Compositional meta-analysis of *Citrus* varieties in the state of São Paulo, Brazil

DANILO EDUARDO ROZANE[1], DIRCEU DE MATTOS JUNIOR[2], SERGE-ETIENNE PARENT[3], WILLIAM NATALE[1] AND LEON ETIENNE PARENT[3]

[1]Departamento de Solos e Adubos, Unesp, Universidade Estadual Paulista, 14870-000, Jaboticabal, São Paulo, Brasil, email: danilorozane@yahoo.com.br; natale@fcav.unesp.br

[2]APTA-Citros, Cordeirópolis-São Paulo-Brasil, e-mail: ddm@centrodecitricultura.br

[3]ERSAM, Département des Sols et de génie agroalimentaire, Université Laval, Québec, Canada G1K 7P4, email: leon-etienne.parent@fsaa.ulaval.ca

Brazil is the largest orange (*Citrus sinensis*) producer worldwide. The nutrient management of orange orchards is designed from experiments on a limited number of varieties. This knowledge is transferred to other varieties by diagnosing tissue nutrient composition. Nutrient diagnostic tools are based on nutrient concentration (critical minimum value or CMV) and ratio (Diagnosis and Recommendation Integrated System or DRIS) norms that disregard the compositional nature of analytical data and the limited number of nutrient ratios that can be diagnosed independently in a given composition. The diagnosis of cationic micronutrients is also biased by contamination from fungicides. Compositional data analysis that can avoid such problems has been first applied to tissue analysis of agricultural crops using centered log ratios (Compositional Nutrient Diagnosis – CND-*clr*). The isometric log ratio (*ilr*) transformation is a new approach based on binary nutrient ratios and the principle of orthogonality (CND-*ilr*). Binary partitions can be defined and varietal nutrient profiles classified based on positive and negative nutrient interactions and meta-analysis. We analyzed 11 nutrients (N, S, P, K, Ca, Mg, B, Cu, Zn, Mn, Fe) in tissue samples across 108 orchard areas, i.e. 31 'Valencia', 22 'Hamlin', 20 'Pêra', and 35 'Natal'. Nutrients were partitioned between macro- and micro-nutrients as well as anionic and cationic species. The effect size of varieties over 'Valencia' was quantified by the mean and standard deviation of *ilr* values across *ilr* coordinates. Specific varietal nutrient profiles and *ilr* norms were defined. To guide correcting nutrient deficiencies by appropriate nutrient management, compositions can be varied by a perturbation vector on nutrients with to the largest and most negative influence on *ilr* differences from *ilr* norms until the Aitchison distance falls below critical value.

1. Introduction

Brazil is the first orange (*Citrus sinensis*) producer worldwide accounting for 27% of the 68,6 Mg total production in 2008, i.e. 2.0 times the US production and 5.5 times that of Spain (FAO, 2010). The nutrient management of orange orchards is based on criteria like nutrient removal through harvest and nutrient concentration in plant tissue along with the monitoring of soil nutrient reserves.

Plant analysis has advantage over surface soil analysis (0-20 cm) as diagnostic tool for a deep rooted fruit crops because the plant have access to nutrient deeper in the soil than would be found through normal soil analytical procedures (Smith et al. 1997). Tissue nutrient diagnosis requires a standardized sampling method, routine analytical methods, standard nutrient values and an interpretation of analytical results that is free of bias for fertilizer recommendations (Kenworthy, 1983). Standard nutrient values are concentrations ranges thought to be adequate for highly productive crops. It is assumed that, although nutrients are interrelated, the effect of a given nutrient on crop yield depends on how close other nutrients are from their optimum. Hence, all nutrients but the ones being diagnosed are assumed to be in sufficient but non excessive amounts. Nevertheless, nutrient concentration standards were

Egozcue, J.J., Tolosana-Delgado, R. and Ortego, M.I. (eds.)
ISBN: 978-84-87867-76-7

criticized for not accounting for nutrient interactions (Bates, 1971) since several dual and multiple interactions have been identified in plants (Bergmann, 1988; Malavolta, 2006). Nutrients levels thus bear relative information often expressed as ratios or products (Walworth and Sumner, 1987).

Beaufils (1973) addressed nutrient interactions using dual ratios in the Diagnosis and Recommendation Integrated System (DRIS). DRIS was applied to 'Valencia' orange in USA (Beverly, 1987a), Venezuela (Rodriguez et al., 1997) and Brazil (Mourão Filho, 2005). DRIS is often viewed as complementary to nutrient concentration standards (Mourão Filho, 2005). On the other hand, tissue analytical data are strictly positive data constrained to a compositional space that includes nutrient concentrations and a filling value constrained to the unit of measurement. Within that closed space, nutrient interaction whereby any nutrient addition may positively or negatively influence the level of another (Malavolta, 2006) is a source of resonance and spurious correlations (Aitchison, 1986).

