

Delineating dambo catenary soil-landscape units using aerial gamma-ray and terrain data: a comparison of classification approaches

Jerome S. Lugumira^a, David J. Brown^{b*}, Philip E. Dennison^c, Mathew K. Hansen^c,
and Lee A. Vierling^d

^aSchool of the Environment, Washington State University, Pullman, WA 99164-2812, USA;
^bDepartment of Crop and Soil Sciences, Washington State University, PO Box 64620, Pullman, WA,
USA; ^cDepartment of Geography, University of Utah, Salt Lake City, UT 84112-9155, USA;
^dDepartment of Forest, Rangeland and Fire Sciences, College of Natural Resources, University of
Idaho, PO Box 441135, Moscow, ID 83844-1135, USA

(Received 3 July 2014; accepted 2 October 2014)

Soil data are largely absent for most of Africa. For landscapes with recognizable catenary elements, this data gap can be filled by mapping the catenary units and assigning them with known soil properties. An example is the landscape map for a region with dambos in central Uganda, which shows the four catenary units in order from well-drained to seasonal wetland: uplands, margins, floors, and bottoms. However, this map was created using optical data, which are cost prohibitive and are also limited by cloud cover. We evaluated the potential of freely available aerial gamma-ray spectrometry (AGRS) data as an alternative source of classification inputs. Analysis of variance based upon field data for a region with dambos in central Uganda showed gamma activity to differ along the catenary sequence, with landscape position explaining an appreciable proportion of variation of potassium (28%), thorium (27%), and uranium (46%). Using the three gamma channels, together with terrain indices derived from the Shuttle Radar Topography Mission (SRTM) digital elevation model (DEM) as inputs, three classifiers were evaluated – conditional inference trees (CITs), random forests (RF), and multinomial-iterative self-organizing data analysis (ISODATA). While untransformed terrain and gamma predictors were used for the first two methods, we applied the ISODATA classification to landscape unit probability maps generated using multinomial principal components regression. For the CIT classification, all decision rules were based on terrain data, which might explain why the map was slightly less accurate (unweighted kappa = 0.61, linear weighted kappa = 0.73) than the map created using a RF classifier (unweighted kappa = 0.63, linear weighted kappa = 0.74), where both terrain and gamma predictors were used. But the existence of artefacts of margins within uplands in the map based on CIT modelling, and not that created using RF, is because the former missed the smoothing effect of gamma, attributed to zonal differences in activity of all three gamma channels. The multinomial-ISODATA predictions were poor (unweighted kappa = 0.56, linear weighted kappa = 0.67), partly because the regression model could not adequately resolve differences between bottoms and floors. However, we did find the probability maps generated using multinomial regression to be useful end products that capture the continuous nature of landscape unit transitions. It is important to note that in this study we used 90 m grid resolution gamma and terrain data to predict features that transition over distances of less than 10 m, so better results might be possible with finer-resolution gamma and terrain data.

*Corresponding author. Email: david_brown@wsu.edu

1. Introduction

Good-quality soil information is required to guide land management decisions (Cambule, Rossiter, and Stoorvogel 2013; Sanchez et al. 2009). It is also increasingly sought by scientists, who use it for many applications (Scull, Franklin, and Chadwick 2005) such as watershed hydrological (Zhu and Mackay 2001) and climate change (Sanchez et al. 2009) modelling. However, soil data are sparse for many parts of the world (Hartemink, Krasilnikov, and Bockheim 2013). Where they are available, the resolution is inappropriate for hydro-ecologic modelling (Quinn, Zhu, and Burt 2005), and is not suitable for land management applications at the field scale (Sanchez et al. 2009).

Africa is host to countries where soil maps not only lack detail (Dewitte et al. 2013), but are also highly fragmented (Dewitte et al. 2012) and outdated (Dewitte et al. 2012; Odeh et al. 2012) and cannot readily be improved because of limits imposed by the cost of soil surveys (see Dewitte et al. 2012; McBratney, Santos, and Minasny 2003; Moore et al. 1993). For many of these areas, mapping catenary units with predictable soil properties would contribute soil information needed to manage the soil resource (see Hansen et al. 2009). Furthermore, it would provide the African Soil Information Services (AfSIS) project with information required to implement digital soil mapping routines (see McBratney, Santos, and Minasny 2003), especially where legacy soil data are scarce (Sanchez et al. 2009).

Classification of soil–landscape elements has always been an integral activity of soil mapping (Schoknecht, Tille, and Purdie 2004), where the correspondence between landform and soil characteristics (Cook et al. 1996; Odeh, Chittleborough, and McBratney 1991; Summerell et al. 2005) is the basis upon which landscape units are demarcated (Irvin, Ventura, and Slater 1997; Schoknecht, Tille, and Purdie 2004). The process has traditionally been accomplished by manually delineating landscape units from aerial photographs (Burrough, van Gaans, and MacMillan 2000; Hengl and Rossiter 2003; Irvin, Ventura, and Slater 1997), although this is subjective (Burrough, van Gaans, and MacMillan 2000), not reproducible (MacMillan et al. 2000), and the rules are not adaptable across a range of scales (Burrough, van Gaans, and MacMillan 2000). Improvements in computing technology and availability of digital elevation models (DEMs) have allowed replacement of this manual approach by automated landscape classification methods (e.g. Burrough, van Gaans, and MacMillan 2000; Drăguț and Blaschke 2006; Hengl and Rossiter 2003; Irvin, Ventura, and Slater 1997; Iwahashi and Pike 2007; MacMillan et al. 2000; McKenzie and Gallant 2007; Schmidt and Hewitt 2004; Summerell et al. 2005), reliant on terrain parameters, purposely selected to mirror the influence of hydrogeomorphic processes on soil development (MacMillan et al. 2000).

Given the complementary effect of information from different sensor systems (Taramelli and Melelli 2009; Wright and Gallant 2007), multi-sensor approaches have recently been tested and used to delineate landscape elements. Saadat et al. (2008) used parameters from a 10 m-resolution DEM together with advanced spaceborne thermal emission and reflection radiometer (ASTER) data to map landforms in a watershed in Iran; Ehsani and Quiel (2009) combined Landsat Thematic Mapper (TM) with morphometric parameters derived from the Shuttle Radar Topographic Mission (SRTM) DEM and mapped landscape elements for a region at the Poland/Slovakia/Ukraine border; and Hansen et al. (2009) used variables derived from Système Pour l'Observation de la Terre (SPOT-4) and SRTM DEM imagery to accurately delineate landscape elements of a dambo in central Uganda. However, use of high-resolution spectral imagery is limited by costs (Adam, Mutanga, and Rugege 2010), and optical data are affected by cloud cover

(Ju and Roy 2008; Roy et al. 2006; Schroeder, Csiszar, and Morisette 2008). A potential substitute for optical remote-sensing data for soil–landscape classification routines is aerial gamma-ray spectrometry (AGRS). This is because (i) data are collected by aircraft flying at a maximum of 225 m above the ground (Pitkin and Duval 1980), hence cloud cover is not an issue, and (ii) data are readily available in countries where geophysical surveys have been conducted.

AGRS is a passive form of remote sensing (Wilford, Bierwirth, and Craig 1997) that involves the use of a gamma-ray spectrometer, mounted on the belly of an aircraft, to measure high-frequency gamma photons emitted during the decay of naturally occurring radioactive isotopes of potassium (K), thorium (Th), and uranium (U) (Pitkin and Duval 1980). For land surface applications, the instrument is able to take measurements integrated over the upper part of the solum (McKenzie and Ryan 1999) – within the first 30–45 cm of the soil surface (Bierwirth 1996; Dickson and Scott 1997; Wilford et al. 2001) – hence providing data reflective of bedrock mineralogy and geochemistry, and soil development (Wilford, Bierwirth, and Craig 1997). However, the amount of energy recorded by the sensor is affected by several factors, including soil moisture and vegetation (Minty 1997). For example, a 10% increase in soil moisture can reduce the amount of radiation by the same proportion (Minty 1997; Wilford, Bierwirth, and Craig 1997); although precipitation can also result in increased concentration of U on the ground, since rain is contaminated by dust particles onto which daughter products of airborne radon (a nuclide in the U decay series) are attached (Minty 1997). The amount of gamma radiation attenuated by vegetation is dependent upon vegetation density (Wilford, Bierwirth, and Craig 1997). Where vegetation is sparse, energy is mostly unaffected (Wilford, Bierwirth, and Craig 1997). However, for densely vegetated areas, gamma radiation reaching an aerial platform is reduced (Wilford, Bierwirth, and Craig 1997), so that correction for vegetation interference is necessary (see Lavreau and Fernandez-Alonso 1991).

K, U, and Th are the only elements of interest in aerial radiometrics because during decay, their radionuclides (or the radionuclides of their daughter products) produce sufficient energy that can be detected using an airborne detector (Minty 1997). It is from the characteristic emittance peak of a radioisotope that the abundance of the element is estimated (Wilford et al. 2001). K is the only element whose concentration is directly measured following the decay of ^{40}K to argon (^{40}Ar) (Dickson and Scott 1997; Minty 1997). The concentrations of U and Th are estimated using their radioactive daughter products (Minty 1997; Wilford 2012), because these elements themselves do not emit gamma rays during decay (Dickson and Scott 1997; Minty 1997). As a result, their measurements are expressed in units of ‘equivalent’ parts per million (e.g. eU and eTh), while K is expressed as a percentage (K%) (IAEA 2003; Wilford et al. 2001).

