agronomy

MDPI

Article

Humboldtian Diagnosis of Peach Tree (*Prunus persica*) Nutrition Using Machine-Learning and Compositional Methods

Débora Leitzke Betemps [1,2], Betania Vahl de Paula [1], Serge-Étienne Parent [3], Simone P. Galarça [4], Newton A. Mayer [5], Gilmar A. B. Marodin [6], Danilo E. Rozane [7], William Natale [8], George Wellington B. Melo [9], Léon E. Parent [1,3,*] and Gustavo Brunetto [1]

1 Departamento dos Solos, Universidade Federal de Santa Maria, Av. Roraima, 1000 Camobi, Santa Maria RS 97105-900, Brazil
2 Campus Cerro Largo, Universidade Federal da Fronteira Sul, Av. Jacob Reinaldo Haupenthal, 1580-Bairro São Pedro, Cerro Largo RS 97900-000, Brazil
3 Department of Soils and Agrifood Engineering, Laval University, Québec, QC G1V 0A6, Canada
4 Ascar Emater—Piratini, Rua 20 de Setembro, 158-Centro, Piratini RS 96490-000, Brazil
5 Embrapa Clima Temperado, Centro de Pesquisa Agropecuária de Clima Temperado, BR 392, km 78, Monte Bonito, Pelotas RS 96010971, Brazil
6 Departemento de Horticultura e Silvicultura, Universidade Federal do Rio Grande do Sul, av. Bento Gonçalves 7712, C.P. 15.100, Agronomia, Porto Alegre RS 91540000, Brazil
7 Departamento de Engenharia Agronômica, Universidade Estadual de São Paulo (UNESP), Campus de Registro, Av. Nelson Brihi Badur, Registro SP 11.900-000, Brazil
8 Departamento de Fitotecnia, Universidade Federal do Ceará (UFC), Av. Mister Hull, 2977-Campus do Pici, Fortaleza CE 60356-000, Brazil
9 Embrapa Uva e Vinho, Rua Livramento, 515, Bento Gonçalves RS 95701-008, Brazil
* Correspondence: leon-etienne.parent@fsaa.ulaval.ca

Received: 1 May 2020; Accepted: 20 June 2020; Published: 24 June 2020

check for updates

Abstract: Regional nutrient ranges are commonly used to diagnose plant nutrient status. In contrast, local diagnosis confronts unhealthy to healthy compositional entities in comparable surroundings. Robust local diagnosis requires well-documented data sets processed by machine learning and compositional methods. Our objective was to customize nutrient diagnosis of peach (*Prunus persica*) trees at local scale. We collected 472 observations from commercial orchards and fertilizer trials across eleven cultivars of *Prunus persica* and six rootstocks in the state of Rio Grande do Sul (RS), Brazil. The random forest classification model returned an area under curve exceeding 0.80 and classification accuracy of 80% about yield cutoff of 16 Mg ha^{-1}. Centered log ratios (*clr*) of foliar defective compositions have appropriate geometry to compute Euclidean distances from closest successful compositions in "enchanting islands". Successful specimens closest to defective specimens as shown by Euclidean distance allow reaching trustful fruit yields using site-specific corrective measures. Comparing tissue composition of low-yielding orchards to that of the closest successful neighbors in two major Brazilian peach-producing regions, regional diagnosis differed from local diagnosis, indicating that regional standards may fail to fit local conditions. Local diagnosis requires well-documented Humboldtian data sets that can be acquired through ethical collaboration between researchers and stakeholders.

Keywords: compositional entity; Humboldtian data sets; centered log ratio; machine learning; random forest; nutrient limitations; local diagnosis; peach trees

1. Introduction

In 2017, peaches and nectarines were produced on 1.5×10^6 ha worldwide [1]. Mainland China accounted for 51.0% of total area, followed by Spain (5.5%) and Italy (4.4%). Brazil ranked 13th with 17,116 ha, and 21th in total production. The states of Rio Grande do Sul (RS), Santa Catarina and Paraná accounted for 72% of Brazilian production [2]. Average Brazilian yield was half that of USA and Europe, and this was attributed in part to regional nutrient guidelines based on a limited number of fertilizer experiments that may not fit local conditions.

The performance of Brazilian peach orchards could be improved by tackling local yield-limiting factors. Because the plant explores the soil in deeper layers than the arable layer sampled for soil testing [3,4], tissue tests are generally more closely related to crop performance than soil tests [5]. Indeed, the plant integrates site-specific growth-impacting genetic, managerial, and environmental factors [6]. Yield, fruit quality and tissue nutrient composition depend on cultivar, rootstock, phenological stage, yield, pedoclimatic conditions and crop management [7,8]. To address several factors simultaneously, Humboldtian system-based data sets integrate records at large scale and as much quantitative data as possible instead of focusing on data collected in isolated studies [9].

Well-documented data sets relating features to crop performance can be processed by machine learning methods of artificial intelligence [10]. On the other hand, compositional features are intrinsically multivariate, and necessitating to address information redundancy and the closure problem of compositions using log ratio transformation methods [11]. Acknowledging the multivariate nature of tissue compositions, Lagatu et al. [12] drew a diagnostic yield contour map within an interactive $N \times P \times K$ ternary diagram. Holland [13] proposed using multivariate data analysis to diagnose tissue nutrients holistically rather than separately but did not demonstrate it explicitly nor did he address nutrient interactions. Assuming data additivity and function reflectivity, Beaufils [14] suggested adding up standardized nutrient ratios to nutrient indices to conduct regional nutrient diagnosis, ignoring the large range of mathematically robust multivariate statistical analysis methods and biasing nutrient standards by non-normal data distribution patterns and false positive specimens.

Aitchison [11] and Egozcue et al. [15] developed the concepts of compositional log ratios amenable to multivariate analyses. Log ratios are additive in the n-dimensional Euclidean space and address compositions as unique combinations of parts or entities rather than parts taken in isolation [16]. Using log ratio techniques, multivariate distances can be computed as Euclidean or Mahalanobis distances between defective and successful compositional entities for diagnostic purposes [17,18].

Machine learning and compositional data analysis methods provide unprecedented tools to conduct nutrient diagnosis of tissue compositional entities at local scale if supported by large data sets. We hypothesized that (1) machine learning methods return accurate classification relating yield to combinations of features that influence the performance of peach orchards, and (2) regional diagnosis using state standards differs from local diagnosis that compares the tissue compositional entities of defective and successful specimens. The objective of this paper was to customize nutrient diagnosis of peach orchards using site-specific information.

2. Material and Methods

2.1. Experimental Data Set

The peach data set comprised 472 mature tree specimens representing eleven cultivars of *Prunus persica* collected between 2009 and 2014 on commercial or experimental farms in Rio Grande do Sul, southern Brazil. Three major peach producing mesoregions were represented (i.e., Nordeste Rio-Grandense (Bento Gonçalves, Caxias do Sul, Farroupilha and Flores da Cunha), Porto Alegre (Eldorado do Sul) and Sudeste Rio-Grandense (Pelotas)). The number of specimens varied widely among cultivars. There were 226 "Maciel", 108 "Chimarrita", 54 'Chiripá', 36 'Eragil', 12 'Pialo', and 6 each of 'Dela Nona', 'BRS Fascínio', 'BRS Kampai', 'PS10711', 'PS25399' and 'Barbosa'. Rootstocks were "Aldrighi", 'Capdeboscq', 'Flordaguard', "Nemaguard", 'Okinawa' or "Japanese apricot" (*P. mume*).

The tree training method was open vase. Plantation density ranged between 770 and 1358 trees ha^{-1}. Orchards were managed following national standards for integrated peach orchard management [19,20]. One of the most sensitive features controlling peach yield is winter chill hour requirement for bud break [21], measured in Brazil as the number of hours where temperature is lower than 7.2 °C during the winter period. Meteorological data were obtained from regional weather stations ([22] in Bento Gonçalves, Farroupilha, Flores da Cunha, Caxias do Sul, [23] in Pelotas and [24] in Eldorado do Sul). Orchards were not irrigated. Soils were Typic Hapludalf and Udorthent [25]. Soil between rows was covered by vegetation year-round.

Fertilization followed the Brazilian guidelines for mature (≥4 years old) peach orchards [26] based on tissue tests (0–110 kg N ha^{-1}, 0–52 kg P ha^{-1} and 0–83 kg K ha^{-1}). Hence, the effects of fertilizer dosage and tissue tests were confounded. By comparison, nitrogen was applied at three occasions (i.e., 50% at bud break, 25% at fruit thinning, and 25% after harvest (except in years of low production or excessive vigor)). In comparison, phosphorus was applied once together with the first application of nitrogen. Potassium was applied once except on coarse-textured soils where K was split-applied. The crop was harvested yearly from November to February. Yield was measured in three central trees in experimental areas.

2.2. Soil and Tissue Analyses

Soil nutrients were extracted for K, P, Cu, Zn and Mn using the Mehlich1 method [27]. The Ca, Mg and Na were extracted using KCl 1 M. The Fe was extracted using DTPA. Exchangeable acidity was measured using the Shoemaker–McLean–Pratt (SMP) method. Organic matter content was determined by oxidization in a sulfo-chromic solution. Soil pH was measured in water. Clay content was determined by sedimentation.

Diagnostic leaves were collected in June from the middle tier of annual growth, dried in at ±65 °C, ground to pass through a 1 mm sieve. A subsample was digested using sulfuric acid and quantified for N by micro-Kjeldahl [27]. Another subsample was digested in a mixture of nitric and perchloric acids and quantified by ICP-OES for S, P, K, Ca, Mg, Zn, Cu, Mn, Fe and B concentrations. Fruit quality was measured as fruit weight, dimension (average length and width), firmness, Brix index and acidity using 30 fruits per experimental unit [28].