Lagatu and Maume (1935) illustrated nutrient interactions in a constrained space using closed ternary diagrams where vertices represented nutrients in proportions varying between 0 and 100%. Because nutrient fractions add up to 100%, there is at least one redundant concentration. The trivial case is a 2-compositional system where the correlation coefficient between changing components is exactly minus one since any change in one component must affect the proportion of the other by exactly the same value (Thomas and Aitchison, 2006). Since uncorrelated proportions are not necessarily independent, true or spurious correlations between proportions cannot be interpreted in a meaningful way (Butler, 2005).

Compositional data intrinsically have non-normal distribution since they are constrained between 0 and 100% rather than being randomly distributed in the real space between $-\infty$ and $+\infty$. Gaussian techniques should not be applied to compositional data because it is impossible that confidence intervals be less than 0% or more than 100% (Diaz-Zorita et al., 2002). As a result, statistical analyses based on the normality assumption such as regression, univariate, and multivariate analyses are misleading (Butler et al., 2005). In contrast, log ratios can take any real value hence projecting the compositional space into the real space where Gaussian laws can be applied (Aitchison, 1986). Tissue nutrient concentrations can be transformed into real values using techniques of compositional data analysis. Aitchison (1986) proposed that compositional data should be converted to additive (*alr*) and centered log ratios (*clr*) before running linear statistical models.

Parent and Dafir (1992) linked DRIS to compositional data analysis using the log-DRIS approach of Beverly et al. (1987b) and *clr* and considered CND-*clr* as a diagnostic tool by itself. Egozcue and Pawlowsky-Glahn (2005) showed that isometric log ratios (*ilr*) are sound transformations for sequential binary partitions of components with orthonormal basis. A binary partition can be a dual ratio or product or a ratio or product involving two sub-compositions within an orthogonal framework. This is appropriate for plant nutrients since two and several nutrients can be arranged into sequential binary partitions such as macro- vs. micronutrients, [N + S] vs. [P], [K] vs. [Ca + Mg], and [Ca] vs. [Mg] (Malavolta, 2006). The *ilr* coordinates are real data that avoids spurious correlations, can be tested for normality and are analyzed by common linear statistical models. The *ilr* coordinates may thus be useful to describe nutrient profiles of orange orchards for classification and nutrient management.

Our objective was to compare the *ilr* coordinates of tissue analytical data for orange varieties used in Brazilian orchards. We compare varieties using Manova and Anova across *ilr* coordinates. Groups of varieties are compared across *ilr* coordinates using meta-analysis (Borenstein et al., 2009). This meta-analysis procedure designated here as 'compositional meta-analysis' is instrumental to form groups of varieties with similar nutrient profiles leading to uniform nutrient management. Knowledge from costly fertilizer trials on a given variety could be transferred to other varieties showing the same nutrient profile.

2. Material and methods

We collected, in non irrigated orange orchards across the state of São Paulo, Brazil, yield data (kg fruit ha^{-1} and kg fruit tree^{-1}) and summer foliar samples on fruit-bearing shoot (3rd to 4th leaf from fruits) when fruits were 2-4 cm in size (Quaggio et al., 1997). The dataset comprised six rootstocks ('Cravo', 'Cleópatra', 'Swingle', 'Outro', 'Volkameriano', and 'Trifoliata') and four varieties ('Valencia', 'Hamlin', 'Pêra', and 'Natal'). In each orchard, uniform areas of 25 trees were documented for yield, leaf analysis, plant density and age. Leaves were not washed after collection despite seasonal fungicide applications. Leaf samples were analyzed for N, S, P, K, Ca, Mg, B, Cu, Zn, Mn, and Fe using standard procedures (Jones and Case, 1990). The original dataset comprised 256 lines but we screened data for replications and outliers (Grubbs test). Where there were several yield measurements for the same foliar composition, we selected the median yield. The final dataset comprised 108 observations as follows: 31 'Valencia', 22 'Hamlin', 20 'Pêra', and 35 'Natal'. The compositional space *x* of nutrient concentrations in the foliar orange tissue was thus described as follows:

$$x = C(N, S, P, K, Ca, Mg, B, Cu, Zn, Mn, Fe, F_v) = 1000g\ kg^{-1} \qquad (1)$$

The filling value F_v was computed by difference between 1000 g kg^{-1} and analytical results. With 11 nutrients and F_v (D = 12) there were *D-1* or 11 *ilr* coordinates. Binary partitions are made of nutrients compounded in the numerator (+) and the remainder in the denominator (-). The *ilr* coordinate is computed as follows (Egozcue and Pawlowsky-Glahn, 2006):

$$x_i^* = \sqrt{\frac{rs}{r+s}}\, ln \frac{g(x_+)}{g(x_-)}, \qquad (2)$$

Where r and s are numbers of components in plus (+) and minus (–) groups, respectively, $g(x_+)$ is geometric mean of components in the plus group x_+ and $g(x_-)$ is geometric mean of components in the minus group x_-. Coefficient $\sqrt{\frac{rs}{r+s}}$ is the balance between the plus and minus groups. The *ilr* are ad hoc partitions based on a conceptual interactive model that facilitates the interpretation of the results. The sequential binary partitions were selected (Table 1) according to current knowledge on nutrient levels and both positive and negative interactions in plants (Malavolta, 2006).