Historically, AGRS has been used in geologic mapping (McBratney, Santos, and Minasny 2003; Wilford, Bierwirth, and Craig 1997), owing to the correspondence between the abundance of radioactive elements and the nature of bedrock mineralogy and geochemistry (Wilford 2012). Because the distribution of these elements is modified by pedologic processes acting on the bedrock (Pitkin and Duval 1980), aerial gamma-ray data are now increasingly integral to soil mapping. To date, AGRS has been used to predict the spatial distribution of soil phosphorus (e.g. McKenzie and Ryan 1999); identify landscape anomalies, such as salt stores (e.g. Wilford et al. 2001); map regolith density (e.g. Beckett 2007); delineate geomorphic (e.g. McKenzie and Gallant 2007) and soil-mapping units (e.g. Triantafilis, Gibbs, and Earl 2013); develop weathering intensity indices (e.g. Wilford 2012); and predict regolith thickness (e.g. Wilford and Thomas 2013). All these studies were conducted outside Africa, despite availability of free AGRS data in most African countries where geophysical surveys have been conducted.

In this study, we used AGRS and topographic data to delineate landscape elements of a dambo in central Uganda. Dambos are seasonally inundated grassy valleys found on headward ends of drainage systems (Ackermann 1936) in seasonally wet tropical and sub-tropical locations in Africa (Meadows 1985). They are channel-less (Boast 1990; Mäkel 1973), linear (Mäkel 1973; Meadows 1985), or lobate (Whitlow 1985) depressions, subdivided into three landscape elements – bottoms, floors, and margins. Margins are adjacent to the interfluvium (uplands). Their widths vary from narrow (Hansen et al. 2009; Mäkel 1973) to broad belts, hundreds of metres in width (Mäkel 1973). They are sandy (Mäkel 1985; Young 1976) and have sparse vegetation, mainly grasses (Bullock 1992; Hansen et al. 2009; von Der Heyden 2004). Floors fringe bottoms, and are the largest portions of dambos (Mäkel 1973). They have clayey subsoils and are more vegetated than the margins (Bullock 1992). At the centre of a dambo is the bottom, the size of which is also variable. Here, soils have more clay, vegetation is herbaceous and denser than at the floors, and the water table is usually close to the surface (Mäkel 1973; Acres et al. 1985; Mäkel 1985). We sought to examine whether aerial gamma-ray data are sensitive to the differences in soil properties across dambo cross-profiles, and therefore whether inclusion of AGRS data improves separability of dambo landscape elements in a classification where inputs also include DEM-derivatives. We compared three classification approaches in order to choose a method that maximizes information from the two data sources. With dambos occupying an estimated 11% of Africa's arable land (Hansen et al. 2009), use of AGRS and DEM as the only sources of predictors of dambo soil-landscape units is an inexpensive way to collate soil information for a significant proportion of Africa's arable land.

2. Methods

2.1. Study area

The study area is found in central Uganda (Figure 1), in a region with undulating topography underlain by Precambrian granitic gneisses. The area experiences a tropical climate characterized by two rainfall seasons, with mean annual precipitation and mean annual temperature of ~120 cm and 23°C, respectively (Survey Department 1967).

Vegetation in the area varies along cross-profiles in a manner similar to what has been observed in other regions with dambos (e.g. Bullock 1992). The low-lying bottoms and floors have sedges (e.g. *Cyperus alba* and *Cyperus denudata*), grasses (e.g. *Commelina subulata* and *Setaria sphacelata*), and forbs (e.g. *Dyschoriste magchena* and *Emilia javanica*), but height and density of cover are lower on the floors, perhaps due to animal grazing (Hansen et al. 2009). Grasses, such as *Paspalum scrobiculatum* and *Hyparrhenia filipendula*, and the forb, *Murdania simplex*, dominate the margin, while in the uplands, woody bushy vegetation of the family Combretaceae dots a mat of grass dominated by *C. subulata* and the drought-tolerant grass, *Brachiaria brizantha* (Hansen et al. 2009).

The study was restricted to the area in Figure 1 because in the northeast and southwest of this area are swampy streams. To include swampy areas would require use of data other than what has been proposed, given that gamma activity recorded over them does not differ from measurements made in the wetter elements of dambos. Attempts are under way to use other data types (e.g. active radar) to differentiate these swampy rivers from the lowest members of dambos in areas where they occur alongside each other.

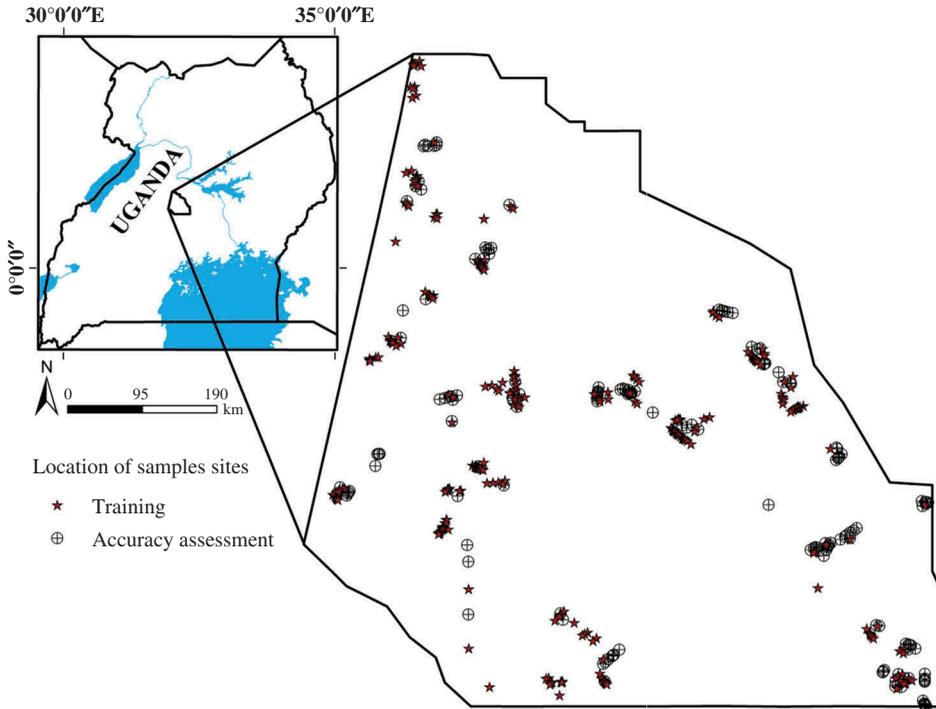


Figure 1. Location of the study area. Included are the sites where training and validation polygons were delimited by Hansen et al. (2009).

2.2. Data sources and classification inputs

2.2.1. Terrain data

A 90 m SRTM DEM was the source of terrain variables used in this study. It was downloaded in its native US Geological Survey (USGS) format from the Global Land Cover Facility (GLCF) website (<http://www.landcover.org/>) and re-projected to the Universal Transverse Mercator spatial reference. The DEM was corrected for elevation anomalies (e.g. tall vegetation and termitaries in the area) by smoothing with a nine-pixel neighbourhood mean (e.g. Hansen et al. 2009). From this DEM, we extracted elevation data and computed terrain indices, namely, topographic wetness index (TWI), relative slope position (RSP), topographic position index (TPI), tangential curvature, and elevation relative rank (RR). These variables were thought to be sufficient to differentiate landscape elements because: (i) some are proxies of hydrologic processes (e.g. TWI and TPI) that account for the differences in soils and vegetation observed in dambos; (ii) others (e.g. elevation, RSP, RR, and tangential curvature) are intended to differentiate elements that cannot otherwise be differentiated using TWI or TPI alone; and (iii) RR successfully distinguished a dambo from the upland when it was used in the study area (see Hansen et al. 2009).

The topographic wetness index is used to predict saturation by quantifying topographic control of hydrology (Wilson and Gallant 2000; Sørensen, Zinko, and Seibert 2006), and is therefore a useful predictor for systems such as dambos where surface hydrology plays an important role in the variability of soil properties (Young 1976; Mäkel 1973, 1985; Hansen et al. 2009) and vegetation composition (Whitlow 1985;

Bullock 1992) along cross-profiles. RSP represents the percentage distance of a location from slope bottom to the nearest ridge top. It is used in particular to enable low-gradient locations found in the bottom of the landscape to be differentiated from their counterparts in the uplands. Where this was not possible, elevation and/or tangential curvature were expected to overcome the ambiguities because uplands comprise the elevated part of the landscape, characterized by convex hilltops. The topographic position index is the difference between a smoothed and unsmoothed DEM (e.g. Roberts, Dowling, and Walker 1997; Guisan, Weiss, and Weiss 1999; Wilson and Gallant 2000). It is a measure of the relative depth of the water table, assuming the water table conforms to the topography (Roberts, Dowling, and Walker 1997). Our use of TPI is premised on the consideration of groundwater level as an alternative discriminator of landscape elements, especially where surface flow conditions for areas with subtle differences in elevation are inappropriately modelled using TWI (Böhner and Selige 2006). RR defines elevation of a location relative to its neighbourhood, and is suitable to classify catenary units for a landscape with inconstant elevation range (see Hansen et al. 2009).

Prior to calculating TWI, and in order to ensure correct hydrologic flow, the fill tool in ArcGIS 10.2 (ESRI 2013) was used to correct the DEM for sinks. TWI was computed using the wetness index tool in SAGA (2.1.0) geographic information systems (GIS) software. This is because for landscapes such as dambos, where terrain is characterized by low amplitude and wide valleys which are also levelled in places, small differences in elevation would result in poorly modelled flows at cells in valleys when a specific catchment area (SCA) is determined using algorithms other than SAGA's modified SCA (see Böhner and Selige 2006). We also used tangential curvature in place of plan curvature because the latter would exaggerate curvature for locations with very low gradient (see Wilson and Gallant 2000).

In the absence of information regarding the version of TPI that best approximates the relationship between topography and water table depth in the study area, 10 versions of TPI were used. They were differentiated by varying the radius of the processing window from 200 to 2000 m. Similarly, 12 versions of RR were used. These were derived by varying the processing window from 11×11 cells to 231×231 cells. This was also intended to enhance and contrast local (small windows) and regional (large windows) topographic conditions (see Hansen et al. 2009).

