BALANCE-BASED NUTRIENT DIAGNOSIS OF NEW ZEALAND KIWIFRUIT ORCHARDS

Serge-Étienne Parent[1], Philip Barlow[2] and Léon E. Parent[1*]

[1] ERSAM, Department of Soils and Agrifood Engineering, Université Laval, Québec, Canada G1V0A6,

[2] Bio Soil & Crop Ltd, 37a Newnham Road, RD4, Tauranga, 3172, New Zealand ; Email philip@biosoil.co.nz

This edition of the paper contains some post conference (International Symposium of Soil & Plant Analysis) adjustments made by P Barlow because of extra information which involved local knowledge not previously shared with L Parent.

ABSTRACT

Tissue nutrient composition (ionome) is widely used to diagnose nutrient problems in fruit crops. However, the current interpretation systems based on critical nutrient ranges disregard the particular nature of concentration data that leads to numerically biased diagnosis. Compositional Nutrient Diagnosis based on centred log ratios (CND-*clr*) accounts for nutrient interactions, but is sensitive to agro-chemical contaminants and limits multivariate analysis. The progressive increase of productivity that followed the development of kiwifruit as a commercial crop, as well as the advent of environmental awareness has caused the need to develop nutrient diagnostic tools for better managing kiwifruit orchards. The isometric log ratio (*ilr*) technique, that structures data as coordinates of linearly independent balances of components in the Euclidean space, allows overcoming these mathematical difficulties. Our objective is to use a CND-*ilr* framework to diagnose nutrient imbalances in kiwifruit orchards. We collected kiwifruit analytical data (N, S, Cl, P, K, Ca, Mg, B, Cu, Zn, Mn, Fe) and metadata in 433 kiwifruit production areas. We developed optimum ranges of *ilr* balances to reach high yield standards and computed the Mahalanobis distance as nutrient imbalance predictor. The Mahalanobis distance must be less than 4.28 to reach *ilr* standards of true high-yielders, delimited by a yield of 43 312 kg ha^{-1}. (>12,000 Trays/Ha class I fruit). **High- and low-yielders differed significantly in 6 of the 12 ilr balances.** This novel approach is promising for guiding the nutrient management of kiwifruit orchards.

Keywords: Isometric log ratio coordinates, Aitchison distance, nutrient norms, kiwifruit

INTRODUCTION

Foliar analysis is a useful method to detect nutrient shortage in fruit crops and guide fertilization choices (Kenworthy, 1983). Indeed, plant analysis has the advantage over soil analysis as diagnostic tool for deep rooted plants that access nutrient deeper in the soil than would be found through normal soil analytical procedures (Smith et al. 1997).

As diagnostic tools, Critical Nutrient Range (CNR) nutrient concentrations and Diagnosis and Recommendation Integrated System (DRIS) (Beaufils E.R 1973) nutrient indices – both commonly used to diagnose fruit crop nutrition – were found moderately to closely correlated (Parent et al., 1993; Parent et al., 1994a; Urano et al., 2007; Serra et al., 2010; Camacho et al., 2012). However, in several studies, they returned conflicting diagnoses (Silva et al., 2004; Blanco-Macias et al., 2009; Huang et al., 2012; Wairegi and van Asten, 2012). In this respect, Parent et al. (2012a) demonstrated that leaf nutrient signature standards developed in terms nutrient ranges or DRIS ratios are numerically biased. Parent and Dafir (1992) rectified DRIS using the centred log-ratio (*clr*) technique – proposed by Aitchison (1986) for compositional data – to conduct Compositional Nutrient Diagnosis (CND-*clr*). The critical nutrient range approach, DRIS and CND-*clr* nutrient indices were found to be moderately to closely related to each other (Parent et al., 1994a; Parent et al., 1994b; Wairegi and van Asten, 2011; Parent, 2011; Wairegi and van Asten, 2012). However, there were still some difficulties with CND-*clr* as follows (1) the occurrence of a singular matrix in multivariate analyses computations (due to closure of indices to a zero-sum) made *clr* an inappropriate transformation, (2) the geometric mean of the whole unstructured vector was affected by large variations in micronutrient concentrations due to fungicide applications.

The isometric log-ratio (*ilr*) transformation, on the other hand, generates linearly independent variables computed as structured balances of components or groups of components (Egozcue et al., 2003). To date, CND-*ilr* has been used to classify the nutrient composition of several fruit crops (Parent, 2011; Parent et al., 2012b; Hernandes et al., 2012; Marchand et al., 2013; Parent et al., 2013).

In recent years, little attention has been given to the diagnosis of nutritional disorders in New Zealand kiwifruit (*Actinidia deliciosa* (A. Chv) C.F. Liang et A.R. Ferguson *var deliciosa*) because the kiwifruit industry was initially confined to the Bay of Plenty (New Zealand) deep volcanic ash soils where few obvious nutritional problems were encountered (Smith et al. 1997). With the expansion of production to other soil types and with improved farming systems giving increased productivity, which has generated new problems related to crop nutrition. These coupled with increase environmental awareness has necessitated the development of improved nutrient diagnostic tools for managing kiwifruit orchards.

Our objective was to design a CND-*ilr* method for coherently assessing kiwifruit nutrient status for orchards in the North Island of New Zealand and we expect that the calculated ideals should be approximately correct for kiwifruit grown in other regions.

THEORY

COMPOSITIONAL ANALYSIS

Concentration data belong to the compositional data class, i.e. data that add up to a constant sum such as 1, 100%, 1000 g kg^{-1} or 1,000,000 as ppm. Because they are strictly positive data closed to a bounded space, and exhibit important numerical properties leading to numerical biases in multivariate analysis as follows:

- Redundancy: the amount of one component can be calculated by the difference between the constant scale and the sum of the others, hence there are D-1 degrees of freedom in a D-parts composition and a D-parts composition has rank D-1 (Aitchison and Greenacre, 2002). On the other hand, because any of the $D\times(D-1)/2$ dual ratios between components can be computed from other ratios, they convey redundant information and spurious correlations. These problems can be solved using D-1 variables. Because log ratios are log contrasts, assigning orthogonal coefficients to log ratios provide D-1 orthogonal, i.e. linearly independent, log contrasts.
- Scale dependency: results differ whether concentration data are scaled on dry or wet basis or any nutrient basis forming a stoichiometric rule (e.g. N, P,K, Ca, Mg as in Ingestad et al., 1987). This is an apparent nonsense for a coherent system where parts are interconnected. Because scale change (e.g. between dry and fresh weight basis) is driven by component removal or addition (e.g., water), scale change depends on the way a composition vector is expanded or not. The addition of a component to the composition just provides an additional dimension to that space, not another scale.
- Non-normal distribution: concentration data constrained to closed space (i.e. measurement unit) are only allowed to range between 0 and the unit of measurement according to a log-

logistic distribution (Bacon-Shone, 2011; Parent et al., 2012a). In contrast, a log ratio between two components allows scanning across the unconstrained real space (±∞). As a result, all log-ratio transformed values cannot produce confidence intervals below 0 and the unit of measurement after their back-transformation to original units.

Those three numerical biases are a consequence of inherently closing the compositional space. The closure operator C computes the constant sum assignment as follows (Aitchison, 1986):

$$S^D = C(c_1, c_2, \ldots, c_D) = \left(\frac{c_1 \kappa}{\sum_{i=1}^{D} c_i}, \frac{c_2 \kappa}{\sum_{i=1}^{D} c_i}, \ldots, \frac{c_D \kappa}{\sum_{i=1}^{D} c_i}\right) \qquad \text{Eq. 1}$$

Where κ is the unit of measurement and c_i is the i^{th} part of a composition containing D parts.

When conducting foliar nutrient diagnosis, it is convenient to include a filling value (*Fv*) computed by difference between κ and the sum of all nutrients because the *ilr* values can be back-transformed to the familiar unit of measurement rather than any sum of nutrients. The main components of the filling value are C, O and H, as found in products of photosynthesis.

Because one component of the simplex can be computed by subtracting the sum of the other components from total sum, there are D-1 degrees of freedom in a D-parts composition (Aitchison and Greenacre, 2002).

Log ratios are log contrasts, i.e. log(A/B) = log(A) – log(B), that, after multiplication by orthogonal coefficients, become orthogonal log contrasts, balances or coordinates, with D-1 degrees of freedom.