Ilr	N	S	P	K	Ca	Mg	B	Cu	Zn	Mn	Fe	F$_v$	r	s	Balance
Sequential orthonormal partitions of components															
1	1	1	1	1	1	1	1	-1	-1	-1	-1	0	7	4	1.595
2	1	1	1	1	1	1	-1	0	0	0	0	0	6	1	0.926
3	1	1	1	-1	-1	-1	0	0	0	0	0	0	3	3	1.225
4	1	1	-1	0	0	0	0	0	0	0	0	0	2	1	0.816
5	1	-1	0	0	0	0	0	0	0	0	0	0	1	1	0.707
6	0	0	0	1	-1	-1	0	0	0	0	0	0	1	2	0.816
7	0	0	0	0	1	-1	0	0	0	0	0	0	1	1	0.707
8	0	0	0	0	0	0	0	1	1	-1	-1	0	3	1	1.000
9	0	0	0	0	0	0	0	1	-1	0	0	0	2	1	0.707
10	0	0	0	0	0	0	0	0	0	1	-1	0	1	1	0.707
11	1	1	1	1	1	1	1	1	1	1	1	-1	11	1	0.957

Table 2. *Ilr* coordinates and filling value (F$_v$) of foliar nutrient composition of orange trees in Brazil.

Nutrient interactions involving [B] vs. macronutrient, anions vs. cations, and [K] vs. [Ca + Mg] (Malavolta, 2006) were translated into *ilr* coordinates for diagnostic purposes. The [Cu + Zn] vs. [Mn +Fe], [Cu] vs. [Zn] and [Fe] vs. [Mn] *ilr* coordinates are related to fungicide applications (Cu, Zn) as well as soil genesis (Mn, Fe) (Embrapa, 2006).

In compositional analysis, the difference between two compositions is computed as the Aitchison distance across *ilr* coordinates computed as follows (Egozcue & Pawlowsky-Glahn, 2006):

$$d_a^2(x,y) = \sum_{i=1}^{D-1}(x_i^* - y_i^*)^2 \text{ and } \mathcal{A} = \sqrt{d_a^2(x,y)}, \qquad (3)$$

Where y_i^* is reference composition and $\mathcal{A}$ is Aitchison distance.

We selected orange trees that produced more than 92 kg fruit tree^{-1} for non-irrigated conditions (Rodriguez et al., 1997). In average those orchards produced 48.5 Mg fruit ha^{-1}, in the low-yield range (40-50 Mg fruit ha^{-1}) but 31 orchards were in the high-yield range (> 60 Mg fruit ha^{-1}) compared to irrigated orchards (Mourão Filho, 2005).

The critical value of Aitchison distance for the high-yield sub-population was obtained using the Cate-Nelson procedure (Nelson and Anderson, 1984). Briefly, the scatter diagram relating a performance criterion (e.g. yield) to the Aitchison distance was partitioned into four quadrants (NW, NE, SW, SE) and the number of points was maximized in the opposite quadrants NW and SE. The NW quadrants represented true high yielders (high yield and small $\mathcal{A}$), the NE, false high yielders (high yield and high $\mathcal{A}$), the SW, false high yielders (low yield and low $\mathcal{A}$), and the SE, true low yielders (low yield and high $\mathcal{A}$). Using the Excel package and the chi-square homogeneity test (Hollander and Wolfe, 1999), the number of successful and unsuccessful classification was compared to an equal distribution between successful and unsuccessful cases.

Compositional analyses were conducted in R (van den Boogaart et al., 2000). We ran Manova and Anova and meta-analysis using the R package. Normality was verified using the Anderson-Darling test (Aitchison, 1986). Not normally distributed *ilr* data were transformed according to Yeo and Johnson (2000) and their analysis compared to non transformed *ilr* data to verify the robustness of the statistical analyses against non normal distribution. We made groups of varieties across rootstocks (e.g. Rodriguez et al., 1997) and tested their homogeneity using meta-analysis in R (Schwarzer, 2010). Since the reference variety was compared to each others and thus repeated in the same analysis, we divided the number of observations in the reference variety by the Box factor (Huynh and Feldt, 1976), *k-1*, where *k* is total number of varieties. Difference between varieties was determined using the random model (Borenstein et al., 2009) and *t*-tests. The degree of heterogeneity was measured by the I^2 statistics: heterogeneity is high if $I^2 > 0.75$, medium if $I^2 \approx 0.50$, and small if $I^2 < 0.25$ (Borenstein et al., 2009). If $I^2 < 0$, I^2 was given a value of zero.

3. Results and discussion

3.1 Statistical analysis

Statistics on foliar composition are presented in Table 2. Like Mourão Filho (2005), we found large variations in Cu, Zn, Mn, and Fe levels possibly due to contamination by fungicide sprays and differential soil chemistry. ‘Hamlin’ apparently received less fungicide sprays than others. There is a wide variety of soil classes in the state of São Paulo such as Argilissols. Cambissols, Latossols, Neossols, and Nitossols (Embrapa, 2006). Besides of differences in nutrient management of orchards, these soils cover a large spectrum of pH values and levels of oxy-hydroxides of Al, Fe, and Mn that may impact on foliar Mn and Fe levels and their interactions with other nutrients.

