University of Naples Federico II

Department of Agricultural Sciences

Ph.D. Food Science

XXXIII Cycle

Geochemical and spectroscopic fingerprinting for authentication and geographical traceability of high-quality lemon fruits

Coordinator: Prof. Amalia Barone

Amalia Barone

Ph.D. Candidate

Ruggiero Luigi

Tutor**: Prof. Paola Adamo**

Paola Adamo

Co-tutor: **Dott. Carmine Amalfitano**

Carmine Amalfitano

Academic year: 2019-2020

Leave this world a little better than you found it

Baden-Powell (1941)

Contents

Abstract

Geochemical (mineral element and Sr isotope ratio) and spectroscopical fingerprinting (Near Infrared Spectroscopy) were proposed to authenticate and track the two high-quality lemon fruits from the Campania region (Limone di Sorrento PGI and Limone Costa d'Amalfi PGI) to protect them from frauds. Considering the geochemical indicators, we built different chemometric discriminant models based on mineral profile and $^{87}Sr/^{86}Sr$ isotope ratio. These two techniques were applied to discriminate fruits from different territorial scales, small territorial scales (region scale), and large territorial scales. The results of different discriminant models applied on mineral profiles of lemon juices, both on a small and large territorially scale, showed good discrimination according to provenance, especially for non-essential elements as Rb, Ba, Sr, Ti, and Co. These same elements have shown a good correlation with cultivation soils and stability between the two production years. It is worth noting that although, the performance of the whole elemental profile gave a better result than the profile of the non-essential elements, the reliability of the two models, calculated as the ratio between the percentage of correctly validated and classification samples, was similar. In addition, the Sr isotope ratio had shown a clear differentiation among the fruits from the Campania region and extra-regional samples, and by analysis of $^{86}Sr/^{87}Sr$ of soils, it was clear that the strontium isotope ratio of lemon juices was closely related to that of the bioavailable fractions of the soil. Furthermore, combining both isotopic and mineral profiles in lemon juices by a low-level data fusion approach, the results showed a better clustering according to geographical origins than the two-determination taken separately, although on an explorative level. In addition, the spectroscopical data (NIR) on intact lemon fruits showed the strong influence of environmental growing conditions on the samples. For this, the application of Linear Discriminant Analysis (LDA) models suggested building the discrimination models according to origins (PGI and not PGI productions) based on one production year. In the same way, the application of MLR models, that showed a strong relationship between quality properties of lemon fruits and NIR spectra, suggested the applicability of this technique to build predictive models for the quality properties. In addition, on a part of the total samples collected only in 2019 (intact lemons and juices), have been successfully applied two different chemometrics models i.e., LDA and Partial Least Square Discriminant Analysis (PLS-DA). The results showed better provenance discrimination using the lemon juices than the intact lemons. Comparing the results obtained, of the two approaches used, the results of geochemical fingerprinting have shown more stability for discriminate lemon fruits derived from two different production years, especially for not essential elements. However, considering the various vantages of the application of NIR spectroscopy (non-destructive, rapid, and cheap) and the results obtained, this technique can be used for rapid screening of samples in order to verify the quality and origins of lemon fruits during the year. The

study of the pedoclimatic features was fundamental to understand the nature of discriminating variables, in both approaches. Additional research should be conducted to include a greater number of lemon farms (or sampling points) in the PGI area and to enlarge the existing database including lemon samples from other regions and validate the models built. These discriminant models based on geochemical and spectroscopical profiles of lemon fruits could substantially contribute to implementing a blockchain system for Campanian lemon traceability, providing real-time information not only to the final consumers but also to manufacturers, distributors, and retailers.

1. INTRODUCTION

The local food is an important resource for the sustainability of territory and local communities. The link between food and sustainability can be explained in terms of benefits about the creation of rural communities and enhancing the local environment, economy, and culture, implementing local tourism. In a defined area, food, can improve sustainable local tourism generating a multiplier effect, which would benefit the local economy, reducing the carbon footprint of the tourism industry by the "buying local" and increasing the attractiveness of the place (Sims, 2009). Moreover, according to the FAO guide for promoting quality linked to geographical origin and sustainable geographical indications (Arfini et al., 2010), the local food generates local identity as an interaction between natural resources and production system in a specific landscape. In this contest, the local products are defined for specific quality attributes that are linked to the territory where they are produced and associated with a geographical indication (GI) that identifies them. In fact, the GIs are a result of technical, social, and economic interactions, including local knowledge and cultural traditions (Belletti et al., 2017). All this represents not only a heritage to be preserved but also an important market value because the consumer becomes increasingly interested in quality linked to geographical origin, traditions, and typicity. Due to this, since 1992, the Council of the European Communities, defined as GI a product originating in a specific region, place or country that possesses a specific quality, reputation attributable to that geographical origin and the production and/or processing of which take place in the defined geographical area (EEC No 2080/92, 1992; EEC No 2081/92, 1992). In addition to PGI (Protected Geographical Indication) products, this regulation defined another quality mark as POD (Protected Designation of Origin). For twenty years, this regulation has undergone not substantial changes and improvements. The first changes were introduced with regulations 509 and 510 of 2006 (ECC No 509/06, 2006; ECC No 510/06, 2006). The latter gave the possibility to third countries (non-EU) to register their products. In 2012, with the regulation EEC No 1151/2012, another quality label as TSG (Traditional Speciality Guaranteed) and two optional quality terms i.e. "mountain products" and "product of island farming" were introduced, leaving the definition of Protected Geographical Indication (PGI) practically unchanged. These specifications give greater typicality to increase the value of products originating from disadvantaged areas such as mountain and island areas. As reported in the EU database for registration of food products (EC, 2020), 748 PGIs have been registered so far. Thirty-four countries (27 EU and 7 extra EU) had their products awarded as PGI. The first three countries that have been awarded the major number of PGI products are France (19.2%), Italy (17.3%) and Spain (12.4%). The high typicality, the small productions, the small communities and the close link with the territories give PGI products a high economic value especially because the consumers are willing to pay more for protected foods, for many reasons including hygiene, safety and food quality (Dias & Mendes, 2018). According to ISMEA (2019) the Italian PGI, PDO, and TSG products, in 2018, were worth €16.2 billion (+6.0% of 2017), with an export value of € 9.0 billion (+2.5% of 2017). Due to their high economic value, PGI products are often subjected to frauds. One of the most common fraudulent phenomena for *Made in Italy* is the *Italian sounding*. The Italian Sounding consists in the use (on labels and packaging) of denominations, geographical references, images, colour combinations and brands that evoke Italy and in particular, some of its most famous typical products, to promote the marketing of products (especially agri-food) leading deceptively into believing that they are authentic Italians (Carreño & Vergano, 2017). In 2019, in Italy, the frauds value was over 301 million euros with about 72 million kg goods seized (ICQRF, 2019). From this derives the need to give a defined authentication of a single product for distinguish it from other fraudulent products. However, the use of GI allows producers to obtain market recognition and often a higher price due to their typically. The false use of GI labels by unauthorised parties is detrimental to consumers and legitimate producers. For this, the use of analytical techniques for determining the geographical origin of products is a key point to protect consumers, agricultural farmers, retailers, and administrative authorities.

1.1 AUTHENTICATE TO TRACK

Food authentication is defined as the process that verifies if the food complies with its label description. This may include origin, production method, or processing technologies (Danezis, et al., 2016). The goal of food authentication, among others, is to fight the food frauds defined as *"any deliberate action of businesses or individuals to deceive others in regard to the integrity of food to gain undue advantage. Types of food fraud include but not limited to adulteration, substitution, dilution, tampering, simulation, counterfeiting, and misrepresentation"* (FAO, 2018). Different techniques can be used for authenticating food products and, on basis of their characteristics, to find the unique profile able to distinguish from products of different nature or provenance. This thesis will give special attention to the concept of food counterfeiting. The choice of the most suitable analytical technique for the determination of food origins can depend on the nature of the food to track. To demonstrate the link between food and territory of origin, there is a need to use appropriate parameters, namely traceability indicators (Lees, 2003). These traceability indicators can be associated with the type of technique used. Thus, the techniques can consider the relationship between food and environmental growing conditions of raw materials, including production methods (metabolomics, spectroscopy, volatile organic compounds, and sensory analyses) or the cultivation soil (isotope analyses, elemental profile) in order to link the food at its geographical provenance, defining a specific terroir of the products.

Isotopic Analyses

The use of stable isotope ratios as potential geochemical tracers in studies of traceability is generally based on two different groups of elements: the light element isotope ratios ($^2H/^1H$, $^{13}C/^{12}C$, $^{15}N/^{14}N$, $^{18}O/^{16}O$, and $^{34}S/^{32}S$) and the heavy element isotope ratios ($^{87}Sr/^{86}Sr$, $^{143}Nd/^{144}Nd$, $^{208}Pb/^{204}Pb$, $^{206}Pb/^{204}Pb$, $^{207}Pb/^{204}Pb$). The theory behind this approach is that the isotopes of a same element have similar chemical properties, because of the same number of protons and electrons, but different physical properties, due to the different mass. The ratio between two different isotopes of the same element can be affected by isotopic fractionation and radioactive decay. The isotope fractionation can be affected during physicochemical processes such as chemical reactions, diffusion, and transformation between the different state of matter (solid, liquid, and gaseous phases). The degree of fractionation is generally dependent on the relative difference in mass between the stable isotopes of a given element and the temperature at which the fractionation process occurs (Nakano, 2016). Light elements, participating of different natural or artificial biogeochemical phenomena, can be affected by large isotopic fractionations. For this, due to the large relative difference in mass between their various isotopes, this technique was one of the most used for both the establishment of geographic authenticity of foodstuffs and identification of food adulteration (Georgiou & Danezis, 2017). Several papers reported this technique useful to detect the nature of water in pressed e tap water diluted fruit juices, for identification of adulteration (Koziet et al., 1993, 1995) or, again in fruits juices, for the establishment of geographic authenticity due to different quality of the harvested fruits that they subjected to significant geographical and seasonal variations depending on growing environmental conditions of the production area (Bat et al., 2012; Rossmann, 2001). In contrast, for heavy element isotope ratios, such as the 87Sr/86Sr isotope ratio, the fractionation and the seasonal and annual variation are more or less negligible during the plant uptake (Zannella et al., 2017), and the bioavailable fraction in soils provides a direct link from soil to the plant. In particular the studies about the isotopes of heavy elements provided important results in several application fields (Georgiou & Danezis, 2017) as food traceability. This applications are based on the principle of radioactive decay and consequently relative abundance of daughter elements (the process of radioactive β decay of ^{87} Rb in ^{87} Sr was illustrated in (Georgiou & Danezis, 2017)). This means that the relative abundances of isotopes of Sr, Nd (α decay), and Pb (α and β decay for ^{206}Pb, ^{207}Pb, and ^{208}Pb, respectively) within rocks and soil minerals are constant in the environmental systems on time scales of $<10^4$ years, but can vary significantly depending on their age of formation (Nakano, 2016). Consequently, products cultivated in soils of different age will have different isotope ratios. For determination of the isotopic ratios, various methods can be used. Isotope Ratio Mass Spectrometry

(Muccio & Jackson, 2009) (IRMS), Multi Collector – Inductively Coupled Plasma – Mass Spectrometry (Wieser et al., 2012) (MC-ICP-MS), and Thermal Ionization Mass Spectrometry (Heumann et al., 1995) (TIMS). For the determination of light isotopes ratios, the IRMS interfaced with Elemental Analyser, Pyrolysis, Equilibration devices, is used, while MC-ICP-MS and TIMS measure heavy isotopes.

Multielement Fingerprinting

Elemental fingerprinting can be applied to food authenticity and traceability as a potential geochemical tracer (Danezis, Tsagkaris, Camin, et al., 2016). In this case, traceability is based on the principle that the mineral composition of agri-food products reflects the bioavailable and mobile elements present in the cultivation soils (Anderson & Smith, 2005; Cheajesadagul et al., 2013; Gonzálvez et al., 2011; Kelly et al., 2002; Laursen et al., 2014; Simpkins et al., 2000). In particular, several papers show the direct relationship with cultivation soil (Oddone et al., 2009; Pepi et al., 2017; Podio et al., 2013; Shen et al., 2013; Zhao, Zhang, et al., 2017). Furthermore, the traceability of animal-derived food (meat, milk, cheese, honey) is based on the principle that their mineral composition is mostly related to the mineral content of fodder and vegetation the animals eat and water they drink (Batista et al., 2012; Erasmus et al., 2018; Franke et al., 2008; Park et al., 2018). Potentially, all chemical elements could be used to define the food's unique multi-element profile able to distinguish agri-products grown in different sites. The multi-element fingerprinting was successfully applied to determine the geographical origin of Calabrian chili peppers (Naccarato et al., 2016), Tropea Red Onion (Furia et al., 2011), Clementine of Calabria (Benabdelkamel et al., 2012), Italian virgin olive oils (Benincasa et al., 2007), Spanish PDO paprika (Palacios-morillo et al., 2014), Welsh Onions (Riyama et al., 2004), orange juices (Simpkins et al., 2000), lemon juices (Pellerano et al., 2008), lemon pulps (Potortì et al., 2018), tea leaves (Zhao, Yu, et al., 2017; Zhao, Zhang, et al., 2017) and asparagus (Richter et al., 2019). In all these studies it was shown that about 40% of all analysed elements discriminated against the samples according to geographical provenance. Among these elements 31% were considered essentials for plant growth as macro (K, Mg, P, Ca) and micronutrients (Cu, Fe, Mn, B, Zn) (Bean et al., 2011), 17% were considered beneficial (at low levels) (Ni, Al, Na) (Kabata-Pendias, 2011; Kronzucker et al., 2013), 53% were not-essential (Tl, U, Li, Sr, Se, Sb, Rb, Pb, Ga, Cs, Cr, Co, Cd, Br, Be, Ba, As) and 15% were rare-earths (Pr, Sc, Sm, Nd, Y, Ho, Gd, La, Eu, Dy, Ce). It is worth noting that the mineral composition of agri-food products is affected by several factors as plant species, agronomic management (fertilization, pest and weed control management), crop rotation, climate, and soil fertility (Rosen & Allan, 2007). In soil, several factors can influence the phytoavailability of elements such as pH, organic matter, soil texture, and chemical form of the element (Kabata-Pendias, 2011). For this reason, the discrimination with macro and micronutrients and beneficial elements could be more affected by external factors such as agronomic management compared to not essential elements and rare-earths. In particular, the rare-earths have been recognized as useful geochemical tracers (Laveuf & Cornu, 2009) and their profiles applied successfully in several studies (Farmaki et al., 2012; Ma et al., 2016; Pisciotta et al., 2017; Spalla et al., 2009; Wyttenbach et al., 1998) due to their variation according to parent material history and weathering state of the soil. The elemental fingerprint of foods can be measured by different analytical techniques. Inductively Coupled Plasma- Mass Spectroscopy(Beauchemin, 2008) (ICP-MS) and Inductively Coupled Plasma- Atomic Emission Spectroscopy(Dean, 2005) (ICP-AES) are the most widely applied.

Metabolomics Fingerprinting

The food metabolome comprises a great variety of components that are very different from a chemical point of view, indeed belonging to a variety of chemical classes as sugars, organic acids, amines, amino acids, peptides, nucleic acids, steroids, or phenolic compounds (Castro-puyana & Herrero, 2013). An accurate metabolomic analyses needs, for the variability of components, to involve a combination of different techniques, as separation techniques such as gas (GC) or liquid (LC) chromatography, and detection techniques such as Mass spectrometry (MS), Nuclear Magnetic Resonance (NMR), Infra-Red Spectroscopy (IR), etc., followed by chemometric analysis of the multivariate data to extract relevant information (Lees, 2003). In this contest, the metabolomic approach can be divided into targeted (a specific group of metabolites), or untargeted determinations (fingerprinting). In this paragraph, we will consider only GC-MS, LC-MS, NMR, and PTR-MS while IR will be considered separately in the next sections. In food, the metabolite profile can be influenced by the environmental growing conditions and ripeness of raw materials as much as the food processing stages. This could be useful for the identification of a unique metabolomic profile of a product able to give a specific identity. This identity could be related to the provenance of the same product (geographical indication) or the distinctive of a specific food process. The metabolomic profile by NMR for geographical origin was successfully applied to Argentine citrus (Salazar et al., 2018), red wine(Mazzei et al., 2010) and royal jelly(Mazzei et al., 2020) . The use of volatile and aromatic compounds (VOCs) Is based on the principle that characteristics of volatile fingerprinting of products (types and contents) could depend on environmental (climate growing condition, culture, and cultivar), genetic(Zhang et al., 2019) and nutritional treatments(Ghanbari et al., 2019). The VOCs profile has been successfully applied for authentication of provenance of coffee(Yener et al., 2014), saffron(Ghanbari et al., 2019), cheese (Boscaini et al., 2003) cocoa beans(Acierno et al., 2019), olive oils (Araghipour et al., 2008), white truffle (Gioacchini et al., 2008) and South African lambs(Erasmus et al., 2017) by PTR-MS and GC-MS.

Near Infrared Spectroscopy Fingerprinting

The Near Infrared spectroscopy (NIR) is based on the principle that in normal conditions the vibrational states of nuclei in constant motion (bending and vibrating) can be altered in precise ways by absorption of radiation (Workman & Weyer, 2007). According to the quantum theory, only certain vibrational alterations are permitted (Downey, 2013). Therefore, the frequencies are characteristic of various bond types and they are dependent on the specific molecular environment in which a molecule is located. For this, the measure of the radiation absorption over the precise wavelength (or frequency) range can potentially provide very specific information about the types and numbers of molecules present in any given sample (Downey, 2013). This specific absorption follows the Beer–Lambert law i.e., the concentration of an absorber is proportional to the sample's absorbance (for homogeneous and non-scattering liquid samples). The theoretical specific information is not so easy to be found for biological samples, as food, due to the nature and number of molecules that are highly variable. Therefore, spectra derived from biological matrixes contain information about the complete molecular composition present in each sample and can be a fingerprint of it. Furthermore, considering the complexity of matrices, in addition to the fundamental molecular absorptions there will also be other absorptions arising from combinations and overtones. The first one occurs when more than two fundamental vibrations are excited simultaneously, while the second is observed when a vibrational state is excited from v=0 to v=2 (first overtone), or v=0 to v=3 (second overtone) (Holmes, 2020). Overtone and combination band absorbances are important in NIR spectra (ranges 12,820-4000 cm^{-1}) and are found at 5917–5698 cm^{-1} (first over- tone), 8873–8547 cm^{-1} (second overtone), and 11,834–11,390 cm^{-1} (third overtone) (Workman & Weyer, 2007). The radiation of NIR spectroscopy ranges between 12820 and 4000 cm^{-1}. It was discovered in 1800, and applied only around 1950s, when with the work of Norris revealed its potential use in food and agriculture fields (Downey, 2013). The major focus of this technique lies with the quantitative determination for authentication and quality parameters (sugar, acidity, pH etc..) of different agri-food products as bayberry fruit (Xie et al., 2011), mango (Jha et al., 2014; Saranwong et al., 2004), orange (Cavaco et al., 2018), pomegrate (Arendse et al., 2017), satsuma mandarin (Gòmez et al., 2006; Kawano et al., 1993; Miyamoto & Kitano, 2009), Arabica coffee (Ribeiro et al., 2011), safflower (Elfadl et al., 2010), Jeju citrus fruits (Lee et al., 2004), wheat flour (Kahriman et al., 2011), tomatoes(Li et al., 2021), lemon (Reddy et al., 2016), fava beans (Wang et al., 2013), strawberry (Amodio et al., 2017), cherry and apricot (Carlini et al., 2000), apples (Gabioud et al., 2008) and Onions(Birth et al., 1985); however this technique can also be applied for the determination of mineral content of agri-products as successfully done on forages(Clark et al., 1987), legumes(Cozzolino & Moron, 2004), red grape(Cozzolino et al., 2011a) and straw(Huang et al., 2009). According to Schmitt et al. (2014), the use of NIR spectroscopy for determination of the mineral composition of food is a good green alternative compared to the traditional determination by ICP techniques (use of acids and chemicals, time consuming, waste, etc..), although, for studies reported in this review, this determination remain in many cases at the academic level because they lack of a deep evaluation of their prediction capabilities. In addition, the use of NIR spectroscopy can give a unique molecular fingerprint able to discriminate samples of different nature or origin (Luykx & van Ruth, 2008). The use of NIR spectroscopy to determine the geographical origin of a food product, has been successfully applied on many agri-products such as honey (Latorre et al., 2013; Woodcock et al., 2009); *dong quai* or female ginseng (Woo et al., 2005); wheat (Zhao et al., 2013); wine (Cozzolino et al., 2011b; Iu et al., 2006); olive oil (Jiménez-carvelo et al., 2019; Hirri et al., 2016); sea cucumber (Guo et al., 2018), black tea (Ren et al., 2013), apple (Eisenstecken et al., 2019), Sorrento walnut (Amendola et al., 2020) and coffee (Monteiro et al., 2018). The highlight of NIR spectroscopy is that it can be used as rapid screening techniques owing to its easy use and low cost, but above all for its on-line application in the food industry (Jha et al., 2014). Usually, the NIR spectra are very similar, they are dominated by the water (overtone bands of the OH-bonds) (Polesello et al., 1981). For these reasons, the NIR spectra need multivariate statistical to extract useful information from spectra (Blanco & Villarroya, 2002). The pre-processing of NIR spectra is the necessary step before the chemometric analysis (Rinnan et al., 2009). As reported in Rinnan et al. (2009), pre-processing techniques can be divided into two categories. The first one, namely the scatter corrections, can be used to reduce the undesirable scatter effects. The most widely used are Multiplicative Scatter Correction (MSC), and Standard Normal Variate (SNV) (Roggo et al., 2007; Blanco & Villarroya, 2002). The second one, namely the derivate correction, can be used for baseline correction, first derivate, and, both baseline correction and linear trend, second derivate. The derivate methods more broadly applied are Norris-Williams derivatives and Savitzky-Golay (SG) polynomial derivates (Rinnan et al., 2009). Both techniques can be used to decrease the detrimental effect on the signal-to-noise ratio to improve both the explorative analysis and the classification and discrimination models.

1.2 CHEMOMETRICS TOOLS FOR TRACEABILITY

The food could be intended as a complex matrix, due to different interactions with many constituent factors. The chemical and physical composition of vegetable food products could be dependent on cultivation soils, climatic conditions, and the physiological status of plants. The chemical and physical composition of animal food products (meats, milks, honey, fishes, and eggs) mostly related to the type of alimentation, farming condition, and animal health status. The chemical and physical composition of manufactured food products (cheese, cold meats, and beverages) could be depended on the nature of raw materials and processing. As even the simplest food is a complex matrix, the way to investigate its chemistry cannot be other than multivariate. The multivariate chemometrics approach, for traceability studies, is not only recommended but essential (Marini, 2013) due to a large number of variables to consider. The multivariate chemometrics approach of chemical and physical fingerprinting of food products can be divided in Explorative Analyses (EA), Discriminant Classification Methods (DCMs), and Class-Modelling Methods (CMMs). The aim of EA is to explore the tendency of natural groupings and eventual natural clusters of samples. One of the first definition of EA was proposed by Tukey (1962) as "*Procedures for analysing data, techniques for interpreting the results of such procedures, ways of planning the gathering of data to make its analysis easier, more precise or more accurate, and all the machinery and results of (mathematical) statistics which apply to analysing data*." Among the different techniques (Descriptive Statistics, Frequency Histograms, Box and Whisker Plots, and Cluster Analyses) used for EA certainly plays an important role the Principal Component Analysis (PCA) considered one of principal tools of exploratory multivariate data analysis. The PCA is a bilinear decomposition technique capable of reducing large variables into few principal components (PCs). This reduction is achieved by a linear transformation preserving data variance. Scatter plots in a Euclidean plane may represent the PCs. The PCs score plot can be used to detect sample differences and similarities whereas the loadings plot (weight of each original variable on the PCs) is useful to distinguish the variables that are major contributors to a PC and that lead to the natural clustering of groups (data not labeled, unsupervised method) within the dataset (Li Vigni et al., 2013; Naccarato et al., 2016). Both DCMs and CMMs aim to build classification methods, based on the results of the explorative analysis, to define the groups according to determinate characteristics (provenance, species, etc...) and find the variables that better discriminate the groups of samples. Unlike the PCA, DCMs and CMMs are supervised methods. Linear discriminant analysis (LDA), first proposed by Fisher in 1936 (Fisher, 1936), is the oldest supervised classification method. It is a linear technique, where the decision boundaries separating the classes in the multidimensional space of the variables can be linear surfaces (hyperplanes). The aim of LDA is to discriminate between classes (two or more) by maximizing the variance between classes (the distance between the centroids) and minimizing the variance within each class (the dispersion of the samples within the categories). LDA defines discriminant functions that are linear combinations of the original variables for optimizing the separation. The model creates a centroid that is the mean position of all points in all directions. The prediction result for the test set is obtained by the projection of the new samples according to the minimal distance to the centroid of each class (Gaiad et al., 2016). One of the principal assumptions of LDA is that the variance/covariance for the different categories are homogeneous. If this condition is met, then it is possible to combine the individual matrices. In case they are not homogeneous, it is necessary to extend the hyperplanes of LDA in a hyperquadrics space considering not the variance/covariance of the pooled matrix of LDA but the variance/covariance matrices of the individual categories (Quadratic discriminant analysis, QDA)(Marini, 2013). The ratio between samples/variables is important to avoid overfitting problems. As suggested in Lees (2003) the number of samples should be more than five times (no less than three times) the number of variables to be estimated. For LDA, which considers a pooled variance/covariance matrix, this means that the total number of objects should be at least five times the number of variables. Whereas the QDA computes a variance-covariance matrix for each category; this makes it a more powerful method than LDA, but this also means that each category should have a number of samples at least five times higher than the number of variables(Lees, 2003). The LDA and QDA are based on matrices where the number of samples is larger than the number of variables. However, in many instrumental methods, as Infrared spectroscopy, the number of variables largely exceeds the number of analysed samples. For this, LDA and QDA cannot be applied to process these matrices. Alternatively, a reduction of variables by PCA can be useful for applying LDA and QDA on PCs scores (Canonical Variates Analysis, CVA). The problem of this approach is that not always the PCs extracted, based on the original variance, are relevant for discrimination. The Partial Least-Squares (PLS) algorithm can be useful to overcome this limitation reformulating the classification problem in terms of a regression equation (Barker & Rayens, 2003).The PLS Discriminant Analysis (PLS-DA) works on the introduction of a dummy dependent matrix Y (number of samples *X* number of class), which codifies for a class belonging of the samples. The variable importance in projection (VIP) scores and the regression coefficients can be used to estimate the contribution of each variable to the model in the PLS-DA. The VIP score is an index that has been specifically introduced to express whether a predictor is significant in the definition of the F latent vectors model for the prediction of a particular response (Marini, 2013). Unlike DCMs, the CMMs regards the individual categories rather than what makes one class different from the others (Marini, 2013). Wold & Sjöström, (1977) presented the first class-modelling method, namely Soft Independent Modelling Class Analogy

(SIMCA). The SIMCA works on the variability of each class, based on PCA, dividing the space into as many subspaces as categories, and each point belongs to one and only one category. For this the samples that will be predicted by such method must belong to one of the categories that have been used to build the models; if not, they will anyway be assigned to one of them (nearer). Different applications of SIMCA methods are described in Lees (2003). The mathematical foundations of all described statistical methods were summarized by (Georgiou & Danezis, 2017; Hair et al., 2014; Lees, 2003; Marini, 2013).

References

Acierno, V., Fasciani, G., Kiani, S., Caligiani, A., & van Ruth, S. (2019). PTR-QiToF-MS and HSI for the characterization of fermented cocoa beans from different origins. *Food Chemistry*, *289*(September 2018), 591–602. https://doi.org/10.1016/j.foodchem.2019.03.095

Amendola, L., Firmani, P., Bucci, R., Marini, F., & Biancolillo, A. (2020). Authentication of Sorrento walnuts by NIR spectroscopy coupled with different chemometric classification strategies. *Applied Sciences (Switzerland)*, *10*(11). https://doi.org/10.3390/app10114003

Amodio, M. L., Ceglie, F. G., Chaudhry, M. M. A., Piazzolla, F., & Colelli, G. (2017). Potential of NIR spectroscopy for predicting internal quality and discriminating among strawberry fruits from different production systems. *Postharvest Biology and Technology*, *125*(March), 112–121. https://doi.org/10.1016/j.postharvbio.2016.11.013

Anderson, K. A., & Smith, B. W. (2005). Use of chemical profiling to differentiate geographic growing origin of raw pistachios. *Journal of Agricultural and Food Chemistry*, *53*(2), 410–418. https://doi.org/10.1021/jf048907u

Araghipour, N., Colineau, J., Koot, A., Akkermans, W., Rojas, J. M. M., Beauchamp, J., Wisthaler, A., Märk, T. D., Downey, G., Guillou, C., Mannina, L., & Ruth, S. van. (2008). Geographical origin classification of olive oils by PTR-MS. *Food Chemistry*, *108*(1), 374–383. https://doi.org/10.1016/j.foodchem.2007.10.056

Arendse, E., Fawole, O. A., Magwaza, L. S., Nieuwoudt, H. H., & Opara, U. L. (2017). Development of calibration models for the evaluation of pomegranate aril quality by Fourier-transform near infrared spectroscopy combined with chemometrics. *Biosystems Engineering*, *159*, 22–32. https://doi.org/10.1016/j.biosystemseng.2017.04.004

Arfini, F., Vandecandelaere, E., Belletti, G., & Marescotti, A. (2010). Linking people, places and products. A guide for promoting quality linked to geographical origin and sustainable geographical indications. In *Food and Agriculture Organization of the United Nations (FAO) and SINER-GI*. www.foodquality-origin.org/guide/guide.pdf

Barker, M., & Rayens, W. (2003). Partial least squares for discrimination. *Journal of Chemometrics*, *17*(3), 166–173. https://doi.org/10.1002/cem.785

Bat, K. B., Vidrih, R., Nečemer, M., Vodopivec, B. M., Mulič, I., Kump, P., & Ogrinc, N. (2012). Characterization of Slovenian apples with respect to their botanical and geographical origin and agricultural production practice. *Food Technology and Biotechnology*, *50*(1), 107–116.

Batista, B. L., da Silva, L. R. S., Rocha, B. A., Rodrigues, J. L., Berretta-Silva, A. A., Bonates, T. O., Gomes, V. S. D., Barbosa, R. M., & Barbosa, F. (2012). Multi-element determination in Brazilian honey samples by inductively coupled plasma mass spectrometry and estimation of geographic origin with data mining techniques. *Food Research International*, *49*(1), 209–215. https://doi.org/10.1016/j.foodres.2012.07.015

Bean, F., Laursen, K. H., Schjoerring, J. K., Olesen, J. E., Askegaard, M., Halekoh, U., & Husted, S. (2011). Multielemental Fingerprinting as a Tool for Authentication of Organic. *Journal of Agricultural and Food Chemistry*, *59*, 4385–4396.

Beauchemin, D. (2008). Inductively Coupled Plasma Mass Spectrometry. *Analytical Chemistry*, *80*(12), 4455–4486.

Belletti, G., Marescotti, A., & Touzard, J. M. (2017). Geographical Indications, Public Goods, and Sustainable Development: The Roles of Actors' Strategies and Public Policies. *World Development*, *98*, 45–57. https://doi.org/10.1016/j.worlddev.2015.05.004

Benabdelkamel, H., Di Donna, L., Mazzotti, F., Naccarato, A., Sindona, G., Tagarelli, A., & Taverna, D. (2012). Authenticity of PGI "clementine of Calabria" by multielement fingerprint. *Journal of Agricultural and Food Chemistry*, *60*(14), 3717–3726. https://doi.org/10.1021/jf2050075

Benincasa, C., Lewis, J., Perri, E., Sindona, G., & Tagarelli, A. (2007). Determination of trace element in Italian virgin olive oils and their characterization according to geographical origin by statistical analysis. *Analytica Chimica Acta*, *585*, 366–370. https://doi.org/10.1016/j.aca.2006.12.040

Birth, G. S., Dull, G. G., Renfroe, W. T., & Kays, S. J. (1985). Nondestructive Spectrophotometric Determination of Dry-Matter in Onions. *Journal of the American Society for Horticultural Science*, *110*(2), 297–303.

Blanco, M., & Villarroya, I. (2002). NIR spectroscopy: A rapid-response analytical tool. *TrAC - Trends in Analytical Chemistry*, *21*(4), 240–250. https://doi.org/10.1016/S0165-9936(02)00404-1

Boscaini, E., Van Ruth, S., Biasioli, F., Gasperi, F., & Märk, T. D. (2003). Gas chromatography-olfactometry (GC-O) and proton transfer reaction-mass spectrometry (PTR-MS) analysis of the flavor profile of Grana Padano, Parmigiano Reggiano, and Grana Trentino cheeses. *Journal of Agricultural and Food Chemistry*, *51*(7), 1782–1790. https://doi.org/10.1021/jf020922g

Carlini, P., Massantini, R., & Mencarelli, F. (2000). Vis-NIR Measurement of Soluble Solids in Cherry and Apricot by PLS Regression and Wavelength Selection. *Journal of Agricultural and Food Chemistry*, *48*, 5236–5242. https://doi.org/10.1021/jf000408f

Carreño, I., & Vergano, P. R. (2017). Geographical indications , “ Food Fraud ” and the Fight Against “ Italian sounding ” Products. *European Journal of Risk Regulation*, *7*(2), 416–420.

Castro-puyana, M., & Herrero, M. (2013). Metabolomics approaches based on mass spectrometry for food safety , quality and traceability. *Trends in Analytical Chemistry*, *52*, 74–87. https://doi.org/10.1016/j.trac.2013.05.016

Cavaco, A. M., Pires, R., Antunes, M. D., Panagopoulos, T., Brázio, A., Afonso, A. M., Silva, L., Lucas, M. R., Cadeiras, B., Cruz, S. P., & Guerra, R. (2018). Validation of short wave near infrared calibration models for the quality and ripening of ‘Newhall’ orange on tree across years and orchards. *Postharvest Biology and Technology*, *141*(March), 86–97. https://doi.org/10.1016/j.postharvbio.2018.03.013

Cheajesadagul, P., Arnaudguilhem, C., Shiowatana, J., Siripinyanond, A., & Szpunar, J. (2013). Discrimination of geographical origin of rice based on multi-element fingerprinting by high resolution inductively coupled plasma mass spectrometry. *Food Chemistry*, *141*(4), 3504–3509. https://doi.org/10.1016/j.foodchem.2013.06.060

Clark, D. H., Mayland, H. F., & Lamb, R. C. (1987). Mineral Analysis of Forages with near Infrared Reflectance Spectroscopy1. *Agronomy Journal*, *79*(3), 485. https://doi.org/10.2134/agronj1987.00021962007900030016x

Cozzolino, D., Cynkar, W., Shah, N., & Smith, P. (2011a). Quantitative analysis of minerals and electric conductivity of red grape homogenates by near infrared reflectance spectroscopy. *Computers and Electronics in Agriculture*, *77*(1), 81–85. https://doi.org/10.1016/j.compag.2011.03.011

Cozzolino, D., Cynkar, W. U., Shah, N., & Smith, P. A. (2011b). Can spectroscopy geographically classify Sauvignon Blanc wines from Australia and New Zealand? *Food Chemistry*, *126*(2), 673–678. https://doi.org/10.1016/j.foodchem.2010.11.005

Cozzolino, D., & Moron, A. (2004). Exploring the use of near infrared reflectance spectroscopy (NIRS) to predict trace minerals in legumes. *Animal Feed Science and Technology*, *111*(1–4), 161–173. https://doi.org/10.1016/j.anifeedsci.2003.08.001

Danezis, G. P., Tsagkaris, A. S., Brusic, V., & Georgiou, C. A. (2016). Food authentication: state of the art and prospects. *Current Opinion in Food Science*, *10*, 22–31. https://doi.org/10.1016/j.cofs.2016.07.003

Danezis, G. P., Tsagkaris, A. S., Camin, F., Brusic, V., & Georgiou, C. A. (2016). Food authentication: Techniques, trends & emerging approaches. *TrAC - Trends in Analytical Chemistry*, *85*, 123–132. https://doi.org/10.1016/j.trac.2016.02.026

Dean, J. R. (2005). Inductively Coupled Plasma – Atomic Emission. In John Wiley & Sons (Ed.), *Practical Inductively Coupled Plasma Spectroscopy*.

Dias, C., & Mendes, L. (2018). Protected Designation of Origin (PDO), Protected Geographical Indication (PGI) and Traditional Speciality Guaranteed (TSG): A bibiliometric analysis. *Food Research International*, *103*(September 2017), 492–508. https://doi.org/10.1016/j.foodres.2017.09.059

Downey, G. (2013). Vibrational spectroscopy in studies of food origin. In *New Analytical Approaches for Verifying the Origin of Food* (pp. 94–116). Woodhead Publishing Limited. https://doi.org/10.1533/9780857097590.2.94

EC. (2020). *eAmbrosia – the EU geographical indications register*. https://ec.europa.eu/agriculture/quality/door/list.html.

ECC No 509/06. (2006). COUNCIL REGULATION (ECC) No 509/2006 on agricultural products and foodstuffs as traditional specialities guaranteed. *Official Journal of the European Union*, *L 93*, 1–11.

ECC No 510/06. (2006). Council regulation (EC/510/2006) on the protection of geographical indications and designations of origin for agricultural products and foodstuffs. *Official Journal of the European Union*, *L93*(510), 12–25.

EEC No 1151/2012. (2012). Regulation (EU) No 1215/2012 of the European Parliament and of the Council on quality schemes for agricultural products and foodstuffs. In *Fundamental Texts On European Private Law* (pp. 1–29). https://doi.org/10.5040/9781782258674.0008

EEC No 2080/92. (1992). COUNCIL REGULATION (EEC) No 2080/92 Instituting a Comunity aid Scheme for Forestry Measures in Agriculture Fearnside. *Official Journal of European Communities*, *2080*, 96–99.

EEC No 2081/92. (1992). COUNCIL REGULATION (EEC) No 2081/92 On the protection of geographical indications and designations of origin for agricultural products and foodstuffs. *Official Journal of the European Communities*, *L208*(L), 1–8.

Eisenstecken, D., Stürz, B., Robatscher, P., Lozano, L., Zanella, A., & Oberhuber, M. (2019). The potential of near infrared spectroscopy (NIRS) to trace apple origin: Study on different cultivars and orchard elevations. *Postharvest Biology and Technology*, *147*(August 2018), 123–131. https://doi.org/10.1016/j.postharvbio.2018.08.019

Elfadl, E., Reinbrechta, C., & Claupein, W. (2010). Development of near infrared reflectance spectroscopy (NIRS) calibration model for estimation of oil content in a worldwide safflower germplasm collection. *International Journal of Plant Production*, *4*(4), 259–270. http://www.gau.ac.ir/Jm/Programs/JurnalMgr/VolumArticle/EN_163_2.pdf

Erasmus, S. W., Muller, M., Alewijn, M., Koot, A. H., van Ruth, S. M., & Hoffman, L. C. (2017). Proton-transfer reaction mass spectrometry (PTR-MS) for the authentication of regionally unique South African lamb. *Food Chemistry*, *233*, 331–342. https://doi.org/10.1016/j.foodchem.2017.04.131

Erasmus, S. W., Muller, M., Butler, M., & Hoffman, L. C. (2018). The truth is in the isotopes: Authenticating regionally unique South African lamb. *Food Chemistry*, *239*, 926–934. https://doi.org/10.1016/j.foodchem.2017.07.026

FAO. (2018). *Discussion Paper On Food Integrity And Food Authenticity,Codex Alimentarius Commission* (Issue August).