2.2.2. Aerial gamma-ray data

The Ministry of Energy and Mineral Development, Uganda, contracted Fugro Airborne Surveys (Pty) Limited to collect and process AGRS data as part of the High Resolution Airborne Geophysical Survey Programme, involving magnetic, radiometric, and electromagnetic surveys of almost three-quarters of the country. Data were collected between 2006 and 2008, and are archived at the Department of Geological Survey and Mines, Ministry of Energy and Mineral Development, Uganda. These were provided in a processed form after removing cosmic and background noise, levelling, correction for height, and amplifying the low signal-to-noise regions of the survey (Fugro 2008). For the selected area, gamma-ray measurements were made in 2007 based on the following survey specifications: 100 m of mean terrain clearance, with traverse (flight-line spacing) and control lines (tie-line spacing) at 500 and 5000 m, respectively. We considered all three channels – K%, eTh, and eU. The spatially geocoded data points of each channel were interpolated using ordinary Kriging in ArcGIS 10.2 (ESRI 2013). The rasters were set to the same spatial resolution and assigned a similar spatial reference as the DEM.

2.2.3. Model training and validation data

The model training and calibration data used in this study are the same as those used by Hansen et al. (2009). They consisted of (i) randomly selected polygons delimited around areas belonging to the same dambo class and (ii) polygons based on locations sampled along selected cross-section transects (see Hansen et al. 2009). These data were intended for use with images gridded at 20 m. Unlike the validation dataset, the training data were modified to suit the data sources used in this study. Since aerial gamma-ray and terrain data have 90 m pixels, the polygons were used to create a new training dataset made up of point features. This was a two-stage process involving a polygon-to-raster conversion, followed by a raster-to-point transformation. To convert raster to points, point features are positioned at centres of pixels, and the distance between points is determined by the raster resolution. We tested several pixel dimensions and found a 50 m pixel raster to be appropriate. This is because the resultant points had a neighbourhood distance of ~71 m (equivalent to the diagonal of a 50 m pixel), which ensured that many training sample locations would not lie in the same 90 m pixel. This minimized the tendency to generate an over-fit conditional inference tree (CIT) – one of the proposed modelling approaches. A total of 869 training sample locations were generated – 174, 189, 227, and 279 locations representing bottom, floor, margin, and upland catenary units, respectively. Onto these, pixel values of the respective predictor variables were extracted and the data used as described in the following section.

In addition to these data, a second set was needed to analyse gamma variability along dambo cross-profiles. This included all locations sampled for gamma-ray activity and bound by the training and accuracy assessment polygons described above. In all, 102 samples were selected: 15 in bottoms, 20 in floors, 29 in margins, and 38 in uplands.

2.3. Statistical analysis and landscape modelling

2.3.1. Variability of gamma activity along a dambo cross-profile

We used R (R Development Core Team 2013) and the 102 samples described in Section 2.2.3 to construct box plots. These were intended to illustrate differences in gamma-ray activity along dambo cross-profiles. To determine whether there is any significant difference in K%, eTh, and eU activity between dambo classes, the same software was used to implement a single-factor analysis of variance (ANOVA), where dambo landscape classes were considered to be the factor levels for which mean (μ_i) responses of K%, eTh, and eU were derived. The null model – K% or eTh or eU concentration is the same across a dambo cross-section – was tested using the F-test for equality of factor levels. If this was not the case (i.e. the null model rejected at $\alpha = 0.05$), differences in mean gamma activity amongst classes were assessed using a modification of the Tukey–Kramer multiple comparison procedure, suitable when a dataset has unequal class sizes (Herberich et al. 2010). In addition, we sought to determine the proportion of variance in K%, eTh, and eU activity attributed to landscape position, by deriving the coefficient of determination (R^2) for each scenario.

2.3.2. Predicting the spatial distribution of dambo catenary units

2.3.2.1. *Terminal classifiers.* The distribution of dambo catenary units was predicted using the approaches shown in Figure 2, which differ based upon the terminal classifier. CITs and random forests (RF) are variants of classification tree analysis (CTA). These

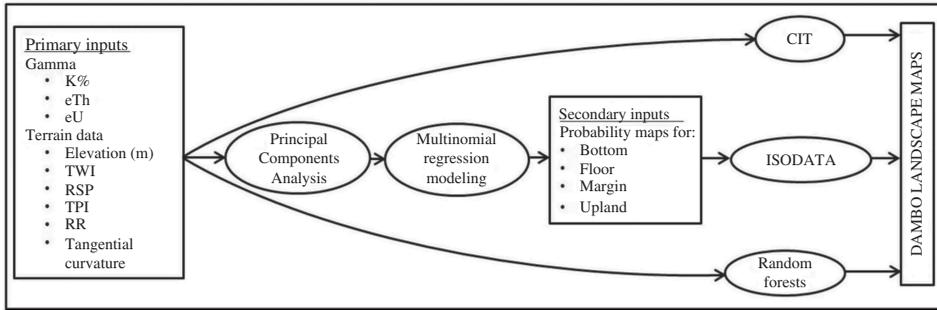


Figure 2. Schematic diagram showing the general flow of the modelling approaches. ISODATA, Iterative Self-Organizing Data Analysis Technique.

involve induction of rules based upon input predictor variables, and use of these rules to repeatedly partition sample data in such a way that at each successive split (node) purer descendant nodes are created – nodes are expected to be dominated by a single class (out of n target classes or objects) at each successive split (Breiman et al. 1984; Franklin, McCullough, and Curtis 2000; Simard, Saatchi, and De Grandi 2000; Xu et al. 2005). The process is terminated using a stopping criterion (Everitt and Hothorn 2010), resulting in what are referred to as terminal nodes (leaves) at the lowest level of the hierarchy (Simard, Saatchi, and De Grandi 2000). A terminal node is assigned a class label with more samples in that node (Simard, Saatchi, and De Grandi 2000). Among the advantages of CTA include its ability to handle both ordered and categorical data in a single suite, its insensitivity to data with outliers, and the ease with which interpretations can be drawn from the structure of the decisions (Breiman et al. 1984). Furthermore, CTA techniques are non-parametric (Breiman et al. 1984; Simard, Saatchi, and De Grandi 2000; Adam, Mutanga, and Rugege 2010) and therefore no normality assumptions have to be made (Wright and Gallant 2007).

CIT is a simpler form of CTA. It differs from other single-tree classifiers based on the selection criteria of features used to split nodes. The decision about which variable to use at a split is made by conducting a test of independence between the response and competing covariates, and a variable is elected based upon a p -value lower than a predetermined α value (Everitt and Hothorn 2010). In regard to RF, an ensemble of trees is grown (Breiman 2001) where each tree is independently constructed using a bootstrap sample of the training data, and at each node a split is effected using randomly selected variables (Breiman 2001; Liaw and Wiener 2002). This increases the possibility of using most of the variables in the classification process. Class predictions for new data are made by aggregating predictions from n number of independently grown trees (Liaw and Wiener 2002).

The iterative self-organizing data analysis (ISODATA) procedure is parametric and is commonly used in unsupervised clustering of sensor data (e.g. Irvin, Ventura, and Slater 1997; Lillesand, Kiefer, and Chipman 2008). The algorithm requires that the user defines the number of clusters or classes, k . It then searches in multidimensional attribute space to assign means to these clusters (Ventura and Irvin 2000). This process is repeated n times, where at each iteration statistics describing the clusters (e.g. standard deviation, distance between mean points of clusters) are evaluated to determine whether they deviate from preset thresholds (Lillesand, Kiefer, and Chipman 2008). If they do, clusters are merged,

split, or deleted (Lillesand, Kiefer, and Chipman 2008). The process is terminated if cluster statistics do not change or when the maximum n is reached (Lillesand, Kiefer, and Chipman 2008). In ArcGIS 10.2, the procedure is accomplished using the Iso Cluster tool to define the clusters, and the maximum likelihood classification (MLC) algorithm to implement the classification, using parameters in the signature file created by the Iso Cluster tool.

2.3.2.2. *Implementation of the classification methods.* The training data described in Section 2.2 was used to build a CIT with the help of the *party* package (Hothorn et al. 2013) in R (R Development Core Team 2013). The emergent decisions were used to classify the input raster data, creating a landscape map of the study area. The *randomForest* package (Liaw and Wiener 2013) in R (R Development Core Team 2013) was used to generate decisions based upon the same training data used in CIT modelling. We maintained the default parameters (e.g. number of variables sampled at a node, minimum samples required to split a node) and constructed forests (e.g. models) with 100, 300, 600, and 900 trees. The model with the lowest out-of-bag error rate (e.g. 600 trees) was used to predict the distribution of landscape units over the entire study area. For this model, we also extracted variable importance measures in order to examine the contribution of each variable to node purity (e.g. mean decrease in Gini coefficient) and overall class prediction (e.g. mean decrease in accuracy) (see Breiman 1996; Breiman 2001; Liaw and Wiener 2002).

The inputs used in ISODATA classification were derived via an intermediate step – multinomial regression modelling (Figure 2). Since the variables described in Section 2.2 are correlated (results not shown), not all could be used in the same multinomial regression model because this would impact the reliability of model coefficients. Using ArcGIS 10.2, principal components (PCs) were derived after standardizing the rasters. Fifteen PCs accounting for ~100% of the total variation were selected (Table 1). The first two PCs do not seem to be dominated by any single variable (Table 1). PCs 5 and 9 are dominated by gamma-ray variables (Table 1) and can therefore be considered to represent variability in K% (e.g. PC 5), eTh, and eU (e.g. PC 9). The remainder are dominated either by terrain-based variables or a combination of terrain and aerial gamma-ray data (Table 1).