Component	‘Hamlin’ (n = 22)	‘Natal’ (n = 35)	‘Pêra’ (n = 20)	‘Valencia’ (n = 31)
	----------------------------------	---------------- $g\ kg^{-1}$	----------------	----------------------------------
N	28.6±3.0	28.3±2.8	26.3±1.4	27.9±2.1
S	3.2±0.7	3.0±0.7	3.0±0.9	3.6±0.7
P	2.0±0.5	1.6±0.6	1.5±0.7	2.0±0.6
K	16.3±3.0	12.9±3.7	13.5±4.5	14.8±2.8
Ca	34.7±8.0	34.3±7.8	34.5±8.8	38.4±6.4
Mg	4.7±1.5	4.0±1.1	3.7±1.3	5.2±1.2
B	0.095±0.021	0.088±0.012	0.080±0.044	0.107±0.045
Cu	0.031±0.035	0.055±0.033	0.061±0.027	0.067±0.034
Zn	0.032±0.017	0.066±0.041	0.070±0.027	0.051±0.031
Mn	0.052±0.034	0.076±0.064	0.095±0.052	0.095±0.063
Fe	0.144±0.074	0.162±0.091	0.173±0.069	0.228±0.084
Filling value	910.1±8.6	915.4±9.2	917.0±13.9	907.6±6.0

Table 2. Statistics (mean±standard deviation) on foliar nutrient concentration and the filling value to 1000 $g\ kg^{-1}$ dry matter for orange varieties producing > 92 kg fruit $tree^{-1}$ in the state of São Paulo, Brazil (N.B. Totals may be different than 1000 due to rounding error).

The Anderson-Darling normality test was significant for *ilr2*, *ilr5*, and *ilr11* ($P > 0.01$). The Manova showed a significant effect of varieties across *ilr* coordinates ($P < 0.01$). Anova showed similar results whether *ilr* or Yeo-Johnson transformed *ilr* were used (Table 3).

Ilr	Orange variety					Partition
	‘Hamlin’	‘Natal’	‘Pêra’	‘Valencia’		
	Mean				SD	
1	7.483a	6.543b	6.113b	6.500b	0.6508	[Macronutrients + B] vs. [Cu, Zn, Mn, Fe]
2	4.202a	4.142a	4.278a	4.162a	0.2348	Macronutrients vs. B
3	-1.084a	-1.039a	-1.087a	-1.104a	0.1241	Anions vs. cations
4	1.282a	1.459ab	1.496b	1.345ab	0.2361	[N, S] vs. P
5	1.564ab	1.610a	1.561ab	1.466b	0.1927	N vs. S
6	0.214a	0.062a	0.133a	0.037a	0.2621	K vs. [Ca, Mg]
7	1.422ab	1.532ab	1.587a	1.421b	0.1909	Ca vs. Mg
8	-1.133c	-0.613a	-0.683ab	-0.968bc	0.4237	[Cu, Zn] vs. [Mn, Fe]
9	-0.276b	-0.161ab	-0.112ab	0.206a	0.6143	Cu vs. Zn
10	-0.730a	-0.589a	-0.449a	-0.682a	0.5667	Mn vs. Fe
11	-6.709b	-6.626b	-6.586ab	-6.455a	0.1841	[Macronutrients, B] vs. [Filling value]

Table 3. Statistics (mean and standard deviation = SD) of *ilr* coordinates of orange varieties. Means followed by the same letter on the same line are not significantly different at the 0.05 level according to the Duncan Multiple Range Test (104 error degrees of freedom).

Varieties had significant effects (*P < 0.05*) on *ilr* 1-4-5-7-8-9-11. 'Hamlin' showed a significantly higher *ilr1* value than others due to lower concentrations of cationic micronutrients (Table 1). The macronutrient/B and anion/cation ratios did not differ between varieties (*ilr2* and *ilr3*). The anionic balance (*ilr4*) was significantly higher in 'Pêra' than 'Hamlin' due to P accumulation in 'Hamlin' (Table 1). The N/S ratio (*ilr5*) was higher in 'Hamlin', 'Natal' and 'Pêra' than in 'Valencia'. The *ilr6* did not differ among varieties. The Ca/Mg ratio was significantly higher in 'Pêra' than in 'Valencia' due to higher leaf Mg in the latter. The *ilr8* was higher in 'Natal' and 'Pêra' than in 'Hamlin' and 'Valencia'. The nutrient balance expressed by *ilr* coordinates thus varied widely among varieties as shown by ANOVA. However, the degree of varietal heterogeneity can be further explored by meta-analysis for use in nutrient management decisions.