2.3. Isometric Log-Ratio Transformation

Isometric log ratios (*ilr*) are orthogonal arrangements of *D* components into *D*-1 balances, the exact number of degrees of freedom available in *D*-part compositional entity [29]. Isometric log ratios were computed as follows [30]:

$$ilr_i = \sqrt{\frac{rs}{r+s}} ln\left(\frac{G_N}{G_D}\right) \tag{1}$$

where *r* and *s* are the numbers of components at numerator and denominator, respectively, and G_N and G_D are the geometric means of components at numerator and denominator, respectively. Components were arranged as meaningful balances in a sequential binary partition or SBP (Table 1). We first contrasted nutrients against the filling value computed by difference between measurement unit and the sum of quantified components. While N, K, Mg, P, S, Cl and Na are phloem-mobile, the Fe, Zn, Cu, B and Mo have intermediate mobility and the Ca and Mn are relatively immobile [31]. Concentrations of Cu, Zn and Mn may vary widely due to fungicide applications [32]. Orthonormal balances allowed computing Mahalanobis distance as follows:

$$\mathcal{M} = \sqrt{\sum_{i=1}^{D}\left(ilr_i - ilr_i^*\right)^T COV^{-1}\left(ilr_i - ilr_i^*\right)} \tag{2}$$

where ilr_i and ilr_i^* are orthonormal balances for the specimen under diagnosis and reference balances, respectively, COV is the covariance matrix, and T indicates that the ilr vector is transposed. The $\mathcal{M}^2$ is distributed like a χ^2 variable.

Table 1. Sequential binary partition of components of the dry tissue mass.

Ilr	N	P	K	Mg	Ca	Cu	Zn	Mn	Fe	F_v	r	s
1	1	−1	0	0	0	0	0	0	0	0	1	1
2	0	0	1	−1	0	0	0	0	0	0	1	1
3	1	1	−1	−1	0	0	0	0	0	0	2	2
4	1	1	1	1	−1	−1	−1	−1	−1	0	4	5
5	0	0	0	0	1	−1	−1	−1	−1	0	1	4
6	0	0	0	0	0	1	−1	0	0	0	1	1
7	0	0	0	0	0	0	0	1	−1	0	1	1
8	0	0	0	0	0	1	1	−1	−1	0	2	2
9	1	1	1	1	1	1	1	1	1	−1	9	1

2.4. Centered Log-Ratio Transformation

The centered log ratio (clr) integrates all pairwise log ratios in a composition [11], as follows for N:

$$clr_N = ln\left(\frac{N}{G}\right) = ln\left(\sqrt[11]{\frac{N}{N} \times \frac{N}{P} \times \frac{N}{K} \times \frac{N}{Mg} \times \frac{N}{Ca} \times \frac{N}{B} \times \frac{N}{Cu} \times \frac{N}{Zn} \times \frac{N}{Mn} \times \frac{N}{Fe} \times \frac{N}{Fv}}\right) \quad (3)$$

where G is geometric mean across components including Fv, and Fv is the filling value computed by difference between measurement unit and the sum of quantified nutrients. The clr transformation has Euclidean geometry. The Euclidean distance ϵ between two D-part compositions of equal length is computed at local scale as follows:

$$\epsilon = \sqrt{\sum_{i=1}^{D}\left(clr_i - clr_i^*\right)^2} = \sqrt{\sum_{i=1}^{D}\left(clr_i - clr_i^*\right)^T I^{-1}\left(clr_i - clr_i^*\right)} \quad (4)$$

where clr_i is the clr transformation of the diagnosed composition, I is the identity matrix and clr_i^* is the clr transformation for reference local compositions. Nutrients are classified in the order of their limitation to yield along the $clr_i - clr_i^*$ gradient and illustrated in histograms.

Nutrient diagnosis can be conducted as Mahalanobis distance at regional scale as follows, assuming independence among clr variables [33]:

$$\mathcal{M} = \sqrt{\sum_{i=1}^{D}\left(clr_i - clr_i^*\right)^T VAR^{-1}\left(clr_i - clr_i^*\right)} \quad (5)$$

where clr_i and clr_i^* refer to diagnosis and reference compositions and VAR is variance matrix excluding the clr value for the filling value to avoid generating a singular matrix. Hence, the reference compositions (Equation (4)) and weighted (Equation (5)) clr differences as well as assumptions differed between local and regional diagnoses.

Machine learning (ML) analysis was run using freeware Orange 3.24. Fruit yield categories were separated at cut off yield of 16 Mg ha^{-1}, the world average in 2017, yet above the 14.5 Mg ha^{-1} average in Brazil [1]. Exploratory analysis was conducted using the classification tree algorithm and the tree viewer. The survey data set was split into training (70%) and testing (30%) sets to test precision and across the data set by cross-validation to select a subset of balanced specimens. Precision metrics were accuracy (proportion of instances predicted as true negative or true positive) and area under curve (AUC) [17]. We expected AUC of 70–90% [34]. In exploratory analysis, random forest (RF), support vector machine, neural networks, adaboost, and stochastic gradient decent models returned similar accuracies in cross-validation (data not shown). However, we selected RF to deal with over-fitting of partition trees, but RF may be affected by data transformation [35]. The significance of the partition in the confusion matrix for the testing data set was assessed as a χ^2 homogeneity test with Yates' correction. Classification prediction and risk analysis for independent specimens were provided by the prediction module of Orange 3.24 using the same features as in the training set. Descriptive statistics were computed using Excel Microsoft 365.

3. Results

3.1. Features

The meteorological indices and soil and tissue tests used to run machine learning models are presented in Tables 2–4. There were large ranges of properties, allowing to model fruit yields across a large range of features. Soil pH varied from 5.0 to 5.9 with a median value of 5.3. High soil P, Cu and Zn contents are due in part to the application of organic residues of diverse origins. Foliar nutrient composition is presented by cultivar in Table 4. Exploratory analysis using the classification tree algorithm indicated that the number of chilling hours, the cultivar and tissue K were driving variables at high yield level (data not shown), indicating genetic–environment–management interactions at local scale.

Table 2. Synthetic meteorological variables and fruit yield in the peach orchard data set of Rio Grande do Sul, Brazil (2009–2014).

	Unit	Minimum	Median	Maximum
			Bento Gonçalves	
Mean annual air temperature	°C	12.95	17.15	22.30
Mean annual precipitations	mm	1401	1810	2043
Average number of chilling hours < 7.2 °C	°C	263	360	435
Fruit yield	Mg ha^{-1}	0.2	7.1	30.4
			Pelotas	
Mean annual air temperature	°C	14.40	17.90	23.68
Mean annual precipitations	mm	1096	1398	1833
Average number of chilling hours < 7.2 °C	°C	173	350	440
Fruit yield	Mg ha^{-1}	0.5	14.1	38.8
			Eldorado do Sul	
Mean annual air temperature	°C	12.77	18.28	24.75
Mean annual precipitations	mm	1333	1530	2011
Average number of chilling hours < 7.2 °C†	°C	282	376	469
Fruit yield	Mg ha^{-1}	0.4	5.0	10.1

† Number of chilling hours available in 2009 and 2010 only.

Table 3. Soil analyses reported in peach orchards of experimental sites, Rio Grande do Sul, Brazil.

	Minimum	**Median**	**Maximum**
		%	
Clay	14	17	25
Organic matter	1.0	1.8	3.9
Base saturation	36	48	78
		$cmol_c\ dm^{-3}$	
Cation exchange capacity	4	11	16
Sum of bases	23	73	92
		$mg\ dm^{-3}$	
K	44	128	330
Ca	160	1400	2000
Mg	48	108	648
Na	3	10	33
P	2	27	84
Cu	4.5	6.3	30.1
Zn	1.8	5.7	16.7
Mn	14	17	48
Fe	1000	2000	5000

3.2. Model Precision

The AUC of the RF model varied between 0.834 and 0.844 in test and 0.894–0.901 in cross-validation, in the range of 0.7–0.9 considered acceptable by Delacour et al. [34] for diagnostic purposes (Table 5). Classification accuracy was close to 80% as reached by most fruit crops [36]. There was no apparent advantage using log-ratio transformations before processing compositional data with RF. At the step of model building, raw compositions were; thus, preferable because they did not require full-length compositions needed to log-ratio transform the data, hence avoiding to impute data, replace values lower than detection limits or remove observations.

The confusion matrix showed 142 true negative (high-yield, well-balanced) specimens producing more than 16 Mg ha^{-1}, providing a diversity of factor combinations leading to high performance of peach orchards. There were 254 true positive (low-yield, imbalanced), 39 false negative (low-yield, well-balanced indicating yield-limiting factors other than nutrients) and 37 false positive (high-yield, imbalanced due to luxury consumption or contamination) specimens. The partition was significant at $p = 0.01$ according to the χ^2 homogeneity test with Yates' correction.

