Journal of Vegetation Science 18: 131-140, 2007

Mapping the floristic continuum: Ordination space position estimated from imaging spectroscopy

Schmidtlein, S.[1,2*]**; Zimmermann, P.**[1]**; Schüpferling, R.**[1] **& Weiß, C.**[3]

[1]*Biogeographie, Universität Bayreuth, DE-95440 Bayreuth, Germany;* [2]*Current address: Geographisches Institut, Universität Bonn, DE-53115 Bonn, Germany;* [3]*Sektion Geographie, Universität München, DE-80333 München, Germany; E-mail c.weiss@iggf.geo.uni-muenchen.de;* [*]*Corresponding author; E-mail s.schmidtlein@uni-bonn.de*

Abstract

Objective: To present a non-classificatory technique of map representation of compositional patterns of vegetation as no two plant species assemblages are completely alike and gradations often occur. Variation is depicted as continuous fields instead of classes.

Location: Murnauer Moos, Bavaria.

Methods: The study combined vegetation ecology and remote sensing methods. The gradual representation of compositional patterns was based on techniques of ordination and regression, instead of mapping class fractions. The floristic field data were collected in relevés and subjected to three-dimensional non-metric multidimensional scaling (NMS). The reflectance information corresponding to plots was gathered from remotely sensed imagery with a high spectral resolution. Reflectance values in numerous wavelengths were related to NMS axes scores by partial least squares regression analysis. The regression equations were applied to the imagery and yielded three grey-scale images, one for each ordination axis. These three images were transformed into a red, green, and blue colour map with a specific colour for each position in the ordination space. Similar colours corresponded to similar species compositions.

Results: Compositional variation was mapped accurately (R^2 = 0.79), using continuous fields. The results took account of various types of stand transitions and of heterogeneities within stands. The map representation featured relatively homogeneous stands and abrupt transitions between stands as well as within-stand heterogeneity and gradual transitions.

Conclusions: The use of NMS in combination with imaging spectroscopy proved to be an expedient approach for non-classificatory map representations of compositional patterns. Ordination is efficiently extended into the geographic domain. The approach in abandoning pre-defined plant communities is able to reconcile mapping practice and complex reality.

Keywords: Floristic composition; Gradient; Gradual transition; Hyperspectral; Pattern; Plant community; Remote sensing; Vegetation mapping.

Nomenclature: Wisskirchen & Haeupler (1998) for phanerogams; Smith (1980) for mosses; Rennwald (2000) for syntaxa.

Abbreviations: NMS = Non-metric Multidimensional Scaling; PLS = Partial least squares regression.

Introduction

Most maps of vegetation at the stand level, including maps derived from remote sensing, consist of classified patches with sharp boundaries dividing them. Such maps do not take into account the various types of stand transition and the variation within stands. Sharp map boundaries do not necessarily follow concrete, local discontinuities – at least when such discontinuities are considered as a within-class variation (Zonneveld 1974). When a compositional within-class variation is ignored, many differences between stands are lost, and the degree of fuzzyness of the concrete boundaries is completely disguised. The fact that conventional vegetation maps reduce imprecise boundaries to mere lines has often been criticized (e.g. Zonneveld 1974; Kent et al. 1997). The 'softness' of a boundary may be an expression of relevant processes like disturbance, invasion, or disintegration; transitional zones are frequently associated with problems of environmental management, and a proper map representation can be advantageous (Weaver & Albertson 1956; van der Maarel 1976; Hobbs 1986; Fortin et al. 2000). Transitional zones are highly affected by environmental change, since species are often at the limit of their tolerance. Accordingly, the movement of vegetation boundaries may indicate environmental change (Fortin et al. 2000). However, for monitoring spatial shifts of 'soft' transitions, linear boundaries are unsuitable because they introduce spatial uncertainty (Foody 1996; Green & Hartley 2000).

A less apparent problem is the classificatory step that precedes the assignment of concrete, local stands to types of vegetation. It takes for granted that plant species assemblages recur, or compositional gradients can be dissected in a helpful way. This issue has caused controversy in the past (see Mueller-Dombois & Ellenberg 1974), which can be traced back to the historical antagonism between the community concept (Humboldt 1807) and alternatives such as the individualistic view or the continuum concept (Gleason 1926; Austin 1985). Since class limits in feature space (and the criteria for

defining them) tend to be a source of dispute and uncertainty, it is an obvious thing to search for complementary mapping methods that do without classification.

In the present paper, we aim to introduce such a complementary, non-classificatory mapping method with the goal of pattern recognition in mixed stands, based on remote sensing and ordination. Some desirable properties would be inevitable by-products of a comprehensive non-classificatory approach: compositional variation would be depicted in continuous fields instead of linear boundaries, and much of the concrete, local variation in plant species composition and the degree of 'softness' of concrete boundaries would be shown. This includes a proper representation of truly abrupt boundaries.

Currently, there are two groups of methods for continuous representations of floristic composition in remote sensing. The first is still classificatory because the mapped values are proportions of classes. Vegetation is thought to form 'grey zones as transitions between presumed vegetation types' (Mucina 1997, p. 758). These approaches refer to the idea of vegetation as a mixture of class fractions or at least as something that can be numerically described as a mixture of class fractions; graduation is seen, but without abandoning abstract classes (van Niel & Lees 2003). Class fractions as continuous fields can be derived with variable success from linear unmixing (Roberts et al. 1998), spectral angle mapper (de Lange et al. 2004), maximum likelihood classifier, or neural networks (Foody 1996). In these algorithms, fractions are computed by relating reflectance in pixels to reference spectra or training pixels that represent classes. As long as single species are defined as constituents, such approaches could be a good starting-point for an *ex post* derivation of compositional patterns. However, these approaches are limited to species-poor environments (Townsend 2000). When the constituents are already mixed vegetation (Foody 1996), the idea of selected spectra representing classes becomes problematic since there is not always a reason for claiming one stand as being more representative than another (Townsend 2000). Here, the selection of single reference spectra representing typical or extreme stands may introduce an inauspicious reduction of deducible compositional states of vegetation. In some approaches, within-class variation is taken into account by multiple reference spectra, which allow for more flexible class limits (Roberts et al. 1998; de Lange et al. 2004; Dennison et al. 2004). Still, these attempts aim to optimize classes.