3.2 Meta-analysis

A heterogeneity test was provided by meta-analysis when a group is compared to a control. We selected 'Valencia' as control since 'Valencia' is the most documented orange variety worldwide. Criteria to make groups of varieties are the heterogeneity test (I^2) and confidence interval about the standardized mean difference. Where confidence intervals of standardized means are strictly above or below zero, group average is significantly different than control. Where I^2 exceeds 0.25, one variety should be removed to recover group homogeneity. Thereafter, the removed variety is compared to control using a *t* test.

Results of the meta-analysis are presented in Table 4 and Fig. 1 when comparing 'Valencia' to other varieties.

Ilr	Q	I^2	C. I.	Proposed grouping (all $I^2 = 0$)
1	17**	0.87	-1.08 to 1.91	[Hamlin] [Valencia, Natal] [Pêra]
2	4.5ns	0.56	-0.84 to 0.65	[Valencia, Hamlin, Pêra] [Natal]
3	0.4ns	0	0.00 to 1.00	[Valencia] [Hamlin, Natal, Pêra]
4	4.3ns	0.53	-0.45 to 1.02	[Valencia, Hamlin] [Natal, Pêra]
5	0.5ns	0	0.23 to 1.24	[Valencia] [Hamlin, Natal, Pêra]
6	1.2ns	0	0.04 to 1.04	[Valencia] [Hamlin, Natal, Pêra]
7	1.8ns	0	0.03 to 1.03	[Valencia] [Hamlin, Natal, Pêra]
8	4.6ns	0.57	-0.53 to 1.00	[Valencia, Hamlin] [Natal, Pêra]
9	0.3ns	0	-1.33 to -0.32	[Valencia] [Hamlin, Natal, Pêra]
10	2.1ns	0.03	-0.29 to 0.71	[Valencia, Hamlin, Natal, Pêra]
11	0.0ns	0	-1.43 to -0.40	[Valencia, Hamlin, Natal, Pêra]

Table 4. Degree of heterogeneity (I^2) and confidence interval (C. I.) of standardized *ilr* coordinates of orange varieties against 'Valencia'. Groupings are proposed to reduce I^2 to zero.

The *ilr1*, the most contrasting among orange varieties, showed that 'Hamlin', of early maturation, was the one where macronutrients and boron loaded most against cationic micronutrients, probably due to less fungicide applications. 'Pêra', of later maturation, was at the other end, indicating that more fungicides (Cu, Zn, Mn) had been applied probably due to differences in susceptibility to fungal diseases by the time of fruit maturation.