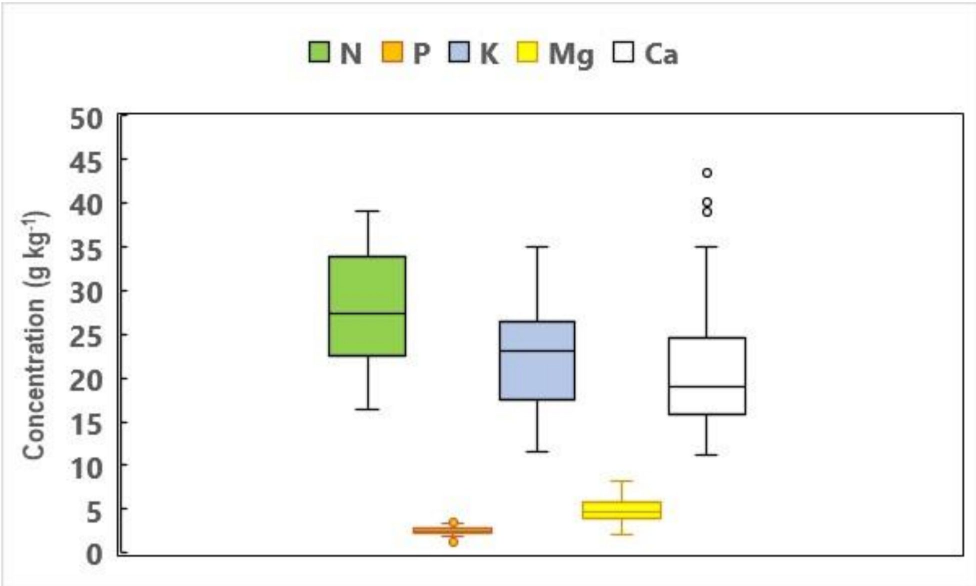
Boxplots of foliar macro- and micro-nutrient concentrations and of centered log ratios of true negative specimens are presented in Figures 1 and 2. There were some outliers among P and Ca expressions. The number of outliers was larger among micronutrient expressions likely due to variable soil composition, site-specific applications of organic amendments, and different timings between tissue sampling and fungicide applications (Zn and Mn in carbamate formulations, copper sulfate).

Table 4. Tissue nutrient concentrations of eleven peach cultivars in the peach orchard data set of Rio Grande do Sul. Brazil.

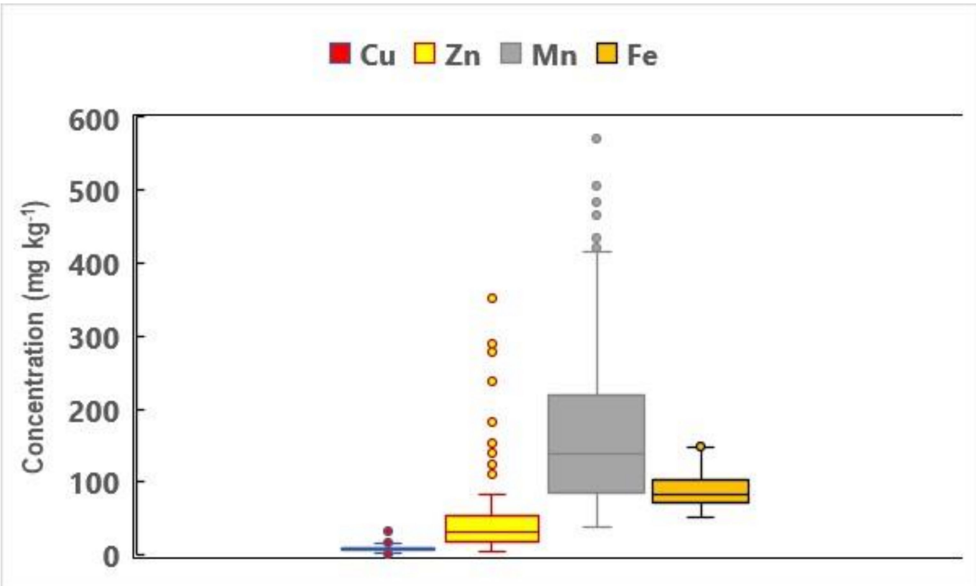
	Minimum	Median	Maximum	Minimum	Median	Maximum	Minimum	Median	Maximum	Minimum	Median	Maximum
		g kg^{-1}			g kg^{-1}			g kg^{-1}			g kg^{-1}	
		Maciel			**Chimarrita**			**Chiripá**			**Eragil**	
N	10.60	29.25	39.04	19.26	27.55	41.32	14.36	21.80	29.24	18.04	22.02	27.40
P	1.20	2.26	3.40	1.36	2.16	2.57	1.92	2.58	3.99	1.99	2.62	3.16
K	1.94	22.5	37.18	14.58	22.5	38.45	12.92	19.45	29.25	13.14	25.80	34.97
Ca	9.43	17.25	32.88	8.07	16.95	31.10	11.42	26.65	43.90	9.70	25.11	33.60
Mg	2.00	4.86	8.20	2.00	3.92	7.45	3.57	5.72	9.16	2.61	5.26	6.85
Cu	0.001	0.005	0.011	0.001	0.005	0.020	0.007	0.010	0.061	0.008	0.011	0.032
Zn	0.003	0.017	0.036	0.013	0.025	0.050	0.025	0.047	0.148	0.030	0.078	0.352
Mn	0.023	0.081	0.262	0.050	0.100	0.200	0.012	0.176	0.535	0.060	0.195	0.482
Fe	0.043	0.097	0.190	0.023	0.144	0.570	0.045	0.074	0.139	0.049	0.077	0.112
		Pialo			**Delanona**			**Fascínio**			**Kampai**	
N	21.54	24.73	26.96	20.66	22.59	24.43	20.31	21.63	22.33	18.04	19.83	21.89
P	2.25	2.42	2.73	2.04	2.18	2.53	2.16	2.33	2.59	2.64	2.75	3.11
K	11.48	14.91	21.69	13.28	14.24	16.03	16.11	17.62	19.06	13.78	15.89	17.76
Ca	11.18	15.64	19.09	21.04	24.13	27.32	12.67	16.85	18.08	13.43	24.12	27.97
Mg	3.48	4.33	5.27	5.08	5.42	6.43	3.77	4.43	4.63	3.96	5.22	5.62
Cu	0.008	0.009	0.011	0.069	0.080	0.103	0.012	0.014	0.018	0.010	0.010	0.011
Zn	0.032	0.038	0.048	0.145	0.192	0.215	0.018	0.020	0.026	0.045	0.054	0.066
Mn	0.179	0.244	0.272	0.428	0.538	0.627	0.095	0.102	0.136	0.127	0.154	0.171
Fe	0.058	0.073	0.082	0.064	0.075	0.094	0.101	0.113	0.134	0.058	0.063	0.083
		PS10711			**PS-Tardia**			**São Barbosa**			**General**	
N	23.73	25.96	27.14	20.23	21.71	22.94	21.89	23.77	24.16	10.6	26.11	41.32
P	2.29	2.80	3.08	2.03	2.10	2.24	2.13	2.17	2.22	1.20	2.30	3.99
K	13.81	14.99	16.62	18.24	20.71	21.63	12.98	18.80	20.27	19.40	21.90	38.45
Ca	13.63	18.13	20.52	32.95	35.27	40.58	26.13	31.06	40.11	8.07	18.42	43.90
Mg	4.57	6.40	7.33	5.08	5.36	6.63	5.35	6.34	7.52	2.00	4.85	9.16
Cu	0.009	0.010	0.012	0.013	0.015	0.016	0.009	0.009	0.010	0.001	0.006	0.103
Zn	0.047	0.054	0.064	0.110	0.141	0.183	0.045	0.058	0.079	0.003	0.026	0.352
Mn	0.186	0.224	0.247	0.321	0.428	0.492	0.134	0.256	0.394	0.012	0.130	0.627
Fe	0.072	0.118	0.224	0.074	0.082	0.112	0.068	0.084	0.099	0.043	0.092	0.204

Table 5. Comparison between expressions for tissue nutrient compositions of peach trees using the random forest model.

Expression	Testing Data (30% of the Data)		Cross-Validation (100% of the Data)	
	Area under Curve	Classification Accuracy	Area under Curve	Classification Accuracy
Raw concentration data	0.844	0.801	0.894	0.826
Centered log ratios	0.834	0.794	0.901	0.835
Isometric log ratios	0.844	0.794	0.901	0.836

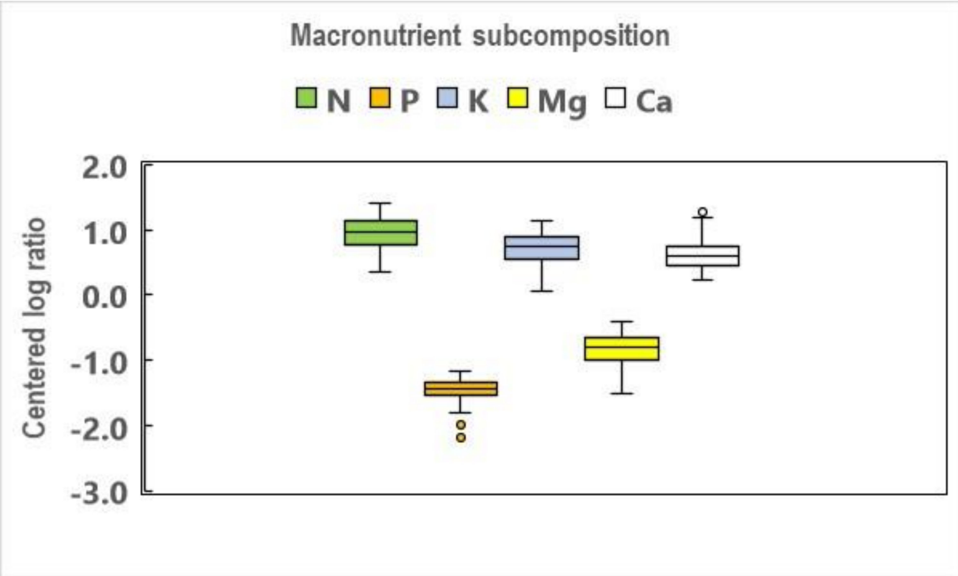


(**a**)