The second group of remote sensing methods for continuous representations of vegetation goes without classifying steps. These methods are frequently used for representations of biochemical or structural canopy properties. These studies (e.g. Smith et al. 2002; Oppelt & Mauser 2004) use regression models where reflectances in wavelengths serve to predict canopy chemistry or structure. The main obstacle in transferring these methods to an application in organismic applications is transforming floristic information into a manageable set of metric variables ready for regression. Ordination methods can transform large species-by-plot matrices into few dimensions, where each plot composition corresponds to a specific range of scores, and these scores can be used as response variables in modelling.

Ordination methods provide quantitative access to transitions in plant species composition without forcing any classification. Unlike field mapping, remote sensing has the potential to extrapolate such quantitative, continuous, floristic information from the plot into the surrounding area. Unfortunately, this potential remained largely neglected. There are applications of Canonical Correspondence Analysis in remote sensing (CCA; van de Ven & Weiss 2001; Ohman & Gregory 2002), an algorithm that performs an ordination of the field data and relates it to predictors (in this case, the image information) in one go. As in multiple regression analysis, multicollinearity in predictors gives rise to unstable solutions (McCune & Mefford 1999). This causes limitations in the use of remote sensing data with a high spectral resolution.

For a more flexible selection of methods, ordination and regression analysis can be separated. Thessler et al. (2005) combined Non-Metric Multidimensional Scaling (NMS; Shepard 1962) with an adaptation of the *k*-nearest neighbour (kNN) procedure. Schmidtlein & Sassin (2004) used a Detrended Correspondence Analysis (DCA; Hill & Gauch 1980) in combination with a partial least squares regression (PLS; Wold 1966). In the present study, NMS and PLS are linked to provide a methodical pathway that combines the advantages of a non-parametric ordination with a regression method capable of taking account of the information in numerous correlated spectral bands of the imagery. The latter ability is essential to make full use of the huge amount of valuable detail provided by contemporary imaging spectroscopy. Imaging spectroscopy is based on a spatially contiguous detection of reflectance signatures. The reflectance in each pixel is measured in adjacent, narrow spectral bands. The technique has been evolving in remote sensing since the late 1980s and is also referred to as hyperspectral imaging (van der Meer & de Jong 2001; Aspinall et al. 2002; Ustin et al. 2004).

Material and Methods

At a glance

The floristic field data were collected in ground relevés and arranged in a virtual similarity space using three-dimensional NMS. Three NMS axis scores were assigned to each plot describing the position in the ordination space. The reflectance spectra corresponding to plots were taken from the remotely sensed imagery and used to compute a PLS regression between reflectance values in different wavelengths (predictors) and axes scores (responses). The resulting regression equations were applied to the imagery, resulting in three grey-scale images, one for each ordination axis. Colour scales (red, green, blue) were assigned to the three axes, and the three greyscale images were transformed into a red-green-blue colour map with a specific colour for each position in the ordination space. Similar species compositions resulted in similar colours.

Study site and data

The study took place in the Murnauer Moos area in Bavaria (at 47°39' N, 11°11' E, ca. 790 m a.s.l.). The area comprises a mosaic of raised bogs, poor fens, tall sedge beds, reed swamps, and wet meadows (Kaule 1974). Most of the area was traditionally mown for litter, and parts of it are now subject to land management contracts between the local conservancy authority and farmers; other parts feature different successional stages or near-natural vegetation. From this mosaic, an area of 20.5 ha was selected where a co-occurrence of distinctive and inconspicuous, gradual and abrupt transitions between different compositions of plant species is found.

The floristic field data were collected in 44 plots, each consisting of three subplots. The subplot design (Fig. 1) served for an estimation of plot homogeneity. It is useful to have information on the homogeneity before relating ground plot data to reflectance data extracted from imagery. The distances between the centres of the circular subplots were 5 m, and each subplot had an area of 4 m^2. All subplots were marked with magnetic markers. The plots were arranged in a systematic grid at distances of ca. 80 m and localized at an accuracy of ± 0.3 m using differential GPS.

The vascular plants and dominant mosses and lichens occurring were registered in the summer of 2004, with their percentage cover in subplots. These cover values were averaged at the plot level and log-transformed. In the following analyses, both the log-transformed data and the untransformed data were used, but the transformed data performed slightly better, so this paper refers only to these results. The plot homogeneity was calculated using Bray-Curtis distances between subplots. Three plots that appeared unusually heterogeneous were excluded from further analyses, leaving 41 plots for the next processing steps.

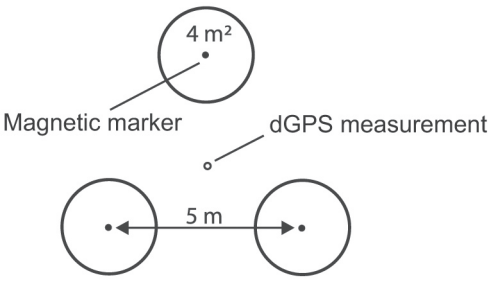

Fig. 1. Plot consisting of three circular subplots. This arrangement allowed for an estimation of plot homogeneity.

The imagery had been collected one year before, using the airborne imager HyMap. This sensor measures reflectance in 126 bands in three wavelength regions between 0.45 and 2.5 mm and with bandwidths of 15-20 nm. Noise-affected bands at 438 nm and 450 nm, 1404-1475 nm, 1795-2009 nm, and 2389-2483 nm were removed, leaving 105 bands for the analysis. Data were acquired on 22 July 2003 at 9:25 (UTC). The imagery was not subjected to atmospheric correction, since spectra for calibration were not taken from independent sources but derived instead from the imagery (Aspinall et al. 2002). All calculations and considerations are based on radiance measured at the sensor. However, for simplicity, we refer to reflectance instead of measured radiance. These data were transformed by log (1/*R*), where *R* was the radiance. This 'pseudo-absorbance' tends to be near-linearly related to the true absorbance of materials (Kumar et al. 2001). For comparison, the untransformed data were also used.