Balances are binary partitions between components of some whole. The *ilr* technique (Egozcue et al., 2003) allows projecting the simplex S^D into a Euclidean space of D-1 non-overlapping orthogonal log-contrasts between the geometric means of two components or subsets of components. Such balance variables are amenable to multivariate analysis without bias (Filzmoser and Hron, 2011). A system of balances can be formalized in a sequential binary partition (SBP) that arranges elements hierarchically into *ad hoc* functional subsystems. A SBP is a (D-1)×D matrix, in which parts labelled "+1" (group numerator) are contrasted with parts labelled "-1" (group denominator) in each ordered row **(see Table 1 for an example)**. A part labelled "0" is excluded from the balance. The composition is partitioned sequentially at every ordered row into two contrasts until the (+1) and (-1) subsets each contain a single part. Fv (filling value) is the unmeasured elements in the leaf analysis which usually include carbon, hydrogen, oxygen and a small amount of silicon. Balances are computed as follows (Egozcue and Pawlowsky-Glahn, 2005):

$$ilr_j = \sqrt{\frac{n_j^+ n_j^-}{n_j^+ + n_j^-}}\, ln \frac{g(c_j^+)}{g(c_j^-)} \qquad \text{Eq. 2}$$

Where, in the j^{th} row of the SBP, n_j^+ and n_j^- are the numbers of components in the plus (+) or group and the minus (–) or group, respectively, $g(c_j^+)$ is the geometric mean of components in the + (plus) group and $g(c_j^-)$ is the geometric mean of components in the – (minus) group. The natural log of the ratio of geometric means is a log-contrast; the preceding coefficient is the orthogonal coefficient assuring orthogonality between *ilr* coordinates.

Table 1 Sequential Binary Partition

J	N	P	S	Cl	K	Ca	Mg	B	Cu	Zn	Mn	Fe	Fv	n^+	n^-	*Ilr* definition [c_- | c_+]
ilr_1	1	1	1	1	1	1	1	1	1	1	1	1	-1	12	1	[Fv | Fe,Mn,Zn,Cu,B,Mg,Ca,K,P,Cl,S,N]
ilr_2	1	1	1	1	1	1	1	1	-1	-1	-1	-1	0	8	4	[Fe,Mn,Zn,Cu | B,Mg,Ca,K,P,Cl,S,N]
ilr_3	1	1	1	1	1	1	1	-1	0	0	0	0	0	7	1	[B | Mg,Ca,K,P,Cl,S,N]
ilr_4	1	1	1	1	-1	-1	-1	0	0	0	0	0	0	4	3	[Mg,Ca,K | P,Cl,S,N]
ilr_5	1	1	-1	-1	0	0	0	0	0	0	0	0	0	2	2	[Cl,S | P,N]
ilr_6	1	-1	0	0	0	0	0	0	0	0	0	0	0	1	1	[P | N]
ilr_7	0	0	1	-1	0	0	0	0	0	0	0	0	0	1	1	[Cl | S]
ilr_8	0	0	0	0	1	-1	-1	0	0	0	0	0	0	1	2	[Mg,Ca | K]
ilr_9	0	0	0	0	0	1	-1	0	0	0	0	0	0	1	1	[Mg | Ca]
ilr_{10}	0	0	0	0	0	0	0	0	1	1	1	-1	0	3	1	[Fe | Mn,Zn,Cu]
ilr_{11}	0	0	0	0	0	0	0	0	1	1	-1	0	0	2	1	[Mn | Zn,Cu]
ilr_{12}	0	0	0	0	0	0	0	0	1	-1	0	0	0	1	1	[Zn | Cu]

n^+ = number of plus signs and n^- = number of minus signs

The SBP presented in Table 1 was elaborated from current knowledge on nutrient interactions in plants (Bergmann, 1988; Marschner, 1995; Malavolta, 2006). It should be mentioned that, due to orthogonality, the way balances are structured does not influence the results of multivariate linear statistical analyses performed on them. In this paper, balances are conventionally noted as [-1 group | +1 group] because in algebra negative numbers are located on the left side of the zero. As a result, when– group loads more, balance leans to the left and when + group loads more, balance leans to the right. For example, the ilr_6 counterpart of the [P | N] partition is $\sqrt{\frac{1}{2}}\, ln\frac{N}{P}$. As N loads more on this *ilr*, the *ilr* increases in value and the [P | N] partition leans to the right due to heavier weight in the N bucket compared to the P.

Dissimilarities

The distance between an observed leaf nutrient assemblage (ionome) and a known reference can be used as imbalance index. In an orthogonal system of axes, like the ones obtained using the *ilr* transformation, it is possible to compute Euclidean distances. Alternativly, the Mahalanobis distance which is widely used in ecology to compare groups of objects to each other (Legendre and Legendre, 2012) can be used to account for the covariance structure of the nutrient balances. Lovell et al. (2011) showed that Euclidean distances between two compositions are always larger when computed across natural log of components compared to their associated *ilrs*. Parent et al. (2012a) showed that the same reasoning applies to Mahalanobis distances. Numerical biases can thus be revealed by positive shift from *ilr*-based distances to natural log based distances.

Classification of nutrient balances

For diagnostic purposes, there is a need to split the resultss into low- and high-productivity groups. There is also a need for a predictor index that allows separating balanced from unbalanced nutrient signatures. The classification procedure used herein is similar to Cate-Nelson's (Nelson and Anderson, 1984) as improved by Parent et al. (2012a). The Cate-Nelson procedure (see figure 2) is commonly used in soil fertility studies to partition soil test into quadrants to illustrate situations where crops are likely to be responsive or non-responsive to added nutrients. In this paper, we interpreted every quadrant in terms of performance indices commonly used in system diagnosis to distinguish noise from signals (Swets, 1988) according to response and predictor delimiters. Each quadrant defined a response class as follows:

- True positive (TP: nutrient imbalance): low yield crops correctly diagnosed as imbalanced (above critical index). At least one nutrient is imbalanced.
- False positive (FP: type I error): high yield crops incorrectly identified as imbalanced (above predictor critical index). FP observations indicate luxury consumption of nutrients by the plant.
- True negative (TN: nutrient balance): high yield crops correctly diagnosed as balanced (below predictor critical index). The nutrient status of the plant is adequate.
- False negative (FN: type II error): low yield crops incorrectly identified as balanced (below critical index). FN observations indicate the impact of other limiting factors on crop performance.

The performance of the classification was measured by five indices:

- Sensitivity: probability that a low yield is imbalanced, as TP/(TP+FN)
- Specificity: probability that a high yield is balanced, as TN/(TN+FP)
- Positive predictive value (PPV): probability that an imbalance diagnosis returns low yield, as TP/(TP+FP)
- Negative predictive value (NPV): probability that a balance diagnosis returns high yield, as TN/(TN+FN)
- Accuracy: probability that an observation is correctly diagnosed as balanced or imbalanced, as (TP+TN)/(TP+TN+FP+FN)

The classification can be optimized using receiving operating characteristic (ROC) curves (Swets, 1988) (see figure 1a). The predictor delimiter corresponds to a compromise between sensitivity and specificity, (see Figure 1b) where the maximal value of sensitivity × specificity is chosen (the nearest point to the [1,1] top right corner of the sensitivity versus specificity plot). The area under the sensitivity versus specificity curve (AUC) can also be used as an accuracy index for the partition (Swets, 1988).

Because crop yield is a continuous variable, a procedure is needed to optimize the response delimiter. In our survey datasets, true negative (TN) specimens represent the reference population. The Mahalanobis distance can be used as strong nutrient response predictor, as computed between observations and the barycentre of the TNs according to the covariance structure of TNs. Because TNs are not defined a priori (assumed value) , an iteration procedure is needed to identify appropriate values. For a given response (crop yield) delimiter, the predictor is initiated using high-yielders as reference points for comparison and thereafter, a predictor value is selected. The delineated TN specimens are then used as the reference population for the computation of the Mahalanobis distance. The TN specimens are resampled and the Mahalanobis distance is recomputed iteratively until two iterations classify observations identically. The Moore-Penrose pseudo-inversion was used to avoid singularities in the inversion of the covariance matrix (Prekopcsák and Lemire, 2012). Procedures to optimize response and predictor delimiters are developed in the Material and methods section.

