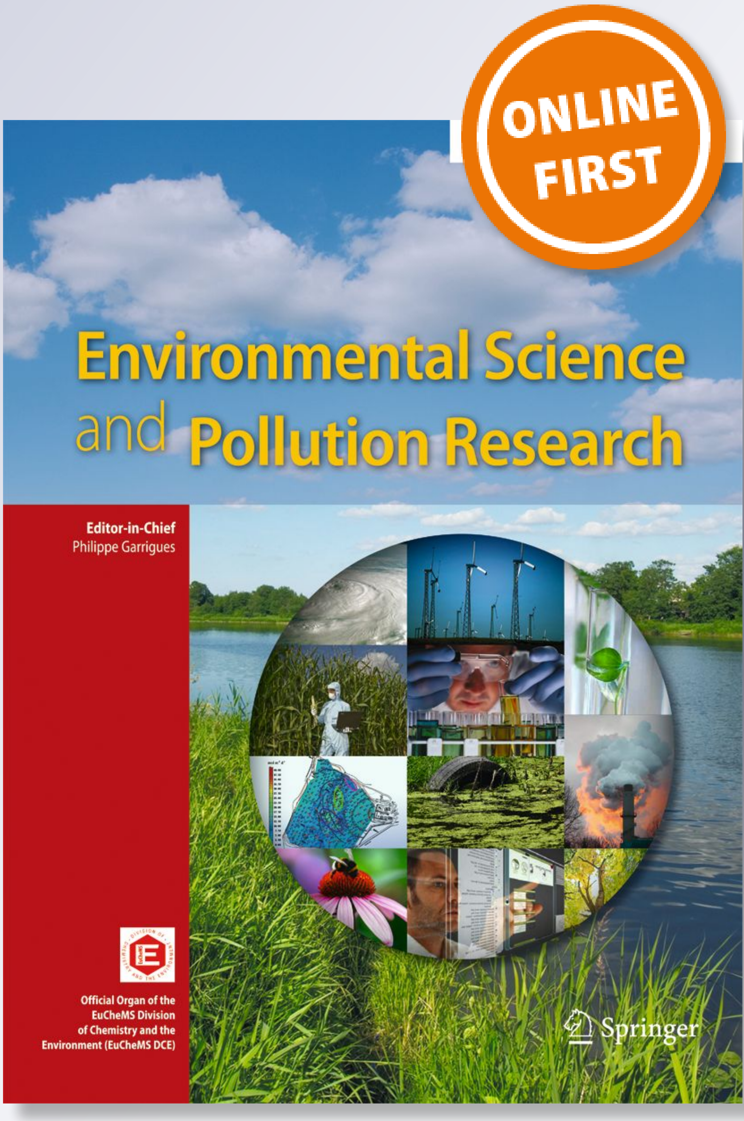
Non-essential element concentrations in brown grain rice: Assessment by advanced data mining techniques

Roxana Villafañe, Melisa Hidalgo, Analía Piccoli, Eduardo Marchevsky & Roberto Pellerano

Environmental Science and Pollution Research

ISSN 0944-1344

Environ Sci Pollut Res
DOI 10.1007/s11356-017-9017-2



Springer

Environ Sci Pollut Res
DOI 10.1007/s11356-017-9017-2

ADVANCES IN ENVIRONMENTAL BIOTECHNOLOGY AND ENGINEERING 2016

Non-essential element concentrations in brown grain rice: Assessment by advanced data mining techniques

Roxana Villafañe[1] **· Melisa Hidalgo**[2] **· Analía Piccoli**[2] **· Eduardo Marchevsky**[1] **· Roberto Pellerano**[2]