Pixel values were extracted from the imagery (15 PCs) onto the training data, to create a set of predictors used in multinomial regression modelling. Generally a multinomial regression model constitutes n logits, where n is the number of categories that define the response variable. In the case of this analysis, four logits were expected – one for each landscape element. The logits for the non-baseline category are defined as

$$\pi_{ij} = \left(\frac{\exp(\boldsymbol{\beta}_j \cdot \mathbf{x}_i^T)}{1 + \sum_{k \neq j^*} \exp(\boldsymbol{\beta}_k \cdot \mathbf{x}_i^T)} \right), \quad (1)$$

where \mathbf{x} represents the predictors (e.g. the most significant PCs), j^* is the baseline category (e.g. bottom), and j denotes the non-baseline categories (e.g. floor, margin, and upland zones), so that $j \neq j^*$ (see Kutner et al. 2005). Predictions for the baseline category = $1 - \sum \pi_{ij}$. To set up these logits, the *mlogit* package (Croissant 2012) in R was used. Each logit was then implemented in ArcGIS 10.2 (ESRI 2013) to create a probability map. The four probability maps were classified using ISODATA classification, resulting in a single soil–landscape unit map.

The accuracy of the maps was tested using the validation data collected by Hansen et al. (2009). The maps based on the three modelling approaches were independently

Table 1. Projection of terrain and aerial gamma-ray variables on the first 15 PCs, accounting for ~99.9% of the variation.

PCs	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
VE†	62.74	19.28	5.47	3.40	2.49	1.89	1.28	0.85	0.69	0.51	0.45	0.39	0.24	0.13	0.08
Elevation	-0.18	-0.21	-0.16	-0.04	-0.04	-0.22	-0.24	-0.17	-0.13	-0.24	-0.69	-0.11	-0.37	-0.16	-0.17
RSP	-0.21	-0.07	0.14	0.06	0.01	0.08	0.24	0.08	-0.22	0.00	-0.14	0.88	-0.07	-0.01	0.01
TWI	0.09	0.26	0.26	-0.04	0.17	-0.58	0.61	-0.28	-0.06	-0.09	-0.06	-0.12	-0.01	0.02	0.00
T. curvature	-0.08	0.19	-0.38	-0.47	0.04	0.47	0.24	-0.52	-0.08	-0.11	0.06	0.01	-0.08	0.05	-0.02
Topographic position index (m)	-0.17	0.23	-0.19	-0.24	0.04	-0.04	0.09	0.28	0.02	0.19	-0.27	-0.04	0.38	-0.44	0.24
200	-0.18	0.24	-0.16	-0.18	0.02	-0.08	0.02	0.28	0.03	0.15	-0.14	-0.02	0.11	0.03	-0.10
400	-0.19	0.23	-0.13	-0.12	0.01	-0.10	-0.02	0.22	0.02	0.11	-0.04	-0.01	-0.05	0.22	-0.17
600	-0.19	0.23	-0.09	-0.06	-0.02	-0.11	-0.06	0.14	0.01	0.08	0.04	0.00	-0.15	0.28	-0.14
800	-0.20	0.21	-0.05	0.01	-0.04	-0.11	-0.10	0.03	-0.01	0.07	0.11	0.01	-0.20	0.25	-0.07
1000	-0.20	0.20	-0.01	0.08	-0.07	-0.09	-0.13	-0.07	-0.02	0.06	0.13	0.02	-0.18	0.15	0.02
1200	-0.21	0.18	0.03	0.13	-0.09	-0.07	-0.14	-0.16	-0.03	0.07	0.13	0.01	-0.11	0.01	0.08
1400	-0.21	0.16	0.06	0.17	-0.11	-0.04	-0.14	-0.22	-0.03	0.08	0.09	0.01	-0.01	-0.12	0.10
1600	-0.21	0.14	0.08	0.19	-0.12	0.00	-0.12	-0.26	-0.03	0.09	0.04	0.00	0.09	-0.22	0.09
1800	-0.21	0.12	0.10	0.21	-0.13	0.03	-0.10	-0.28	-0.03	0.10	-0.02	-0.02	0.18	-0.26	0.05
2000	-0.19	0.19	-0.05	0.03	0.00	-0.01	0.01	0.28	0.10	-0.81	0.27	0.01	-0.10	-0.24	0.15
Elevation relative rank	31	-0.21	0.06	0.14	-0.06	0.23	0.10	-0.04	0.06	-0.26	-0.20	-0.10	0.47	0.19	-0.52
31	-0.21	-0.03	0.16	0.17	-0.04	0.27	0.23	0.08	0.09	-0.02	-0.25	-0.20	0.01	0.32	0.26
51	-0.22	-0.09	0.16	0.06	-0.03	0.21	0.23	0.11	0.09	0.10	-0.10	-0.20	-0.25	0.05	0.31
71	-0.21	-0.14	0.14	-0.03	-0.01	0.12	0.19	0.10	0.07	0.13	0.05	-0.15	-0.27	-0.16	0.02
91	-0.21	-0.17	0.12	-0.10	0.01	0.04	0.13	0.06	0.05	0.11	0.14	-0.10	-0.17	-0.22	-0.21
111	-0.20	-0.19	0.10	-0.14	0.03	-0.02	0.07	0.02	0.03	0.08	0.18	-0.05	-0.05	-0.18	-0.28
131	-0.20	-0.20	0.08	-0.17	0.04	-0.07	0.01	-0.01	0.02	0.04	0.18	-0.02	0.06	-0.09	-0.22
151	-0.20	-0.21	0.06	-0.18	0.05	-0.11	-0.03	-0.04	0.00	-0.01	0.15	0.00	0.14	0.01	-0.10
171	-0.20	-0.21	0.04	-0.18	0.05	-0.13	-0.07	-0.06	-0.01	-0.04	0.10	0.00	0.18	0.11	0.07
191	-0.19	-0.22	0.00	-0.17	0.06	-0.14	-0.10	-0.08	-0.03	-0.08	0.04	0.01	0.20	0.20	0.23
211	-0.19	-0.22	0.00	-0.16	0.06	-0.15	-0.13	-0.09	-0.04	-0.10	-0.03	0.01	0.20	0.25	0.35
231	0.02	-0.21	-0.34	0.04	-0.83	-0.20	0.30	0.02	0.07	-0.01	0.08	0.02	0.06	0.04	0.01
K%	-0.12	-0.16	-0.42	0.39	0.27	-0.01	0.19	0.12	-0.66	0.04	0.17	-0.19	0.04	-0.01	0.00
eTh	-0.11	-0.16	-0.44	0.34	0.37	-0.13	0.12	-0.15	0.66	0.08	0.01	0.16	-0.01	-0.01	0.00
eU															

Note: †Percentage of total variance (VE) explained by the PC.

assessed and compared using the usual estimates of class accuracy (e.g. producer's and user's accuracies) and overall prediction accuracy (e.g. overall accuracy and linearly weighted kappa). Linearly weighted kappa was preferred over unweighted kappa because dambo landscape elements are ordinal – owing to the prevailing environmental gradients (e.g. bottoms are wetter than the outward-lying floors, margins, and uplands). Given that Hansen et al. (2009) used unweighted kappa to assess the accuracy of their map, we also derived a similar statistic for each of the three maps and compared the results to the statistics in the article by Hansen et al. (2009).

3. Results

3.1. Gamma activity along a dambo cross-profile

Gamma concentration is shown to vary along dambo cross-profiles (Figure 3). K activity is highest at margins and lowest at uplands, while eTh and eU activity increases toward the uplands (Figure 3 and Table 2). These trends mirror a dependence of gamma activity on landscape position, particularly eU, ~50% of whose variability is explained by landscape position (Table 2). However, it is eTh with concentrations that significantly differ in most landscape elements – the concentrations between bottom and margin, bottom and upland, floor and upland, and margin and upland, are significantly different (Figure 4). In regard to K% and eU, differences are only significant between uplands and the remainder of the landscape (Figure 4).

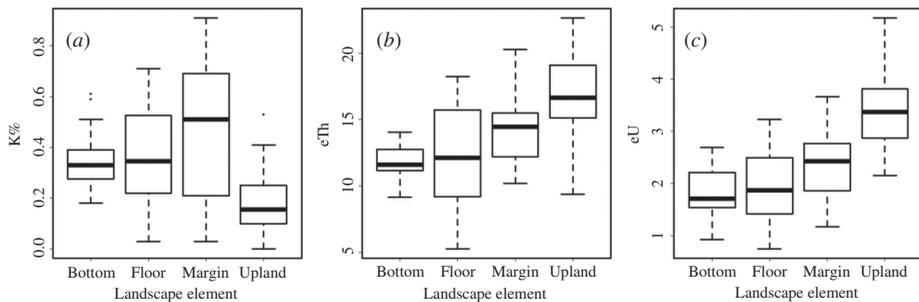


Figure 3. Variability in gamma activity along a dambo cross-profile. (a) is K%, (b) is eTh, and (c) is eU.

Table 2. Results of ANOVA analysis.

Gamma	ANOVA P -values	Adjusted	Predicted mean (μ_i) for each zone			
		R^2	Bottom	Floor	Margin	Upland
%K	<0.0001	0.282	0.350	0.358	0.458	0.175
eTh	<0.0001	0.273	11.9	12.4	14.5	16.7
eU	<0.0001	0.460	1.87	1.92	2.34	3.37

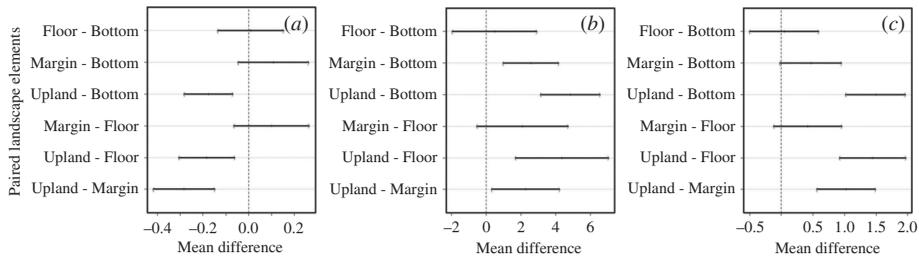


Figure 4. Multiple comparison of differences in mean activity of K% (a), eTh (b), and eU (c) at 95% family-wise confidence level.

