Nutr Cycl Agroecosyst (2014) 98:155–167
DOI 10.1007/s10705-014-9603-3

ORIGINAL ARTICLE

Using multivariate analysis of soil fertility as a tool for forest fertilization planning

Jesús Fernández-Moya · Alfredo Alvarado · Manuel Morales · Alfonso San Miguel-Ayanz · Miguel Marchamalo-Sacristán

Received: 10 October 2013 / Accepted: 22 January 2014 / Published online: 29 January 2014

Abstract The design of fertilization plans to cover large areas is complex, due to the considerable number of soil samples and soil fertility variables that must be taken into account. Classifying forest stands in groups according to their soil fertility (i.e. in nutrient management areas) can be very helpful to this respect and it is considered to be a first step in what has been called precision forestry. For this paper, we explore the capability of multivariate analyses of topsoil data to be used as tools for evaluating and classifying soil fertility. A case study from a teak (*Tectona grandis* L.f.) plantation in Costa Rica was used to evaluate and illustrate how to use multivariate analysis with these aims. A topsoil (0–20 cm) database with soil test results assembled by Panamerican Woods Ltd. was used. Different multivariate techniques [Principal Component Analysis, Non-metric Multidimensional Scaling (NMDS), Cluster analysis] were performed and compared. Cluster analysis resulted as an appropriate tool for grouping soil samples into soil fertility classes. Therefore, it is considered as a promising tool which would help to design a fertilization program to meet the specific needs of each group of stands with relatively homogeneous soil fertility properties. NMDS is also a suitable complementary tool to graphically explore the similarities within groups and the differences between them. The application of procedures similar to those being reported may help to optimize the design of nutritional and fertilization plans across large forest plantations, by using multivariate analysis to establish fertilization regimes that are appropriate to groups of stands of more homogeneous soil fertility.

Keywords Forest nutrition · Planted forests · Soil fertility · *Tectona grandis* · Site-specific management · Nutrition management areas

J. Fernández-Moya (✉) · A. San Miguel-Ayanz
Dpto. Silvopascicultura, E.T.S.I. Montes, Universidad Politécnica de Madrid (UPM), Ciudad Universitaria s/n, 28040 Madrid, Spain
e-mail: jesusfmoya@gmail.com

J. Fernández-Moya · A. Alvarado
Centro de Investigaciones Agronómicas (CIA), Universidad de Costa Rica (UCR), San José, Costa Rica

M. Morales
Panamerican Woods S.A., P.O. Box 7842-1000, San José, Costa Rica

M. Marchamalo-Sacristán
Dpto. Ingeniería y Morfología del Terreno, E.T.S.I. Caminos, Canales y Puertos, Universidad Politécnica de Madrid (UPM), Ciudad Universitaria s/n, 28040 Madrid, Spain

Introduction

Forest plantation areal extent has globally increased during recent decades, and now it covers 264×10^6 ha, 7 % of global forest area, in response to the growing global demand for timber, pulp, energy and other goods (Evans 2009; FRA 2010). Meanwhile, forest managers have been increasingly concerned about maintaining high productivity rates through several rotations, especially in short-rotation plantations, and the relationship between forest nutrition, soil management and sustainable timber production (e.g., Nambiar 1995; Fox 2000). It has long been recognized that forest growth depends on the ability of soil to maintain a supply of required nutrients. However, soil nutrient availability can be modified by management practices, such as fertilizer use (e.g. Rennie 1955; Miller 1981; Fox 2000). A requirement for fertilization regimes, to compensate for nutrient export through timber extraction, is the long-standing specification as indicated by some authors (e.g. Rennie 1955; Worrel and Hampson 1997). However, as Fölster and Khanna (1997) emphasised, such fertilization provision has been traditionally neglected. Nowadays, in order to enhance forest productivity, sustain site fertility, and avoid soil nutrient depletion, fertilization is utilized for intensively managed forests across the globe (Ballard 1984; Gonçalves et al. 1997).

Assuming that deficient soil nutrients have been identified, fertilization programs should be designed considering the following aspects: (a) what fertilizer to use; (b) when to apply it; (c) how much is needed; (d) how often to apply it; and, (e) by what method to apply it (Ballard 1984; Bertsch 1998). The current situation in Central America is that fertilization programmes for forest plantations in most cases have been designed taking into account general rules and a quick interpretation of soil analyses, based on non-specific critical levels (Bertsch 1998). A single fertilization recommendation is usually applied to large plantations of several square kilometres, without taking into account any soil fertility heterogeneity. An important consequence of the precision agriculture approach was a trend towards heterogeneity of crop fertilizer application, with modification of formula and rate according to changing requirements within individual fields, rather than simply considering each field as a whole (Robert 2002). Such precision farming can be established around (1) site-specific management, e.g., management focused principally upon soil type heterogeneity within each field, assuming their microclimate can be considered homogeneous, or (2) management zones with treatments specified only at a greater scale, across groups of sites, for cases when budgetary or other restrictions limit the scope for a wider range of management treatments. A major barrier for site-specific management is the economic cost of generating a satisfactory soil map (Robert 2002). Analogously, Fox (2000) observed that "site-specific management is the key to sustaining soil quality and long-term site productivity" for intensively-managed forest plantations. The delimitation of 'stands' is one of the basic principles of forest management. A 'stand' is regarded as a homogeneous group of trees growing together on a sufficiently uniform site. Forest management is not as intensive as agriculture can be, and in practice, little consideration is given to establishing stand-specific nutritional plans. However, forest sites are amenable to grouping by similarity in their soil fertility, through which managers could delineate nutritional management areas (groups of stands), and therefore facilitate more efficient and productive management.

In this study, it was evaluated how effectively multivariate statistical analysis can contribute to decision-making when used as a tool for analysing soil fertility databases, to classify the stands of a forest plantation according to their soil fertility, and thereby a specific fertilization program could be designed for each of the defined groups. The intention is to expose a case study that illustrates how these analyses can be performed, using data from a specific forest plantation in Costa Rica. However, the objectives are not to make an interpretation of the results in terms of nutritional status and quantitative fertilization needs of the plantations, evaluate possible growth responses after the fertilization, or elaborate maps of soil fertility of the plantations. The aim is just to explore the capabilities of the multivariate techniques and show the possibility of making groups of the already existing stands according to their soil fertility similarities, in order to be easily used to improve forest fertilization programs.