Received: 25 November 2016 / Accepted: 10 April 2017

Abstract The concentrations of 17 non-essential elements (Al, As, Ba, Be, Cd, Ce, Cr, Hg, La, Li, Pb, Sb, Sn, Sr, Th, Ti, and Tl) were determined in brown grain rice samples of two varieties: Fortuna and Largo Fino. The samples were collected from the four main producing regions of Corrientes province (Argentina). Quantitative determinations were performed by inductively coupled plasma mass spectrometry (ICP-MS), using a validated method. The contents of As, Be, Cd, Ce, Cr, Hg, Pb, Sb, Sn, Th, and Tl were very low or not detected in most samples. The non-essential element levels detected were in line with studies conducted in rice from different parts of the world. In order to characterize the influence of geographical origin in the samples, the following classification methods were carried out: linear discriminant analysis (LDA), k-nearest neighbors (k-NN), partial least squares discriminant analysis (PLS-DA), support vector machine (SVM) and random forests (RF). The best performance was obtained by using RF (96%) and SVM (96%). The results reported here showed the variation in the non-essential element profiles in rice grain depending on the geographical origin.

Keywords *Oryza sativa* · ICP-MS · Mineral content · LDA · K-NN · PLS-da · RF · SVM

Responsible editor: Philippe Garrigues

✉ Roberto Pellerano
roberto.pellerano@comunidad.unne.edu.ar

[1] INQUISAL (CONICET), Av. Ejército de los Andes, 950 San Luis, Argentina

[2] IQUIBA-NEA (UNNE-CONICET), Av. Libertad, 5470 Corrientes, Argentina

Introduction

Rice (*Oryza sativa* L.) is one of the most important staple food crops, and represents a fundamental ingredient for the diet in Asia. Currently, the development of valued-added rice products that capture the nutritional and functional attributes of rice, expand domestic markets of this cereal at worldwide. In this context, it is known that brown rice is rich in minerals and vitamins, making it a nutritionally valuable food. However, the process that produces brown rice removes only the outermost layer, preserving the most part of mineral contents in the outer kernels, including essential and non-essential trace elements (Hansen et al. 2012).

Ideally, rice should be free from contaminants such as non-essential or potentially toxic elements, which can be harmful to humans, causing serious health problems. These elements are ubiquitous in the environment, but can be locally concentrated by anthropogenic or other natural factors. In general, the non-essential elements are of concern because they are characterized by their long biological half-lives, non-biodegradable nature and persistent accumulation in different body parts. In addition, rice is grown mainly under flooded conditions, a practice that can expose rice plants to high-levels of possible pollutants that could be present in the irrigating water or added agrochemicals (Khan et al. 2010; Van Geen et al. 2006).

In the last few years, global trade has expanded with advances in production technologies leading to record yields in high-quality rice. Among Latin American countries, Argentina, together with Brazil and Paraguay, appears as one of the leading rice-producing and exporting countries (USDA, 2017). The possible contamination of rice during the growing process is an important issue in the food industry. At the moment, reported information on the trace element compositions of rice from Argentina and its correlation with their site of

Published online: 20 April 2017

production is unavailable. In this context, data mining techniques combined with mineral content is an effective way to characterize and classify agricultural products from different regions (Potorti et al., 2013, Salvo et al. 2014; Di Bella et al. 2015; Barbosa et al. 2016; Maione et al. 2016).

The objectives of this study were to (1) quantify the levels of non-essential elements in samples of brown grain rice produced in the northeastern region of Argentina; and (2) evaluate the compositional data by advanced data mining techniques to characterize the influence of the geographical origin on multi-elemental profile.

Materials and methods

Rice samples

A total of 84 brown rice samples of two varieties—Fortuna and Largo Fino—were analyzed for their non-essential element contents. The samples were provided from local growers in ten rice fields grouped according to their producing areas from Corrientes province (Argentina), namely: Ituzaingó (IT) (27° 35′ S, 56° 41′ W), Mercedes (ME) (29° 11′ S, 59° 04′ W), Santa Lucía (SL) (28° 59′ S 59° 06′ W), and Santo Tome (ST) (28° 34′ S, 56° 06′ W). Information on the different soil types is available in the soil map published by National Agricultural Technology Institute of Argentina (INTA, 2017). Currently, Corrientes province is the major rice-producing region of Argentina, with 44.5% of total production in 2014/2015 (Ministerio de Agroindustria—Argentina, 2016). The samples were obtained between March and June 2014–2015. After collection, all samples were immediately stored in zipped bags until analysis.

