

## Temporal changes of forest species composition studied by compositional data approach

Milan Kobal <sup>(1)</sup>,  
Damijana Kastelec <sup>(2)</sup>,  
Klemen Eler <sup>(2)</sup>

Many ecological data are compositional and different quantitative techniques have been used to analyze such data, albeit some of them being methodologically wrong. The aim of this contribution is to apply the compositional data approach to forestry data and demonstrate the strengths of this method for percentage or relative data with infrequent zero values. Basal areas of three dominant tree species (*Abies alba*, *Picea abies*, *Fagus sylvatica*) in 119 forest compartments in some of the *Omphalodo-Fagetum* forests in Slovenia in 1954 and 2004 were used to investigate the dynamics of forest species composition over a 50-year period. For the investigated period some additional data about geomorphology and harvesting rates within the compartments were used as explanatory variables of compositional change. The species composition of each forest compartment was subjected to several methods within a compositional analysis framework: descriptive, ternary diagram-based graphical presentations, significance of compositional differences between management classes, significance of perturbation differences (the indicator of forest compositional change) and relation of the compositional change with the explanatory variables by means of compositional linear model. Results indicated that the silver fir was the dominant species in both years, but a clear reduction in silver fir proportion was observed after 50 years. The perturbation differences indicated comparatively large relative increase in the proportion of Norway spruce between 1954 and 2004. Subsequently, the perturbation differences were subjected to isometric log-transformation (*ilr*) and two derived *ilr* coordinates were further used as dependent variables in the multivariate linear model. The initial stand structure correlated well with the perturbation differences. These were also significantly correlations with salvage cutting, a consequence of silver fir decline in the 1954-2004 period. This study demonstrated that the compositional data approach can be successfully used to study forest dynamics yielding some insights into data which are not possible or even not valid using some alternative methods.

**Keywords:** Percentage Data, Data Transformation, Compositional Change, Compositional Linear Model, Forest Dynamics, Vegetation Shift, *Omphalodo-Fagetum*

### Introduction

Many ecological data are compositional (composed of only relative information) or may be regarded as compositional. The forest species composition is expressed as the number of trees, the basal area, or the biomass of each species in a given area. These data may be viewed as absolute values when studying biomass increment, tree density, or diameter structure (Condit et al.

1996, Duchesne & Ouimet 2008, Janik et al. 2008, Firm et al. 2009) and are related to environmental factors or management regime. Often, it is more reasonable to consider such data in a relative manner, *i.e.*, as part of a whole. Vegetation data may be directly assessed or estimated with percentages or percentage classes, such as the Braun-Blanquet, Londo, and Domin cover scale (Dierschke 1994). This ap-

proach is prevalent in community ecology where the structure of biological community is primarily characterized by the relative abundance of species (Van Der Maarel 2005).

Different quantitative techniques have been used to analyze patterns in community species composition and dynamics (Legendre & Legendre 1998, Philippi et al. 1998). Different dimension reduction (ordination) techniques are commonly utilized to analyze multiple species in a community; these approaches include principal component analysis (PCA – Sagar et al. 2003, Heiri et al. 2009), correspondence analysis (CA – Pallardy et al. 1988, Donnegan & Rebertus 1999, Woods 2000a, 2000b), and non-metric multidimensional scaling (NMDS – Philippi et al. 1998, De Grandpre et al. 2000, Woods 2000a, Schuler 2004, Lawrence et al. 2005, Elliott & Swank 2008, Fule et al. 2009). These techniques have been widely applied in community ecology, for which data are often characterized by scarcity

□ (1) University of Ljubljana, Biotechnical faculty, Večna pot 83, 1000 Ljubljana (Slovenia);  
(2) University of Ljubljana, Biotechnical faculty, Jamnikarjeva 101, 1000 Ljubljana (Slovenia)

@ Klemen Eler ([klemen.eler@bf.uni-lj.si](mailto:klemen.eler@bf.uni-lj.si))

Received: Aug 02, 2016 - Accepted: May 10, 2017

**Citation:** Kobal M, Kastelec D, Eler K (2017). Temporal changes of forest species composition studied by compositional data approach. *iForest* 10: 729-738. - doi: [10.3832/ifor2187-010](https://doi.org/10.3832/ifor2187-010) [online 2017-08-01]

Communicated by: Luca Salvati

(many zero values – Legendre & Legendre 1998). Other approaches reduce the community composition to certain values that can be analyzed by univariate statistics, such as various diversity indices (Sagar et al. 2003) and importance values (Chapman et al. 2006, Heiri et al. 2009). Both ordination and univariate indices are the methods of choice when a large number of species are studied and data on many infrequent species are scarce (Condit et al. 1996). The main disadvantage of these predominantly descriptive approaches is the large loss of information that occurs when the community composition is reduced to a few dimensions or one value (Billheimer et al. 2001).

However, zeros are very rare and do not occur in some instances, such as during the investigation of the occurrence of a small number of dominant tree species, plant functional groups, or diameter classes. In these cases, it is desirable to use all of the information available in the data, *i.e.*, to use linear modeling, which can more powerfully explain community composition, dynamics, and species interactions (Billheimer et al. 2001). Past studies utilized linear models to study shifts in dominant species composition, and each species was separately analyzed with repeated ANOVA or paired *t*-tests preceded by an arcsin square root transform of percentage data (Pallardy et al. 1988, Elliott & Swank 2008, Pinto et al. 2008). As will be discussed, the forest species composition is compositional by nature and demands an approach specific to these data. Compositional data are intrinsically multivariate. One or more other components inevitably exist when only one component is known, and thus special multivariate methods are needed (Aitchinson 1986, Filzmoser et al. 2009).

Compositional data contain only relative information and present certain peculiarities that preclude traditional statistical methods. The mathematical foundations of compositional data analysis and terminology are explained in Aitchinson (1986), and Egozcue & Pawlowsky-Glahn (2005). These peculiarities arise from use of the closed data, *i.e.*, data summing to a constant, mainly 1 or 100%. Compositional data are always positive and constrained to a constant sum. The parts do not vary between  $-\infty$  and  $+\infty$  which is generally the assumption when using ordinary statistical methods. Any increase of one component inevitably leads to the decrease in one or more other components. This bias toward negative correlations leads to unusual properties of the variance-covariance matrix (Pearson 1897, Chayes 1960). The distribution of compositional data are often rightly skewed (Filzmoser et al. 2009) and differs considerably from the normal distribution central to classical statistical estimation. Calculation of the arithmetic mean for right-skewed data will result in a biased and excessively high estimate of the central value. Another issue is the subcomposi-

tional incoherence that results from the changed correlations between the raw values of components following the analysis of some parts of a composition (*i.e.*, subcomposition – Aitchinson 1986, Greenacre 2011). The approach presented in this paper successfully addresses all of these issues.

The purpose of this contribution is to introduce the compositional data approach to forest science and demonstrate its applicability in forest dynamics research. The forest species composition of uneven-aged forests in the Snežnik area in southwest Slovenia was analyzed using different methods within compositional data framework. A brief theory of compositional data analysis is first presented, followed by an example of its application to 3-component tree species forest compositions. The graphical representations of the compositional data and descriptive statistics are presented and interpreted. Statistical analyses were applied in a compositional context, including MANOVA and compositional linear modeling.

## Material and methods

### Brief theory of compositional data analysis

#### Basic concepts

Based on the nature of compositional data, which seeks to describe the parts of some whole and provide only relative information between components, three principles can be derived (Aitchinson 1986, Egozcue & Pawlowsky-Glahn 2011):

- The results of any statistical analysis must be the same if all components of the composition are scaled with a positive number (scale invariance). This relationship is ensured by closing the data to 1 or 100% prior to any analysis. Changing the units of the components does not change the results of the compositional data analysis.
- The results of any statistical analysis must be the same if previously unmeasured components are added to the composition (subcompositional coherence); the statistical result for some components may not be dependent on the other reported components.
- Any statistical result for compositions may not be dependent on the sequence of the reported components (permutation invariance).

Composition is presented as a real vector  $\mathbf{x} = [x_1, x_2, \dots, x_D]$  of positive components, in which the sum of positive constant  $\kappa$  is usually taken as 1 or 100% (eqn. 1):

$$\mathbf{x} = C[x_1, x_2, \dots, x_D] = \left[ \frac{x_1 \kappa}{\sum_{i=1}^D x_{ij}}, \frac{x_2 \kappa}{\sum_{i=1}^D x_{ij}}, \dots, \frac{x_D \kappa}{\sum_{i=1}^D x_{ij}} \right] \quad (1)$$

where  $C$  is called the closure operator, *i.e.*, each component of the vector is divided by the sum of all its components and multi-

plied by a constant  $\kappa$  (Aitchinson 1986). The closure operation ensures the principle of scale invariance. When investigating soil texture, for example, the total soil volume is irrelevant and is only used to calculate the relative values of components. Changing the total soil volume cannot change the relative values of the soil texture classes. Two compositions are considered identical when proportional, so the expression of a composition in different units does not change the results of the analysis of such data. For identical compositions, the equation holds:  $\mathbf{x} = \mathbf{y}$ , if  $\mathbf{x} = c \mathbf{y}$ , where  $c$  is a positive real number.