The spatial resolution in the area of investigation was about 6 m × 4 m. Geolocation and image geocorrection were based on a differential GPS, an inertial measurement unit (both on board) and a digital elevation model (DEM). Owing to the almost flat terrain, a course DEM resolution of 50 m was considered sufficient. The final registration to an ortho-image from the land surveying administration (resolution 0.4 m) resulted in a root-mean-squared error (RMSE) of position of 0.66 m.

The reflectance spectra associated with the plots were taken from the imagery. To this end, the spectra of the pixels that hit the corresponding subplot centres were averaged. Before averaging, the homogeneity of reflectance was checked as expressed by the Euclidean distances between the spectra of a plot. One heterogeneous plot was removed, leaving 40 plots for the final analysis.

Ordination

The similarity patterns among plots were derived using a three-dimensional NMS. This ordination method arranges plots so that the rank order of similarities corresponds to the rank order of Euclidean distances in the ordination space. In order to avoid local minima, the best solution out of 9999 iterations was selected. The underlying distance measure for the original data was the Bray-Curtis distance. For ease of interpretation, the NMS data space was rotated by principal component rotation so that the first axis expressed the maximum floristic variation (Clarke 1993). This did not affect the relative position of plots in the ordination. The variance explained by NMS was assessed by computing the coefficients of determination between distances in the ordination space and in the original species data. The ordination and test were implemented using PC-ORD software (v. 4, MjM Software, Gleneden Beach, OR).

Regression

The NMS ordination transformed the complex floristic field data into three metric variables that could be regressed against reflectance data. The algorithm used was the PLS regression (Wold 1966; ter Braak & de Jong 1998). Only one ordination axis was modelled at a time (PLS1 regression). PLS1 is based on the decomposition of the predictors into latent variables that are used afterwards in the regression step. Unlike principal component regression (PCR), the latent variables are calculated in a way that combines a good representation of predictors and response. Like PCR, PLS can handle numerous, correlated predictors – a typical situation in remote sensing applications with numerous wavelength bands. Because of the low number of samples, regression models were validated by full leave-one-out cross-validation. In this procedure, several submodels are computed in which all samples are left out one by one for validation (Efron & Tibshirani 1993). The PLS regression method is affected by overfitting when too many latent variables (PLS principal components) are included in the regression. Overfitting was avoided by repeating the described validation procedure, adding more and more latent variables. The model with the number of latent variables causing the smallest validation errors was selected.

A proper variable selection could enhance the model results. The wavelength bands were selected that gave significant results in Martens' uncertainty test (Martens & Martens 2000; Davies 2001). This test was based on a validation of regression coefficients in the sub-models of the cross-validation. The differences between the coefficients in the sub-models and the coefficients in the full model indicate the stability of a variable. Removal of variables with unstable coefficients reduced the validation errors of the final models. The PLS regression analyses were implemented using Unscrambler software (v. 8.0, Camo, Oslo).

The resulting three regression equations (one for each axis) were applied to the digital aerial imagery, resulting in predictions of NMS axes scores for every pixel in the image or in three greyscale images representing three NMS axes. Even when the NMS ordination space has been rotated to fit the first axis to the largest variation in the data, the interpretation of a single axis often makes little sense with NMS. While an interpretation of single axes would be best when referring to the three separate grey-scale images, the interpretation of relative positions in the ordination space is facilitated by a colour composite of all three images.

Hence, a colour scale was assigned to each axis image, and the three images were combined into a red-green-blue colour composite. Each colour corresponded to a particular species composition, and similar colours represented similar compositions. A coloured map legend representing the NMS ordination space allowed for a direct comparison between colours, the relative positions in the ordination space, and the distributions of diagnostic species. It is also possible to relate a classification to the ordination space and to show this relation in the legend. This does not affect the map itself, so the non-classificatory nature of the map remains untouched. In order to show how a locally accepted classification scheme can be assigned to the legend, plots were classified by phytosociological 'table work' (Mueller-Dombois & Ellenberg 1974) and assigned to syntaxa using Oberdorfer et al. (1993-1998).

In order to estimate the combined errors of NMS and PLS, the validation output from PLS regression was compared with the original, untransformed species data. The variance explained was assessed by computing the coefficients of determination between distances in the predicted Euclidean ordination space and in the original species data (Bray-Curtis distances).

Results

Ordination

In field sampling, between six and 34 plant taxa per plot and a total of 112 taxa were registered. The three-dimensional NMS of the species data reflected 98.1% of the total variation in the original distance matrix. After the principal component rotation, 70.7% of variation was related to the first axis, 18.7% to the second, and 8.7% to the third.

The distributions of characteristic species in the

ordination space reflected the underlying compositional patterns. The largest floristic differences (largely reflected by the first axis of the rotated NMS) were found along an ombrotrophic–minerotrophic gradient between raised bogs (high cover of *Sphagnum magellanicum* and *Eriophorum vaginatum*) and fallow, species-poor wetlands with *Molinia caerulea* and *Phragmites australis*. At the extreme end of the gradient, the latter stands were mixed with nitrophilous species like *Urtica dioica*. The transitional vegetation in the centre of the ordination was characterized by cover of *Trichophorum cespitosum*. In the minerotrophic section of the ordination, there was an additional, detached group of plots, which represented mown mesotrophic wetlands with sparse stands of tall sedges (e.g. *Carex elata*, *C. lasiocarpa*) and representatives of poor calcareous fens (e.g. *Carex lepidocarpa*, *C. davalliana*, and *Campylium stellatum*). This minor gradient between fallow and mown, minerotrophic wetlands was roughly paralleled by the second axis. The third dimension did not describe a particular additional gradient but emphasized some of the variation already observed along the first and second axes.