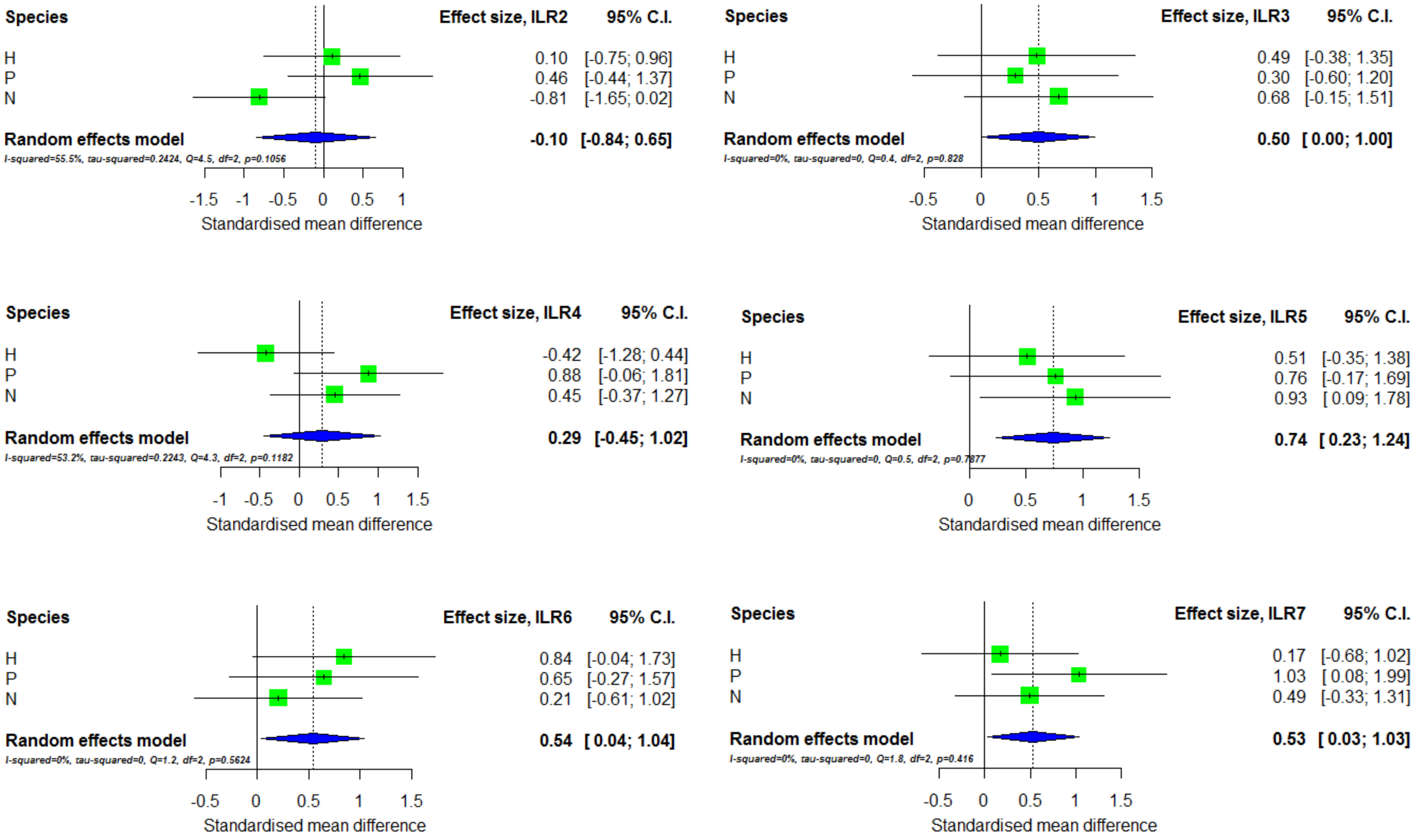

Fig. 1. Forest plots of the meta-analysis of ilr2 to ilr7 across orange varieties (H = 'Hamlin'; P = 'Pêra'; N ='Natal') against 'Valencia'.

Nutrient status was similar across varieties for *ilr2*, *ilr3*, *ilr6*, and *ilr10*. 'Hamlin' and 'Valencia' formed one group while 'Natal' and 'Pêra' formed another one for *ilr4*, *ilr7*, and *ilr8*. 'Valencia' contrasted with other cultivars for *ilr5*, *ilr9*, and *ilr11*. These results show that nutrient management of 'Valencia' may be suitable for other orange varieties across certain dimensions (i.e. *ilr* coordinates) of plant nutrition but unsuitable across other dimensions.

3.3 Nutrient norms and diagnosis

Nutrient norms for the reference composition (y_i^*) were derived from mean *ilr* values where effect sizes were homogeneous across varieties (Table 5). The *ilr4* to *ilr7* coordinates were critical for macronutrient management. 'Natal' and 'Pêra' required a higher [N, S] to [P] ratio than 'Hamlin' and 'Valencia' (*ilr4*). Target value for the [N] to [S] ratio (*ilr5*) was found to be lower for 'Valencia' than others. Although balance between K and divalent cations was the same across cultivars, the [Ca] to [Mg] ratio was lower in 'Hamlin' and 'Valencia'.

Ilr	Orange variety			
	'Hamlin'	'Natal'	'Pêra'	'Valencia'
	Mean			
1	7.483	6.521	6.113	6.521
2	4.182	4.182	4.182	4.182
3	-1.084	-1.084	-1.084	-1.084
4	1.314	1.478	1.478	1.314
5	1.564	1.564	1.564	1.466
6	0.098	0.098	0.098	0.098
7	1.422	1.560	1.560	1.422
8	-1.051	-0.648	-0.648	-1.051
9	-0.137	-0.137	-0.137	-0.137
10	-0.636	-0.636	-0.636	-0.636
11	-6.606	-6.606	-6.606	-6.606

Table 5. Proposed nutrient norms using median values for homogeneous varietal groupings with zero I^2.

The Aitchison distance was computed between a given composition (x_i^*) and the reference one. If the Aitchison distance across the selected set of *ilr* values exceeded a critical value, the diagnosed composition was declared 'imbalanced'. We excluded *ilr1* and *ilr8* to *ilr11* since cationic micronutrients were more related to soil genesis and disease control than plant nutrition. The critical $\mathcal{A}$ value was found to be 0.47 for yields exceeding 154 kg tree^{-1} (Fig. 1). With an average plantation density of 357 tree ha^{-1}, high yielders produced more than 55 Mg fruit ha^{-1}, close to 60 Mg fruit ha^{-1} considered as high yield level by Mourão Filho (2005). True high yielders and true low yielders made up 64% of this classification (69 successful and 39 unsuccessful cases for classification). Compared to an equal distribution of 50% successful and 50% unsuccessful cases and using a chi-square homogeneity test (Hollander and Wolfe, 1999), the null hypothesis was rejected (χ_1^2 = 3.93, $P < 0.05$). Hence, the critical Aitchison distance of 0.47 showed diagnostic potential to detect nutrient imbalance in orange orchards. However, independent fertilizer trials must be conducted to ascertain the critical Aitchison distance of 0.47.