MATERIAL AND METHODS

DATA SET

The metafile comprised 431 observations with plant compositions and metadata collected in commercial kiwifruit orchards ('Hayward') in the North Island of New-Zealand across two farming systems (organic and conventional) and several soil types. From 2003 to 2010, 2 to 3 recently matured, fully expanded leaves taken from the 2nd lateral of 32 vines (excluding young vines and sick leaves) were taken within 0 to 4 weeks after flowering. The criterion of crop performance were yield (kg·ha^{-1} of grade one export quality fruit). Because of the 2011 outbreak of PSA (*Pseudomonus Syringii actinidia*), copper sprays are today widely used. However, the present dataset is pre-PSA and therefore only includes orchards with zero or low use of fungicide following approved Global-gap and Certified Organic practices (when applicable).

FOLIAR ANALYSIS

The P, K, Ca, Mg, Zn, Cu, Mn, Fe, and B concentrations were determined in plant tissues by IPC-OES after microwave digestion (Blackmore et al., 1987). Chlorine was extracted with 0.01 M $CaSO_4$ and quantified by ion chromatography. Total N and S were determined by dry combustion using a Leco 2000 CNS analyser.

OPTIMIZATION OF DELIMITERS IN BINARY CLASSIFICATION

As mentioned in the theory section, there is a need for an iterative procedure to define predictors because in survey analyses the definition of the reference population depends on the Mahalanobis distance computed from the reference population ($\mathcal{M}_{TN}$). The iterative procedure is described as follow.

- For a given response (crop yield) delimiter, the predictor is initiated using high-yielders as reference specimens for computing $\mathcal{M}_{HY}$.
- A predictor delimiter is selected and its barycenter and covariance are computed among newly delineated TN specimens to solve $\mathcal{M}_{TN}$.
- The $\mathcal{M}_{TN}$ is iterated until two iterations classifies observations identically.

The Moore-Penrose pseudo-inversion was used to avoid singularities in the inversion of the covariance matrix (Prekopcsák and Lemire, 2012).

In a scatter of n observations, there are n possible response delimiters and n possible predictor delimiters, resulting in n×n possible binary classifications. The yield response delimiter returning the largest area under the sensitivity versus specificity curve (AUC) was selected.

STATISTICAL ANALYSIS

Statistical computations were conducted in the R statistical environment (R Development Core Team 2012). The "compositions" package (van den Boogaart et al., 2013) was used for *ilr* computations. The "robCompositions" (Templ et al., 2013) package was used to robustly impute missing concentrations values. The "mvoutlier" package (Filzmoser and Gschwandtner, 2013) was used to detect outliers across the dataset. Discriminent analyses were performed using the "ade4" package (Chessel et al., 2012).

RESULTS

DATA PRE-TREATMENT

A number of 17 missing values had to be imputed, essentially Cl concentrations. The outlier detection procedure eliminated 18 observations in the data set, leaving 413 observations for further analyses.

CLASSIFYING OBSERVATIONS

The area under the ROC curve (AUC) reached a peak at 0.92, corresponding to a yield delimiter of 46 339 kg ha^{-1} (**Error! Reference source not found.**a).

The ROC curve (sensitivity versus specificity relationship) obtained a realisable high yield optimum value of 46 339 kg ha^{-1} is shown by the thick black line in **Error! Reference source not found.**b.

The ROC curve did not show a regular decrease of sensitivity as specificity increased, as usually observed in ROC diagnoses. This phenomenon is due to the re-sampling procedure (see methodology section) that is generally not needed in conventional clinical studies.

The closest point to the [1,1] corner corresponded to a specificity of 100%, a sensitivity of 91% and a Mahalanobis distance ($\mathcal{M}_{TN}$) of 4.21. This critical $\mathcal{M}_{TN}$ was used as predictor delimiter between balanced and unbalanced specimens.

Figure 1a Receiving Operating Characteristic Curve

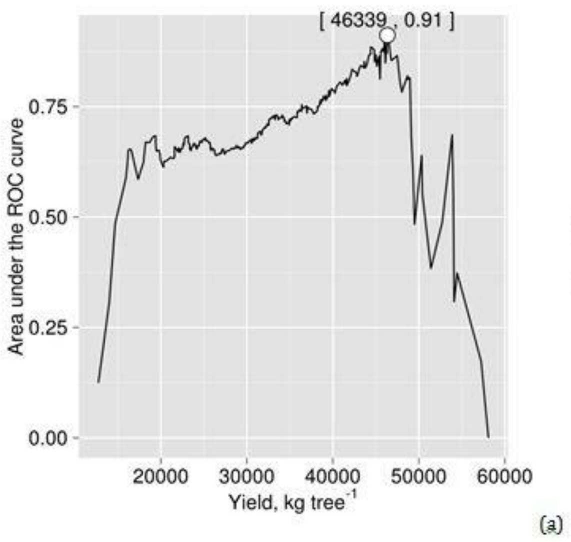


Figure 1b Sensitivity Vs Specificity

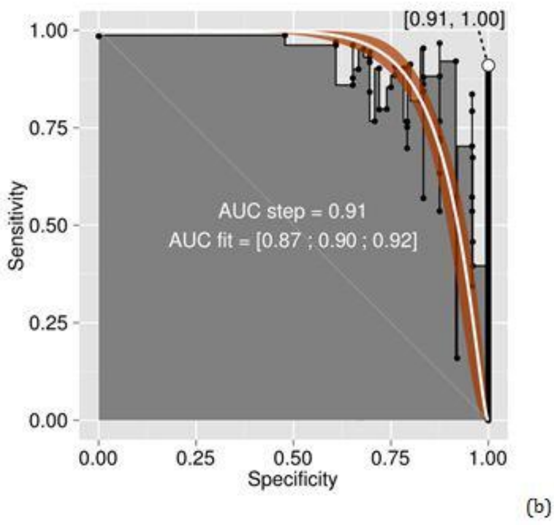


Figure 2 shows data partitioned into four quadrants. The semi-transparent ellipse encloses 95% of the theoretical distribution of points. Statistics corresponding to this partition are presented in Table 2.

All specimens declared imbalanced yielded less than cut-off yield value (PPV = 100%). On the other hand, nearly 2 balanced specimens out of 5 yielded more than cut-off yield value (NPV=42%).

Figure 2 Cate Nelson plot of Kiwifruit Ionome Mahalanobis distance Vs Productivity

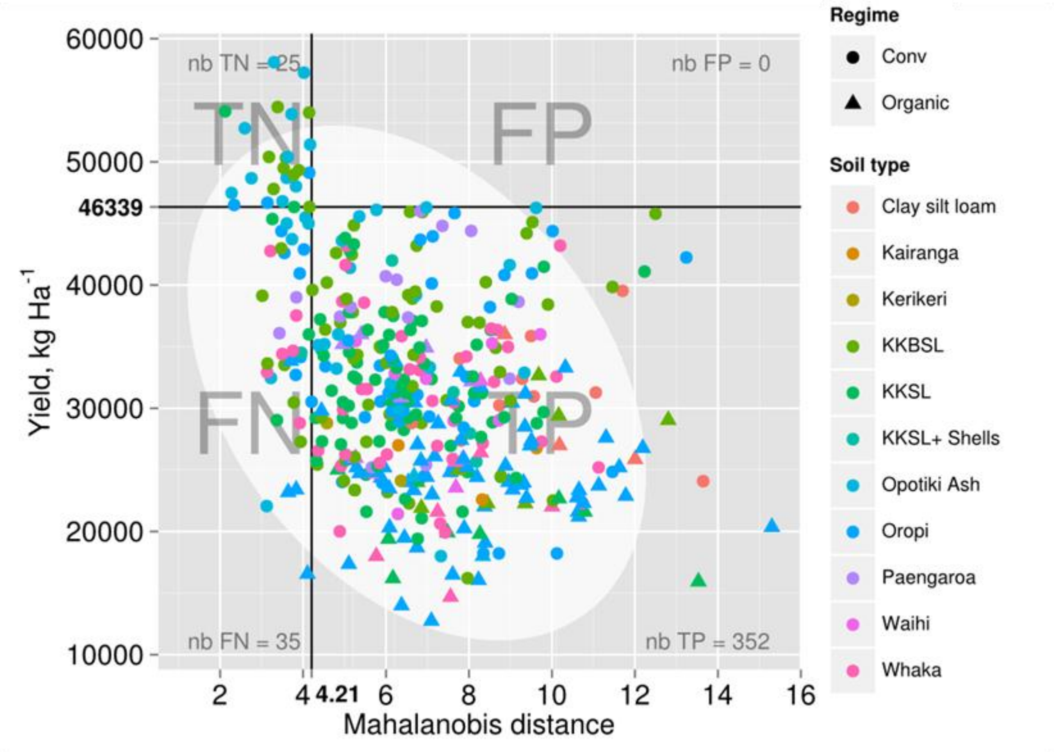


TABLE 2. DATA CLASSIFICATION STATISTICS

Predictor (Mahalanobis distance)	4.21
Response (kg ha^{-1})	46 339
TN counts	25
FN counts	35
TP counts	354
FP counts	0
Sensitivity	91%
Specificity	100%
PPV	100%
NPV	42%
Accuracy	92%

The contingency bar plot in

Figure 33 shows counts in quadrants containing observations (TN, FN and TP) by soil type (see table 5) and farming system (Conventional Vs Organic).