#### Simplex and simplicial geometry

The sample space of compositions is the  $D$ -part simplex ( $S^D$ ), which is a restricted space of a real vector space (eqn. 2):

$$S^D = \{[x_1, \dots, x_D] : x_i > 0, x_1 + \dots + x_D = 1\} \quad (2)$$

where  $i = 1, \dots, D$ . In the case of a 2-dimensional real space, the simplex within this space is a line segment (1-simplex); in the case of a 3-dimensional real space, the simplex is a 2-dimensional triangle, and so on. The closure operation is the projection of a vector with positive components onto the simplex. The only relevant information given by the composition is the relative one, *i.e.*, the ratios between its parts. This relation gives the  $D$ -part composition an intrinsically multivariate nature of  $D-1$  degrees of freedom.

Simplicial geometry does not follow ordinary Euclidean rules observed in real vector space. Nevertheless, concepts such as distance between points, straight lines, orthogonality, norm, and inner vector product exist for a simplex vector space, although these concepts are highly non-intuitive (see Egozcue & Pawlowsky-Glahn 2011 for details). These concepts, however, are needed when we investigate the compositional difference between compositions or a linear compositional process.

The compositional operations perturbation and powering satisfy the properties required to give a simplex  $S^D$  the vector space structure (Aitchinson 1986, Pawlowsky-Glahn & Egozcue 2002).

Perturbation plays the same role in  $S^D$  as the sum of vectors in real space. Given two  $D$ -part compositions  $\mathbf{x} = [x_1, x_2, \dots, x_D]$  and  $\mathbf{y} = [y_1, y_2, \dots, y_D]$ , with  $\mathbf{x}, \mathbf{y} \in S^D$ , their perturbation  $\mathbf{x} \oplus \mathbf{y}$  is defined as follows (eqn. 3):

$$\mathbf{x} \oplus \mathbf{y} = \frac{[x_1 y_1, \dots, x_D y_D]}{(x_1 y_1 + \dots + x_D y_D)} = C[x_1 y_1, \dots, x_D y_D] \quad (3)$$

The neutral element of the simplicial vector space is the identity composition  $\mathbf{1} = C[1, \dots, 1] = [1/D, \dots, 1/D]$  (in the case of a 2-simplex in the barycenter of the triangle), and the inverse  $\mathbf{x}^{-1}$  of a composition  $\mathbf{x} = [x_1, \dots, x_D]$  is given by the composition  $\mathbf{x}^{-1} = C[1/x_1, \dots, 1/x_D]$ .

The perturbation of a composition  $\mathbf{x}$  with

the inverse of a composition  $\mathbf{y}$  is denoted by symbol “ $\ominus$ ” (eqn. 4):

$$\begin{aligned} \mathbf{x} \ominus (\mathbf{y}^{-1}) &= \mathbf{x} \ominus [(-1) \odot \mathbf{y}] = \\ &= \mathbf{x} \odot \mathbf{y} = \\ &= C[x_1/y_1, \dots, x_D/y_D] \end{aligned} \quad (4)$$

This operation is equivalent to subtraction in real vector spaces and is called the perturbation difference. For instance, the inference about mean difference between paired compositions is based on the inference about perturbation differences (Aitchison & Ng 2005).

The powering of composition  $\mathbf{x} = [x_1, x_2, \dots, x_D]$  is an operation by a real number or scalar  $c$ , defined by (eqn. 5):

$$c \odot \mathbf{x} = C[x_1^c, x_2^c, \dots, x_D^c] \quad (5)$$

Powering plays the same role in  $S^D$  as multiplication by scalars in real space.

**Transformation of compositional data and the use of coordinates**

Aitchison (1986) pioneered the analysis of compositional data and introduced two transformations, additive log-ratios (*alr*) and centered log-ratios (*clr*), which transform values constrained by simplex into real space and facilitate the application of classical statistics. Both transformations, however, have important drawbacks: the *alr* transformation is permutation invariant and not isometric (the distances between points are not preserved), and the *clr* transformation lacks subcompositional coherence and yields a singular covariance matrix. Egozcue et al. (2003) introduced an isometric log-ratio transformation that successfully overcomes the above-mentioned problems. This approach reduces compositional operations to ordinary vector operations when representing compositions by their coordinates. Egozcue et al. (2003) showed that a simplex is a part of the real space that has a different geometric structure. An orthonormal base (perpendicular, unit vectors) can be defined in simplex and compositions can be expressed as coordinates with respect to this base. Fortunately, these coordinates have an ordinary Euclidean vector structure, permitting the application of ordinary statistics such as linear modeling and hypothesis testing, among other statistical methods.

The principle of working in coordinates states that each composition in  $D$ -part simplex ( $\mathbf{x} \in S^D$ ) can be expressed as a perturbation linear combination (Mateu-Figueras et al. 2011 – eqn. 6):

$$\mathbf{x} = (x_1^* \odot \mathbf{e}_1) \oplus (x_2^* \odot \mathbf{e}_2) \oplus \dots \oplus (x_{D-1}^* \odot \mathbf{e}_{D-1}) \quad (6)$$

where coefficients  $x_i^*$  ( $i = 1, \dots, D-1$ ) are coordinates with respect to the orthonormal basis  $[\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_{D-1}]$  of the simplex  $S^D$ . The isometric log-ratio transformation (*ilr*) assesses coordinates  $\mathbf{x}^*$  to  $\mathbf{x}$ . Given a basis, the coordinates of  $\mathbf{x}$  are uniquely determined. The vector  $\mathbf{x}^*$  of  $D-1$  coordinates is

the real vector in  $\mathbf{R}^{D-1}$ , and the values of coordinates are not constrained to be positive or less than 1. In a 2-simplex, the composition is expressed with two *ilr* coordinates, while a 3-simplex using three *ilr* coordinates, and so on. Similar to real space, many orthogonal bases exist; no unique approach dictates the definition of the most appropriate base. The main criterion to select an orthonormal basis is the enhancement of interpretability of the representation in coordinates. Some special orthonormal bases are associated with the sequential binary partition of the compositional vector (Egozcue & Pawlowsky-Glahn 2005), in which the parts of compositions are divided into 2 groups of parts (subcompositions) in  $D-1$  sequential steps. The method of partitioning the groups of parts can be related to the nature of the parts. In that case, the generated coordinates, or balances, are more easily interpreted.

The corresponding general equation for isometric log-ratio transformation from simplex to real space  $\mathbf{y} = \text{ilr}(\mathbf{x})$  is defined using orthonormal basis,  $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_{D-1}$  (Egozcue et al. 2003 – eqn. 7):

$$\mathbf{e}_i = C \left[ \exp \left( \sqrt{\frac{1}{i(i+1)}}, \dots, \sqrt{\frac{1}{i(i+1)}}, -\sqrt{\frac{1}{i+1}}, 0, \dots, 0 \right) \right] \quad (7)$$

with  $i = 1, 2, \dots, D-1$ . The components of  $\mathbf{y} = \text{ilr}(\mathbf{x})$  with respect to orthonormal basis of eqn. 7 are (eqn. 8):

$$y_i = \sqrt{\frac{1}{i+1}} \ln \left[ \frac{g(x_1, \dots, x_i)}{x_{i+1}} \right] \quad (8)$$

where  $g(x_1, \dots, x_i) = g(\mathbf{x}) = [x_1 x_2 \dots x_i]^{1/i}$  is the geometric mean of the parts of composition  $\mathbf{x}$  with  $D$  parts, and  $i = 1, 2, \dots, D-1$ .

**Descriptive statistics of compositional data**

Standard descriptive statistics, such as the arithmetic mean and the variance of the individual components, do not fit with Aitchison geometry as measures of central tendency and dispersion. The closed sample geometric mean is a measure of the central tendency for compositional data. For a data set  $n$  compositions, all with  $D$  parts, the central tendency is called the centre and is defined as follows (eqn. 9):

$$\mathbf{g} = C[g_1, \dots, g_D] = \frac{g_1 \dots g_D}{g_1 + \dots + g_D} \quad (9)$$

with  $g_i$  ( $i = 1, \dots, D$ ) as the geometric mean of all the parts of a composition.

The variability of a set of compositions can be expressed using various measures. One can either use metric (or total) variance based on simplicial distances between a composition and the geometric mean of a set of compositions (Egozcue & Pawlowsky-Glahn 2011). Another option is to use either variation matrix or variance matrix of centered log-ratios (Aitchison 1986) both of which contain information on codependence between compositional

parts but lack simpler one-value representation.