Farmaki, E. G., Thomaidis, N. S., Minioti, K. S., Ioannou, E., Georgiou, C. A., & Efstathiou, C. E. (2012). Geographical Characterization of Greek Olive Oils Using Rare Earth Elements Content and Supervised Chemometric Techniques. *Analytical Letters*, *45*(8), 920–932.

https://doi.org/10.1080/00032719.2012.655656

Fisher, R. A. (1936). The use of multiple measurements in taxonomic problems. *Ann Eugenics*, *7*, 179–186.

Franke, B. M., Koslitz, S., Micaux, F., Piantini, U., Maury, V., Pfammatter, E., Wunderli, S., Gremaud, G., Bosset, J. O., Hadorn, R., & Kreuzer, M. (2008). Tracing the geographic origin of poultry meat and dried beef with oxygen and strontium isotope ratios. *European Food Research and Technology*, *226*(4), 761–769. https://doi.org/10.1007/s00217-007-0588-x

Furia, E., Naccarato, A., Sindona, G., Stabile, G., Tagarelli, A., Chimica, D., C, V. P. B. C., Arcavacata, I.-, & Cs, R. (2011). *Multielement Fingerprinting as a Tool in Origin Authentication of PGI Food Products : Tropea Red Onion*. 8450–8457. https://doi.org/10.1021/jf201556e

Gabioud, S., Baumgartner, D., Gasser, F., Kneubühler, Y., Lattmann, S., & Höhn, E. (2008). Non destructive quality measurements on apples. *Acta Horticulturae*, *796*(August), 217–224. https://doi.org/10.17660/ActaHortic.2008.796.29

Gaiad, J. E., Hidalgo, M. J., Villafañe, R. N., Marchevsky, E. J., & Pellerano, R. G. (2016). Tracing the geographical origin of Argentinean lemon juices based on trace element profiles using advanced chemometric techniques. *Microchemical Journal*, *129*, 243–248. https://doi.org/10.1016/j.microc.2016.07.002

Georgiou, C. A., & Danezis, G. P. (2017). *Food Authentication: Management, Analysis and Regulation* (Constantinos A. Georgiou and Georgios P. Danezis (ed.); Frist Edit). Jonh Willy & Sons Ltd.

Ghanbari, J., Khajoei-Nejad, G., Erasmus, S. W., & van Ruth, S. M. (2019). Identification and characterisation of volatile fingerprints of saffron stigmas and petals using PTR-TOF-MS: Influence of nutritional treatments and corm provenance. *Industrial Crops and Products*, *141*(April), 111803. https://doi.org/10.1016/j.indcrop.2019.111803

Gioacchini, A. M., Menotta, M., Guescini, M., Saltarelli, R., Ceccaroli, P., Amicucci, A., Barbieri, E., Giomaro, G., & Stocchi, V. (2008). Geographical traceability of Italian white truffle (Tuber magnatum Pico) by the analysis of volatile organic compounds. *Rapid Communications in Mass Spectrometry*, *22*, 3147–3153. https://doi.org/10.1002/rcm.3714

Gòmez, A. H., He, Y., & Annia Garcìa Pereira. (2006). *Non-destructive measurement of acidity ,* *soluble solids and firmness of Satsuma mandarin using Vis / NIR-spectroscopy techniques*. *77*, 313–319. https://doi.org/10.1016/j.jfoodeng.2005.06.036

González, A., Armenta, S., & De La Guardia, M. (2011). Geographical traceability of “arrs de Valencia” rice grain based on mineral element composition. *Food Chemistry*, *126*(3), 1254–1260. https://doi.org/10.1016/j.foodchem.2010.11.032

Guo, X., Cai, R., Wang, S., Tang, B., Li, Y., & Zhao, W. (2018). Non-destructive geographical traceability of sea cucumber (*Apostichopus japonicus*) using near infrared spectroscopy combined with chemometric methods. *Royal Society Open Science*, *5*(1), 170714. https://doi.org/10.1098/rsos.170714

Hair, F. J. . J., Black, W. C. ., Babin, B. J., & Anderson, R. E. (2014). *Multivariate Data Analysis* (Pearson Education Limited (ed.); Seventh Ed).

Heumann, K. G., Eisenhut, S., Gallus, S., Hebeda, H., Nusko, R., Vengosh, A., & Walczyk, T. (1995). Recent Developments in Thermal Ionization Mass Spectrometric Techniques for Isotope Analysis A Review. *Analyst*, *120*(May).

Hirri, A., Bassbasi, M., & Platikanov, S. (2016). FTIR Spectroscopy and PLS-DA Classification and Prediction of Four FTIR Spectroscopy and PLS-DA Classification and Prediction of Four Commercial Grade Virgin Olive Oils from Morocco. *Food Analytical Methods*, *9*(August), 974–981. https://doi.org/10.1007/s12161-015-0255-y

Holmes, A. (2020). *Combination Bands, Overtones and Fermi Resonances*. https://chem.libretexts.org/@go/page/1853

Huang, C., Han, L., Yang, Z., & Liu, X. (2009). Exploring the use of near infrared reflectance spectroscopy to predict minerals in straw. *Fuel*, *88*(1), 163–168. https://doi.org/10.1016/j.fuel.2008.07.031

ICQRF. (2019). Report 2019 Department of central inspectorate for fraud repression and quality protection of the agri-food products and foodstuffs (ICQRF). *Resting Energy Expenditure and Oxygen Consumption in Critically Ill Patients With Vs Without Sepsis*, *28*(2), 136–141.

ISMEA. (2019). *Rapporto 2019 Ismea-Qualivita sulle produzioni agroalimentari e vitivinicole italiane DOP, IGP e STG*.

Jha, S. N., Narsaiah, K., Jaiswal, P., Bhardwaj, R., Gupta, M., Kumar, R., & Sharma, R. (2014). Nondestructive prediction of maturity of mango using near infrared spectroscopy. *Journal of Food Engineering*, *124*, 152–157. https://doi.org/10.1016/j.jfoodeng.2013.10.012

Jiménez-carvelo, A. M., Lozano, V. A., & Olivieri, A. C. (2019). Comparative chemometric analysis of fluorescence and near infrared spectroscopies for authenticity con firmation and geographical origin of Argentinean extra virgin olive oils. *Food Control*, *96*(August 2018), 22–28. https://doi.org/10.1016/j.foodcont.2018.08.024

Kabata-Pendias, A. (2011). Trace elements in soils and plants. In *CRC Press*. https://doi.org/10.1201/b10158-25

Kahriman, F., Egesel, C. O., Onsekiz, Ç., & Agriculture, F. (2011). Development of a Calibration Model To Estimate Quality Traits in Wheat Flour Using Nir (Near Infrared Reflectance) Spectroscopy. *Journal of Agricultural Science*, *43*(3), 392–400.

Kawano, S., Fujiwara, T., & Iwamoto, M. (1993). Nono destructive determination of Sugar content in Satuma Mandarin using Near Infrared (NIR) Transmittance. *Journal of the Japanese Society for Horticultural Science*, *62*(2), 465–470.

Kelly, S., Baxter, M., Chapman, S., Rhodes, C., Dennis, J., & Brereton, P. (2002). The application of isotopic and elemental analysis to determine the geographical origin of premium long grain rice. *European Food Research and Technology*, *214*(1), 72–78. https://doi.org/10.1007/s002170100400

Koziet, J., Rossmann, A., Martin, G. J., & Ashurts, P. R. (1993). Determination of carbon-13 content of sugars of fruit and vegetable juices. *Analytica Chimica Acta*, *271*(1), 31–38. https://doi.org/10.1016/0003-2670(93)80548-y

Koziet, J., Rossmann, A., Martin, G. J., & Johnson, P. (1995). Determination of the oxygen-18 and deuterium content of fruit and vegetable juice water. An European inter-laboratory comparison study. *Analytica Chimica Acta*, *302*(1), 29–37. https://doi.org/10.1016/0003-2670(94)00424-K

Kronzucker, H. J., Coskun, D., Schulze, L. M., Wong, J. R., & Britto, D. T. (2013). Sodium as nutrient and toxicant. *Plant and Soil*, 1–23. https://doi.org/10.1007/s11104-013-1801-2

Latorre, C. H., Crecente, R. M. P., Martín, S. G., & García, J. B. (2013). A fast chemometric procedure based on NIR data for authentication of honey with protected geographical indication.

Food Chemistry, *141*(4), 3559–3565. https://doi.org/10.1016/j.foodchem.2013.06.022

Laursen, K. H., Schjoerring, J. K., Kelly, S. D., & Husted, S. (2014). Authentication of organically grown plants - advantages and limitations of atomic spectroscopy for multi-element and stable isotope analysis. *TrAC - Trends in Analytical Chemistry*. https://doi.org/10.1016/j.trac.2014.04.008

Laveuf, C., & Cornu, S. (2009). A review on the potentiality of Rare Earth Elements to trace pedogenetic processes. *Geoderma*, *154*(1–2), 1–12. https://doi.org/10.1016/j.geoderma.2009.10.002

Lee, K. J., Kim, G. Y., Kang, S. W., Son, J. R., Choi, D. S., & Choi, K. H. (2004). Measurement of Sugar Contents in Citrus Using Near Infrared Transmittance. *Key Engineering Materials*, *270–273*, 1014–1019. https://doi.org/10.4028/www.scientific.net/KEM.270-273.1014

Lees, M. (2003). Food authenticity and traceability. In *LWT - Food Science and Technology* (Vol. 38, Issue 1). https://doi.org/10.1016/j.lwt.2004.04.009

Li Vigni, M., Durante, C., & Cocchi, M. (2013). Exploratory Data Analysis. In *Data Handling in Science and Technology* (1st ed., Vol. 28). Copyright © 2013 Elsevier B.V. All rights reserved. https://doi.org/10.1016/B978-0-444-59528-7.00003-X

Li, X., Tsuta, M., Hayakawa, F., Nakano, Y., Kazami, Y., & Ikehata, A. (2021). Estimating the sensory qualities of tomatoes using visible and near-infrared spectroscopy and interpretation based on gas chromatography–mass spectrometry metabolomics. *Food Chemistry*, *343*, 128470. https://doi.org/10.1016/j.foodchem.2020.128470

Liu, L., Cozzolino, D., Cynkar, W. U., Gishen, M., & Colby, C. B. (2006). Geographic Classification of Spanish and Australian Tempranillo Red Wines by Visible and Near-Infrared Spectroscopy Combined with Multivariate Analysis. *Journal of Agricultural and Food Chemistry*, *54*, 6754–6759.

Luykx, D. M. A. M., & van Ruth, S. M. (2008). An overview of analytical methods for determining the geographical origin of food products. *Food Chemistry*, *107*(2), 897–911. https://doi.org/10.1016/j.foodchem.2007.09.038

Ma, G., Zhang, Y., Zhang, J., Wang, G., Chen, L., Zhang, M., Liu, T., Liu, X., & Lu, C. (2016). Determining the geographical origin of Chinese green tea by linear discriminant analysis of trace metals and rare earth elements: Taking Dongting Biluochun as an example. *Food Control*, *59*, 714–720. https://doi.org/10.1016/j.foodcont.2015.06.037

Marini, F. (2013). *Data handling in science and technology: Chemometrics in Food Chemistry*.

Mazzei, P., Francesca, N., Moschetti, G., & Piccolo, A. (2010). NMR spectroscopy evaluation of direct relationship between soils and molecular composition of red wines from Aglianico grapes. *Analytica Chimica Acta*, *673*(2), 167–172. https://doi.org/10.1016/j.aca.2010.06.003

Mazzei, P., Piccolo, A., Brescia, M., & Caprio, E. (2020). Assessment of geographical origin and production period of royal jelly by NMR metabolomics. *Chemical and Biological Technologies in Agriculture*, 1–13. https://doi.org/10.1186/s40538-020-00190-8

Miyamoto, K., & Kitano, Y. (2009). Non-Destructive Determination of Sugar Content in Satsuma Mandarin Fruit by near Infrared Transmittance Spectroscopy. *Journal of Near Infrared Spectroscopy*, *3*(4), 227–237. https://doi.org/10.1255/jnirs.73

Monteiro, P. I., Santos, J. S., Alvarenga Brizola, V. R., Pasini Deolindo, C. T., Koot, A., Boerrigter-Eenling, R., van Ruth, S., Georgouli, K., Koidis, A., & Granato, D. (2018). Comparison between proton transfer reaction mass spectrometry and near infrared spectroscopy for the authentication of Brazilian coffee: A preliminary chemometric study. *Food Control*, *91*(10), 276–283. https://doi.org/10.1016/j.foodcont.2018.04.009

Muccio, Z., & Jackson, G. P. (2009). Isotope ratio mass spectrometry. *Analyst*, *134*(2), 213–222. https://doi.org/10.1039/b808232d

Naccarato, A., Furia, E., Sindona, G., & Tagarelli, A. (2016). Multivariate class modeling techniques applied to multielement analysis for the verification of the geographical origin of chili pepper. *Food Chemistry*, *206*, 217–222. https://doi.org/10.1016/j.foodchem.2016.03.072

Nakano, T. (2016). Potential uses of stable isotope ratios of Sr, Nd, and Pb in geological materials for environmental studies. *Proceedings of the Japan Academy, Series B*, *92*(6), 167–184. https://doi.org/10.2183/pjab.92.167

Oddone, M., Aceto, M., Baldizzone, M., Musso, D., & Osella, D. (2009). Authentication and traceability study of hazelnuts from piedmont, Italy. *Journal of Agricultural and Food Chemistry*, *57*(9), 3404–3408. https://doi.org/10.1021/jf900312p

Palacios-morillo, A., Marcos, J., Alcázar, Á., & Pablos, F. De. (2014). Geographical characterization of Spanish PDO paprika by multivariate analysis of multielemental content. *Talanta*, *128*, 15–22. https://doi.org/10.1016/j.talanta.2014.04.025

Park, Y. M., Lee, C. M., Hong, J. H., Jamila, N., Khan, N., Jung, J. H., Jung, Y. C., & Kim, K. S. (2018). Origin discrimination of defatted pork via trace elements profiling, stable isotope ratios analysis, and multivariate statistical techniques. *Meat Science*, *143*(August 2017), 93–103. https://doi.org/10.1016/j.meatsci.2018.04.012

Pellerano, R. G., Mazza, S. S., Marigliano, R. A., & Marchevsky, E. J. (2008). Multielement analysis of argentinean lemon juices by instrumental neutronic activation analysis and their classification according to geographical origin. *Journal of Agricultural and Food Chemistry*, *56*(13), 5222–5225. https://doi.org/10.1021/jf073555n

Pepi, S., Sansone, L., Chicca, M., & Vaccaro, C. (2017). Relationship among geochemical elements in soil and grapes as terroir fingerprintings in Vitis vinifera L. cv. "Glera." *Chemie Der Erde*, *77*(1), 121–130. https://doi.org/10.1016/j.chemer.2017.01.003

Pisciotta, A., Tutone, L., & Saiano, F. (2017). Distribution of YLOID in soil-grapevine system (Vitis vinifera L.) as tool for geographical characterization of agro-food products. A two years case study on different grafting combinations. *Food Chemistry*, *221*, 1214–1220. https://doi.org/10.1016/j.foodchem.2016.11.037

Podio, N. S., Baroni, M. V., Badini, R. G., Inga, M., Ostera, H. A., Cagnoni, M., Gautier, E. A., García, P. P., Hoogewerff, J., & Wunderlin, D. A. (2013). Elemental and isotopic fingerprint of argentinean wheat. Matching soil, water, and crop composition to differentiate provenance. *Journal of Agricultural and Food Chemistry*, *61*(16), 3763–3773. https://doi.org/10.1021/jf305258r

Polesello, A., Giangiacomo, R., & Dull, G. G. (1981). Application of near infrared spectrophotometry to the nondestructive analysis of foods : A review of experimental results. *Food Science and Nutrition*, *18*(3), 203–230.

Potortì, A. G., Di Bella, G., Mottese, A. F., Bua, G. D., Fede, M. R., Sabatino, G., Salvo, A., Somma, R., Dugo, G., & Lo Turco, V. (2018). Traceability of Protected Geographical Indication (PGI) Interdonato lemon pulps by chemometric analysis of the mineral composition. *Journal of Food Composition and Analysis*, *69*(March), 122–128. https://doi.org/10.1016/j.jfca.2018.03.001

Reddy, N. S., D, N., & Yadav, B. K. (2016). Non-Destructive Quality Assessment of Citrus Fruits Using Ft-Near-Infrared Spectroscopy. *International Journal of Scicence Environment and TEchnology*, *5*(4), 1850–1860.

Ren, G., Wang, S., Ning, J., Xu, R., Wang, Y., Xing, Z., Wan, X., & Zhang, Z. (2013). Quantitative analysis and geographical traceability of black tea using Fourier transform near-infrared spectroscopy (FT-NIRS). *Food Research International*, *53*(2), 822–826. https://doi.org/10.1016/j.foodres.2012.10.032

Ribeiro, J. S., Ferreira, M. M. C., & Salva, T. J. G. (2011). Chemometric models for the quantitative descriptive sensory analysis of Arabica coffee beverages using near infrared spectroscopy. *Talanta*, *83*(5), 1352–1358. https://doi.org/10.1016/j.talanta.2010.11.001

Richter, B., Gurk, S., Wagner, D., Bockmayr, M., & Fischer, M. (2019). Food authentication: Multi-elemental analysis of white asparagus for provenance discrimination. *Food Chemistry*, *286*(August 2018), 475–482. https://doi.org/10.1016/j.foodchem.2019.01.105

Rinnan, Å., Berg, F. van den, & Engelsen, S. B. (2009). Review of the most common pre-processing techniques for near-infrared spectra. *TrAC - Trends in Analytical Chemistry*, *28*(10), 1201–1222. https://doi.org/10.1016/j.trac.2009.07.007

Riyama, K. A., Orita, H. H., & Asui, A. Y. (2004). Chemometric Techniques on Inorganic Elements Composition for the Determination of the Geographic Origin of Welsh Onions. *Analytical Sciences*, *20*.

Roggo, Y., Chalus, P., Maurer, L., Lema-martinez, C., & Jent, N. (2007). *A review of near infrared spectroscopy and chemometrics in pharmaceutical technologies*. *44*, 683–700. https://doi.org/10.1016/j.jpba.2007.03.023

Rosen, C. J., & Allan, D. L. (2007). *Exploring the Benefits of Organic Nutrient Sources for Crop Production and Soil Quality*. *17*(December).

Rossmann, A. (2001). Determination of stable isotope ratios in food analysis. *Food Reviews International*, *17*(3), 347–381. https://doi.org/10.1081/FRI-100104704

Salazar, M. O., Pisano, P. L., González Sierra, M., & Furlan, R. L. E. (2018). NMR and multivariate data analysis to assess traceability of argentine citrus. *Microchemical Journal*, *141*(May), 264–270. https://doi.org/10.1016/j.microc.2018.05.037

Saranwong, S., Sornsrivichai, J., & Kawano, S. (2004). Prediction of ripe-stage eating quality of mango fruit from its harvest quality measured nondestructively by near infrared spectroscopy. *Postharvest Biology and Technology*, *31*(2), 137–145. https://doi.org/10.1016/j.postharvbio.2003.08.007

Schmitt, S., Garrigues, S., & de la Guardia, M. (2014). Determination of the Mineral Composition of Foods by Infrared Spectroscopy: A Review of a Green Alternative. *Critical Reviews in Analytical Chemistry*, *44*(2), 186–197. https://doi.org/10.1080/10408347.2013.835695

Shen, S., Xia, L., Xiong, N., Liu, Z., & Sun, H. (2013). Determination of the geographic origin of rice by element fingerprints and correlation analyses with the soil of origin. *Analytical Methods*, *5*(21), 6177–6185. https://doi.org/10.1039/c3ay40700d

Simpkins, W. A., Louie, H., Wu, M., Harrison, M., & Goldberg, D. (2000). Trace elements in Australian orange juice and other products. *Food Chemistry*, *71*(4), 423–433. https://doi.org/10.1016/S0308-8146(00)00150-3

Sims, R. (2009). Food, place and authenticity: Local food and the sustainable tourism experience. *Journal of Sustainable Tourism*, *17*(3), 321–336. https://doi.org/10.1080/09669580802359293

Spalla, S., Baffi, C., Barbante, C., Turretta, C., Cozzi, G., Beone, G. M., & Bettinelli, M. (2009). Determination of rare earth elements in tomato plants by inductively coupled plasma mass spectrometry techniques. *Rapid Communications in Mass Spectrometry*, *23*(20), 3285–3292. https://doi.org/10.1002/rcm.4244

Tukey, J. W. (1962). The Future of Data Analysis. *Ann. Math. Stat.*, *33*(1), 1–67. https://projecteuclid.org/download/pdf_1/euclid.aoms/1177704711

Wang, J., Liu, H., & Ren, G. (2013). Near-infrared spectroscopy (NIRS) evaluation and regional analysis of Chinese faba bean (Vicia faba L .). *The Crop Journal*, *2*(1), 28–37. https://doi.org/10.1016/j.cj.2013.10.001

Wieser, M., Schwieters, J., & Douthitt, C. (2012). Multi-Collector Inductively Coupled Plasma Mass Spectrometry. In F. Vanhaecke & P. Degryse (Eds.), *Isotopic Analysis: Fundamentals and Applications Using ICP-MS* (First Edit, pp. 77–91).

Wold, S., & Sjöström, M. (1977). SIMCA: A Method for Analyzing Chemical Data in Terms of Similarity and Analogy. *In Chemometrics: Theory and Application; Kowalski, B.; ACS*

Symposium Series; American Chemical Society: Washington, DC, 243–282. https://doi.org/10.1021/bk-1977-0052.ch012

Woo, Y. A., Kim, H. J., Ze, K. R., & Chung, H. (2005). Near-infrared (NIR) spectroscopy for the non-destructive and fast determination of geographical origin of Angelicae gigantis Radix. *Journal of Pharmaceutical and Biomedical Analysis*, *36*(5), 955–959. https://doi.org/10.1016/j.jpba.2004.08.037

Woodcock, T., Downey, G., & O'Donnell, C. P. (2009). Near infrared spectral fingerprinting for confirmation of claimed PDO provenance of honey. *Food Chemistry*, *114*(2), 742–746. https://doi.org/10.1016/j.foodchem.2008.10.034

Workman, J. . J., & Weyer, L. (2007). Practical Guide to Interpretive Near-Infrared Spectroscopy. In *Practical Guide to Interpretive Near-Infrared Spectroscopy*. https://doi.org/10.1201/9781420018318

Wyttenbach, A., Furrer, V., Schleppi, P., & Tobler, L. (1998). Rare earth elements in soil and in soil-grown plants. *Plant and Soil*, 267–273.

Xie, L., Ye, X., Liu, D., & Ying, Y. (2011). Prediction of titratable acidity , malic acid , and citric acid in bayberry fruit by near-infrared spectroscopy. *FRIN*, *44*(7), 2198–2204. https://doi.org/10.1016/j.foodres.2010.11.024

Yener, S., Romano, A., Cappellin, L., Märk, T. D., Del Pulgar, J. S., Gasperi, F., Navarinic, L., & Biasioli, F. (2014). PTR-ToF-MS characterisation of roasted coffees (C. arabica) from different geographic origins. *Journal of Mass Spectrometry*, *49*(9), 929–935. https://doi.org/10.1002/jms.3455

Zannella, C., Carucci, F., Aversano, R., Prohaska, T., Vingiani, S., Carputo, D., & Adamo, P. (2017). Genetic and geochemical signatures to prevent frauds and counterfeit of high-quality asparagus and pistachio. *Food Chemistry*, *237*, 545–552. https://doi.org/10.1016/j.foodchem.2017.05.158

Zhang, W., Chen, T., Tang, J., Sundararajan, B., & Zhou, Z. (2019). Tracing the production area of citrus fruits using aroma-active compounds and their quality evaluation modles. *Journal of the Science of Food and Agriculture*, *September*. https://doi.org/10.1002/jsfa.10026

Zhao, H., Guo, B., Wei, Y., & Zhang, B. (2013). Near infrared reflectance spectroscopy for determination of the geographical origin of wheat. *Food Chemistry*, *138*(2–3), 1902–1907.

https://doi.org/10.1016/j.foodchem.2012.11.037

Zhao, H., Yu, C., & Li, M. (2017). Effects of geographical origin, variety, season and their interactions on minerals in tea for traceability. *Journal of Food Composition and Analysis*, *63*(July), 15–20. https://doi.org/10.1016/j.jfca.2017.07.030

Zhao, H., Zhang, S., & Zhang, Z. (2017). Relationship between multi-element composition in tea leaves and in provenance soils for geographical traceability. *Food Control*, *76*, 82–87. https://doi.org/10.1016/j.foodcont.2017.01.006

OBJECTIVES

The focal point of this thesis was to study the food origin of high value agri-food products, specifically, the authentication and traceability of Limone di Sorrento PGI and Limone Costa d'Amalfi PGI. In this context, the main aim was to link the product to its territory of provenance using primary and secondary traceability indicators. Special attention was given to the relationship between fresh food products and the environmental (mainly pedoclimate) conditions of production areas, to define a unique character profile distinguishing the product origin using different analytical techniques.

The investigation was conducted on eight orchards' farms in a small territorial scale (regional scale). Three farms located in the Limone di Sorrento PGI area, three in the Limone Costa d'Amalfi PGI area and, two outside the PGI areas (specifically the experimental fields of the Agricultural Sciences Department of University of Naples Federico II, in Portici, Napoli, and the Regional experimental agricultural farm "Improsta", in Eboli, Salerno). were selected In addition, to compare the Campanian lemon fruits with lemons of other origin (large scale), other four groups of fruits were taken in Italian (Calabria region, Italy) and U.S. (Berkeley, California, U.S.) markets.

In **Chapter 1**, we explored the possibility to build a chemometric discriminant model to authenticate and track Sorrento lemons by multi-element fingerprinting. The applied chemometric approach (explorative analysis on mineral profiles of lemon juices, correlations between soil and plant, annual-based discriminant models, and discriminant model on two production years) showed the stability and reliability of the not-essential elements for discrimination of products provenance at a small geographical scale (regional scale).

In **Chapter 2**, the Sorrento lemon discriminant model at regional scale was extended to Amalfi lemon juices (small scale) and to lemons of foreign origin (large scale). In addition to multi-element fingerprinting, the 87Sr/86Sr isotope signature was used for discriminating the samples of different provenance. The results confirmed the not-essential elements as good discriminants for lemon origin also at a large geographical scale and reinforced the relation between the 87Sr/86Sr isotope ratio in lemon and the bioavailable fractions of cultivation soil.

In **Chapter 3** we explored the potential use of Near Infrared Spectroscopy (NIR) on intact lemons to describe the annual variation of juice quality, due to different pedoclimatic conditions, the relationship between NIR spectra and quality properties of the juices, the possible discrimination both different cultivar of lemon and their geographical provenances

In Chapter 4, we applied the Near Infrared Spectroscopy (NIR) on intact lemons and lemon juices to build different discrimination models according to geographical origins of samples

Chapter 1: Provenance discrimination of Sorrento lemon with Protected Geographical Indication (PGI) by multi-elemental fingerprinting

Food Chemistry 362 (2021): 130168

Luigi Ruggiero[1*], Maria Chiara Fontanella[2], Carmine Amalfitano[1*], Gian Maria Beone[2], Paola Adamo[1].

[1]Department of Agricultural Sciences, University of Naples Federico II, via Università 100, 80055 Portici (NA), Italy

[2]Department for Sustainable Process, Faculty of Agriculture, Food and Environmental Science, University Cattolica del Sacro Cuore of Piacenza, 29212 Piacenza, Italy

* Corresponding author.

Email addresses: **luigi.ruggiero@unina.it** (L. Ruggiero), **amalfita@unina.it** (C. Amalfitano).

Abstract

Multielement analysis and chemometric methods were proposed to discriminate the *Sorrento lemon* (PGI) juices according to geographical origin. In 2018 and 2019, 169 fruits from three farms in PGI area and two in not-PGI area were collected and analysed for essential and not-essential elements by ICP-MS. The PCA of multielement fingerprinting grouped lemon juices from PGI farms revealing a strong differentiation at small geographical scale. The S-LDA discriminated lemon juices for Mo, Ba, Rb, Mg, Co, Ca, Fe, Sr on the two production years, giving 97.7% correct classification, 98.5% accuracy and 93.8% external validation. The good correlation lemon juice *vs* cultivation soil and the soil discrimination by not-essential elements suggested the use of these elements as reliable indicators of lemon juices provenance. Despite lowering the number of variables, constituted by not-essential elements Ba, Rb, Ti, Co, the use of S-QDA discriminated the lemons juices with 87.5% accuracy and 83.9% validation.

Keywords: Limone di Sorrento PGI; Multielement traceability; Lemon juice; Soil; Multivariate analysis; not essential elements.

1 Introduction

In the last decades EU has seen a growth of trade in high value agri-foods and origin-based marketing strategies that have generated a strong need to protect the food with "Geographical Indication" labelling (GI). Protected Geographical Indication label (PGI), introduced by European Community Regulation (EC N° 510/2006; provided in Appendix A. Supplementary references), is a tool to protect resources and indigenous knowledge, increase the sustainability of production sites, and help the development of rural communities (Aprile et al., 2012). GIs identify food products that are the result of technical, social, and economic interactions of locally specific resources (Belletti et al., 2017). As reported in the EU database for registration of food products (DOOR, 2019; provided in Appendix A. Supplementary references), 748 PGIs have been registered so far. Thirty-four countries (27 EU and 7 Extra EU) have their products awarded as PGI. France (19.2%), Italy (17.3) and Spain (12.4) have been granted the highest number of PGI products. Because of their high economic value, PGI products are often subjected to fraud with products of different origins. This damages consumer trust and whole native agricultural sector and territory. Seven lemon fruits in the EU have been awarded PGI so far, out of which six are in Italy, the second-largest citrus fruit producer in Europe after Spain (Clever & Valverde, 2019). All six arise from the southern regions of Italy, i.e., Sicily, Calabria, Puglia, and Campania. The two PGI products from Campania are *Limone di Sorrento and Limone Costa D'Amalfi*. *Limone di Sorrento PGI* refers to the lemons of cultivar Femminello *Ovale di Sorrento*, produced in the Sorrento peninsula and the Capri Island. Only few studies report on lemon traceability, and in particular with the use of multi-element fingerprinting of lemon juice (Amenta et al., 2016; Gaiad et al., 2016; Pellerano et al., 2008; Potortì et al., 2018). Multielement fingerprinting is one of the most widely used technique to discriminate the geographical origin of food (Luykx & van Ruth, 2008). It takes advantage of the fact that mineral elements can be transferred from soil to agricultural products and thus any differences in elements distribution between different geographic regions are reflected in agro products. Nevertheless, the mineral composition of agri-food products depends not only on the mineral composition of cultivation soils but also on the environmental conditions (i.e.: climatic factors and agronomic management) influencing mineral element uptake by plants. Indeed, several papers report on multielement fingerprinting for estimation of the geographic origin of both animal-derived and plant-derived food (Anderson & Smith, 2005; Batista et al., 2012; Cheajesadagul et al., 2013; G. Di Bella et al., 2012, 2013, 2016; Franke et al., 2007; Gonzálvez et al., 2011; Kelly et al., 2002; Potortì et al., 2017, 2020, 2021; Simpkins et al., 2000; Zhao et al., 2011 provided in Appendix A. Supplementary references) and in all cases, the discriminant elements are suggested to be related to both soils and environmental conditions. Multielement data must be processed by chemometric tools to determine the geographical origin, among them unsupervised, principal component analysis and cluster analysis, and supervised, linear discriminant analysis and partial least squares discriminant analysis, methods can be used (Barbosa et al., 2014). The relationship of agro food multielement fingerprinting with cultivation soil can be investigated by the use of different chemometric tools, i.e.: *t*-test, and Pearson correlation analysis (Oddone et al., 2009; Pepi et al., 2017; Podio et al., 2013; Shen et al., 2013; Zhao et al., 2017b). The building of proper and reliable chemometric models can contribute towards development of information-sharing tools to track the food from farm to fork, such as blockchain (George et al., 2019). It can improve food authentication along food chain and consumer trust. The main aim of this work was to develop a discriminant model to authenticate and track the lemon juice of *Limone di Sorrento* PGI on a small geographical scale (regional scale) by multi-element fingerprinting to protect it from fraud. Out of five lemon orchards chosen for this study, three are in the PGI cultivation area of the Sorrento PGI lemon, with the remaining two at a short distance from the PGI area in the same region (Campania). The pseudo total and bioavailable contents of nineteen elements (Al, Ba, Cd, Co, Cr, Cu, Fe, K, Mg, Mn, Mo, Na, Ni, Pb, Rb, Sr, Ti, Zn, and Ca) in the cultivation soils were determined and their spatial variability was evaluated to explore relationships with lemon element fingerprinting. Multivariate statistics were used to explore the natural grouping and to select discriminating elements of lemon juices, while univariate statistic was applied to determine the significant difference between orchard soils. The results of multivariate analyses and of correlations between the multielement fingerprinting of lemon juices and cultivation soils were discussed with special emphasis on the discriminant ability of not-essential elements.

2 Materials and Methods

2.1. Study area and sampling

The Sorrento PGI lemon refers to the lemon ecotype Femminello *Ovale di Sorrento* (Citrus lemon L., Burms.), also known as *Limone di Massa Lubrense* or *Massese*. The PGI area is within the territory of the following municipalities of the Campania region: Vico Equense, Meta, Piano di Sorrento, Sant'Agnello, Sorrento, Massa Lubrense, Capri, and Anacapri. The production area extends around 139 hectares with 265 registered producers (ISTAT 2017; provided in Appendix A. Supplementary references). Due to the geomorphological conformation of the territory, the size of lemon orchards ranges between 0.5-1.5 hectares. The cultivation method follows the typical and traditional method of production and the terraces are very common. The trees are cultivated with the "free vase" system and covered with the "pergolato sorrentino", which is made of chestnut trellises and reed curtains known as "pagliarelle". The *Ovale di Sorrento* is not the only ecotype growing in the PGI area, however, 15% of other varieties are allowed in PGI orchards as reported in the production specification (GU 2011; provided in Appendix A. Supplementary references). The other most diffused Femminello cultivars, are *Zagara Bianca*, *Siracusano 2KR*, *Sfusato Amalfitano*, *Adamo* and *Cerza*. In the Sorrento peninsula, the annual rainfall recorded in the last ten years ranges between 692 and 1160 mm with the monthly maximum of 96 mm usually recorded in November, and the minimum of 19 mm in July-August. As generally for the Mediterranean coast and the geographical conformation of the territory, the amount of rainfall increases along the elevation gradient and toward more inland areas providing for different microclimates and habitats (Savo et al., 2016). As reported in Calcaterra & Santo, (2004) the Sorrento peninsula comprises of Mesozoic carbonates as limestones, dolomitic limestones, and dolostones, locally covered by discontinuous pyroclastic deposits with a thickness less than 1 m (Picarelli et al., 2008). In this setting, the coast is susceptible to landslides and consequently land degradation (Savo et al., 2016). In this geomorphological conformation, traditional agriculture, mostly on terraces, plays an important role and is a key element in maintaining soil stability and hydrogeological equilibrium (Lal, 2001). The peculiar soil and climatic features and the traditional terraced cultivations have contributed not only to produce high-quality lemon fruit but also to protect the landscape.Lemon fruits, belonging to the most diffuse genotypes grown in Sorrento peninsula, were collected between March and April 2018 and 2019, from the same selected trees in both years, in four lemon orchard farms located in the Campania region; in 2019 another orchard farm was selected in PGI area (Fig. 1). Each farm had homogeneous morphology with sparse cover of spontaneous herbaceous plants under trees. The genotype *Ovale di*

Sorrento was collected in all farms, while the other mentioned genotypes were collected based on availability in the farms. In each farm lemon trees and lemons (63 and 106 fruits in 2018 and 2019, respectively) were randomly selected to ensure the greatest variability among sampling areas. Only in some cases, due to the distribution of the different cultivars, lemon fruits were sampled from nearby trees. The three PGI farms were geopedologically in a quite homogeneous northern part of the Sorrento peninsula (ISPRA, 2015a; provided in Appendix A. Supplementary references) and all these farms were covered with the typical pergola culture of Limone di Sorrento and undergone the other requirements of production specifications (GU 2011; provided in Appendix A. Supplementary references). Farm locations, the total numbers and cultivars of collected lemons, in both production years are as follow:

- MS-PGI (40°36’28’’N; 14°20’36’’E) in the municipality of Massa Lubrense, cv. Femminello *Ovale di Sorrento*, n° 31 fruits;
- PZ-PGI (40°37’51’’N; 14°24’22’’E) in the municipality of Piano di Sorrento, cv. Femminello *Ovale di Sorrento* n°39 fruits, cv. Femminello *Zagara Bianca* n° 6 fruits, cv. Femminello *Siracusano 2KR* n° 6 fruits;
- SR-PGI (40°36’59’’N; 14°23’10’’E), in the municipality of Sorrento, production only in the 2019, cv. Femminello “Ovale di Sorrento” n°6 fruits, cv. Femminello *Zagara Bianca* n° 7 fruits, cv. Femminello *Siracusano 2KR* n° 7 fruits; cv. Femminello *Sfusato Amalfitano* n°7 fruits.

The whole extension of these orchard farms was 3.2 hectares which represented 2.3% of all PGI farm area. Their soil characteristics, considering the similarity of geological substrates (ISPRA, 2015a provided in Appendix A. Supplementary references), can be considered representative of most of the PGI orchard farms, excluding Capri island where, however lemon production is scarce. Although all lemon orchard farms ware located in the north side of Sorrento peninsula, a certain degree of microclimate inhomogeneity cannot be excluded. The other two farms, fields hosting lemon germplasm collections, were located outside the Sorrento PGI area. As for PGI farms, they had homogeneous geomorphology and sparse cover of spontaneous herbaceous plants under trees, and trees were not covered. Their specifications are:

- PR-NOPGI (40°48’54’’N; 14°20’49’’ E) in the municipality of Portici, Department of Agricultural Sciences, University of Naples Federico II; cv. Femminello *Ovale di Sorrento* n°12 fruits, cv. Femminello *Sfusato Amalfitano* n°12 fruits, cv. Femminello *Adamo* n° 6 fruits, and Femminello *Cerza* n° 6 fruits;

- EB-NOPGI (40°33’21’’ N; 14°58’53’’E) in the municipality of Eboli, Improsta Campania region experimental farm; cv. Femminello *Ovale di Sorrento* n°12 fruits, cv. Femminello *Sfusato Amalfitano* n°12 fruits.

In all orchard farms, around each selected tree or in some cases around two nearby trees, three soil samples (about 500g each) were collected at the depth of 0-20 cm and 40-50 cm (top-soil and sub-soil, respectively). The three soil samples were then mixed to obtain composite top-soil and sub-soil samples reaching a total of 14 composite soil samples for each soil depth.

2.2. Samples preparation

After sampling lemon fruits were washed with tap water and then with deionized water for two times. Lemon fruits were cut with disposable plastic knives and squeezed with a plastic domestic juicer. Juices were centrifuged at 2000 *g* for 20 min to separate from pulps and stored at -20°C before the analyses. Whole soil samples were air-dried to constant weight, sieved at 2 mm, homogenized, and quartered in four subsamples. About 50 g of a subsample for each soil were pulverized in agata ball mill (PM 200, Retsch Gmbh, Haan, Germany) for the determination of pseudo total elements.