3.2. Performance of modelling approaches

The root node of the CIT attests to the observation that dambos are generally low-lying (Figure 5). It is also shown that the drier uplands can be reliably separated from the remainder of the landscape using proxies of potential saturation (e.g. TPI and RSP) (Figure 5). The remainder of the landscape elements were differentiated using a combination of elevation and potential of saturation of land (Figure 5). For example, three decisions based on elevation were used to differentiate bottoms from either of floors and margins, and one decision based on hydrologic position was used to separate bottoms from the remainder of the landscape (Figure 5). However, none of the aerial gamma-ray channels were used in CIT modelling, despite their potential (Table 2). RF variable-importance measures show that gamma activity strongly influences the prediction of floors (e.g. eTh and K%), margins, and uplands (e.g. eU and eTh), and the overall strength of the RF model (Figures 6 and 7). Terrain variables are equally important. As an example, the accuracy of bottoms, floors, and margins would severely be affected if RR11 was not part of the inputs (Figure 7), while removing TPI400 would lower the accuracy of uplands by over 15% (Figure 7). Jointly, TPI400 is an important predictor (Figure 6) and, like TPI600, TPI800, RR11, and TPI1000, the nodes attributed to it are purer (Figure 6).

For the multinomial-ISODATA approach (Figure 2), the final multinomial model is based on 13 PCs (Table 3) because initial analysis showed PCs 6 and 11 to be insignificant across all logits (results not shown). In this model, most PCs are significant in more than one logit (Table 3). As was the case with RF, PCs dominated by terrain variables strongly influenced predictions in all logits (coefficients are positive and relatively large). However, they were complemented by PCs whose variability is attributed to aerial gamma-ray channels (Table 3). For instance, PC 5 is dominated by K% and influenced the prediction of uplands (where K% activity is lowest) while PC 9 is jointly influenced by eTh and eU and contributed to predictions of margins and uplands (where the concentration of Th and U is highest). Principal component 7, which is generally influenced by K% and TWI, had a similar effect on the prediction of margins and uplands (Table 3).

Predictions resulting from the four logits are shown in Figure 8. While margins and uplands appear to be well predicted, the probability maps of bottoms and floors show difficulties in separating the two classes. For example, in the north and east of the study area, some locations tend to have approximately equal probability of membership to the two classes (Figure 8). This explains the positive correlation between the two maps (Table 4) and the tendency to confuse bottoms with floors (Table 5 and Figure 9). This problem compounds the limits imposed by the difficulty to discriminate marginal zones.

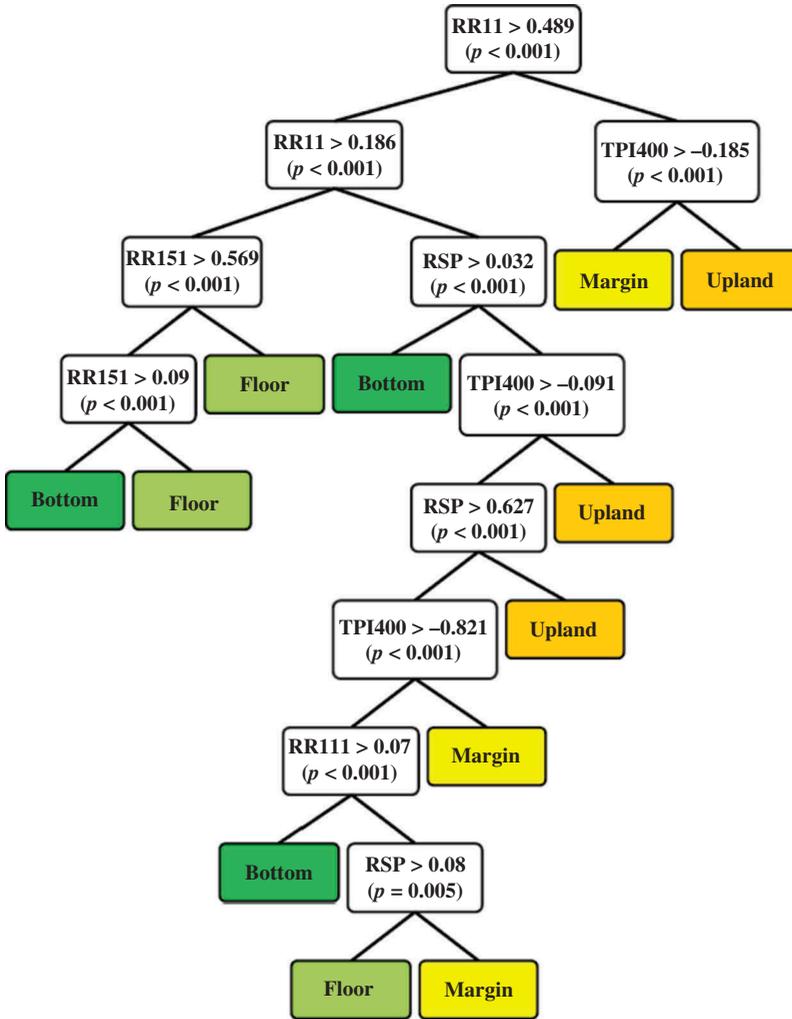


Figure 5. Decision tree based on CIT modelling.

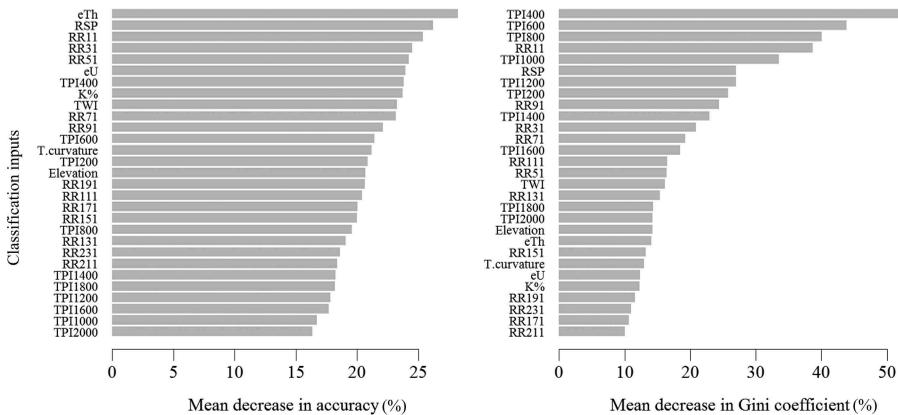


Figure 6. Importance of variables to the overall prediction of dambo landscape units using RF.

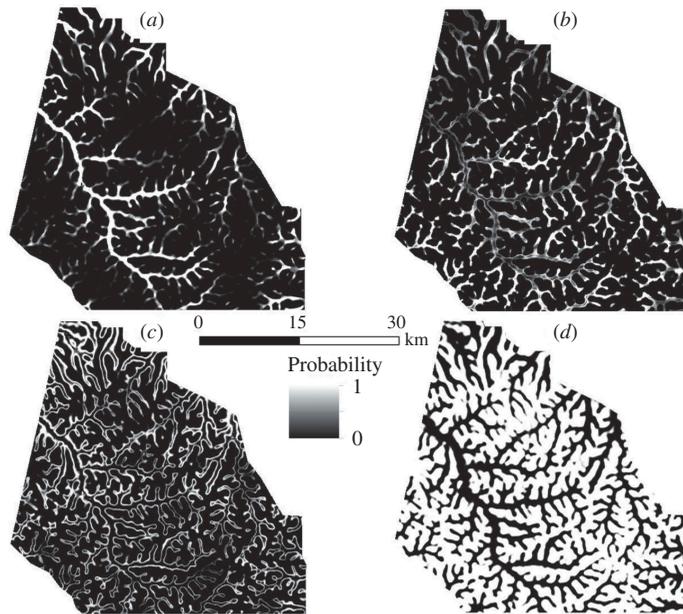


Figure 8. Results of multinomial regression modelling showing the potential occurrence of bottom (a), floor (b), margin (c), and upland (d) landscape elements.

Table 4. Correlation between probability maps.

Probability maps	Floor	Margin	Upland
Bottom	0.204	-0.381	-0.506
Floor		-0.100	-0.633
Margin			-0.416

Overall, margins and uplands were fairly well classified using the three modelling approaches. This is the reason why weighted kappa is high, especially for the CIT and RF methods (Table 6). In general, omission and commission errors are comparable to those of the output by Hansen et al. (2009) (Table 5), and accuracy statistics (overall accuracy and kappa) are not very different (Table 6) save for the ISODATA approach, which poorly classified bottoms and floors (Table 5). The difference between CIT and RF methods is seen in the way the former classified margins. These were clearly over-represented, resulting in high omission and commission errors for the margin and upland classes (Table 5). This translated to a weighted kappa difference of 1% (Table 6) but does not seem to be significant. However, a close visual inspection of the map created using CIT reveals artefacts of margins within upland zones (Figure 9).

4. Discussion

The observed variability of gamma activity along dambo cross-profiles (Table 2 and Figure 3) corresponds with the frequently reported differences in soil properties, particularly when inferences are limited to subsurface soils readily sampled by a spectrometer on board an aerial platform. The low concentration of K measured in the uplands is attributed

Table 5. Error matrices based on the three classification methods.