Sample pretreatment

Rice samples were homogenized in a blender and sieved through a nylon sieve of 1 mm mesh size before acid digestion. Microwave-assisted acid digestion was carried out using an Ethos-One microwave reaction system (Milestone, Chicago, USA) equipped with segmented high-pressure rotor that accommodates 10 PTFE vessels at a time. Powdered samples (~500 mg) were weighed into the microwave oven PTFE vessels and 6 mL of HNO_3 (≥69% *w*/*w*) plus 2 mL of H_2O_2 (≥30% *v*/*v*) were added to each vessel. The procedure was completed by using the following microwave program: 250 W for 1 min, 0 W for 2 min, 250 W for 5 min, 400 W for 5 min, and 600 W for 5 min. Sample solutions were diluted to 25 mL in decontaminated volumetric flasks with deionized water, after cooling.

Reagents

Analytical reagent grade HNO_3 65% and ultrapure grade 30% (*m*/*m*) H_2O_2 were acquired from Sigma (St. Louis, MO, USA). Nitric acid was additionally cleaned by sub-boiling distillation. Standard solutions used for external calibrations were prepared from single-element solutions acquired from Sigma-Aldrich (St Loius, MO, USA). Ultrapure deionized water (18.3 MΩ cm^{-1}) was obtained from a Milli-Q system (Millipore, Bedford, MA, USA).

Instrumentations

The quadrupole inductively coupled plasma mass spectrometer used in this work was an Agilent 7700 (Agilent Technologies, Santa Clara, CA, USA) equipped with a Peltier-cooled quartz spray chamber and a standard torch (2.5 mm i.d). This instrument was equipped with a MicroMist glass concentric pneumatic nebulizer. The ICP-MS instrument operational parameters were as follows: RF power—1550 W; plasma gas flow—14 L/min; auxiliary gas flow—0.8 L/min; nebulizer flow rate—0.95 L/min. The following isotopes (*m*/*z* ratios) were selected for analytical determinations after an interference study: ^{7}Li, ^{9}Be, ^{27}Al, ^{47}Ti, ^{52}Cr, ^{75}As, ^{88}Sr, ^{111}Cd, ^{118}Sn, ^{121}Sb, ^{137}Ba, ^{139}La, ^{140}Ce, ^{202}Hg, ^{205}TI, ^{208}Pb, and ^{232}Th. Rhodium was added as internal standard to a final concentration of 10 μg/L. The instrument was tuned daily for maximum signal sensitivity, stability and for low oxides and doubly charged ion formation using tuning solutions for ICP-MS (Agilent; 10 μg/L of Li, Y, Ce, Tl, and Co in 2% HNO_3).

Quality assurance

Five multi-element standard solutions were used for external calibration (0 to 20 μg/L). The limits of detection (LODs) were determined for each element as 3.3 times the standard deviation of blank signal. The limits of quantification (LOQs) were calculated as 10 times the standard deviation blank signal. Precision and trueness studies were carried out performing 10 independent determinations on a Certified Reference Material (CRM 1573a). In routine analyses, two replicates of each sample were analyzed and the trace element concentrations were evaluated as mean of two measurements, with less than 10% repeatability value. In every mineralizing cycle, blank and CRM were analyzed as quality control assay.

Statistical analysis

Data analysis was performed using data mining methods with caret package (Kuhn 2012) in R-project software platform (R Development Core Team 2014).