Table 1. PLS regression models between ordination scores of 40 plots and reflectance (R) or pseudo-absorbance, $\log_{10}(1/R)$. Min, Max = absolute range of the ordination axis values; # PC = number of principal components used in the final regression models; # Bands = number of predictor wavelengths; 1 PC ($RMSE_{val}$), 2 PC ($RMSE_{val}$), etc. = root mean squared errors in cross-validations of models with increasing numbers of principal components included (* = final model); R^2_{cal} = coefficient of determination in model calibration; R^2_{val} = coefficient of determination in cross-validation (all coefficients are significant at the 0.001 level (***)).

	NMS1	NMS2	NMS3
Reflectance	$\log_{10}(1/R)$	$\log_{10}(1/R)$	R
Min	–0.84	–0.87	–0.74
Max	1.27	1.10	0.60
# PC	4	5	4
# Bands	9	20	71
1 PC ($RMSE_{val}$)	0.73	0.20	0.10
2 PC ($RMSE_{val}$)	0.19	0.17	0.08
3 PC ($RMSE_{val}$)	0.09	0.14	0.05
4 PC ($RMSE_{val}$)	0.05*	0.08	0.02*
5 PC ($RMSE_{val}$)	0.05	0.03*	0.02
6 PC ($RMSE_{val}$)	0.05	0.03	0.03
R^2_{cal}	0.94 ***	0.87 ***	0.82 ***
R^2_{val}	0.92 ***	0.82 ***	0.78 ***

Regression

When interpreting NMS, it is advisable to ask for positions in the multi-dimensional data cloud rather than for positions along single axes. However, the separate treatment of the three dimensions allowed the positional information to be projected into the geographic space. The parameters of the PLS regression models are summarized in Table 1. The total variation in the predicted NMS scores (from cross-validation) expressed 78.7% of the original variation in the species-by-plot matrix (percentage of variance in distance matrix). As compared with the variance explained by the original NMS (98.1%) there was a loss of 19.4%, which is due to the modelling process. As for the PLS regression, the R^2 in cross-validation was highest with the first axis ($R^2 = 0.92$, Fig. 2). However, the results for the second and third axis were still very satisfactory ($R^2 = 0.82$, 0.78, Fig. 2). For the first two axes, better model results were observed with pseudo-absorbance; for the third, the untransformed reflectance was used. The lower the predictive power of axes regarding floristic variation, the weaker was the interrelation with the reflectance or pseudo-absorbance. This is also expressed by the linear correlations between reflectance and axis scores (Fig. 3a) and by the number of wavelengths necessary for an optimum prediction. The first axis was best predicted by only nine spectral bands, the second axis by 20, and the third by 71. The typical errors of the models (RMSE) as measured in cross-validation were about one-tenth of an axis range.

Linear correlations between reflectance and axes scores (Fig. 3a) are useful for an interpretation of the general behaviour of reflectance in the ordination space.

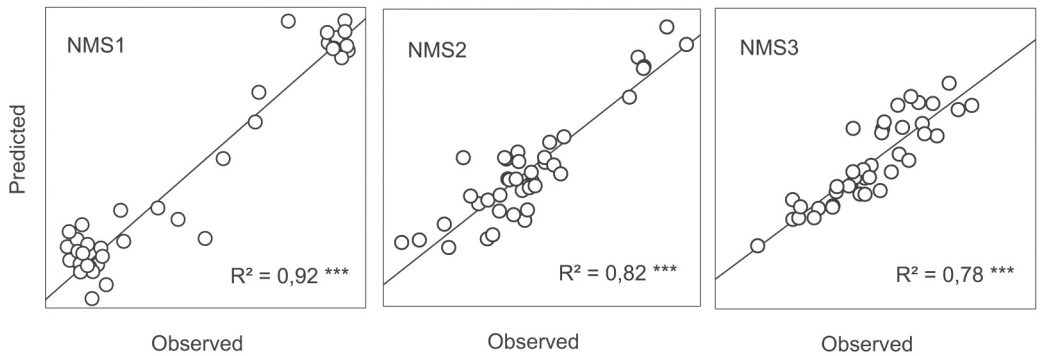

Fig. 2. Cross-validation of the predicted scores of plots on the three NMS axes.

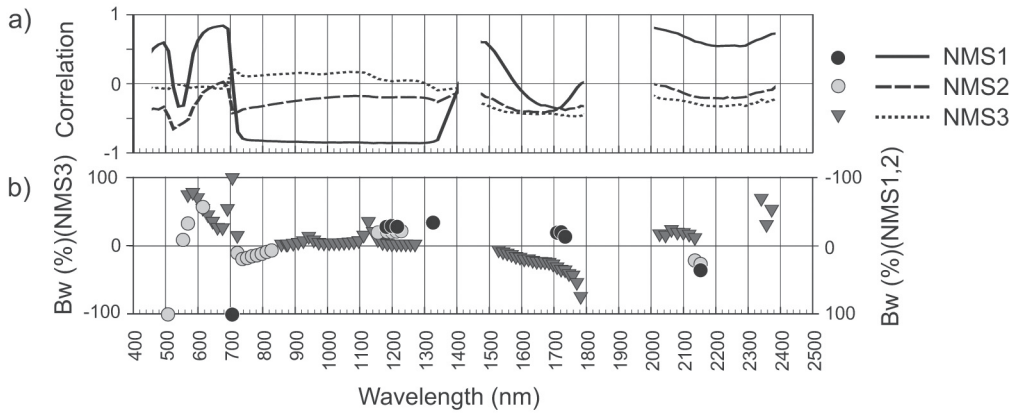

Fig. 3. a. Linear correlations between untransformed reflectance and NMS axes scores of plot samples. **b.** Weighted regression coefficients (Bw) of the wavelengths (original values of reflectance or pseudo-absorbance standardized, the most extreme values set to 100 or –100%). The larger the distance to 0%, the greater was the relative importance of that specific wavelength for the prediction. Unlike the third axis, NMS1 and NMS2 were modelled by pseudo-absorbance. Hence, for better comparability, the *y*-axis was mirrored for NMS1 and NMS2: symbols in the upper part of the graph always indicate a positive influence of the original reflectance on the prediction.