	B	F	M	U	Total	UA (%)	B	F	M	U	Total	UA (%)
CIT	B	353	179	26	0	558	63.3	RF	86	0	698	50.4
	F	112	624	175	0	911	68.5		100	6	790	68.0
	M	53	294	1001	287	1635	61.2		1040	169	1528	68.1
	U	0	0	95	1124	1219	92.2		71	1236	1307	94.6
	Total	518	1097	1297	1411	4323		518	1297	1411	4323	
	PA (%)	68.2	57.0	77.2	79.7		68.0	49.0	80.2	87.6		
Multinomial-ISODATA	B	269	299	108	16	692	39.0	Hansen et al. (2009)	88	22	772	67.4
	F	159	564	179	34	936	60.3		232	0	1364	73.5
	M	90	234	917	168	1409	65.1		1159	294	1767	65.6
	U	0	0	93	1193	1286	92.8		102	1396	1498	93.2
	Total	518	1097	1297	1411	4323		706	1402	1581	1712	5401
	PA (%)	52.0	51.4	70.7	84.6		73.7	71.5	73.3	81.5		

Included is the matrix for the map created by Hansen et al. (2009). Shown are accuracy statistics based on the number of correctly classified pixels. B, bottom; F, floor; M, margin; U, upland; UA, user's accuracy; PA, producer's accuracy.

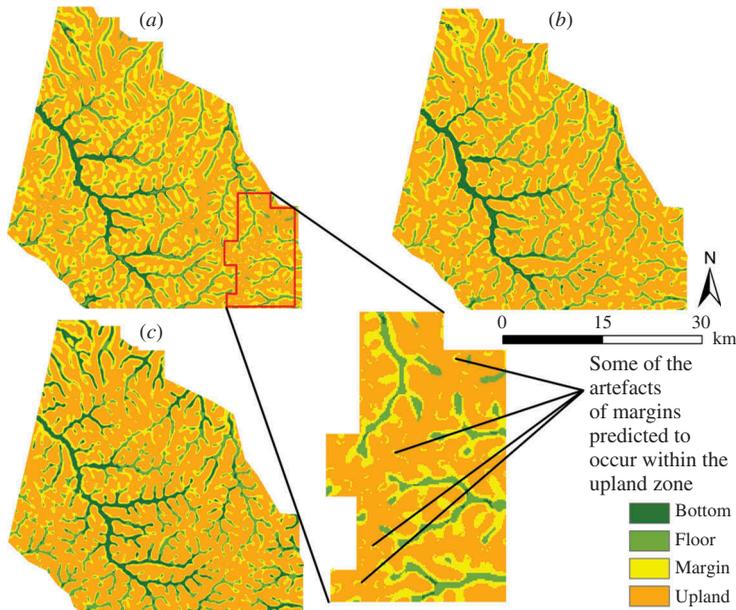


Figure 9. Dambo landscape maps created using: (a) CIT – included is an inset showing artefacts of margins within the upland zone; (b) RF; and (c) ISODATA.

Table 6. Accuracy level of the respective mapping methods.

	Overall accuracy (%)	Kappa	
		Unweighted	Weighted
CIT	71.8	0.61	0.73
RF	73.2	0.63	0.74
Multinomial-ISODATA	68.1	0.56	0.67
Hansen et al. (2009)	75.5	0.67	–

to losses of K that accompany extensive weathering (Dickson and Scott 1997; Wilford, Bierwirth, and Craig 1997; Wilford 2012), which is characteristic of this zone (Brown, Helmke, and Clayton 2003; Mäkel 1973, 1985; Young 1976). In dambo margins, the dominance of sand (Hansen et al. 2009; Mäkel 1973, 1985; Young 1976) derived from felsic gneisses constituting the area's geology (Brown, Helmke, and Clayton 2003) accounts for the increased K% activity. Since sand is also washed to the floor (Young 1976), there is an equivalent increase in K concentration relative to the bottom, where the concentration of clay is highest (Young 1976; Hansen et al. 2009).

Thorium and eU activity are highest in uplands (Table 2 and Figure 3) because clay and residual oxides are present (see Young 1976; Hansen et al. 2009) onto which these radio-elements preferentially adsorb (see Dickson and Scott 1997; Kiss, De Jong, and Bettany 1988; Wilford, Bierwirth, and Craig 1997; Wilford 2012). The concentration of these elements is low where soils are weakly weathered (e.g. margins). In regard to the more low-lying members of the landscape (e.g. bottoms and floors), it appears that the

presence of water and/or dense vegetation act to reduce the amount of gamma energy recorded by the sensor (Cook et al. 1996; Minty 1997) although the trends fits our expectation – this is a zone with soils not as developed as those found at the interfluvium (Mäkel 1973), where K% activity is lowest and eTh and eU activity is highest.

Owing to differences in gamma activity along dambo cross-profiles, gamma variables complemented terrain data, resulting in a fairly accurate map (e.g. based on RF modelling). However, the accuracy level of this map was not very different from that of the map created using a CIT (where terrain variables masked the contribution of gamma inputs to the classification), implying that information in gamma variables is redundant in the presence of terrain data, especially when a simpler classifier is used. However, the presence of artefacts of margins in uplands in this map, and not the map based on RF modelling (Figure 9), signifies the value of gamma – gamma had a smoothing effect on the RF output, hence correcting anomalous class assignment, particularly at the interfluviums. Furthermore, the minimized tendency to confuse uplands for margins (Table 5) in this map is partly attributed to eTh and eU (Figure 7), elements the activity of which significantly differs at the two landscape locations (Figure 4). Therefore, in the interest of using all relevant information in the classification process, a robust classifier (e.g. RF) is favoured.

Probability surfaces created by multinomial regression modelling can also be used because of the flexibility of the modelling process, that is, the ability to objectively select the most important variables prior to regression modelling. In addition, the approach allows an examination of the magnitude of influence of each variable on the logits. Thus the outputs can be interpreted in the context of this relationship. Further, the outputs or probability maps are optimized and therefore readily classifiable (Figure 8), although this depends on how well each variable defines the landscape class. This also depends on how well the model fits. In our case, it seems to be fair but not adequate (~62%), with the result that two probability maps (i.e. bottom and floor) were positively correlated. This impacted the ability of ISODATA to discriminate bottoms from floors (Table 5), because the probability that a location is either class was largely indeterminate (Table 4). Otherwise ISODATA classification would have been appropriate if the probability that a location was a given dambo class tended to be 1 for that class and 0 for all other landscape classes. In this way, feature clouds representing catenary units would be perfectly separated in multi-dimensional space, lying nearer to the terminus of the respective axes.

A major limitation of this study was the spatial resolution of the data used. At 90 m, some of the narrow dambo zones are difficult to represent. Moreover, dambo landscape elements have indistinct environmental boundaries and are therefore prone to misclassification. The generally high omission and commission errors associated with bottoms and floors are indicative of this. Therefore, better results might be obtained if finer-resolution aerial gamma-ray and terrain data are used.

5. Conclusion

We found landscape position to have a proportionate influence on the variability of eU (46), K% (28%), and eTh (27), arising from differences in activity of these elements along dambo cross-profiles, due in part to variable soil properties. We sought to exploit these differences by using these data together with terrain variables to delineate dambo landscape elements, with the help of three classifiers: CITs, RF, and ISODATA. For the last classifier, inputs were landscape probability maps resulting from multinomial regression modelling of PCs, derived from gamma-ray and terrain variables.

The maps were generally accurate – CIT (unweighted kappa = 0.61, weighted kappa = 0.73), RF (unweighted kappa = 0.63, weighted kappa = 0.74), and multinomial-ISODATA (unweighted kappa = 0.56, weighted kappa = 0.67) – and did not differ much from the map created by Hansen et al. (2009). Results also show that terrain data provided more information. In fact on their own, terrain data allowed creation of an output that was only slightly different from that based on both data sources although it had errors in the upland class, which did not appear when gamma-ray imagery was included. A look at the variable importance statistics that accompanied the RF classification showed a strong contribution of eTh and eU to the prediction of uplands. This is the reason that artefacts of margins are not in uplands of the RF map, because the concentrations of both radio-elements significantly differ at margins and uplands. Although for this study the information contained in aerial gamma-ray imagery is largely redundant in the presence of terrain data, especially where a simple classifier is used, there may be other applications where aerial gamma-ray data provide more unique information.

Generally, it was difficult to distinguish bottoms from floors because the imagery used had low spatial resolution (90 m), which cannot allow separation of most of the narrow structures of these landscape elements. A convenient solution is to merge these landscape elements, given that they were mostly separable from margins and uplands. We should also consider using probability maps created using multinomial regression modelling. Given that these illustrate a location's probability of class membership and not the definitive class, they are most suited to show the distribution of the continuous catenary units of dambos.

Acknowledgement

Aerial gamma-ray data were freely provided by the Department of Geological Survey and Mines, Ministry of Energy and Mineral Development, Uganda.

Funding

This work was supported by the National Science Foundation Geography and Regional Science Program [grant number 0620206]. The first author benefited from the Fulbright Foreign Student Program doctoral training grant.

References

- Ackermann, E. 1936. "Dambos in Nordrhodesien." *Wiss. Veroff. Dt. Mus. Landerkde. Leipzig NF* 4: 147–175.
- Acres, B. D., A. B. Rains, R. B. King, R. M. Lawton, J. B. Mitchell, and L. J. Rackham. 1985. "African Dambos: Their Distribution, Characteristics and Use." *Zeitschrift Fur Geomorphologie Supplementband* 52: 63–86.
- Adam, E., O. Mutanga, and D. Rugege. 2010. "Multispectral and Hyperspectral Remote Sensing for Identification and Mapping of Wetland Vegetation: A Review." *Wetlands Ecology and Management* 18 (3): 281–296. doi:10.1007/s11273-009-9169-z.
- Beckett, K. 2007. "Mapping Porosity and Density Changes in Soil and Regolith from 256-Channel Radiometric Data." *ASEG 2007*, Perth, WA, November 18–23.
- Bierwith, P. N. 1996. "Gamma-Radiometrics: A Remote Sensing Tool for Understanding Soils." *Australian Collaborative Land Evaluation Program Newsletter* 5: 12–14.
- Boast, R. 1990. "Dambos: A Review." *Progress in Physical Geography* 14 (2): 153–177. doi:10.1177/030913339001400201.