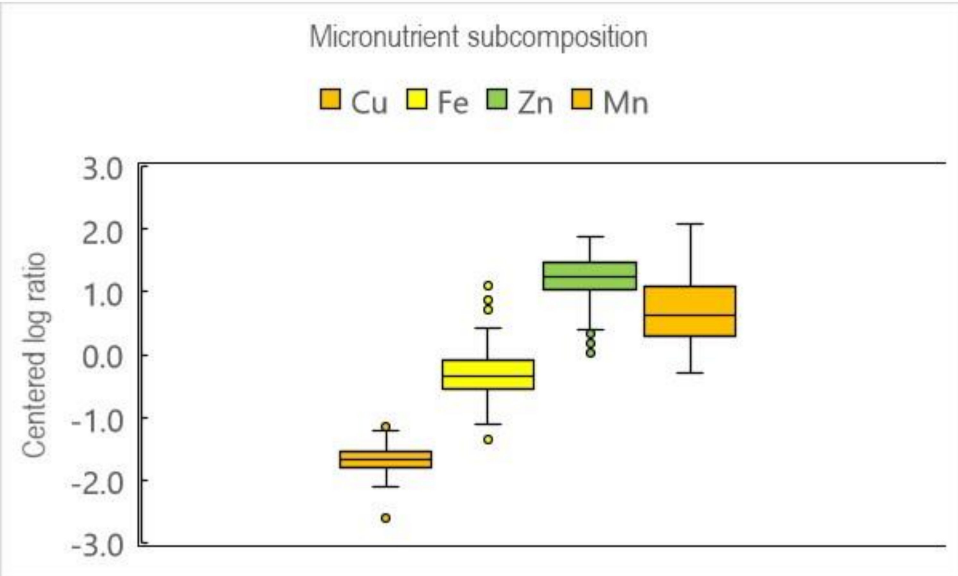


(**b**)

Figure 1. Boxplots of foliar concentrations of (a) macronutrient and (b) micronutrient subcompositions of true negative specimens peach orchards in Rio Grande do Sul, Brazil.



(**a**)



(**b**)

Figure 2. Boxplots of foliar centered log ratios of (**a**) macronutrient and (**b**) micronutrient subcompositions of true negative specimens peach orchards in Rio Grande do Sul, Brazil. Minus signs indicate that that nutrient concentration is lower than geometric mean.

Ranges of nutrient concentrations, centered log ratios and isometric log ratios in boxplots are presented in Table 6. Among macronutrients, the lower and upper limits of N boxplots differed the most from Brazilian standards that could lead likely to N over-fertilization. The P ranges were similar between standards and boxplots, while the ranges of K, Mg and Ca concentrations were wider. Macronutrients, which showed narrower ranges of concentrations compared to micronutrients, were diagnosed as a separate subset to facilitate comparison with Brazilian standards (Table 6). The means and covariance matrix across 181 balanced (TN + FN) specimens are presented in Table 7.

Table 6. Foliar nutrient levels for state standards and in 142 true negative specimens of peach orchards in Rio Grande do Sul, Brazil, compared to current standards.

Nutrient	State Standards (Brunetto et al [26])			True Negative Specimens			Centered Log Ratio		
	g kg^{-1}						Unitless		
	Insufficient	Normal	Excessive	Minimum	Median	Maximum	Minimum	Median	Maximum
N	<20.0	33.0–45.0	>60.0	16.5	27.4	39.0	0.344	0.965	1.422
P	<0.5	1.5–3.0	>4.0	1.2	2.5	3.4	−1.804	−1.337	−1.171
K	<5.0	14.0–20.0	>28.0	11.5	23.1	35.0	0.063	0.755	1.144
Mg	<2.0	5.0–8.0	>12.0	2.1	4.7	8.3	−1.511	−0.804	−0.393
Ca	<6.5	17.0–26.0	>36.0	11.2	19.1	35.0	0.236	0.590	1.270
	mg kg^{-1}								
Cu	?	6–30	>50	2	6	18	−2.112	−1.671	−1.209
Fe	<50	100–230	>330	53	83	148	−1.122	−0.346	0.430
Zn	<10	24–37	>50	4	31	84	0.404	1.245	1.874
Mn	<20	30–160	>400	38	139	422	−0.281	0.637	2.087
B	<3	30–60	>90	-	-	-	-	-	-

Table 7. Covariance matrix and means of isometric log ratios (balances) of 181 nutritionally balanced specimens of peach trees in Rio Grande do Sul, Brazil.

Balance	P|N	Mg|K	K,Mg|N,P	Ca|N,P,K,Mg
		Mean		
	1.692	1.081	−0.202	−0.702
		Covariance matrix		
P|N	0.04991	0.04247	0.01101	0.03489
Mg|K	0.04247	0.10014	0.02391	0.04505
K,Mg|N,P	0.01101	0.02391	0.03912	0.03483
Ca|N,P,K,Mg	0.03489	0.04505	0.03483	0.06803

3.3. Regional vs. Local Diagnosis

The *ilr* values of state standards [26] and the Mahalanobis distance from regional standards computed in the present study were measured using state standard median (M), first quartile (Q1), third quartile (Q3) and six sequential combinations thereof as Q1M, MQ3, Q3M, Q1Q3 and Q3Q1 (Table 7). The $\mathcal{M}^2$ values from the present TN subset were 13.36 (MM), 13.76 (Q1Q1), 13.07 (Q3Q3), 14.20 (MQ1), 12.71 (MQ3), 12.96 (Q1M), 13.76 (Q3M), 12.35 (Q1Q3) and 14.63 (Q3Q1). The Q1Q3 sequence (N_{Q1}, P_{Q3}, K_{Q1}, Mg_{Q3}, Ca_{Q1}) that showed the smallest Mahalanobis distance was retained to compare regional to local diagnosis (Table 8).

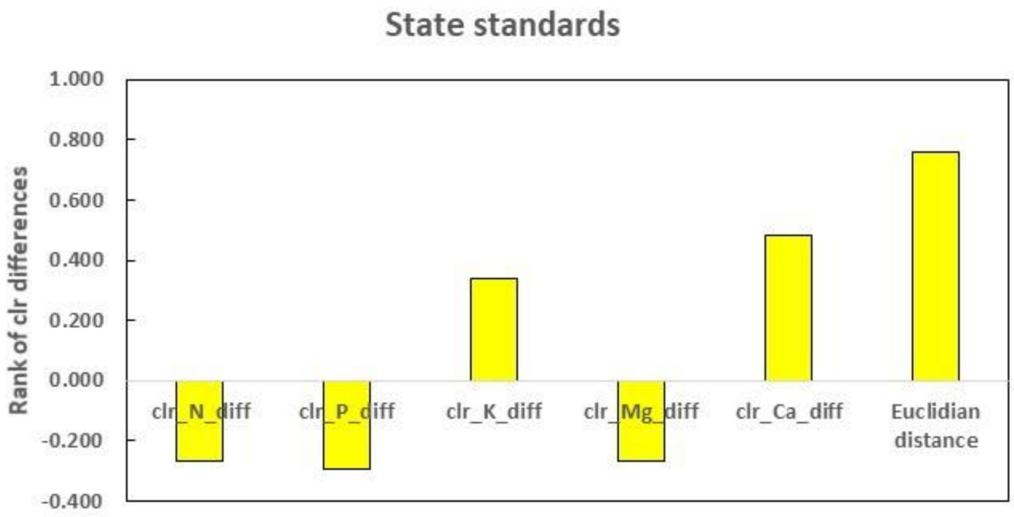
The closest successful Euclidean neighbors were detected by comparing foliar compositions and other features of TN specimens (municipality, cultivar, rootstock, number of chilling hours and some soil analyses where available in the data set) to those of the diagnosed specimens. For the compared defective and successful peach orchards at Bento Gonçalves and Pelotas, soil texture and classification, clay and organic matter contents, and number of chilling hours were similar, but yield, cultivar, rootstock and tissue composition differed.

A defective specimen of "Chimarrita" grafted on "Aldrighi" (8.9 Mg ha^{-1}) was grown in Bento Gonçalves. The closest successful orchards (29.5–30.4 Mg ha^{-1}) were "Chimarrita" and "Maciel" grafted on "Nemaguard". Because both successful orchards returned similar diagnosis, the "Chimarrita" orchard was selected as the closest successful neighbor. Regional diagnosis across factors indicated N, K and Mg shortage and P sufficiency (Table 8). The diagnosed tissue specimen was classified as true positive with χ^2_5 value (squared Mahalanobis distance) of 17.47 across *ilr* variables and a highly significant probability to respond to a more appropriate fertilization regime. At local scale, the nearest neighbor returned an inverse K and Mg diagnosis (Figure 3), indicating site-specific factor interactions that were not depicted by nutrient standards at regional scale. While the apparent K:Mg imbalance detected at local scale may also be attributed not only to different rootstocks ("Aldrighi" vs. "Nemaguard"), comparable rootstock for successful "Chimarrita" was not available, emphasizing the importance of acquiring larger data sets.