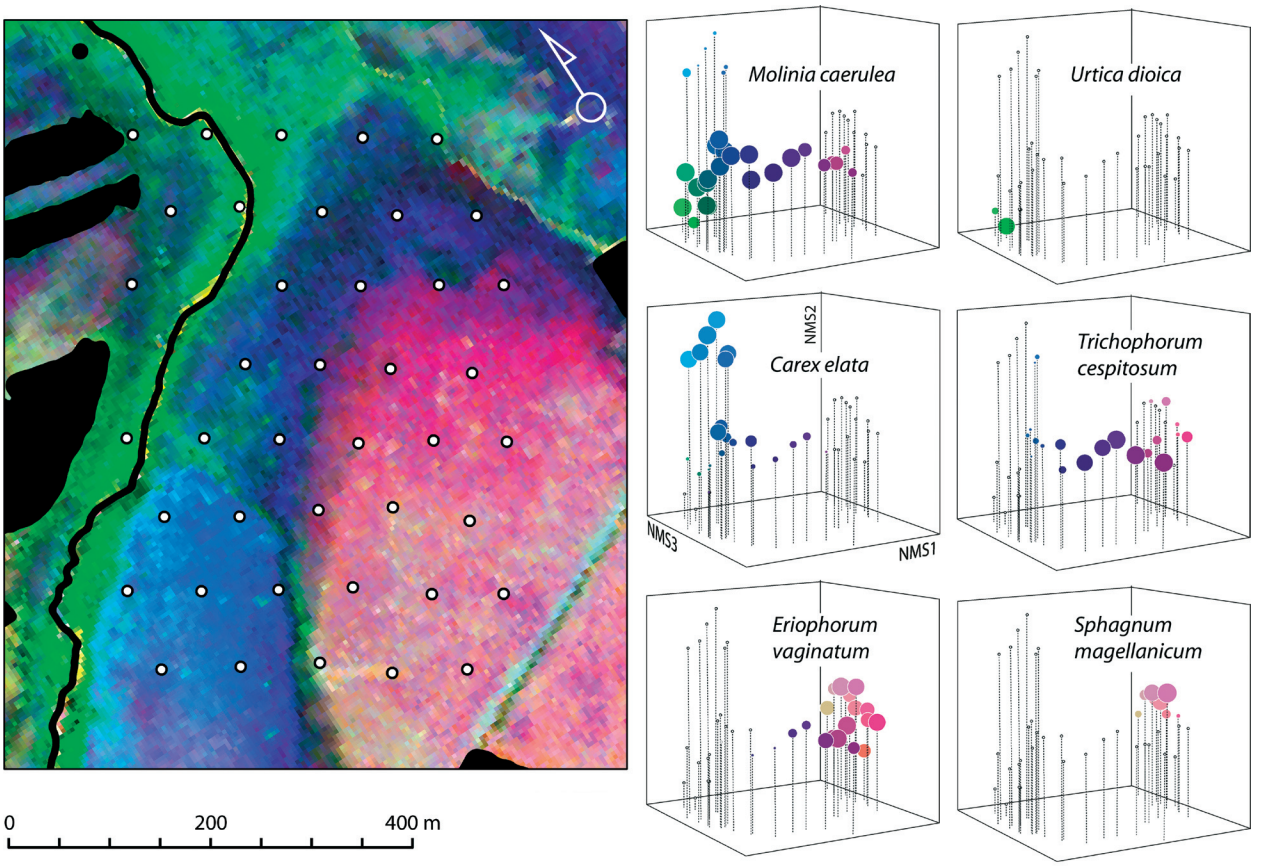

Fig. 4. Spatial representation of the NMS ordination. Similar colours indicate similar plant species composition. Masked (black) sections = forests and water bodies. Circles = sampling plots used in regression. The NMS ordination diagrams display the distributions of diagnostic species in plots (coloured dots). The plots are arranged according to floristic similarities. The sizes of the dots symbolize the log-transformed cover of species (scaled to unit range); small, white dots indicate absence. The diagrams allow for an association between map colours and diagnostic species, the latter representing certain plant communities (see text). For example, *Eriophorum vaginatum* is diagnostic for raised bog vegetation and is limited to one corner of the ordination space. The pink colour of plots with *Eriophorum vaginatum* equals the colour of the map pixels assigned to these plots; hence, pink in the map indicates raised bog vegetation.

They show that the typical spectral features of lush vegetation (Kumar et al. 2001), such as high reflectance in the green light section (around 540 nm) and in the near infrared (720-1400 nm), were negatively related to scores on the first and second axes. Correlations with the third axis were weakly opposed. The behaviour of the first axis is an expression of the association of raised bog vegetation and high scores. Along the second axis, low scores were associated with fallow meadows with more biomass and often a greater nutrient supply, resulting in correlation patterns similar to the first axis. These nutrient-rich stands are also found at the top of axis 3, thus explaining the weakly opposite correlations for this axis. The wavelengths used for the prediction of axes scores were distributed all along the spectrum (Fig. 3b). The weighted regression coefficients show that the wavelengths with the most weight for the prediction were situated in the transition between red and near infrared (at 707 nm: axes 1, 3) and in the green area (at 508 nm: axis 2). Other important predictor wavelengths for axis 3 were registered in the yellow light section (at 585 nm) and in the short-wave infrared (at 1795 and 2325 nm).

Map

The spatially explicit representations of the three NMS axes were combined into a colour map (Fig. 4). The map shows the area of investigation with abrupt and gradual transitions between plant species compositions. The actual occurrence of these transitions could be verified by field observations. The interpretation of the colours is possible by examining the distribution of diagnostic species in the coloured ordination space (Fig. 4, legend). As an alternative, degrees of memberships are shown in Fig. 5 for the examples of the bog associations *Sphagnetum magellanicum* and *Eriophoro-Trichophoretum cespitosi*. *Molinia caerulea* (Fig. 4) represents the fallow, species-poor, mesotrophic wetlands with *Molinia caerulea* and *Phragmites australis*. These stands are shown in dark blue and green. Within these wetlands, the greenest spots represent richer sites along brooks (diagnostic species: *Urtica dioica*). *Carex elata* is a representative of the mown mesotrophic wetlands with sparse stands of tall sedges and plants of poor calcareous fens. This community appears in light blue tones. *Eriophorum vaginatum* is a typical plant of the raised bogs, which are shown in pinks. The wettest parts (with the lightest pink) feature a high dominance of *Sphagnum magellanicum*. Forests and water bodies were masked, since no predictions were intended for these land-cover types.