- Böhner, J., and T. Selige. 2006. "Spatial Prediction of Soil Attributes Using Terrain Analysis and Climate Regionalization." In *SAGA – Analysis and Modeling Applications*, edited by J. Böhner, K. R. McCloy, and J. Strobl, 13–27. Göttingen: Verlag Erich Goltze GmbH.
- Breiman, L. 1996. "Bagging Predictors." *Machine Learning* 24 (2): 123–140. doi:10.1007/BF00058655.
- Breiman, L. 2001. "Random Forests." *Machine Learning* 45 (1): 5–32. doi:10.1023/a:1010933404324.
- Breiman, L., J. H. Friedman, R. A. Olshen, and C. J. Stone. 1984. *Classification and Regression Trees*. Belmont, CA: Wadsworth International Group.
- Brown, D. J., P. A. Helmke, and M. K. Clayton. 2003. "Robust Geochemical Indices for Redox and Weathering on a Granitic Laterite Landscape in Central Uganda." *Geochimica et Cosmochimica Acta* 67 (15): 2711–2723. doi:10.1016/S0016-7037(03)00104-2.
- Bullock, A. 1992. "Dambo Hydrology in Southern Africa—Review and Reassessment." *Journal of Hydrology* 134 (1–4): 373–396. doi:10.1016/0022-1694(92)90043-u.
- Burrough, P. A., P. F. M. van Gaans, and R. A. MacMillan. 2000. "High-Resolution Landform Classification Using Fuzzy K-Means." *Fuzzy Sets and Systems* 113 (1): 37–52. doi:10.1016/S0165-0114(99)00011-1.
- Cambule, A. H., D. G. Rossiter, and J. J. Stoorvogel. 2013. "A Methodology for Digital Soil Mapping in Poorly-Accessible Areas." *Geoderma* 192: 341–353. doi:10.1016/j.geoderma.2012.08.020.
- Cook, S. E., R. J. Corner, P. R. Groves, and G. J. Grealish. 1996. "Use of Airborne Gamma Radiometric Data for Soil Mapping." *Australian Journal of Soil Research* 34 (1): 183–194. doi:10.1071/SR9960183.
- Croissant, Y. 2012. "Package 'Mlogit' for R (Version ≥ 2.10)." <http://www.r-project.org>
- Dewitte, O., A. Jones, H. Elbelrhiti, S. Horion, and L. Montanarella. 2012. "Satellite Remote Sensing for Soil Mapping in Africa: An Overview." *Progress in Physical Geography* 36 (4): 514–538. doi:10.1177/0309133312446981.
- Dewitte, O., A. Jones, O. Spaargaren, H. Breuning-Madsen, M. Brossard, A. Dampha, J. Deckers, T. Gallali, S. Halletti, R. Jones, M. Kilasara, P. Le Roux, E. Micheli, L. Montanarella, L. Thiombiano, E. Van Raust, M. Yemefack, and R. Zougmore. 2013. "Harmonisation of the Soil Map of Africa at the Continental Scale." *Geoderma* 211: 138–153. doi:10.1016/j.geoderma.2013.07.007.
- Dickson, B. L., and K. M. Scott. 1997. "Interpretation of Aerial Gamma-Ray Surveys-Adding the Geochemical Factors." *AGSO Journal of Australian Geology and Geophysics* 17 (2): 187–200.
- Drăguț, L., and T. Blaschke. 2006. "Automated Classification of Landform Elements Using Object-Based Image Analysis." *Geomorphology* 81(3–4): 330–344. doi:10.1016/j.geomorph.2006.04.013.
- Ehsani, A. H., and F. Quiel. 2009. "A Semi-Automatic Method for Analysis of Landscape Elements Using Shuttle Radar Topography Mission and Landsat ETM+ Data." *Computers & Geosciences* 35 (2): 373–389. doi:10.1016/j.cageo.2007.09.019.
- ESRI. 2013. *ArcGIS Desktop: Release 10.2*. Redlands, CA: Environmental Systems Research Institute (ESRI).
- Everitt, B., and T. Hothorn. 2010. *A Handbook of Statistical Analyses Using R*. Boca Raton, FL: CRC Press.
- Franklin, J., P. McCullough, and G. Curtis. 2000. "Terrain Variables Used for Predictive Mapping of Vegetation Communities in Southern California." In *Terrain Analysis: Principles and Applications*, edited by J. P. Wilson and J. C. Gallant, 331–352. New York: John Wiley & Sons.
- Fugro. 2008. "Logistics Report: Fixed Wing Borne Magnetic and Radiometric Survey over the Project Area, Block 1 in Uganda." In *Mineral Resources Management and Capacity Building Project*. Entebbe, Uganda: Department of Geological Survey and Mines.
- Guisan, A., S. B. Weiss, and A. D. Weiss. 1999. "GLM versus CCA Spatial Modeling of Plant Species Distribution." *Plant Ecology* 143 (1): 107–122. doi:10.1023/a:1009841519580.
- Hansen, M. K., D. J. Brown, P. E. Dennison, S. A. Graves, and R. S. Brickleyer. 2009. "Inductively Mapping Expert-Derived Soil-Landscape Units within Dambo Wetland Catenae Using Multispectral and Topographic Data." *Geoderma* 150: 72–84. doi:10.1016/j.geoderma.2009.01.013.
- Hartemink, A. E., P. Krasilnikov, and J. G. Bockheim. 2013. "Soil Maps of the World." *Geoderma* 207–208: 256–267. doi:10.1016/j.geoderma.2013.05.003.

- Hengl, T., and D. G. Rossiter. 2003. "Supervised Landform Classification to Enhance and Replace Photo-Interpretation in Semi-Detailed Soil Survey." *Soil Science Society of America Journal* 67 (6): 1810–1822. doi:10.2136/sssaj2003.1810.
- Herberich, E., J. Sikorski, T. Hothorn, and F. Rapallo. 2010. "A Robust Procedure for Comparing Multiple Means under Heteroscedasticity in Unbalanced Designs." *PLoS ONE* 5 (3): e9788. doi:10.1371/journal.pone.0009788.
- Hothorn, T., K. Hornik, C. Strobl, and A. Zeileis. 2013. "Package 'Party' for R (Version \geq 2.14.0)." <http://cran.r-project.org/web/packages/party/index.html>.
- IAEA. 2003. *Guidelines for Radioelement Mapping Using Gamma Ray Spectrometry Data*. Vienna: International Atomic Energy Agency (IAEA).
- Irvin, B. J., S. J. Ventura, and B. K. Slater. 1997. "Fuzzy and Isodata Classification of Landform Elements from Digital Terrain Data in Pleasant Valley, Wisconsin." *Geoderma* 77 (2–4): 137–154. doi:10.1016/s0016-7061(97)00019-0.
- Iwahashi, J., and R. J. Pike. 2007. "Automated Classifications of Topography from Dems by an Unsupervised Nested-Means Algorithm and a Three-Part Geometric Signature." *Geomorphology* 86 (3–4): 409–440. doi:10.1016/j.geomorph.2006.09.012.
- Ju, J., and D. P. Roy. 2008. "The Availability of Cloud-Free Landsat ETM Plus Data over the Conterminous United States and Globally." *Remote Sensing of Environment* 112 (3): 1196–1211. doi:10.1016/j.rse.2007.08.011.
- Kiss, J. J., E. de Jong, and J. R. Bettany. 1988. "The Distribution of Natural Radionuclides in Native Soils of Southern Saskatchewan, Canada." *Journal of Environment Quality* 17 (3): 437–445. doi:10.2134/jeq1988.00472425001700030016x.
- Kutner, M. H., C. J. Nachtsheim, J. Neter, and W. Li. 2005. *Applied Linear Statistical Models*. New York: McGraw-Hill.
- Lavreau, J., and M. Fernandez-Alonso. 1991. "Correcting Airborne Radiometric Data for Water/Vegetation Screening Using Landsat Thematic Mapper Imagery." *8th Thematic Conference on Geologic Remote Sensing*, Denver, CO, April 29–May 2.
- Liaw, A., and M. Wiener. 2002. "Classification and Regression by randomForest." *R News* 2 (3), 18–22 (citeulike-article-id:1121494).
- Liaw, A., and M. Wiener. 2013. "Package 'randomForest' for R (Version \geq 2.5.0)." <http://stat-www.berkeley.edu/users/breiman/RandomForests>.
- Lillesand, T. M., R. W. Kiefer, and J. W. Chipman. 2008. *Remote Sensing and Image Interpretation*. Hoboken, NJ: Wiley.
- Mäckel, R. 1973. "Dambos: A Study in Morphodynamic Activity on the Plateau Regions of Zambia." *Catena* 1: 327–365. doi:10.1016/S0341-8162(73)80018-9.
- Mäckel, R. 1985. "Dambos and Related Landforms in Africa: An Example of the Ecological Approach to Tropical Geomorphology." *Zeitschrift Fur Geomorphologie Supplementband 52*: 1–23.
- MacMillan, R. A., W. W. Pettapiece, S. C. Nolan, and T. W. Goddard. 2000. "A Generic Procedure for Automatically Segmenting Landforms into Landform Elements Using Dems, Heuristic Rules and Fuzzy Logic." *Fuzzy Sets and Systems* 113 (1): 81–109. doi:10.1016/s0165-0114(99)00014-7.
- McBratney, A. B., M. L. M. Santos, and B. Minasny. 2003. "On Digital Soil Mapping." *Geoderma* 117 (1–2): 3–52. doi:10.1016/s0016-7061(03)00223-4.
- McKenzie, N. J., and J. C. Gallant. 2007. "Digital Soil Mapping with Improved Environmental Predictors and Models of Pedogenesis." In *Digital Soil Mapping: An Introductory Perspective*, edited by P. Lagacherie, A. B. McBratney, and M. Voltz, 327–349. Amsterdam: Elsevier Scientific.
- McKenzie, N. J., and P. J. Ryan. 1999. "Spatial Prediction of Soil Properties Using Environmental Correlation." *Geoderma* 89 (1–2): 67–94. doi:10.1016/s0016-7061(98)00137-2.
- Meadows, M. E. 1985. "Dambos and Environmental Change in Malawi, Central Africa." *Zeitschrift Fur Geomorphologie Supplementband 52*: 147–169.
- Minty, B. R. S. 1997. "Fundamentals of Airborne Gamma-Ray Spectroscopy." *AGSO Journal of Australian Geology and Geophysics* 17 (2): 39–50.
- Moore, I. D., P. E. Gessler, G. A. Nielsen, and G. A. Peterson. 1993. "Soil Attribute Prediction Using Terrain Analysis." *Soil Science Society of America Journal* 57 (2): 443–452. doi:10.2136/sssaj1993.03615995005700020026x.