2.3. Soil and fruit analysis

All the reagents used were analytical grade or better. Ultrapure water with resistivity >18 MΩcm (PURELAB, ELGA, Bucks, UK), 0.45 µm filtered (DigiFILTER, SCP Science, Champlain, NY, USA) was used for the analysis procedures with inductively coupled plasma spectrometers. The soil chemical and physical properties, i.e. pH (H_2O), Cation Exchange Capacity (C.E.C), total Carbonates (tot. Carb.), Organic Carbon (O.C.), soil texture (sand, silt, and clay), were determined according to the Italian official methods of soil analysis (MiPAF 2000 provided in Appendix A. Supplementary references). The soil pseudo total elements Al, Ba, Cd, Co, Cr, Cu, Fe, K, Mg, Mn, Mo, Na, Ni, Pb, Rb, Sr, Ti, Zn, and Ca were determined digesting pulverized soil (0.5 g) with aqua regia (7.5 mL of 37% HCl and 2.5 mL of 65% HNO_3) at 95°C for 2 hours in a block system (DIGIPREP, SCP Science, Champlain, NY, USA) equipped by polypropylene tubes (digiTUBES, SCP Science, Champlain, NY, USA). The soil bioavailable contents of the same elements were determined by EDTA 0.05 M extraction (Agrelli et al., 2017). The elements in the mineralized and EDTA solutions were determined by inductively coupled plasma-optical emission spectrometer (ICP-OES 5100 Agilent Technologies, Santa Clara, CA, USA) whose operating conditions are listed in Table S1 (provided in Appendix A. Supplementary materials). All the calibration solutions were prepared in the same acid matrix used for soil samples and measurements were in triplicate. Analytical quality control was periodically carried out in triplicate with certified reference materials: BCR 141R (calcareous loam soil) for pseudo total element following ISO 11466:1995, and BCR 700 (organic-rich soil) for bioavailable element (Rauret et al. 2001). For lemons, 5 g of juice were mineralized in a Teflon tube with 4 mL of ultrapure HNO_3 65% and 2 mL of H_2O_2 30% using a microwave system (Mars 5 Express, CEM) following power ramp to 1600W for 5-5-10 min at 100-150-200 °C, respectively. After cooling, the mixtures were completed to 50 mL final volume with ultrapure water and elements were determined by inductively coupled plasma-mass spectrometer (ICP-MS, 7800 Agilent Technologies, Santa Clara, CA, USA) whose operating conditions are listed in Table S2 (provided in Appendix A. Supplementary materials). A recovery test using a spike solution was performed for lemon juices. Blank samples were included in each procedure, and MDL (Method Detection Limit) for soil and lemon juice treatments was calculated as three times the standard deviation of six analytical blanks prepared and diluted in the same way of samples (Spalla et al., 2009).Triplicate analysis were performed on a randomly chosen 10% daily batch of samples and the average of their relative standard deviations was attributed to all samples of that batch. Plant essential elements (Fe, Mn, Cu, Zn, Mo, Ca, K, Mg, including Na, as ubiquitous beneficial element in high concentration in plants) and not-essential elements (Ti, Co, Rb, Sr and Ba) were chosen on the basis of detectability and reproducibility of the determinations.

2.4. Statistical analysis

The statistical analysis was performed using SPSS 23.0 statistical software (SPSS Inc., Chicago, IL) and XLSTAT version 2016.02.284551 (Addinsoft, NY, USA). The normal distribution of each variable was checked by the Shapiro-Wilk test ($p<0.05$). Univariate statistic by Kruskal–Wallis test with a post-hoc Dunn test, analysis of variance by ranks (Dunn, 1964), was performed on soil samples ($p<0.05$). Multivariate statistics were carried out on lemon juices. Before these, the dataset was variance scaled, mean-centred and normalized with log transformation. Principal Component Analysis (PCA) was carried out for exploring the natural grouping of the dataset according to geographical origins and to different cultivars. Before PCA, Kaiser-Meyer-Olkin sampling adequacy (KMO) and Bartlett’s sphericity tests were performed to evaluate the data matrix suitability (Pellerano et al., 2008). The discriminant supervised methods, Stepwise Linear Discriminant Analysis (S-LDA) and Stepwise Quadratic Discriminant Analysis (S-QDA), were carried out. LDA creates a linear combination of extracted functions to obtain maximal separation among groups and compactness inside groups (Pellerano et al., 2008), where the variance/covariance matrices for different groups are homogeneous. In QDA the different groups are hyperquadrics (multidimensional space) and the class separation based, like for LDA, on Wilks' lambda. Wilks' lambda values close to zero indicate that models can perfectly discriminate (Zhao et al., 2017a). The stepwise procedure selects only the best discriminant variables based on Wilks' lambda. The main differences between LDA and QDA are that the first works on a pooled variance-covariance matrix, the second on a variance-covariance inverse matrix for each group. Thus, for LDA the ratio between total sample number and their associated variables is suggested to be not lower than three, better if it overcomes five, for QDA that ratio has to be considered for each group, therefore QDA is more exigent and reliable than LDA (Lees, 2003; Marini, 2013). The goodness of discriminant models was evaluated by the percentage of correct classification and accuracy, by leave-one-out cross-validation and by external validation (Lees, 2003). To investigate the linear relationships between element contents in soil and lemon juices the Pearson Correlation analysis was applied. Since the number of lemon samples was higher than that of the corresponding sampled soils, the correlation analysis was carried out associating soils with corresponding lemon juices, randomly chosen, and reiterating this process five times in manner to consider all lemon juice samples at least one time. Each correlation coefficient was calculated as the mean of five coefficients, and it was evaluated for significance ($p<0.05$). The frequencies of significant coefficients from the five iterations were calculated as percentage.

3. Results and Discussion

3.1. Variability of mineral element composition of lemon juices

The mean element contents of lemon juices according to provenance are shown in Table 1 and the validation parameters of element determinations are reported in Table S3 (provided in Appendix A. Supplementary references). Similar element profiles were observed in all lemon juices and element concentrations were in the following order: K>> Mg> Ca > Na > Rb >> Zn > Fe > Cu > Sr > Mn > Ba >> Mo > Ti > Co. The concentrations of Fe, Zn, Mn, Cu and Sr were in the same order of magnitude found in centrifuged cv. Femminello lemon juices from Sicily (Amenta et al., 2016) but all elements were slightly lower than those in no-centrifuged lemon juices and lemon pulps from Argentina and Sicily (Gaiad et al., 2016; Pellerano et al., 2008, and Potortì et al., 2018). An exploratory analysis by PCA on the 169 sampled lemon fruits was carried out for screening the natural grouping of samples. KMO test, giving a result of 0.718, and Bartlett sphericity test, showing p-value < 0.0001, indicated the adequacy of samples to apply the PCA (Li et al., 2018). The PCA showed that the first five extract factors (eigenvalues >1), PCs, explained 70.4% of the total variance, and the first three explained 27.6%, 17.1%, and 10.6%, giving a cumulate variance of 55.3%. The score plot of PCA showed that all cultivars from PGI area (MS-PGI, PZ-PGI, and SR-PGI) were closer to each other in the centre of the plot respect those from NO-PGI sites that arranged at the margins (Fig.2a), consequently the groupings for each cultivar of Femminello present in different areas were not evident (Fig. 2b). Similar results were obtained by Pellerano et al. (2008) by discriminant analysis, which showed the overlapping of different botanical varieties grouped in three different production Argentinian regions. This suggests that juice mineral composition did not depend strongly by cultivars, while the natural grouping of juices by provenance suggests that mainly the features of cultivation environments of Sorrento peninsula influenced the elemental profile of juices. Considering the first three PCs, the highest loading weight (elements) were on PC1 for Ca (15.3%), Cu (11.3%), Mg (11.2%), Rb (9.5%), Sr (8.3%), Co (7.9%) and Na (6.1%); on PC2 for Mo (20.4%), Ba (20.0%) and Ti (15.0%), and on PC3 for Fe (30.7%), Zn (14.8%), K (14.2%) and Mn (8.3%). It is worth noting that, among these elements, Ba, Na, K, Rb, Sr and Mg were also those discriminating cultivation soils (see paragraph 3.2). The results of the exploratory analysis suggested a possibility to build a supervised discriminant model dividing the data set in three groups according to provenance, i.e.: MS- PZ- SR-PGI (PGI area), PR-NOPGI and EB-NOPGI.

3.2. Variability of the main properties and mineral element composition of cultivation soils

The main physical and chemical properties, as mean values, for all soils sampled in each orchard, along with their standard deviations and the results of nonparametric Kruskal–Wallis test ($p>0.05$) between top-soil (0-20 cm) and sub-soil (40-50 cm) and among orchard farms, are reported in Table S4 (provided in Appendix A. Supplementary materials). In general, except for total carbonates, in each orchard soil layer, soil property variations, as standard deviations on mean values in percent, were not higher than 51%, (recorded for O.C. in MS-PGI top-soil), and the mean of property variations in each orchard soil layer was not higher than 21% (recorded in top-soil of SR-PGI). Considering all orchards soil layers overall the mean of propriety variations resulted higher for top-soil (17%) than sub-soil (12%) indicating a higher homogeneity of the properties in sub-soils, i.e., the soil portion more explored by lemon tree roots (Castle and Krezdorn, 1975). The variation of carbonates was wider, between around 100% (for SR-PGI on average in top-soil and sub-soil), and 7% (for PR-NOPGI on average in top-soil and sub-soil), however the mean of total carbonate variations in all orchards was no higher than 45% on overage in top and sub soil. These results suggest that the sampled lemon trees were experiencing quite similar soil agronomic properties during their growth. In each orchard farm, in most of the cases, O.C as well as C.E.C. were higher in top-soil respect to sub-soil while pH showed an opposite trend in agreement with the typical shallow minimum tillage and mowing of orchard soils preserving soil organic matter. Total carbonates varied differently, being higher in top-soil respect to sub-soil in MS-PGI and SR-PGI, showing opposite trend in PR-NOPGI and no significant differences in the other orchards. Regarding particle-size distribution, clay content was higher in sub-soils than in top-soils in PR-NOPGI and EB-NOPGI, the opposite was in PZ-PGI, and no differences were found in MS-PGI and SR-PGI. The other particle-size fractions did not show differences with soil depth in all orchard farms suggesting a certain homogeneity of physical characteristics between the two soil layers in all orchard farms. In the orchards O.C. was generally sufficient and showed small differences (between 11 g kg^{-1} and 19 g kg^{-1} in top-soils and 5 g kg^{-1} and 11 g kg^{-1} in sub soil). The top-soils had similar slight alkaline pH (on average 7.7), except in MS-PGI which showed a lower value (7.1). PGI sub-soils had similar slight alkaline pH (on average 7.6), whereas NOPGI sub-soils showed more alkaline pH (7.9 for EB-NOPGI, 8.1 for PR-NOPGI). In top-soils C.E.C. was similar in all PGI orchards farms (on average 23.5 cmol(+) kg^{-1}), while the lowest value, although similar to MS-PGI, was in PR-NOPGI (12.5 cmol(+) kg^{-1}), the highest, similar to PZ-PGI and SR-PGI, in EB-NOPGI (27.3 cmol(+) kg^{-1}). In the sub-soils of all orchard farms generally the same C.E.C. trend of top-soil was observed, except the significant difference between MS-PGI, and SR-PGI, however the mean C.E.C of PGI sub-soil (21.1 cmol(+) kg^{-1}) was lower than in top-soil. Top-soils and sub-soils of all orchard farms were generally poor in total carbonates (<25 g kg^{-1}) and their content was not ever similar in both soil layers. The highest value was in SR-PGI top-soil (31.2 g kg^{-1}). However, NOPGI orchard farms generally showed higher total carbonates than PGI farms in both layers. According to USDA (Soil Survey Laboratory Staff, 2004, provided in Appendix A. Supplementary references), top-soil and sub-soils had sandy loam texture in orchards of PGI area, sandy in PR-NOPGI and clay loam in EB-NOPGI. The recorded differences in soil agronomic properties, suggested that the availability of most mineral elements could be different among orchard farms, although all soils did not evidence problems for plant nutrition. In soils pseudo total elements (*PTE*) and bioavailable elements (*BAE*) were determined with relative standard deviations lower than 3% and 1%, as checked with BCR 141R and BCR 700 certified reference materials, respectively (Tables S5 and S6, provided in Appendix A. Supplementary references). For most of elements, there were no significant differences between the contents of *PTE* in top-soil and sub-soil of all sites, with some exceptions as follows (Table S7, provided in Appendix A. Supplementary materials). In top-soil, Ca was higher in MS-PGI and PZ-PGI, and Cu in MS-PGI than in the others. The contents of Ti, Co, and Mg were lower in PZ-PGI, and Cd and Na in SR-PGI and PR-NOPGI, respectively. About bioavailable elements (*BAE*), generally, a considerable number of *BAE* was more bioavailable in top-soils than in sub-soils in all farms. The differences in *PTE* and *BAE* contents among orchards for both top-soil and sub-soil did not follow a similar trend for all elements and were generally diffused and not wide. Among PGI orchard farms the similarities regarded 37% and 42% of *PTE*-top-soil and *PTE*-sub-soil, respectively; and 63% and 53% of *BAE*-top-soil and *BAE*-sub-soil. Appling the univariate statistic and grouping MS, PZ, SR element profiles as PGI area, the comparison of the multielement profile of PGI area with PR-NOPGI (Table 2) showed that the percentage for *PTE* with similar contents was 37% in top-soil and 53% in sub-soil, and for *BAE* 42% in top-soil and 74% in sub-soil, while with EB–PGI, for *PTE* 26% in top-soil and 26% in sub-soil, and for *BAE* 26% in top-soil and 37% in sub-soil. Furthermore, the similarity between NO-PGI orchard farms was for *PTE* 16% top-soil and 26% sub-soil, and for *BAE* 53% top-soil and 58% sub-soil. These results indicate that the similarities between orchard sites were generally higher for *BAE* than for *PTE* contents. Further, the similarity between PGI area and PR-NOPGI was quite high in some cases exceeding that found among orchard farms in PGI area. This finding is likely to derive from the fact that the Sorrento lemon PGI area and the site of PR-NOPGI orchard are not so distant and that both were affected in the past by deposition of volcanic materials from eruptions of Mount Vesuvius. Indeed, it is well known that volcanic materials deriving from old eruptions of Mt Vesuvius are widely spread in fertile soils of Campania region well known by ancient Roman time as *Campania felix*. In contrast, EB-NOPGI soil was more different, especially for *PTE* contents, respect to the soil of the other areas because in that site the soil is mainly derived by alluvial sediments (ISPRA, 2015a, b, and c; provided in Appendix A. Supplementary references). All this implied a low number of elements discriminating soils, especially between PGI area and PR-NOPGI, that for the three areas they were the *PTE* Ba, K, Na, Rb, and Sr in the top and sub-soil, and the *BAE* Ba, Cd and Mg in top-soil, and only Mg in sub-soil.

3.3. Relationship between soil and lemon mineral element contents

The relationships between the contents of 14 mineral elements (essential elements Fe, Mn, Cu, Zn, Mo, Ca, K, Mg, including Na, and not-essential elements Ti, Co, Rb, Sr and Ba) measured both in lemon juices and in the related top-soil and sub-soils (*PTE* and *BAE*) were investigated by Pearson correlation analysis. The elements showing significant ($p < 0.05$) mean correlation coefficients (on five iterated Pearson analysis) are given in Table 3 together with the frequencies (number of significant coefficients *per* number of iterations %) as an indication of correlation reliability. Only half of the 14 elements showed a significant correlation, namely Mo, Mn, Ba, Rb, Mg, Co, and Sr. Most macro and micronutrients, such as K, Ca, Na, Cu, Zn and Fe, did not significantly correlated or when correlating (Mg, Mn and Mo) showed low correlation or frequency values. This result could depend by a high variability of lemon nutritional status due to differences in environment and plant growing characteristics in orchards influencing specific nutrient uptake. The lowest number of correlating elements was found in *BAE*-top-soil dataset (Mn, Rb, Mg, Co). The same elements plus

Sr were even better correlated in *BAE*-sub-soil. These results might be due to differences in the nutrient uptake of lemon trees from upper and deep soil layers. Changes in soil physical, chemical, or biological characteristics could affect not only the citrus distribution of roots, but also their growth and activity (Castle and Krezdorn, 1975). Deeper roots are generally more efficient in water uptake, but no clear relationship between soil depth and nutrient uptake has been shown (Castle, 1978). Movement of mobile elements from upper to deeper soil layers might also affect root uptake in different soil layers. The significant coefficients of correlations found in *PTE* datasets for Mn, Rb, Sr, Ba and Mo were quite similar for top-soil and sub-soil; this could depend on the diffused similarity of element contents in both layers (section 3.1) rather than on uptake from both soil layers. The correlations of Ba and Mo, two elements exhibiting low soil mobility (De Andrade et al., 2014; Williams & Thornton, 1973), suggested a contribution to lemon nutrition also of elements apparently not potentially available as assessed by EDTA extraction. In any case the correlation and frequency values with *PTE* were generally slightly lower than those with *BAE*. In *PTE*, Mn and Mo positively correlated, while Sr, Ba and Rb negatively correlated. Similar relationship between multi-element composition in tea leaves with provenance soils were shown by Zhao et al. (2017b). The negative correlations might suggest root ability to uptake elements not readily available from particular soil mineral component.

3.4. Chemometric model for discriminating the provenance of lemon juices

In order to discriminate the provenance of lemon juices, a stepwise linear discriminant analysis (S-LDA, after log transformation) was applied. The 169 lemon juices samples were divided into two sets: the calibration set (CS) formed by 88 samples (52%), and the validation set (VS) formed by 81 samples (47%) randomly chosen for no more than 50% of each cultivar in each orchard. The CS samples were represented as follows: PGI area (43%), PR-NOPGI (32%), and EB-NOPGI (25%). The chemometric model obtained by S-LDA showed a λWilks value proximal to zero (0.038) suggesting that group means were different (Pellerano et al., 2008). The S-LDA (probability of F to remove 0.10 and to enter 0.05) selected eight elements, Mo, Ba, Rb, Mg, Co, Ca, Fe and Sr, as discriminant variables on the two discriminant functions, for FD1 (76.7%): 0.870 Ca, -0.782 Sr, 0.710 Ba, 0.684 Rb, -0.542 Fe, 0.487 Co, -0.451 Mo, -0.323 Mg, and for FD2 (23.2%): -1.160 Sr, 0.761 Mg, 0.409 Mo, 0.366 Co, 0.297 Ca, 0.100 Ba, 0.083 Fe, 0.027 Rb. The sample scores of the two discriminant functions are shown in Fig. 3. The score plot explained a clear discrimination of EB-NOPGI samples from those of the other two groups, while there was a slight overlapping between PR-NOPGI and PGI area samples. The results of S-LDA showed 97.7% correctly classified samples (100% for PGI area, 92.8% for PR-NOPGI, and 100% for EB-NOPGI) and 98.4% accuracy. The model was leave-one-out cross-validate obtaining 95.4% correctly validated samples (97.3% for PGI area, 89.3% for PR-NOPGI, and 100% for EB-NOPGI). Further, the external validation of the model with VS samples showed 93.8% correct validation (95.71% for MS-PZ-SR-PGI, 77.7% for PR-NOPGI, and 100% for EB-NOPGI). Both validations agreed with the slight overlapping between lemon juices from PGI area and those from PR-NOPGI, and with the clear discrimination of lemon juices from EB-NOPGI. However, the S-LDA application on all samples of each sampling year, 2018 and 2019 (63 and 106 for 2018 and 2019 year, respectively), although with less variability due to reduced and less homogeneous sampling, and lack of validation sets, gave λWilks values 0.012 and 0.080, respectively, correct classifications 100% and 93.4% and leave-one-out cross-validations 98.4% and 91.5%, quite similar to those seen with the samples of entire biennium. However, the lists of discriminating elements were different: for the 1st year they were Cu, Ba, Rb, Mg, Co, Ca, Na, Fe and Sr, while for the 2nd year they were Mo, Mn, Ba, Rb, Ca, Ti and Sr; thus, only Ca, Ba, Rb and Sr were discriminant elements in both years. The difference between the discriminant profiles could depend by the annual variations of environmental conditions (such as climate, pathogen and insect attacks, water availability, etc.) and cultivation practices in orchards influencing element uptake. It is worth noting that only one (Ca) of the nine determined essential elements discriminates and coincides in both years (11%), while well three (Ba, Sr and Rb) out the five not-essential elements do this (60%). It is likely that in cultivated soils the uptake of Ba, Sr and Rb, being generally in low amount in soil and consequently also low available (Kabata-Pendias, 2011), is more linear with their availability than competitive essential elements, e.g. Ca for Ba and Sr and K for Rb (El-Sheikh & Ulrich, 1970; Jackman, 1965; Rediske & Selders, 1953). In juices the cations, including not-essential elements, can be mainly related to the physiological role for osmotic regulation and salification rather than for biochemical processes such as in plant tissues. Thus, it is likely that essential and not-essential element composition of juices changes with lemon ripening stage. However, in our study, the processed lemons were all at the same ripening stage as proved by homogeneous pH, treatable acidity and soluble solid content of juices (data not shown). Therefore, the element composition of juices, especially the not-essential element composition, should be affected mainly by the possibility for these elements to be intercepted and uptaken by the roots of the plant, also, during maturation some translocation of not-essential elements to the fruits from other plant organs could be negligible respect to essential elements. Among the elements discriminating both soil and lemon juices, only four, Ba, Rb, Sr, and Mg, showed soil vs juices correlation but Mg did not discriminate lemon juice in the 2nd year while the not-essential elements Ba, Rb and Sr discriminated for each year. Among the other elements discriminating only for lemon juices, Ca and Fe did not show correlation, and Mo and Co did not discriminate soils. This suggested that essential elements, including Mg, ware less consistent than not-essential elements in discriminating geographical provenience of Sorrento PGI lemon juices. As suggested before, this might be explained assuming that the content of essential elements in juices depends more on annual environment variations and orchard managements compared to not-essential elements. Thus, the use of not-essential elements could be more consistent than essential ones to trace the geographical provenance of lemon juices. It is worth noting that reducing the number of variables to the five not-essential elements (Ba, Rb, Sr, Ti and Co), the ratio samples/variables for each group overcomes the limit of three suggested by Lees (2003) and Marini (2013), and the stepwise QDA, more restrictive than S-LDA, could be applied. The S-QDA showed a λWilks value of 0.173 ($p<0.0001$), 87.5% correct classification and 83.9% external validation. The percentage of correct classification was lower than that obtained by S-LDA on total variables (97.7%). However, the ratio between the percentages of external validation and classification for both S-LDA and S-QDA were equal to 0.96, this would indicate that the reliability of the two models was similar. The not-essential elements, Ba, Rb and Sr, already found by S-LDA as discriminating elements for lemon juice geographical provenience, were confirmed also by S-QDA with the addition of the other not essential elements, Ti and Co, which affected the model more than Sr. It is worth noting that Ti and Co did not discriminate cultivation soils by univariate statistics (see paragraph 3.2) and that Ti discriminated lemon juices only in the 2nd year and Co only in the 1st year by S-LDA.

4. Conclusions

The results of this paper showed that multielement fingerprinting of lemon juices from Sorrento peninsula PGI area was an useful tool to discriminate lemon provenance on a small geographical scale (region scale), but it was not able to discriminate the PGI Femminello *Ovale di Sorrento* from the other Femminello cultivars (*Zagara Bianca*, *Siracusano 2KR*, *Sfusato Amalfitano*, *Adamo* and *Cerza*) growing in the same area. This indicates a prevalent influence of growing environment and of soil mineral element composition and availability on the element composition of lemon fruits. The correlations found between in lemon juices and in related soil (pseudo total and available) reinforce this conclusion. Eight elements (namely Mo, Ba, Rb, Mg, Co, Ca, Fe and Sr) among the 14 determined resulted to discriminate (by S-LDA) lemon juices of Sorrento peninsula PGI area from those of two NO-PGI areas, Portici and Eboli, located in the same region at no more than 60 km away from PGI area. Lemon juices were identified by S-LDA with 97.7% correct classification, 98.4% accuracy and 93.8% external validation. However, a certain degree of annual variation in the content of essential elements discriminating lemon juices and the good correlation between the content of not-essential elements in lemon juices and in corresponding soils suggested the use of only not-essential elements (Ti, Co, Rb, Ba, and Sr) as stable fingerprinting to discriminate the geographical provenance of lemon juices. With only their use, despite the reduction of the number of variables, PGI lemon juices were identified by S-QDA with 87.5% correct classification and 83.9% validation. The use of a lower number of variables allows applying QDA to predict the origin of all samples without increasing the number of samples; this could be suitable for routine approach in quality protection and fraud repression. The traceability based on not-essential elements could be even more effective on large than on small geographical scale, considering that cultivated soils might differ more for content and availability of not-essential elements than for essential elements often supplied by fertilizations. While this pilot study was successful in achieving its aims, there are still limitations with respect to the coverage of natural pedoclimatic variability of the *Limone di Sorrento* PGI area. Additional research should be conducted to include a greater number of lemon farms (or sampling points) in the PGI area and to enlarge the existing database including lemon samples from other regions. In this way, the model for authenticating and protecting the *Limone di Sorrento* PGI lemon could substantially contribute to implement a blockchain information system for Sorrento lemon traceability, providing real-time information not only to the final consumers but also to manufacturers, distributors, and retailers.

5. Acknowledgments

The authors would like to thank Dr. Paolo Lodigiani (University Cattolica del Sacro Cuore, Piacenza, Italy) for analytical work support, Prof. Claudio Di Vaio (University of Naples Federico II) for providing PR-NOPGI and EB-NOPGI soil and lemon samples, Dr. Andrea Chierchia and Cooperativa Solagri for helping in soil and lemon sampling in the Sorrento peninsula PGI area and Dr. Paolo Nasta (University of Naples Federico II) for providing climatic data.The present study is part of the PhD project of Luigi Ruggiero financially supported by University of Naples Federico II.

6. References

Agrelli, D., Adamo, P., Cirillo, T., Duri, L. G., Duro, I., Fasano, E., … Fagnano, M. (2017). Soil versus plant as indicators of agroecosystem pollution by potentially toxic elements. Journal of Plant Nutrition and Soil Science, 180(6), 705–719. **https://doi.org/10.1002/jpln.201700050**

Amenta, M., S., Fabroni, Costa, C., Rapisarda, P., Fabroni, S., Costa, C., & Rapisarda, P. (2016). Traceability of "Limone di Siracusa PGI" by a multidisciplinary analytical and chemometric approach. Food Chemistry, 211, 734–740. **https://doi.org/10.1016/j.foodchem.2016.05.119**

Aprile, M. C., Caputo, V., & Nayga, R. M. (2012). Consumers' valuation of food quality labels: The case of the European geographic indication and organic farming labels. International Journal of Consumer Studies, 36(2), 158–165. **https://doi.org/10.1111/j.1470-6431.2011.01092.x**

Barbosa, R. M., Batista, B. L., Varrique, R. M., Coelho, V. A., Campiglia, A. D., & Barbosa, F. (2014). The use of advanced chemometric techniques and trace element levels for controlling the authenticity of organic coffee. Food Research International, 61, 246–251. **https://doi.org/10.1016/j.foodres.2013.07.060**

Belletti, G., Marescotti, A., & Touzard, J. M. (2017). Geographical Indications, Public Goods, and Sustainable Development: The Roles of Actors' Strategies and Public Policies. World Development, 98, 45–57. **https://doi.org/10.1016/j.worlddev.2015.05.004**

Calcaterra, D., & Santo, A. (2004). The January 10, 1997 Pozzano landslide, Sorrento Peninsula, Italy. Engineering Geology, 75(2), 181–200. https://doi.org/10.1016/j.enggeo.2004.05.009

Castle, W. S. (1978). Citrus root systems: their structure, function, growth, and relationship to tree performance. Proceedings of the International Society of Citriculture 1978., (JANUARY 1978), 62–69.

Castle W. S. and A. H. Krezdorn. 1975. Effect of citrus rootstocks distribution and leaf mineral content of 'Orlando' tangelo trees . .J. A mer. Soc. Hart. Sci. 100:1-4.

Clever, J., & Valverde, C. (2019). EU-28 Citrus Semi-annual Rebound in EU Production Expected to Lower EU Citrus Imports. Global Agricultural Information Network, SP1929(June).

De Andrade, A. A. F. M., Do Amaral Sobrinho, N. M. B., Dos Santos, F. S., Magalhães, M. O. L., Tolón-Becerra, A., & Da Silva Lima, L. (2014). EDTA-induced phytoextraction of lead and barium by brachiaria (B. decumbens cv. Basilisk) in soil contaminated by oil exploration drilling waste. Acta Scientiarum - Agronomy, 36(4), 495–500. **https://doi.org/10.4025/actasciagron.v36i4.18172**

Dunn, O. J. (1964). Multiple Comparisons Using Rank Sums. Technometrics, 6(3), 241–252

El-Sheikh, A. M., & Ulrich, A. (1970). Interactions of Rubidium, Sodium, and Potassium on the Nutrition of Sugar Beet Plants. Plant Physiology, 46, 645–649. **https://doi.org/10.1104/pp.46.5.645**

Gaiad, J. E., Hidalgo, M. J., Villafañe, R. N., Marchevsky, E. J., & Pellerano, R. G. (2016). Tracing the geographical origin of Argentinean lemon juices based on trace element profiles using advanced chemometric techniques. Microchemical Journal, 129, 243–248. **https://doi.org/10.1016/j.microc.2016.07.002**

George, R. V., Harsh, H. O., Ray, P., & Babu, A. K. (2019). Food quality traceability prototype for restaurants using blockchain and food quality data index. Journal of Cleaner Production, 240, 118021. **https://doi.org/10.1016/j.jclepro.2019.118021**

ISO 11466, 1995. 1995, Soil quality – Extraction of trace elements soluble in aqua regia.

Jackman, R. H. (1965). The Uptake of rubidium by the roots of some graminaceous and leguminous plants T. New Zealand Journal of Agricultural Research ISSN:, 8(4), 763–777.

Kabata-Pendias, A. (2011). Trace elements in soils and plants. In CRC Press. **https://doi.org/10.1201/b10158-25**

Lal, R. (2001). Managing world soils for food security and environmental quality. Advances in Agronomy, 74(1), 155–192. **https://doi.org/10.1016/s0065-2113(01)74033-3**

Lees, M. (2003). Food authenticity and traceability. In LWT - Food Science and Technology (Vol. 38). Woodhead Publishing Limited, 2003 **https://doi.org/10.1016/j.lwt.2004.04.009**

Li, L., Wen, B., Zhang, X., Zhao, Y., Duan, Y., Song, X., … Zhu, X. (2018). Geographical origin traceability of tea based on multi-element spatial distribution and the relationship with soil in district scale. Food Control, 90, 18–28. **https://doi.org/10.1016/j.foodcont.2018.02.031**

Luykx, D. M. A. M., & van Ruth, S. M. (2008). An overview of analytical methods for determining the geographical origin of food products. Food Chemistry, 107(2), 897–911. **https://doi.org/10.1016/j.foodchem.2007.09.038**

Marini, F. (2013). Data handling in science and technology: Chemometrics in Food Chemistry.

Oddone, M., Aceto, M., Baldizzone, M., Musso, D., & Osella, D. (2009). Authentication and traceability study of hazelnuts from piedmont, Italy. Journal of Agricultural and Food Chemistry, 57(9), 3404–3408. **https://doi.org/10.1021/jf900312p**

Pellerano, R. G., Mazza, S. S., Marigliano, R. A., & Marchevsky, E. J. (2008). Multielement analysis of argentinean lemon juices by instrumental neutronic activation analysis and their classification according to geographical origin. Journal of Agricultural and Food Chemistry, 56(13), 5222–5225. **https://doi.org/10.1021/jf073555n**

Pepi, S., Sansone, L., Chicca, M., & Vaccaro, C. (2017). Relationship among geochemical elements in soil and grapes as terroir fingerprintings in Vitis vinifera L. cv. “Glera.” Chemie Der Erde, 77(1), 121–130. **https://doi.org/10.1016/j.chemer.2017.01.003**

Picarelli, L., Santo, A., Crescenzo, G., & Olivares, L. (2008). Macro-zoning of areas susceptible to flowslide in pyroclastic soils in the Campania region. Landslides and Engineered Slopes. From the Past to the Future, (June), 1951–1957. **https://doi.org/10.1201/9780203885284-c271**

Podio, N. S., Baroni, M. V., Badini, R. G., Inga, M., Ostera, H. A., Cagnoni, M., … Wunderlin, D. A. (2013). Elemental and isotopic fingerprint of argentinean wheat. Matching soil, water, and crop composition to differentiate provenance. Journal of Agricultural and Food Chemistry, 61(16), 3763–3773. **https://doi.org/10.1021/jf305258r**

Potortì, A. G., Di Bella, G., Mottese, A. F., Bua, G. D., Fede, M. R., Sabatino, G., … Lo Turco, V. (2018). Traceability of Protected Geographical Indication (PGI) Interdonato lemon pulps by chemometric analysis of the mineral composition. Journal of Food Composition and Analysis, 69(March), 122–128. **https://doi.org/10.1016/j.jfca.2018.03.001**

Rauret, G., López-Sánchez, J. F., Bacon, J., Gómez, A., Muntau, H., & Quevauviller, P. (2001). Certification of the contents (mass fractions) of Cd, Cr, Cu, Ni, Pb and Zn in an organic-rich soil following harmonised EDTA and acetic acid extraction procedures, BCR-700. In Report EUR (Vol. 19774).

Rediske, J. . H., & Selders, A. . A. (1953). The Absorption and Translocation of Strontium by Plants. Plant Physiology, 28(4), 594–605.

Savo, V., Salvati, L., & Caneva, G. (2016). In-between soil erosion and sustainable land management: climate aridity and vegetation in a traditional agro-forest system (Costiera Amalfitana, southern Italy). International Journal of Sustainable Development and World Ecology, 23(5), 423–432. **https://doi.org/10.1080/13504509.2015.1132282**

Shen, S., Xia, L., Xiong, N., Liu, Z., & Sun, H. (2013). Determination of the geographic origin of rice by element fingerprints and correlation analyses with the soil of origin. Analytical Methods, 5(21), 6177–6185. **https://doi.org/10.1039/c3ay40700d**

Spalla, S., Baffi, C., Barbante, C., Turretta, C., Cozzi, G., Beone, G. M., & Bettinelli, M. (2009). Determination of rare earth elements in tomato plants by inductively coupled plasma mass spectrometry techniques. Rapid Communications in Mass Spectrometry, 23(20), 3285–3292. **https://doi.org/10.1002/rcm.4244**

Williams, C., & Thornton, I. (1973). The use of soil extractants to estimate pant-available molybdenum and selenium in potentially toxic soils. Plant and Soil, (39), 149–159.

Zhao, H., Yu, C., & Li, M. (2017a). Effects of geographical origin, variety, season and their interactions on minerals in tea for traceability. Journal of Food Composition and Analysis, 63(July), 15–20. **https://doi.org/10.1016/j.jfca.2017.07.030**

Zhao, H., Zhang, S., & Zhang, Z. (2017b). Relationship between multi-element composition in tea leaves and in provenance soils for geographical traceability. Food Control, 76, 82–87. https://doi.org/10.1016/j.foodcont.2017.01.006

Table 1. Mineral composition (Mean ± standard deviation, SD) of lemon juices according to provenance; MS-PGI, PZ-PGI and SR-PGI from the "Sorrento PGI area"; PR-NOPGI from the germplasm field, Portici; EB-NOPGI from germplasm field, Eboli. n: total number of samples

Elements	**PGI area**						**NO-PGI area**		**NO-PGI area**	
	MS-PGI (n 31)		**PZ-PGI (n 51)**		**SR-PGI (n 26)**		**PR-NOPGI (n 37)**		**EB-NOPGI (n 24)**	
mg kg^{-1}	**Mean**	**±SD**	**Mean**	**±SD**	**Mean**	**±SD**	**Mean**	**±SD**	**Mean**	**±SD**
K	1393	161	1336	250	1263	111	1526	209	2147	1705
Mg	65.7	9.3	58.9	4.5	57.8	6.4	79.1	14.8	70.4	9.7
Ca	49.5	9.3	36.8	10.8	37.5	10.2	46.1	10.9	58.6	14.1
Na	7.2	2.8	6.0	3.6	3.9	1.2	4.0	1.2	8.2	7.0
µg kg^{-1}										
Rb	1269	315	949.9	149.3	864.5	121.1	943.7	256.2	1810	576
Zn	415.0	81.5	436.7	114.6	309.3	74.6	479.7	111.6	345.9	162.4
Cu	268.9	59.6	219.7	67.5	241.4	59.6	299.8	97.4	261.5	93.1
Fe	277.2	104.7	257.0	256.0	192.8	116.3	310.6	98.6	226.3	169.2
Sr	210.5	59.3	133.6	45.9	161.3	46.2	129.8	45.3	136.5	49.5
Mn	98.8	33.7	104.1	35.7	83.4	30.0	94.1	24.8	124.1	64.3
Ba	55.9	18.2	46.3	24.2	68.3	18.0	27.2	16.8	87.7	23.2
Mo	6.1	3.8	5.6	3.2	5.4	1.6	12.9	4.3	5.1	3.6
Ti	8.0	3.8	5.5	4.7	3.2	1.1	10.8	5.3	4.5	2.4
Co	0.6	0.3	0.4	0.2	0.6	0.4	0.7	0.2	1.2	0.6

Table 2. Mean contents ± standard deviation (SD) of pseudo total (*PTE*) and bioavailable (*BAE*) elements in soils. A nonparametric multiple test (Kruskal–Wallis test) and Dunn's comparison for post hoc were applied. Lowercase letters (a, b, c) and uppercase letters (A, B, C) were used to compare the mean ranks of distributions of EDTA and AR contents, respectively. n: total number of samples.