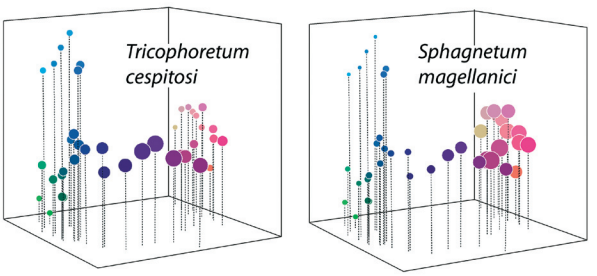

Fig. 5. These NMS Ordination diagrams display an *ex post* affiliation of plots to vegetation classes of a locally accepted classification scheme. In these examples, the size of the dots symbolizes relative plot memberships to classes. Euclidean distances between plots and class centroids were used as simple membership measures.

Discussion

The present paper aims to introduce a non-classificatory mapping method for mixed stands. The goals were to depict compositional variation as continuous fields and to show as much as possible of the local variation in plant species composition. The results from the present study form one such representation of mixed stands. The map is, of course, a simplified picture of the true patterns, but it depicts much of the relationships between individual species compositions. Each pixel has the chance to be unique, and no typifying step took place prior to the interpretation of the resulting pattern.

The high coefficients of determination of the models may partly be caused by the presence of 'long' floristic gradients associated with considerable differences in growth forms and structural characteristics. Another factor affecting the overall validation results is the selection of the validation method. Ideally, the test data would come from a new set of samples, but in the present study, there were not enough data available. In such cases, cross-validation is a standard tool. With very small data sets, as in the present study, the omission of too many samples causes a pessimistic bias (Kohavi 1995). However, the selected method of omitting only one sample at a time preserves the spatial autocorrelation structure of the data. This means that the validation may have caused overly optimistic error estimates compared with a test with spatially independent data (Labovitz 1986). The spatial autocorrelation in the present data is caused by a dense sampling scheme (in other words, distances between plots are short as compared with the distance decay in floristic composition). This scheme was selected in order to thin out plots afterwards and to test the necessary sampling intensity. In our example, models based on half the plots still reached 84-97% of the predictive power of the full models (as expressed by the R^2 in cross-validation). For future investigations, we suggest starting with a small number of relevés and adding relevés until a sufficient

model quality is reached.

There was a gap of one year between the flight and ground investigations, and the year of image-data acquisition was particularly dry. If the spectral information associated with the calibration plots is extracted from the image itself, as in this study, there is no strict necessity to carry out the field work at the same time as image-data acquisition. However, the floristic composition should change as little as possible. Since successional processes in the investigated systems are usually much slower, this is considered a minor problem and the accuracy of the results, as indicated by cross-validation, seems to confirm this assumption. Since no independent reference spectra were used, and no comparison with other images took place, no atmospheric correction was applied (Aspinall et al. 2002). However, in investigations of larger areas and in multitemporal analyses, atmospheric corrections are advisable, even if spectra are collected from the imagery (Aspinall et al. 2002).

The use of spectra that are not collected from the imagery itself is probably not expedient. There are too many degrees of freedom in the relationship between floristic composition and reflectance to put confidence in spectral libraries. Typical examples of factors acting upon the relationship are phenological processes (Dennison & Roberts 2003) or short-term drought stresses (Carter 1993). Field work will never be completely replaced by this kind of investigation, since calibration with samples from the area under investigation will remain a crucial step in deriving such maps.

The spatial resolution of the imagery results in a mixture of different floristic composition at sub-pixel levels. The subplot design took account of this fact in preventing vegetation with considerably differing positions in the ordination space from being related to one single reflectance spectrum. Nevertheless, some inevitable mixture should be expected, especially in the raised bog areas with their fine-scale mosaics consisting of hummocks and hollows. In these areas, recurring ordination scores refer to recurring vegetation mosaics.

In addition to the dimension of pixels, the value of remotely sensed data depends on the spectral resolution. Imaging spectroscopy is the leading technique in this regard (Kumar et al. 2001; van der Meer & de Jong 2001; Ustin et al. 2004). Further improvements could be achieved by multitemporal image data that takes account of phenological differences. Specific temporal patterns may reveal compositional differences that are otherwise difficult to map (Townsend & Walsh 2001).

In some cases, it might be advantageous to skip the a priori ordination and to use individual species models, which can then be aggregated in a post-processing step (van Niel & Lees 2003). However, in a case like ours, with 112 plant taxa, this means considerable effort, and if a representation of general compositional patterns is the goal, as in our case, some type of dimensionality reduction would be necessary anyway.

Earlier studies have already shown that gradients depicted by DCA can be similarly modelled with the aid of imaging spectroscopy (Schmidtlein & Sassin 2004). However, in reading such gradient maps, gradients or axes per se played little role: in the end, map interpretation was based on distances in colour space. Thus, a method that relies on distances in the similarity space was considered as more suited than a method that relies on gradients. NMS has an additional advantage: it provides an optimal representation of the floristic variation with a limited number of dimensions. For the purpose of a colour composite, three dimensions fit the needs. In DCA, parts of the information are lost to minor axes.

Of course, the presented method is not appropriate for all applications of conventional vegetation maps. For example, in conservation, it may be crucial if a place does or does not belong to a certain protected habitat type. In such cases, a combined representation with clear class limits may be appropriate. In some cases, the heterogeneity may be so high that one ordination is not able to represent the variation in a comprehensible number of dimensions (Whittaker 1972).

Strictly field-based methods are not suitable for providing similar spatial information at a comparable level of detail. In the field, mapping compositional patterns as continuous fields is an almost impossible task. The method presented here may be suitable in many situations where pattern recognition is useful, and a representation of stand transitions and heterogeneities is needed. It may also be useful in cases where maps are needed, but no clear types can be identified.