There were 25 high-yielding observations (TN quadrant), all grown under conventional farming.

Most true high yielders (TN) were grown in the Opotiki ash and KKBSL soil types (20/25).

On the other hand, 41% of the true low-yielders (TP) were grown in the KKSL soil type under conventional practices (72/352) and Oropi soil type under organic practices (64/352).

Most false low-yielders (FN) were found under conventional practices in the Oropi (7/35), as well as KKBSL , Opotiki ash (6/35) and Whaka (6/35) soil types.

FIGURE 3.
CONTINGENCY BAR PLOT OF COUNTS IN EACH QUADRANT (TN, FN, TP, FP) BY SOIL TYPE AND FARMING SYSTEM.

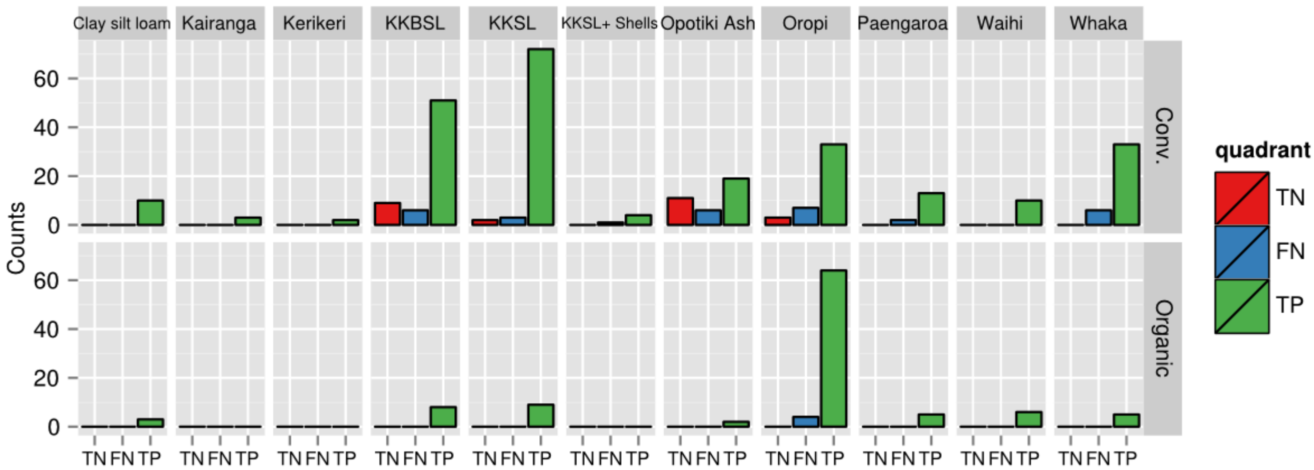


TABLE 5.
BRIEF DESCRIPTION OF SOIL TYPES ASSOCIATED WITH THIS DATA SET

Site	# obs.	Soil type
Clay silt loam	21	Fluvent derived mainly from shale (Greywake) alluvial deposits, drained by ditches; flooding is now rare.
Kairanga	3	Fluvent derived mainly from Rhyolite alluvial deposits; naturally deficient in boron.
Kerikeri	2	Udand derived from older Basaltic volcanic activity; gravelly coarse soil naturally high in iron.
KKBSL	75	Katikati black sandy loam; andesitic Udand middle aged, moderately high in aluminum with aeolian volcanic ash content.
KKSL	85	Katikati sandy loam, same andesitic Udand as KKBSL except it has never been under water in the harbor.
KKSL + shells	4	Fine textured andesitic Udand as KKBSL except it contains sea shells.
Opotiki Ash	38	Udand, fine textured mixed andesitic and rhyolitic ash receiving less rainfall than other Udand soils.
Oropi	104	Intermediate Udand /vitrand, fine textured andesitic and rhyolitic ash.
Paengaroa	20	Very young soils Vitrand mixed textured rhyolitic with Pongakawa soil inclusions that tend to be more pumice rich.
Waihi	16	Mature andesitic ash Udand, fine textured.
Whaka	46	Whakamarama, mature fne textured andesitic ash Udand, more heavily leached than other soils.

Median *ilr* values of TN specimens, associated with their confidence intervals, are presented in Table 1 below.

TABLE 1.
CONFIDENCE INTERVALS OF ILR VALUES ($\pm t_{0.025}$) FOR TRUE NEGATIVE (TN) SPECIMENS (N = 33) IN THE NEW ZEALAND KIWIFRUIT DATA SET (LL = LOWER LIMIT; UL = UPPER LIMIT).

Ilr definition	TN			TP		
	LL	Median	UL	LL	Median	UL
[Fv | Elements]	-6.683	-6.654	-6.626	-6.677	-6.666	-6.656
[Oligo | B+Macro]	7.979	8.070	8.161	7.974	7.996	8.019
[B | Macro]	4.946	4.989	5.032	4.934	4.949	4.963
[Mg,Ca,K | P,Cl,S,N]	-1.044	-0.980	-0.916	-0.995	-0.978	-0.961
[Cl,S | P,N]	0.372	0.438	0.504	0.443	0.460	0.476
[P | N]	1.618	1.661	1.704	1.594	1.607	1.620
[Cl | S]	-0.418	-0.361	-0.304	-0.315	-0.298	-0.282
[Mg,Ca | K]	0.735	0.786	0.838	0.759	0.775	0.791
[Mg | Ca]	1.316	1.352	1.388	1.387	1.398	1.409
[Fe | Mn,Zn,Cu]	-0.807	-0.743	-0.678	-0.689	-0.670	-0.652
[Mn | Zn,Cu]	-1.383	-1.313	-1.243	-1.304	-1.280	-1.257
[Zn | Cu]	-0.724	-0.663	-0.603	-0.707	-0.683	-0.658

NOTE THAT THE BALANCES ARE CONVENTIONALLY NOTED AS [-1 GROUP | +1 GROUP].

NUTRIENT BALANCE COMPARISONS BETWEEN TN AND TP SPECIMENS

Tukey's test allowed detecting in which balance significant differences occurred between TN and TP specimens (

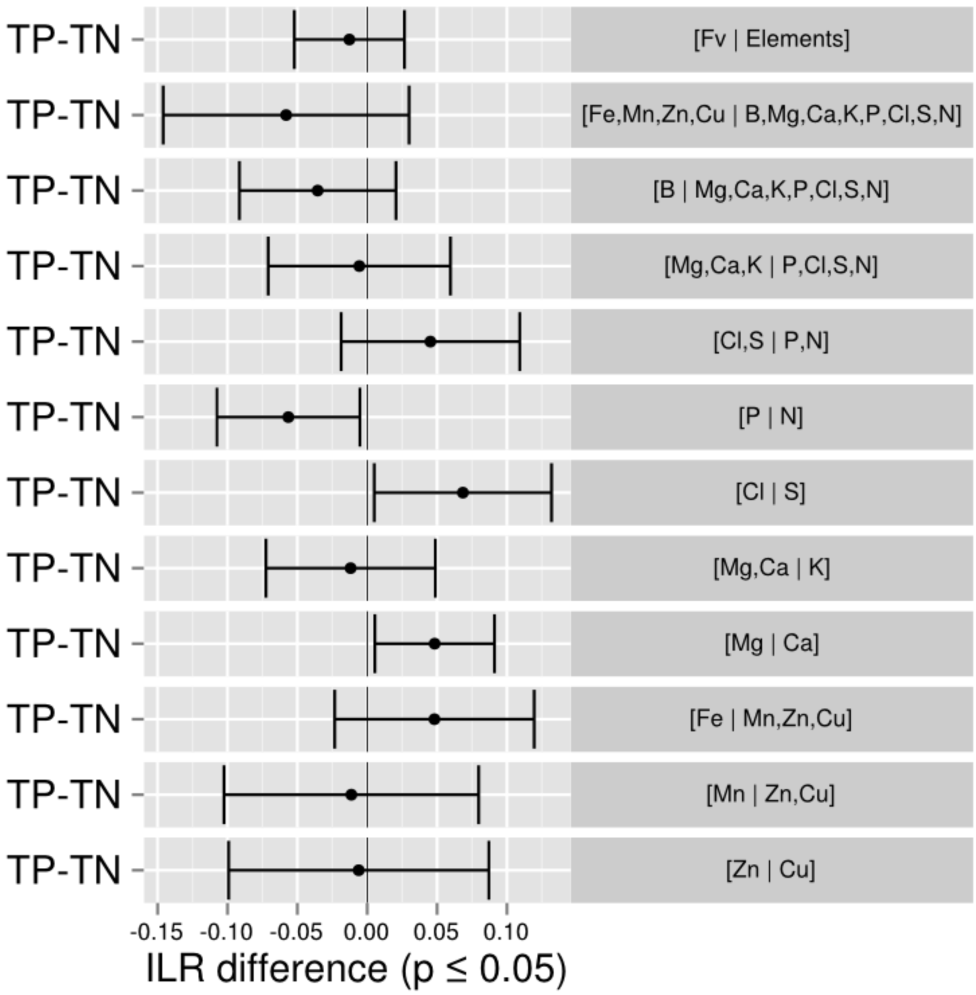
Figure **4**).