Elements (mg/kg)	**depths (cm)**	**PGI-area**						**NO-PGI area**						**NO-PGI area**					
		(MS; PZ;SR)-PGI (n 16)						**PR-NOPGI (n 4)**						**EB-NOPGI (n 4)**					
		PTE			***BAE***			***PTE***			***BAE***			***PTE***			***BAE***		
		Mean	***±SD***		***Mean***	***±SD***		***Mean***	***±SD***		***Mean***	***±SD***		***Mean***	***±SD***		***Mean***	***±SD***	
Al	**0-20**	47310	6680	A	116.5	31.0	a	48019	3184	A	84.32	7.46	b	34672	4565	B	113.8	1.0	ab
	40-50	47640	7330	A	106.2	21.8	a	49507	2453	A	80.45	11.49	b	33515	776	B	96.41	5.58	ab
Fe	**0-20**	27078	2634	B	135.3	43.6	a	28310	2182	A	100.3	12.0	b	26595	1315	B	103.4	6.7	ab
	40-50	27509	2783	A	116.4	39.8	b	29222	1152	A	102.5	9.3	ab	26632	1265	A	76.63	3.77	a
K	**0-20**	16455	7061	B	1015	110	a	31972	3151	A	554.2	11.9	b	7154	8241	C	1030	226	a
	40-50	16502	7459	B	993.5	121.7	a	33405	1855	A	1313	822	a	4741	228	C	585.2	42.0	b
Ca	**0-20**	14468	5051	B	4914	2804	a	19482	697	A	4078	839	a	14487	3770	B	5823	1594	a
	40-50	13389	4410	B	4281	2483	a	19926	607	A	4084	804	a	15316	4655	AB	5325	1369	a
Mg	**0-20**	5442	825	B	238.1	71.2	b	8635	1734	A	150.7	12.7	c	6166	976	B	678.8	9.6	a
	40-50	5522	907	AB	244.1	69.7	b	8272	321	A	160.6	2.3	c	6043	865	B	554.3	66.7	a
Na	**0-20**	6285	3138	B	117.8	36.6	b	10760	592	A	158.0	33.4	a	1630	2860	C	136.1	1.5	a
	40-50	6377	3307	B	136.8	43.5	ab	11113	413	A	171.3	17.9	a	829	52	C	119.4	9.3	b
Ti	**0-20**	1671	171	A	6.49	2.03	a	1758	146	A	3.57	0.50	b	1056	224	B	4.80	0.08	b
	40-50	1687	208	A	6.05	1.92	a	1821	62	A	3.21	0.83	b	1003	54	B	3.63	0.12	b
Mn	**0-20**	869.0	79.6	B	40.1	29.6	b	647.5	38.5	C	25.84	7.26	b	1405	292	A	364.8	27.7	a
	40-50	880.3	76.5	B	44.39	31.20	b	662.0	30.9	C	24.12	9.06	b	1502	192	A	293.8	4.0	a
Ba	**0-20**	495.0	82.9	B	103.7	178.2	b	638.3	58.7	A	29.76	3.92	c	340.4	82.9	C	120.7	31.1	a
	40-50	499.3	91.9	B	243.8	325.7	a	657.4	37.5	A	29.34	4.20	b	321.1	20.0	C	65.29	4.15	a

Cu	**0-20**	161.8	127.3	A	77.68	28.59	a	104.0	8.0	A	35.72	5.92	b	41.28	21.30	B	20.49	3.56	b
	40-50	126.8	52.5	A	44.71	20.07	a	105.3	7.4	A	40.49	0.75	a	31.42	7.63	B	6.80	0.38	b
Rb	**0-20**	382.5	152.1	B	3.19	0.48	a	636.8	68.7	A	1.93	0.42	b	229.7	125.7	C	3.72	0.32	a
	40-50	384.8	159.9	B	3.07	0.72	b	666.8	48.8	A	3.38	1.64	ab	196.0	5.4	C	4.46	0.56	b
Sr	**0-20**	184.7	44.5	B	24.39	5.04	a	322.1	25.2	A	20.55	4.06	ab	77.81	73.13	C	19.57	2.04	b
	40-50	185.8	47.5	B	25.70	6.16	b	332.5	12.5	A	21.08	3.37	ab	57.59	4.94	C	15.89	2.02	b
Zn	**0-20**	118.0	25.1	A	19.99	6.10	a	90.67	7.00	B	5.25	0.36	b	63.46	8.32	B	5.71	1.96	b
	40-50	114.1	23.1	A	14.42	8.53	a	91.06	7.13	AB	5.83	0.34	ab	60.69	2.92	B	1.94	0.39	a
Pb	**0-20**	65.47	13.80	A	15.69	3.25	a	56.04	8.11	A	12.48	1.99	ab	43.17	6.75	B	9.87	0.44	b
	40-50	63.64	13.71	A	21.68	22.26	a	58.95	9.53	A	15.21	1.13	a	41.92	2.98	B	8.74	0.29	b
Ni	**0-20**	19.11	4.59	A	0.24	0.10	b	17.70	1.08	A	0.20	0.02	b	18.27	0.92	A	2.80	0.11	a
	40-50	19.67	4.63	A	0.23	0.18	b	18.32	0.61	A	0.36	0.22	ab	18.65	1.16	A	2.45	0.28	a
Cr	**0-20**	13.98	4.88	B	0.11	0.05	b	13.31	1.51	B	0.07	0.01	b	19.73	2.63	A	0.30	0.01	a
	40-50	14.15	4.81	B	0.08	0.05	b	12.95	0.71	B	0.32	0.28	ab	20.87	1.03	A	0.28	0.02	a
Co	**0-20**	13.83	1.62	B	0.16	0.12	b	15.74	1.34	A	0.08	0.04	b	17.20	1.45	A	4.64	0.37	a
	40-50	14.09	1.64	B	0.19	0.15	b	16.33	0.57	A	0.21	0.11	b	17.56	1.67	A	3.68	0.55	a
Mo	**0-20**	1.23	0.23	B	0.03	0.06	b	2.27	0.25	A	0.04	0.01	ab	1.23	0.35	B	0.20	0.03	a
	40-50	1.23	0.25	B	0.04	0.05	b	2.35	0.16	A	0.15	0.13	a	1.17	0.13	B	0.19	0.04	a
Cd	**0-20**	0.28	0.10	B	0.10	0.03	b	0.24	0.07	B	0.08	0.00	c	0.42	0.09	A	0.31	0.02	a
	40-50	0.28	0.09	B	0.07	0.02	b	0.21	0.04	B	0.20	0.16	ab	0.45	0.03	A	0.28	0.05	a

Table 3. Significant correlation coefficients (different from 0 with a significance level α=0.05) between elements in lemon juices and pseudo-total (PTE) and bioavailable (BAE) elements in soils at two depths and frequencies in percent of significant coefficients from five iterations. ns: not significant.

Depths	**Elements**	***PTE***	***Freq. %***	***BAE***	***Freq. %***
Top-soil (0-20 cm)	Mg	*ns*		0.350	*80*
	Rb	-0.335	*100*	0.414	*100*
	Sr	-0.381	*60*	*ns*	
	Mn	0.387	*60*	0.457	*80*
	Ba	-0.574	*80*	*ns*	
	Mo	0.428	*80*	*ns*	
	Co	*ns*		0.410	*60*
Sub-soil (40-50 cm)	Mg	*ns*		0.373	*80*
	Rb	-0.392	*80*	0.446	*100*
	Sr	-0.306	*80*	0.432	*100*
	Mn	0.429	*60*	0.478	*80*
	Ba	-0.503	*80*	*ns*	
	Mo	0.525	*80*	*ns*	
	Co	*ns*		0.395	*60*

Figure 1. Location of orchards of soils and lemons sampling.

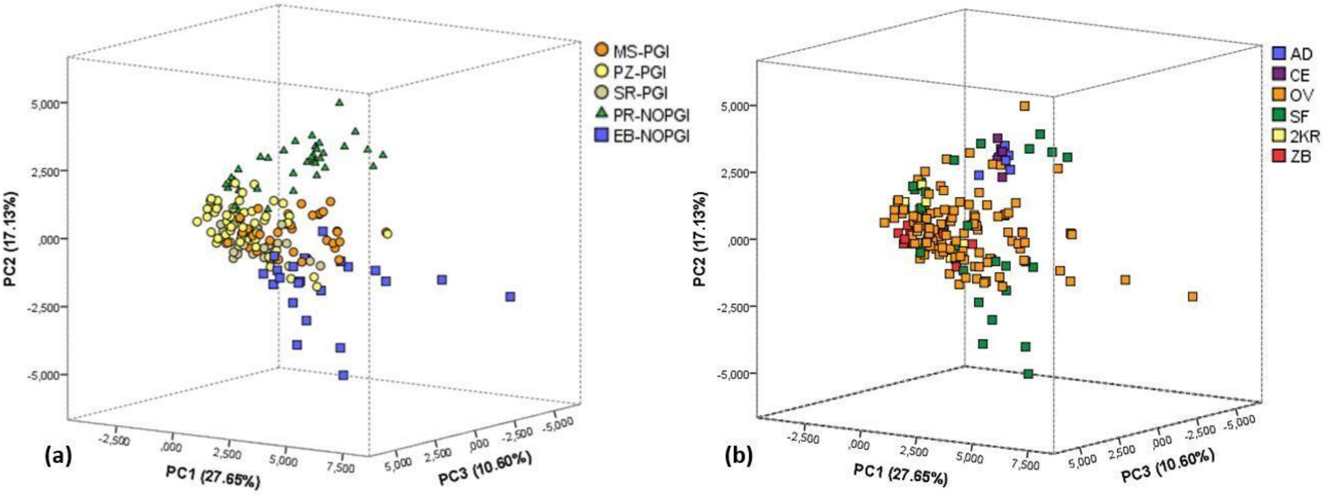


Figure 2. Score plot of Principal Component Analysis (PCA) of lemon juices labelled according to: (a) orchard farms: MS-PGI, PZ-PGI, SR-PGI, PR-NOPGI and EB-NOPGI; (b) lemon cultivars: Ovale di Sorrento (OV), Sfusato Amalfitano (SF), Adamo (AD), Cerza (CE), Siracusano 2KR (2KR), Zagara Bianca (ZB).

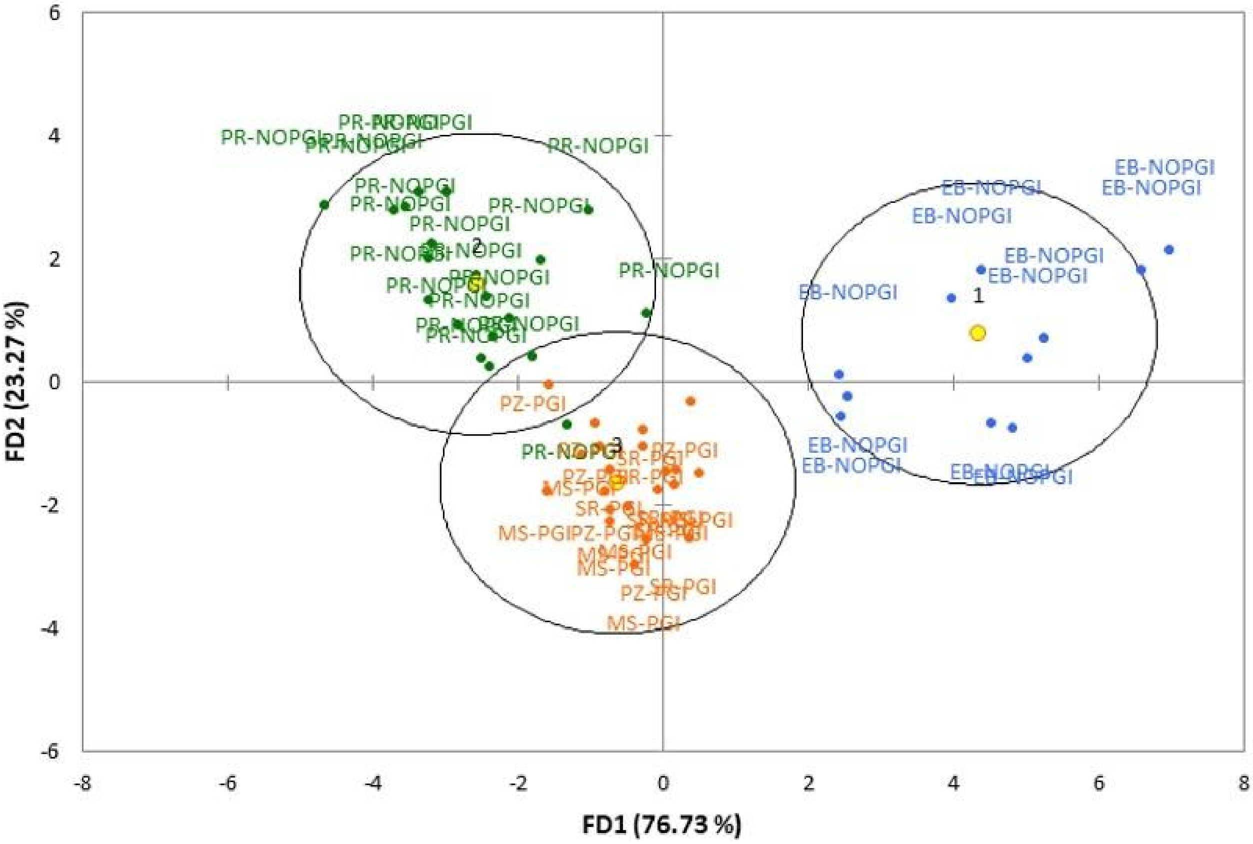


Figure 3. S-LDA score plot for the three groups of production area (14 elements in 88 lemon juice samples).

Tab. S1. ICP-OES operating conditions.

Instrument parameters	**Setting**
Read time (s)	15
Replicates	3
Sample uptake delay (s)	25
Fast pump	yes
Stabilisation time (s)	15
Pump speed (rpm)	12
RF power (W)	1450
Aux flow rate (L min^{-1})	1.00
Plasma flow rate (L min^{-1})	12.0
Nebulizer flow rate (L min^{-1})	0.70
Elements (nm)	Al (396.152)* – Ba (493.408)* – Ca(317.933) – Cd(228.802) – Co(228.615) – Cr(267.716) – Cu(327.395) – Fe(238.204)* – K(766.491)* – Mg(280.270)* – Mn(257.610)* – Mo(202.032) – Na(588.995) – Ni(231.604) – Pb(220.353) – Rb(780.026) – Sr(407.771) – Ti(334.941) – Zn(231.857)
Viewing modes	Axial, radial(*)
Viewing height (mm)	8
Background correction	fitted

Tab. S2. ICP-MS operating conditions.

Instrument parameters	**Settings**
RF Power (W)	1550
Sampling Depth (mm)	8.0
Plasma flow rate (L min^{-1})	15.0
Carrier Gas (L min^{-1})	Ar: 1.05
Dilution Gas (L min^{-1})	Ar: 0.10
Tune mode	Helium mode
He flow (ml min^{-1})	4.0
Peltier-cooler multipass condenser	2°C
Oxide/Doubly Charged Ratio	$^{140}Ce^{16}O^+/^{140}Ce^+$: 0.3%
	$^{140}Ce^{2+}/^{140}Ce$: 2.3%
Internal standards	^{6}Li, ^{45}Sc, ^{72}Ge, ^{89}Y, ^{115}In, ^{159}Tb
Selected isotopes	^{47}Ti, ^{56}Fe, ^{55}Mn, ^{59}Co, ^{63}Cu, ^{66}Zn, ^{85}Rb, ^{88}Sr, ^{95}Mo, ^{137}Ba, ^{43}Ca, ^{44}Ca, ^{39}K, ^{24}Mg, ^{23}Na

Tab. S3. Validation parameters for the determinations of elements in lemon juices: method detection limits (MDL), dynamic range, dependent coefficient of calibration curves (R^2), values of spike, recovery on spike, and relative standard deviation (RSD).

	MDL (µg/L)	**Dynamic range** (µg/L)	**R^2**	**Spike values** (µg/L)	**Recovery** (%)	**RSD** (%)
Ti	2.0	0 - 10	0.9992	1	106 ± 3.7	3.5
Fe	75	0 - 30	0.9989	10	103 ± 12	12
Mn	12	0 - 5	0.9996	1	105 ± 7.2	6.9
Co	0.2	0 - 1	0.9999	1	94 ± 4.0	4.2
Cu	16	0 - 30	1.0000	5	105 ± 11	12
Zn	220	0 - 30	0.9999	5	92 ± 11	12
Rb	2.1	0- 50	0.9993	5	100 ± 12	12
Sr	20.2	0 - 5	0.9999	5	110 ± 2.6	2.4
Mo	1.4	0 - 5	1.0000	1	112 ± 6.4	5.8
Ba	8.5	0 - 10	0.9999	5	109 ± 4.5	5.0
Ca	6×10^4	$0 – 1\text{x}\ 10^4$	0.9999	2×10^3	104 ± 7.9	7.6
K	$1\text{x}10^5$	$0 - 2\times10^5$	1.0000	$1\text{x}10^5$	85 ± 3.5	4.1
Mg	6×10^4	$0 - 10^4$	0.9999	2×10^3	93 ± 8.0	8.6
Na	3×10^3	$0 - 2\times10^3$	1.0000	2×10^3	106 ± 7.3	6.9

Tab. S4. Means ± standard deviation (SD) of chemical and physical characteristics of soil at different depths (0-20 cm and 40-50 cm). Nonparametric multiple tests (p values of Kruskal–Wallis Test) was applied. Lowercase letters (a, b, c) and uppercase letters (A, B, C) were used to compare the mean ranks of distributions between two depths and among the five different orchards respectively. n° = number of top and sub soil samples

Properties	depths (cm)	PGI-area									NO-PGI area			NO-PGI area		
		MS-PGI (n°4)			**PZ-PGI (n°10)**			**SR-PGI (n°6)**			**PR-NOPGI (n°4)**			**EB-NOPGI (n°4)**		
		Mean	**±SD**		**Mean**	**±SD**		**Mean**	**±SD**		**Mean**	**±SD**		**Mean**	**±SD**	
pH *(H_2O)*	**0-20**	7.13	*0.08*	b-B	7.61	*0.23*	a-A	7.66	*0.20*	a-A	7.73	0.17	b-A	7.67	*0.08*	b-A
	40-50	7.43	*0.09*	a-B	7.69	*0.26*	a-B	7.67	*0.27*	a-B	8.10	0.04	a-A	7.93	*0.09*	a-AB
C.E.C (cmol(+) kg^{-1})	**0-20**	21.02	*1.47*	a-BC	24.53	*1.23*	a-AB	25.00	*4.84*	a-AB	12.53	1.28	a-C	27.35	*0.13*	a-A
	40-50	16.73	*0.93*	b-BC	21.13	*2.09*	b-AB	25.50	*3.65*	a-A	9.96	3.21	a-C	23.80	*0.13*	b-A
tot. Carb. (g kg^{-1})	**0-20**	5.56	*0.60*	a-BC	3.99	*1.09*	a-C	31.23	*40.41*	a-A	9.04	0.32	b-BC	19.64	*6.25*	a-AB
	40-50	4.21	*0.26*	b-BC	3.70	*2.76*	a-C	15.92	*15.05*	b-AB	14.74	1.56	a-AB	17.43	*5.75*	a-A
O.C. (g kg^{-1})	**0-20**	18.97	*9.74*	a-A	14.44	*3.32*	a-A	18.59	*6.30*	a-A	11.66	2.25	a-A	12.10	*0.30*	a-A
	40-50	5.73	*0.17*	b-C	11.52	*2.23*	b-A	10.18	*1.19*	b-AB	9.23	1.69	a-ABC	7.74	*1.34*	b-BC
sand (g kg^{-1})	**0-20**	675.21	*25.12*	a-BC	737.86	*22.91*	a-AB	619.06	*94.35*	a-BC	870.81	8.03	a-A	414.69	*55.73*	a-C
	40-50	657.30	*25.12*	a-BC	721.34	*31.02*	a-AB	591.90	*52.20*	a-C	847.10	42.66	a-A	416.95	*29.76*	a-C
silt (g kg^{-1})	**0-20**	164.66	*45.12*	a-BC	198.15	*15.07*	a-ABC	227.12	*41.07*	a-AB	125.44	5.93	a-C	271.30	*39.35*	a-A
	40-50	207.34	*45.12*	a-AB	202.27	*28.75*	a-AB	225.63	*12.51*	a-A	145.90	43.49	a-B	238.25	*23.33*	a-A
clay (g kg^{-1})	**0-20**	160.14	*34.38*	a-AB	63.99	*18.24*	a-BC	153.82	*55.36*	a-AB	3.75	2.09	b-C	314.01	*16.37*	b-A
	40-50	135.36	*34.38*	a-AB	76.39	*8.41*	b-BC	182.47	*40.66*	a-A	7.00	0.84	a-C	344.81	*15.13*	a-A

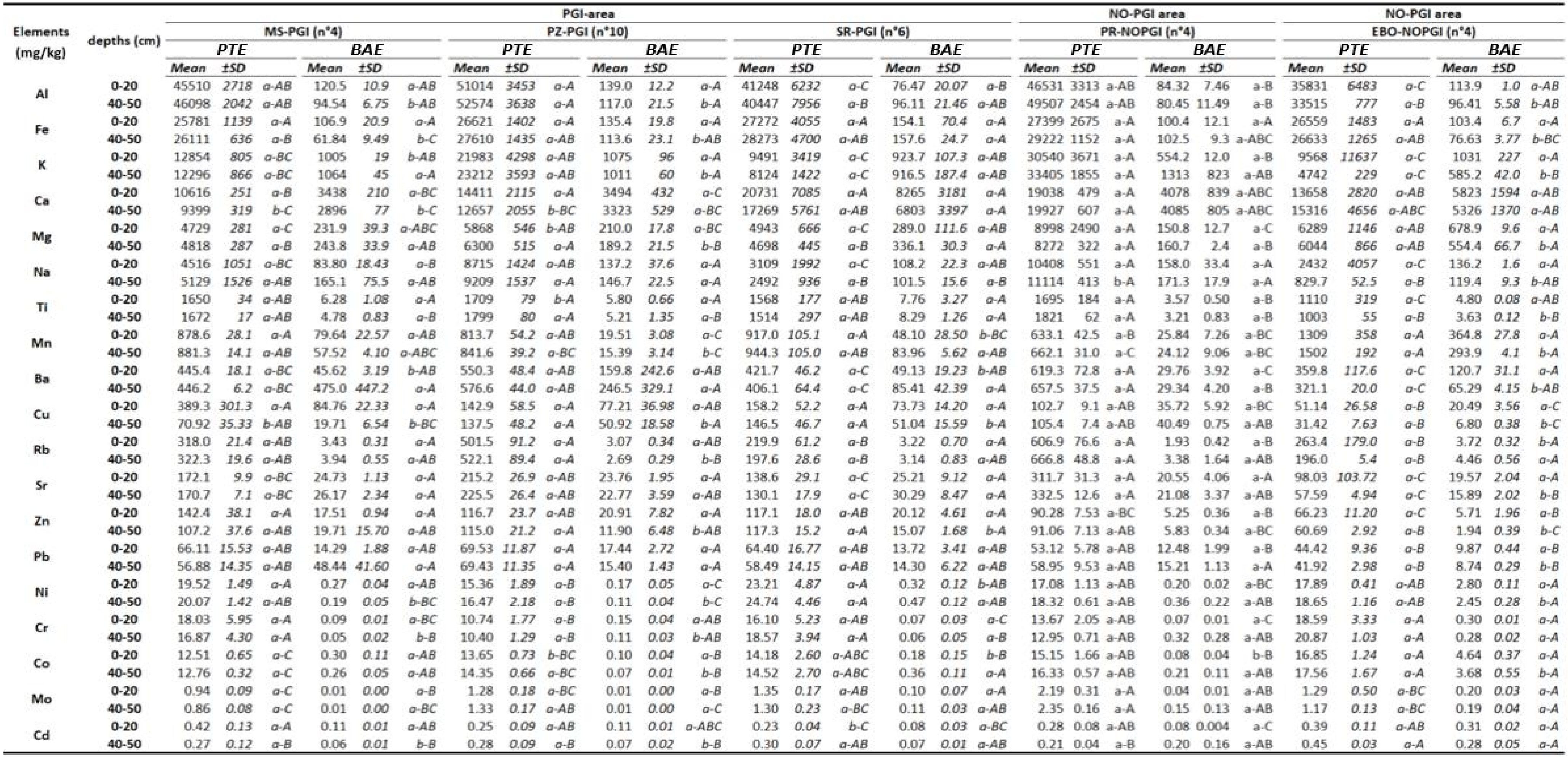
Tab. S5. Validation parameters for the determinations of pseudo total elements in soils: method detection limits (MDL), dynamic range, dependent coefficient of calibration curves (R^2), certified reference material values (BCR 141R), recovery on certified soils, and relative standard deviation (RSD).

	MDL (mg/L)	**Dynamic range** (mg/L)	**R^2**	**BCR 141R** (mg/kg)	**Recovery** (%)	**RSD** (%)
Al	5.1	0-500	1.000			
Ba	0.04	0-5	1.000			
Ca	5.2	0-500	0.99999			
Cd	0.04	0-1.0	1.000	14.0 ± 0.4	97 ± 1.4	1.5
Co	0.11	0-1.0	1.000	9.2 ± 0.5	85 ± 1.8	2.1
Cr	0.05	0-1.0	1.000	138 ± 5	80 ± 0.5	1.5
Cu	0.15	0-5.0	1.000	46.9 ± 1.8	98 ± 0.8	1.6
Fe	0.24	0-500	1.0000			
K	7.6	0-500	0.99997			
Mg	0.31	0-500	1.0000			
Mn	0.01	0-40.0	0.99999	653 ± 16	96 ± 0.6	1.8
Mo	0.14	0-1.0	1.0000			
Na	0.24	0-500	0.99999			
Ni	0.12	0-1.0	1.0000	94 ± 5	85 ± 0.5	1.3
Pb	0.44	0-1.0	1.0000	51.3 ± 2.0	89 ± 0.6	1.6
Rb	1.5	0-40.0	0.99807			
Sr	0.09	0-5.0	0.99983			
Ti	0.01	0-40.0	1.0000			
Zn	0.05	0-5.0	1.0000	270 ± 8	96 ± 0.6	1.6

Tab. S6. Validation parameters for the determinations of bioavailable elements in soils: method detection limits (MDL), dynamic range, dependent coefficient of calibration curves (R^2), values of certified reference materials (BCR 700), recovery on certified soils and relative standard deviation (RSD).

	MDL (mg/L)	**Dynamic range** (mg/L)	R^2	**BCR 700** (mg/kg)	**Recovery** (%)	**RSD** (%)
Al	0.19	0-500	1.000			
Ba	0.0021	0-5	1.000			
Ca	0.047	0-500	1.000			
Cd	0.0021	0-1.0	1.000	65 ± 4	95 ± 0.6	0.6
Co	0.0060	0-1.0	1.000			
Cr	0.011	0-1.0	1.000	10 ± 1	94 ± 0.5	0.6
Cu	0.0050	0-5.0	1.000	89 ± 3	93 ± 0.6	0.6
Fe	0.085	0-500	1.0000			
K	0.22	0-500	0.99997			
Mg	0.046	0-500	1.0000			
Mn	0.0021	0-40.0	0.99999			
Mo	0.012	0-1.0	1.0000			
Na	0.10	0-500	0.99999			
Ni	0.0090	0-1.0	1.0000	53 ± 3	86 ± 0.5	0.6
Pb	0.019	0-1.0	1.0000	103 ± 5	89 ± 0.3	0.3
Rb	0.002	0-40.0	0.99807			
Sr	0.0010	0-5.0	0.99983			
Ti	0.011	0-40.0	1.0000			
Zn	0.028	0-5.0	1.0000	510 ± 17	88 ± 0.5	0.6

Tab.S7. Means ± standard deviation (SD) of the contents of pseudo total elements (*PTE*) and bioavailable elements (*BAE*) in soils at different depths (0-20 cm and 40-50 cm). A nonparametric multiple test (p values of Kruskal–Wallis Test) was applied. Lowercase letters (a, b, c) and uppercase letters (A, B, C) were used to compare the mean ranks of distributions of depths and different orchards, respectively. n°: number of top and subsoil samples



Elements (mg/kg)	depths (cm)	PGI-area																		NO-PGI area						NO-PGI area					
		MS-PGI (n°4)						PZ-PGI (n°10)						SR-PGI (n°6)						PR-NOPGI (n°4)						EBO-NOPGI (n°4)					
		PTE			BAE			PTE			BAE			PTE			BAE			PTE			BAE			PTE			BAE		
		Mean	±SD		Mean	±SD		Mean	±SD		Mean	±SD		Mean	±SD		Mean	±SD		Mean	±SD		Mean	±SD		Mean	±SD		Mean	±SD	
Al	0-20	45510	2718	a-AB	120.5	10.9	a-AB	51014	3453	a-A	139.0	12.2	a-A	41248	6232	a-C	76.47	20.07	a-B	46531	3313	a-AB	84.32	7.46	a-B	35831	6483	a-C	113.9	1.0	a-AB
	40-50	46098	2042	a-AB	94.54	6.75	b-AB	52574	3638	a-A	117.0	21.5	b-A	40447	7956	a-B	96.11	21.46	a-AB	49507	2454	a-AB	80.45	11.49	a-B	33515	777	a-B	96.41	5.58	b-AB
Fe	0-20	25781	1139	a-A	106.9	20.9	a-A	26621	1402	a-A	135.4	19.8	a-A	27272	4055	a-A	154.1	70.4	a-A	27399	2675	a-A	100.4	12.1	a-A	26559	1483	a-A	103.4	6.7	a-A
	40-50	26111	636	a-B	61.84	9.49	b-C	27610	1435	a-AB	113.6	23.1	b-AB	28273	4700	a-AB	157.6	24.7	a-A	29222	1152	a-A	102.5	9.3	a-ABC	26633	1265	a-AB	76.63	3.77	b-BC
K	0-20	12854	805	a-BC	1005	19	b-AB	21983	4298	a-AB	1075	96	a-A	9491	3419	a-C	923.7	107.3	a-AB	30540	3671	a-A	554.2	12.0	a-B	9568	11637	a-C	1031	227	a-A
	40-50	12296	866	a-BC	1064	45	a-A	23212	3593	a-AB	1011	60	b-A	8124	1422	a-C	916.5	187.4	a-AB	33405	1855	a-A	1313	823	a-AB	4742	229	a-C	585.2	42.0	b-B
Ca	0-20	10616	251	a-B	3438	210	a-BC	14411	2115	a-A	3494	432	a-C	20731	7085	a-A	8265	3181	a-A	19038	479	a-A	4078	839	a-ABC	13658	2820	a-AB	5823	1594	a-AB
	40-50	9399	319	b-C	2896	77	b-C	12657	2055	b-BC	3323	529	a-BC	17269	5761	a-AB	6803	3397	a-A	19927	607	a-A	4085	805	a-ABC	15316	4656	a-ABC	5326	1370	a-AB
Mg	0-20	4729	281	a-C	231.9	39.3	a-ABC	5868	546	b-AB	210.0	17.8	a-BC	4943	666	a-C	289.0	111.6	a-AB	8998	2490	a-A	150.8	12.7	a-C	6289	1146	a-AB	678.9	9.6	a-A
	40-50	4818	287	a-B	243.8	33.9	a-AB	6300	515	a-A	189.2	21.5	b-B	4698	445	a-B	336.1	30.3	a-A	8272	322	a-A	160.7	2.4	a-B	6044	866	a-AB	554.4	66.7	b-A
Na	0-20	4516	1051	a-BC	83.80	18.43	a-B	8715	1424	a-AB	137.2	37.6	a-A	3109	1992	a-C	108.2	22.3	a-AB	10408	551	a-A	158.0	33.4	a-A	2432	4057	a-C	136.2	1.6	a-A
	40-50	5129	1526	a-AB	165.1	75.5	a-AB	9209	1537	a-A	146.7	22.5	a-A	2492	936	a-B	101.5	15.6	a-B	11114	413	b-A	171.3	17.9	a-A	829.7	52.5	a-B	119.4	9.3	b-AB
Ti	0-20	1650	34	a-AB	6.28	1.08	a-A	1709	79	b-A	5.80	0.66	a-A	1568	177	a-AB	7.76	3.27	a-A	1695	184	a-A	3.57	0.50	a-B	1110	319	a-C	4.80	0.08	a-AB
	40-50	1672	17	a-AB	4.78	0.83	a-B	1799	80	a-A	5.21	1.35	a-B	1514	297	a-AB	8.29	1.26	a-A	1821	62	a-A	3.21	0.83	a-B	1003	55	a-B	3.63	0.12	b-B
Mn	0-20	878.6	28.1	a-A	79.64	22.57	a-AB	813.7	54.2	a-AB	19.51	3.08	a-C	917.0	105.1	a-A	48.10	28.50	b-BC	633.1	42.5	a-B	25.84	7.26	a-BC	1309	358	a-A	364.8	27.8	a-A
	40-50	881.3	14.1	a-AB	57.52	4.10	a-ABC	841.6	39.2	a-BC	15.39	3.14	b-C	944.3	105.0	a-AB	83.96	5.62	a-AB	662.1	31.0	a-C	24.12	9.06	a-BC	1502	192	a-A	293.9	4.1	b-A
Ba	0-20	445.4	18.1	a-BC	45.62	3.19	b-AB	550.3	48.4	a-AB	159.8	242.6	a-AB	421.7	46.2	a-C	49.13	19.23	b-AB	619.3	72.8	a-A	29.76	3.92	a-C	359.8	117.6	a-C	120.7	31.1	a-A
	40-50	446.2	6.2	a-BC	475.0	447.2	a-A	576.6	44.0	a-AB	246.5	329.1	a-A	406.1	64.4	a-C	85.41	42.39	a-A	657.5	37.5	a-A	29.34	4.20	a-B	321.1	20.0	a-C	65.29	4.15	b-AB
Cu	0-20	389.3	301.3	a-A	84.76	22.33	a-A	142.9	58.5	a-A	77.21	36.98	a-AB	158.2	52.2	a-A	73.73	14.20	a-A	102.7	9.1	a-AB	35.72	5.92	a-BC	51.14	26.58	a-B	20.49	3.56	a-C
	40-50	70.92	35.33	b-AB	19.71	6.54	b-BC	137.5	48.2	a-A	50.92	18.58	b-A	146.5	46.7	a-A	51.04	15.59	b-A	105.4	7.4	a-AB	40.49	0.75	a-AB	31.42	7.63	a-B	6.80	0.38	b-C
Rb	0-20	318.0	21.4	a-AB	3.43	0.31	a-A	501.5	91.2	a-A	3.07	0.34	a-AB	219.9	61.2	a-B	3.22	0.70	a-A	606.9	76.6	a-A	1.93	0.42	a-B	263.4	179.0	a-B	3.72	0.32	b-A
	40-50	322.3	19.6	a-AB	3.94	0.55	a-AB	522.1	89.4	a-A	2.69	0.29	b-B	197.6	28.6	a-B	3.14	0.83	a-AB	666.8	48.8	a-A	3.38	1.64	a-AB	196.0	5.4	a-B	4.46	0.56	a-A
Sr	0-20	172.1	9.9	a-BC	24.73	1.13	a-A	215.2	26.9	a-AB	23.76	1.95	a-A	138.6	29.1	a-C	25.21	9.12	a-A	311.7	31.3	a-A	20.55	4.06	a-A	98.03	103.72	a-C	19.57	2.04	a-A
	40-50	170.7	7.1	a-BC	26.17	2.34	a-A	225.5	26.4	a-AB	22.77	3.59	a-AB	130.1	17.9	a-C	30.29	8.47	a-A	332.5	12.6	a-A	21.08	3.37	a-AB	57.59	4.94	a-C	15.89	2.02	b-B
Zn	0-20	142.4	38.1	a-A	17.51	0.94	a-A	116.7	23.7	a-AB	20.91	7.82	a-A	117.1	18.0	a-AB	20.12	4.61	a-A	90.28	7.53	a-BC	5.25	0.36	a-B	66.23	11.20	a-C	5.71	1.96	a-B
	40-50	107.2	37.6	a-AB	19.71	15.70	a-AB	115.0	21.2	a-A	11.90	6.48	b-AB	117.3	15.2	a-A	15.07	1.68	b-A	91.06	7.13	a-AB	5.83	0.34	a-BC	60.69	2.92	a-B	1.94	0.39	b-C
Pb	0-20	66.11	15.53	a-AB	14.29	1.88	a-AB	69.53	11.87	a-A	17.44	2.72	a-A	64.40	16.77	a-AB	13.72	3.41	a-AB	53.12	5.78	a-AB	12.48	1.99	a-B	44.42	9.36	a-B	9.87	0.44	a-B
	40-50	56.88	14.35	a-AB	48.44	41.60	a-A	69.43	11.35	a-A	15.40	1.43	a-A	58.49	14.15	a-AB	14.30	6.22	a-AB	58.95	9.53	a-AB	15.21	1.13	a-A	41.92	2.98	a-B	8.74	0.29	b-B
Ni	0-20	19.52	1.49	a-A	0.27	0.04	a-AB	15.36	1.89	a-B	0.17	0.05	a-C	23.21	4.87	a-A	0.32	0.12	b-AB	17.08	1.13	a-AB	0.20	0.02	a-BC	17.89	0.41	a-AB	2.80	0.11	a-A
	40-50	20.07	1.42	a-AB	0.19	0.05	b-BC	16.47	2.18	a-B	0.11	0.04	b-C	24.74	4.46	a-A	0.47	0.12	a-AB	18.32	0.61	a-AB	0.36	0.22	a-AB	18.65	1.16	a-AB	2.45	0.28	b-A
Cr	0-20	18.03	5.95	a-A	0.09	0.01	a-BC	10.74	1.77	a-B	0.15	0.04	a-AB	16.10	5.23	a-AB	0.07	0.03	a-C	13.67	2.05	a-AB	0.07	0.01	a-C	18.59	3.33	a-A	0.30	0.01	a-A
	40-50	16.87	4.30	a-A	0.05	0.02	b-B	10.40	1.29	a-B	0.11	0.03	b-AB	18.57	3.94	a-A	0.06	0.05	a-B	12.95	0.71	a-AB	0.32	0.28	a-AB	20.87	1.03	a-A	0.28	0.02	a-A
Co	0-20	12.51	0.65	a-C	0.30	0.11	a-AB	13.65	0.73	b-BC	0.10	0.04	a-B	14.18	2.60	a-ABC	0.18	0.15	b-B	15.15	1.66	a-AB	0.08	0.04	b-B	16.85	1.24	a-A	4.64	0.37	a-A
	40-50	12.76	0.32	a-C	0.26	0.05	a-AB	14.35	0.66	a-BC	0.07	0.01	b-B	14.52	2.70	a-ABC	0.36	0.11	a-A	16.33	0.57	a-AB	0.21	0.11	a-AB	17.56	1.67	a-A	3.68	0.55	b-A
Mo	0-20	0.94	0.09	a-C	0.01	0.00	a-B	1.28	0.18	a-BC	0.01	0.00	a-B	1.35	0.17	a-AB	0.10	0.07	a-A	2.19	0.31	a-A	0.04	0.01	a-AB	1.29	0.50	a-BC	0.20	0.03	a-A
	40-50	0.86	0.08	a-C	0.01	0.00	a-BC	1.33	0.17	a-AB	0.01	0.00	a-C	1.30	0.23	a-BC	0.11	0.03	a-AB	2.35	0.16	a-A	0.15	0.13	a-AB	1.17	0.13	a-BC	0.19	0.04	a-A
Cd	0-20	0.42	0.13	a-A	0.11	0.01	a-AB	0.25	0.09	a-AB	0.11	0.01	a-ABC	0.23	0.04	b-C	0.08	0.03	a-BC	0.28	0.08	a-AB	0.08	0.004	a-C	0.39	0.11	a-AB	0.31	0.02	a-A
	40-50	0.27	0.12	a-B	0.06	0.01	b-B	0.28	0.09	a-B	0.07	0.02	b-B	0.30	0.07	a-AB	0.07	0.01	a-AB	0.21	0.04	a-B	0.20	0.16	a-AB	0.45	0.03	a-A	0.28	0.05	a-A

Supplementary References

Anderson, K. A., & Smith, B. W. (2005). Use of chemical profiling to differentiate geographic growing origin of raw pistachios. Journal of Agricultural and Food Chemistry, 53(2), 410–418. https://doi.org/10.1021/jf048907u

Batista, B. L., da Silva, L. R. S., Rocha, B. A., Rodrigues, J. L., Berretta-Silva, A. A., Bonates, T. O., Gomes, V. S. D., Barbosa, R. M., & Barbosa, F. (2012). Multi-element determination in Brazilian honey samples by inductively coupled plasma mass spectrometry and estimation of geographic origin with data mining techniques. Food Research International, 49(1), 209–215. https://doi.org/10.1016/j.foodres.2012.07.015

Cheajesadagul, P., Arnaudguilhem, C., Shiowatana, J., Siripinyanond, A., & Szpunar, J. (2013). Discrimination of geographical origin of rice based on multi-element fingerprinting by high resolution inductively coupled plasma mass spectrometry. Food Chemistry, 141(4), 3504–3509. https://doi.org/10.1016/j.foodchem.2013.06.060

Di Bella, G., Lo Turco, V., Potortì, A. G., Luppino, R. R., Fotia, V., Conte, F., & Mo Dugo, G. (2012). Classification of the geographical origin of Italian donkey's milk based on differences in inorganic anions. Food Additives and Contaminants - Part A Chemistry, Analysis, Control, Exposure and Risk Assessment, 29(7), 1021–1029. https://doi.org/10.1080/19440049.2012.674979

Di Bella, G., Turco, V. Lo, Potortì, A. G., Rando, R., Licata, P., & Dugo, G. (2013). Statistical analysis of heavy metals in Cerastoderma edule glaucum and Venerupis aurea laeta from Ganzirri Lake, Messina (Italy). Environmental Monitoring and Assessment, 185(9), 7517–7525. https://doi.org/10.1007/s10661-013-3116-4

Di Bella, Giuseppa, Naccari, C., Bua, G. D., Rastrelli, L., Lo Turco, V., Potortì, A. G., & Dugo, G. (2016). Mineral composition of some varieties of beans from Mediterranean and Tropical areas. International Journal of Food Sciences and Nutrition, 67(3), 239–248. https://doi.org/10.3109/09637486.2016.1153610

DOOR 2019, URL **https://ec.europa.eu/agriculture/quality/door/list.html** . Accessed 25.11.2019

EC No 510/2006, Council Regulation (EC) No 510/2006 of 20 March 2006 URL **https://eur-lex.europa.eu/legal-content/en/ALL/?uri=CELEX%3A32006R0510**). Accesed 17.08.2020

Franke, B. M., Haldimann, M., Reimann, J., Baumer, B., Gremaud, G., Hadorn, R., Bosset, J. O., & Kreuzer, M. (2007). Indications for the applicability of element signature analysis for the determination of the geographic origin of dried beef and poultry meat. European Food Research and Technology, 225(3–4), 501–509. https://doi.org/10.1007/s00217-006-0446-2

GU 2011, *Gazzetta Ufficiale Della Repubblica Italiana*, Ser. Gen. N° 24, Italy 31-1-2011 p.31. URL http://www.agricoltura.regione.campania.it/tipici/disciplinari/Disciplinare_di_produzione__Limone_di_Sorrento_IGP.pdf Accessed 27.11.2019

Gonzálvez, A., Armenta, S., & De La Guardia, M. (2011). Geographical traceability of "arrs de Valencia" rice grain based on mineral element composition. Food Chemistry, 126(3), 1254–1260. https://doi.org/10.1016/j.foodchem.2010.11.032

ISPRA. (2015) c. *NOTE ILLUSTRATIVE della CARTA GEOLOGICA D'ITALIA alla scala 1:50.000 Foce del Sele. Foglio 486*. Retrieved from https://www.isprambiente.gov.it/Media/carg/486_FOCE_DELSELE/Foglio.html

ISPRA. (2015) a. *NOTE ILLUSTRATIVE della CARTA GEOLOGICA D'ITALIA alla scala 1:50.000 Sorrento-Termini. Fogli 466-485*. Retrieved from https://www.isprambiente.gov.it/Media/carg/466_485_SORRENTO_TERMINI/Foglio.html

ISPRA. (2015) b. *NOTE ILLUSTRATIVE della CARTA GEOLOGICA D'ITALIA alla scala 1:50.000 Ercolano. Foglio 448*. Retrieved from https://www.isprambiente.gov.it/Media/carg/448_ERCOLANO/Foglio.html

ISTAT 2017, URL **http://dati.istat.it/Index.aspx?QueryId=36787&lang=en**. Accesed 10.07.2020

Kelly, S., Baxter, M., Chapman, S., Rhodes, C., Dennis, J., & Brereton, P. (2002). The application of isotopic and elemental analysis to determine the geographical origin of premium long grain rice. European Food Research and Technology, 214(1), 72–78. https://doi.org/10.1007/s002170100400

MiPAF (2000): Metodi di analisi chimica del suolo. Franco Angeli Ed., Milan, Italy.