Materials and methods

Study area, sites and field sampling

Teak (*Tectona grandis* L.f.) has been extensively used for forest plantations in Central America, originally in Costa Rica and Panama (De Camino et al. 2002), and more recently in Guatemala, El Salvador and Nicaragua. Across the region, teak plantations are intensively

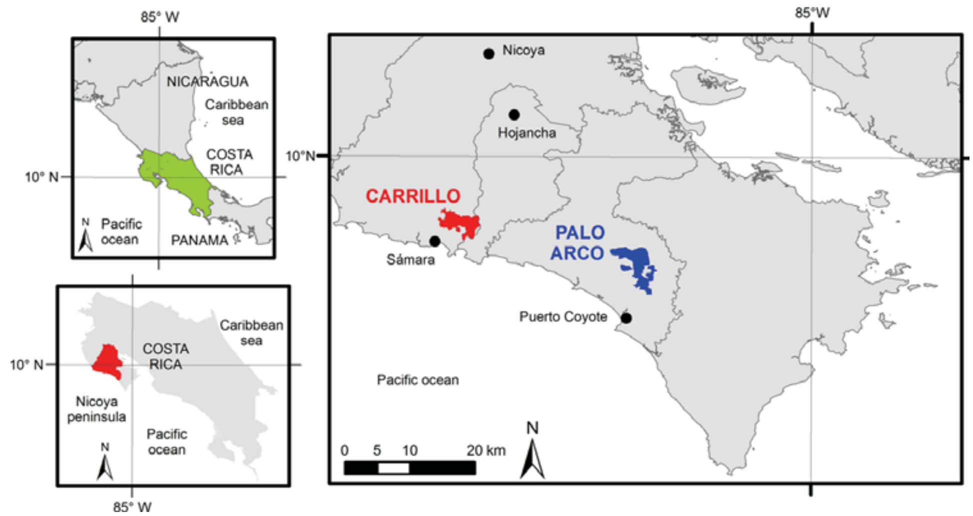

Fig. 1 Location of the two study sites, Carrillo and Palo Arco, on the north Pacific coast of Costa Rica (Nicoya Peninsula), comprising two teak (*Tectona grandis* L.f.) plantations owned by Panamerican Woods Ltd.

managed in rotations of 20–25 years, usually in carefully selected productive sites, with commercial volume expected to be around 10 m^3 ha^{-1} $year^{-1}$ (Pandey and Brown 2000; De Camino et al. 2002). Forest fertilization at establishment has become a common practice for intensively managed forest plantations in Central America, but fertilization at an intermediate or even mature age is not a common practice in the region. Notwithstanding the primary importance of site selection as an issue for teak plantation management, subsequent fertilization is also necessary. Such amelioration can fulfil the high nutrient demand of teak trees, thereby maintaining the high nutrient concentrations they exhibit (Drechsel and Zech 1991; Fernández-Moya et al. 2013), and promoting the productivity and sustainability of production sites (e.g., Prasad et al. 1986; Liang et al. 2005; Zhou et al. 2012).

The case study was located on the North Pacific coast of Costa Rica (Fig. 1), in a company managed teak plantation[1] which is divided in two sites: Carrillo (1,040 ha) and Palo Arco (1,488 ha). The climate of the region is classified as tropical wet forest, following Holdridge's life zones, with a mean annual rainfall of 2,500 mm, and a dry season of 4–6 months. Most common soils are fertile reddish clayey (Table 1), described as Typic Rhodustalfs mixed with Typic Dystrustepts in Carrillo, and Typic Haploustalfs mixed with Vertic Haploustepts in Palo Arco, with small clusters of other soils. Soils are derived from sedimentary limestone and basalt parent material.

The plantations were chosen to be representative of properly-managed teak-planted forests in Central America. In general, management of these plantations consists of continuous forestry management activities: fertilization at establishment; weed control; pruning; and thinning (from approximately 800 trees ha^{-1} at establishment to 150–200 trees ha^{-1} by final felling). The use of clones has become common in recent years. An expected commercial volume of 100–150 m^3 is expected for this kind of plantation, over a rotation of approximately 20 years.

Through the company's routine activity, a database was created for the plantations under study, comprising a total of 195 samples of topsoil (0–20 cm) from across all the different stands, 75 and 129 from the Carrillo and Palo Arco plantations, respectively. Topsoil (0–20 cm) nutrient availability estimates are commonly used for forest fertilization planning in Central America, as fine root absorption is reported to be most active in this soil layer, whether in plantations of teak, or those of other species (Srivastava et al. 1986; Gonçalves et al. 1997; Behling 2009). Soil samples were analysed at the Centro de Investigaciones Agronómicas from the University of Costa Rica (CIA-UCR), to determine the following variables: pH (in water), exchangeable Ca, Mg, K, P, Fe, Cu, Zn, Mn and acidity. pH was determined in water 10:25;

[1] The teak plantations used as a study case in this work are both owned by the Panamerican Woods Ltd. company (hereafter 'PAW').