The [P | N], balance showed a (TP-TN) difference significantly lower than 0, because N was exceedingly larger than P in TN specimens.

The [Cl | S] balance was significantly higher in TP specimens, indicating higher S over Cl log-ratio in TP specimens.

Finally, there was a significant difference in the [Mg | Ca] balance, where the weight of Mg Ca was significantly greater in TN specimens. Differences are significant ($P < 0.05$) where they do not overlap zero.

Figure 4. Tukey test (P = 0.05) for *ilr* differences between true positive (TP) and true negative (TN).



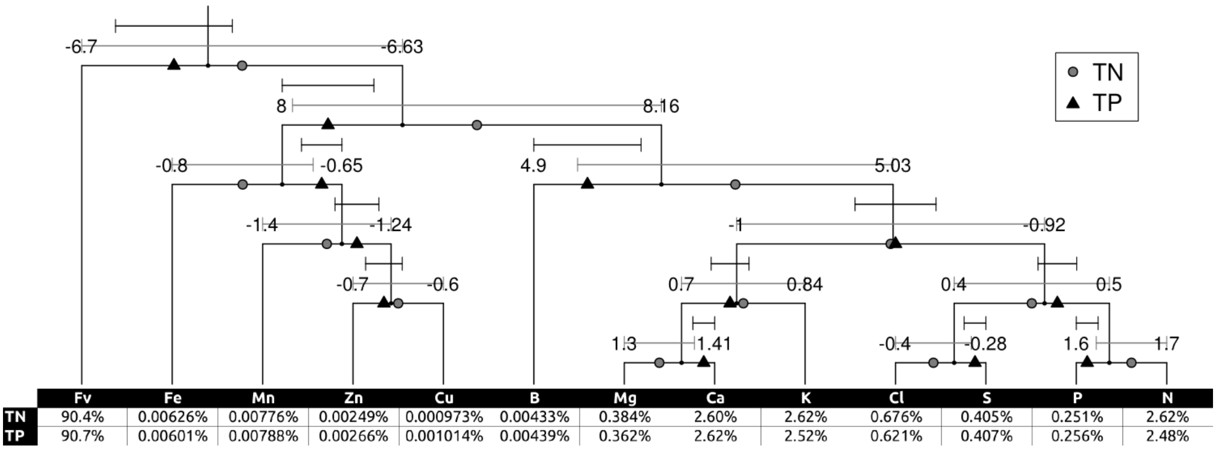
NOTE THAT THE BALANCES ARE CONVENTIONALLY NOTED AS [-1 GROUP | +1 GROUP].

PAN BALANCE REPRESENTATION

Balances can be represented metaphorically using a stand-alone mobile diagram with fulcrums and weighing pans, where changing nutrient concentrations or contents in buckets impact directly on nutrient balances at fulcrums. **Error! Reference source not found.**5 presents a balance dendrogram derived from the SBP with TN and TP median *ilr* values at fulcrums as well as their confidence intervals. The average *ilr* values at fulcrums are used for diagnostic purposes while the back-transformed TN *ilr* values to concentrations are laid down in weighing pans to support the interpretation of balances in terms of relative shortage, sufficiency or excess of contributing nutrients. However, the analyst should be reminded that sufficiency or excess of any nutrient may only be diagnosed in relation to other nutrients in the balance system. The weighing pans facilitate adjusting the balances correctly (through fertilisation) by shifting the fulcrums towards the TN depicted as a grey circle. For example, the lower the [Mg | Ca] balance in TN specimens can be interpreted as a combination of lower Ca and higher Mg concentrations compared to TP specimens. In general, differences between TN and TP were relatively small in terms of concentrations. Overall, the pan balance representation of the ionomes showed that N, Cl and Mg concentrations were lower in TP compared to the TN specimens.

FIGURE 5.

MOBILE-AND-FULCRUMS SCHEME OF A BALANCE SYSTEM FOR THE KIWIFRUIT IONOME.



AVERAGE *ILR* VALUES ACROSS SPECIMENS ARE LOCATED AT FULCRUMS. DEPARTURE FROM TN RANGE INDICATES RELATIVE NUTRIENT IMBALANCE. CONCENTRATIONS LOCATED IN WEIGHING PANS ARE BACK-TRANSFORMED AVERAGE *ILR* VALUES FOR TN AND TP SPECIMENS. NOTE THAT THE BALANCES ARE CONVENTIONALLY NOTED AS [-1 GROUP | +1 GROUP].

INFLUENCE OF PRODUCTION SYSTEM ON NUTRIENT BALANCE IN KIWIFRUIT

Discriminant analyses performed across *ilr* balances indicated significant differences between score means (boxes) of farming systems (6) and soil types (7).

FIGURE 6.

DIFFERENCE IN IONOME BETWEEN FARMING SYSTEMS.

CONFIDENCE INTERVALS (P=0.05) OF SINGLE AXIS DISCRIMINANT SCORES ABOUT POPULATION (THIN LINE – STANDARD DEVIATION) AND ABOUT MEAN (THICK LINE – STANDARD ERROR) FOR TWO FARMING SYSTEMS.

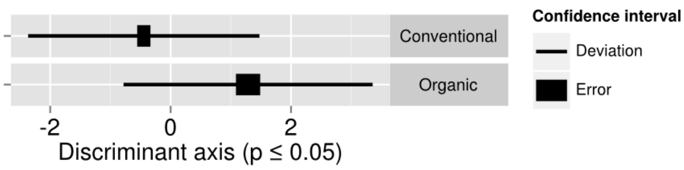


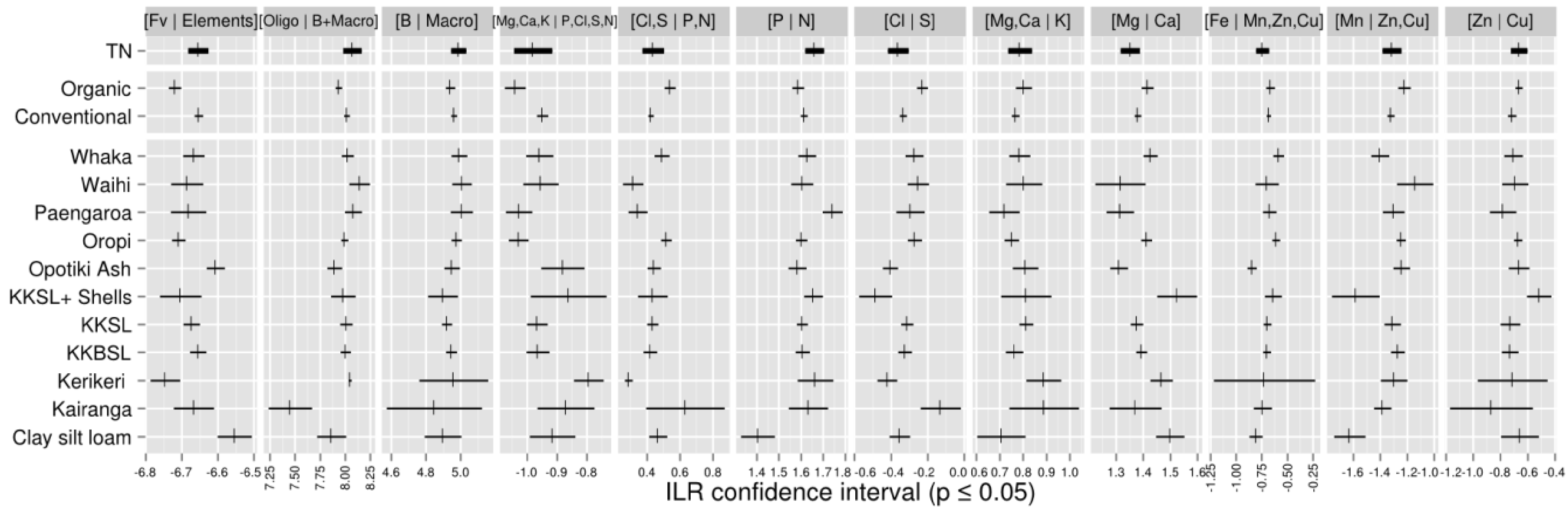
Figure 1.