Conclusions

The use of NMS of species data in combination with imaging spectroscopy proved to be an expedient approach. In abandoning pre-defined plant communities during the mapping process, the authors tried to reconcile mapping practice and a reality where no two species assemblages are completely alike and where concrete stands sometimes grade into one another. The method presented here was judged to be satisfactory, since each pixel had the chance to represent an individual species composition without referring to any class membership. If more or less homogeneous areas emerge (as in our case), they are a true expression of community structure and not a matter of threshold decisions as in conventional vegetation maps. Ordination maps provide intuitive access to similarity patterns, which are difficult to discern from categorical maps.

While conventional vegetation maps fit the community concept, the present method conforms to both major paradigms of vegetation ecology. It fits the community concept because it is able to show recurring composition and relatively homogeneous areas. On the other hand, the method satisfies the presumptions of the individualistic or continuum concepts because each pixel has the chance to represent an individual composition. The use of more such maps would allow further studies on the conceptual aspects in vegetation ecology to be carried out, since ordination methods, which are normally used for revealing patterns in the coenospace, are extended into the geographic domain. Palmer & White (1994, p. 281) proposed 'not to map the world onto our mental structures, but to re-wire our mental structures to reflect the world'. We hope that the approach presented here will be useful in this sense.

Acknowledgements. This project was funded by the German Science Foundation (DFG), the Deutsche Bundesstiftung Umwelt (DBU) and the University of Bayreuth. We gratefully acknowledge the cooperation of the German Aerospace Centre (especially A. Müller and S. Holzwarth). Thanks are also due to C. Beierkuhnlein, E. Hertel, F. Wieneke, and the local conservancy authorities. M. Austin and two anonymous referees provided substantial feedback.

References

Aspinall, R.J., Marcus, W.A. & Boardman, J.W. 2002. Considerations in collecting, processing, and analysing high spatial resolution hyperspectral data for environmental investigations. *J. Geogr. Syst*. 4: 15-29.

Austin, M.P. 1985. Continuum concept, ordination methods, and niche theory. *Annu. Rev. Ecol. Syst*. 16: 39-61.

Carter, G.A. 1993. Responses of leaf spectral reflectance to plant stress. *Am. J. Bot. 80*: 239-243.

Clarke, K.R. 1993. Non-parametric multivariate analyses of changes in community structure. *Aust. J. Ecol*. 18: 117-143.

Davies, H.C. 2001. Uncertainty testing in PLS regression. *Spectroscopy Europe* 13: 16-19.

de Lange, R., van Til, M. & Dury, S. 2004. The use of hyperspectral data in coastal zone vegetation monitoring. *EARSeL eProc*. 3: 143-153.

Dennison, P.E. & Roberts, D.A. 2003. The effects of vegetation phenology on endmember selection and species mapping in southern California chaparral. *Remote Sens. Environ*. 87: 295-309.

Dennison, P.E., Halligan, K.Q. & Roberts, D.A. 2004. A comparison of error metrics and constraints for multiple endmember spectral mixture analysis and spectral angle mapper. *Remote Sens. Environ*. 93: 359-367.

Efron, B. & Tibshirani, R.J. 1993. *An introduction to the bootstrap*. Chapman & Hill, New York, NY, US.

Foody, G.M. 1996. Fuzzy modelling of vegetation from remotely sensed imagery. *Ecol. Model*. 85: 3-45.

Fortin, M.J., Olson, R.J., Ferson, S., Iverson, L., Hunsaker, C., Edwards, G., Levine, D., Butera, K. & Klemas, V. 2000. Issues related to the detection of boundaries. *Landscape Ecol*. 15: 453-466.

Gleason, H. 1926. The individualistic concept of the plant association. *Bull. Torrey Bot. Club* 53: 7-26.

Green, D.R. & Hartley, S. 2000. Integrating photointerpretation and GIS for vegetation mapping: Some issues of error. In Millington, A.C. (ed.) *Vegetation mapping*, pp. 103-134. Wiley & Sons, Chichester, UK.

Hill, M.O. & Gauch, H.G. 1980. Detrended correspondence analysis: an improved ordination technique. *Vegetatio* 42: 47-58.

Hobbs, E.R. 1986. Characterizing the boundary between California annual grassland and coastal sage scrub with differential profiles. *Plant Ecol*. 65: 115-126.

Humboldt, A. von 1807. Essai sur la géographie des plantes. *Voyage de Humboldt et Bonpland*, F. Schoell, Paris, FR.

Kaule, G. 1974. *Die Übergangs- und Hochmoore Süddeutschlands und der Vogesen*. Cramer, Lehre, DE.

Kenkel, N.C., Juhász-Nagy, P. & Podani, J. 1989. On sampling procedures in population and community ecology. *Vegetatio*, 83: 195-207.

Kent, M., Gill, W.J., Weaver, R.E. & Armitage, R.P. 1997. Landscape and plant community boundaries in biogeography. *Progr. Phys. Geogr.* 21: 315-353.

Kohavi, R. (ed.) 1995. A study of cross-validation and bootstrap for accuracy estimation and model selection. *International Joint Conference on Arteficial Intelligence (IJCAI)* 2, pp. 1137-1145. Morgan Kaufmann Publishers, San Francisco, CA, US.

Kumar, L., Schmidt, K., Dury, S. & Skidmore, A. 2001. Imaging spectrometry and vegetation science. In: van der Meer, F.D. & de Jong, S.M. (eds.), *Imaging spectroscopy. Basic principles and prospective applications*, pp. 111-155. Kluwer Academic Publishers, Dordrecht, NL.

Labovitz, M.L. 1986. Issues arising from sampling designs and band selection in discriminating ground reference attributes using remotely sensed data. *Photogramm. Engin. Remote Sens*. 52: 201-211.

Martens, H. & Martens, M. 2000. Modified Jack-knife estimation of parameter uncertainty in bilinear modelling by partial least squares regression (PLSR). *Food Quality and Preference* 11: 5-16.