Table 1 Summary of analysed soil properties for the Panamerican Woods Ltd. teak (*Tectona grandis* L.f.) plantations on the north Pacific coast of Costa Rica

	Carrillo (n = 75)			Palo Arco (n = 120)			General (n = 195)		
	Mean	SE	CV (%)	Mean	SE	CV (%)	Mean	SE	CV (%)
pH	5.8	0.06	9.7	6.0	0.04	6.4	5.9	0.03	7.9
Ca [cmol (+) L^{-1}]	27	1.11	35.5	26.4	0.74	30.6	26.6	0.62	32.5
Mg [cmol (+) L^{-1}]	7.7	0.44	49.2	8.7	0.35	43.7	8.3	0.27	45.9
K [cmol (+) L^{-1}]	0.2	0.01	63.7	0.2	0.02	84.8	0.2	0.01	77.2
Acidity [cmol (+) L^{-1}]	0.2	0.01	42.4	0.1	0.00	32.3	0.2	0.01	43.3
ECEC [cmol (+) L^{-1}][a]	35.1	1.36	33.6	35.4	0.98	30.2	35.3	0.80	31.5
P (mg L^{-1})	1.5*	0.29	168.3	3.2*	0.43	145.5	2.5*	0.29	159.4
Cu (mg L^{-1})	7.0	1.21	150.1	12.6	0.61	52.7	10.4	0.62	83.9
Fe (mg L^{-1})	25.6	2.78	94.2	38.3	3.58	102.3	33.4	2.48	103.7
Mn (mg L^{-1})	41.4	1.95	40.7	25.3	1.79	77.5	31.5	1.44	63.9
Zn (mg L^{-1})	15.7	1.99	109.8	3.1	0.25	86.7	7.9	0.89	157.2
A. S. (%)	0.6	0.05	73.0	0.4	0.02	59.6	0.5	0.03	72.1
Ca S. (%)	76.8	0.76	8.6	74.8	0.55	8	75.6	0.45	8.3
Mg S. (%)	22.0	0.75	29.7	24.3	0.56	25.2	23.4	0.46	27.2
K S. (%)	0.6	0.06	88.7	0.5	0.04	88.4	0.5	0.03	88.5

Means, standard errors (SE) and coefficients of variation (CV) are provided. Values marked with * are lower than adequate reference soil levels (after Bertsch 1998). The 'General' column shows the values when calculated across all the samples for both plantation sites

[a] *ECEC* effective cation exchange capacity

acidity, Al, Ca and Mg in KCl solution 1 M 1:10; P, K, Zn, Fe, Mn and Cu in modified Olsen solution pH 8,5 ($NaHCO_3$ 0.5 N, EDTA 0.01 M, Superfloc 127) 1:10. The effective cation exchange capacity (ECEC) was calculated as the addition of Ca, Mg, K and acidity [ECEC = Ca + Mg + K + acidity]. Ca saturation (Ca S.), Mg saturation (Mg S.), K saturation (K S.) and acidity saturation (A. S.) were calculated as the percentage of ECEC relative to each of the components.

Multivariate statistical methods

Different multivariate analysis methods were employed for simplifying the data, either through graphic representation of similarities between plot points (ordination methods), or through grouping of similar samples into discrete classes (classification techniques) (Oksanen 2010). Both approaches are based on methods to estimate the similarities or dissimilarities between different objects, based on the values of a set of variables measured on each of the objects. Selecting the dissimilarity measure is of primary importance to multivariate analysis (Oksanen 2011). Several distance measures, such as Bray–Curtis, have been considered appropriate in various ecological community studies, but Euclidean distance is considered to be the best-disposed dissimilarity measure for this study, as it fulfils the metric properties, is based upon squared differences, and is dominated by single large differences (Oksanen 2011). Data standardization and transformation are critical in the process of selecting between different methodologies (Kenkel 2006).

Principal component analysis (PCA) and non-metric multidimensional scaling (NMDS) are the most commonly used ordination methods. PCA is based on orthogonal axes, Euclidean space and linear rotation, with an assumption of normally-distributed data, being analogous to simple linear regression (Kenkel 2006). NMDS does not require any underlying assumptions of linearity, and so has emerged as one of the more robust and widely-used techniques, especially in ecology and related disciplines

(Minchin 1987; Kenkel 2006; Oksanen 2011). Conversely, NMDS does present some disadvantages, in particular: (a) NMDS is unable to interpret the relative importance of the ordination axes when summarizing the variation of the data; and (b) NMDS cannot produce a true ordination bi-plot, as variable weights are not determined (Kenkel 2006). While these disadvantages have caused some authors to refrain from advocating adoption of NMDS (Kenkel 2006), in the context of the objectives and data structure of this study, the disadvantages were considered to be of negligible importance. Cluster analysis is a customary classification method which incorporates the calculation of a distance matrix (similar to that used for ordination methods), from which objects can accordingly be classified. Complete linkage (or farthest neighbour) hierarchic clustering was considered the best option for our data and objectives, as this method is based upon maximizing the distance between groups or clusters (Oksanen 2010).

Data analysis

The topsoil database was used to perform different multivariate analyses of the soil test data in order to group the sampled soils according to similarities between the measured properties. The use of different multivariate analysis methods allowed comparing their usefulness for grouping similar soil samples (Table 2). One set of analyses was carried out with the soil test variables centred using their means, along with an alternative set of analyses that instead used the soil test critical levels to centre the variables (Bertsch 1998). In both cases, each variable was standardized using its standard deviation. PCA was performed with the entire dataset, comprising the 195 samples from both plantation sites. NMDS was also performed with the general dataset (this analysis is hereafter referred to as the 'G-NMDS'). Additionally, two NMDS analyses were constructed: (a) one analysis for the 75 samples from the Carrillo plantation ('C-NMDS'); and (b) a second analysis for the 120 samples from Palo Arco ('PA-NMDS'). Five cluster analyses were carried out using the entire dataset (195 samples from both plantations), in order to distinguish: (a) two groups, (b) three groups, (c) four groups, (d) five groups, and (e) six groups. Four additional cluster analyses were computed, two for the Carrillo and two for the Palo Arco plantations, respectively, in order to distinguish two and three groups for each plantation.

The coefficient of variation (hereafter 'CV') was calculated for each variable in each of the constructed soil groups, and the average CV for each group was determined as:

$$\mathrm{CV}_j = average\,\mathrm{CV}_{ij}$$

where $\mathrm{CV}_{i\,j}$ is the CV for each of the study variables (i) for each group (j).