Confidence intervals (p=0.05) for *ilr* nutrient balances for soil types and farming regimes.

Figure 1 presents the confidence interval (p=0.05) about the mean of each balance for both farming systems compared to the TN specimens.

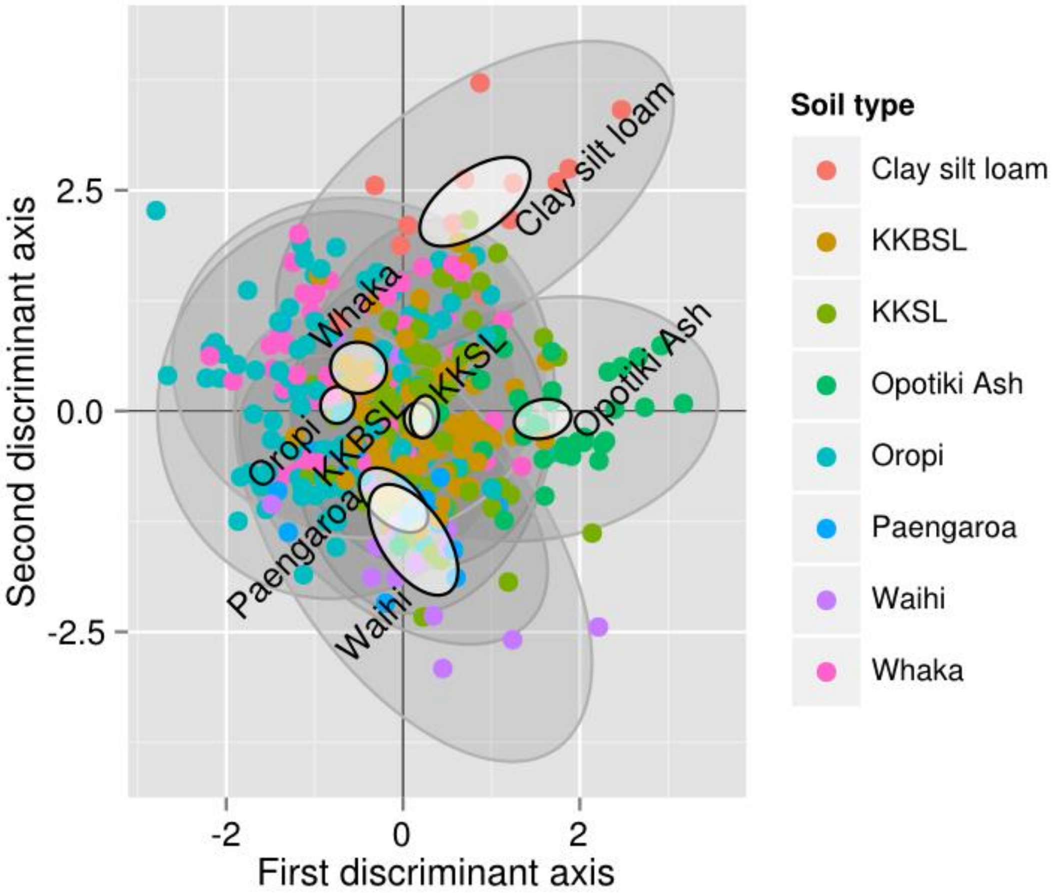
Organic farming resulted in less nutrients accumulations than the filling value compared to conventional farming and TN. The dissimilarity is attributable to micronutrients, as shown by the significant difference in the [Cationic micronutrients | Macronutrients+B] balance. Other significant differences between organic farming and TNs were found in the [Cl,S | P,N], [P | N], [Cl | S] and [Mg | Ca] balances. Overall, all balances in conventional farming practices did not differ significantly from TN, compared to half of them for organic farming practices.

Figure 1 also showed the state of imbalance for each soil type compared to TN. Notably, soils related to the highest number of balances that do not differ significantly from TN were KKSL, KKBSL, Paengaroa and Waihi (12/12), Opotiki flats (11/12) as well as Kairanga and KKSL+Shells (10/12). On the other hand,



Clay silt loam soils and Kerikeri showed the largest number of balances that differed significantly from TN (4/12), followed by Oropi (3/12).

Figure 8.
DISCRIMINANT ANALYSIS OF NUTRIENT *ILR* BALANCES BY SOIL TYPE



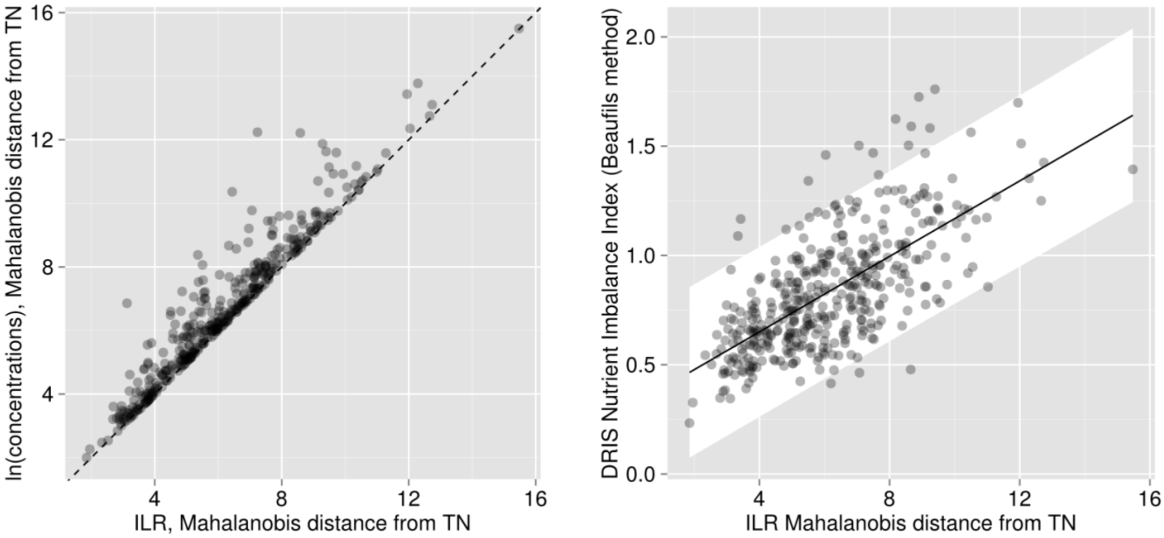
Large semitransparent ellipses that enclose swarms of data points represent regions that include 95% of the theoretical distribution of canonical scores for each soil type. Smaller plain white ellipses represent confidence regions about means of canonical scores at 95% confidence level. The optimum TN point is where the zero lines cross.

The discriminant analysis performed across soil types (Figure 8) distinguished four groups. One group comprised KKSBL and KKSL soils, another the Waihi and the Paengoroa soils, then Oropi and Whaka soils and, finally the Clay silt loam and the Opotiki ash. Confidence intervals (p=0.05) about the mean of each balance across soil types

Numerical biases in concentration values and DRIS indices

In CND-*ilr*, there is no conflicting interpretation of nutrient levels and balances as could be the case when interpreting the results of critical value and DRIS diagnoses separately. Numerical biases of critical ilr concentration ranges (a) and DRIS (b) are shown in figure 9a & 9b as departure from unbiased Mahalanobis distance using *ilr* coordinates.

FIGURE 9.
NUMERICAL BIASES IN ILR CONCENTRAITON VALUES AND DRIS INDICES.