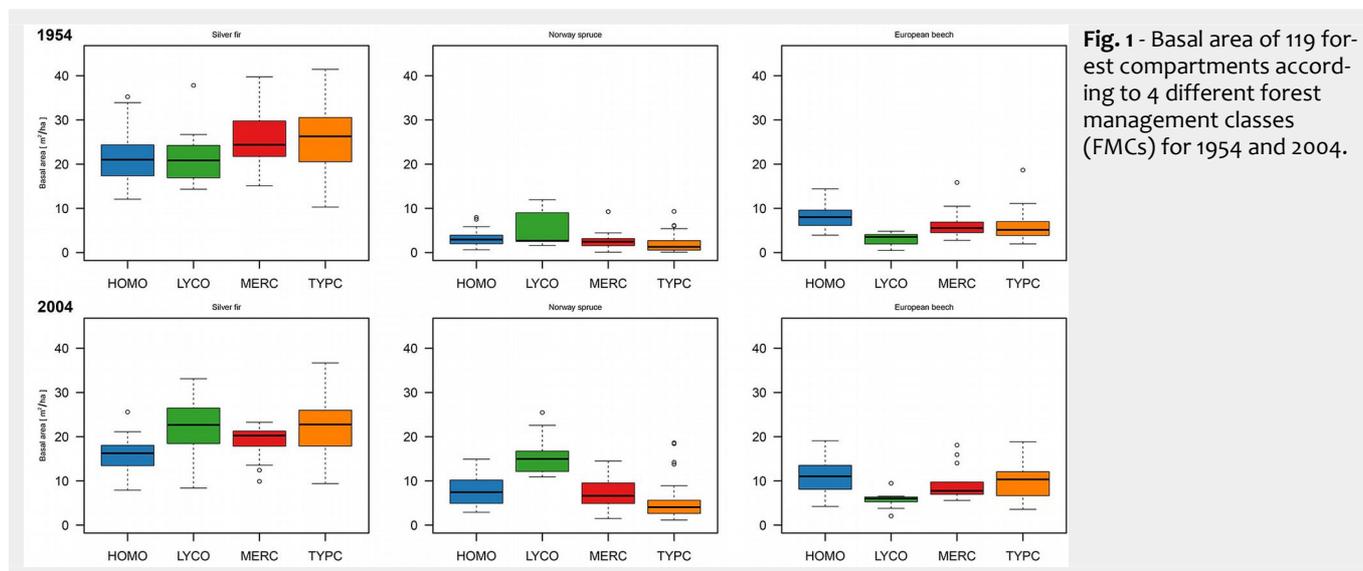
**Compositional linear modeling**

Compositions can evolve dependence on one or more external co-variables such as time, space, or different environmental variables. The explaining variable can be continuous, discrete, real, or even compositional data. The strategy of linear modeling of compositional processes in the simplex can be replaced with linear modeling of compositional coordinates in real space (Pawlowsky-Glahn 2003). The use of log-ratio transformed compositions (e.g., *ilr* and *alr*) instead of raw values can make the linear combinations meaningful with respect to the geometry of the simplex (Pawlowsky-Glahn & Egozcue 2002, Tolosana-Delgado & Van Den Boogart 2011). Aitchison (1986) argues that the conditional distribution is the normal distribution of the simplex. The multivariate nature of the data set is preserved when the applied methods are linear (Tolosana-Delgado & Van Den Boogart 2011).

*Application of the approach*

**Study area**

The study site is situated in the predominantly forested Natura 2000 Snežnik area in southwest Slovenia (Lon. 14° 26' E, Lat. 45° 35' N). The area is located in the Dinaric Mountains, which range from Slovenia to Montenegro. The forest management unit Leskova dolina (2,392 ha) was chosen as the research area. The altitude ranges from 610 m to 1796 m a.s.l. The topography of the area is very diverse, with abundant sinkholes typical of high Karst geology with limestone and dolomite as the parent material. Due to the high variability in topography, different pedogenetic associations developed primarily litosols, rendzic leptosols, cambisols, and luvisols. The soil depth varies between 10 and 300 cm, depending on the micro-topographic position, and precipitation is evenly distributed throughout the year, with a mean annual precipitation of 2150 mm. The average mean temperature is 6.5 °C, and late spring and early autumn frosts are common. The prevalent plant community in the study area is the Dinaric silver fir – European beech forest (*Omphalodo-Fagetum* Mar. 93). The main tree species are silver fir (*Abies alba* Mill.), Norway spruce (*Picea abies* Karst.), and European beech (*Fagus sylvatica* L.). Sycamore (*Acer pseudoplatanus* L.) and elm (*Ulmus glabra* Huds.) are also present but in a much lower percentage. Most of the stands are managed with a selective cuts (single-tree or group) or shelterwood system, leading to considerable within-stand variation in tree age and tree species composition. The close-to-nature forest management prevents forest clear cuts or intensive forest shifts, making the Snežnik forests very suitable for the study of forest community structure and



**Fig. 1** - Basal area of 119 forest compartments according to 4 different forest management classes (FMCs) for 1954 and 2004.

dynamics. Tree ingrowth largely results from natural processes rather than tree planting.

**Data sources**

The analysis utilized the basal area, expressed as  $m^2 ha^{-1}$ , which represents the area of a given section of land that is occupied by the cross-section of silver fir, Norway spruce and European beech trunks measured at the diameter at breast height (1.3 m above ground) per forest compartment in 1954 and 2004. In the 1890s, the forests in the Snežnik area were divided into forest compartments, and each has relatively homogenous site characteristics and stand type. The shapes of the compartments have not changed throughout the last century. Forest compartments with similar site and stand characteristics (parent material, soil, vegetation community, stand structure, etc.) determine forest management system and are thus grouped into forest management classes (FMCs). These classes were named after the prevalent vegetation units in this area of Dinaric silver fir and European beech forests: *Omphalodo-Fagetum typicum* (TYPC), *Omphalodo-Fagetum homogynetosum* (HOMO),

*Omphalodo-Fagetum mercurialeetosum* (MERC), and *Omphalodo-Fagetum lycopodiotosum* (LYCO – FMP 2004).

In each compartment, all trees with a diameter greater than 10 cm were measured (full calipering) at breast height in 1954, and the trees were classified into 5-cm DBH classes for each tree species (FMP 1954). The basal area of each tree species was calculated from the DBH frequency distribution using the means of DBH classes as class representatives and dividing the values by the compartment size. Personnel limitation necessitated a different, indirect approach of basal area determination in 2004. In each forest compartment, several homogenous forest stands were delineated with orthophotographies at 1:5000 scale (FMP 2004). In each forest stand, several temporary sampling plots (minimum 7 per stand) were established. The basal area of each tree species in each plot was estimated with Bitterlich’s angle-count method with basal area factor  $k = 2$ . Sample plot data were averaged for each forest stand. To obtain information on tree species composition in each forest compartment, the basal area of certain tree species within a forest stand belonging to

the specific compartment was weighted according to the proportion of the forest stand in the specific forest compartment. The validity of 2004 indirect basal area estimation per forest compartment was additionally tested using permanent sampling plots of forest inventory ( $n = 471$  plots). No significant deviation between indirect basal area estimation and forest inventory data were ascertained. Fig. 1 presents the basal area data for both observed periods among FMCs. In addition, explanatory forest structure variables and terrain topography were defined (Tab. 1) for each compartment.

**Analysis of sample data**

We utilized several consecutive steps to perform the compositional analysis of this example data set. First, some descriptive statistics, such as means and variation, were shown and coupled with general graphic data representation. A ternary diagram was most suitable for our 3-part compositions. In a ternary diagram, each point represents a 3-part composition in a forest compartment. The vertices of the ternary diagram correspond to the three components or, in this study, species. The components are proportional to the lengths of the perpendicular segments from the vertex to the opposite side of the triangle.

Second, we separately assessed the relative differences in the mean species compositions between the 4 FMCs for each year (1954 and 2004) by classical MANOVA of the *ilr* transformed data. The first *ilr* coordinate (*ilr*<sub>1</sub>) represents the balance between silver fir and Norway spruce, and the second *ilr* (*ilr*<sub>2</sub>) represents the balance between conifers and broadleaves. The MANOVA assumptions were verified. The Box test was used to test the equality of the covariance matrices and univariate normality of the individual *ilr* coordinates to ensure multivariate normality. The assumptions are only marginally accepted because of the outliers. A univariate ANOVA was performed to assess the significance of in-

**Tab. 1** - Explanatory variables and data sources of forest structure and terrain topography specified for each forest compartment. (FMP 1954, 1964, 1974, 1994, 2004).