A data matrix (84 × 8) whose rows are the different samples analyzed (cases) and whose columns correspond to the content of the trace elements (variables, Li, Al, Ti, Cr, Sr, Ba, La, and Pb) was built. To verify the existence of relationships between samples, an unsupervised dimensionality reduction method was used, principal component analysis (PCA), aiming to represent high-dimensional data in lower-dimensional spaces in a faithful way. Then, in order to build a model of possible patterns in data, five supervised learning algorithms were compared according to their performance. Specifically, a supervised learning algorithm takes a set of labeled input data and known responses to the data (output), and trains a model to generate reasonable classifications to new data. The learning algorithms used were from different types, two linear models given the possibility of generating easily interpretable models, and three non-linear models that generally have better performances for the classification of samples such as those analyzed in this work. The linear methods performed were linear discriminant analysis (LDA) and partial least squares discriminant analysis (PLS-DA). The non-linear methods were: a distance-based method as k-nearest neighbors (k-NN), a tree-based method as random forest (RF) and the support vector machine (SVM) with radial bases kernel (RF).

In order to compare the performance of the studied classification algorithms, the dataset of results was divided into two subsets, training and testing, the selection of the samples for each group was realized by stratified sampling. The training set was formed by 70% of the total samples (n = 59), and the remaining 30% samples (n = 25) constituted the testing set. The samples included in the training set are used exclusively to optimize the parameters that are necessary for each method. The samples included in the test set are used exclusively to evaluate the performance of each method against an unknown set of samples. It is important not to use the testing dataset in any way in building or tuning the built models, because the accuracy rate of the model is not an unbiased estimation (Williams 2011). All these learning methods, excepting LDA, need to optimize several parameters during training stage; to avoid bias. Tenfold cross validation (repeated five times) on the training set was used to build optimal classifiers. In the tenfold cross validation, the data matrix is randomly split into 10 mutually exclusive subsets, tenfold cross validation in this case. The classifier is trained and tested 10 times, and then, the cross-validation estimate of metrics is over-tested 10 times.

LDA is a typical discriminating method with hard modeling features, this method has proven to work well in practice to tackle multiclass problems and have no-hyperparameters to tune.

PLS-DA is a method based on partial least squares regression for discrimination analysis. This technique finds the latent variables that best discriminate the different groups (Moncayo et al. 2015). The number of significant components was optimized for the PLS regression (ncomp = 5).

K-NN is a non-parametric technique which stores all cases and classify new cases based on similarity measures, the classification of new cases is based on the most of the k-nearest neighbor (Huang et al. 2014). In this work, the best average accuracy was obtained when using k = 5 obtained by tenfold cross validation optimization stage.

SVM is a learning machine method that creates a mapping of the training data into a high-dimensional space, and constructs a classifier in that space (Batista et al. 2012). The classification performance of SVM model depends on the kernel parameter, settings C and cost parameter γ that was tuned for this data set (Bona et al. 2016). After cross validation during training step, the parameters C = 0.5 and γ = 0.096 were used for classification.

RF is an ensemble learning method that combines a bootstrap aggregation to form unlimited sample sets, and generate a tree decision predictors for classification (Hernández-Pereira et al. 2015). RF has one tuning parameter, *mtry*. The mtry could range from 1 to p, p being the total number of variables. In this work, 500 trees were built, with an mtry = 3 obtained by tenfold cross validation.

Finally, the performance of the models were evaluated in terms of the following metrics per-class (Williams 2011): sensitivity (proportion of true positives that are correctly identified), specificity (proportion of true negatives that are correctly identified as negative cases); and global accuracy (average of total cases correctly identified among the total number of cases examined).

Results and discussion

Validation of results

Acceptable analytical figures of merit, i.e., linearity, limits of detection (LODs), limits of quantification (LOQs), precision and trueness, were obtained (Table 1). Accuracy was measured analyzing a certified reference material (CRM 1573a) and the obtained recoveries percentages were ranged between 95 and 103%. Precision, expressed as percent relative standard deviation, ranged between 2.1 and 6.2%.