Discussion

Classification

The area under the ROC curve (AUC) of 0.91 (**Error! Reference source not found.**a) is comparable to the AUC for fairly informative tests (0.80-0.98) in medical sciences (Swets, 1988). The accuracy of 92% was comparable to values > 80-90% reported by Baxter et al. (2008) in plant nutrition. Values projected in the FN quadrant could have been partially hidden by factors external to the ionome (e.g. climate, diseases, etc.), while the three observations projected into the FP quadrant may have been cases of luxury consumption.

Pan balance diagnosis

Walworth and Sumner (1987) and Marschner (1995) argued that optimal ratios between nutrients are insufficient criteria for diagnostic purposes, because it would be impossible to determine whether a nutrient level is too high (excessive), adequate (sufficient) or too low (deficient) in the ratio. Indeed, although concentrations and ratios (or balances) portray the same status, they should not be interpreted separately as commonly done, possibly leading to conflicting interpretation. This problem of interpretation is solved easily by the pan balance as metaphor for coherent concepts relating nutrients and balances to each other. The design of the balance system can be derived from plant physiology, soil biogeochemistry or crop management to facilitate interpretations (**Error! Reference source not found.**5). The mean *ilr* values of TN specimens back-transformed to concentrations in weighing pans allow interpreting the analytical results of specimens in relative shortage, sufficiency or excess, as already diagnosed by *ilr* and the Mahalanobis distance. However, relative shortage, sufficiency or excess diagnoses should be based on balances rather than concentrations alone.

As a result, any deviation in concentration from the ones indicated in weighing pans must affect balances directly, hence avoiding misinterpreting diagnoses conducted independently on concentrations and ratios. In the present study, , for a sample mapped at the TP mean (Figure 12), a shift of the [Mg | Ca] fulcrum to the left by adding more Mg should rebalance the cationic balances, but at risk to misbalance [Mg,Ca,K,P,Cl,S,N | B] in this complex system. Shifted balances should thus be monitored regularly for possible adjustment. In the same perspective, increasing Cl could shift [Cl | S] to the left, [P,N | Cl,S] to the right and so on for the higher-level balances in the hierarchy of the balance system. On the other hand, increasing N could not only shift [P | N] to the right, but also [Cl,S | P,N], which might be already too far in the right direction.

Proper N and Cl management is central to the kiwifruit production. Hasey et al. (1997) found that foliar N was lower while leaf Cl and Na were higher in organic orchards although all nutrient levels were within acceptable concentration ranges. However, high N may increase vine yield and average fruit weight and produce higher proportions of over-ripened and rotten fruit at harvest (Tagliavini et al., 1995; Costa et al., 1997), or may show no effect on fruit yield, size or fruit quality at harvest and be associated with fruit softness during storage (Johnson et al., 1997). In fertilizer experiments, Prasad et al. (1993) found foliar Cl ranging between 0.6 and 2.1% with toxic threshold at 1.5% in five 'Hayward' kiwifruit orchards on the North Island of New Zealand. We found that high yield of kiwifruit was associated with average Cl concentration of 0.71% in TNs compared to 0.61% in TPs.

The pan balance model provides an overall view on how nutrient additions may contribute rebalancing nutrient relationships in kiwifruit orchards. The approach is intuitive and applicable to any natural system. Numerical biases explain why the critical concentration ranges that do not account for nutrient interactions and DRIS that has unstructured geometry often produce conflicting diagnoses. In contrast, the nutrient balance concept is a stand-alone diagnostic system of linearly independent and organically linked variables within the same setup where the analytical results and their balances can be interpreted coherently.

INFLUENCE OF PRODUCTION SYSTEM ON NUTRIENT BALANCE IN KIWIFRUIT

Discriminant analyses confirmed that expanding the kiwifruit production to other soil types and farming systems increased productivity problems related to crop nutrition. Results reflected the large differences between agro-ecosystems and could be used to identify in what direction efforts should be directed to alleviate nutrient imbalance in kiwifruit orchards. Agro-ecosystems under organic practices were less productive compared to conventional farming systems, likely due to nutrient imbalance. In our dataset, equilibrating the nutrient status of these organic orchards might be performed by increasing the Cl content in leaves, such that balance [Cl | S] is decreased to reach the associated TN range. On the other hand, sufficiently increasing Mg content could rebalance [Mg | Ca] in TP specimens and reach the TN range. The addition of N would help rebalancing [P | N] in the TPs. The addition of Cl, Mg and N would also shift the [B | Mg,Ca,K,Cl,S,P,N] balance to the right, and this should be monitored.

Because most organic orchards are grown on Oropi soil, it was thus not possible to attribute imbalance to farming system or soil type.

A key factor that will have caused differences in ilr confidence intervals (Figures 6 & 7) is because bud break enhancers like HiCane™ is routinely used in New Zealand conventional orchards but it is prohibited under organic certification, thus organic orchards produced less fruit for 2 reasons.

1. On organic orchards it is normal that a less number of buds break dormancy (about 4%) than those grown with conventional methods.
2. Where bud break enhancers are used the bud break is advanced by about four weeks giving the conventional vines the advantage of a longer growing season

In order to compensate for later maturing leaves in the organic orchards the mean leaf sampling date is about four weeks after the conventional orchards. Because certain nutrient concentrations tend to increase whilst others decrease as spring progresses the ionome will appear to be different according to regime unless week number is compensated for.

The KKSL and KKBSL soil group agro-ecosystems, where kiwifruit was traditionally grown, met all nutrient balance requirements for producing high yield of kiwifruits (

Figure 1). The KKBSL and KKSL soils both contain ash from the Mayor Island eruption. The KKBSL is nearly all adjacent to Tauranga Harbour and is probably uplifted harbour floor. The region where these two soils occur consistently suffer from very strong westerly spring winds coming from over the Kaimai range, which does much damage to the young growing shoots reducing the yield potential significantly. Conversely Opotiki is well known to have high sunshine hours over the course of the growing season to the benefit of fruit production. Because of these confounding differences of micro-climate it was not possible to determine to a significant level whether nutrient imbalance and low yield correspond to soil type.

To the best of our knowledge DRIS system has never previously been developed for kiwifruit, however after a consideration of figures 9a & 9b it is quite apparent that the development of ilr balanced nutrient diagnoses is a significant improvement over the DRIS system and is a huge leap forward from the previous CNR method.

CONCLUSION

This paper presents a novel stand-alone balance approach to diagnose nutrient imbalance in kiwifruit orchards. The pan balance model differs markedly from the traditional critical nutrient range approach illustrated by Liebig's barrel and from the DRIS; when conducted separately, the critical value approach and DRIS may yield conflicting results. The *ilr* concept reflects balances between two or more nutrients that facilitate interpreting the diagnosis. The diagnosis is conducted in three steps: 1) compute the Mahalanobis distance as a measure of general nutrient imbalance; 2) in case of imbalance, select the balances differing significantly from TN specimens; 3) diagnose concentration values in terms of relative shortage, sufficiency or excess.

The nutrient pan balance model is intuitive and coherent and allows nutrient balances and concentrations to be interpreted simultaneously, hence avoiding numerical biases and conflicting interpretations. We found that kiwifruit nutrition varied widely in New Zealand and that Black Sandy Loam soil type was the most properly balanced agro-ecosystems.

ACKNOWLEDGEMENTS

The authors acknowledge the financial support of the Natural Sciences and Engineering Council of Canada (NSERC-DG 2254 and CRDPJ 385199 - 09). We thank Alan McCurran and Amelia Barlow for data collection.

REFERENCES

Aitchison, J. 1986. The statistical analysis of compositional data. Chapman and Hall, London, 416 pp.

Aitchison, J., and M. Greenacre. 2002. Biplots of compositional data. J. Royal Stat. Soc. Series C (Appl. Stat.) 51(4):375-392.

Bacon-Shone, J. 2011. A short history of compositional data analysis. p. 3-11 in Compositional data analysis: Theory and Applications, Pawlowsky-Glahn V., and Buccianti A. (eds), John Wiley and Sons, NY.

Baxter, I. R., Vitek, O., Lahner, B., Muthukumar, B., Borghi, M., Morrissey, J., Guerinot, M. L., and Salt, D. E. 2008. The leaf ionome as a multivariable system to detect a plant's physiological status. PNAS 105 :12081–12086.