Data source	Explanatory variable per forest compartment	Units
Calculated from full calipering from 1954 data	Median silver fir diameter at DBH	cm
	Median Norway spruce diameter at DBH	cm
	Median European beech diameter at DBH	cm
	Species composition of trees with diameter < 30 cm	-
Records of harvested trees in 1954-2004	Harvesting rates	$m^3$
	Salvage logging relative to total harvesting rates	%
	Broadleaf logging relative to total harvesting rates	%
	Conifers logging relative to total harvesting rates	%
Digital elevation model 12.5 × 12.5 m	Mean altitude	m
	Mean terrain curvature	-
	Mean slope	degrees

dividual *ilr* coordinates. The Waller-Duncan K-ratio post-hoc test was used for multiple comparisons between FMC means. Because of the marginally accepted assumptions for MANOVA, we additionally utilized the Kruskal-Wallis non-parametric analysis of variance and *post-hoc* test for medians of the *ilr* coordinates between FMCs, and we obtained the same results.

Third, the compositional change was evaluated with a calculation of perturbation differences (eqn. 4). In the case of no change in tree species composition, the perturbation difference should match the identity composition (composition with equal portions of each part). In our research, the perturbation differences were tested against the identity  $I = (1/3, 1/3, 1/3)$ .

Finally, to obtain a deeper explanation of the compositional change, we related the compositional change to the explanatory variables by regression of the perturbation-differences against these variables. We modeled the perturbation differences against FMC; tree species composition of thin trees (DBH < 30 cm) in 1954; median DBH of silver fir, Norway spruce, and European beech; and percentage of salvage cut in 1954-2004 period. The dependent variables in the multivariate linear models included both *ilr* coordinates of perturbation differences that measure forest compositional changes between 1954 and 2004. Prior to the analysis, the tree species composition of thin trees (DBH < 30 cm) in 1954 was transformed by isometric log-ratio transformation (*ilr*).

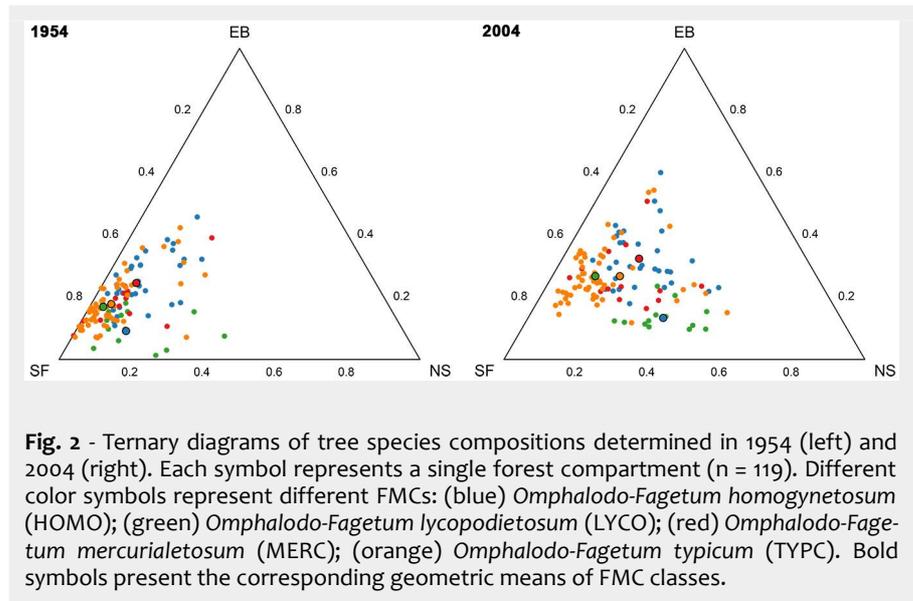
The data were analyzed in the R environment (R Core Team 2012) using the package “compositions” (Van Den Boogart & Tolosana-Delgado 2006, 2008).

## Results

### Descriptive statistics

The subcomposition of the three dominant tree species included in this analysis represents 97.6% of the total basal area. The tree species abundance in each compartment was expressed as part of the total basal area of all trees. The  $D = 3$ -part composition describes the relative abundance of three tree species, silver fir, Norway spruce, and European beech, in 119 forest compartments. Our data did not include zero values, indicating that no tree species was absent in any forest compartment.

Fig. 2 presents the forest compositions in two ternary diagrams. The centers (closed geometric means) of the compositions for each FMC are presented in the ternary diagrams and in Tab. 2. Silver fir is generally the dominant species in both years, but a clear reduction of the silver fir portion is observed after 50 years. There is an increase in Norway spruce and European beech. The total relative variability expressed by the total variance (Tab. 2) decreased from 1954 to 2004 in all FMCs (Tab. 2).



**Fig. 2** - Ternary diagrams of tree species compositions determined in 1954 (left) and 2004 (right). Each symbol represents a single forest compartment (n = 119). Different color symbols represent different FMCs: (blue) *Omphalodo-Fagetum homogynetosum* (HOMO); (green) *Omphalodo-Fagetum lycopodietosum* (LYCO); (red) *Omphalodo-Fagetum mercurialetosum* (MERC); (orange) *Omphalodo-Fagetum typicum* (TYPC). Bold symbols present the corresponding geometric means of FMC classes.

### Comparing species compositions among FMCs

We applied a classical MANOVA test to the *ilr*-transformed data to assess the relative differences in mean species compositions between the 4 FMCs for each year (1954 and 2004). The differences in species composition between FMCs were statistically significant ( $p_{1954} < 0.001$ ,  $p_{2004} < 0.001$ ) in both investigated periods. The results indicate that the ratio between silver fir and Norway spruce in 1954 is significantly higher in TYPC and MERC compared to HOMO and LYCO. The second *ilr* coordinate represents the ratio between the geometric mean of the subcomposition of silver fir and Norway spruce (conifers) and European beech (broadleaves). This ratio in LYCO is significantly different from that in other FMCs in 1954 and 2004. TYPC in 2004 displays the significantly highest ratio between silver fir and Norway spruce, but the ratio is smaller than in 1954. In short, LYCO is distinctly different from other FMCs be-

cause of a lower proportion of European beech.

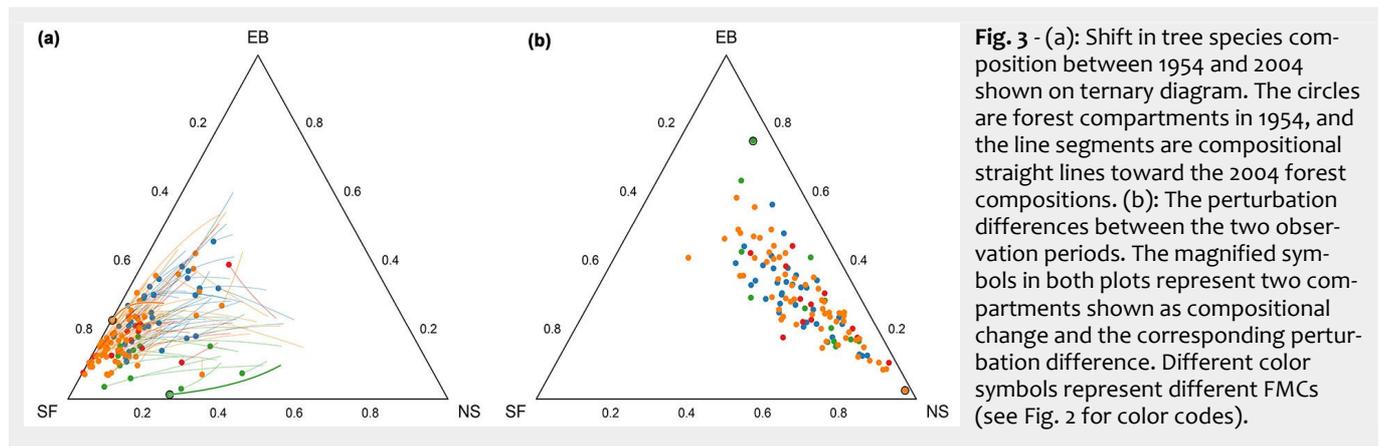
### Temporal changes in tree species composition

The perturbation differences for all forest compartments were calculated to analyze the difference (distance) between forest compositions in 1954 and 2004 (Fig. 3). The perturbation differences are concentrated in the right part of the ternary graph, with the mean perturbation (silver fir = 0.14, Norway spruce = 0.57, European beech = 0.29) indicating the comparatively large relative change in Norway spruce. This result is caused by the low proportion of Norway spruce in the majority of the forest compartments in 1954 and by the increase of the abundance of spruce in the study period.