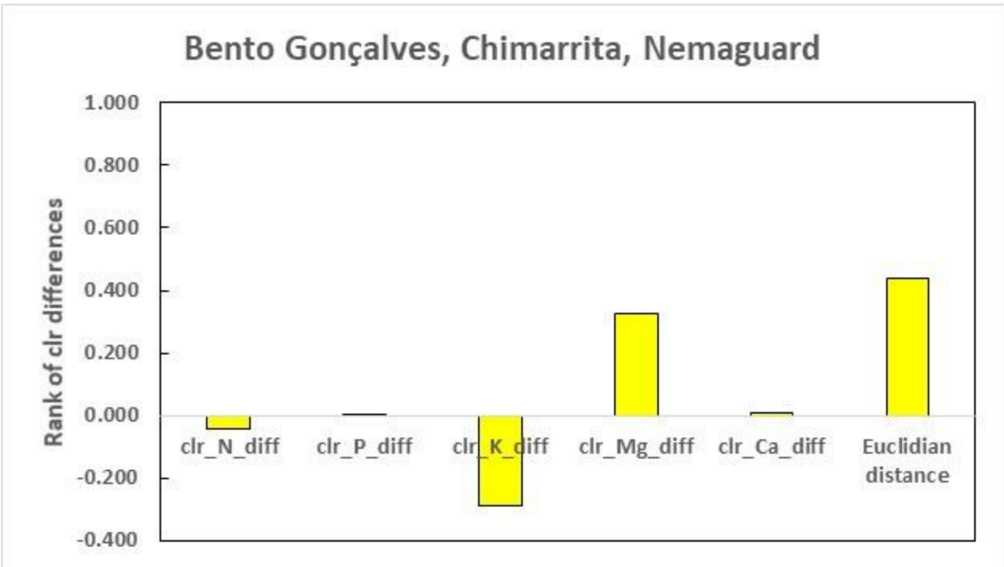
Table 8. Reference concentration and *clr* values for macronutrients from state standards and closest true negative specimens in two contrasting regions of Rio Grande do Sul (RS), Brazil.

Reference Values	**Concentration Values (g kg^{-1})**					**Centered Log Ratios (Unitless)**					**Fruit Yield**
	N	**P**	**K**	**Mg**	**Ca**	**clr_N**	**clr_P**	**clr_K**	**clr_Mg**	**clr_Ca**	**Mg kg^{-1}**
State "normal" concentrations (Rio Grande do Sul)											
Median value †	39.0	2.3	17.0	6.5	21.5	1.214	−1.639	0.384	−0.578	0.619	-
Q1_Q3 †	36.0	2.6	15.5	7.3	19.3	1.138	−1.481	0.295	−0.465	0.512	-
Bento Gonçalves—RS (Nordeste Rio-Grandense)											
Defective trees	23.2	1.7	18.3	4.7	26.2	0.873	−1.772	0.635	−0.731	0.995	8.9
Closest successful neighbors	25.0	1.7	25.3	3.5	26.8	0.917	−1.773	0.926	−1.056	0.986	30.4
Pelotas—RS (Sudeste Rio-Grandense)											
Defective trees	31.4	2.8	18.9	3.9	11.7	1.200	−1.218	0.692	−0.886	0.212	0.4
Closest successful neighbors	33.4	3.2	23.2	4.5	13.8	1.120	−1.226	0.755	−0.885	0.236	21.5

† Median, Q1, Q3 = median and alternate values of the first and third quartiles, respectively, for regional "normal" standards (Brunetto et al., 2016).



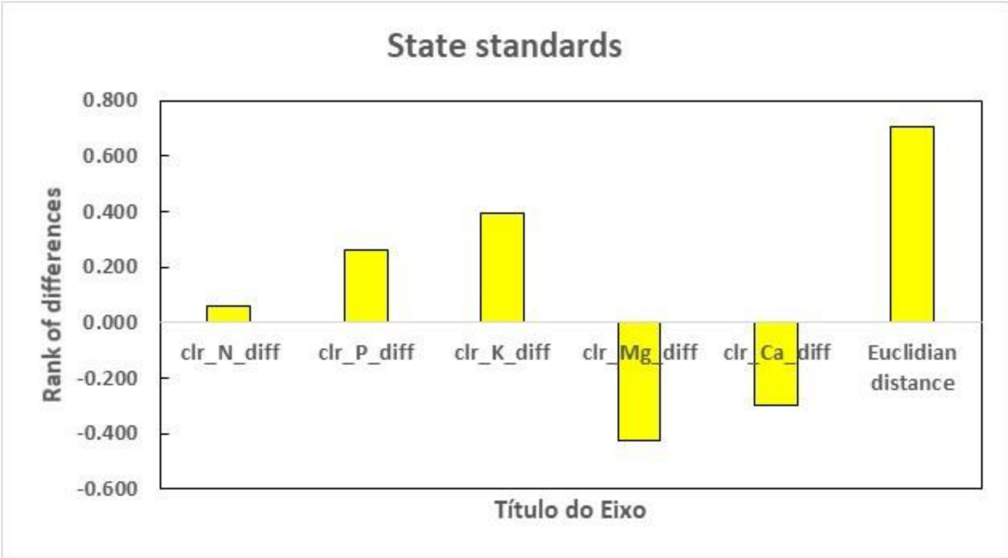
(**a**)



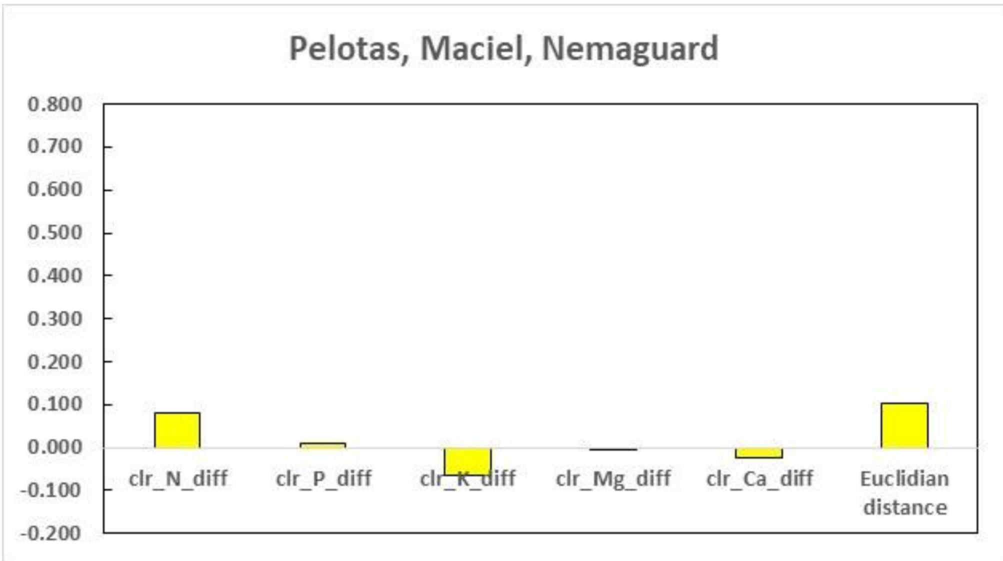
(**b**)

Figure 3. Histograms showing contrasting macronutrient diagnoses of a low-yielding peach orchard between (**a**) state standards and (**b**) the closest neighbor in Bento Gonçalves, Rio Grande do Sul, Brazil (same location and cultivar, similar soil properties and number of chilling hours). The centered log ratio (*clr*) differences are computed as *clr* values of the diagnosed specimen (yield of 8.9 Mg ha^{-1}) minus the corresponding *clr* values of (left) state standards (no yield target indicated) or (right) the closest successful specimen (yield target of 30.4 Mg ha^{-1}). The Euclidian distance is computed across *clr* differences. Negative differences between defective and successful specimens indicate relative shortage. Positive differences indicate relative excess.

In Pelotas, state standards indicated considerable nutrient imbalance (Figure 4). Hence, regional diagnosis classified the specimen as true positive potentially responsive to K and Ca additions and to the reduction the N, P and Mg fertilization. At local scale, in contrast, the nearest successful neighbor, where other features were close to those of the diagnosed specimen, indicated negligible nutrient imbalance. Hence other factors likely limited yield at local scale but this was not indicated by the regional diagnosis.



(**a**)



(**b**)

Figure 4. Histograms showing contrasting macronutrient diagnoses of a low-yielding peach orchard between (**a**) state standards and (**b**) the closest neighbor in Pelotas, Rio Grande do Sul, Brazil (same location and cultivar, similar soil properties and number of chilling hours). The centered log ratio (*clr*) differences are computed as *clr* values of the diagnosed specimen (yield of 8.9 Mg ha^{-1}) minus the corresponding *clr* values of (left) state standards (no yield target indicated) or (right) the closest successful specimen (yield target of 30.4 Mg ha^{-1}). The Euclidian distance is computed across *clr* differences. Negative differences between defective and successful specimens indicate relative shortage. Positive differences indicate relative excess.

4. Discussion

4.1. Compositions as Separate Parts or Interactive Systems?

Johnson et al. [37] reported that nutrient concentration ranges, the Diagnosis and Recommendation Integrated System or DRIS [14] and the deviation from optimum percentage or DOP [38] have been used with some success to diagnose the nutrient status of peach trees. Normally distributed nutrient concentration ranges addressed simultaneously to diagnose tissue nutrient status collapse in the ellipsoidal multivariate hyperspace of nutrients and are thus useless as the number of diagnosed nutrients increases [39]. On the other hand, dual ratios such as P:Zn [40], N:P, N:S and S:P [41] are important in peach tree nutrition. Nutrient interactions such as N × P synergism and K × Mg antagonism [32,42] should also be considered. However, D-part tissue compositions can return up to $D \times (D-1)/2$ dual ratios, most of them being redundant, hence useless for correlation analysis with yield. For example, the N:P, N:S and S:P ratios are redundant because $\frac{N}{P} = \frac{N}{S} \times \frac{S}{P}$. Tissue compositions should be rather viewed as entities (i.e., unique combinations of nutrients). Aitchison [11] integrated pairwise log ratios into centered log ratios to secure the unique character of combinations of components in a composition (Equation (3)). The tissue nutrient *clr* variables are multivariate in nature and affected by farm nutrient management, climate and soil composition that vary widely regionally. Moreover, because adding one nutrient through fertilization may affect several others by resonance within the compositional space of tissue dry matter, other nutrients are also impacted by fertilization. Downscaling regional *clr* descriptive statistics (mean, variance) to site-specific level may be hazardous. Direct comparison between two equal-length compositions (Equation (4)) lumped into the Euclidean distance [11] at local scale where soil, management and meteorological factors are comparable; thus, appeared to be a more appropriate diagnostic method than computing *clr* indices using means and standard deviations at regional scale. In addition, differences between *clr* values adding up to the Euclidean distance allow classifying nutrients numerically in the order of their apparent limitation to yield [16]. The perturbation vector between two compositions computed as nutrient-wise ratios $(X_{i_{defective}}/X_{i_{successful}})$ between defective and successful compositions is an alternative expression to rank nutrients in a numerical order at local scale [43].