Beaufils, E.R. 1973. Diagnosis and Recommendation Integrated System (DRIS). Soil Sci. Bull. Vol. 1, p. 1-132

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

Blackmore, L.C, Searle, P.L., and Daly, B.K. 1987. Methods for chemical analysis of soils. NZ Soil Bureau Scientific Report 80. NZ Soil Bureau, DSIR.

Blanco-Macías, F., Magallanes-Quintanar, R., Valdez-Cepeda, R.D., Vázquez-Alvarado, R., Olivares-Sáenz, E., Gutiérrez-Ornelas, E., and Vidales-Contreras, J.A. 2009. Comparison between CND norms and boundary-line approach nutrient standards: Opuntia ficus indica L. case. Revista Chapingo Serie Horticultura 15(2): 217-223.

Camacho, M.A., Silveira, M.V. da, Camargo, R.A., and Natale, W. 2012. Normal nutrient ranges by the Chm, DRIS and CND methods and critical level by method of the reduced normal distribution for orange Pera. Bras. Ci. Solo, 36:193-200.

Chessel, D., Dufour, A.-B. and Dray, S. 2012. ade4: Analysis of Ecological Data : Exploratory and Euclidean methods. R package version 1.5-1. http://CRAN.R-project.org/package=ade4

Costa, G., Lain, O., Vizzotto, G., and Johnson, S. 1997. The effect of nitrogen fertilization on the fruiting and vegetative performance, fruit quality and post-harvest life of kiwifruit cv Hayward. Acta Hort. 444: 279-284.

Egozcue, J.J. and Pawlowsky-Glahn, V. 2005. Groups of parts and their balances in compositional data analysis. Math. Geol. 37: 795-828.

Egozcue, J. J., Pawlowsky-Glahn, V., Mateu-Figueras, G., and Barceló-Vidal, C. 2003. Isometric log-ratio transformations for compositional data analysis. Math. Geol. 35, 279-300.

Filzmoser, P. and Gschwandtner, M. 2013. mvoutlier: Multivariate outlier detection based on robust methods. R package version 1.9.9. http://CRAN.R-project.org/package=mvoutlier

Filzmoser, P. and Hron, K. 2011. Robust statistical analysis. p. 57-72n Compositional data analysis: Theory and Applications, Pawlowsky-Glahn V., and Buccianti A. (eds), John Wiley and Sons, NY.

Hasey, J.K., Johnson, R.S., Meyer, R.D., and Klonsky, K. 1997. An organic versus a conventional farming system in kiwifruit. Acta Hort. 444(1): 223-228.

Hernandes, A., Parent, S. E., Natale, W., and Parent, L. E. 2012. Balancing guava nutrition with fertilization and liming. Rev. Bras. Frutic., 34(4): 224-1234.

Huang, H., Xiao Hu, C., Tan, Q., Hu, X., Sun X., and Bi, L. 2012. Effects of Fe–EDDHA application on iron chlorosis of citrus trees and comparison of evaluations on nutrient balance with three approaches. Scientia Hort. 146:137–142.

Johnson, R.S., Mitchell, F.G., Crisosto, C.H., Olsen, W.H., and Costa, G. 1997. Nitrogen influences kiwifruit storage life. Acta Horticulturae 444: 285-289.

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, Soil Science Society of America, Madison, WI.

Legendre, P. and Legendre, L. 2012. Numerical Ecology – 3rd Ed. Elsevier. 1006 p.

Lovell, D., Müller, W., Taylor, J., Zwart , A., and Helliwell, C. 2011. Proportions, percentages, ppm: do the molecular biosciences treat compositional data right? p. 193-207 in: Pawlowsky-Glahn V. and Buccianti A. (eds), Compositional Data Analysis: Theory and Applications, John Wiley and Sons, NY.

Malavolta, E. Manual de nutrição de plantas. 2006. Pav. Chimica, ESALQ and Ed. Agron. CERES, São Paulo, Brazil, 631 p..

Marchand, S., Parent, S.-É., Deland, J.P., and Parent, L. E. 2013. Nutrient signature of Quebec (Canada) cranberry (Vaccinium macrocarpon Ait.). Rev. Bras. Frutic., 35(1) (in press)

Marschner, H. 1995. Mineral nutrition of higher plants. Academic Press, NY, 674 p.

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

Parent, LE. 2011. Diagnosis of the nutrient compositional space of fruit crops. Revi. Bras. Frut.33: 321-334.

Parent, L.E., Cambouris, A.N., and Muhawenimana, A. 1994a. Multivariate Diagnosis of Nutrient Imbalance in Potato Crops. Sci. Soc. Am. J. 58:1432-1438.

Parent, L.E., Isfan, D., Tremblay, N., and Karam, A. 1994b. Multivariate Nutrient Diagnosis of the Carrot Crop. J. Amer. Soc. Hort. Sci. 119(3):420–426.

Parent, L.E. and Dafir, M. 1992. A theoretical concept of compositional nutrient diagnosis. J. Amer. Soc. Hort. Sci. 117: 239-242.

Parent, L.E., Karam, A., and Visser, S.A. 1993. Compositional Nutrient Diagnosis of the Greenhouse Tomato. HortScience 28(10): 1041–1042.

Parent S.-É., Parent, L. E. Rozane, D. E. Hernandes, A. and Natale, W. 2012a. Nutrient balance as paradigm of plant and soil chemometrics. p. 83-114 in Issaka, R.N. (ed.) Soil Fertility, InTech Publ., http://www.intechopen.com/books/soil-fertility.

Parent, L.E., Parent, S.É., Rozanne, D.E., Amorim, D.A., Hernandes, A., and Natale, W. 2012b. Unbiased approach to diagnose the nutrient status of guava. Proc. 3rd Int. Symp. on Guava and Other Myrtaceae. ISHS 2012. Eds.: C.A.F. Santos et al. Acta Hort. 959:145-159.

Parent, S.-É., Parent, L.E., Egozcue, J.J., Rozane, D.E., Hernandes, A., Lapointe, L., Hebert Gentile, V., Naess, K., Marchand, S., Lafond, J., Mattos Jr, D., Barlow, P., and Natale, W. 2013. The plant ionome revisited by the nutrient balance concept. Front. Plant Sci. 4:39. doi: 10.3389/fpls.2013.00039

Prasad, M., Burge, G.K., Spiers, T.M. and Fietje, G. 1993. Chloride induced leaf breakdown in kiwifruit. J. Plant Nutr. 16: 999-1012.

Prekopcsák, Z. and Lemire, D. 2012. Time series classification by class-specific Mahalanobis distance measures. Advances in Data Analysis and Classification, 6(3):185-200.

R Development Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing Version 2.15.1; 2012. http://www.R-project.org (accessed 14 March 2013).

Serra, A.P., Marchetti, M.E., Vitorino, A.C.V.T., Novelino, J.O., and Camacho, M.A. 2010. Determination of normal nutrient ranges for cotton by the ChM, CND and DRIS methods. R. Bras. Ci. Solo, 34:105-113.

Silva, G.G.C. da, Neves, J.C.L., Alvarez V.V.H., and Leite, F.P. 2004. Nutritional diagnosis for Euvcalypt by DRIS, M-DRIS , and CND. Sci. Agric. (Piracicaba, Braz.) 61(5):507-515.

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

Swets, J., 1988. Measuring the accuracy of diagnostic systems. Science, 240 :1285–1293.

Tagliavini, M., Toselli, M., Marangoni, B., Stampi, G., Ravenna, C. A. P.,and Pelliconi, F. 1995. Nutritional status of kiwifruit affects yield and fruit storage. Acta Hort. 383:227-237.

Templ, M., Hron, K., and Filzmoser, P. 2013. robCompositions: Robust Estimation for Compositional Data. R package version 1.6.3. http://CRAN.R-project.org/package=robCompositions

Urano, E.O.M., Kurihara, C.H., Maeda, S., Vitorino, A.C.T., Manoel Carlos Gonçalves, M.V., and Marchetti, M.E. 2007. Determination of optimal nutrient contents for soybean by the mathematical change, diagnosis and Recommendation Integrated System and Compositional Nutrient Diagnosis methods. Rev. Bras. Ci. Solo, 31:63-72.

van den Boogaart, K. G., R. Tolosana-Delgado and M. Bren, M. 2013. compositions: Compositional Data Analysis in R package version 1.30-1, http://CRAN.R-project.org/package=compositions

Wairegi, L. and van Asten, P. 2011. Norms for multivariate diagnosis of nutrient imbalance in the East African highland bananas (Musa spp. AAA). J. Plant Nutr. 34:1453–1472.