Potortì, A. G., Bua, G. D., Lo Turco, V., Ben Tekaya, A., Beltifa, A., Ben Mansour, H., Dugo, G., & Di Bella, G. (2020). Major, minor and trace element concentrations in spices and aromatic herbs from Sicily (Italy) and Mahdia (Tunisia) by ICP-MS and multivariate analysis. Food Chemistry, 313(April 2019), 126094. https://doi.org/10.1016/j.foodchem.2019.126094

Potortì, A. G., Lo Turco, V., & Di Bella, G. (2021). Chemometric analysis of elements content in Algerian spices and aromatic herbs. Lwt, 138(November 2020), 110643. https://doi.org/10.1016/j.lwt.2020.110643

Potortí, A. G., Lo Turco, V., Saitta, M., Bua, G. D., Tropea, A., Dugo, G., & Di Bella, G. (2017). Chemometric analysis of minerals and trace elements in Sicilian wines from two different grape cultivars. Natural Product Research, 31(9), 1000–1005. https://doi.org/10.1080/14786419.2016.1261341

Simpkins, W. A., Louie, H., Wu, M., Harrison, M., & Goldberg, D. (2000). Trace elements in Australian orange juice and other products. Food Chemistry, 71(4), 423–433. https://doi.org/10.1016/S0308-8146(00)00150-3

Zhao, H., Guo, B., Wei, Y., Zhang, B., Sun, S., Zhang, L., & Yan, J. (2011). Determining the geographic origin of wheat using multielement analysis and multivariate statistics. Journal of Agricultural and Food Chemistry, 59(9), 4397–4402. https://doi.org/10.1021/jf200108d

Chapter 2: Traceability of high-quality lemon juices by multielement composition and strontium isotope ratio analyses

1 Introduction

Local food is an important resource for the territorial sustainability and economic welfare of local communities (Sims, 2009). It generates local identity as an interaction between natural resources and production system in a specific landscape (Arfini et al., 2010). In this contest the local food can generate benefits as creation of rural communities and implementation of local tourism (Sims, 2009). The Sorrento Peninsula is one of the most successful destinations for international tourism, for many reasons, such as unique landscape of tuff cliffs, gentle climate, citrus orchards and olive grove (Agovino et al., 2021). A special place in increasing the local tourism in the area is occupied by the lemon cultivations well-known all over the world for their quality. Currently, in Campania, two high-quality products have been certified and registered as Protected Geographical Indication (PGI) products, i.e. *Limone di Sorrento* (G.U.n°24, 2011) and *Limone Costa D'Amalfi* (G.U.n°178, 2001), both located in Sorrento Peninsula. These PGI products are specifically marketed for fresh consumption (intact lemon or squeezed lemon juice) or to produce typical liquor called "Limoncello". The lemon orchards follow a traditional cultivation system. They are cultivated on terraces and covered with "pergolato sorrentino", a system made from vertical wooden canes that covers the whole lemon grove with reeds known as "pagliarelle" (Giare & Giuca, 2008). Due to geomorphological conformation of Sorrento peninsula, the traditional agriculture, mostly on terraces, contributes to the soil stability and hydrogeological equilibrium of slopes (Lal, 2001). Furthermore, this type of territorial geographical conformation, can produce different microclimates and habitats along the slopes (Savo et al., 2016). These peculiar environmental features contribute not only to produce high-quality lemon fruits but also to protect the precious and fragile landscape. Apart from the higher economic costs associated with the laborious traditional production systems, the Sorrento and Amalfi PGI lemons also carry a story and tradition that is valued by consumers who are willing to pay a high price/premium for the product (Giare & Giuca, 2008). Given the difficulty of subjectively distinguish different valued lemons, it is vital to develop suitable techniques to detect frauds. Considering the strong link of this product with soil and climatic features the choice of authentication and traceability techniques naturally falls on geochemical tracers as the $^{87}Sr/^{86}Sr$ isotope ratio and multielement fingerprinting of lemons. The isotope ratio fingerprinting is based on the principle that the different isotopes of the same element have similar chemical properties, due to the same number of protons and electrons, but different physical properties, due to the different mass. Radiogenic isotopes, such as $^{87}Sr/^{86}Sr$, are invariant from biological fractionation and annual variation, during the plant uptake, compared to the isotopes of the light elements as C, H, N, O, and S ($^{2}H/^{1}H$, $^{13}C/^{12}C$, $^{15}N/^{14}N$, $^{18}O/^{16}O$, and $^{34}S/^{32}S$), because their measurements only record the variations due to a single factor: the radioactive decay of ^{87}Rb into ^{87}Sr (see methods). Given the very long time scale of this process, the $^{87}Sr/^{86}Sr$ ratio within rocks and soil minerals is constant in the environmental systems on time scales of $<10^4$ years (Nakano, 2016). Hence, the $^{87}Sr/^{86}Sr$ isotope ratio was successfully applied to discriminate different food products as a soil-derived traceability marker (Choi et al., 2008; Kawasaki et al., 2002; Petrini et al., 2015; Rummel et al., 2010; Vinciguerra et al., 2016; West et al., 2009). Moreover, several authors have shown that the $^{87}Sr/^{86}Sr$ ratio of plants closely reflects that of the soil solution in which they grow and therefore suggest measurement of the isotopic composition of Sr bioavailable fraction present in soil instead of total Sr (Brunner et al., 2010; Swoboda et al., 2008; Zannella et al., 2017).Some authors (Bong et al., 2012; Trincherini et al., 2014) reported that in rare cases, the use of only the $^{87}Sr/^{86}Sr$ isotope ratio could be not sufficient to discriminate different areas showing similar geological features. Thus a multidisciplinary approach, as the addition of multielement profile, can be useful for improving the discrimination power of Sr isotopic signature (Bong et al., 2012; Epova et al., 2018; García-Ruiz et al., 2007; Hiraoka et al., 2016; Zampella et al., 2011; Zannella et al., 2017). The multielement fingerprinting, as the Sr isotope signature, is considered a soil-derived geochemical tracer (Danezis et al., 2016). It is based on the principle that the mineral composition of food products reflects in a certain way the element arrays of cultivation soils (Anderson & Smith, 2005; Cheajesadagul et al., 2013; Gonzálvez et al., 2011; Kelly et al., 2002; Laursen et al., 2014; Simpkins et al., 2000). It is worth noting that the mineral composition of agri-food products is affected by several factors as agronomic management (fertilization, pest and weed control management), crop rotation, climate, in addition to natural soil fertility (Rosen & Allan, 2007). Moreover, in soil several factors can influence the bioavailability of elements such as pH, organic matter, texture, and chemical form of the elements (Kabata-Pendias, 2011). In a previous study (Chapter 1) we developed a chemometric discriminant model to track the juice of Sorrento PGI lemon on a small geographical scale (Campania regional scale) by multi-element fingerprinting. The results showed that the mineral profile of lemon juices was able to discriminate for the geographical origins despite the different cultivars, and with some elements (most of them not essential for plant) well correlating with cultivation soils. In this work we extend the previous model to a larger scale, adding samples gathered from local markets outside Campania (mainly from Calabria region) and in California, combining stable Sr isotope ratio and multi-element analysis of lemon juices and cultivation soils.

2. Materials and Methods

2.1. The study area and samples details

Eight lemon orchards were sampled in Campania, south Italy. Six orchards were located in the Limone di Sorrento and Limone Costa d'Amalfi PGI areas, respectively on the northern and southern slopes of the Sorrento Peninsula. The other two were located outside the PGI areas in the experimental fields of the Agricultural Sciences Department of University of Naples Federico II, in Portici (NA), and in the Regional experimental agricultural farm "Improsta", in Eboli (SA). In order to compare the Campanian lemon fruits with lemon of other origins (large scale), other four groups of fruits were taken from Italian and U.S. local markets. The Sorrento Peninsula is a headland located in south Italy that separates the Gulf of Naples to the north from the Gulf of Salerno to the south. Dominated by Lattari Mountains, the Sorrento Peninsula is mainly characterized by Mesozoic dolomite limestone, Triassic to Upper Cretaceous in age, with Quaternary pyroclastic deposits unevenly distributed over the area (Savo et al., 2014). These deposits are constituted by ash-flows and tuffs derived from the activity of the Phlegrean Fields, well preserved in the piedmont plain, and Somma-Vesuvius, preserved on mountain slopes (Apuzzo et al., 2013). The soils in the area mainly derive from the pyroclastic cover, characterized by the overlapping of different deposits among which pedogenic processes developed. Primary deposits prevail at the top of Mountains, whereas reworked sediments are essentially found at lower elevations, in concavities and in karstic depressions. Along the slopes, undifferentiated colluvium with volcanic and limestone fragments crops out, the soil profiles are frequently truncated or even totally removed, and due to the presence of low-grade crystallinity clays, the soils are characterized by thixotropic properties (Terribile et al., 2007). Traditionally, terraces are built on slopes to obtain arable land. The construction of these terraces modifies the landscape, the hydro-geological equilibrium and the vegetation patterns. Terraces are typically dried stone structures, sometimes, with minimum slopes, steps are made of earth and covered by grass. The experimental fields of Agricultural Sciences Department of Naples University are located in the Royal Park of Portici (NA) (latitude 40° 48.860' N, longitude 14° 20.851' E) at the south-western foothills of Mt. Vesuvius, between 20 and 90 m a.s.l. The soils of the Park, like those of all the Vesuvius area, derive from volcanic substrata such as lava flows, ash and lapilli, have moderate profile differentiation and moderately or weakly expressed Andic properties reflecting the activities of the primary vitreous materials (Di Gennaro, 2002; Di Gennaro & Terribile, 1999). The Regional experimental agricultural farm "Improsta", is located in the alluvial plain of Sele river. The Eboli area is characterized by a sequence of alluvial clastic deposits outcropping along the southern border of the Picentini Mountains called “conglomerati di Eboli” (Cinque et al., 1988). The soils in the area fall within the Cartographic Unit 3 of the Soil Map (scale 1: 50,000) of the right bank of the Sele river, realised by the SeSIRCA Department of Agriculture of Campania Region. UC3 has a considerable areal extension and is widespread mostly in the municipal area of Battipaglia and Eboli, downstream of the Tusciano river. It includes deep soils of areas with very modest slopes, between 1 and 3%, with rectilinear or slightly undulating morphology, formed on fine alluvium-colluvial deposits, stratified by channel, with reworked cinerite contributions. According to the USDA they are classified as Typic Haploxerepts or Pachic Haploxerolls. In all sampling sites (Sorrento peninsula, Portici, Eboli), considering the relatively small scale and the vicinity to the sea, the climate features are typically Mediterranean with a cold wet season (October to March) and a hot dry season (April to September). In total, eighty-eight mature lemon fruits (Citrus limon (L.) Burm.) were randomly collected (5-7 lemons were collected per tree) in 2019 from the eight orchard farms located in Campania region. The details on sampling sites and samples are given in table 1. In each orchard, around the trees sampled for lemons, three soil samples (about 500g each) were collected at the depth of 40-50 cm. The single soil samples taken from each orchard were then mixed to obtain composite samples reaching a total of 8 composite soil samples. It is worth noting that the samples bought in the local markets (USA and CAL) did not have specific indications of provenance on the labels, except for CAL2 referring to Citrus limon (L.) Burm. Femminello *comune Antico di Rocca Imperiale IGP* cultivated on the Ionic coast of Calabria region.

2.2. Determination soil elemental composition

The composite soil samples were air-dried, sieved at 2 mm, homogenized, and quartered in four subsamples. About 50 g of a subsample for each soil were pulverized in agata ball mill (PM 200, Retsch Gmbh, Haan, Germany). Soil powders were subjected to four different Sr-extraction procedures: for the determination of total Sr fraction, about 100 mg of sample was digested in a mixture of 2 ml of 38% Ultrapure HF + 1 ml 65% Ultrapur HNO_3, left for 7 days on a hot-plate at 80 °C, and daily subjected to ultrasonic bath to ensure complete dissolution of silicate fraction; for pseudototal Sr fraction, 200 mg of pulverized soils was digested in aqua regia (3 mL of 38% Ultrapure HCl and 1 mL of 65% Ultrapur HNO_3), put on a hot plate at 95 °C for 2 hours and subjected to ultrasonic bath each half hour. In addition, soil potentially and readily bioavailable fractions were determined by EDTA 0.05 M and 1 M NH_4NO_3 extraction following procedure described in Agrelli et al. (2017).

2.3. Determination of lemon juice elemental composition

Plant essential (Fe, Mn, Cu, Zn, Mo, Ca, K, Mg, including Na) and not-essential elements (Ti, Co, Rb, Sr and Ba) were determined in lemon juices. Five grams of juices were mineralized in a Teflon tube with 4 mL of HNO_3 65% ultrapure and 2 mL of H_2O_2 30 % using a microwave system (Mars 5 Express, CEM) following power ramp to 1600W for 5-5-10 min with a temperature of 100-150-200 °C, respectively. After cooling, the mixtures were completed to 50 mL final volume with ultrapure water and element were determined by inductively coupled plasma-mass spectrometer (ICP-MS, 7800 Agilent Technologies, Santa Clara, CA, USA). More details about ICP-MS operating conditions and lemon juice sample preparation are reported in (Chapter 1).

2.4. Determination of Sr isotopic ratio

Twenty-three juice mineralized solutions (about 20% of all samples) were used for Sr isotopic analysis. The samples for Sr isotope analyses were randomly chosen inside the different provenance areas in order to explain the major variability inside each area i.e. 8 samples for SR, 5 for AM, 2 for EB, 2 for PR, 4 for CAL and 2 for USA. In addition, for eight selected soil samples, Sr isotopic analyses were performed on total Sr, pseudototal Sr fraction, as well as EDTA and NH_4NO_3 extraction solution. The Sr isotopic analysis were performed using a ThermoFisher Neptune Plus MC-ICP-MS at the IGG-CNR of Pisa, in 2% HNO3 solution containing 20-200 $ng*g^{-1}$ of analyte after Sr extraction from matrix with Sr-spec resin in class 100 clean room. The instrument was equipped with a combined cyclonic and Scott-type quartz spray chamber, Ni-cones and a MicroFlow PFA 100 µl/min self-aspiring nebulizer. To eliminate the possible interferences, due to the presence of other elements, in particular Ca and Rb, the Sr was separated using a specific Sr-spec resin (Eichrom Sr). $^{87}Sr/^{86}Sr$ isotope ratios were corrected for mass bias fractionation using the $^{88}Sr/^{86}Sr$ ratio (= 8.375209), thus discarding also any effect due to mass dependent isotope fractionation eventually occurring in natural processes, and for mass interference using the ratios $^{83}Kr/^{84}Kr$ (= 0.201750), $^{83}Kr/^{86}Kr$ (= 0.664740) and $^{85}Rb/^{87}Rb$ (= 2.592310). The results of four repeated analyses of $^{87}Sr/^{86}Sr$ of reference material NIST SRM 987 carried out along with unknown samples was 0.710281 ± 4, samples were adjusted for a value of 0.710248.

2.5. Statistical analysis

The data analysis was carried out by SPSS 23.0 statistical software (SPSS Inc., Chicago, IL) and XLSTAT version 2016.02.284551 (Addinsoft, NY, USA). The Shapiro-Wilk test ($p < 0.05$) was applied to check the normal distribution of each variable. Multivariate statistic was used on mineral compositions of juices. Before the multivariate analysis, the whole dataset was variance scaled, mean-cantered, and normalized with log transformation. The Principal Component Analysis (PCA) was used for exploring the natural grouping of samples. The PCA is a bilinear decomposition technique capable of reducing large variables into few principal components (PCs), preserving data variance. The PCs score plot can be used to detect sample differences and similarities whereas the loadings plot (weight of each original variable on the PCs) is useful to distinguish the variables that are major contributors to a PC and that lead to the natural clustering of groups (data not labelled, unsupervised method) within the dataset (Li Vigni et al., 2013). Before PCA, to measure sampling adequacy and evaluate the suitability of the data matrix, the Kaiser-Meyer-Olkin test (KMO test), and Bartlett's test of sphericity were performed (Pellerano et al., 2008). Based on the results of PCA, the Stepwise Linear Discriminant Analysis (S-LDA) was carried out. LDA is a supervised (groups defined a priori) discriminant method that defines discriminant functions that are linear combinations of the original variables for optimizing the separation. The aim of LDA is to discriminate among groups by maximizing the variance between groups and minimizing the variance within each group (Gaiad et al., 2016). The results of LDA were evaluated based on Wilks' lambda. This value, close to zero indicates that the models can perfectly discriminate (Zhao et al., 2017). In addition, the application of Stepwise procedure selects only the best discriminant variables based on F-statistic factor (Rencher, 2003). Before the LDA, whole dataset was randomly divided in calibration set (75% of all samples) and external validation set (25%). The rightness of LDA was evaluated by the percentage of correct classification (accuracy), and the percentage of correct external validation. The juice $^{87}Sr/^{86}Sr$ isotope ratio matrix was combined whit the mineral profile of lemon juice by a low-level data fusion approach to investigate the contribution of each single analysis ($^{87}Sr/^{86}Sr$ and mineral profile) in discriminating the samples according to provenances. The data low-level data fusion approach consists to create a single matrix combining data from different analytical techniques (Borràs et al., 2015). In our case, the matrix of Sr isotope ratio was unsuitable to build a discrimination model because the sizes of different provenance class were unbalanced. So we only applied an explorative PCA analysis. The use of this approach allowed to put in evidence the natural clustering of samples and the contribution of each variable.

3. Results and Discussion

3.1. Strontium isotope ratio in soils and lemon juices

The $^{87}Sr/^{86}Sr$ isotopic ratios of total Sr (*TotSr*), pseudototal Sr (*PTSr*), EDTA extractable Sr (*EDTASr*), and NH_4NO_3 extractable Sr (*NH4Sr*) measured in soils and in lemon juices from the Campania region are given in Fig. 1. The isotopic ratio of *TotSr* in PR soil (0.70744) was the lowest among all orchards and within the interval of Campanian Plio-Quaternary volcanic rocks (0.7063-0.7080) reported by Peccerillo (2005). The highest ratio was found in EB (0.70861), likely attributed to the heterogeneity of alluvial clastic sediments probably of silicate origin (Nisi, 1998) of Sele plain sediments. In Sorrento Peninsula, the values of $^{87}Sr/^{86}Sr$ isotopic ratio of *TotSr* ranging between 0.70756 and 0.70822 are in good agreement with the values reported by Peccerillo (2005) for Campanian Plio-Quaternary volcanic rocks of Somma Vesuvius (0.7063-0.7080) and Phlegraean Fields (0.7065-0.7086). Considering the geological substrate of Sorrento peninsula mainly constituted by Dolomite carbonates of different age with $^{87}Sr/^{86}Sr$ values ranging between 0.707385 and 0.7760 (0.70738-0.70739 for Amalfi side and 0.707471-0.70776 for Sorrento side) as recorded by Iannace et al. (2011) our results denote an overall predominance of volcanic materials deposited on the carbonate substrate (Calcaterra & Santo, 2004), in the $^{87}Sr/^{86}Sr$ isotopic ratios in the soils of Sorrento peninsula. The $^{87}Sr/^{86}Sr$ values for *PTSr* were in all cases lower than for *TotSr*. This can be explained with the occurrence of a high ^{87}Sr component of the silicates fraction to *TotSr* (e.g. the aforementioned clastic sediments). The $^{87}Sr/^{86}Sr$ isotope ratio recorded in lemon juices, with the only exception of AM1, was always closer to that of the bioavailable soil fractions (*EDTASr* and *NH4Sr*) than to the total fractions (*TotSr* and *PTSr*). Indeed, in most cases, the *NH4Sr* values of soils were very close or, when considering the error bars, overlapping those of juices. This confirms once more that the Sr isotope ratio of the mobile Sr forms in soil is a better predictor of the $^{87}Sr/^{86}Sr$ ratio of lemon juices compared with the ratio of the soil total and pseudototal Sr (Brunner et al., 2010; Swoboda et al., 2008; Zannella et al., 2017). The only outliers are represented by juices AM2 and SR2, whose high Sr isotope ratio can be explained by the proximity of both orchards to the sea and thus by the contribution of marine spray (87Sr/86Sr ratio of 0.709202 ± 0.000003; Kuznetsov et al., 2012) to juice ratio. The distribution of $^{87}Sr/^{86}Sr$ isotope ratios in lemon juices is reported in Fig. 2. Apart from these two outliers, the Campanian samples had ratios significantly lower than samples from other origins (CAL and USA). However, inside the Campanian group, there was a general overlapping among samples with $^{87}Sr/^{86}Sr$ values recorded in samples from EB (0.70768-0.70769) being significatively lower than those recorded in SR area (0.70776-0.70829) and PR (0.70785-

0.70792), but slightly overlapping the values recorded in AM area (0.70769-0.70850). This observation could denote the large influence of carbonate deposits in both EB and AM, due to the occurrence in the two areas of alluvial clastic deposits (ISPRA Foglio 486, 2015) and Mesozoic dolomite limestone (ISPRA Fogli 466-485, 2015), respectively. The lemon samples originating from Calabria region had $^{87}Sr/^{86}Sr$ ratio ranging between 0.70864 and 0.70891 in agreement with values found by Rummel et al. (2010) in orange juices (on average $^{87}Sr/^{86}Sr$: 0.7090) and attributed to Calabria Cenozoic sediments consisting of a mixture of Mesozoic sediments and crystalline rocks. The high variability observed between the two lemon samples originating from California ($^{87}Sr/^{86}Sr$ isotope ratio of 0.70902 for USA1 and 0.70831 for USA2), might be related to very heterogeneous geological and lithological conditions of South California state (San Joaquin Valley, Fresno, Kern, Tulare, Ventura, and Riverside) (Hall, 2007) where citrus fruit is mainly produced (Geisseler & Horwath, 2016). All above results confirm the Sr isotopic analysis as a powerful analytical tool to discriminate the geographical provenance of lemon juices, especially at large territorial scale. It is worth noting that for samples originating from not too far away areas (i.e. AM, SR and PR) the discrimination was not so clear due to the similarity in geological and soil features. Although the Sr isotope ratio in juices reflected not only that present in rocks but also in soils, the use of mineral profile could give more information on samples discrimination, not only for a small territorial scale, already demonstrated only for Limone di Sorrento PGI (Chapter 1) but also on a large scale, due to the possible differences in the pool of elements in soils (different concentration, presence-absence). In fact, in the early stages of pedogenic processes, the concentration of elements in soils are close to that of parental materials; with time, the element content in soil will change due to pedogenic and anthropogenic processes (Kabata-Pendias, 2011). Although, in the same way, also the $^{87}Sr/^{86}Sr$ isotope ratio can be influenced by different factors as precipitation, soil parent material, and surface or groundwater (Capo et al., 1998), in soils located in a small-scale with similar geological feature and climatic condition, the variation of soil elemental composition is likely to be higher than the variation of $^{87}Sr/^{86}Sr$ isotope ratio.

3.2. Mineral composition of lemon juices

The contents (mean ± standard deviation) of essential (K, Mg, Ca, Na, Zn, Fe, Cu, Mn, Mo) and not-essential (Rb, Sr, Ba, Ti, Co) elements in lemon juices is shown in table 2. An exploratory analysis by PCA on the whole dataset (14 elements and 89 lemon juices) was carried out for screening the natural grouping of samples. The KMO test, giving a result of 0.671, greater than recommended value of 0.500 (Potortì et al., 2020), and the Bartlett sphericity test, showing p-value < 0.0001, indicated the adequacy of the matrix to apply the PCA (Li et al., 2018). PCA showed that the first six extract components, eigenvalues >1 (Kasier, 1960), explained 81.5% of the total variance. The score plot of the first two PCs (Fig.3a), explaining the 48.4% of the cumulate variance (33.5% and 14.9%), showed a general overlapping among the different groups according to provenance, although a slight differentiation was shown among Campanian (AM, SR, PR, EB), Californian (USA1 and USA2) and Calabrian samples (CAL1 and CAL2) on PC1 (pink and black labels, respectively). However, all Campania samples seem to be overlapped and only a slight trend of grouping was detected. In particular, the samples of PR and EB were differentiated on PC2, and both PGI areas (SR and AM) seem to be close to each other. According to the loadings, for PC1 they were Sr (16.2%), Co (16.1%), Mn (15.6%), Ca (12.8%), Mg (11.9%) and Ba (11.1%) and for PC2: Na (33.4%), K (23.9%) and Rb (22.1%). Two others explorative PCAs were carried, to explore the contribution of essential and not-essential elements for discrimination of origins, on the same samples. The KMO test was 0.517 and 0.658 for essential and not-essential elements, respectively, and the p-value of Bartlett sphericity in both cases was < 0.0001. For essential elements, five components (83.3% of total variance) were extracted. The score plots of the first two PCs (Fig.3b), explained the 49.3% of the total variance (29.7% and 19.7%) and showed, in the same way for the dataset with all elements, a general overlapping among the different sites of sampling, with a very slight differentiation among Campania, Calabrian and Californian samples. On the contrary, inside the Campanian groups, it seems to be no grouping between the two PGI areas (orange and red labels), but the differentiation between PR and EB seems to be clear. For not-essential elements, three components, with 84.2% of the total variance, were extracted. The score plots of the first two PCs (Fig.3c) explained 66.18% of the cumulative variance (45.2% and 21.0%), and showed a clear grouping, compared with the other two datasets, also for the groups of the Campania region with only a slight overlapping each other. It is worth noting that the two groups from the Calabria region, the two Californian groups and the two Campanian PGI groups, were clearly separated. These results suggested that the contribution of not essential elements was predominant in the discrimination of lemon juices samples, in particular way at small territorial scale, with the possibility to build supervised discriminant models (all elements, essential and not-essential elements) dividing the data set according to provenance (AM, SR, PR, EB, CAL, and USA).

3.3. Discriminant analyses

Based on the results of different exploratory analyses, three different S-LDA on 75% of the samples (66 lemon juice samples) were built according to geographical provenance. The 25% remaining samples (23 lemon juice samples) were used for external validation. The calibration set was divided in AM (15), SR (17), PR (10), EB (9), CAL (9), USA (6). In all cases, the S-LDA models showed significant differences between groups with Wilks' values near zero, namely 0.001, 0.012, 0.020 for all, essential and not-essential elements, respectively. The model based on all elements gave the 98.45% of accuracy and the 90.9% of external samples correctly validated. The stepwise procedure selected the 78.5% of analysed elements as best discriminants. Among them there were all the analysed not-essential elements (Sr, Rb, Ba, Ti, and Co) and only 66.6% of the essential elements (K, Ca, Mg, Fe, Mn, and Mo). The model based on the essential elements gave the 86.3% of correct classification and the 72.7% of the samples correctly validated. The stepwise procedure selected 77.7% of all considered elements (K, Mg, Na, Cu, Zn, Mn, Mo). The model based on not-essential elements gave the 92.4 % of samples correctly classified with 86.3% of correct validation. In this case, the stepwise procedure selected all not-essential elements as discriminants (Sr, Rb, Ba, Ti, and Co). Comparing the best discriminants selected in all models, it is worth noting that among essential elements only variables as K, Mg, Mn, and Mo were in common with the all-element model, with Fe and Ca selected only for the all-element model and Na, Cu, and Zn only for the essential-element model, while all not-essential elements were selected in both cases (the all-element and the not-essential element models). All the above results suggest that the use of the all-element profile was the best way to discriminate the lemon juices of different geographical origin. However, the reliability of the not-essential element model, expressed as the ratio between the percentages of external validation and classification (Chapter 1), was more similar to that of the all-element model (0.92 and 0.93, respectively) than that of the essential element model (0.84). This would indicate that the contribution of not-essential elements was higher than that of essential elements in the all-element profiles for discrimination according to geographical origins. However, in the all-element profile, mainly depending on composition of cultivation soil, the contents of the not-essential elements could be more effective than the essential elements, because their contents in soils could be more affected by cultivation management, primarily through their addition by fertilization practices. It is worth noting that the uptake from the soil of not-essential elements as Ba, Sr, Rb is more linear than that of competitive essential elements, (Ba and Sr for Ca and Rb for K) (El-Sheikh & Ulrich, 1970; Jackman, 1965; Rediske & Selders, 1953), due to their low amount and consequently bioavailability in soil (Kabata-Pendias, 2011). In addition, in the juices, the role of all cations can be more related to osmotic regulation than to biochemical roles compared to what happens in plant tissues. Thus, the not-essential element composition of juices should be more affected by the root uptake than the essential elements, also affected by translocation from other plant organs.

3.4. Fusion of ^{87}Sr/^{86}Sr isotope ratio and mineral profile of lemon juices

Based on the results of the discriminative models on lemon juice mineral profile and ^{86}Sr/^{87}Sr isotope signature, two exploratory data analysis using PCA were carried out considering only the twenty-three samples analysed for the Sr isotope ratio. The first PCA was applied on the combined matrix (^{86}Sr/^{87}Sr + multielement profile), the second on the elemental profile matrix alone. Before the PCA, the KMO test was applied, giving 0.517 for the combined matrix and 0.657 for the elemental profile matrix. In both cases, the Bartlett sphericity test showed a p-value < 0.0001, and the first 4 PCs were extracted (eigenvalues >1) giving 74.59% and 74.46% of the cumulative variance for the combined and elemental profile matrix, respectively. The biplot (Fig.4a) of the first two PCs (55.55% of explained variance) of the lemon juice multielement profile showed a natural grouping similar to the one seen for the entire dataset (Fig.3a) with a slight differentiation among the samples from Campania, Calabria, and California. Also, Zn (13.2%), Sr (13.08%), Cu (10.8%), Mn (10.7%), Co (9.7%) and Ti (8.0%) showed high contribution on PC1 and K (26.8%), Na (21.16%), Mo (14.3%) Co (9.4%) and Rb (8.1%) on PC2. For the combined matrix, the biplot (Fig.4a) of the first two PCs, explaining the 54.6% of the variance, showed a clear clustering on the large territorial scale on PC2, with the highest contribution of 86Sr/87Sr (15.0%), and a slightly better differentiation of AM-PGI on PC1 compared with the biplot considering only the elemental profile, with the highest contribution of Sr (13.6%) and Co (10.33%). Comparing the two biplots (Fig.4), the data fusion approach showed a clearer grouping of samples compared to the mineral profile alone, especially on a large territorial scale.

4. Conclusions

The multi analytical approach applied in this paper based on $^{87}Sr/^{86}Sr$ isotopic signature and mineral composition of lemon juice was able to discriminate the samples of different geographical origin either at small and large territorial scale. The strontium isotope ratio of lemon juices was more closely related to that of the bioavailable fractions of soil and shown differences among samples from different provenances better at a large than at a small territorial scale. It is widely accepted that the Sr isotopic ratio of agri-products reflects the geological characteristics of the product provenience sites. Our results clearly show that the $^{87}Sr/^{86}Sr$ isotope ratio in lemon juices was closely related to the cultivation soils having values of $^{87}Sr/^{86}Sr$ ratio slightly different from that of the underlying geological substrates (mainly limestone), and this behavior was likely due to the occurrence in the study area of foreign contributions such as volcanic depositions. The results of exploratory analysis applied on all, only essential, and only not-essential element composition of lemon juices showed a potential ability to discriminate among samples of different origin. The S-LDA model based on the all-element profile gave the 98.45% of accuracy with the 90.9% of external samples correctly validated, indicating as best discriminant variables all not-essential elements (Sr, Rb, Ba, Ti, and Co) and only 66.6% of the analysed essential elements (K, Ca, Mg, Fe, Mn, and Mo). Thus, applying the same model on two groups of variables (essential and not-essential elements), we showed that the contribution of not-essential elements was more reliable than that of essential elements more affected by cultivation management. The combination of Sr isotope signature and multielement profile by PCA explorative analysis (data fusion approach) produced promising results in terms of discrimination of lemon juice geographical origin. Nevertheless, in this study, due to statistical limits (unbalanced groups), it was not possible to apply to the combined matrix, supervised discriminant models as LDA. This might be done in future increasing the number of samples analysed for Sr isotope ratio. The validity of the multi analytical approach proposed in this study deserves future research involving replication across growing seasons and a larger sample size to encompass the natural pedoclimatic variability of the study area.

5. References

Agovino, M., Musella, G., Pisano, S., & Scaletti, A. (2021). Assessing the impact of tourist flows on emergency department treatment speed for residents and tourists. The case of Sorrento. *Socio-Economic Planning Sciences*, *August 2020*, 101004. https://doi.org/10.1016/j.seps.2020.101004

Agrelli, D., Adamo, P., Cirillo, T., Duri, L. G., Duro, I., Fasano, E., Ottaiano, L., Ruggiero, L., Scognamiglio, G., & Fagnano, M. (2017). Soil versus plant as indicators of agroecosystem pollution by potentially toxic elements. *Journal of Plant Nutrition and Soil Science*, *180*(6), 705–719. https://doi.org/10.1002/jpln.201700050

Anderson, K. A., & Smith, B. W. (2005). Use of chemical profiling to differentiate geographic growing origin of raw pistachios. *Journal of Agricultural and Food Chemistry*, *53*(2), 410–418. https://doi.org/10.1021/jf048907u

Apuzzo, D., De Vita, P., Palma, B., & Calcaterra, D. (2013). Approaches for mapping susceptibility to rockfalls initiation in carbonate rock-masses : a case study from the Sorrento coast (southern Italy). *Itallian Juornal of Geosciences*, *132*(3), 380–393. https://doi.org/10.3301/IJG.2013.12

Arfini, F., Vandecandelaere, E., Belletti, G., & Marescotti, A. (2010). Linking people, places and products. A guide for promoting quality linked to geographical origin and sustainable geographical indications. In *Food and Agriculture Organization of the United Nations (FAO) and SINER-GI*. www.foodquality-origin.org/guide/guide.pdf

Bong, Y. S., Shin, W. J., Gautam, M. K., Jeong, Y. J., Lee, A. R., Jang, C. S., Lim, Y. P., Chung, G. S., & Lee, K. S. (2012). Determining the geographical origin of Chinese cabbages using multielement composition and strontium isotope ratio analyses. *Food Chemistry*, *135*(4), 2666–2674. https://doi.org/10.1016/j.foodchem.2012.07.045

Borràs, E., Ferré, J., Boqué, R., Mestres, M., Aceña, L., & Busto, O. (2015). Data fusion methodologies for food and beverage authentication and quality assessment - A review. *Analytica Chimica Acta*, *891*, 1–14. https://doi.org/10.1016/j.aca.2015.04.042

Brunner, M., Katona, R., Stefánka, Z., & Prohaska, T. (2010). Determination of the geographical origin of processed spice using multielement and isotopic pattern on the example of Szegedi paprika. *European Food Research and Technology*, *231*(4), 623–634. https://doi.org/10.1007/s00217-010-1314-7

Calcaterra, D., & Santo, A. (2004). The January 10, 1997 Pozzano landslide, Sorrento Peninsula, Italy. *Engineering Geology*, *75*(2), 181–200. https://doi.org/10.1016/j.enggeo.2004.05.009

Capo, R. C., Stewart, B. W., & Chadwick, O. A. (1998). Strontium isotopes as tracers of earth surface processes: theory and methods. *Geoderma*, *82*, 197–225.

Cheajesadagul, P., Arnaudguilhem, C., Shiowatana, J., Siripinyanond, A., & Szpunar, J. (2013). Discrimination of geographical origin of rice based on multi-element fingerprinting by high resolution inductively coupled plasma mass spectrometry. *Food Chemistry*, *141*(4), 3504–3509. https://doi.org/10.1016/j.foodchem.2013.06.060

Choi, S., Lee, H., Lee, G., & Han, J. (2008). *Food Chemistry Determination of the strontium isotope ratio by ICP-MS ginseng as a tracer of regional origin*. *108*, 1149–1154. https://doi.org/10.1016/j.foodchem.2007.11.079

Cinque, A., Guida, F., Russo, F., & Santangelo, N. (1988). Dati cronologici e stratigrafici su alcuni depositi continentali della Piana del Sele (Campania): i “conglomerati di Eboli’’.” *Geografia Fisica e Dinamica Quaternaria*, *11*(1), 39–44.