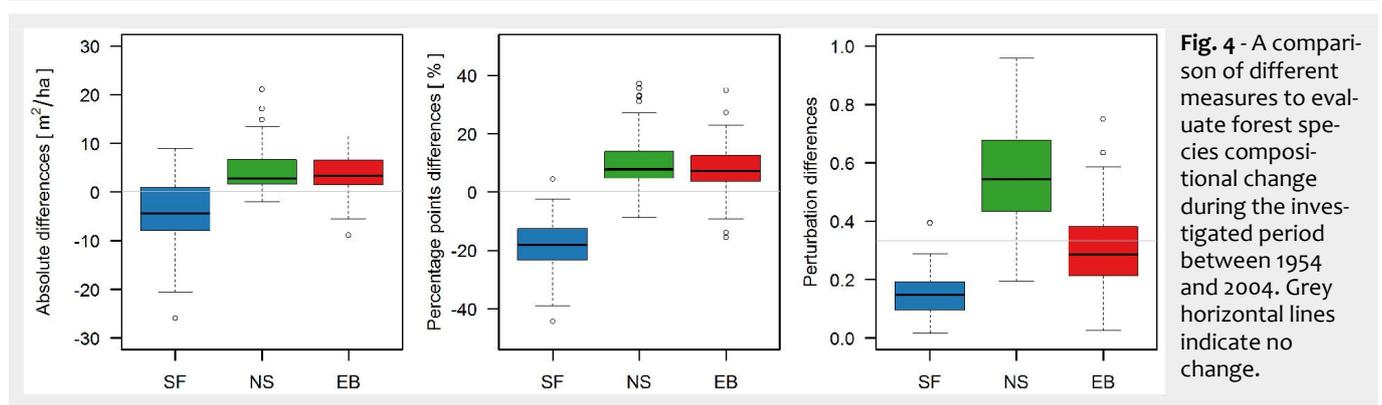
The significant compositional change is quite evident in Fig. 3. However, we tested the significance with the Hotelling T2 test, which is a multivariate extension of the

**Tab. 2** - Mean percentages per tree species (SF: silver fir; NS: Norway spruce; EB: European beech), total variance of the dataset, and significance of differences between four FMCs in 1954 and 2004. Significant differences between FMCs are described for each *ilr*-coordinate separately using letters; FMCs with same letters are not significantly different at 0.05 confidence level for a given *ilr*-coordinate.

Year	Parameter	Species / <i>irl</i>	HOMO	LYCO	MERC	TYPC
-	No of compartments	-	33	13	13	60
1954	Mean portion of	Silver fir (SF)	0.66	0.77	0.77	0.79
		Norway spruce (NS)	0.09	0.14	0.05	0.04
		European beech (EB)	0.25	0.09	0.18	0.17
	Total variance	-	0.43	0.98	1.30	1.04
Homogenous groups	<i>ilr1</i> (SF/NS)	a	a	b	b	
	<i>ilr2</i> [(SF*NS)/EB]	a	b	a	a	
2004	Mean portion of	Silver fir (SF)	0.46	0.49	0.54	0.61
		Norway spruce (NS)	0.21	0.37	0.19	0.12
		European beech (EB)	0.32	0.13	0.27	0.27
	Total variance	-	0.30	0.14	0.49	0.46
Homogenous groups	<i>ilr1</i> (SF/NS)	b	a	b	c	
	<i>ilr2</i> [(SF*NS)/EB]	a	b	a	a	



**Fig. 3** - (a): Shift in tree species composition between 1954 and 2004 shown on ternary diagram. The circles are forest compartments in 1954, and the line segments are compositional straight lines toward the 2004 forest compositions. (b): The perturbation differences between the two observation periods. The magnified symbols in both plots represent two compartments shown as compositional change and the corresponding perturbation difference. Different color symbols represent different FMCs (see Fig. 2 for color codes).



**Fig. 4** - A comparison of different measures to evaluate forest species compositional change during the investigated period between 1954 and 2004. Grey horizontal lines indicate no change.

paired *t*-test. The Hotteling T2 test indicated that the compositional change (ratios between species percentages) is highly significant ( $F\text{-stat} = 244.08$ ,  $p = 0.000$ ).

We also compared change in 1954-2004 period of each of the dominant three species using three different measures, namely change in absolute terms (in basal area per hectare), in percentage points and in perturbation differences. Fig. 4 shows that the magnitude of change for each species is much different for perturbation difference than for the other two measures of change. Change in absolute terms and in percentage points was the largest for silver fir (decrease), whereas perturbation difference was the largest for Norway spruce and smallest for silver fir.

#### Linear modeling of compositional change

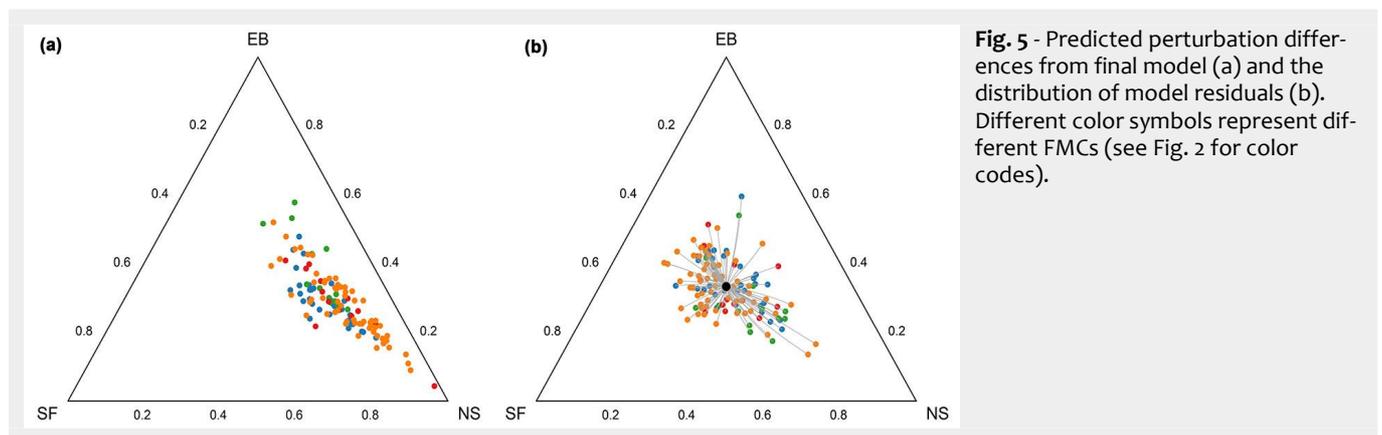
The initial stand structure was well corre-

lated with the perturbation differences. The site characteristics, expressed as FMC, seem to play an important role in the forest species composition dynamics. In both cases, the first and second *ilr* coordinates are significantly correlated with salvage cuts, as a consequence of silver fir decline. The compositional coefficient of determination ( $R^2$ ) is 0.56, and residuals are uniformly distributed around the barycenter of the ternary diagram (Fig. 5).

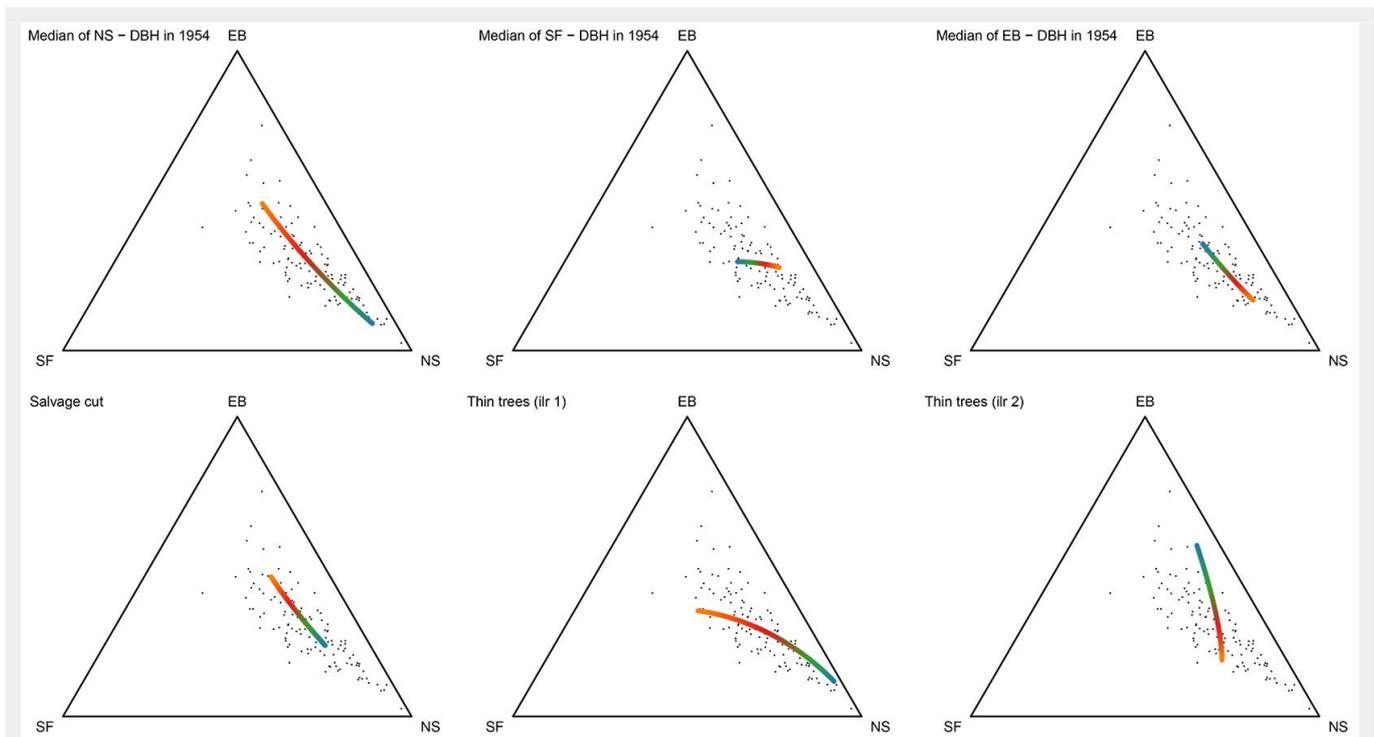
The effect of unique variables on perturbation differences is presented in Fig. 6 and Tab. 3. The changes in the share of Norway spruce are highest in the forest compartments with the lowest median DBH of Norway spruce in 1954, while an increase in the European beech portion is observed in compartments with larger DBH of Norway spruce trees in 1954. In these compartments, there is no ingrowth