Compositional data distribution of successful specimens needs not have a specific shape. Successful specimens in "enchanting islands" [10,43] may be even harbored close to the composition of defective specimens but outside the regional critical hyper-ellipsoid. Regional and local diagnoses thus involve different references (regional centroids vs. local enchanting islands), and weighting matrices (identity, variance, covariance) that may lead to contrasting nutrient diagnoses (Figures 3 and 4). An additional benefit of selecting the closest successful neighbors (smallest Euclidean distance) is to provide reliable means to correct controllable growth-limiting factors and reach the trustful potential yields recorded in comparable surroundings.

4.2. From Regional to Local Diagnosis

In the early 1800's, Alexander von Humboldt championed the principles of quantitative biogeography [9] that illuminated a cascade of key concepts in agronomy and soil science such as Boussingault's nutrient budgets [44], Sprengel's law of the minimum, Liebscher's law of the optimum, Mitscherlich's law of diminishing returns [45], as well as Dokutchaev's morphogenetic soil classification system. Bernhard Baule combined interactive nutrients into a multiplicative law of diminishing returns that was later extended by Wallace and Wallace [46] to ≈70 multiplicative growth factors, a concept known as the law of the maximum. At local scale, it is unlikely that 70 growth factors can reach non-limiting conditions but in some illusory "Gardens of Eden". However, several near-optimum conditions could be reached in enchanting islands showing uncontrollable factors comparable to those found in defective orchards.

Difficulties to fit deterministic models to facts and data and to derive economically optimum nutrient dosage led to the development of empirical polynomial models by economists [47]. However, to make predictions, calculations required not only response models but also assumptions on the likelihood of future events and of uncontrollable and controllable factors [48,49]. Kyvegyga et al. [48,50,51] showed that the historical difficulties to tackle optimum nutrient rates using a limited number of fertilizer experiments could be alleviated by collecting large amounts of on-farm data.

Natale et al. [7] emphasized the importance of considering local conditions for nutrient management of orchards. The low-performing "Chimarrita" in Bento Gonçalves was grafted on

"Aldrighi" and the high-performing one on "Nemaguard", indicating possible nutrient imbalance attributable either to inadequate nutrient management or to difference in rootstock. Mestre et al. [52] and Jimenez et al. [53] showed that rootstock could regulate the nutrition of peach trees. In contrast, Mayer et al. [54] did not find any difference in leaf nutrient content of "Maciel" grafted on "Nemaguard" and "Aldrighi", although nutrient levels were generally below state standards. Galarça et al. [55] concluded, from field trials on "Chimarrita" and "Maciel" grafted on "Aldrighi", 'Capdeboscq', 'Flordaguard' or "Nemaguard", that scion, rootstock and soil nutrient supply can impact on leaf content of peach trees but not necessarily on tree performance. Only well-documented data sets can fully capture the combined effects of yield-driving variables at local scale.

It may be argued that nutrient dosage has been optimized by curve fitting in a few well-conducted fertilizer experiments but has not been optimized in growers' enchanting islands. Successful specimens provide nutrient dosage at local scale where uncontrollable and controllable growth factors interact and where controllable factors have been combined successfully. It is common that growers compare unhealthy and healthy specimens on their own property and in comparable surroundings. Trustful data sets and effective data-processing methods can allow growers to compare defective to well-documented successful specimens, avoiding extra analytical costs. Proximity between defective and successful specimens makes corrective measures more trustful. However, regional standards insure more protection against outliers that may contaminate some unsupervised enchanting islands (several enchanting islands in the TN data subset should be compared as compositional references for defective specimens). Nutrient diagnosis at regional scale then becomes a subsidiary tool in the decision-making process. Because soil fertility classes are established across growth-limiting factors, such as soil texture, compaction, and stoniness as well as soil conservation measures, regional guidelines could be upgraded or downgraded to adjust fertilization to local conditions.

4.3. Machine Learning and Big Data

In this study, we compared defective and successful compositional entities at local scale where all factors but the ones limiting yield were comparable. While yield cut-off was fixed at 16 Mg fruit ha^{-1}, local organizations may select another yield cut-off and a minimum set of features to run their own machine learning and compositional models. As data sets build up, machine learning methods could assess more accurately the contribution of each feature to crop yield by removing them sequentially to test parsimoniously their impact on yield prediction (razor of Occam). To facilitate data collection, minimum data sets can be selected from meaningful quantitative and qualitative data easily available at farm level.

Nowadays, large data sets can be processed by machine learning and compositional data analysis methods directly from data input rather than being supervised by deterministic response models to conduct nutrient diagnosis. Given local uncontrollable factors such as climate, soil depth, stoniness and texture, enchanting islands may be documented where controllable factors have been already addressed successfully by local growers. It should be noted that any change in fertilization regimes of peach orchards may take more than one season to be effective because nutrient reserves accumulated in off years can be remobilized in large amounts in fruiting years [31,56] and at high rate [57].

4.4. Citizen Science and Precision Farming

The concept of site-specific nutrient management has been developed to increase crop yield and quality at local scale [54,58] and to minimize environmental damages from unwise fertilization [39,59]. Precision maps indicated that fruit quality may decrease at high-yield level [60]. [61] demonstrated the importance to adopt profitable site-specific nutrient management and disease control in Brazilian fruit orchards.

Citizen science to collect high-quality data is challenging because it requires close and ethical collaboration between researchers and stakeholders to build trustful and informative data sets [62,63]. Data sets are developing rapidly in North America from continental [64] to regional [9,51,65,66] and local [50] scales. A recent survey showed positive attitude of American fruit growers toward precision agriculture if supported by research and extension programs [67]. Researchers and growers can document, store and track analytical and managerial records. Spectroscopic techniques may facilitate collecting proximate soil analyses at low cost [68–70]. While plant tissue analysis has long been non-competitive with soil analysis for price and the facility of data collection and interpretation, high-throughput inductively-coupled plasma (ICP) technology [71], low-cost visible-infrared-ultraviolet (VIS-IR-UV) spectroscopy [68] and laser-induced breakdown technology [72] may increase the use of both plant and soil analysis in the near future. Large data sets can be processed rapidly by machine learning and compositional methods to tackle local production problems. The larger and more diversified the data set, the more accurate the prediction. Our study combined the efforts of Brazilian growers and research institutions to build knowledge on the site-specific nutrient management of peach orchards.

5. Conclusions

There is a great challenge in Brazil and many other fruit-producing countries to increase the production of high-quality fruits by improving nutrient management of orchards at local scale. Up till now, regional nutrient standards based on field trials have been used to interpret the results of soil and tissue analyses. In the present study, machine learning models relating fruit yield to tissue composition returned classification accuracy >80% from a set of growth-impacting features at yield cutoff of 16 Mg ha^{-1}. The collection of state-wide data sets from experimental farms and commercial orchards allowed setting apart nutritionally balanced specimens to provide updated tissue nutrient standards from ever-growing data sets.

At regional scale, site attributes are assumed to be equal and yield targets are not documented. At local scale, several attributes are reported, and trustful yield targets and corrective measures are provided in close enchanting islands. Nutrient imbalance diagnosis at regional scale may; thus, differ from local diagnosis. Such discrepancy may explain in part why several Brazilian peach orchards produced deceiving yields using the present regional standards. Due to high cost of field trials, local diagnosis requires a close and ethical collaboration between researchers and stakeholders to acquire large-size and diversified sets of high-quality trustful data. As data sets mature in size and diversity, machine learning and compositional methods could solve more complex and subtle factor interactions at local scale. This will be possible only by combining the efforts of researchers, extension specialists, crop advisers and growers.

Author Contributions: D.L.B. organized the data set, ran computations, co-wrote the paper. B.V.d.P. organized the data set, ran computations, co-wrote the paper. S.-É.P. developed the concept of successful neighbors, revised the paper. S.P.G. provided growers' data, revised the paper. N.A.M. provided growers' data, revised the paper. G.A.B.M. provided growers' data, revised the paper. D.E.R. provided growers' data, revised the paper. W.N. revised the paper. G.W.B.M. revised the paper. L.E.P. supervised computations, organized and co-wrote the paper. G.B. revised the paper. All authors have read and agreed to the published version of the manuscript.

Funding: We thank the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), the CMPC Impresa and the Natural Science and Engineering Research Council of Canada (NSERC-2254) for financial support.

Conflicts of Interest: The authors have declared that no competing interests exist.

References

1. FAOSTAT. Production Quantities of Peaches and Nectarines by Country. 2019. Available online: http://www.fao.org/faostat/en/#data/QC/visualize (accessed on 22 June 2020).
2. FNP; Pêssego. *FNP Consultoria & Comércio*; Agrianual edition; Agribuisiness Intelligence: São Paulo, Brazil, 2018.