The reduction of the CV for each variable in each of the constructed soil groups relative to the original CV for the 195 samples ($\mathrm{CV}_{i\ general}$) was estimated as:

$$\Delta\mathrm{CV}_{ij} = \mathrm{CV}_{ij}/\mathrm{CV}_{i\,general}$$

The average reduction of the CV of the study variables for each group was calculated as:

$$\Delta\mathrm{CV}_j = average\ \Delta\mathrm{CV}_{ij} \tag{1}$$

The CV calculations allowed an estimate of the homogeneity for each of the groups identified; a comparison to be made against the null hypothesis of 'no-groups'; and to identify which method resulted in the best grouping.

NMDS and cluster analysis were done using the Vegan library in R (R Development Core Team 2011). Euclidean distance was used as the measure of dissimilarity. No rotation was used for the PCA or the NMDS analyses. For the NMDS analyses, the number of k dimensions was set to k = 2. A Shepard diagram 'stress-plot' was constructed as a measure of the goodness of fit for the NMDS analysis (Oksanen 2011).

Results and discussion

No important disparities were found (data not shown) between the results of multivariate analyses obtained using the mean or the critical value as a reference for centring the data (Bertsch 1998). Therefore, we hereafter describe only the results of the former analyses, i.e., from normally-standardized data that used mean and standard deviation as references. Similarly, minor differences were observed between the PCA and the NMDS constructed through the 'general' analyses, using the data from both plantations (data not shown). The NMDS provided the best representation of the differences between soil samples,

Table 2 Summary of the different multivariate analyses performed in the study

Type of analysis	Origin of the data	Number of samples	Name	Number of groups	Reference for centering
PCA	General	195	G-PCA	–	Average
				–	Critical value
NMDS	General	195	G-NMDS	–	Average
				–	Critical value
	Carrillo	75	C-NMDS	–	Average
				–	Critical value
	Palo Arco	120	PA-NMDS	–	Average
				–	Critical value
Cluster	General	195	G-2	2	Average
					Critical value
			G-3	3	Average
					Critical value
			G-4	4	Average
					Critical value
			G-5	5	Average
					Critical value
			G-6	6	Average
					Critical value
	Carrillo	75	C-2	2	Average
					Critical value
			C-3	3	Average
					Critical value
	Palo Arco	120	PA-2	2	Average
					Critical value
			PA-3	3	Average
					Critical value

and the Shepard plot showed a non-metric fit to be better than a linear fit (non-metric fit pseudo $R^2 = 0.978$; linear fit pseudo $R^2 = 0.936$), consolidating our interpretation of NMDS as the better ordination method. Hence, only NMDS analysis was carried out for the Carrillo and Palo Arco plantation data independently.

Palo Arco plantation has generally been considered to exhibit higher soil fertility than Carrillo plantation, to the extent that different nutritional management plans have been designed for the two plantations. However, the average soil data results for the plantations in Carrillo and Palo Arco showed similar values (Table 1), with the possible exception of the P and Zn. Furthermore, the soil samples from Carrillo could not be differentiated from those of Palo Arco, when we investigated the similarities between soil samples using the 'general' NMDS analyses (Fig. 2). This contradiction shows how the traditional methods being used nowadays in many large forest plantations in Central America can be improved using new techniques and how this improvement could result in a more appropriate soil and nutrient management in those ecosystems.

Cluster analyses were used to distinguish the following: two, three, four, five and six groups of soil samples from the entire dataset in general; and, two and three groups from independent analyses of Carrillo and Palo Arco data, respectively (Table 2). Figure 2 represents the 'G-NMDS' general analysis across all 195 samples, plotted in accordance with the plantation field (Carrillo or Palo Arco), while Fig. 3 does it according to the groups defined by cluster analysis. Table 3 summarizes the trend in CV that was evident as more groups were differentiated by cluster analysis: with increasing the number of groups, each

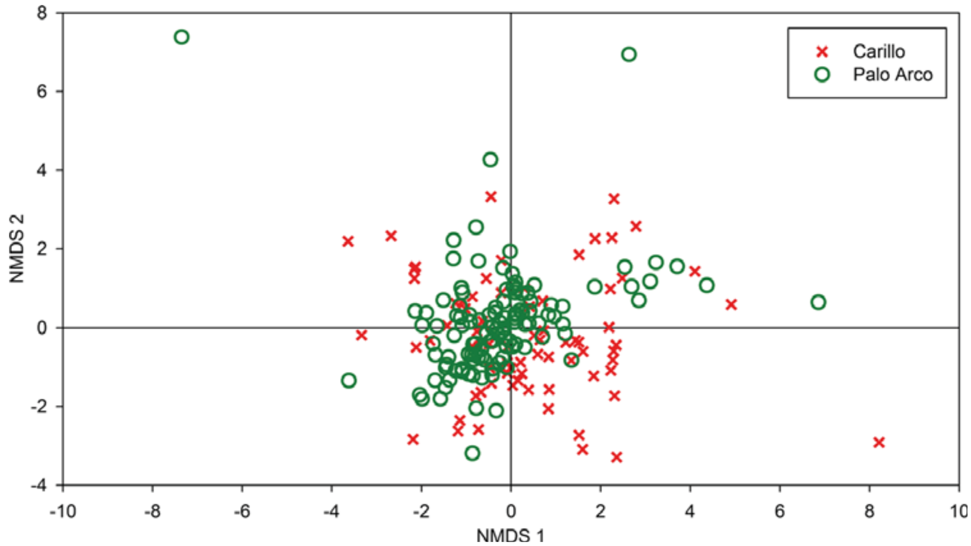

Fig. 2 Non-metric multidimensional scaling for the 195 topsoil samples from the teak (*Tectona grandis* L.f.) plantations owned by Panamerican Woods Ltd., on the north Pacific coast of Costa Rica: Carrillo (*times symbol*) and Palo Arco (*circle*)

group became more homogeneous, and the CV for each variable diminished. As repeated cluster analyses progressively distinguished more groups, each corresponding group could then be diagrammatically isolated within the NMDS (Fig. 3). Hence, a small increase in operational cost would allow an improvement of fertilizer efficiency and it would translate into a higher economic return. However, from a theoretical point of view, at some number of these nutritional groups, the increase (marginal) in benefit should be equal to the increase (marginal) in cost and further increase in the number of groups should result in negative increments of the benefits. In addition to this economical reason, the amount of groups cannot be higher than a reasonable number in order to be practical for the company managers; groups in excess of this number would ultimately contribute to generate disproportionate complexity in this approach to forest management, to the extent that we could anticipate abandonment of such practices. We judged that a maximum of six groups was an appropriate number of soil groups for a 2,500 ha plantation.