McCune, B. & Mefford, M.J. 1999. *PC-ORD. Multivariate analysis of ecological data, Version 4*. MjM Software Design, Gleneden Beach, OR, US.

Mucina, L. 1997. Classification of vegetation: Past, present and future. *J. Veg. Sci*. 8: 751-760.

Mueller-Dombois, D. & Ellenberg, H. 1974. *Aims and methods of Vegetation Ecology*. John Wiley and Sons, New York, NY, US.

Oberdorfer, E. (ed.) 1993-1998. *Süddeutsche Pflanzengesellschaften*. 5 Vols. Fischer, Jena, DE.

Ohman, J.L. & Gregory, M.J. 2002. Predictive mapping of forest composition and structure with direct gradient analysis and nearest-neighbor imputation in coastal Oregon, U.S.A.

Can. J. For. Res. 32: 725-741.
Oppelt, N. & Mauser, W. 2004. Hyperspectral monitoring of physiological parameters of wheat during a vegetation period using AVIS data. *Int. J. Remote Sens*. 25: 145-159.
Palmer, M.W. & White, P.S. 1994. On the existence of ecological communities. *J. Veg. Sci*. 5: 279-282.
Rennwald, E. 2000. *Verzeichnis und Rote Liste der Pflanzengesellschaften Deutschlands*. Schriftenreihe für Vegetationskunde, 35, Bundesamt für Naturschutz, Bonn-Bad Godesberg, DE.
Roberts, D.A., Gardner, M., Church, R., Ustin, S., Scheer, G. & Green, R.O. 1998. Mapping chaparral in the Santa Monica Mountains using multiple endmember spectral mixture models. *Remote Sens. Environ*. 65: 267-279.
Robinson, G.D., Gross, H.N. & Schott, J.R. 2000. Evaluation of two applications of spectral mixing models to image fusion. *Remote Sens. Environ*. 71: 272-281.
Schmidtlein, S. & Sassin, J. 2004. Mapping of continuous floristic gradients in grasslands using hyperspectral imagery. *Remote Sens. Environ*. 92: 126-138.
Shepard, R.D. 1962. The analysis of proximities: Multidimensional scaling with an unknown distance function. *Psychometrika* 27: 125-140, 219-246.
Smith, A.J.E. 1980. *The moss flora of Britain and Ireland*. Cambridge University Press, Cambridge, UK.
Smith, M.-L., Ollinger, S.V., Martin, M.E., Aber, J.D., Hallett, R.A. & Goodale, C.L. 2002. Direct estimation of aboveground forest productivity through hyperspectral remote sensing of canopy nitrogen. *Ecol. Appl*. 12: 1286-1302.
ter Braak, C.J.F. 1986. Canonical correspondence analysis: a new eigenvector technique for multivariate direct gradient analysis. *Ecology* 67: 1167-1179.
ter Braak, C.J.F. & de Jong, S.M. 1998. The objective function of partial least squares regression. *J. Chemometrics* 12: 41-54.
Thessler, S., Ruokolainen, K., Tuomisto, H. & Tomppo, E. 2005. Mapping gradual landscape-scale floristic changes in Amazonian primary rain forests by combining ordination and remote sensing. *Global Ecol. Biogeogr.* 14: 315-325.
Townsend, P.A. 2000. A quantitative fuzzy approach to assess mapped vegetation classifications for ecological applications. *Remote Sens. Environ*. 72: 253-267.
Townsend, P.A. & Walsh, S.J. 2001. Remote sensing of forested wetlands: application of multitemporal and multispectral satellite imagery to determine plant community composition and structure in southeastern USA. *Plant Ecol*. 157: 129-149A.
Ustin, S., Roberts, D.A., Gamon, J.A., Asner, G.P. & Green, R.O. 2004. Using imaging spectroscopy to study ecosystem processes and properties. *BioSci*. 54: 523-533.
van de Ven, C.M. & Weiss, S.B. 2001. Mapping arid vegetation species distributions in the White Mountains, eastern California, using AVIRIS, topography, and geology. In: *AVIRIS Proceedings: May 20 and 21, 1991*. Jet Propulsion Laboratory, Pasadena, CA, US.
van der Maarel, E. 1976. On the establishment of plant community boundaries. *Ber. Dtsch. Bot. Ges*. 89: 415-443.
van der Meer, F.D. & de Jong, S.M. (eds.) 2001. *Imaging spectrometry. Remote sensing and digital image processing,* Vol. 4. Kluwer Academic Publishers, Dordrecht, NL.
van Niel, K.P. & Lees, B.G. (eds.) 2003. The case for a data model to reconcile geographical and ecological paradigms. *Proc. 4th International Conference on Integrating Geographic Information Systems and Environmental Modeling: Problems, prospectus, and needs for research*. University of Colorado, Boulder, CO, US.
Weaver, J.A. & Albertson, F.W. 1956. *Grasslands of the Great Plains: their nature and use*. Johnsen, Lincoln, NE, US.
Whittaker, R.H. 1972. Convergences of ordination and classification. In: van der Maarel, E. & Tüxen, R. (eds.) *Grundfragen und Methoden der Pflanzensoziologie*, pp. 39-57. W. Junk, The Hague, NL.
Wisskirchen, R. & Haeupler, H. 1998. *Standardliste der Farn- und Blütenpflanzen Deutschlands*. Ulmer, Stuttgart, DE.
Wold, H. 1966. Estimation of principal components and related models by iterative least squares. In: Krishnaiah, P.R. (ed.) *Multivariate analysis*, pp. 391-420. Academic Press, New York, NY, US.
Zonneveld, I.S. 1974. On abstract and concrete boundaries, arranging and classification. In: Tüxen, R. (ed.) *Tatsachen und Probleme der Grenzen in der Vegetation. Bericht über das Internationale Symposion der Internationalen Vereinigung für Vegetationskunde in Rinteln 8.-11. April 1968*, pp. 17-43. Cramer, Lehre, DE.

Received 5 December 2005;
Accepted 21 September 2006;
Co-ordinating Editor: M. Austin.