3. Parent, L.E.; Parent, S.É.; Rozane, D.E.; Amorim, D.A.; Hernandes, A.; Natale, W. Unbiased Approach to Diagnose the Nutrient Status of Guava, Petrolina, Pernambuco, Brazil. In *III International Symposium on Guava and other Myrtaceae*; International Society for Horticultural Science (ISHS): Petrolina, Brazil, 2012; Volume 959, pp. 145–159.
4. Sousa, R.O.; Ermani, P.R. *Manual de Calagem e Adubação para os Estados do Rio Grande do Sul e de Santa Catarina*; Comissão de Química e Fertilidade do Solo—RS/SC (Sociedade Brasileira de Ciência do Solo—Núcleo Regional Sul): Santa Maria, Brazil, 2016.
5. Van Maarschalkerweerd, M.; Husted, S. Recent in Fast Spectroscopy for Mineral Plant Analysis. *Front. Plant Sci.* **2015**, *6*, 169. [CrossRef]
6. Munson, R.D.; Nelson, W.L. Principles and Practices in Plant Analysis. In *Soil Testing and Plant Analysis*; Westerman, R.L., Ed.; Soil Science Society of America: Madison WI, USA, 1990; pp. 359–387.
7. Natale, W.; Neton, A.J.L.; Rozane, D.E.; Parent, L.E.; Corrêa, M.C. Mineral Nutrition Evolution in the Formation of Fruit Tree Rootstocks and Seedlings. *Rev. Bras. Frutic.* **2018**, 46. [CrossRef]
8. Rombolà, A.D.; Sorrenti, G.; Marodin, G.A.B.; De Pieri, A.Z.; Barca, E. Nutrition and Soil Management in Stone Fruit Trees in Temperate Regions. *Ciências Agrárias* **2012**, *33*, 639–654.
9. Keppel, G.; Kreft, H. Integration and Synthesis of Quantitative Data: Alexander von Humboldt's Renewed Relevance in Modern Biogeography and Ecology. *Front. Biogeogr.* **2019**, *11*, e43187. [CrossRef]
10. Coulibali, Z.; Cambouris, A.N.; Parent, S.-É. Cultivar-Specific Nutritional Status of Potato (*Solanum Tuberosum* L.) Crops. *PLoS ONE*, 2020; 15, e0230458. [CrossRef]
11. Aitchison, J. *The Statistical Analysis of Compositional Data; Monographs on Statistics and Applied Probability*; Chapman & Hall Ltd.: London, UK, 1986.
12. Lagatu, H.; Maume, L. Le diagnostic foliaire de la pomme de terre. *Ann. L'école Natl. Agron. Montp.* **1934**, *22*, 50–158.
13. Holland, D.A. The interpretation of leaf analysis. *J. Hortic. Sci.* **1966**, *41*, 311–329. [CrossRef]
14. Beaufils, E.R. *Diagnosis and Recommendation Integrated System (DRIS)*; Soil Science Bulletin #1; Department of Soil Science and Agrometeorology at the University of Natal: Pietermaritzburg, South Africa, 1973.
15. Egozcue, J.J.; Pawlowsky-Glahn, V.; Mateu-Figueras, G.; Barceló-Vidal, C. Isometric Logratio Transformations for Compositional Data Analysis. *Math. Geol.* **2003**, *35*, 279–300. [CrossRef]
16. Parent, L.E.; Dafir, M. A Theoretical Concept of Compositional Nutrient Diagnosis. *J. Am. Soc. Hortic. Sci.* **1992**, *117*, 239–242. [CrossRef]
17. Parent, S.É.; Parent, L.E.; Rozane, D.E.; Natale, W. Plant Ionome Diagnosis Using Sound Balances: Case Study with Mango (Mangifera Indica). *Front. Plant Sci.* **2013**, *4*. [CrossRef]
18. De Deus, J.A.L.; Neves, J.C.L.; M.C.M.C.; Parent, S.-É.; Natale, W.; Parent, L.E. Balance design for robust foliar nutrient diagnosis of 'Prata' banana (*Musa* spp.). *Sci. Rep.* **2018**, *2018*, 15040. [CrossRef]
19. MAPA. Normas Técnicas Específicas Para a Produção Integrada de Pêssego—NTEPI-Pêssego. 2008. Available online: https://www.normasbrasil.com.br/norma/instrucao-normativa-37-2008_76973.html (accessed on 22 June 2020).
20. May-De Mio, L.L.; Monteiro, L.B.; Motta, C.C.V.; Cuquel, F.L.; Serrat, B.M.; Kowata-Dresch, L.S. Nutrição, danos e produção de pessegueiro em sistema de Produção Integrada. *Rev. Bras. Ciências Agrárias* **2014**, *9*, 512–518. [CrossRef]
21. Campoy, J.A.; Darbyshire, R.; Dirlewanger, E.; Quero-García, J.; Wenden, B. Yield Potential Definition of the Chilling Requirement Reveals Likely Underestimation of the Risk of Climate Change on Winter Chill Accumulation. *Int. J. Biometeorol.* **2019**, *63*, 183–192. [CrossRef] [PubMed]
22. Embrapa. Uva e Vinho. Available online: https://www.embrapa.br/en/uva-e-vinho/dados-meteorologicos (accessed on 19 April 2020).
23. Embrapa. Clima-Temperado—Laboratório de Agrometeorologia. Available online: http://agromet.cpact.embrapa.br/ (accessed on 19 April 2020).
24. UFRGS. Série Meteorológica Da Estação Experimental Agronômica. Available online: https://hospedagemphp.ufrgs.br/agronomia/joomla/index.php/eea-pesquisa (accessed on 19 April 2020).
25. De Santos, H.G.; Jacomine, P.K.T.; Anjos, L.H.C.; de Oliveira, V.A.; Lumbreras, J.F.; Coelho, M.R.; de Almeida, J.A.; de Araujo Filho, J.C.; de Oliveira, J.B.; Cunha, T.J. *Brazilian Soil Classification System*; Embrapa: Brasília, Brazil, 2018.

26. Brunetto, G.; Ermani, P.R.; de Melo, G.W.B.; Nava, G.F. *Manual de calagem e adubação para os estados do Rio Grande do Sul e de Santa Catarina*; Silva, L.S., Gatiboni, L.C., Anghinoni, I., Sousa, R.O., Emani, P.R., Eds.; Comissão de Química e Fertilidade do Solo,RS/SC (Sociedade Brasileira de Ciência do Solo—Núcleo Regional Sul): Santa Maria, Brazil, 2016; pp. 189–233.
27. Embrapa. *Manual de Análises Químicas de Solos, Plantas e Fertilizantes*; Silva, F.C., Ed.; Embrapa A Agroindústria Tropical, Empresa Brasileira de Pesquisa Agropecuária: Fortaleza, Brazil, 2009.
28. Instituto Adolfo Lutz. Métodos físico-químicos para análise de alimentos /coordenadores Odair Zenebon, Neus Sadocco Pascuet e Paulo Tiglea. 2008. Available online: http://www.ial.sp.gov.br/resources/editorinplace/ial/2016_3_19/analisedealimentosial_2008.pdf (accessed on 22 June 2020).
29. Aitchison, J.; Greenacre, M. Biplots of Compositional Data. *Appl. Stat.* **2002**, *51*, 375–392. [CrossRef]
30. Egozcue, J.J.; Pawlowsky-Glahn, V. Groups of Parts and Their Balances in Compositional Data Analysis. *Math. Geol.* **2005**, *37*, 795–828. [CrossRef]
31. Tagliavini, M.; Zavalloni, C.; Rombolà, A.D.; Quartieri, M.; Malaguti, D.; Mazzanti, F.; Millard, P.; Marangoni, B. Mineral Nutrient Partitioning to Fruits of Deciduous Trees. *Acta Hortic.* **2000**, *512*, 131–140. [CrossRef]
32. Rietra, R.P.J.J.; Heinen, N.; Dimkpa, C.O.; Bindraban, P.S. Effects of Nutrient Antagonism and Synergism on Yield and Fertilizer Use Efficiency. *Commun. Soil Sci. Plant Anal.* **2017**, *48*, 1895–1920. [CrossRef]
33. Badra, A.; Parent, L.E.; Allard, G.; Tremblay, N.; Desjardins, Y.; Morin, N. Effect of leaf nitrogen concentration versus CND nutritional balance on shoot density and foliage colour of an established Kentucky bluegrass (*Poa pratensis* L.) turf. *Can. J. Plant Sci.* **2006**, *86*, 1107–1118. [CrossRef]
34. Delacour, H.; Servonnet, A.; Perrot, A.; Vigezzi, J.F.; Ramirez, J.M. La courbe ROC (receiver operating characteristic): Principes et principales applications en biologie clinique. *Ann. Biol. Clin.* **2005**, *63*, 145–154.
35. Tolosana-Delgado, R.; Talebi, H.; Khodadadzadeh, M.; van den Boogaart, K.G. On machine learning algorithms and compositional data. In Proceedings of the 8th International Workshop on Compositional Data Analysis, Terrassa, Spain, 3–8 June 2019; Egozcue, J.J., Graffelman, J., Ortego, M.I., Eds.; pp. 172–175, ISBN 978-84-947240-2-2.
36. Parent, L.E.; Rozane, D.E.; Deus, J.A.L.; Natale, W. Composition in Fruit Crops: Latest Developments. In *Fruit Crops. Diagnosis and Management of Nutrient Constraints*; Chapter 12; Srivastava, A., Hue, C., Eds.; Elsevier: New York, NY, USA, 2019.
37. Johnson, R.S. Nutrient and Water Requirements of Peach Trees. In *The Peach: Botany, Production and Uses*; Layne, D.R., Bassi, D., Eds.; CABI: Wallingford, UK, 2008; pp. 303–331.
38. Montañéz, L.; Heras, L.; Abadia, J.; Sanz, M. Plant Analysis Interpretation Based on a New Index: Deviation from Optimum Percentage (DOP). *J. Plant Nutr.* **1993**, *16*, 1289–1308. [CrossRef]
39. Nowaki, R.D.H.; Parent, S.É.; Cecílio Filho, A.B.; Rozane, D.E.; Meneses, N.B.; Silva, A.S.; Natale, W.; Parent, L.E. Phosphorus Over-Fertilization and Nutrient Misbalance of Irrigated Tomato Crops in Brazil. *Front. Plant Sci.* **2017**. [CrossRef]
40. Tagliavini, M.; Hogue, E.J.; Nielsen, G.H. Influence of Phosphorus Nutrition and Root Zone Temperature on Growth and Mineral Uptake of Peach Seedlings. *J. Plant Nutr.* **1991**, *14*, 1267–1276. [CrossRef]
41. Sinclair, A.G.; Morrison, J.D.; Smith, L.C.; Dodds, K.G. Determination of Optimum Nutrient Element Ratios in Plant Tissue. *J. Plant Nutr.* **1997**, *20*, 1069–1083. [CrossRef]
42. Wilkinson, S.R. Nutrient Interactions in Soil and Plant Nutrition. In *Handbook of Soil Science*; Sumner, M.E., Ed.; CRC Press: Boca Raton, FL, USA, 2000; pp. 89–112.
43. Parent, S.É. Why We Should Use Balances and Machine Learning to Diagnose Ionomes. *Authorea* **2020**. [CrossRef]
44. Epstein, E.; Bloom, A.J. *Mineral Nutrition of Plants: Principles and Perspectives*; Sinauer Assoc.: Sunderland, MA, USA, 2005.
45. De Wit, C.T. Resource Use in Agriculture. *Agric. Syst.* **1992**, *40*, 125–151. [CrossRef]
46. Wallace, A.; Wallace, G.A. Limiting Factors, High Yields, and Law of the Maximum. *Hortic. Rev.* **1993**, *15*, 409–448. [CrossRef]
47. Heady, E.O.; Pesek, J.T.; Brown, W.G. *Crop Response Surfaces and Economic Optima in Fertilizer Use*; Iowa Agriculture and Home Economics Experiment Station Research Bulletin: Ames, IO, USA, 1955.
48. Kyveryga, P.; Blackmer, T.M.; Morris, T.F. Disaggregating Model Bias and Variability When Calculating Economic Optimum Rates of Nitrogen Fertilization for Corn. *Agron. J.* **2007**, *99*, 1048–1056. [CrossRef]