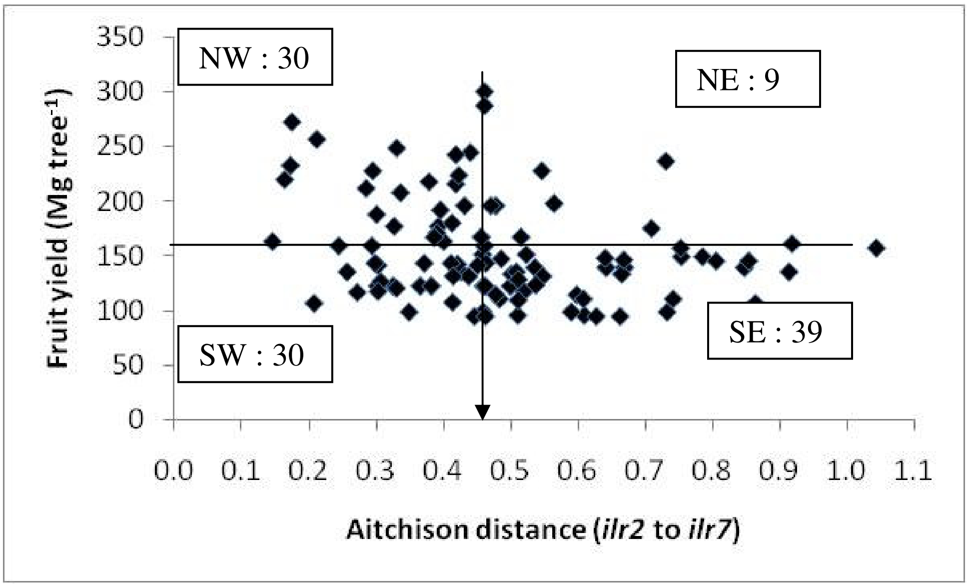

Fig. 1. Cate-Nelson partitioning of the relationship between the Aitchison distance across *ilr2* to *ilr7* and fruit yield of orange orchards in the state of São Paulo, Brazil (number of observations per quadrant in parentheses).

Since the Aitchison distance can be decomposed into ilr differences (Eq. 3), the main cause of imbalance is shown by the greatest ilr difference while the direction of imbalance is shown by the sign of that difference. To correct the most imbalanced nutrients, composition of the diagnosed tissue can be perturbed until the Aitchison distance falls below 0.40.

In compositional data analysis, a perturbation is a multiplicative operation whereby the original composition x is converted into a perturbed composition X through perturbation vector u as follows (Aitchison, 1986):

$$X = u \oplus x = \mathcal{C}(x_1u_1, \dots, x_Du_D) \qquad (4)$$

Where $\mathcal{C}$ is the closure operator. Tissue composition can be varied for nutrients contributing most to *ilr* differences from *ilr* norms using a perturbation vector *u* to identify the most deficient nutrients. Compositions are changed by vector u until proper Aitchison distance is recovered, i.e. below critical value. Nutrient management is adjusted by applying a source of deficient nutrients at the rate and timing sought appropriate to re-establish nutrient balance.

4. Conclusion

This paper combines compositional data analysis and meta-analysis to define the nutrient profile of orange varieties and guide nutrient management practices. Sequential binary partitions provided a framework for nutrient profiles based on current knowledge in plant physiology and pathology and in soil genesis.

Conducting a meta-analysis on a survey dataset collected in the state of São Paulo, Brazil, we found that the nutrient profile of 'Valencia' differed from others for *ilr3* and *ilr5*. The varietal nutrient profiles were similar for *ilr2-6-7-8-9-10-11*. 'Hamlin' and 'Valencia' showed *ilr4* and *ilr7* nutrient profiles different from 'Natal' and 'Pêra'. On the other hand, the *ilr* coordinates involving cationic micronutrients may reflect differential applications of fungicides, hence the degree of varietal resistance to fungal diseases.

Our results showed that the nutrient profile of orange varieties could be classified into homogeneous groups to take advantage of fertilizer trials conducted on varieties of the same group. The Aitchison distance and a perturbation vector could be instrumental for diagnostic purposes and nutrient management.

5. References

Aitchison, J. (1986) *The statistical analysis of compositional data.* Chapman and Hall, London, 416 p.

Bates, T.E. (1971) Factors affecting critical nutrient concentrations in plants and their evaluation: a review. *Soil Science 112*, 116-130.

Beaufils, E.R. (1973) *Diagnosis and Recommendation Integrated System (DRIS)*. Soil Science Bulletin #1, University of Natal, Pietermaritzburg, South Africa.

Bergmann, W. (1988) *Ernährungs-störungen bei Kulturpflanzen*. 2. Auflage. Gustav Fischer Verlag, Stuttgart, 762 p.

Beverly, R.B. (1987a) Modified DRIS method for simplified nutrient diagnosis of 'Valencia' oranges. *Journal of Plant Nutrition 10*, 1401-1408.

Beverly, R.B. (1987b) Comparison of DRIS and alternative nutrient diagnostic methods for soybean. Journal of Plant Nutrition 10, 901-920.

Borenstein, M., L. V. Hedges, J. P. T. Higgins, and H. R. Rothstein (2009) *Introduction to meta-ahalysis*. John Wiley and Sons, NY. 421 p.