Danezis, G. P., Tsagkaris, A. S., Camin, F., Brusic, V., & Georgiou, C. A. (2016). Food authentication: Techniques, trends & emerging approaches. *TrAC - Trends in Analytical Chemistry*, *85*, 123–132. https://doi.org/10.1016/j.trac.2016.02.026

Di Gennaro, A. (2002). *I sistemi di terre della Campania.Carta 1:250.000 e Legenda.*

Di Gennaro, A., & Terribile, F. (1999). *I suoli della provincia di Napoli. Carta 1:75.000 e Legenda.*

El-Sheikh, A. M., & Ulrich, A. (1970). Interactions of Rubidium, Sodium, and Potassium on the Nutrition of Sugar Beet Plants. *Plant Physiology*, *46*, 645–649. https://doi.org/10.1104/pp.46.5.645

Epova, E. N., Bérail, S., Zuliani, T., Malherbe, J., Sarthou, L., Valiente, M., & Donard, O. F. X. (2018). 87Sr/86Sr isotope ratio and multielemental signatures as indicators of origin of European cured hams: The role of salt. *Food Chemistry*, *246*(August 2017), 313–322. https://doi.org/10.1016/j.foodchem.2017.10.143

G.U.n°178. (2001). *Disciplinare di produzione dell ’ Indicazione Geografica Protetta " Limone Costa d ’ Amalfi ”.*

G.U.n°24. (2011). *Disciplinare di produzione dell'Indicazione Geografica Protetta "Limone di Sorrento."* https://ec.europa.eu/agriculture/quality/door/documentDisplay.html?chkDocument=1063_1_it

Gaiad, J. E., Hidalgo, M. J., Villafañe, R. N., Marchevsky, E. J., & Pellerano, R. G. (2016). Tracing the geographical origin of Argentinean lemon juices based on trace element profiles using advanced chemometric techniques. *Microchemical Journal*, *129*, 243–248. https://doi.org/10.1016/j.microc.2016.07.002

García-Ruiz, S., Moldovan, M., Fortunato, G., Wunderli, S., & García Alonso, J. I. (2007). Evaluation of strontium isotope abundance ratios in combination with multi-elemental analysis as a possible tool to study the geographical origin of ciders. *Analytica Chimica Acta*, *590*(1), 55–66. https://doi.org/10.1016/j.aca.2007.03.016

Geisseler, D., & Horwath, W. R. (2016). *Citrus Production in California*. https://apps1.cdfa.ca.gov/FertilizerResearch/docs/Citrus_Production_CA.pdf

Giare, F., & Giuca, S. (2008). Cultura, tradizioni e qualità degli agrumi. *Istituto Nazionale Di Economia Agraria*, *Quaderno n*, 135.

Gonzálvez, A., Armenta, S., & De La Guardia, M. (2011). Geographical traceability of "arrs de Valencia" rice grain based on mineral element composition. *Food Chemistry*, *126*(3), 1254–1260. https://doi.org/10.1016/j.foodchem.2010.11.032

Hall, C. A. (2007). *Introduction to the Geology of Southern california and Its native plants*. University of California Press.

Hiraoka, H., Morita, S., Izawa, A., Aoyama, K., Shin, K. C., & Nakano, T. (2016). Tracing the geographical origin of onions by strontium isotope ratio and strontium content. *Analytical Sciences*, *32*(7), 781–788. https://doi.org/10.2116/analsci.32.781

Iannace, A., Capuano, M., & Galluccio, L. (2011). "Dolomites and dolomites" in Mesozoic platform carbonates of the Southern Apennines: Geometric distribution, petrography and geochemistry. *Palaeogeography, Palaeoclimatology, Palaeoecology*, *310*(3–4), 324–339. https://doi.org/10.1016/j.palaeo.2011.07.025

ISPRA Fogli 466-485. (2015). *NOTE ILLUSTRATIVE della CARTA GEOLOGICA D'ITALIA alla scala 1:50.000 Sorrento-Termini. Fogli 466-485.*

https://www.isprambiente.gov.it/Media/carg/466_485_SORRENTO_TERMINI/Foglio.html

ISPRA Foglio 486. (2015). *NOTE ILLUSTRATIVE della CARTA GEOLOGICA D'ITALIA alla scala 1:50.000 Foce del Sele. Foglio 486*. https://www.isprambiente.gov.it/Media/carg/486_FOCE_DELSELE/Foglio.html

Jackman, R. H. (1965). The Uptake of Rubidium by the Roots of Some Graminaceous and Leguminous Plants. *New Zealand Journal of Agricultural Research*, *8*(4), 763–777. https://doi.org/10.1080/00288233.1965.10423712

Kabata-Pendias, A. (2011). Trace elements in soils and plants. In *CRC Press*. https://doi.org/10.1201/b10158-25

Kasier, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, *XX*(1), 141–151.

Kawasaki, A., Oda, H., & Hirata, T. (2002). Determination of strontium isotope ratio of brown rice for estimating its provenance. *Soil Science and Plant Nutrition*, *48*(5), 635–640. https://doi.org/10.1080/00380768.2002.10409251

Kelly, S., Baxter, M., Chapman, S., Rhodes, C., Dennis, J., & Brereton, P. (2002). The application of isotopic and elemental analysis to determine the geographical origin of premium long grain rice. *European Food Research and Technology*, *214*(1), 72–78. https://doi.org/10.1007/s002170100400

Kuznetsov, A. B., Semikhatov, M. A., & Gorokhov, I. M. (2012). The Sr isotope composition of the world ocean, marginal and inland seas: Implications for the Sr isotope stratigraphy. *Stratigraphy and Geological Correlation*, *20*(6), 501–515. https://doi.org/10.1134/S0869593812060044

Lal, R. (2001). Managing world soils for food security and environmental quality. *Advances in Agronomy*, *74*(1), 155–192. https://doi.org/10.1016/s0065-2113(01)74033-3

Laursen, K. H., Schjoerring, J. K., Kelly, S. D., & Husted, S. (2014). Authentication of organically grown plants - advantages and limitations of atomic spectroscopy for multi-element and stable isotope analysis. *TrAC - Trends in Analytical Chemistry*. https://doi.org/10.1016/j.trac.2014.04.008

Li, L., Wen, B., Zhang, X., Zhao, Y., Duan, Y., Song, X., Ren, S., Wang, Y., Fang, W., & Zhu, X.

(2018). Geographical origin traceability of tea based on multi-element spatial distribution and the relationship with soil in district scale. *Food Control*, *90*, 18–28. https://doi.org/10.1016/j.foodcont.2018.02.031

Li Vigni, M., Durante, C., & Cocchi, M. (2013). Exploratory Data Analysis. In *Data Handling in Science and Technology* (1st ed., Vol. 28). Copyright © 2013 Elsevier B.V. All rights reserved. https://doi.org/10.1016/B978-0-444-59528-7.00003-X

Nakano, T. (2016). Potential uses of stable isotope ratios of Sr, Nd, and Pb in geological materials for environmental studies. *Proceedings of the Japan Academy, Series B*, *92*(6), 167–184. https://doi.org/10.2183/pjab.92.167

Nisi, B. (1998). Geochimica isotopica delle acque del bacino dell'Arno. In *Indagine geochimica ed isotopica del carico disciolto nelle acque di scorrimento superficiale della valle dell'Arno: valutazione del contributo naturale ed antropico.* (pp. 52–58).

Peccerillo, A. (2005). *Plio-Quaternary Volcanism in Italy Petrology, Geochemistry, Geodynamics*. Springer Berlin Heidelberg.

Pellerano, R. G., Mazza, S. S., Marigliano, R. A., & Marchevsky, E. J. (2008). Multielement analysis of argentinean lemon juices by instrumental neutronic activation analysis and their classification according to geographical origin. *Journal of Agricultural and Food Chemistry*, *56*(13), 5222–5225. https://doi.org/10.1021/jf073555n

Petrini, R., Sansone, L., Slejko, F. F., Buccianti, A., Marcuzzo, P., & Tomasi, D. (2015). The 87Sr/86Sr strontium isotopic systematics applied to Glera vineyards: A tracer for the geographical origin of the Prosecco. *Food Chemistry*, *170*, 138–144. https://doi.org/10.1016/j.foodchem.2014.08.051

Potortì, A. G., Bua, G. D., Lo Turco, V., Ben Tekaya, A., Beltifa, A., Ben Mansour, H., Dugo, G., & Di Bella, G. (2020). Major, minor and trace element concentrations in spices and aromatic herbs from Sicily (Italy) and Mahdia (Tunisia) by ICP-MS and multivariate analysis. *Food Chemistry*, *313*(April 2019), 126094. https://doi.org/10.1016/j.foodchem.2019.126094

Rediske, J. H., & Selders, A. A. (1953). The absorption and translocation of strontium by plants. *Plant Physiology*, *28*(4), 594–605.

Rencher, A. C. (2003). Methods of Multivariate Analysis. In *Methods of Multivariate Analysis*. John

Wiley & Sons, Inc., Hoboken, New Jersey. Published. https://doi.org/10.1002/0471271357

Rosen, C. J., & Allan, D. L. (2007). *Exploring the Benefits of Organic Nutrient Sources for Crop Production and Soil Quality*. *17*(December).

Rummel, S., Hoelzl, S., Horn, P., Rossmann, A., & Schlicht, C. (2010). The combination of stable isotope abundance ratios of H, C, N and S with 87 Sr/ 86 Sr for geographical origin assignment of orange juices. *Food Chemistry*, *118*(4), 890–900. https://doi.org/10.1016/j.foodchem.2008.05.115

Savo, V., Caneva, G., McClatchey, W., Reedy, D., & Salvati, L. (2014). Combining environmental factors and agriculturalists' observations of environmental changes in the traditional terrace system of the Amalfi Coast (Southern Italy). *Ambio*, *43*(3), 297–310. https://doi.org/10.1007/s13280-013-0433-3

Savo, V., Salvati, L., & Caneva, G. (2016). In-between soil erosion and sustainable land management: climate aridity and vegetation in a traditional agro-forest system (Costiera Amalfitana, southern Italy). *International Journal of Sustainable Development and World Ecology*, *23*(5), 423–432. https://doi.org/10.1080/13504509.2015.1132282

Simpkins, W. A., Louie, H., Wu, M., Harrison, M., & Goldberg, D. (2000). Trace elements in Australian orange juice and other products. *Food Chemistry*, *71*(4), 423–433. https://doi.org/10.1016/S0308-8146(00)00150-3

Sims, R. (2009). Food, place and authenticity: Local food and the sustainable tourism experience. *Journal of Sustainable Tourism*, *17*(3), 321–336. https://doi.org/10.1080/09669580802359293

Swoboda, S., Brunner, M., Boulyga, S. F., Galler, P., Horacek, M., & Prohaska, T. (2008). Identification of Marchfeld asparagus using Sr isotope ratio measurements by MC-ICP-MS. *Analytical and Bioanalytical Chemistry*, *390*(2), 487–494. https://doi.org/10.1007/s00216-007-1582-7

Terribile, F., Basile, A., De Mascellis, R., Iamarino, M., Magliulo, P., Pepe, S., & Vingiani, S. (2007). Landslide processes and Andosols: the case study of the Campania region, Italy. In *Soils of Volcanic Regions in Europe 2007* (pp. 545–563). Springer-Verlag.

Trincherini, P. R., Baffi, C., Barbero, P., Pizzoglio, E., & Spalla, S. (2014). Precise determination of strontium isotope ratios by TIMS to authenticate tomato geographical origin. *Food Chemistry*, *145*, 349–355. https://doi.org/10.1016/j.foodchem.2013.08.030

Vinciguerra, V., Stevenson, R., Pedneault, K., Poirier, A., Hélie, J. F., & Widory, D. (2016). Strontium isotope characterization of wines from Quebec, Canada. *Food Chemistry*, *210*, 121–128. https://doi.org/10.1016/j.foodchem.2016.04.017

West, J. B., Hurley, J. M., Dudás, F. Ö., & Ehleringer, J. R. (2009). The stable isotope ratios of marijuana. II. strontium isotopes relate to geographic origin. *Journal of Forensic Sciences*, *54*(6), 1261–1269. https://doi.org/10.1111/j.1556-4029.2009.01171.x

Zampella, M., Quétel, C. R., Paredes, E., Asfaha, D. G., Vingiani, S., & Adamo, P. (2011). Soil properties, strontium isotopic signatures and multi-element profiles to authenticate the origin of vegetables from small-scale regions: Illustration with early potatoes from southern Italy. *Rapid Communications in Mass Spectrometry*, *25*(19), 2721–2731. https://doi.org/10.1002/rcm.5081

Zannella, C., Carucci, F., Aversano, R., Prohaska, T., Vingiani, S., Carputo, D., & Adamo, P. (2017). Genetic and geochemical signatures to prevent frauds and counterfeit of high-quality asparagus and pistachio. *Food Chemistry*, *237*, 545–552. https://doi.org/10.1016/j.foodchem.2017.05.158

Zhao, H., Yu, C., & Li, M. (2017). Effects of geographical origin, variety, season and their interactions on minerals in tea for traceability. *Journal of Food Composition and Analysis*, *63*(July), 15–20. https://doi.org/10.1016/j.jfca.2017.07.030

Table 1. Details of samples and sampling sites. AM-PGI, *Limone Costa D'Amalfi* PGI area; *Limone di Sorrento* PGI area (SR-PGI), local markets in Calabria region, Italy (CAL-NOPGI) and local markets in Berkeley, California, U.S. (USA-NOPGI)

Labels	*Locations*	*Citrus limon (L.) Burm cv.*	*n° fruits*
AM-PGI			
AM1	40°40'27"N,14°38'53"E	Femminello *Sfusato Amalfitano*	6
AM2	40°39'32"N,14°38'48"E	Femminello *Sfusato Amalfitano*	7
AM3	40°38'32"N,14°39'31"E	Femminello *Sfusato Amalfitano*	6
SR-PGI			
SR1	40°36'28"N;14°20'36"E	Femminello *Ovale di Sorrento*	8
SR2	40°37'51"N;14°24'22"E	Femminello *Ovale di Sorrento*	10
SR3	40°36'59"N;14°23'10"E	Femminello *Ovale di Sorrento*	4
		Femminello *Sfusato Amalfitano*	3
PR-NOPGI	40°48'54"N;14°20'49"E	Femminello *Ovale di Sorrento*	6
		Femminello *Sfusato Amalfitano*	6
EB-NOPGI	40°33'21"N;14°58'53"E	Femminello *Ovale di Sorrento*	6
		Femminello *Sfusato Amalfitano*	5
CAL-NOPGI			
CAL1	local market, Italy	Femminello *Siracusano* 2KR	6
CAL2	local market, Italy	Femminello C*omune Antico di Rocca Imperiale IGP*	5
USA-NOPGI			
USA1	local market, U.S.	Lisbon	4
USA2	local market, U.S.	Eureka	6

Table 2. Mineral composition (mean ± standard deviation) of lemon juices according to provenance; AM1, AM2 and AM3 from the "Amalfi PGI area"; S1, S2 and S3 from the "Sorrento PGI area"; PR-NOPGI from the germplasm field, Portici (NA); EB-NOPGI from germplasm field, Eboli (SA); CAL1 and CAL2 from Italian markets in Calabria region (CAL-NOPGI); USA1 and USA2 from local markets in Berkeley, CA, U.S. n: total number of samples.

Elements	**AM-PGI *(n°19)***			**SR-PGI *(n°25)***			**EB-NOPGI *(n°11)***	**PR-NOPGI *(n°12)***	**CAL-NOPGI *(n°11)***		**USA-NOPGI *(n°10)***	
	AM1 *(n°6)*	**AM2 *(n°7)***	**AM3 *(n°6)***	**S1 *(n°8)***	**S2 *(n°10)***	**S3 *(n°7)***			**CAL1 *(n°6)***	**CAL2 *(n°5)***	**USA1 *(n°4)***	**USA2 *(n°6)***
mg kg^{-1}												
K	1381 ± *207*	1352 ± *111*	1125 ± *83*	1420 ± *222*	1304 ± *136*	1214 ± *113*	1482 ± *99*	1314 ± *121*	1219 ± *78*	1086 ± *118*	1098 ± *67*	1002 ± *41*
Mg	56.4 ± *5.0*	62.1 ± *5.0*	40.8 ± *6.0*	59.8 ± *8.2*	60.1 ± *5.4*	55.7 ± *8.1*	67.3 ± *5.5*	60.4 ± *3.3*	57.9 ± *7.0*	63.4 ± *6.7*	74.4 ± *5.8*	99.5 ± *3.8*
Ca	35.0 ± *9.0*	44.8 ± *10.4*	33.6 ± *5.2*	53.8 ± *12.0*	32.7 ± *7.8*	36.4 ± *13.4*	51.7 ± *10.8*	34.8 ± *5.0*	39 ± *6.6*	66.5 ± *7.5*	72.9 ± *20.6*	74.1 ± *6.9*
Na	5.9 ± *2.0*	7.5 ± *3.0*	5.8 ± *1.9*	7.1 ± *3.3*	4.7 ± *1.2*	3.5 ± *1.1*	8.8 ± *8.3*	3.0 ± *0.6*	2.5 ± *0.4*	4.5 ± *1.4*	4.8 ± *1.6*	2.9 ± *0.9*
µg kg^{-1}												
Zn	353.7 ± *32.8*	363.2 ± *80.4*	346.7 ± *119.6*	423.2 ± *71.4*	476.2 ± *123.5*	369.8 ± *59.2*	413.0 ± *207.5*	481.6 ± *144.5*	431.8 ± *83.9*	437.4 ± *114.4*	590.1 ± *251.3*	572.0 ± *164.9*
Cu	219.7 ± *46.8*	219.9 ± *27.9*	278.3 ± *40.7*	260.6 ± *52.5*	208.4 ± *51.3*	243.1 ± *34.0*	291.6 ± *109.1*	205.6 ± *21.9*	315.3 ± *79.2*	358.2 ± *116.5*	234.1 ± *31.3*	341 ± *83.1*
Fe	277.5 ± *75.2*	371 ± *105.7*	539.1 ± *290.8*	345.3 ± *83.1*	235.7 ± *123.8*	249.9 ± *88.5*	346.1 ± *149.2*	257.3 ± *44.8*	347.9 ± *306.5*	181.3 ± *147.3*	255.8 ± *85.9*	344.2 ± *111.9*
Mn	49.2 ± *7.6*	41.8 ± *17.5*	72.2 ± *19.3*	95.4 ± *17.3*	84 ± *17.9*	97.7 ± *26.7*	155.7 ± *78.2*	88.3 ± *13.9*	111.8 ± *29.1*	130.7 ± *36.2*	130.7 ± *43.5*	227.3 ± *36.7*
Mo	5.0 ± *2.3*	2.8 ± *0.7*	6.3 ± *2.1*	4.6 ± *1.4*	5.3 ± *3.4*	5.8 ± *1.9*	3.3 ± *0.9*	9.8 ± *2.7*	1.7 ± *0.7*	2.8 ± *1.3*	8 ± *2.5*	8.6 ± *2.5*
Rb	1145 ± *231*	574 ± *79*	853 ± *128*	1358 ± *354*	960 ± *135*	893 ± *79*	2035 ± *684*	792 ± *103*	366 ± *64*	1026 ± *175*	1504 ± *548*	871 ± *340*
Sr	30.1 ± *7.8*	58.7 ± *16.6*	64.2 ± *20.9*	206.6 ± *60.8*	122 ± *36.7*	173.5 ± *54.5*	123.0 ± *54.9*	89.8 ± *13.3*	132.5 ± *37.2*	315.7 ± *76.4*	465.7 ± *211.0*	474.3 ± *112.5*
Ba	71.0 ± *15.5*	41.7 ± *9.6*	31.7 ± *12.5*	52.7 ± *14.2*	44.4 ± *14.8*	73.6 ± *23.0*	87.9 ± *27.9*	24.7 ± *21.7*	40.1 ± *14.4*	61.0± *16.4*	40.5 ± *23.4*	902.8 ± *241.2*
Ti	10.9 ± *1.0*	15.6 ± *4.1*	17.5 ± *13.3*	4.8 ± *1.4*	6.9 ± *7.9*	2.7 ± *0.7*	3.2 ± *1.7*	7.9 ± *5.9*	6.0 ± *2.4*	3.9 ± *1.4*	13.6 ± *5.4*	11.4 ± *4.1*
Co	0.2 ± *0.1*	0.4 ± *0.1*	0.4 ± *0.2*	0.6 ± *0.2*	0.4 ± *0.2*	0.8 ± *0.3*	1.3 ± *0.9*	0.7 ± *0.2*	1.4 ± *0.3*	2.5 ± *1.0*	2.4 ± *0.4*	3.2 ± *0.9*

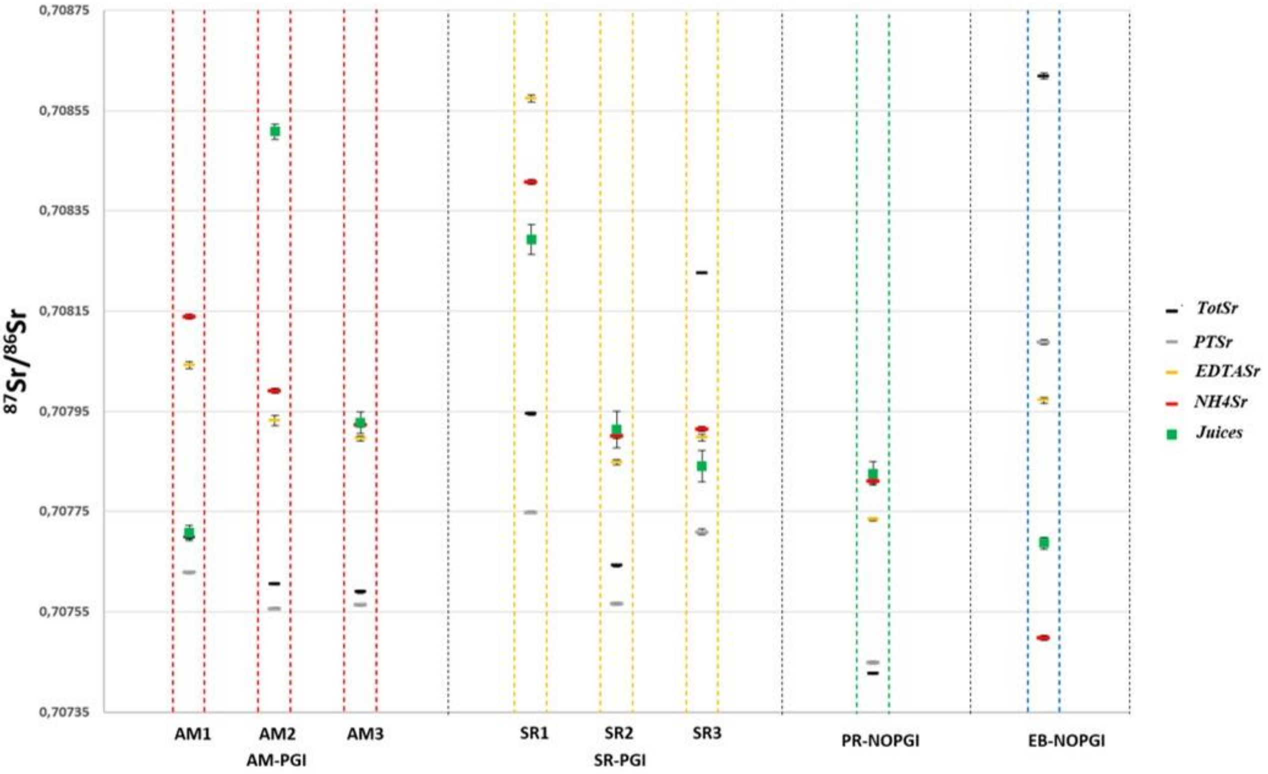


Figure 4 The distribution of 87Sr/86Sr isotopic ratios of total Sr (*TotSr*), pseudototal Sr (*PTSr*), EDTA extractable Sr (*EDTASr*), and NH4NO3 extractable Sr (*NH4Sr*) in soils and in juices from Campania

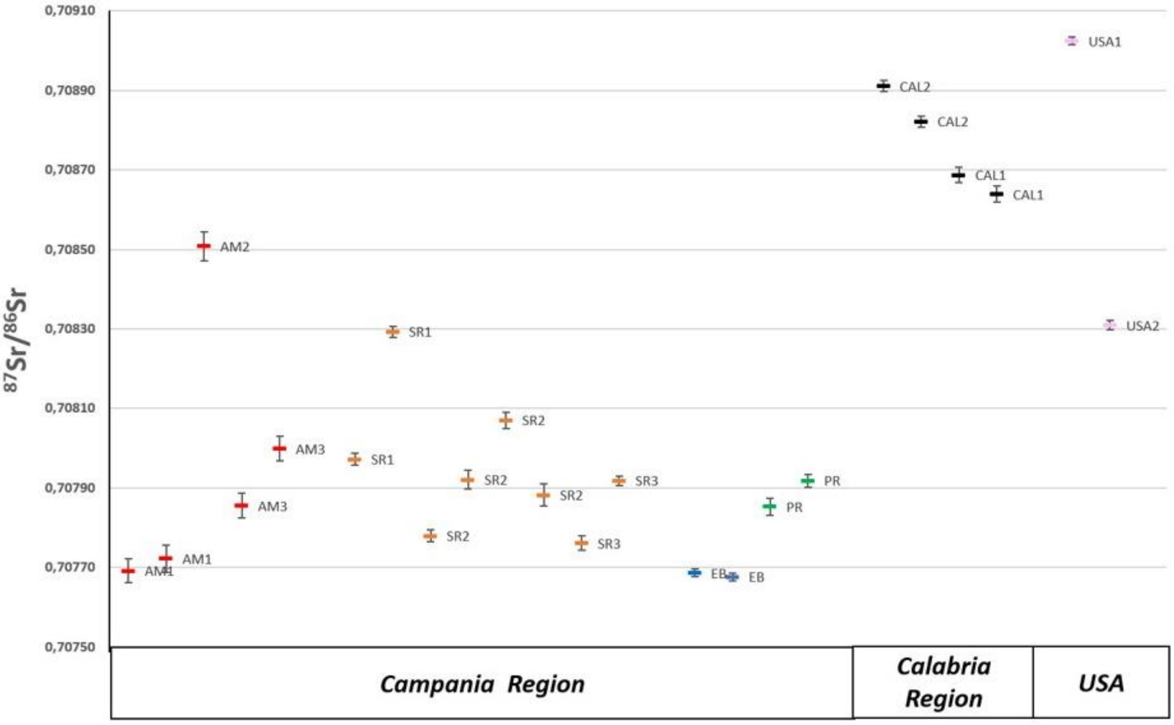


Figure 5 The distributions of 87Sr/86Sr isotopic ratio in lemon juice according to provenance.

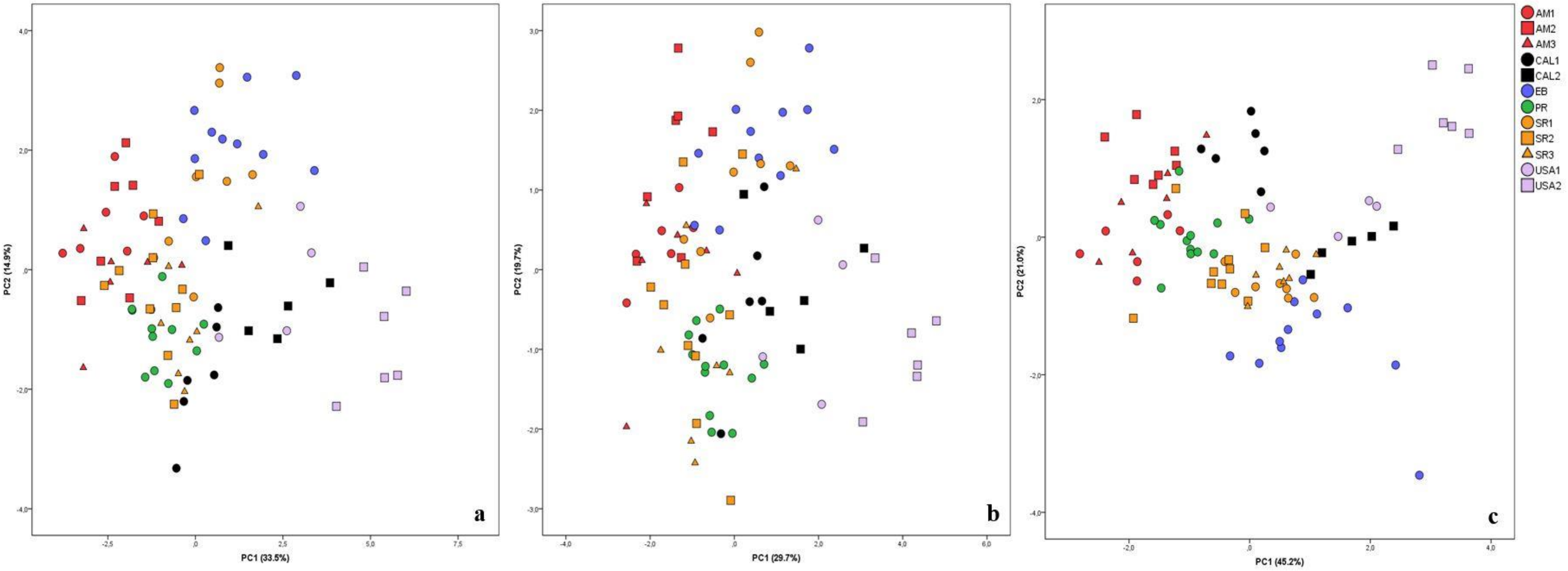


Figure 6 Score plot of the exploratory Principal Component Analysis of lemon juices labelled according to origin: (a) all elements; (b) essential elements; (c) not-essential elements.

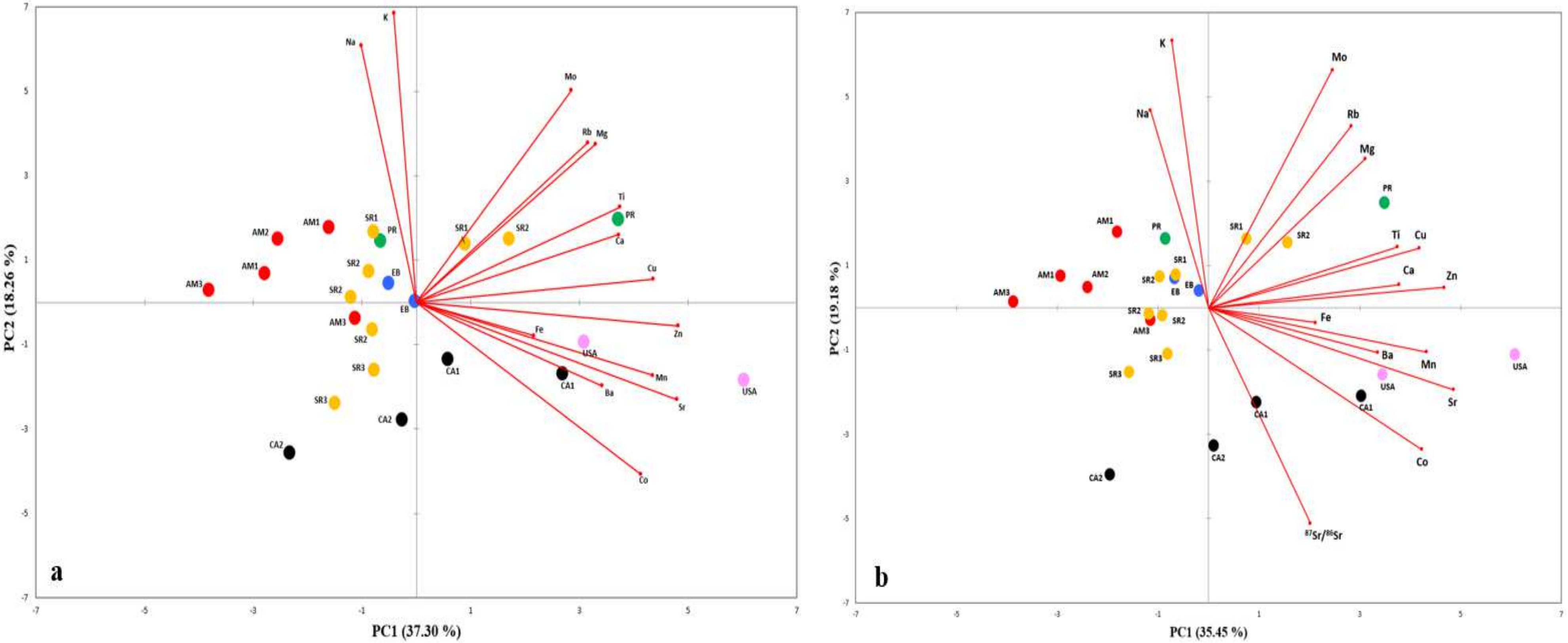


Figure 7 Explorative PCA biplots. Biplot of lemon juice multielement profile (a); biplot of multielement profile and 87Sr/86Sr isotope ratio in lemon juice

Chapter 3: Potential use of Near-Infrared Spectroscopy on intact lemon fruits for authentication and traceability evaluated by chemometrics

1 Introduction

Lemon (*Citrus limon* L. Burm.), belong to the family Rutaceae, is one of the most worldwide cultivated citrus species mainly in tropical and subtropical areas (Sun et al. 2019; Shafqat et al. 2021). The lemon fruits are rich in nutrients, micronutrients, vitamins (as vitamin C), acids (mainly citric acid) and essential oils (Di Matteo et al., 2021; Sun et al., 2019). The quality of citrus fruits depends by different factors among which plant nutrition. About nutrient effects, nitrogen increased availability adversely affect fruit-keeping quality, while potassium is considered indispensable for good quality production and yields although its excess increases acid content (Ladaniya, 2010). Also the availability of water has significant impact on organoleptic proprieties, its shortage can increase total soluble solid and titratable acidity (García-Tejero et al., 2010). Globally the production of lemons, together limes (Citrus aurantiifolia), are around 23 MLN of tons per year with a harvest area of 1.4 MLN of ha. The European production represents 6% worldwide production concentrated for 98% in the Mediterranean basin, where 51% refers to a European country (FAOSTAT, 2021). The geomorphological and climatic characteristics of the entire Mediterranean coastal area are suitable to produce excellent lemons which are prevalently devoted to fresh consumption (whole lemon and lemon juices), unlike those from other countries, such as Brazil and the United States, mostly used to produce concentrated juices (Giare and Giuca 2008). In this area the Italian lemon production, the second in Europe after Spain (Clever & Valverde, 2019), is concentrated mostly in the southern regions (Sicily, Calabria, Puglia, and Campania). Six lemon fruit from Italy were awarded Protected Geographical Indication (PGI) labels since 1992, in agreement with the European policies for the protection of geographical origin of high-quality products (EEC No 2080/92, 1992; EEC No 2081/92, 1992). Among them, two arise from Sorrento peninsula in Campania region, the Limone di Sorrento IGP (cv. Femminello Ovale di Sorrento) from the northern slopes of the peninsula and Capri Island, and the Limone Costa d'Amalfi IGP (cv. Femminello Sfusato Amalfitano) from the southern slope. These PGI cultivations occupy around 260 hectares, mostly on terraces, with 550 producers registered (ISTAT, 2017), and play an important role for the production, but overall for territory stability guaranteed by traditional management, and tourism attractiveness due to suggestive landscapes (Savo et al., 2016). Ovale di Sorrento and, Sfusato Amalfitano, as all those widely cultivated in Italy of Feminello variety, produce fruit during the whole year usually called, in relation to the pickup period. Primofiore (September-November), Limone, properly (December-May), Bianchetto (April-June, and Verdello (May-July) (Ladaniya, 2010). Such high value lemon productions need protection and enhancement forms, and these can be supported by the proposal of methods and protocols applied on lemon as such to evaluate product quality and provenance, they can be also useful to avoid frauds with lemons from other source. For this purpose, instrumental methods as Near Infrared Spectroscopy (NIRS) could be useful. It is a wide used technique for quality control in all production line of food industry. It is characterised by cheapness and applicability on materials as such (Jha et al., 2014; Luykx & van Ruth, 2008). NIR spectroscopy has been successfully applied for the determination of qualitative properties of different fruits (Xie et al., 2011; Jha et al., 2014; Saranwong et al., 2004; Ribeiro et al., 2011; Elfadl et al., 2010; Kahriman et al., 2011; Li et al., 2021; Wang et al., 2013; Amodio et al., 2017; Carlini et al., 2000; Gabioud et al., 2008; Birth et al., 1985), among which citrus fruits (Ana M. Cavaco et al., 2018; Arendse et al., 2017; Gòmez et al., 2006; Kawano et al., 1993; Miyamoto & Kitano, 2009; K. J. Lee et al., 2004), and, specifically lemon fruits (Reddy et al., 2016). In addition, it has been applied for the determination of the geographical origins of several agri-food products (Latorre et al., 2013; Woodcock et al., 2009; Woo et al., 2005; Zhao et al., 2013; Cozzolino et al., 2011; Iu et al., 2006; Jiménez-carvelo et al., 2019; Hirri et al., 2016; Guo et al., 2018; Ren et al., 2013; Eisenstecken et al., 2019; Amendola et al., 2020; Monteiro et al., 2018), and also of citrus fruits (Amenta et al., 2016; Fabroni et al., 2015). NIR spectroscopy investigates materials in the wavelengths region between 780 nm and 2500 nm (wavenumber between 12820 and 4000 cm-1), and the typical spectra obtained is characterised by absorption bands due to the overtone vibrational and rotational transitions of molecular bonds (Nicolaï et al., 2007). The NIR spectra of entire fruits are generally acquired in diffuse reflectance more appropriate for opaque and thick samples for which IR ray transmission is very difficult, However, the fruit probing depth of light scattered in diffuse reflectance measurements remain questionable (A. M. Cavaco et al., 2021). The NIR radiation decreases exponentially with the depth into fruit tissues (Arendse et al., 2018) and, especially for thick rind fruit such as citrus, pulp could be poorly explored by ray. Generally NIR spectra show low resolution and for sample of the same nature differences are difficult to appreciate (Polesello et al., 1981). Thus, to compare and obtain information from NIR spectra, their treatment by multivariate statistic can be very effective (Blanco & Villarroya, 2002). Such chemometric analysis often require a pre-processing of spectra, typically, scatter correction, namely Multiplicative Scatter Correction (MSC) and Standard Normal Variate (SNV), and derivate correction, first and second derivate as Norris-Williams derivatives and Savitzky-Golay (SG) polynomial derivates (Rinnan et al., 2009). The aim of this work was to investigate the potential application of NIR spectroscopy on intact lemons for fruit quality and geographical origin assessment. For this purpose, lemons were collected in two production years and in seven different orchards of the Campania PGI areas of Limone di Sorrento and Limone Costa d'Amalfi cultivations, and in two sites outside the PGI area but in the same region. Fruits were analysed for quality properties and their NIR spectra were acquired. The NIR spectroscopy reliability for quantitative analysis and traceability purposes was evaluated by multivariate statistical analyses.

