
54286 Trier, GERMANY

Lars Landschreiber (lars.landschreiber@uni-hamburg.de)
Institute of Soil Science
University of Hamburg
Allende-Platz 2
20146 Hamburg, GERMANY

* Corresponding author