Mineral contents in rice samples

In this work, we determined 17 trace elements (Al, As, Ba, Be, Cd, Ce, Cr, Hg, La, Li, Pb, Sb, Sn, Sr, Th, Ti, and Tl) in brown grain rice samples collected from four production areas in Corrientes province (Argentina). The analytical results are showed (fresh weight basis) in Table 2.

Based on the analytical results obtained, the trace elements Be, Ce, Hg, Sb, Sn, Th, and Tl, were undetectable (below LOD) in all samples, while As and Cd were detectable only in five samples from ME at very low levels. On the other hand,

Table 1 Analytical figures of merit of ICP-MS for rice samples and quality control based on the analysis of certified reference material

Trace elements	R^2	LOD (ng/g)	LOQ (ng/g)	Recovery (%)
Al	0.9845	92	278	97
As	0.9945	7	22	98
Ba	0.9867	11	35	103
Be	0.9934	1	4	97
Cd	0.9812	2	6	95
Ce	0.9978	1	3	99
Cr	0.9956	5	15	101
Hg	0.9976	12	36	99
La	0.9978	2	7	99
Li	0.9956	3	9	100
Pb	0.9856	4	13	99
Sb	0.9956	1	4	97
Sn	0.9989	1	4	98
Sr	0.9898	9	28	98
Th	0.9987	1	3	98
Ti	0.9956	2	8	97
Tl	0.9975	1	3	99

the non-essential trace elements Al, Ba, La, Li, Sr, and Ti were detected in all samples.

Aluminum was the most abundant element, detected at levels above 1 mg/kg in all samples. This element occurs naturally in the environment. The contents of Al in this work ranged between 3.35–6.87 mg/kg, the major concentrations were detected in ME samples and the lowest in samples from SL. These levels were higher than rice consumed in France (Millour et al. 2011), and brown rice from Spain and Portugal (Pinto et al. 2016).

Barium and Sr were detected at levels higher than 0.01 mg/kg. The highest Ba contents were determined in rice grains collected from IT, while Sr contents were higher in samples from SL. The average of Ba concentration was at similar levels determined in rice samples from Brazil (Maione et al. 2016), but slightly lower than levels in rice from France (Millour et al. 2012). Regarding the Sr contents were at similar levels than samples from Spain and Portugal (Pinto et al. 2016).

Table 3 Loadings of experimental variables on significant principal components (PCs) for rice samples

Variables	PC1 (39%)[a]	PC2 (17%)[a]	PC3 (12%)[a]
Al	0.52	0.06	0.01
Ba	−0.25	−0.29	0.09
Cr	0.01	0.54	0.66
La	−0.13	−0.40	0.72
Li	0.41	−0.40	0.10
Pb	0.49	−0.08	0.18
Sr	−0.35	−0.43	0.04
Ti	−0.34	0.33	0.05

[a] Percentage of total variance

The trace elements La and Li were at lower levels than 0.12 mg/kg in all samples. The highest La concentrations were detected in IT samples while the lowest values were detected in ST samples. In the same way, Li highest contents were detected in IT samples and, the lowest values were measured in SL samples.

The contents of Ti ranged from 1.84 mg/kg in ME samples to 5.30 mg/kg in SL samples. These average contents were at similar levels than white rice, parboiled rice and wild rice from Spanish and Portuguese markets, but lower than brown rice (Pinto et al. 2016).

Chromium can be either toxic or essential to human body as trace element, depending on its oxidation state. While Cr(VI) is a pollutant and potentially toxic, Cr(III) is essential for humans and it is present in most vegetal tissues at low