2 Materials and Methods

2.1 Characteristics of sampling sites and soil analyses

In Campania region (Italy) five orchards were chosen in the PGI area of Limone Costa d'Amalfi, on the southern slope of Sorrento peninsula overlooks the gulf of Salerno, and in PGI area of Limone di Sorrento, on northern slope of the peninsula overlook the gulf of Naples (Table 1). Geologically the Sorrento peninsula is essentially entirely constituted of Mesozoic carbonates covered by discontinuous layers (from less than 1 m to 2 m) of pyroclastic deposits mainly related to the air-fall products of the 79 AD Vesuvius eruption (Iannace et al., 2011; Picarelli et al., 2008). In both PGI areas the lemon cultivations are on terraces with "pagliarelle" coverings. Two further experimental orchards were chosen in Campania region, in Portici at the centre of Naples Gulf, that is characterised by typical volcanic soils being at the foot of Mt.Veuvius, and in Eboli, that is characterised by sedimentary soils being in the plain of Sele river flowing into Salerno gulf (Table 1). Further site characteristics are reported in Ruggiero et al., (2021). The climate of all sites is typically Mediterranean with dry summers and mild winter, the mean annual rainfall in a triennium 2016-18 are reported in Table 1. The temperature is available only for Portici and Eboli (13-22 °C on overage 2017-2019), they should not be so different for the Sorrento peninsula given the proximity to the mentioned sites.Soil in all orchards were sampled at the 40-50 cm depth and the sieved soil fractions < 2 were analysed for pH (H2O), Cation Exchange Capacity (C.E.C), total Carbonates (tot. Carb.), Organic Carbon (O.C.), bioavailable K, Mg, Ca and Na, and soil texture (sand, silt, and clay) as reported in Ruggiero et al., (2021). On 2 mm sieved soils, also total nitrogen was determined by elemental analyser (Elementar UNICUBE, Germany) after sample pulverization in agate ball mill, and available phosphorus by colorimetric method (MiPAF, 2000); both determinations were carried out in triplicate.

2.2 Lemon sampling and lemon quality properties

One hundred nineteen full ripened lemon fruits, *Citrus limon* (L.) Burm. Femminello, were collected from seven different orchards in March and first two weeks of April of 2018 and 2019 production years (Table 1). At least one lemon for tree from no less than three trees for orchard was randomly collected. In particular, in 2019, 19 fruits of *cv* Sfusato Amalfitano were collected from the three orchards in Limone Costa d'Amalfi PGI area and 11 fruits from the two germplasm fields in Portici and Eboli; 32 fruits of *cv*. Ovale di Sorrento were collected from the two orchards in the Limone di Sorrento PGI area, and 12 fruits from the germplasm fields in Portici and Eboli. In 2018, 11 fruits of *cv* Sfusato Amalfitano were collected from the germplasm fields in Portici and Eboli; 22 fruits of *cv*. Ovale di Sorrento were collected from the two orchards in Limone di Sorrento PGI area, and 12 fruits from the germplasm fields in Portici and Eboli. All lemons were collected with a mean equatorial diameter of 66 mm (CV 32%) with the lowest diameter of 51 mm, thus above the minimum diameter (45 mm) for commercialization (EU 428/2019, 2019). Lemon fruits were washed with tap-water, then with deionized water, dried with a paper towel, weighted, and stored at 4 °C for no more than one week till analyses. Fruit chromaticity was determined by Minolta Chroma Meter CR-400 (Minolta Corp, Japan) in CIELAB colour space on four opposite spots along the equatorial circumference of lemon and L*, a* and b* mediated. The values L*, a* and b* describe a three-dimensional colour space, where L* corresponds to lightness/darkness (0 black, 100 white), a* represents the redness/greenness (-100 green, +100 red) and b* corresponds to yellowness/blueness (-100 blue, +100 yellow) (Roy Choudhury, 2015). After, NIR spectrum were acquired as described below, then lemon fruits were cut with a plastic knife along equatorial circumference and the pericarp thickness (THK) measured by handheld calliper in two opposite spots of a lemon section and mean thickness calculated. The lemon sections were squeezed with plastic juicer and the juice centrifugated at 2000 *g* for 20 min, to separate it from suspended solid tissues. The residue pericarp was cut in raw pieces and on a fraction of them the water content (PW) was determined by lyophilization. The centrifuged juice volume (ml) was measured to calculate juice yield expressed in percent. Juices were analysed in duplicates for total soluble solid content (SSC) by digital refractometer (HI 96814, Hanna Instruments, RI, USA), titratable acidity (TA), expressed as citric acid, by 0.1M NaOH titration, and pH by pH-meter (pH 211 Hanna Instruments, RI, USA). The mineral composition (K, Ca, Na and Mg) of lemon juices was determined as reported in Ruggiero et al., (2021) and the molar concentration of positive charge from the cations (C^+) was calculated.

2.3 Acquisition of NIR spectra

NIR spectra were acquired by FT-IR/NIR spectrometer (Frontier, PerkinElmer, CT, USA) equipped with a reflectance accessory (NIRA, PerkinElmer, CT, USA), over the range 10000 cm^{-1} to 4000 cm^{-1} (1000 nm to 2500 nm), with a resolution of 4 cm^{-1} and 64 scans for each spectrum (sample and background). Before analysis, the fruits were brought at room temperature of 20 ± 1 °C and then placed in direct contact with detector window. For each lemon, along its equatorial circumference, eight NIRA spectra were acquired each at about one eighth of circumference to mediate inhomogeneity of exocarp and of scattering effects (Fig. S1 Appendix A supplementary material). In addition, on a lemon fruit, triplicate spectra, replacing on detector the sample each time, were acquired on a marked spot on the fruit equatorial circumference, on the total peel (flavedo + albedo) around the spot removed by thin blade, and separately on the flavedo and albedo obtained dividing transversally the removed total peel. The spectra of total peel and the flavedo and albedo parts were acquired taking care to hold their original curvature. Triplicate spectra of the juice (2 ml) from the same lemon were acquired into a glass petri dish (30 mm diameter and 10 mm depth) with trasfectance adaptor (Fig. S1 Appendix A supplementary material). These last spectra of a lemon fruit were pre-processed with 1st derivate (Savitzky-Golay, 2nd order polynomial, 15 points)

2.4 Statistical analysis

Unscrambler X10.4 (Camo Software AS., Norway) and XLSTAT version 2016.02.284551 (Addinsoft, NY, USA) were used for statistical analyses. A mean spectrum, calculated by PerkinElmer Spectrum software (Version 10.03.06, 2011, CT, USA), deriving from all eight spectra acquired for each lemon, was used for multivariate analysis by Principal Component Analysis (PCA). Before the application of PCA, the NIR spectra were mean centered. Multiple Linear Regression (MLR) were applied to investigating the relationship between spectra and quality properties of lemons by using all the extracted PCs from the PCA for reducing the high number of variables of NIR spectra (Cowe & McNicol, 1985; Eisenstecken et al., 2019). The performances of MLR models were evaluated by significance ($\alpha = 0.05$) of regression coefficient (R^2) and coefficient of variation (CV) calculated as a percentage of ratio Root Mean Square Error (RMSE) and mean of each quality property. The normal distribution of each lemon quality property was checked by the Shapiro-Wilk test ($p \leq 0.05$) and non-parametric Kruskal–Wallis and Dunn rank tests (Dunn, 1964) were performed for $p \leq 0.05$. Spearman correlation analysis (rank-based correlation) was carried out on lemon quality properties to investigate the relationship among them. The Stepwise Linear Discriminant Analysis (S-LDA) was applied on quality properties according to provenances. For application of Linear Discriminant Analysis (LDA) models, two spectra for sample (each as a mean of four spectra of the eight, acquired at about 90° among them respect equatorial axis), were used to include also some spectral variability of each lemon sample. Varius LDA models were built to discriminate grouping for cultivars, production years, and provenances. Each different group of lemon spectra was randomly divided in calibration set (not less than 70% of samples) and external validation set (not less than 20% of samples). The LDA models were evaluated based on accuracy, external validation accuracy, correct classification and validation percent. The ratios of accuracies (external validation/calibration) were used to compare the model performances (Ruggiero et al., 2021). In general no spectrum pre-processing techniques (Rinnan et al., 2009) were used to optimize the various multivariate analyses, this was a choice for an homogeneous comparison of statistical results from spectroscopical data as such.

3 Results and Discussion

3.1 Management, soil, and climate of orchards

All Limone Costa d' Amalfi PGI and Limone di Sorrento PGI orchards were managed in agreement to lemon production specifications, also not PGI orchards essentially observed such specifications but without cover on lemon trees (Table 1) (G.U.n°178, 2001; G.U.n°24, 2011). All orchards did not show in-homogeneity. The main properties of orchard soils are reported in Table S1 (provided in Appendix A. Supplementary materials). All soils had a good level of available nutrients (N, K, Ca, Mg) except for phosphorus (lowest value for PR and EB), the pH ranged from alkaline to subalkaline (7.2-8.1 in A1 and PR, respectively). Total carbonate content ranged from 3.6 g/kg and 252 g/kg with values higher in the orchards of Limone Costa D'Amalfi PGI area than in the other farms, similar trend was observed also for O.C. and C.E.C that were on average 27.4 g/kg and 29.7 cmol(+) kg–1 , respectively in the above area respect the other orchards, in which they were no more than 11.7 g/kg (S2) and 23.8 cmol(+) kg–1 (EB). The soils had sandy loam texture in orchards of Limone di Sorrento PGI area, loamy sand in orchards of Limone Costa d'Amalfi PGI area, sandy in PR orchard and clay loam in EB orchard. The rainfall was very similar in all orchards areas given the proximity of the sites (Fig. S2a provided in Appendix A. Supplementary material). In the period of development of collected lemon fruits, i.e., from fruit setting, approximately in June, to colour break, approximately in January, in all considered orchard areas, a lower rainfall was recorded for 2018 production year (618 mm averaging on all areas), than for 2019 (724 mm). In the same period the average temperature was quite similar for both production years, but there were monthly differences. In the period between

June to August of 2018 production, the temperatures were higher than of 2019 with a mean difference of less than 2 °C; an opposite trend was observed until December with even smaller mean difference; again, a higher temperature (of about 5 °C) was recorded in January for 2018 production than for 2019 (Fig. S2b provided in Appendix A. Supplementary material). In the maturation period, about from February to the first two weeks of April, same differences were recorded between production year especially for raining that was higher for 2018 production (on overage 416 mm) than for 2019 (on average 149 mm), while for temperature the mean differences in the period were negligible. It is worth noting that the climate trend of this last period should be less incisive on the ripening characteristics of lemons, at most it can influence the moment of reaching full maturation, than the previous period during fruit developing. In addition, climate characteristics in spring, that were different in the two production years (concerning more the rainfalls than the temperatures) may influence mainly flowering and fruit setting amount, thus essentially the fruit yields. (Spiegel-Roy & Goldschmidt, 1996).

3.2 Lemon quality properties

The lemon quality properties with univariate statistic results are reported in Table 2. The juice yields were in the range of 24% and 46% (34% on average. For all cultivars the highest values, often statistically significant, were observed in 2018 production year (42% in 2018 vs 29% in 2019 on average), and this behaviour was markedly evident in the comparison of annuity for each orchard with biennial harvesting. The values of pericarp thickness (THK) followed an opposite trend (3.6 mm in 2018, 6.9 mm in 2019 on average), with no substantial THK differences between cultivars in each year. This suggests that in 2019 a higher weight contribute of more thickness pericarp, also showing often tendentially higher water content (PW) than in 2018, to the entire lemon weights was the cause of juice lower yields in 2019. The brightness, L*, of all lemons showed little variation among all cultivars and orchards, as well as between production years for each orchard. The colour parameters a*, related to green colour component, and b*, related to yellow component, which increase with maturation, were widely variable among all cultivars and orchards, but frequently showed higher values, both or one of them, in 2018 than in 2019 for each orchard with biennial harvesting such as for OS-S1, OS-S2, OS–EB, and SA-EB. The values of a* and b* only showed an opposite trend for SA-PR or not significant difference for OS-PR. The SSC of all lemon juices from all cultivars and orchards (7.6 °Brix on overage) varied between 6.6 °Brix, for SA-A1-2019 and 8.7 °Brix, for OS-EB-2018, and pH (2.3 on overage) between 2.2, for SA-PR, and 2.6, for OS-S2, both in 2019 with many similarities among orchards and cultivars. By the comparison of production years for each orchard, the SSC were similar except for SA-PR, while pH showed occasional difference. The TA of all juices (57 g/L on average) showed lower variability among cultivars and orchards than the before mentioned properties, but frequently significant or trendily values higher in 2018 (74.2 g/L on average) than in 2019 (53.5 g/L on average) appeared, it was evident by the comparison by production years for each orchard. Further the TA in 2019 for SA-A1, SA-A2 and SA-A3 were below the level indicated in the production specifications for Limone Costa d'Amalfi PGI (≥3.5%) (G.U.n°178, 2001). The cation charge concentrations (C^+) in juices, that followed the trend of TA, showed some significant differences, more evident in the comparison between production years for each orchard, suggesting a higher influx of cations in lemon endocarp for pH homeostasis in 2018 than in 2019. However, except for the few cases of low juice yields and TA, all the lemon properties agreed with the required properties of the PGI production specifications of Limone di Sorrento PGI and Limone Costa d'Amalfi PGI. In general, the observed annual variations of quality properties appeared homogeneous among cultivars and orchards. Such variations could be attributed mainly to climatic differences between the two production years, in particular regarding rainfall during lemon developing period (section 3.1) (Iglesias et al., 2007; Spiegel-Roy & Goldschmidt, 1996), rather than eventual soil and agronomical practices differences, also taking into account the similarity of climate affecting orchards falling in restricted territory. During the development of fruits collected in 2019, the higher water availability and consequently also that of nutrients, such as N and K, could have caused the higher thickness of pericarp and the lower a* and b* than in 2018 (Fhatuwani 2001; Lee and Chapman 1988; Goldschmidt, 1988). In the production of 2018, the higher TA could depend by water shortage respect to 2019 as observed for citrus subject to summer water stress (Levy et al., 1978; Spiegel-Roy & Goldschmidt, 1996). Further, the higher TA and salifying cations in 2018 production were likely needed agreed to rise osmotic potential in fruits, which was probably mainly controlled by salt concentration rather than by SSC that remained similar in both the production years. The temperatures could also have influenced the maturation characteristics, such as TA, since gene expression related to citrate metabolism seems to be thermo regulated (Lin et al., 2016). However, the difference of temperatures between the two production years during the developing of fruit were little and with swinging trend. Further temperature trend in the production of 2019 was very similar to those of 2017 in PR area (data not shown) which yet showed the quality properties of fruits (Di Matteo et al., 2021) more similar to those of 2018 confirming a scarce importance of the yearly variations of temperatures. Instead, the rainfall data in PR area for production in 2017 (data not shown) were more similar during summer to those in 2018 than in 2019, supporting the influence of water availability on the observed differences of lemon properties. It is worth noting that for the productions of 2018 and 2019 the rainfall and some temperature annual differences from spring flowering until fruit set phases, as well as during ripening period, from breaking colour and full ripening, although mainly affecting production levels, could have same consequence on the water and nutrient needs during fruit development, and full ripe characteristics. Also the winter climate in preceding years could have influenced successive fruit yields by flowering control and consequent plant needing for fruit maturation (Spiegel-Roy & Goldschmidt, 1996). In any case, a care was spent to pick lemons that appeared at full maturation level in the two-production years, on the other hand they were collected. In relative not so shorth period in both years. Thus, their different ripening characteristics could mainly be attributed to annual environmental variations.

3.3 NIR spectroscopy of intact lemons

The epicarp of citrus, the outer part of fruit pericarp is made by flavedo, and externally epidermis and hypodermis (henceforth simply called flavedo). The mesocarp, the inner part of pericarp, is made by albedo and inner hypodermis (henceforth simply called albedo). Epicarp is composed of several cell layers that become progressively thicker in the internal part and in epidermic layer also few stomata are present. Among the many types of epicarp cells there are oil-gland cells, containing essential oils, and chromoplasts. At maturation carotenoids (mostly xanthophylls) accumulate and chlorophylls essentially disappear in the chromoplasts. This hormonally controlled process determines the change of colour from green to yellow of fruits upon maturation. In addition to essential oils (mostly terpenoids) and pigments (carotenoids, chlorophylls), the pericarp contains paraffin waxes on surface, steroids, fatty acids, flavonoids, and enzymes, but it is mostly composed of cellulosic material and pectic substances. The albedo is essentially constituted of cells containing abundant sugar within the vacuoles organized to form a spongy tissue rich in pectin and is separated by fruit endocarp by thick-walled cells of inner epidermis covered by a thin cuticle (Dimopoulou et al., 2019; Scott & Baker, 1947; Spiegel-Roy & Goldschmidt, 1996). The NIR reflectance spectra of intact lemon (Fig.1a) were mainly representative of spectroscopic flavedo characteristics and less affected by albedo and still less by endocarp tissues as suggested by PCA score plot of spectra of lemon parts which showed closer groupings of intact lemon, flavedo and entire peel (4.5 mm) respect to the spectra of albedo and of juice which were markedly separated (Fig.1b). The dominating broad absorption bands of lemon spectra (Fig.1a) were at 5173 cm^{-1} and 6876 cm^{-1}, the strongest, and at 5580 cm^{-1} attributable mainly to the 2nd and the 1st overtone and the combination absorptions, respectively, of free and trapped water, in agreement with the high PW in entire peel (about 80%). The absorption of alcoholic OH and C=O, mainly due to polysaccharides (cellulose, pectin, etc.), also contributed to same of those wavenumbers. The shoulder at higher wavenumbers of the band at 6876 cm^{-1} suggests an overlapping of a peak at about 7300 cm^{-1} attributable to 1st overtone combination of CH, CH_2 and CH_3 groups. The band at 8332 cm^{-1} could be attributable to 2nd overtone of CH linkages arising from CH_2 and CH groups, mainly in polysaccharides as well as in oils, waxes, etc. The shoulder to the previous band at 8583 cm^{-1} could be due to 2nd overtone band of CH_3 groups mainly arising from methyl in alkyl chains and methoxy pectin groups. The PCA, performed on the raw spectra of all intact lemons of 2018 and 2019 productions, extracted seven components and the 98% of total variance was explained by the first two components (eigenvalues >1; PC1 90% and PC2 8%). The PCA score plot showed a tendency to separate lemons for production year (Fig. 2), the PR, S2 and S1 groups mainly separated along PC2 for, being the score distribution for the most part in the III and IV quadrants for 2018, and in I and II quadrants for 2019. In these last quadrants also the samples from Amalfi PGI area (A1, A2 and A3) of 2019 production were grouped. Instead, the two production years of EB samples mainly separated along PC1, being the score distribution for the most part in II and III quadrants for 2019 production and in the I and IV quadrants for 2018 production. The higher contributions (loadings) on PC1 were at the wavenumbers of 7266 cm^{-1} and 7412 cm^{-1} (Fig.3) which were around the band at 7300 cm^{-1} in lemon spectra attributed to methyl, methylene, methyl and methoxy absorptions. The further relevant loadings at 6069 cm^{-1} and at 6176 cm^{-1} were in the typical interval of 1st overtone of CH, also in unsaturated structure, these absorptions could be totally overlapped to the 6876 cm^{-1} wide band in lemon spectra. The minimum loadings were recorded for 5164 cm^{-1}and 6882 cm^{-1} wavenumbers quite corresponding to the wide 5173 cm^{-1} and 6876 cm^{-1} bands attributed to peel water in lemon. This suggests that the main discriminating property of lemon spectra on PC1 depends by the difference in alkyl chains, like those of carotenoids, and methoxy groups, such as those of pectin. This agrees with the wide variability of the colour properties a* and b* of lemons (Tab. 2) indicating the variation of chlorophyll and carotenoid contents related to the maturation status to which also a variation of pectin methoxylation degree can correspond (Spiegel-Roy & Goldschmidt, 1996). The loadings on PC2 were evident at the frequencies of 5223 cm^{-1} and 6885 cm^{-1}, and less marked at 5577 cm^{-1}. All these loadings coincided with the relevant signal of lemon spectra attributed to the OH of water and polysaccharides. A variability of polysaccharide spectral characteristics could depend mainly to the extension of peel explored by NIR rays being the peel thickness among the properties of lemons with widest variability (Tab. 2). On this view the groupings of lemon NIR spectra by production year could depend mainly on accumulation of chromatic substances in agreement with the general annual differences of the chromatic properties a* and b* of lemons, particularly for OS-EB and SA-EB which differentiated mainly along PC1, and for water and carbohydrates variation due to annual peel thickness differences for lemons from the other orchards which differentiated mainly along PC1. This is supported by the S-LDAs of lemon quality properties grouped by years for EB orchard and for the remaining orchards with biennial productions (Table S2, provided in Appendix A. Supplementary materials). For lemons from EB, b* was among the best S-LDA discriminating variables, followed also by a* (100% correct classification and cross validation), for the other orchards, THK was among the best (97% correct classification and 96% validation). The TA well discriminated in both the S-LDAs, in agreement with the observed significant differences by years for each orchard (Table 2), as well as the pH for the S-LDA for only the grouped orchards. These properties, however, should not be directly measured by NIR. These results indicate that the different maturation characteristics expressed in the two years are perceived by NIR spectra of intact lemon and that such differences could depend on combination of detected compositional and probably also morphological factors mainly of peels. However, the relative weight of factors, although concurrent, could be different, as suggested by the seasonal variation of the NIR spectra of lemons from EB respect to those from the other orchards. This suggests that although the different seasonal macro-climatic characteristics, quite homogeneous in all orchard areas, could have impress differences in maturation proprieties, they are expressed differently and perceived by NIR spectroscopical analysis. The expressions of this differences could depend by the whole environment, in addition to climate also plant nutrient uptake which depends by soil nutrient availability that could be more similar in the PGI areas and Portici respect to Eboli.

3.4 Relationships between NIR spectra and quality properties

The multiple linear regression between PCs from the NIR spectra of intact lemons and the quality properties were performed considering all lemons from the two production yeas, for each year and, with the same criterion, separating cultivars (Tab. 3). The regression significances could be negatively influenced by a reduced number of samples as resulted evident for SA of 2018 with the lowest number of samples (11) which implied frequent not significant correlations, although with higher R^2. In the most part of the case the regressions with all samples of two production years did not improve, both in terms of R^2 and CV, respect to the best result recorded for a single production year. The regression models could suggest consideration on the use of NIR spectra for predicting lemon properties. About chromatic parameters, the correlations of spectra were significant for b* in all cases examined, values of R^2 between 0.985 and 0.387, and for a*, R^2 between 0.662 and 0.308, except that of SA in 2018 although with R^2 0.910. This is in agreement with the ability of NIR spectroscopy on entire lemon to detect the epicarp contents of chloroplast/chromoplast pigments. The correlations with the brightness parameter L* were significant except for all cases related to 2018 production year. This may depend by roughness of rind surfaces which could have incongruously interfered on NIR and/or colorimeter measurements for the production of 2018 rather than on that of 2019. The roughness and thickening of rind have been suggested to be induced by high levels of endogenous GA3 and cytokinin, these lasts could decrease in mild drought conditions (Cronjé et al., 2013; Nelissen et al., 2018) as for 2018 lemon productions. For 2018 lemon productions THK significantly correlated with spectra, except for the SA although with quite high R^2, while mainly for 2019 lemon productions THK did not correlate or expressed low R^2, this could depend by inaccuracies in NIR measurement due to a high attenuation of IR ray through the thicker peels than those of 2018 lemons. The correlation trend of NIR spectra with PW quite recalled that with THK but often with higher R^2. The evaluation of this intensive property could agree with NIR direct measurement of water content in peels and of related peel volume or an its proportional property such as THK. For PW, the not significant correlation for 2019 OV lemons and the low R^2 of all correlations including 2019 lemons could depend by less accurate measure in thicker peel as discussed before. All the percentage of juices were significantly correlated with NIR spectra, except for the SA in 2018 probably for the above mentioned low number of samples. Such successful correlations, also showing R^2 values quite higher, could be due to a combination of significant correlations with different lemon measurable properties such as THK, as supported by Spearman significant correlation (R -0.755) between both properties (Tab. S3, provided in supplementary material), and other shape characteristics associated to the lemon volume that NIRA detector measured, as proved by correlations of NIR spectra with geometrical characteristics of some spherical and ellipsoidal objects of variable dimension placed on the detector (data not shown). This last is probably due to the ability to measure reflectance effects determined by the curvature angle of these objects respect the flat window of detector (Ø 14 mm). The pH showed only few significant correlations with NIR spectra and with low R^2 values, this depends mainly by the poor detectability of hydronium ion (Chang et al., 2018) but also by the suggested NIR ray attenuation reaching pulp. The linear regression of NIR spectra with the SSC is generally significant and with quite high values of R^2 except for OS and SA together in 2019 and for OS in 2019 (not considering the not significance for SA of 2018 attributable to the low number of samples), instead, the regression with TA was significant in 2019 but not or with low R^2 in 2018. Such trend appeared incongruous since TA was the major contributor to SSC (Tab. 2). This could be explained by the fact that SSC and TA characteristics were not directly measured, due thick lemon peels, but indirectly from some property of peel, detectable by NIR, quite correlated, but not perfectly harmonized, with SSC and TA. The C+ concentrations significantly correlated except for the group of SA in 2018 and in the combined production years 2018-2019, they could only be indirectly measured being cations in solution not detectable by NIR. The indirect NIR measurement of TA was supported by Spearman correlation analysis indicating a significant correlation of this property with the colour parameters measured by NIR, a* (R 0.601) and b* (R 0.393) (Tab. S3, supplementary material). At the same time SCC and C+, although it did not correlate with the colour parameters, were significantly correlated with TA

(R 0.361 and 0.367 respectively). In a certain measure also the occasional correlation of pH with NIR spectra could be justified bay and indirect measurement through the correlate properties TA (R -0.442). and SCC (R -0.215). The best expected results for the prediction accuracy of quality properties evaluated using NIR spectra, evidenced by the CV values of regressions (Tab. 3), were for b* (CV ≤ 8.3%), and SSC (CV ≤ 10.9%) although in some cases the regressions were not significant. For pH the CV values are still better (CV ≤ 7.6%) but the regressions were unreliable because not significant or with very low R^2. The other properties showed CV no higher than 36.4% except for the highest value for a* that already recorded a high variability of the replicates of colorimetric measurements on each lemon. In general, the CV of the properties were quite similar considering each cultivar, OS or SA, and both together, OS and SA, rather the accuracy difference appeared to depend by the production years and in general for the production years together the CV was more similar to the worst CV recorded between the two-production year. For the most part of the properties the CV resulted better for the production of 2018 than of 2019 although with a higher number of samples. All this suggest the use of the annual calibrations with intact lemon NIR spectra for measurement of lemon quality properties rather than to implement a predictive calibration model with further data from other lemon production years.

3.5 Discriminant models with NIR spectra

Cultivar discrimination - The discrimination of cultivars SA and OS by LDA models with all NIR spectra of intact lemons (C1) showed the accuracies of calibration 69% and validation 63%, further their ratio was 0.92. The LDA cultivar discrimination carried out for single production year gave for 2018 production (C2) the accuracies of calibration 85% and validation 66% with the ratio 0.78 and for 2019 (C3) calibration 78% and validation 63% with the ratio 0.81 (Table 4). The calibration accuracy of C1 model was lower than those of C2 and C3 models suggesting that on the biennial model the annual variation of lemon properties can have had a certain weight on detectable spectroscopical characteristics proper of the cultivars attenuating their discrimination. However, the validation accuracies and the accuracy ratios of all the models were not with the same trend of calibration accuracy probably caused by the inhomogeneous sizes of cultivar groups in models. It is worth noting that apparently the model grouping the different cultivars in different production years (C4) showed the best results in calibration and validation accuracies 90% and 72%, respectively (Table 4), but this was due to the amplification of the difference between cultivars arising from production year variations of spectral properties similarly influencing both cultivars, as evidenced by PCA of spectra and lemon quality properties, these last quite detectable by NIR spectroscopy.

This suggest that the discrimination of Ovale di Sorrento from Sfusato Amalfitano by NIR spectroscopy on the intact lemons could be not so effective due the high influence of growing conditions on the spectroscopical response which overcome those arising from the typicality of each cultivar.

Provenance discrimination - To essay the ability of NIR spectra of intact lemons to discriminate provenance, LDA models was explored grouping lemons for production PGI area of Sfusato Amalfitano, SA-PGI, and Ovale di Sorrento, OS-PGI *vs* the outside grouped sites, EB and PR, in which each cultivar growth. The LDA for OS of the 2018 and 2019 productions correctly classified 84% and validated 72% with 0.86 accuracy ratio (P1). Separating the OS groups by production years the LDA models, P2 for 2018 and P3 for 2019 production years gave both 100% correctly classified and 100% and 81% validation accuracy, with 1 and 0.81 accuracy ratios, respectively. Thus, the model based on lemons from the two production years was with lower performance than the models based on a single production year. Similarly, the LDA model for SA discriminate more efficiently in the single production year 2019, P5 100% calibration 74% validation accuracies, ratio 0.74, than considering the production years together, P4 84% calibration 44% validation accuracies, ratio 0.52, although this model lack of SA-PGI of 2018. As seen for cultivar discrimination, the observed NIR spectral differences due to annual environmental variations could have obscured the spectral differences determined by fruit properties arising from the cultivation areas.

4 Conclusions

The NIR spectroscopy with chemometric data treatment on intact lemons of two PGI lemons, Limone di Sorrento (cv Ovale di Sorrento) and Limone Costa D'Amalfi (cv Sfusato Amalfitano), evidenced it potential ability to explore the cultivar difference, and fruit provenance and quality properties, but some awareness in the use has to be recommended. The PCA of the intact lemon NIR spectra of both cultivars evidenced a natural clustering of samples according mainly to the production years, 2018 and 2019. This was attributed to the annual variation of lemon properties caused by the differences of cultivation environment, mainly rainfall, that varied in the two years quite homogeneously for all production sites due to the restrict investigated production areas and influenced similarly the quality properties for both cultivars. This grouping indicated the ability of NIR spectroscopy to detect many of morphological and chemical properties of lemons. In particular, the NIR spectra represented mainly the properties related to peel since the NIR ray appeared to not penetrate beyond peel thickness, thus poorly detecting pulp properties. Relationship of spectra with lemon quality properties evaluated by MLR models with principal components extracted from NIR spectra of intact lemons showed many significant correlations suggesting an applicability for also predicting the quality of juices. However, juice properties evaluated by NIR have to be considered as derived essentially from correlated NIR measurable lemon peel properties quite harmonized with juice composition. In this way, the best NIR prediction results were obtained with linear regression of lemon spectra from production years separately being more homogeneous the lemon spectroscopic properties influenced by ripening environment. Discrimination of the two Femminello cultivars with LDA models based on NIR spectra of intact lemons appeared to be strongly influenced by the growing environment which as evidenced before influence NIR response as much to reduce selectivity of cultivar difference detectable by NIR spectra. As for prediction use of NIR spectra, the differences in cultivars were more reliable if Ovale di Sorrento and Sfusato Amalfitano were compared in the same production year. Further, also LDA models based on NIR spectra of intact lemons of each cultivar according to provenances, PGI area vs relatively close not PGI sites, resulted better for single production year than for two production years together. Although the NIR spectroscopy on entire lemon fruit appeared a powerful analytical method, its use for prediction properties, for origin and cultivar discrimination need caution in the use and in the choice of sampling set for measurements because its ability to grasp many development characteristics of fruit depending by ripening environment can prevail on the result of measure that has set.

5 References

Amendola, L., Firmani, P., Bucci, R., Marini, F., & Biancolillo, A. (2020). Authentication of Sorrento walnuts by NIR spectroscopy coupled with different chemometric classification strategies. *Applied Sciences (Switzerland)*, *10*(11). https://doi.org/10.3390/app10114003

Amenta, M., S.,Fabroni, Costa, C., Rapisarda, P., Fabroni, S., Costa, C., & Rapisarda, P. (2016). Traceability of “Limone di Siracusa PGI” by a multidisciplinary analytical and chemometric approach. *Food Chemistry*, *211*, 734–740. https://doi.org/10.1016/j.foodchem.2016.05.119

Amodio, M. L., Ceglie, F. G., Chaudhry, M. M. A., Piazzolla, F., & Colelli, G. (2017). Potential of NIR spectroscopy for predicting internal quality and discriminating among strawberry fruits from different production systems. *Postharvest Biology and Technology*, *125*(March), 112–121. https://doi.org/10.1016/j.postharvbio.2016.11.013

Arendse, E., Amos, O., Samukelo, L., & Linus, U. (2018). Non-destructive prediction of internal and external quality attributes of fruit with thick rind : A review. *Journal of Food Engineering*, *217*, 11–23. https://doi.org/10.1016/j.jfoodeng.2017.08.009

Arendse, E., Fawole, O. A., Magwaza, L. S., Nieuwoudt, H. H., & Opara, U. L. (2017). Development of calibration models for the evaluation of pomegranate aril quality by Fourier-transform near infrared spectroscopy combined with chemometrics. *Biosystems Engineering*, *159*, 22–32. https://doi.org/10.1016/j.biosystemseng.2017.04.004

Birth, G. S., Dull, G. G., Renfroe, W. T., & Kays, S. J. (1985). Nondestructive Spectrophotometric Determination of Dry-Matter in Onions. *Journal of the American Society for Horticultural Science*, *110*(2), 297–303.

Blanco, M., & Villarroya, I. (2002). NIR spectroscopy: A rapid-response analytical tool. *TrAC - Trends in Analytical Chemistry*, *21*(4), 240–250. https://doi.org/10.1016/S0165-9936(02)00404-1

Carlini, P., Massantini, R., & Mencarelli, F. (2000). Vis-NIR Measurement of Soluble Solids in Cherry and Apricot by PLS Regression and Wavelength Selection. *Journal of Agricultural and Food Chemistry*, *48*, 5236–5242. https://doi.org/10.1021/jf000408f

Cavaco, A. M., Passos, D., Pires, R. M., Antunes, M. D., & Guerra, R. (2021). Nondestructive Assessment of Citrus Fruit Quality and Ripening by Visible–Near Infrared Reflectance Spectroscopy. *Citrus*. https://www.intechopen.com/books/advanced-biometric-technologies/liveness-detection-in-biometrics

Cavaco, Ana M., Pires, R., Antunes, M. D., Panagopoulos, T., Brázio, A., Afonso, A. M., Silva, L., Lucas, M. R., Cadeiras, B., Cruz, S. P., & Guerra, R. (2018). Validation of short wave near infrared calibration models for the quality and ripening of 'Newhall' orange on tree across years and orchards. *Postharvest Biology and Technology*, *141*(March), 86–97. https://doi.org/10.1016/j.postharvbio.2018.03.013

Chang, K., Shinzawa, H., & Chung, H. (2018). Concentration determination of inorganic acids that do not absorb near-infrared (NIR) radiation through recognizing perturbed NIR water bands by them and investigation of accuracy dependency on their acidities. *Microchemical Journal*, *139*, 443–449. https://doi.org/10.1016/j.microc.2018.03.019

Clever, J., & Valverde, C. (2019). EU-28 Citrus Semi-annual Rebound in EU Production Expected to Lower EU Citrus Imports. *Global Agricultural Information Network*, *SP1929*(June).

Cowe, I. A., & McNicol, J. W. (1985). Use of Principal Components in the Analysis of Near-Infrared Spectra. *Applied Spectroscopy*, *39*(2), 257–266. https://doi.org/10.1366/0003702854248944

Cozzolino, D., Cynkar, W. U., Shah, N., & Smith, P. A. (2011). Can spectroscopy geographically classify Sauvignon Blanc wines from Australia and New Zealand? *Food Chemistry*, *126*(2), 673–678. https://doi.org/10.1016/j.foodchem.2010.11.005

Cronjé, P. J. R., Stander, O. P. J., & Theron, K. I. (2013). Fruit splitting in citrus. In *Horticultural Reviews* (First, Vol. 41, pp. 177–200). John Wiley & Sons, Inc., Hoboken, New Jersey. Published. https://doi.org/10.1002/9781118707418.ch04

Di Matteo, A., Di Rauso Simeone, G., Cirillo, A., Rao, M. A., & Di Vaio, C. (2021). Morphological characteristics, ascorbic acid and antioxidant activity during fruit ripening of four lemon (Citrus limon (L.) Burm. F.) cultivars. *Scientia Horticulturae*, *276*(September 2020), 109741. https://doi.org/10.1016/j.scienta.2020.109741

Dimopoulou, M., Alba, K., Campbell, G., & Kontogiorgos, V. (2019). Pectin recovery and characterization from lemon juice waste streams. *Journal of the Science of Food and Agriculture*, *99*(14), 6191–6198. https://doi.org/10.1002/jsfa.9891

Dunn, O. J. (1964). Multiple Comparisons Using Rank Sums. *Technometrics*, *6*(3), 241–252.

EEC No 2080/92. (1992). COUNCIL REGULATION (EEC) No 2080/92 Instituting a Comunity aid

Scheme for Forestry Measures in Agriculture Fearnside. *Official Journal of European Communities*, *2080*, 96–99.

EEC No 2081/92. (1992). COUNCIL REGULATION (EEC) No 2081/92 On the protection of geographical indications and designations of origin for agricultural products and foodstuffs. *Official Journal of the European Communities*, *L208*(L), 1–8.

Eisenstecken, D., Stürz, B., Robatscher, P., Lozano, L., Zanella, A., & Oberhuber, M. (2019). The potential of near infrared spectroscopy (NIRS) to trace apple origin: Study on different cultivars and orchard elevations. *Postharvest Biology and Technology*, *147*(August 2018), 123–131. https://doi.org/10.1016/j.postharvbio.2018.08.019

Elfadl, E., Reinbrechta, C., & Claupein, W. (2010). Development of near infrared reflectance spectroscopy (NIRS) calibration model for estimation of oil content in a worldwide safflower germplasm collection. *International Journal of Plant Production*, *4*(4), 259–270. http://www.gau.ac.ir/Jm/Programs/JurnalMgr/VolumArticle/EN_163_2.pdf

EU 428/2019. (2019). *Modifica del regolamento di esecuzione (UE) n. 543/2011 per quanto concerne le norme di commercializzazione nel settore degli ortofrutticoli*. http://www.confcommercio.cs.it/spaw2/uploads/files/allegati.pdf

Fabroni, S., Amenta, M., Allegra, M., Sorrentino, G., Roccuzzo, G., & Rapisarda, P. (2015). Traceability of Citrus Fruit Using Isotopic and Chemical Markers. *Acta Horticulturae ·*, *January*. https://doi.org/10.17660/ActaHortic.2015.1065.183

FAOSTAT. (2021). *Food and Agriculture Organization of the United Nations*. http://www.fao.org/faostat/

Fhatuwani, M. N. (2001). *Yield and fruit quality of citrus species relative to foliar sprays of macronutrients*. University of Stellenbosch.