49. Nelson, L.A.; Anderson, R.L. Partitioning of soil test-crop response probability. In *Soil Testing: Correlating and Interpreting the Analytical Results*; Stelly, M., Ed.; American Society of Agronomy: Madison, WI, USA, 1984; pp. 19–28.
50. Kyveryga, P.; Blackmer, T.M.; Morris, T.F. Alternative Benchmarks for Economically Optimal Rates of Nitrogen Fertilization for Corn. *Agron. J.* **2007**, *99*, 1057–1065. [CrossRef]
51. Kyveryga, P.; Caragea, P.C.; Kaiser, M.S.; Blackmer, T.M. Predicting Risk of Reducing Nitrogen Fertilization Using Hierarchical Models and On-Farm Data. *Agron. J.* **2013**, *105*, 85–94. [CrossRef]
52. Mestre, L.; Reig, G.; Bertrán, J.A.; Moreno, M.A. Influence of Plum Rootstocks on Agronomic Performance, Leaf Mineral Nutrition and Fruit Quality of 'Catherina' Peach Cultivar in Heavy Calcareous Soil Conditions. *Span. J. Agric. Res.* **2017**, *15*, e0901. [CrossRef]
53. Jimenez, I.M.; Mayer, N.A.; dos Santos Dias, C.T.; Scarpare Filho, J.A.; da Silva, S.R. Influence of Clonal Rootstocks on Leaf Nutrient Content, Vigor and Productivity of Young 'Sunraycer' Nectarine Trees. *Sci. Hortic.* **2018**, *235*, 279–285. [CrossRef]
54. Mayer, A.N.; Ueno, B.; da Silva, V.A.L. Teores de nutrientes foliares de pessegueiro em cinco porta-enxertos. *Rev. Bras. Frutic.* **2015**, *37*, 1045–1052. [CrossRef]
55. Galarça, S.P.; Lima, C.S.M.; Fachinello, J.C.; Pretto, A.; Vahl, L.C.; Betemps, D.L. Influence of Several Rootstocks on Foliar Nutrition in Peach. *Acta Hortic.* **2015**, *1084*, 75–84. [CrossRef]
56. Taylor, B.K.; van den Ende, B. The Nitrogen Nutrition of the Peach Tree. IV. Storage and Mobilization of Nitrogen in Mature Trees. *Aust. J. Agric. Res.* **1969**, *20*, 869–881. [CrossRef]
57. Cruz, A.F.; de Almeida, G.M.; Wadt, P.G.S.; de Carvalho Pires, M.; Ramos, M.L.G. Seasonal Variation of Plant Mineral Nutrition in Fruit Trees. *Braz. Arch. Biol. Technol.* **2019**, *62*, e19180340. [CrossRef]
58. Meyer-Aurich, A.; Weersink, A.; Gandorfer, M.; Wagner, P. Optimal Site-Specific Fertilization and Harvesting Strategies with Respect to Crop Yield and Quality Response to Nitrogen. *Agric. Syst.* **2010**, *103*, 478–485. [CrossRef]
59. Pellerin, A.; Parent, L.E.; Fortin, J.; Tremblay, C.; Khiari, L.; Giroux, M. Environmental Mehlich-III Soil Phosphorus Saturation Indices for Quebec Acid to near Neutral Mineral Soils Varying in Texture and Genesis. *Can. J. Soil Sci.* **2006**, *86*, 711–723. [CrossRef]
60. Gemtos, T.; Fountas, S.; Tagarakis, A.; Liakos, V. Precision Agriculture Application in Fruit Crops: Experience in Hand-Picked Fruits. *Procedia Technol.* **2013**, *8*, 324–332. [CrossRef]
61. Molin, J.P.; Colaço, A.F.; Carlo, E.F.; de Mattos, D., Jr. Mapping Yield, Soil Fertility and Tree Gaps in an Orange Orchard. *Rev. Bras. Frutic.* **2012**, *34*, 1256–1265. [CrossRef]
62. Appenfeller, L.R.; Lloyd, S.; Szendrei, Z. Citizen Science Improves Our Understanding of the Impact of Soil Management on Wild Pollinator Abundance in Agroecosystems. *PLoS ONE* **2020**, *15*, e0230007. [CrossRef]
63. Gibson, K.J.; Streich, M.K.; Topping, T.S.; Stunz, G.W. Utility of Citizen Science Data: A Case Study in Land-Based Shark Fishing. *PLoS ONE* **2019**, *14*, e0226782. [CrossRef]
64. Tremblay, N.; Bouroubi, Y.M.; Bélec, C.; Mullen, R.W.; Kitchen, N.R.; Thomason, W.E.; Ebelhar, S.; Mengel, D.B.; Raun, W.R.; Francis, D.D.; et al. Corn Response to Nitrogen Is Influenced by Soil Texture and Weather. *Agron. J.* **2012**, *104*, 1658–1671. [CrossRef]
65. Gallardo, R.K.; Grant, K.; Brown, D.J.; McFerson, J.R.; Lewis, K.M.; Einshorn, T.; Sazo, M.M. Perceptions of Precision Agriculture Technologies in the U.S. Fresh Apple Industry. *HortTechnology* **2019**, *29*, 151–162. [CrossRef]
66. Morris, T.F.; Murrell, T.S.; Beegle, D.B.; Camberato, J.J.; Ferguson, R.B.; Grove, J.; Ketterings, Q.; Kyveryga, P.M.; Laboski, C.A.; McGrath, J.M.; et al. Strengths and Limitations of Nitrogen Rate Recommendations for Corn and Opportunities for Improvement. *Agron. J.* **2018**, *110*, 1–37. [CrossRef]
67. Parent, S.É.; Dossou-Yovo, W.; Ziadi, N.; Tremblay, G.; Pellerin, A.; Parent, L.E. Corn Response to Banded P Fertilizers with or without Manure Application in Eastern Canada. *Agron. J.* **2020**. [CrossRef]
68. Abdi, D. Predicting Soil Phosphorous and Other Properties Using near Infrared Spectroscopy. *Soil Sci. Soc. Am. J.* **2012**, *76*, 2318–2326. [CrossRef]
69. Ge, Y.; Thomasson, J.A.; Sui, R. Remote Sensing of Soil Properties in Precision Agriculture: A Review. *Front. Earth Sci.* **2011**, *5*, 229–238. [CrossRef]
70. Nduwamungu, C.; Ziadi, N.; Parent, L.E.; Tremblay, G.F.; Thuriès, L. Opportunities for, and Limitations of, near Infrared Reflectance Spectroscopy Applications in Soil Analysis: A Review. *Can. J. Soil Sci.* **2009**, *89*, 531–541. [CrossRef]

71. Salt, D.E.; Baxter, I.; Lahner, B. Ionomics and the Study of the Plant Ionome. *Annu. Rev. Plant Biol.* **2008**, *59*, 709–733. [CrossRef]
72. Villas-Boas, P.; Franco, M.A.; Martin-Neto, L.; Gollany, H.T.; Milori, B.M.D.P. Applications of Laser-induced Breakdown Spectroscopy for Soil Characterization, Part II: Review of Elemental Analysis and Soil Classification. *Eur. J. Soil Sci.* **2019**. [CrossRef]