Multivariate statistical techniques have been widely applied in soil sciences, notably in the analysis of metal contamination (e.g., Yay et al. 2008), but also in precision agriculture, delineation of site-specific management zones, and soil classification and mapping (e.g., Theocharopoulos et al. 1997; Kalähne et al. 2000; Jaynes et al. 2005; Ortega and Santibáñez 2007; Yan et al. 2007; Fu et al. 2010; Arrouays et al. 2011). Fu et al. (2010) identified clustering as an appropriate analytical technique to delineate soil nutrient management zones, and therefore it provides an effective basis to establish variable-rate fertilization regimes for precision agriculture. However, Fu et al. (2010) also noted that clustering methods were sensitive to the iterative initial value. For the Vegan package in the R-software environment (R Development Core Team 2011), this issue can be addressed by using the metaMDS function, which allows establishing several random starts, and selects from similar solutions with smallest stresses (Oksanen 2011). Ortega and Santibáñez (2007) identified cluster analysis as a better technique for delineating homogeneous management zones, relative to alternative methods. Multivariate techniques have also been applied to precision agriculture in association with geostatistical techniques (e.g., Castrignanò et al. 2005; Morari et al. 2009; Arrouays et al. 2011). However, in the present context of fertilization management for Central American forest plantations, nowadays we do not consider analysis from this perspective to be justified, given the degree of complexity associated with the techniques. Rather, we consider that the proposed strategy, classifying stands into groups with similar soil properties, affords greater scope for organizing the already existing stands into management zones, given that it readily facilitates identification of a limited number of nutritional management groups. As the stands are considered as a homogeneous unit, no further detail taking into account geostatistics, geographical location or spatial analysis is considered necessary at this time. The delineation of intra-field management zones, i.e., zones of uniform

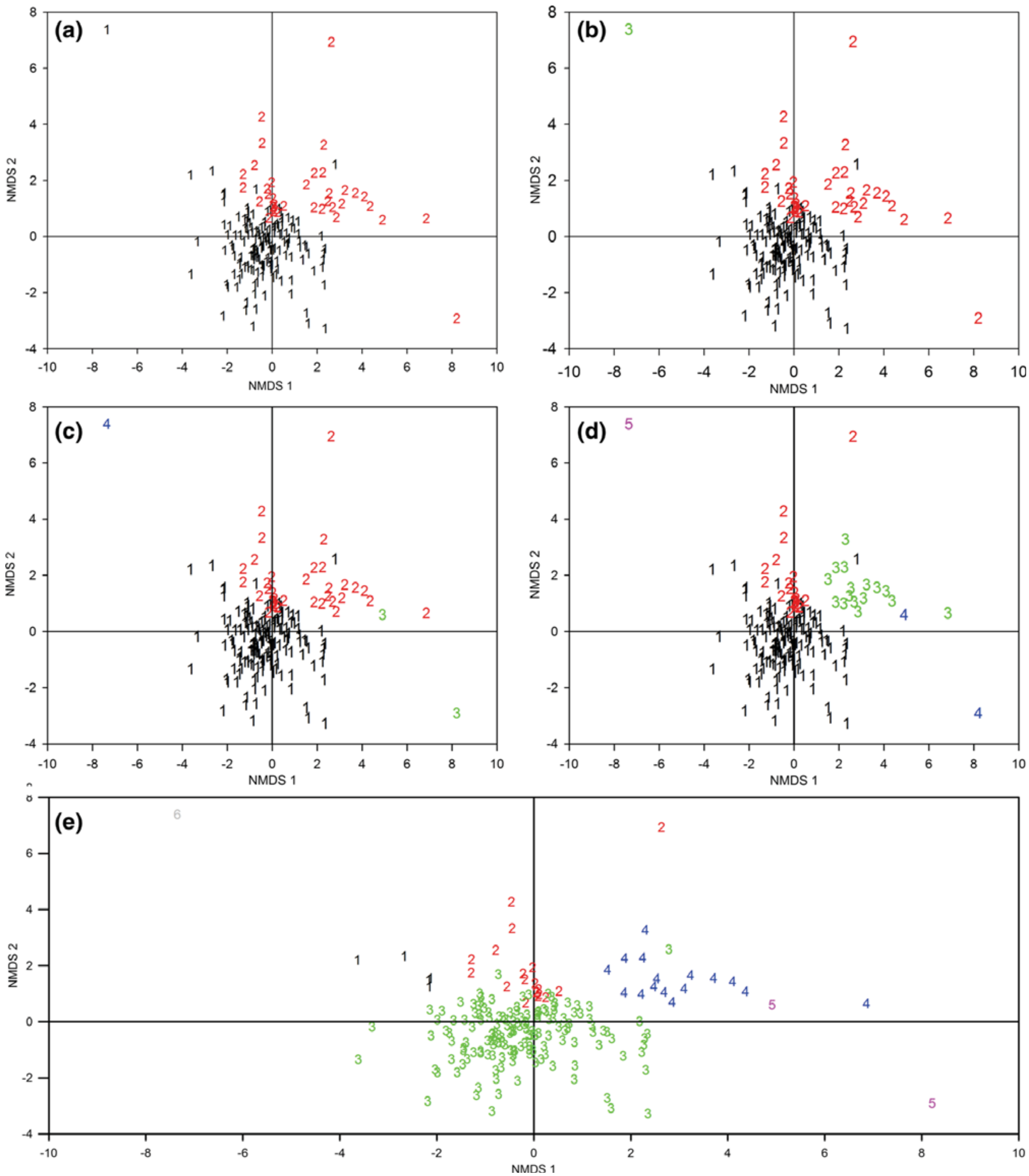

Fig. 3 Graphical representation of **a** two, **b** three, **c** four, **d** five, and **e** six groups, as defined by cluster analysis, based on the spatial scores of the NMDS for the 195 topsoil samples from the teak (*Tectona grandis* L.f.) plantations owned by Panamerican Woods Ltd., on the north Pacific coast of Costa Rica

management, has been assessed as an important initial stage in the implementation of site-specific nutrient management (Ortega and Santibáñez 2007).