of young silver fir trees, bigger Norway spruces trees were cut, and the share of European beech increased (triangle 1 in Fig. 6). The salvage cut created a similar effect. Greater incidences of salvage cut increase the proportion of European beech.

Compartments with a small silver fir DBH do not change greatly (a colored line begins near the center of the ternary diagram on Fig. 6), but compartments with a higher silver fir DBH value tend to have an increasing share of Norway spruce. The proportion of Norway spruce increases in compartments with a high mean value of European beech trees. This result is most likely due to the selective cutting of the larger European beech trees and the more efficient regeneration and ingrowth of Norway spruce compared to silver fir. The first *ilr* coordinate of tree species composition of thinner trees ( $DBH < 30$  cm) in 1954 rep-



**Fig. 5** - Predicted perturbation differences from final model (a) and the distribution of model residuals (b). Different color symbols represent different FMCs (see Fig. 2 for color codes).



**Fig. 6** - The component effect for each co-variable in a linear model of perturbation differences in the forest composition between 2004 and 1954. The small points indicate perturbation differences. The color ramp represents the component effect (orange and blue colors represent high and low values of the explanatory variable, respectively).

**Tab. 3** - Estimates of the parameters, residual standard error, and adjusted coefficients of determination for modeling perturbation differences.

Coordinate	Coefficients	Estimate	Std. Error	t value	p value	Res. Std. Error	Adj. R <sup>2</sup>
ilr 1	(Intercept)	-0.692	0.434	-1.592	0.114	0.358	0.588
	TT ilr 1	-0.488	0.051	-9.531	0.000		
	TT ilr 2	-0.078	0.064	-1.223	0.224		
	NS median	-0.039	0.009	-4.642	0.000		
	SF median	0.038	0.011	3.553	0.001		
	EB median	0.056	0.017	3.377	0.001		
	Salvage cut	-0.004	0.001	-2.861	0.005		
ilr 2	(Intercept)	0.069	0.379	0.182	0.856	0.312	0.437
	TT ilr 1	0.034	0.045	0.754	0.452		
	TT ilr 2	-0.354	0.056	-6.336	0.000		
	NS median	0.039	0.007	5.236	0.000		
	SF median	0.008	0.009	0.905	0.367		
	EB median	-0.063	0.015	-4.324	0.000		
	Salvage cut	0.004	0.001	3.742	0.000		

resents the ratio between Norway spruce and silver fir and can be interpreted as follows: in forest compartments in which the Norway spruce/silver fir ratio is high, the perturbation differences are minor (the line ends near the center of the ternary diagram), but a low Norway spruce/silver fir ratio leads to a high relative increase in Norway spruce.

### Discussion

We have outlined the fundamentals of compositional data analysis and have presented some examples of its application in the study of forest structure and forest dynamics. Aitchinson (1986) noted that the ratios, not the differences, between parts provide information when using closed data. Our research utilized a 3-component

forest species composition. Prior to the application of traditional methods, the isometric log-ratio transform (*ilr*) of the data was utilized to determine the differences between uncorrelated samples (differences in tree species composition between FMC), paired samples (temporal changes in tree species composition), and linear modeling of time changes expressed as the perturbation difference of compositions dependent on covariables that describe the forest condition. The *ilr* transformation preserves the distances between compositions when mapping the data from the Aitchinson geometry to the Euclidean geometry, in which the calculations and graphical representations are normally calculated (Egozcue et al. 2003). Within the Aitchinson geometry, other methods such as clus-

ter analysis or principal component analysis may be applied to the untransformed closed data (Van Den Boogart & Tolosana-Delgado 2006).

Compositional data analysis presents two common challenges: (i) characterizing a change in composition; and (ii) testing hypotheses concerning the nature of the change. In general, a lack of change in a value indicates that the expected difference between two paired samples is zero. For compositions, the perturbation difference is the corresponding measure of change. In our 3-species Sneznik forest composition, we showed that the forest composition change over the last 50 years is highly significant. The interpretation of perturbation differences as compositions is somewhat complicated when investigating

the bidirectional dynamics of forests, in which trees are simultaneously eliminated (felled or dying trees) and growing (recruitment of new trees). The perturbation difference characterizes the net change in species composition. The forest species composition in the forest compartment 10a in 1954 is  $\mathbf{x}_{1954} = [0.77, 0.16, 0.07]$  and  $\mathbf{x}_{2004} = [0.48, 0.24, 0.29]$  in 2004, where the first, second, and third parts are silver fir, Norway spruce, and European beech, respectively. The compositional change expressed as the perturbation difference is  $\mathbf{x}_{\text{diff}} = C[0.48/0.77, 0.24/0.16, 0.29/0.07] = [0.10, 0.24, 0.66]$ , indicating a 6.6-fold larger change in European beech than silver fir and a 2.4-fold larger change in Norway spruce than in silver fir. Compositional data analysis results require careful interpretation. The perturbations in studies of temporal dynamics do not single-handedly indicate anything about the absolute change. Like any other composition, perturbation only possesses relative information, and only the ratios between parts are important and interpretable. Additional information external to compositional data (e.g., forest cut inventory or ingrowth estimation) are needed to reach conclusions about absolute change. Utilization of linear modeling necessitates the interpretation of the regression coefficients within a simplex vector space.

The species composition of the investigated forests displayed relative increases in Norway spruce and European beech and decreases in silver fir. To explain this change some historical facts should be presented. Mature forests of European beech and silver fir prevailed in the Dinaric parts of Slovenia and beyond in the middle of 19<sup>th</sup> century, and young trees below the canopy mainly consisted of shade-tolerant silver fir, which was capable of effective regeneration due to low ungulate densities. Adult beech trees were intensively harvested for firewood and charcoal production until the middle of the 20<sup>th</sup> century, while silver fir became a dominant tree species, desired for its timber and suitability for selection management (Gašperšič 1967, Bončina et al. 2003). The abundance of silver fir faced the turning point in the last decades of 20<sup>th</sup> century with substantial silver fir dieback. Recent studies have offered some possible explanations for this event (Tikvič et al. 2008, Nagel et al. 2015). Those studies suggest that silver fir decline occurred due to interplay of different abiotic effects (air pollution with sulphur and nitrogen compounds, acid rain leading to nutrient deficiencies and imbalances and drought effects). Besides, natural regeneration and recruitment of silver fir was largely hindered due to increasingly high densities of ungulates, browsing young silver fir trees (Diaci et al. 2010). The decrease of silver fir might also be at least partly explained by the observed natural alternation between silver fir- and European beech-dominated phases. During silver fir

dieback Norway spruce was promoted as its substitute.

Using linear modeling we showed which direct and indirect variables were substantially related to forest compositional change. The perturbation differences identified in this study demonstrate the advantage of this approach for ecological studies. Fig. 6 shows that the relative change is highest for Norway spruce, indicating that the method is sensitive to changes in rare, minor tree species, which are usually difficult to be assessed (Woods 2000a).

In general, in the case of forest species composition, the total sum of forest basal area ( $\text{m}^2 \text{ha}^{-1}$ ) is not irrelevant, since the total wood biomass or tree coverage of a site tells a lot about forest productivity and structure (Skovsgaard & Vanclay 2007). Our analysis of compositional data considered that the total of parts of composition is irrelevant, since relative amounts of tree species are of larger importance (Billheimer et al. 2001) in forest community studies. When dealing with the basal area per hectare, we can think of it being the composition of  $n$ -parts ( $n$  represents number of all tree species) plus the remaining part which is the area, not covered by tree trunks. All the parts sum to a constant (a forest hectare), which indicates that the data are closed according to physical laws (Shurtz 2003). The subcompositional coherence principle of compositional data analysis indicates that the results of any subcomposition analysis (e.g., the three common forest species in our study) are the same in a relative context.