- Odeh, I. O. A., D. J. Chittleborough, and A. B. McBratney. 1991. "Elucidation of Soil-Landform Interrelationships by Canonical Ordination Analysis." *Geoderma* 49 (1–2): 1–32. doi:10.1016/0016-7061(91)90089-c.
- Odeh, I. O. A., J. Leenars, A. E. Hartemink, and I. Amapu. 2012. "The Challenges of Collating Legacy Data for Digital Soil Mapping in Nigeria." In *Digital Soil Mapping and Beyond*, edited by B. Minasny, B. P. Malone, and A. B. McBratney, 453–458. London: Taylor & Francis.
- Pitkin, J. A., and J. S. Duval. 1980. "Design Parameters for Aerial Gamma-Ray Surveys." *Geophysics* 45 (9): 1427–1439. doi:10.1190/1.1441131.
- Quinn, T., A.-X. Zhu, and J. E. Burt. 2005. "Effects of Detailed Soil Spatial Information on Watershed Modeling across Different Model Scales." *International Journal of Applied Earth Observation and Geoinformation* 7 (4): 324–338. doi:10.1016/j.jag.2005.06.009.
- R Development Core Team. 2013. "R: A Language and Environment for Statistical Computing." R Foundation for Statistical Computing, Vienna. <http://www.R-project.org/>
- Roberts, D. W., T. I. Dowling, and J. Walker. 1997. "FLAG: A Fuzzy Landscape Analysis GIS Method for Dryland Salinity Assessment." CSIRO Land and Water Tech. Report 8/97, Canberra, Jul 1997.
- Roy, D. P., P. Lewis, C. B. Schaaf, S. Devadiga, and L. Boschetti. 2006. "The Global Impact of Clouds on the Production of MODIS Bidirectional Reflectance Model-Based Composites for Terrestrial Monitoring." *IEEE Geoscience and Remote Sensing Letters* 3 (4): 452–456. doi:10.1109/lgrs.2006.875433.
- Saadat, H., R. Bonnell, F. Sharifi, G. Mehuys, M. Namdar, and S. Ale-Ebrahim. 2008. "Landform Classification from a Digital Elevation Model and Satellite Imagery." *Geomorphology* 100: 453–464. doi:10.1016/j.geomorph.2008.01.011.
- Sanchez, P. A., S. Ahamed, F. Carre, A. E. Hartemink, J. Hempel, J. Huising, P. Lagacherie, A. B. McBratney, N. J. McKenzie, M. Mendonça-santos, B. Minasny, L. Montanarella, P. Okoth, C. A. Palm, J. D. Sachs, K. D. Shepherd, T. Vagen, B. Vanlauwe, M. G. Walsh, L. A. Winowiecki, and G. Zhang. 2009. "Digital Soil Map of the World." *Science* 325 (5941): 680–681. doi:10.1126/science.1175084.
- Schmidt, J., and A. Hewitt. 2004. "Fuzzy Land Element Classification from Dtms Based on Geometry and Terrain Position." *Geoderma* 121 (3–4): 243–256. doi:10.1016/j.geoderma.2003.10.008.
- Schoknecht, N., P. Tille, and B. Purdie. 2004. *Soil-Landscape Mapping in South-Western Australia: An Overview of Methodology and Outputs*. South Perth, WA: Department of Agriculture.
- Schroeder, W., I. Csiszar, and J. Morissette. 2008. "Quantifying the Impact of Cloud Obscuration on Remote Sensing of Active Fires in the Brazilian Amazon." *Remote Sensing of Environment* 112 (2): 456–470. doi:10.1016/j.rse.2007.05.004.
- Scull, P., J. Franklin, and O. A. Chadwick. 2005. "The Application of Classification Tree Analysis to Soil Type Prediction in a Desert Landscape." *Ecological Modelling* 181 (1): 1–15. doi:10.1016/j.ecolmodel.2004.06.036.
- Simard, M., S. S. Saatchi, and G. De Grandi. 2000. "The Use of Decision Tree and Multiscale Texture for Classification of JERS-1 SAR Data over Tropical Forest." *IEEE Transactions on Geoscience and Remote Sensing* 38 (5): 2310–2321. doi:10.1109/36.868888.
- Sørensen, R., U. Zinko, and J. Seibert. 2006. "On the Calculation of the Topographic Wetness Index: Evaluation of Different Methods Based on Field Observations." *Hydrology and Earth System Sciences* 10 (1): 101–112. doi:10.5194/hess-10-101-2006.
- Summerell, G. K., J. Vaze, N. K. Tuteja, R. B. Grayson, G. Beale, and T. I. Dowling. 2005. "Delineating the Major Landforms of Catchments Using an Objective Hydrological Terrain Analysis Method." *Water Resources Research* 41: 12. doi:10.1029/2005wr004013.
- Survey Department. 1967. *Atlas of Uganda*. Entebbe, Uganda: Government Printer.
- Taramelli, A., and L. Melelli. 2009. "Map of Deep Seated Gravitational Slope Deformations Susceptibility in Central Italy Derived from SRTM DEM and Spectral Mixing Analysis of the Landsat ETM+ Data." *International Journal of Remote Sensing* 30 (2): 357–387. doi:10.1080/01431160802339449.
- Triantafylis, J., I. Gibbs, and N. Earl. 2013. "Digital Soil Pattern Recognition in the Lower Namoi Valley Using Numerical Clustering of Gamma-Ray Spectrometry Data." *Geoderma* 192: 407–421. doi:10.1016/j.geoderma.2012.08.021.

- Ventura, S. J., and B. Irvin. 2000. "Automated Landform Classification Methods for Soil-Landscape Studies." In *Terrain Analysis: Principles and Applications*, edited by J. P. Wilson and J. C. Gallant, 267–294. New York: John Wiley & Sons.
- von der Heyden, C. J. 2004. "The Hydrology and Hydrogeology of Dambos: A Review." *Progress in Physical Geography* 28 (4): 544–564. doi:10.1191/0309133304pp424oa.
- Whitlow, J. R. 1985. "Dambos in Zimbabwe: A Review." *Zeitschrift Fur Geomorphologie Supplementband* 52: 115–146.
- Wilford, J., D. L. Dent, T. I. Dowling, and R. Brataten. 2001. "Rapid Mapping of Soils and Salt Stores: Using Airborne Radiometric and Digital Elevation Models." *AGSO Research Letters* 34: 33–40.
- Wilford, J. R. 2012. "A Weathering Intensity Index for the Australian Continent Using Airborne Gamma-Ray Spectrometry and Digital Terrain Analysis." *Geoderma* 183–184 (0): 124–142. doi:10.1016/j.geoderma.2010.12.022.
- Wilford, J. R., P. N. Bierwirth, and M. A. Craig. 1997. "Application of Airborne Gamma-Ray Spectroscopy in Soil/Regolith Mapping and Applied Geomorphology." *AGSO Journal of Australian Geology and Geophysics* 17 (2): 201–216.
- Wilford, J. R., and M. Thomas. 2013. "Predicting Regolith Thickness in the Complex Weathering Setting of the Central Mt Lofty Ranges, South Australia." *Geoderma* 206: 1–13. doi:10.1016/j.geoderma.2013.04.002.
- Wilson, P. J., and C. J. Gallant. 2000. "Secondary Topographic Attributes." In *Terrain Analysis: Principles and Applications*, edited by J. P. Wilson and J. C. Gallant, 87–131. New York: John Wiley & Sons.
- Wright, C., and A. Gallant. 2007. "Improved Wetland Remote Sensing in Yellowstone National Park Using Classification Trees to Combine TM Imagery and Ancillary Environmental Data." *Remote Sensing of Environment* 107 (4): 582–605. doi:10.1016/j.rse.2006.10.019.
- Xu, M., P. Watanachaturaporn, P. K. Varshney, and M. K. Arora. 2005. "Decision Tree Regression for Soft Classification of Remote Sensing Data." *Remote Sensing of Environment* 97 (3): 322–336. doi:10.1016/j.rse.2005.05.008.
- Young, A. 1976. *Tropical Soils and Survey*. Cambridge, NY: Cambridge University Press.
- Zhu, A. X., and D. S. Mackay. 2001. "Effects of Spatial Detail of Soil Information on Watershed Modeling." *Journal of Hydrology* 248 (1–4): 54–77. doi:10.1016/s0022-1694(01)00390-0.