In the context of this study, NMDS emerged as a better ordination method than PCA (Figs. 2, 3). However, it was important to initially consider both PCA and

Table 3 Reduction in the average coefficient of variation (CV) for soil fertility variables in each group, distinguished by the cluster analysis treatments, relative to the null hypothesis ('no-grouping', i.e., one single group of data encompassing both Carrillo and Palo Arco together)

	Group	Average CV (%)	Δ average CV (%)	Number of soil samples in the group
Null hypothesis (no-grouping)		66.8	–	195
Grouping by plantation	Carrillo	66.5	−0.4	75
	Palo Arco	58.3	−12.7	120
G-2	Group 1	60.2	−10.5	158
	Group 2	56.2	−9.7	37
G-3	Group 1	55.8	−14.1	157
	Group 2	56.2	−9.7	37
	Group 3	–	–	1
G-4	Group 1	60.2	−10.5	157
	Group 2	52.3	−17.4	35
	Group 3	51.8	−22.5*	2
	Group 4	–	–	1
G-5	Group 1	60.2	−10.5	157
	Group 2	40.2	−40.9**	19
	Group 3	38.2	−36.6**	16
	Group 4	51.8	−22.5*	2
	Group 5	–	–	1
G-6	Group 1	35.7	−45.5**	5
	Group 2	40.2	−40.9**	19
	Group 3	55.0	−17.7	152
	Group 4	38.2	−36.6**	16
	Group 5	51.8	−22.5*	2
	Group 6	–	–	1
C-2	Group 1	−2.8	−2.7	74
	Group 2	–	–	1
C-3	Group 1	−19.2	−16.3	42
	Group 2	−17.9	−21.4*	32
	Group 3	–	–	1
PA-2	Group 1	−14.0	−26.8*	110
	Group 2	−13.3	−29.0*	10
PA-3	Group 1	–	–	1
	Group 2	−17.4	−31.4*	109
	Group 3	−13.3	−29.0*	10

Values marked with * show a relative CV reduction of between 20 and 35 %. Values marked with ** show a relative CV reduction of more than 35 %

NMDS for these analyses, in order to identify the method most appropriate for the data and objectives in question, as the best option can be anticipated to vary on a case-specific basis (Kenkel 2006). When we distinguished six groups from the entire dataset in general, the groups were more homogeneous than those that emerged when independently deriving three groups from the Carrillo data and three from Palo Arco (Table 3). As no notable differences were evident when making comparisons between Carrillo and Palo Arco data (Fig. 2), analysing these data independently was not considered a useful basis for further similar analyses.

Relatively high microelement concentrations are typically required in order to maintain an appropriate nutritional status for trees in teak plantations, and indeed other forest plantations globally (Gonçalves et al. 1997; Lehto et al. 2010; Zhou et al. 2012; Fernández-Moya et al. 2013). However, little attention has been paid to Zn and B in other studies of teak nutrition. Tropical soils are usually characterized as

Table 4 Average and standard error (SE) of soil attributes for the six stand groups defined for the Panamerican Woods teak (*Tectona grandis* L.f.) plantations in Guanacaste (Costa Rica) based on their similarities in the analysed soil fertility attributes

	Group 1 (n = 5)			Group 2 (n = 19)			Group 3 (n = 152)			Group 4 (n = 16)			Group 5 (n = 2)			Group 6 (n = 1)		
	Mean and SE	CV (%)	Δ CV (%)	Mean and SE	CV (%)	Δ CV (%)	Mean and SE	CV (%)	Δ CV (%)	Mean and SE	CV (%)	Δ CV (%)	Mean and SE	CV (%)	Δ CV (%)	Mean and SE	CV (%)	Δ CV (%)
pH	7.0 (0.1)	4	−44	6.1 (0)	3	−61	5.9 (0)	7	−12	5.4 (0.1)*	6	−23	5.1 (0.1)*	2	−70	7.5 (–)	–	–
Ca (cmol + L^{-1})	42.8 (1.72)	9	−71	25.2 (0.93)	16	−51	27.8 (0.59)	26	−19	13.3 (0.76)	23	−29	4.9 (0.56)	16	−52	53.2 (–)	–	–
Mg (cmol + L^{-1})	3.6 (0.58)	36	−22	5.3 (0.38)	31	−32	9.5 (0.28)	36	−21	3.3 (0.16)	19	−59	2.3 (0.59)	36	−23	4.3 (–)	–	–
K (cmol + L^{-1})	0.2 (0.03)	45	−41	0.4 (0.05)	50	−36	0.1 (0.01)*	61	−21	0.2 (0.02)	65	−15	0.2 (<0.01)	0	−100	0.6 (–)	–	–
A (cmol + L^{-1})	0.1 (0.02)	31	−29	0.2 (0.01)	27	−39	0.2 (0.01)	45	4	0.2 (0.02)	43	−2	0.2 (0.09)	61	40	0.1 (–)	–	–
CICE (cmol + L^{-1})	46.7 (1.5)	7	−78	31.0 (1.2)	17	−45	37.6 (0.8)	26	−19	16.9 (0.8)	20	−37	7.6 (1.2)	23	−28	58.2 (–)	–	–
P (mg L^{-1})	2 (0.3)*	35	−78	5 (1.3)*	114	−28	2 (0.2)*	114	−29	2 (0.3)*	54	−66	2 (2)*	140	−12	41 (–)	–	–
Cu (mg L^{-1})	1 (0.2)	41	−51	11 (1.5)	58	−30	9 (0.5)	67	−20	32 (2)	25	−70	8 (6.7)	119	42	8 (–)	–	–
Fe (mg L^{-1})	7 (1.6)*	52	−50	32 (3.6)	49	−53	25 (1.5)	73	−29	121 (13)	43	−58	38 (22.8)	85	−18	22 (–)	–	–
Mn (mg L^{-1})	22 (0.6)	6	−91	24 (3.6)	66	4	30 (1.5)	61	−5	58 (5.5)	38	−41	58 (32.8)	80	26	5 (–)	–	–
Zn (mg L^{-1})	1 (0.7)*	146	−7	4 (0.6)	69	−56	9 (1.2)	160	2	8 (1.8)	91	−42	18 (16.8)	132	−16	3 (–)	–	–
A S. (%)	0.3 (0)	34	−53	0.5 (0)	32	−55	0.4 (0)	52	−28	0.9 (0.1)	51	−29	2.6 (0.8)	41	−43	0.2 (–)	–	–
Ca S. (%)	91.5 (1.2)	3	−60	81.5 (0.7)	4	−51	74.1 (0.4)	7	−13	78.1 (1.2)	6	−32	65.3 (3.2)	7	−15	91.4 (–)	–	–
Mg S. (%)	7.9 (1.3)	38	39	16.9 (0.7)	18	−34	25.1 (0.4)	22	−21	20.2 (1.1)	21	−21	29.9 (3)	14	−50	7.4 (–)	–	–
K S. (%)	0.3 (0.1)	47	−47	1.3 (0.1)	49	−45	0.4 (0)	69	−22	0.9 (0.2)	67	−24	2.2 (0.4)	23	−75	1 (–)	–	–