This paper does not consider null values, which are the principal problem that can arise in compositional data analysis of ecological data (Martin-Fernandez et al. 2003). The ratios cannot be computed in this situation. Null values are a main characteristic of ecological data, particularly species composition data (Van Der Maarel 2005). To overcome the null value problem, plant community dynamics research traditionally utilizes multivariate methods that are not sensitive to missing data (e.g., correspondence analysis or CA). These essentially descriptive methods, however, lack the predictive power of linear models. Subcompositional incoherence is a key pitfall when utilizing CA instead of the compositional data approach. This problem occurs when the addition of parts (species) into the analysis subsequently changes the result of the analysis. Much of the specialized literature discusses different methods of null replacement (Martin-Fernandez et al. 2003, 2011), but a common assumption argues that a sensitivity analysis should follow value imputation. Greenacre (2011) proposed a method to actually evaluate the subcompositional incoherence and decide whether this factor is small enough to be neglected.

## Conclusions

This paper presents a natural and mathe-

matically clean framework to deal with closed data, particularly when the data are subjected to physical constraints, such as mass or area conservation (Shurtz 2003). Multiple reports have cautioned against analysis of raw percentage data or closed data (Buccianti et al. 2006) with the possibility of wrong conclusions when analyzing compositions directly. Even traditional transformations (e.g., log or arcsin-sqrt) and separate component analyses are now regarded as outdated. Basically, the same statistical methods, such as descriptive statistics, linear modeling, spatial analysis, time series analysis or classification methods can also be used on compositional data, but only if such data is properly (*ilr*) transformed. However, some difficulties still persist with the approach used in our study; the *ilr* coordinates, gained after the transformation, are somewhat difficult to interpret and the problem of zeros and missing values is still not fully resolved.

We showed that certain aspects of forest dynamics such as species composition can be successfully evaluated using compositional data approach. For the investigated forests of SW Slovenia we showed significant compositional change in a 50-year period caused by the interplay between environmental factors and forest regime with their differential effects on three dominant tree species. In absolute terms, the amount of silver fir changed the most; decrease could be attributed to fir dieback and poor regeneration/recruitment due to ungulates. In relative terms, expressed in perturbation differences, largest change was showed for Norway spruce with its increase being predominantly the effect of promotion through forest management measures.

## Acknowledgements

We thank the anonymous reviewers for their valuable comments which greatly contributed to the quality of the paper. We acknowledge the financial support from the Slovenian Research Agency (Young Researcher's programme [MK] and research core funding No. P4-0059 and No. P4-0085).

## References

- Aitchinson J (1986). The statistical analysis of compositional data. Monographs on Statistics and Applied Probability, Chapman and Hall Ltd, London, UK, pp. 416.
- Aitchison J, Ng KW (2005). The role of perturbation in compositional data analysis. *Statistical Modelling* 5: 173-185. - doi: [10.1191/1471082X05st0910a](https://doi.org/10.1191/1471082X05st0910a)
- Billheimer D, Guttorp P, Fagan WF (2001). Statistical interpretation of species composition. *Journal of the American Statistical Association* 96: 1205-1214. - doi: [10.1198/016214501753381850](https://doi.org/10.1198/016214501753381850)
- Bončina A, Gasperšič F, Diaci J (2003). Long-term changes in tree species composition in the Dinaric mountain forests of Slovenia. *Forestry Chronicle* 79: 227-232. - doi: [10.5558/tfc79227-2](https://doi.org/10.5558/tfc79227-2)
- Buccianti A, Mateu-Figueras G, Pawlowsky-Glahn