Table 2 Concentrations of trace element in whole grain rice samples

Elements (mg/kg)	Sampling region (mean ± SD)			
	Ituzaingó (IT)	Mercedes (ME)	Santa Lucía (SL)	Santo Tomé (ST)
Al	5.46 ± 1.86	6.87 ± 1.41	3.35 ± 0.83	6.51 ± 1.75
As	ND	0.008 ± 0.004	ND	ND
Ba	0.32 ± 0.20	0.22 ± 0.06	0.20 ± 0.01	0.25 ± 0.04
Cd	ND	0.010 ± 0.002	ND	ND
Cr	0.028 ± 0.006	0.045 ± 0.012	0.04 ± 0.020	0.05 ± 0.008
La	0.02 ± 0.003	0.02 ± 0.004	0.02 ± 0.001	0.02 ± 0.003
Li	0.09 ± 0.02	0.08 ± 0.01	0.02 ± 0.01	0.03 ± 0.02
Pb	0.02 ± 0.004	0.03 ± 0.005	0.008 ± 0.001	0.03 ± 0.007
Sr	0.36 ± 0.06	0.23 ± 0.07	0.38 ± 0.04	0.25 ± 0.06
Ti	2.74 ± 0.34	2.50 ± 0.48	3.37 ± 0.83	3.12 ± 0.43

ND not detected

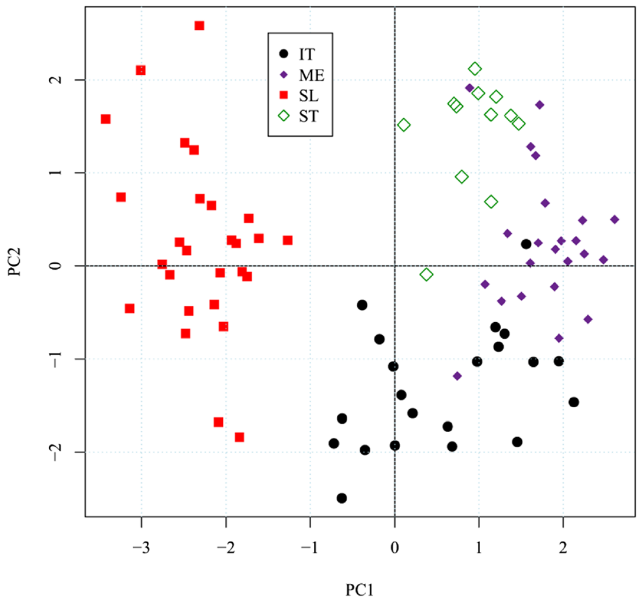


Fig. 1 Principal component analysis on elemental concentrations based on 8 variables in 84 rice samples from different geographical origins

concentrations. The ICP-MS technique cannot determine the oxidation state of this element in samples. The Cr contents ranged from 0.03–0.05 mg/kg in analyzed samples. The highest Cr concentrations were obtained from brown grain rice samples produced in ST, whereas the lowest was in samples produced in IT. The mean concentrations were consistent with different types of rice from Arabia (Shraim 2014), but lower than brown rice from Spain (Pinto et al. 2016) and brown rice from Jamaican market (Antoine et al. 2012).

Finally, the presence of Pb, which is a toxic element, deserves special attention. Lead is a natural occurring chemical compound. Lead can also occur as a residue in food because of its presence in the environment. The bioaccumulation of this metal in the body can lead to harmful effects over time. The lowest and the highest Pb concentrations were 0.008 mg/kg in samples from SL and 0.03 mg/kg in samples from ST. These results were at lower levels than the reported by EFSA for rice grains commercialized in Europe, indicating that consumption is safe (EFSA 2010). At this point, it is important to note that in this study the samples analyzed were brown grains of rice while in normal diets this cereal is consumed as polished grains. In general, it is demonstrated that the rice polishing process tends to diminish the contribution of minerals from them. In addition, the average Pb concentration was in accordance with the results of Arabian white rice (Shraim 2014), and Brazil (Maione et al. 2016) but higher than all types of grain rice from Spain (Pinto et al. 2016).