Values marked with * are lower than adequate reference soil levels (Bertsch 1998). Coefficients of variation (%) for each mean are also provided. Δ CV (%) is the reduction in the coefficient of the variation for the variable when comparing the proposed grouping against the null hypothesis of one single group that includes all the stands in the dataset highly-weathered, and rich in Fe or Mn, but generally deficient in Zn, B, Cu and Mo (Barker and Pilbeam 2006). B is typically deficient in soils on a global scale, and is difficult to evaluate in routine soil fertility analyses (Lehto et al. 2010). There is therefore still a requirement to implement specific evaluations of B and Zn status for forest plantations throughout the tropics. The advantages of multivariate analysis techniques are of particular relevance in this respect. Multivariate analyses can be used to process a large number of variables, and can therefore readily incorporate the range of micronutrients that fertilization planning must take into account.

"The amount of fertilizer to be applied to a given species at a particular site will depend on the level of soil fertility and productivity" (Gonçalves et al. 1997). However, practical management of any fertilization program established on an explicitly stand-specific basis, applying a different fertilization formula and dosage to each stand, is generally regarded as impractical by forest plantation managers. As a contrasting approach, we propose that grouping stands by similarities in soil fertility represents a more practical strategy, in that it facilitates the allocation of sites into a manageable range of soil fertility classes. This process of classification promotes the design of a versatile fertilization regime that is sufficiently proximate to differing soil requirements across all the sites in question.

Analysis of soil data in the context of grouped samples allows us to carry out soil fertility diagnosis with greater precision, and establish a basis for improved nutritional and fertilization planning. This is exemplified by the improvement in precision presented by the results in Tables 3 and 4. In Table 1, which deals with traditional methods for fertilization planning, the soils are presented only as being P-deficient. Hence, if forest managers only take this into consideration, they would design a fertilization programme to solve this deficiency (e.g. application of a phosphorus fertilizer) to the entire plantation area. In comparison, the proposed more detailed grouping analysis (Table 4) indicates that most groups exhibit additional deficiencies. Group 1 shows P, Fe and Zn deficiencies; groups 4 and 5 have low pH values in addition to low values of P; group 3 (representing the majority of the samples) shows low K and P content. Conversely, group 6 indicates exceptional soil that is extraordinarily high in all nutrients. Only group 2 still accords with the results for Table 1, in that it demonstrates a deficiency only in P but with some relatively high pH; thus it is the only group which would have a similar fertilization compared with the initial scenario. On the other hand, a fertilizer formula with P, Fe and Zn would be applied for the stands in group 1; a common N–P–K formula would probably be applied to group 3, while some specific P fertilizer would be needed for groups 4 and 5 with a relatively high basicity index in order to solve the relative low pH values or a common P fertilizer can be applied with previous liming of those stands. Hence, with the traditional methods for fertilization planning the majority of the stands would have a hidden nutrient deficiency that would be lowering the productivity of the plantations, except group 6 (a single stand) that shows high soil fertility with no need to be fertilized. This single stand could have been considered as a statistical outlier; however, it has been considered that a whole stand cannot be removed from the analysis as in the decision-making process done by forest managers; something needs to be done with every stand, even if it is quite different to the others.

Conclusions

Cluster analysis can be applied as an appropriate tool for grouping forest stands according to their soil fertility status and, consequently, for designing fertilizer use programs appropriate to the disparate requirements across differing groups of stands, where each group exhibits relatively homogeneous soil fertility properties. Non Multidimensional Scaling represents a useful complementary tool for graphically exploring the similarities of soil fertility within groups of forest stands, and the differences between those stands.

Multivariate analysis provides techniques to classify soil groups by integrating a large number of soil fertility variables, such as micronutrient concentration values, from across a large number of soil samples. By designing forest fertilization plans for groups of stands, where each group comprises stands with homogeneous soil fertility properties, fertilizer application can therefore be implemented with much greater efficiency and productivity.

Acknowledgments The authors would like to acknowledge the collaboration of Panamerican Woods Ltd. in regard to providing access to the soil database. The authors also thank Paul Robertson, Adam Collins and the personnel of the Natural Resources Laboratory at CIA (UCR) for their help and comments about English language, manuscript format and for their assistance about the contents of this paper. One of the authors, M. Morales, is an employee of a forest plantation company (Panamerican Woods) but none of the authors show any kind of conflict of interests. The present paper was conducted under the MACOSACEN project, financed by PCI-AECID. The authors also thank to two anonymous reviewers for their comments to the present manuscript.