Butler, A., S. Bierman, and G. Marion. 2005. *Statistical methods for environmental risk assessment*. Compositional data module. Biomathematics and Statistics Scotland, The University of Edinburgh, James Clerk Maxwell Building, The King's Buildings, Edinburgh EH9 3JZ. (htpp://www.bioss.ac.uk/staff.html).

Diaz-Zorita, M., E. Perfect, and J. H. Grove. 2002. Disruptive methods for assessing soil structure. *Soil & Tillage Research 64*, 3-22.

Egozcue, J. J. and V. Pawlowsky-Glahn (2005). Groups of parts and their balances in compositional data analysis. *Mathematical Geology 37*, 795-828.

Egozcue, J. J. And V. Pawlowsky-Glahn (2006). Simplicial geometry for compositional data. In Buccianti, A., Mateu-Figueras, G., Pawlowsky-Glahn, V. (Editors). *Compositional data analysis in the geosciences: from theory to practice.* Geological Society, London, Special Publication 264, p. 145-159.

Embrapa. 2006. *Sistema Brasileiro de classificação de solos*. 2nd Ed., Embrapa Solos, Rio de janeiro, Brazil, 306 p.

Food and Agriculture Organization – FAO (2010) Available in: FAO – <http:// faostat.fao.org>. Accessed on December 28, 2010.

Hollander, M. and D. A. Wolfe. 1999. *Nonparametric statistical methods*. 2nd Ed., John Wiley & Sons, NY, 787 p.

Huynh H. and L. S. Feldt (1976) Estimation of the Box correction for degrees of freedom from sample data in randomized block and split-plot designs. *Journal of Educational Statistics 1* (1), 69-82.

Jones, J. B. Jr. and V. W. Case. (1990) Sampling, handling, and analyzing plant tissue samples. p. 389-427 *In* R.L. Westerman (ed). *Soil testing and plant analysis*. 3rd edition, Soil Science Society of America Book Series 3. Madison, WI.

Kenworthy, A.L. (1983) Leaf analysis as an aid in fertilizing orchards. p. 381-392 *in* L.M. Walsh and J.D. Beaton (eds.) *Soil testing and plant analysis*. Revised edition, 5th Printing, Soil Soil Science Society of America, Madison, WI.

Lagatu, H. and L. Maume (1935) The variations of the sum N + P205 + K20 per 100 parts of dry material of the leaf of a cultivated plant. Comptes Rendus de l'Académie D'Agriculture de France 21, 85-92.

Malavolta, E. 2006. *Manual de nutrição de plantas*. Pav. Chimica, ESALQ and Ed. Agron. CERES, São Paulo, Brazil, 638 p.

Mourão Filho, F. D. A. (2005) DRIS and Sufficient Range Approaches in Nutritional Diagnosis of Valencia Sweet Orange on Three Rootstocks. *Journal of Plant Nutrition 28*(4), 691-705.

Nelson, L. A. and R. L. Anderson (1984) Partitioning of soil test-crop response probability. p. 19-38 *in* M. Stelly, ed. *Soil testing: Correlating and interpreting the analytical results*. American Society of Agronomy Special Publication 29. Madison, WI.

Parent, L. E. And M. Dafir (1992) A theoretical concept of compositional nutrient diagnosis. *Journal of American Society of Horticultural Science 117*, 239-242.

Quaggio, J. A., B. Raij, van, and C. T. Piza Junior (1997) Frutíferas. p. 121-153 *in* Raij, B. van, Cantarella, H., Quaggio, J. A., Furlani, A. M. C. (Eds). Recomendações de adubação e calagem para o Estado de São Paulo. Campinas: Instituto Agronômico/Fundação IAC.

Rodriguez, O., E. Rojas, and M. E. Sumner (1997) Valencia orange DRIS norms for Venezuela. *Communications in Soil Science and Plant Analysis 28*(15-16), 1461-1468.

Smith, G. S., G. J. Asher, and C. J. Clark (1997) Kiwifruit Nutrition diagnosis of nutritional disorders. Originally published 1985 ISBN 0-9597693-0-7, revised 1987, republished for HortNET 1997.

Thomas, C. W. And J. Aitchison (2006) Log-ratios and geochemical discrimination of Scottish Dalradian limestones: a case study. In Buccianti, A., Mateu-Figueras, G., Pawlowsky-Glahn, V. (Editors). *Compositional data analysis in the geosciences: from theory to practice*. Geological Society, London, Special Publication 264, 25-41.

Schwarzer, G. (2010). Meta: Meta-Analysis. *R package version 1.6-1*. Accessed: <http://CRAN.R-project.org/package=meta>.

van den Boogaart, K. G, Tolosana, R. and Bren, M. (2010). Compositions: Compositional Data Analysis. *R package version 1.10-1*. Accessed: <http://CRAN.R-project.org/package=compositions>.

Walworth, J. L. and M.E. Sumner (1987) The Diagnosis and Recommendation Integrated System (DRIS). *Advances in Soil Science 6*, 149-188.

Yeo, I. K. and R. Johnson. (2000). A new family of power transformations to improve normality or symmetry. *Biometrika 87*, 954-959.