Statistical analysis

Firstly, PCA was performed on standardized data to ensure the variables have an equal contribution to the results. The correlation matrices were suitable for PCA (KMO values of 0.695; Bartlett's test of sphericity values equal to 274.3 with p values <0.001). The first three components accounted for 68% of total variance. PC1 showed higher positive contributions from Al and Pb; and higher negative loadings for Sr and Ti. PC2 indicated high contributions from Cr and Ti and negative contributions for Sr, Li and La. PC3 showed high positive contributions of La and Cr (Table 3).

Fig 1 shows the obtained PCA score-plot. This figure shows the sample distribution in 2-D space of the first two PCs (56% of total variance). As can be seen, samples are distributed into two principal groups. One group composed only of samples from SL with negative values on PC1, while the other group corresponded to samples from other regions. In the second group, high overlapping among clusters is observed; however, the IT cluster shows a slight difference from the others. The results indicated that the rice samples from SL have different trace element profile, correlated with higher Sr and Ti contents and low Al and Pb concentrations.

Once the exploratory analysis was completed, advanced data mining techniques were used to characterize the influence of the geographical origin on the multi-elemental profile of different rice samples. Table 4 shows the results obtained for the different algorithms. The highest global accuracy was achieved by the most accurate classification algorithms, the RF and SVM models. However, for some groups, LDA achieved an accuracy of 100% (SL), for the other groups (ME and ST) the accuracy was considerably lower. These results confirm that SL rice showed unique characteristics from trace element composition point of view. The average

Table 4 Performance measures detailed for LDA, k-NN, PLS-DA, SVM, and RF models computed

Method	LDA		k-NN		PLS-DA		RF		SVM	
	Sens	Spec	Sens	Spec	Sens	Spec	Sens	Spec	Sens	Spec
IT	83%	94%	66%	100%	83%	88%	83%	100%	83%	100%
ME	100%	94%	83%	100%	66%	94%	100%	94%	100%	94%
SL	100%	100%	100%	86%	100%	100%	100%	100%	100%	100%
ST	66%	100%	100%	95%	100%	100%	100%	100%	100%	100%
Accuracy	92%		87%		87%		96%		96%	

LDA linear discriminant analysis, *k-NN* k-nearest neighbors, *PLS-DA* partial least squares discriminant analysis, *RF* random forest, *SVM* support vector machine accuracy achieved by LDA was 92%, which is an acceptable result for a linear method. The global accuracies of PLS-DA and k-NN were at the lower levels, because the two techniques presented problems to discriminate the samples from IT and ME. Finally, SVM and RF misclassified only one sample from IT as ME in the test set. From these results, we demonstrated here that the generated information on non-essential element contents in brown grain rice from Corrientes province can be modeled according to the geographical origin of samples, enabling to build predictive models for origin authentication. However, more samples from these regions are necessary to develop more accurate models.

Conclusions

In the present study, the contents of As, Be, Cd, Ce, Cr, Hg, Pb, Sb, Sn, Th, and Tl were very low or not detected in most samples. The non-essential elements contents were generally in accordance with data reported by other works conducted in rice from different parts of the world. The exploratory analysis of the results was able to find significant differences among samples. The methods RF and SVM showed good performance to classify samples according to their geographical origin. Finally, the results here reported showed the variations in the non-essential element profiles in rice grain depending on the geographical origin.

Acknowledgments The authors thank Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Universidad Nacional del Nordeste (SGCyT-UNNE), and Universidad Nacional de San Luis (UNSL) for financial support and fellowships.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

Antoine JMR, Hoo Fung LA, Grant CN, Dennis HT, Lalor GC (2012) Dietary intake of minerals and trace elements in rice on the Jamaican market. J Food Compos Anal 26:111–121

Barbosa RM, Silva de Paula E, Paulelli AC, Moore AF, JMO S, Batista BL, Campiglia AD, Barbosa F Jr (2016) J Food Compos Anal 45:95–100

Batista BL et al (2012) Multi-element determination in Brazilian honey samples by inductively coupled plasma mass spectrometry and estimation of geographic origin with data mining techniques. Food Res Int 49:209–215