References

Arrouays D, Saby NPA, Thioulouse J, Jolivet C, Boulonne L, Ratié C (2011) Large trends in French topsoil characteristics are revealed by spatially constrained multivariate analysis. Geoderma 161:107–114

Ballard R (1984) Fertilization of plantations. In: Bowen GD, Nambiar EKS (eds) Nutrition of plantation forests. Academic Press, London, pp 327–360

Barker AV, Pilbeam DJ (2006) Handbook of plant nutrition. CRC Press, Boca Raton, FL

Behling M (2009) Nutrição, partição de biomassa e crescimento de povoamentos de Teca em Tangará da Serra-MT. PhD thesis. Universidade Federal de Viçosa, Brasil

Bertsch F (1998) La fertilidad de los suelos y su manejo. Asociación Costarricense de la Ciencia del Suelo, San José

Castrignanò A, Cherubini C, Giasi CI, Castore M, Di Mucci G, Molinari M (2005) Using multivariate geostatistics for describing spatial relationships among some soil properties. In: ISTRO conference, Brno, Czech Republic

De Camino RV, Alfaro MM, Sage LFM (2002) Teak (*Tectona grandis*) in Central America. Forest plantations working paper 19. Forest Resources Development Service, Forest Resources Division, FAO, Rome

Drechsel P, Zech W (1991) Foliar nutrient levels of broad-leaved tropical trees: a tabular review. Plant Soil 131(1):29–46

Evans J (2009) Planted forests: uses. Impacts and sustainability. CAB International, FAO, Rome

Fernández-Moya J, Murillo R, Portuguez E, Fallas JL, Ríos V, Kottman F, Verjans JM, Mata R, Alvarado A (2013) Nutrient concentrations age dynamics of teak (*Tectona grandis* L.f.) plantations in Central America. For Syst 22(1):123–133

Fölster H, Khanna PK (1997) Dynamics of nutrient supply in plantation soils. In: Nambiar EKS, Brown AG (eds) Management of soil, nutrients and water in tropical plantation forests. Australian Centre for International Agricultural Research, Australia, pp 339–379

Fox TR (2000) Sustained productivity in intensively managed forest plantations. For Ecol Manag 138:187–202

FRA (2010) Global forest resources assessment 2010. FAO, Rome

Fu Q, Wang Z, Jiang Q (2010) Delineating soil nutrient management zones based on fuzzy clustering optimized by PSO. Math Comput Model 51:1299–1305

Gonçalves JLM, Barros NF, Nambiar EKS, Novais RF (1997) Soil and stand management for short-rotation plantations. In: Nambiar EKS, Brown AG (eds) Management of soil, nutrients and water in tropical plantation forests. Australian Centre for International Agricultural Research, Australia, pp 379–418

Jaynes DB, Colvin TS, Kaspar TC (2005) Identifying potential soybean management zones from multi-year yield data. Comput Electron Agric 46:309–327

Kalähne R, Amin M, Sobottka J, Sauerbrey R (2000) Methods of multivariate data analysis applied to the investigation of fen soils. Anal Chim Acta 420:205–216

Kenkel NC (2006) On selecting an appropriate multivariate analysis. Can J Plant Sci 86:663–676

Lehto T, Ruuhola T, Dell B (2010) Boron in forest trees and forest ecosystems. For Ecol Manag 260:2053–2069

Liang KN, Pan YF, Liu WM (2005) Multi-factors fertilization trial of *Tectona grandis* in seedling stage. For Res 188(5):535–540

Miller HG (1981) Forest fertilization: some guiding concepts. Forestry 54(2):157–167

Minchin PR (1987) An evaluation of the relative robustness of techniques for ecological ordination. Theory Models Veg Sci 69:89–107

Morari F, Castrignanò A, Pagliarin C (2009) Application of multivariate geostatistics in delineating management zones within a gravelly vineyard using geo-electrical sensors. Comput Electron Agric 68:97–107

Nambiar EKS (1995) Sustained productivity of plantations: science and practice. Bosque 16(1):3–8

Oksanen J (2010) Cluster analysis: tutorial with R. University of Oulu, Oulu

Oksanen J (2011) Multivariate analysis of ecological communities in R: vegan tutorial. University of Oulu, Oulu

Ortega RA, Santibáñez OA (2007) Determination of management zones in corn (*Zea mays* L.) based on soil fertility. Comput Electron Agric 58:49–59

Pandey D, Brown C (2000) Teak: a global overview. Unasylva 51(2):1–15

Prasad R, Sah AKY, Bhandari AS (1986) Fertilizer trial in ten and twenty years old teak plantations in Nadhya Pradesh. J Trop For 2(1):47–52

R Development Core Team (2011) R: a language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria. http://www.r-project.org

Rennie PJ (1955) The uptake of nutrients by mature forest growth. Plant Soil 8(1):49–95

Robert PC (2002) Precision agriculture: a challenge for crop nutrition management. Plant Soil 247:143–149

Srivastava SK, Singh KP, Upadhyay RS (1986) Fine root growth dynamics in teak (*Tectona grandis* Linn F.). Can J Plant Sci 16:1360–1364

Theocharopoulos SP, Petrakis PV, Trikatsoula A (1997) Multivariate analysis of soil grid data as a soil classification and mapping tool: the case study of a homogeneous plain in Vagia, Viotia, Greece. Geoderma 77:63–79

Worrel R, Hampson A (1997) The influence of some forest operations on the sustainable management of forest soils—a review. Forestry 70(1):61–85

Yan L, Zhou S, Feng L, Hong-Yi L (2007) Delineation of site-specific management zones using fuzzy clustering analysis in a coastal saline land. Comput Electron Agric 56:174–186

Yay OD, Alagha O, Tuncel G (2008) Multivariate statistics to investigate metal contamination in surface soil. J Environ Manag 86:581–594

Zhou Z, Liang K, Xu D, Zhang Y, Huang G, Ma H (2012) Effects of calcium, boron and nitrogen fertilization on the growth of teak (*Tectona grandis*) seedlings and chemical property of acidic soil substrate. New For 43:231–243