- V (2006). Compositional data analysis in the geoscience: from theory to practice. The Geological Society, London, UK, pp. 212. [online] URL: <http://books.google.com/books?id=pJ5Xf n-90KcC>
- Chapman RA, Heitzman E, Shelton MG (2006). Long-term changes in forest structure and species composition of an upland oak forest in Arkansas. *Forest Ecology and Management* 236: 85-92. - doi: [10.1016/j.foreco.2006.08.341](https://doi.org/10.1016/j.foreco.2006.08.341)
- Chayes F (1960). On correlation between variables of constant sum. *Journal of Geophysical Research* 65: 4185-4193 - doi: [10.1029/JZ065i012p04185](https://doi.org/10.1029/JZ065i012p04185)
- Condit R, Hubbell SP, Foster RB (1996). Assessing the response of plant functional types to climatic change in tropical forests. *Journal of Vegetation Science* 7: 405-416. - doi: [10.2307/3236284](https://doi.org/10.2307/3236284)
- De Grandpre L, Morissette J, Gauthier S (2000). Long-term post-fire changes in the northeastern boreal forest of Quebec. *Journal of Vegetation Science* 11: 791-800. - doi: [10.2307/3236549](https://doi.org/10.2307/3236549)
- Diaci J, Rozenberger D, Bončina A (2010). Stand dynamics of Dinaric old-growth forest in Slovenia: are indirect human influences relevant? *Plant Biosystems* 144: 194-201. - doi: [10.1080/11263500903560785](https://doi.org/10.1080/11263500903560785)
- Dierschke H (1994). *Pflanzensoziologie [Phytosociology]*. Ulmer, Stuttgart, Germany, pp. 683. [in German]
- Donnegan JA, Rebertus AJ (1999). Rates and mechanisms of subalpine forest succession along an environmental gradient. *Ecology* 80: 1370-1384. - doi: [10.1890/0012-9658\(1999\)080\[1370:RAMOSF\]2.0.CO;2](https://doi.org/10.1890/0012-9658(1999)080[1370:RAMOSF]2.0.CO;2)
- Duchesne L, Ouimet R (2008). Population dynamics of tree species in southern Quebec, Canada: 1970-2005. *Forest Ecology and Management* 255: 3001-3012. - doi: [10.1016/j.foreco.2008.02.008](https://doi.org/10.1016/j.foreco.2008.02.008)
- Egozcue JJ, Pawlowsky-Glahn V, Mateu-Figueras G, Barcelo-Vidal C (2003). Isometric logratio transformations for compositional data analysis. *Mathematical Geology* 35: 279-300. - doi: [10.1023/A:1023818214614](https://doi.org/10.1023/A:1023818214614)
- Egozcue JJ, Pawlowsky-Glahn V (2005). Groups of parts and their balances in compositional data analysis. *Mathematical Geology* 37: 795-828. - doi: [10.1007/s11004-005-7381-9](https://doi.org/10.1007/s11004-005-7381-9)
- Egozcue JJ, Pawlowsky-Glahn V (2011). Basic concepts and procedures. In: "Compositional Data Analysis: Theory and Applications" (Pawlowsky-Glahn V, Buccianti A eds), John Wiley and Sons, Chichester, UK, pp. 12-28. [online] URL: <http://books.google.com/books?id=s9L5TrM1gYEC>
- Elliott KJ, Swank WT (2008). Long-term changes in forest composition and diversity following early logging (1919-1923) and the decline of American chestnut (*Castanea dentata*). *Plant Ecology* 197: 155-172. - doi: [10.1007/s11258-007-9352-3](https://doi.org/10.1007/s11258-007-9352-3)
- Filzmoser P, Hron K, Reimann C (2009). Univariate statistical analysis of environmental (compositional) data: problems and possibilities. *Science of the Total Environment* 407: 6100-6108. - doi: [10.1016/j.scitotenv.2009.08.008](https://doi.org/10.1016/j.scitotenv.2009.08.008)
- Firm D, Nagel TA, Diaci J (2009). Disturbance history and dynamics of an old-growth mixed species mountain forest in the Slovenian Alps. *Forest Ecology and Management* 257: 1893-1901. - doi: [10.1016/j.foreco.2008.09.034](https://doi.org/10.1016/j.foreco.2008.09.034)
- FMP (1954). Gozdnogospodarski načrt gozdnogospodarske enote Leskova dolina 1954-1963 [Forest management plan for FMU Leskova dolina 1954-1963]. Gozdno gospodarstvo Postojna, Postojna, Slovenia, pp. 215. [in Slovene]
- FMP (1964). Gozdnogospodarski načrt gozdnogospodarske enote Leskova dolina 1964-1973 [Forest management plan for FMU Leskova dolina 1964-1973]. Gozdno gospodarstvo Postojna, Postojna, Slovenia, pp. 207. [in Slovene]
- FMP (1974). Gozdnogospodarski načrt gozdnogospodarske enote Leskova dolina 1974-1983 [Forest management plan for FMU Leskova dolina 1974-1983]. Gozdno gospodarstvo Postojna, Postojna, Slovenia, pp. 198. [in Slovene]
- FMP (1994). Gozdnogospodarski načrt gozdnogospodarske enote Leskova dolina 1994-2003 [Forest management plan for FMU Leskova dolina 1994-2003]. Zavod za gozdove Slovenije, Ljubljana, Slovenia, pp. 223. [in Slovene]
- FMP (2004). Gozdnogospodarski načrt gozdnogospodarske enote Leskova dolina 2004-2013 [Forest management plan for FMU Leskova dolina 2004-2013]. Zavod za gozdove Slovenije, Ljubljana, Slovenia, pp. 201. [in Slovene]
- Fule PZ, Korb JE, Wu R (2009). Changes in forest structure of a mixed conifer forest, southwestern Colorado, USA. *Forest Ecology and Management* 258: 1200-1210. - doi: [10.1016/j.foreco.2009.06.015](https://doi.org/10.1016/j.foreco.2009.06.015)
- Gašperšič F (1967). Razvojna dinamika mešanih gozdov jelke-bukve na Snežniku v zadnjih 100 letih [Developmental dynamics of mixed fir-beech forests in Snežnik in the last 100 years]. *Gozdarski vestnik* 7-8: 202-237. [in Slovene]
- Greenacre M (2011). Measuring subcompositional incoherence. *Mathematical Geosciences* 43: 681-693. - doi: [10.1007/s11004-011-9338-5](https://doi.org/10.1007/s11004-011-9338-5)
- Heiri C, Wolf A, Rohrer L, Bugmann H (2009). Forty years of natural dynamics in Swiss beech forests: structure, composition, and the influence of former management. *Ecological Applications* 19: 1920-1934. - doi: [10.1890/08-0516.1](https://doi.org/10.1890/08-0516.1)
- Janik D, Adam D, Vrska T, Hort L, Unar P, Kral K, Samonil P, Horal D (2008). Tree layer dynamics of the Cahnov-Soutok near-natural floodplain forest after 33 years (1973-2006). *European Journal of Forest Research* 127: 337-345. - doi: [10.1007/s10342-008-0210-1](https://doi.org/10.1007/s10342-008-0210-1)
- Lawrence D, Suma V, Mogeja JP (2005). Change in species composition with repeated shifting cultivation: Limited role of soil nutrients. *Ecological Applications* 15: 1952-1967. - doi: [10.1890/04-0841](https://doi.org/10.1890/04-0841)
- Legendre P, Legendre L (1998). *Numerical Ecology*. Elsevier, Amsterdam, Boston, London, pp. 853.
- Martin-Fernandez JA, Barcelo-Vidal C, Pawlowsky-Glahn V (2003). Dealing with zeros and missing values in compositional data sets using nonparametric imputation. *Mathematical Geology* 35: 253-278. - doi: [10.1023/A:1023866030544](https://doi.org/10.1023/A:1023866030544)
- Martin-Fernandez JA, Palarea-Albaladejo J, Olea RA (2011). Dealing with zeros. In: "Compositional Data Analysis: Theory and Applications" (Pawlowsky-Glahn V, Buccianti A eds), John Wiley and Sons, Chichester, UK, pp. 43-58. [online] URL: <http://books.google.com/books?id=s9L5TrM1gYEC>
- Mateu-Figueras G, Pawlowsky-Glahn V, Egozcue JJ (2011). Principle of working on coordinates. In: "Compositional Data Analysis: Theory and Applications" (Pawlowsky-Glahn V, Buccianti A eds), John Wiley and Sons, Chichester, UK, pp. 31-42. [online] URL: <http://books.google.com/books?id=s9L5TrM1gYEC>
- Nagel TA, Diaci J, Jerina K, Kobal M, Rozenberger D (2015). Simultaneous influence of canopy decline and deer herbivory on regeneration in a conifer-broadleaf forest. *Canadian Journal of Forest Research* 45: 266-275. - doi: [10.1139/cjfr-2014-0249](https://doi.org/10.1139/cjfr-2014-0249)
- Pallardy SG, Nigh TA, Garrett HE (1988). Changes in forest composition in Central Missouri - 1968-1982. *American Midland Naturalist* 120: 380-390. - doi: [10.2307/2426010](https://doi.org/10.2307/2426010)
- Pawlowsky-Glahn V, Egozcue JJ (2002). BLU estimators and compositional data. *Mathematical Geology* 34: 259-274. - doi: [10.1023/A:1014890722372](https://doi.org/10.1023/A:1014890722372)
- Pawlowsky-Glahn V (2003). Statistical modelling on coordinates. Website. [online] URL: [http://ima.udg.edu/Activitats/CoDaWork03/paper\\_Ver aPG.pdf](http://ima.udg.edu/Activitats/CoDaWork03/paper_Ver aPG.pdf)
- Pearson K (1897). Mathematical contributions to the theory of evolution. On a form of spurious correlation which may arise when indices are used in the measurement of organs. *Proceedings of the Royal Society of London*, pp. 409-502
- Philippi TE, Dixon PM, Taylor BE (1998). Detecting trends in species composition. *Ecological Applications* 8: 300-308. - doi: [10.1890/1051-0761\(1998\)008\[0300:DTISC\]2.0.CO;2](https://doi.org/10.1890/1051-0761(1998)008[0300:DTISC]2.0.CO;2)
- Pinto F, Romaniuk S, Ferguson M (2008). Changes to preindustrial forest tree composition in central and northeastern Ontario, Canada. *Canadian Journal of Forest Research* 38: 1842-1854. - doi: [10.1139/X08-034](https://doi.org/10.1139/X08-034)
- R Core Team (2012). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. [online] URL: <http://www.r-project.org>
- Sagar R, Raghubanshi AS, Singh JS (2003). Tree species composition, dispersion and diversity along a disturbance gradient in a dry tropical forest region of India. *Forest Ecology and Management* 186: 61-71. - doi: [10.1016/S0378-1127\(03\)00235-4](https://doi.org/10.1016/S0378-1127(03)00235-4)
- Schuler TM (2004). Fifty years of partial harvesting in a mixed mesophytic forest: composition and productivity. *Canadian Journal of Forest Research* 34: 985-997. - doi: [10.1139/x03-262](https://doi.org/10.1139/x03-262)
- Shurtz RF (2003). Compositional geometry and mass conservation. *Mathematical Geology* 35: 927-937. - doi: [10.1023/B:MATG.0000011586.09198.a0](https://doi.org/10.1023/B:MATG.0000011586.09198.a0)
- Skovsgaard JP, Vanclay JK (2007). Forest site productivity: a review of the evolution of dendrometric concepts for even-aged stands. *Forestry* 81 (1): 13-31. - doi: [10.1093/forestry/cpm041](https://doi.org/10.1093/forestry/cpm041)
- Tikvić I, Seletković Z, Ugarković D, Posavec S, Španjol Z (2008). Dieback of Silver Fir (*Abies alba* Mill.) on Northern Velebit (Croatia). *Periodicum Biologorum* 110: 137-143.
- Tolosana-Delgado R, Van Den Boogart KG (2011). Linear models with compositions in R. In: "Compositional Data Analysis: Theory and Applications" (Pawlowsky-Glahn V, Buccianti A

- eds). John Wiley and Sons, Chichester, UK, pp. 356-371. [online] URL: <http://books.google.com/books?id=s9L5TrM1gYEC>
- Van Den Boogart KG, Tolosana-Delgado R (2006). Compositional data analysis with “R” and the package “compositions”. In: “Compositional Data Analysis: From Theory to Practice” (Buccianti A, Mateu-Figueras, G, Pawlowsky-Glahn V eds). Geological Society, London, UK, pp. 119-127. [online] URL: <http://books.google.com/books?id=s9L5TrM1gYEC>
- Van Den Boogart KG, Tolosana-Delgado R (2008). “compositions”: a unified R package to analyze compositional data. *Computers and Geosciences* 34: 320-338. - doi: [10.1016/j.cageo.2006.11.017](https://doi.org/10.1016/j.cageo.2006.11.017)
- Van Der Maarel E (2005). *Vegetation ecology*. Blackwell Publishing, Malden, Oxford, Carlton, pp. 411. [online] URL: <http://books.google.com/books?id=mmPoH9d72AoC>
- Woods KD (2000a). Dynamics in late-successional hemlock-hardwood forests over three decades. *Ecology* 81: 110-126. - doi: [10.1890/0012-9658\(2000\)081\[0110:DILSHH\]2.o.CO;2](https://doi.org/10.1890/0012-9658(2000)081[0110:DILSHH]2.o.CO;2)
- Woods KD (2000b). Long-term change and spatial pattern in a late-successional hemlock-northern hardwood forest. *Journal of Ecology* 88: 267-282. - doi: [10.1046/j.1365-2745.2000.00448.x](https://doi.org/10.1046/j.1365-2745.2000.00448.x)