Bona E et al (2016) Support vector machines in tandem with infrared spectroscopy for geographical classification of green Arabica coffee. LWT - Food Sci Technol doi. doi:10.1016/j.lwt.2016.1004.1048

Di Bella G et al (2015) Geographical discrimination of Italian honey of multi-element analysis with a chemometric approach. J Food Compos Anal 44:25–35

EFSA (2010) Panel on contaminants in the food chain (CONTAM). Scientific opinion on lead in food. EFSA J 8:1570–1717

Hansen TH et al (2012) Losses of essential mineral nutrients by polishing of rice differ among genotypes due to contrasting grain hardness and mineral distribution. J Cereal Sci 56:307–315

Hernández-Pereira EM, Álvarez-Estévez D, Moret-Bonillo V (2015) Automatic classification of respiratory patterns involving missing data imputation techniques. Biosyst Eng 138:65–76

Huang X, Teye E, Owusu-Sekyere JD, Takrama J, Sam-Amoah LK, Yao L, Firempong CK (2014) Simultaneous measurement of titratable acidity and fermentation index in cocoa beans by electronic tongue together with linear and non-linear multivariate technique. Food Anal Methods 7:2137–2144

Instituto Nacional de Tecnología Agropecuaria (INTA) (2017) Mapa de suelos de la República Argentina. http://intagobar/sites/default/files/script-tmp-inta-200-suelospdf Accesed 19 March 2017

Khan MA, Islam MR, Panaullah GM, Duxbury JM, Jahiruddin M, Loeppert RH (2010) Accumulation of arsenic in soil and rice under wetland condition in Bangladesh. Plant Soil 333:263–274

Kuhn M (2012) Caret: classification and regression training. R package version 5:15–044

Maione C, Batista BL, Campiglia AD, Barbosa F Jr, Barbosa RM (2016) Classification of geographic origin of rice by data mining and inductively coupled plasma mass spectrometry. Comput Electron Agric 121:101–107

Millour S et al (2012) Strontium, silver, tin, iron, tellurium, gallium, germanium, barium and vanadium levels in foodstuffs from the second French total diet study. J Food Compos Anal 25:108–129

Millour S et al (2011) Pb, Hg, Cd, As, Sb and Al levels in foodstuffs from the 2nd French total diet study. Food Chem 126:1787–1799

Moncayo S, Manzoor S, Navarro-Villoslada F, Caceres JO (2015) Evaluation of supervised chemometric methods for sample classification by laser induced breakdown spectroscopy. Chemom Intell Lab Syst 146:354–364

Pinto E, Almeida A, Ferreira IMPLVO (2016) Essential and non-essential/toxic elements in rice available in the Portuguese and Spanish markets. J Food Compos Anal 48:81–87

Potorti AG, Di Bella G, Lo Turco V, Rando R, Dugo G (2013) Non toxic and potentially toxic elements in Italian donkey milk by ICP-MS and multivariate analysis. J Food Compos Anal 31:161–172

The United States Department of Agriculture (USDA) (2017) Grain: World Markets and Trade https://apps.fas.usda.gov/psdonline/circulars/grain-rice.pdf. Accessed 19 March 2017

R Development Core Team (2014) R: A Language and Environment for Statistical Computing, 3.1.1 edn. R Foundation for Statistical Computing Vienna, Austria

Salvo A et al (2014) Statistical characterization of heavy metal contents in *Paracentrotus lividus* from Mediterranean Sea. Nat Prod Res 28:718–726

Shraim AM (2014) Rice is a potential dietary source of not only arsenic but also other toxic elements like lead and chromium. Arab J Chem. doi:10.1016/j.arabjc.2014.1002.1004

Van Geen A et al (2006) Impact of irrigating rice paddies with groundwater containing arsenic in Bangladesh. Sci Total Environ 367:769–777

Williams G (2011) Data mining with rattle and R: the art of excavating data for knowledge discovery. Springer Science & Business Media, New York