G.U.n°178. (2001). *Disciplinare di produzione dell ' Indicazione Geografica Protetta " Limone Costa d ' Amalfi ".*

G.U.n°24. (2011). *Disciplinare di produzione dell'Indicazione Geografica Protetta "Limone di Sorrento."* https://ec.europa.eu/agriculture/quality/door/documentDisplay.html?chkDocument=1063_1_it

Gabioud, S., Baumgartner, D., Gasser, F., Kneubühler, Y., Lattmann, S., & Höhn, E. (2008). Non destructive quality measurements on apples. *Acta Horticulturae*, *796*(August), 217–224. https://doi.org/10.17660/ActaHortic.2008.796.29

García-Tejero, I., Romero-Vicente, R., Jiménez-Bocanegra, J. A., Martínez-García, G., Durán-Zuazo, V. H., & Muriel-Fernández, J. L. (2010). Response of citrus trees to deficit irrigation during different phenological periods in relation to yield, fruit quality, and water productivity. *Agricultural Water Management*, *97*(5), 689–699. https://doi.org/10.1016/j.agwat.2009.12.012

Gòmez, A. H., He, Y., & Annia Garcìa Pereira. (2006). *Non-destructive measurement of acidity , soluble solids and firmness of Satsuma mandarin using Vis / NIR-spectroscopy techniques*. *77*, 313–319. https://doi.org/10.1016/j.jfoodeng.2005.06.036

Guo, X., Cai, R., Wang, S., Tang, B., Li, Y., & Zhao, W. (2018). Non-destructive geographical traceability of sea cucumber (*Apostichopus japonicus*) using near infrared spectroscopy combined with chemometric methods. *Royal Society Open Science*, *5*(1), 170714. https://doi.org/10.1098/rsos.170714

Hirri, A., Bassbasi, M., & Platikanov, S. (2016). FTIR Spectroscopy and PLS-DA Classification and Prediction of Four FTIR Spectroscopy and PLS-DA Classification and Prediction of Four Commercial Grade Virgin Olive Oils from Morocco. *Food Analytical Methods*, *9*(August), 974–981. https://doi.org/10.1007/s12161-015-0255-y

Iannace, A., Capuano, M., & Galluccio, L. (2011). “Dolomites and dolomites” in Mesozoic platform carbonates of the Southern Apennines: Geometric distribution, petrography and geochemistry. *Palaeogeography, Palaeoclimatology, Palaeoecology*, *310*(3–4), 324–339. https://doi.org/10.1016/j.palaeo.2011.07.025

Iglesias, D. J., Cercós, M., Colmenero-flores, J. M., Naranjo, M. A., Carrera, E., Ruiz-rivero, O., Lliso, I., Morillon, R., Tadeo, F. R., & Talon, M. (2007). Physiology of citrus fruiting. *Braz. J. Plant Physiol.*, *19*(4), 333–362.

ISTAT. (2017). *Dataset : PDO, PGI and TSG quality products*. http://dati.istat.it/Index.aspx?QueryId=36787&lang=en

Jha, S. N., Narsaiah, K., Jaiswal, P., Bhardwaj, R., Gupta, M., Kumar, R., & Sharma, R. (2014). Nondestructive prediction of maturity of mango using near infrared spectroscopy. *Journal of Food Engineering*, *124*, 152–157. https://doi.org/10.1016/j.jfoodeng.2013.10.012

Jiménez-carvelo, A. M., Lozano, V. A., & Olivieri, A. C. (2019). Comparative chemometric analysis of fluorescence and near infrared spectroscopies for authenticity con firmation and geographical origin of Argentinean extra virgin olive oils. *Food Control*, *96*(August 2018), 22–28. https://doi.org/10.1016/j.foodcont.2018.08.024

Kahriman, F., Egesel, C. O., Onsekiz, Ç., & Agriculture, F. (2011). Development of a Calibration Model To Estimate Quality Traits in Wheat Flour Using Nir (Near Infrared Reflectance) Spectroscopy. *Journal of Agricultural Science*, *43*(3), 392–400.

Kawano, S., Fujiwara, T., & Iwamoto, M. (1993). Nono destructive determination of Sugar content in Satuma Mandarin using Near Infrared (NIR) Transmittance. *Journal of the Japanese Society for Horticultural Science*, *62*(2), 465–470.

Ladaniya, M. S. (2010). *Citrus fruit: biology, technology and evaluation*. Accademic press.

Latorre, C. H., Crecente, R. M. P., Martín, S. G., & García, J. B. (2013). A fast chemometric procedure based on NIR data for authentication of honey with protected geographical indication. *Food Chemistry*, *141*(4), 3559–3565. https://doi.org/10.1016/j.foodchem.2013.06.022

Lee, K. J., Kim, G. Y., Kang, S. W., Son, J. R., Choi, D. S., & Choi, K. H. (2004). Measurement of Sugar Contents in Citrus Using Near Infrared Transmittance. *Key Engineering Materials*, *270–273*, 1014–1019. https://doi.org/10.4028/www.scientific.net/KEM.270-273.1014

Lee, L. S., & Chapman, J. C. (1988). *Yield and fruit quality responses of Ellendale mandarins to different nitrogen and potassium fertiliser rates*. 143–148.

Levy, Y., Bar-Akiva, A., & Vaadia, Y. (1978). Influence of irrigation and environmental factors on grapefruit acidity. *Journal of the American Society for Horticultural Science*, *1*(103), 73–76.

Li, X., Tsuta, M., Hayakawa, F., Nakano, Y., Kazami, Y., & Ikehata, A. (2021). Estimating the sensory qualities of tomatoes using visible and near-infrared spectroscopy and interpretation based on gas chromatography–mass spectrometry metabolomics. *Food Chemistry*, *343*, 128470. https://doi.org/10.1016/j.foodchem.2020.128470

Lin, Q., Qian, J., Zhao, C., Wang, D., & Liu, C. (2016). Low Temperature Induced Changes in Citrate Metabolism in Ponkan (Citrus reticulata Blanco cv . Ponkan) Fruit during Maturation. *PLoS ONE*, 1–15. https://doi.org/10.1371/journal.pone.0156703

Liu, L., Cozzolino, D., Cynkar, W. U., Gishen, M., & Colby, C. B. (2006). Geographic Classification of Spanish and Australian Tempranillo Red Wines by Visible and Near-Infrared Spectroscopy

Combined with Multivariate Analysis. *Journal of Agricultural and Food Chemistry*, *54*, 6754–6759.

Luykx, D. M. A. M., & van Ruth, S. M. (2008). An overview of analytical methods for determining the geographical origin of food products. *Food Chemistry*, *107*(2), 897–911. https://doi.org/10.1016/j.foodchem.2007.09.038

MiPAF. (2000). *Metodi di analisi chimica del suolo* (Franco Angeli (ed.)).

Miyamoto, K., & Kitano, Y. (2009). Non-Destructive Determination of Sugar Content in Satsuma Mandarin Fruit by near Infrared Transmittance Spectroscopy. *Journal of Near Infrared Spectroscopy*, *3*(4), 227–237. https://doi.org/10.1255/jnirs.73

Monteiro, P. I., Santos, J. S., Alvarenga Brizola, V. R., Pasini Deolindo, C. T., Koot, A., Boerrigter-Eenling, R., van Ruth, S., Georgouli, K., Koidis, A., & Granato, D. (2018). Comparison between proton transfer reaction mass spectrometry and near infrared spectroscopy for the authentication of Brazilian coffee: A preliminary chemometric study. *Food Control*, *91*(10), 276–283. https://doi.org/10.1016/j.foodcont.2018.04.009

Nelissen, H., Sun, X. H., Rymen, B., Jikumaru, Y., Kojima, M., Takebayashi, Y., Abbeloos, R., Demuynck, K., Storme, V., Vuylsteke, M., De Block, J., Herman, D., Coppens, F., Maere, S., Kamiya, Y., Sakakibara, H., Beemster, G. T. S., & Inzé, D. (2018). The reduction in maize leaf growth under mild drought affects the transition between cell division and cell expansion and cannot be restored by elevated gibberellic acid levels. *Plant Biotechnology Journal*, *16*(2), 615–627. https://doi.org/10.1111/pbi.12801

Nicolaï, B. M., Beullens, K., Bobelyn, E., Peirs, A., Saeys, W., Theron, K. I., & Lammertyn, J. (2007). Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: A review. *Postharvest Biology and Technology*, *46*(2), 99–118. https://doi.org/10.1016/j.postharvbio.2007.06.024

Picarelli, L., Santo, A., Crescenzo, G., & Olivares, L. (2008). Macro-zoning of areas susceptible to flowslide in pyroclastic soils in the Campania region. *Landslides and Engineered Slopes. From the Past to the Future*, *June*, 1951–1957. https://doi.org/10.1201/9780203885284-c271

Polesello, A., Giangiacomo, R., & Dull, G. G. (1981). Application of near infrared spectrophotometry to the nondestructive analysis of foods : A review of experimental results. *Food Science and Nutrition*, *18*(3), 203–230.

Reddy, N. S., D, N., & Yadav, B. K. (2016). Non-Destructive Quality Assessment of Citrus Fruits Using Ft-Near-Infrared Spectroscopy. *International Journal of Scicence Environment and TEchnology*, *5*(4), 1850–1860.

Ren, G., Wang, S., Ning, J., Xu, R., Wang, Y., Xing, Z., Wan, X., & Zhang, Z. (2013). Quantitative analysis and geographical traceability of black tea using Fourier transform near-infrared spectroscopy (FT-NIRS). *Food Research International*, *53*(2), 822–826. https://doi.org/10.1016/j.foodres.2012.10.032

Ribeiro, J. S., Ferreira, M. M. C., & Salva, T. J. G. (2011). Chemometric models for the quantitative descriptive sensory analysis of Arabica coffee beverages using near infrared spectroscopy. *Talanta*, *83*(5), 1352–1358. https://doi.org/10.1016/j.talanta.2010.11.001

Rinnan, Å., Berg, F. van den, & Engelsen, S. B. (2009). Review of the most common pre-processing techniques for near-infrared spectra. *TrAC - Trends in Analytical Chemistry*, *28*(10), 1201–1222. https://doi.org/10.1016/j.trac.2009.07.007

Roy Choudhury, A. K. (2015). Principles of Colour and Appearance Measurement. Volume 2: Visual Measurement of Colour, Colour Comparison and Management. In *Principles of Colour and Appearance Measurement* (Vol. 2).

Ruggiero, L., Fontanella, M. C., Amalfitano, C., Beone, G. M., & Adamo, P. (2021). Provenance discrimination of Sorrento lemon with Protected Geographical Indication (PGI) by multi-elemental fingerprinting. *Food Chemistry*, *362*. https://doi.org/10.1016/j.foodchem.2021.130168

Saranwong, S., Sornsrivichai, J., & Kawano, S. (2004). Prediction of ripe-stage eating quality of mango fruit from its harvest quality measured nondestructively by near infrared spectroscopy. *Postharvest Biology and Technology*, *31*(2), 137–145. https://doi.org/10.1016/j.postharvbio.2003.08.007

Savo, V., Salvati, L., & Caneva, G. (2016). In-between soil erosion and sustainable land management: climate aridity and vegetation in a traditional agro-forest system (Costiera Amalfitana, southern Italy). *International Journal of Sustainable Development and World Ecology*, *23*(5), 423–432. https://doi.org/10.1080/13504509.2015.1132282

Scott, M. F., & Baker, K. C. (1947). Anatomy of Washington Navel Orange Rind in Relation to Water Spot. *Botanical Gazette*, *108*(4), 459–475.

Shafqat, W., Naqvi, S. A., Maqbool, R., Haider, M. S., Jaskani, M. J., & Khan, I. A. (2021). Climate Change and Citrus. In *Citrus*. IntechOpen. https://www.intechopen.com/books/advanced-biometric-technologies/liveness-detection-in-biometrics

Spiegel-Roy, P., & Goldschmidt, E. E. (1996). The Biology of Citrus. *The Biology of Citrus*. https://doi.org/10.1017/cbo9780511600548

Sun, Y., Singh, Z., Tokala, V. Y., & Heather, B. (2019). Harvest maturity stage and cold storage period influence lemon fruit quality. *Scientia Horticulturae*, *249*(July 2018), 322–328. https://doi.org/10.1016/j.scienta.2019.01.056

Wang, J., Liu, H., & Ren, G. (2013). Near-infrared spectroscopy (NIRS) evaluation and regional analysis of Chinese faba bean (Vicia faba L .). *The Crop Journal*, *2*(1), 28–37. https://doi.org/10.1016/j.cj.2013.10.001

Woo, Y. A., Kim, H. J., Ze, K. R., & Chung, H. (2005). Near-infrared (NIR) spectroscopy for the non-destructive and fast determination of geographical origin of Angelicae gigantis Radix. *Journal of Pharmaceutical and Biomedical Analysis*, *36*(5), 955–959. https://doi.org/10.1016/j.jpba.2004.08.037

Woodcock, T., Downey, G., & O'Donnell, C. P. (2009). Near infrared spectral fingerprinting for confirmation of claimed PDO provenance of honey. *Food Chemistry*, *114*(2), 742–746. https://doi.org/10.1016/j.foodchem.2008.10.034

Xie, L., Ye, X., Liu, D., & Ying, Y. (2011). Prediction of titratable acidity , malic acid , and citric acid in bayberry fruit by near-infrared spectroscopy. *FRIN*, *44*(7), 2198–2204. https://doi.org/10.1016/j.foodres.2010.11.024

Zhao, H., Guo, B., Wei, Y., & Zhang, B. (2013). Near infrared reflectance spectroscopy for determination of the geographical origin of wheat. *Food Chemistry*, *138*(2–3), 1902–1907. https://doi.org/10.1016/j.foodchem.2012.11.037

Table 1 Sampling locations, annual rainfall, and number of the samples of Femminello lemons cv Ovale di Sorrento (OS) and cv Sfusato Amalfitano (SA) collected in 2018 and 2019.

Orchard locations		**ID**	**Rainfall[+] (mm/year)**	**2018**		**2019**	
				OS	**SA**	**OS**	**SA**
Massa Lubrense[&]	40°36'28"N, 14°20'36"E	**S1**	517	10	-	13	-
Piano di Sorrento[&]	40°37'51"N, 14°24'22"E	**S2**	763	12	-	19	-
Tramonti[#]	40°40'27"N, 14°38'53"E	**A1**	1154	-	-	-	6
Maiori[#]	40°39'32"N, 14°38'48"E	**A2**	905	-	-	-	7
Maiori[#]	40°38'32"N, 14°39'31"E	**A3**		-	-	-	6
Portici*	40°48'54"N, 14°20'49"E	**PR**	680	6	6	6	6
Eboli[§]	40°33'21"N, 14°58'53"E	**EB**	804	6	5	6	5

&Limone di Sorrento PGI area. #Limone Costa d' Amalfi PGI area.

*Experimental field Dept. of Agricultural Sciences, Univ. of Naples Federico II. §Regional experimental farm Improsta.

+Annual mean rainfall 2016-2018 (available: http://centrofunzionale.regione.campania.it)

Table 2. Chemical and physical properties of lemons according to different orchards, cultivars, and production years. Different letters indicate significative differences among the columns, while (*) the differences between the two production years using non-parametric Kruskal–Wallis and Dunn rank tests ($p \leq 0.05$). Peel chromaticity (L*, a*, b*); thickness of pericarp (THK); peel water (PW); juice yield (Juices%); soluble solid content (SSC), titratable acidity (TA), molar concentration of positive charge from the cations (C^+)

	OS-S1		OS-S2		OS-PR		OS-EB		SA-PR		SA-EB		SA-A1	SA-A2	SA-A3
	2018	2019	2018	2019	2018	2019	2018	2019	2018	2019	2018	2019	2019	2019	2019
L*	72.6 ± 1.7ab	74.7 ± 3.0a*	73.1 ± 1.3ab	73.5 ± 1.8ab	69.9 ± 8.1ab	71.3 ± 1.2b	72.9 ± 1.5ab	70.3 ± 2.0b	72.8 ± 2.3ab	72.5 ± 0.7ab	72.9 ± 1.0ab	70.9 ± 1.6b	72.2 ± 2.4ab	71.5 ± 1.0b	74.2 ± 1.3ab
a*	0.1 ± 1.4ab*	-4.6 ± 0.8e	0.2 ± 1.7ab*	-2.1 ± 1.1bcde	-0.9 ± 1.5abcd	-0.4 ± 2.4abc	1.7 ± 1.0a*	-3.4 ± 2.5de	-3.5 ± 0.8de	-1.9 ± 1.9bcde*	0.4 ± 1.0ab	-1.1 ± 3.9bcde	-3 ± 1.9cde	-3.3 ± 0.9cde	-2.3 ± 0.4bcde
b*	70 ± 1.7ab*	58.4 ± 3.4cd	71.9 ± 2.1a*	64.3 ± 1.9abc	57 ± 6.9cd	61.7 ± 4.6abcd	75.6 ± 2.5a*	48.5 ± 5.6d	55 ± 2.4cd	60.8 ± 3.4abcd*	72.8 ± 1.0a*	57.7 ± 8.7cd	62.6 ± 3.4abcd	59.7 ± 2.0bcd	64.3 ± 1.9abcd
THK mm	3.7 ± 0.1cd	5.6 ± 0.1abcd*	4.1 ± 0.1bcd	5.0 ± 0.1abcd*	3.3 ± 0.1d	6.7 ± 0.1abc*	3.1 ± 0.1d	5.8 ± 0.1abcd*	3.1 ± 0.1d	7.6 ± 0.1ab*	4.1 ± 0.1abcd	5.4 ± 0.1abcd*	7.3 ± 0.2abc	9.4 ± 0.2a	9.5 ± 0.2a
PW %	78.7 ± 2.5bc	80.0 ± 5.8abc	80.8 ± 1.8abc*	78.7 ± 2.6bc	81.3± 2.2abc	81.1 ± 1.3abc	71.7 ± 8.1c	79.2± 4.8abc	80.5 ± 0.7abc*	84.8 ± 2.1c	78,9 ± 2.7abc	80.1 ± 1.6abc	81.7 ± 1.1abc	82.7 ± 1.0a	82.6± 0.7ab
Juices %	46.4 ± 6.6a*	29.5 ± 5.3def	42.0 ± 5.9a*	35.4 ± 6.7bcde	39.9 ± 5abc*	35.1 ± 4.8bcdef	42.3 ± 2ab*	25.9 ± 7.5f	37.9 ± 3.7abcd*	32.2 ± 2.3bcdef	46.8 ± 3.1a*	32.0 ± 5.1cdef	26.8 ± 2.7ef	25.1 ± 4.0f	27.4 ± 4.4def
pH	2.3 ± 0.1bc	2.3 ± 0.1ab	2.2 ± 0.1bc	2.6 ± 0.1a*	2.2 ± 0.1bc	2.2 ± 0.2bc	2.3 ± 0.1bc	2.3 ± 0.1ab*	2.3 ± 0.1bc*	2.2 ± 0.1c	2.3 ± 0.1bc	2.3 ± 0.1abc	2.4 ± 0.1abc	2.3 ± 0.1bc	2.3 ± 0.1ab
SSC °Brix	7.6 ± 0.5abc	7.6 ± 0.6abc	6.9 ± 0.6bc	6.7 ± 0.4c	7.8 ± 0.2abc	7.7 ± 0.4abc	8.7 ± 0.5a	7.9 ± 0.7abc	8.1 ± 0.3ab*	6.9 ± 0.3bc	7.8 ± 0.1abc	8.2 ± 1.2abc	6.6 ± 0.5c	6.8 ± 0.4bc	8.5 ± 0.5a
TA g/L	88.6 ± 5.3a*	38.7 ± 9.4b	74.0 ± 16.4a*	42.9 ± 8.9b	62.7 ± 2.5ab	63.1 ± 1.1ab	73.4 ± 9.3a*	61.0 ± 4.5ab	63.2 ± 4.2ab*	55.8 ± 2.4ab	83.4 ± 3.2a*	59.2 ± 3.6ab	32.9 ± 9.2b	28.6 ± 5.9b	27.3 ± 3.3b
C^+ mmol	45.5 ± 5.1abc	42.9 ± 5.8abc	47.6 ± 8.8ab*	40.6 ±3.2bc	50.8 ± 1.5a*	42.3 ± 3.7abc	158.3 ± 106.4a	46 ± 3.6ab	56.4 ± 6.4a*	38.5 ± 1.1bc	53.7 ± 8.2a	46.8 ± 2.2ab	42 ± 5.4abc	42.3 ± 3.4abc	34.1 ± 2.8c

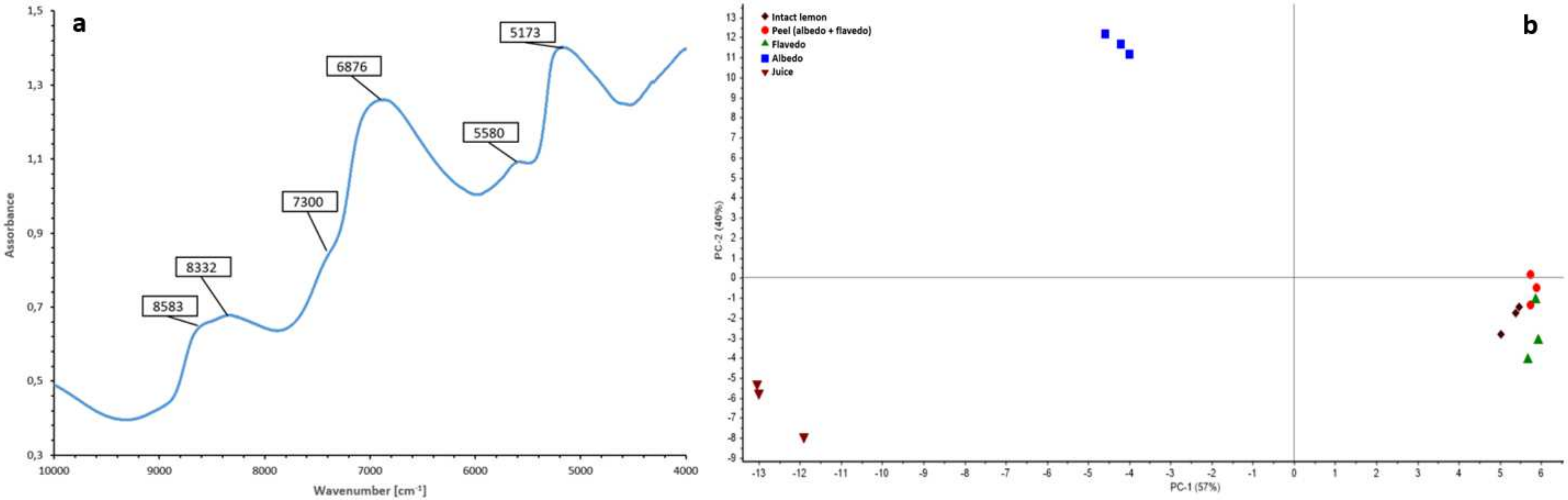


Figure 8 The typical NIR reflectance spectrum of intact lemon (a). PCA score plot of the NIR spectra of different fruit parts (b).

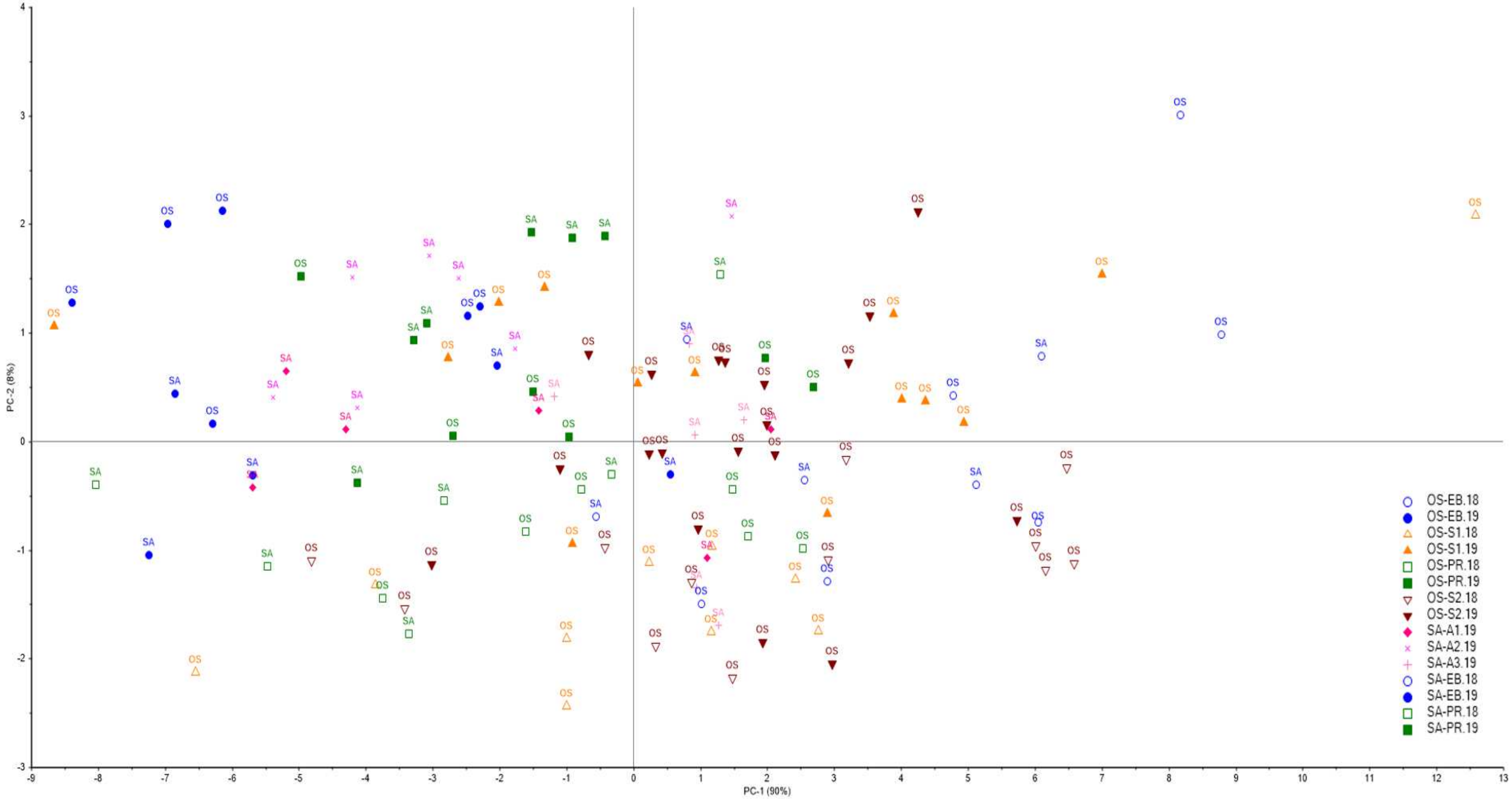


Figure 9 PCA score plot of raw NIR spectra. cv Ovale di Sorrento (OS), cv Sfusato Amalfitano (SA).

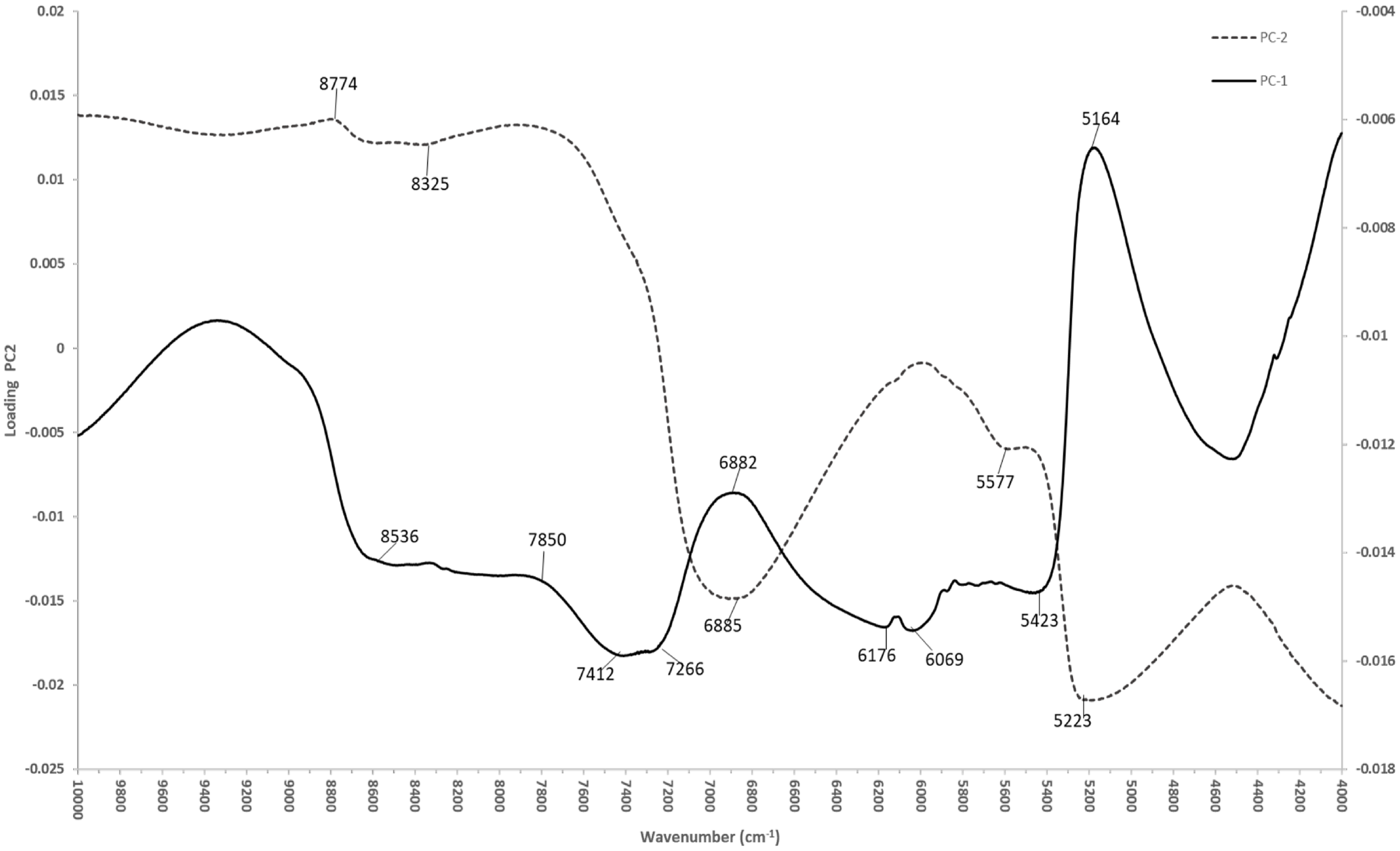


Figure 10 PC1 and PC2 loadings line plots of raw NIR spectra.

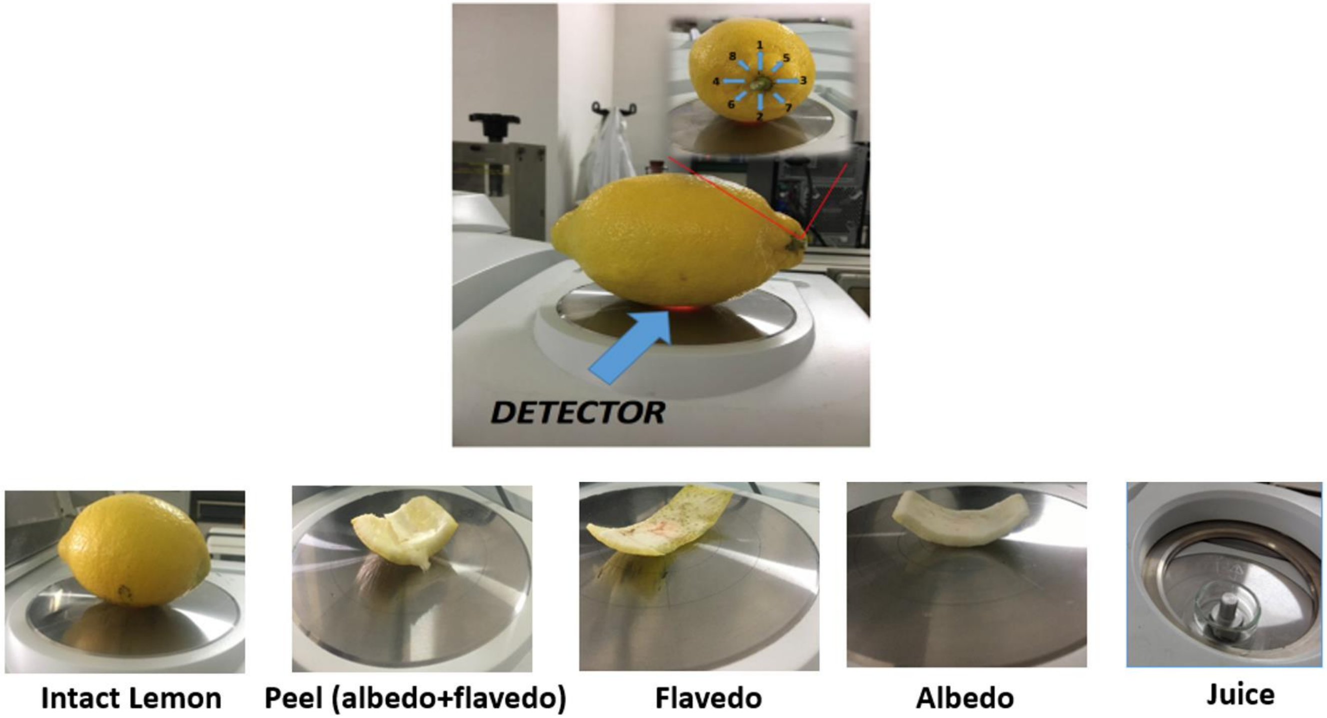
Table 3 Regression coefficient (R^2) and CV% (values below) by MLR models between principal components of NIR spectra and quality properties. Values in bold (R^2) are different from 0 at significance level α = 0.05

Cultivar	**Production years** ***(n° of samples)***	**L***	**a***	**b***	**THK (mm)**	**PW (%)**	**Juices (%)**	**pH**	**SSC (°Brix)**	**TA (g/L)**	**C^+ (cmol$^+$)**
OS & SA	**2018 (*45*)**	0.247	**0.562**	**0.813**	**0.542**	**0.716**	**0.444**	0.180	**0.776**	**0.367**	**0.365**
		4.2	*-569.0*	*5.5*	*14.9*	*12.4*	*11.1*	*3.9*	*4.9*	*15.3*	*21.7*
	2019 (*74*)	**0.580**	**0.308**	**0.387**	**0.265**	0.164	**0.459**	**0.205**	0.159	**0.419**	**0.373**
		2.4	*-64.5*	*7.6*	*29.6*	*17.9*	*16.7*	*6.5*	*10.9*	*24.9*	*9.7*
	2018-2019 (*119*)	**0.323**	**0.485**	**0.640**	**0.339**	**0.373**	**0.591**	**0.130**	**0.200**	**0.399**	**0.231**
		3.3	*-96.1*	*7.4*	*34.6*	*16.5*	*16.6*	*6.2*	*10.0*	*29.2*	*17.9*
OS	**2018 (*34*)**	0.315	0.389	**0.716**	**0.540**	**0.752**	**0.413**	0.263	**0.809**	0.384	**0.521**
		4.6	596.1	6.0	16.2	13.2	11.7	3.8	5.3	16.1	21.0
	2019 (*44*)	**0.549**	**0.326**	**0.460**	0.221	0.063	**0.521**	0.179	0.261	**0.495**	**0.468**
		3.0	-65.7	8.3	22.4	20.4	16.9	7.6	9.2	20.4	8.5
	2018-2019 (*78*)	**0.311**	**0.451**	**0.659**	**0.299**	**0.339**	**0.537**	**0.179**	**0.323**	**0.461**	**0.373**
		3.9	-126.3	7.6	26.5	18.2	16.2	7.0	8.9	25.2	16.6
SA	**2018 (*11*)**	0.896	0.910	**0.985**	0.764	0.863	0.771	0.863	0.896	0.947	0.322
		1.4	-69.8	3.4	15.6	6.6	11.9	2.8	1.9	6.5	19.1
	2019 (*30*)	**0.788**	**0.619**	**0.533**	0.295	**0.449**	**0.526**	0.404	**0.640**	**0.668**	**0.587**
		1.3	-53.4	5.9	27.6	9.5	13.0	3.2	9.2	24.1	9.6
	2018-2019 (*41*)	**0.722**	**0.662**	**0.697**	**0.354**	**0.526**	**0.585**	0.320	**0.408**	**0.405**	0.239
		1.4	-55.1	6.1	36.4	9.5	17.7	3.3	10.1	35.1	18.7

Table 4 Results of LDA of six different models. Number of spectra used for calibration and validation (Spectra); calibration and validation accuracy (Cal.Acc and Val.Acc); Accuracy ratio (Acc. Ratio)

ID Models	Groups	Calibration		Validation		
		Spectra	Cal Acc. (%)	Spectra	Val Acc. (%)	Acc. Ratio
C1	OS 18-19	120	69	36	63	0.92
	SA 18-19	68		14		
C2	OS-18	39	85	29	66	0.78
	SA-18	14		6		
C3	OS-19	72	78	16	63	0.81
	SA-19	48		14		
C4	OS-18	52	90	16	72	0.80
	SA-19	62		12		
P1	OS-PGI-18_19	88	84	20	72	0.86
	OS-NOPGI-18_19	36		12		
P2	OS-PGI-18	34	100	10	100	1.0
	OS-NOPGI-18	18		6		
P3	OS-PGI-19	52	100	10	81	0.81
	OS-NOPGI-19	18		6		
P4	SA-PGI-19	30	84	8	44	0.52
	SA-NOPGI-18_19	34		10		
P5	SA-PGI-19	30	100	8	74	0.74
	SA-NOPGI-19	20		4		

Appendix A. Supplementary data



Intact Lemon

Peel (albedo+flavedo)

Flavedo

Albedo

Juice

Figure S11. Schematic of NIR spectra acquisition of intact lemons and lemon parts. The numbers (upper figure) follow the order of acquisition.

Appendix A. Supplementary data

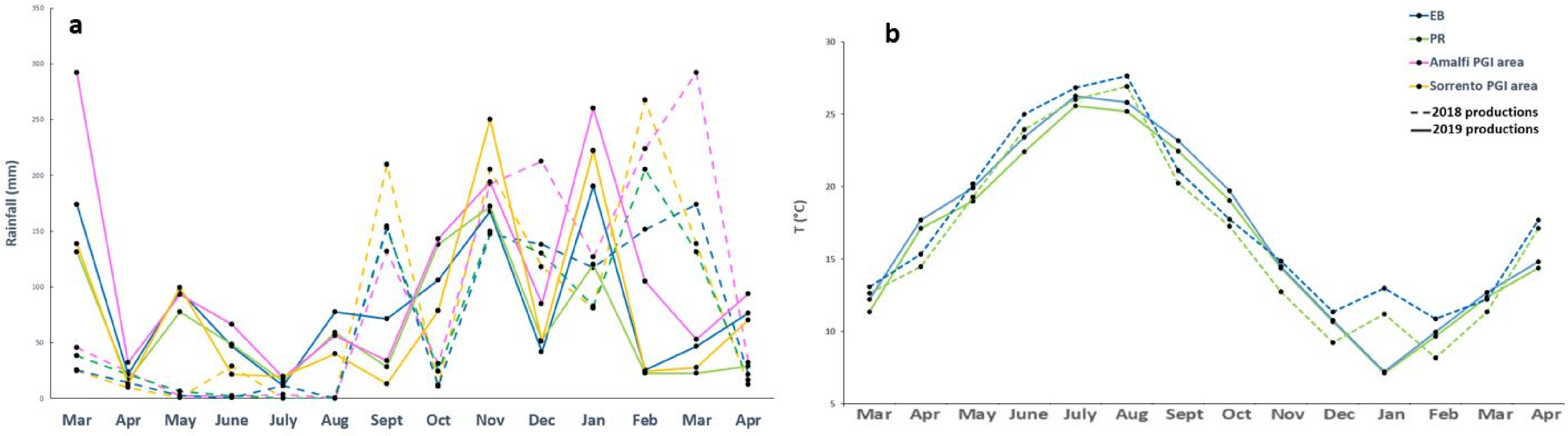
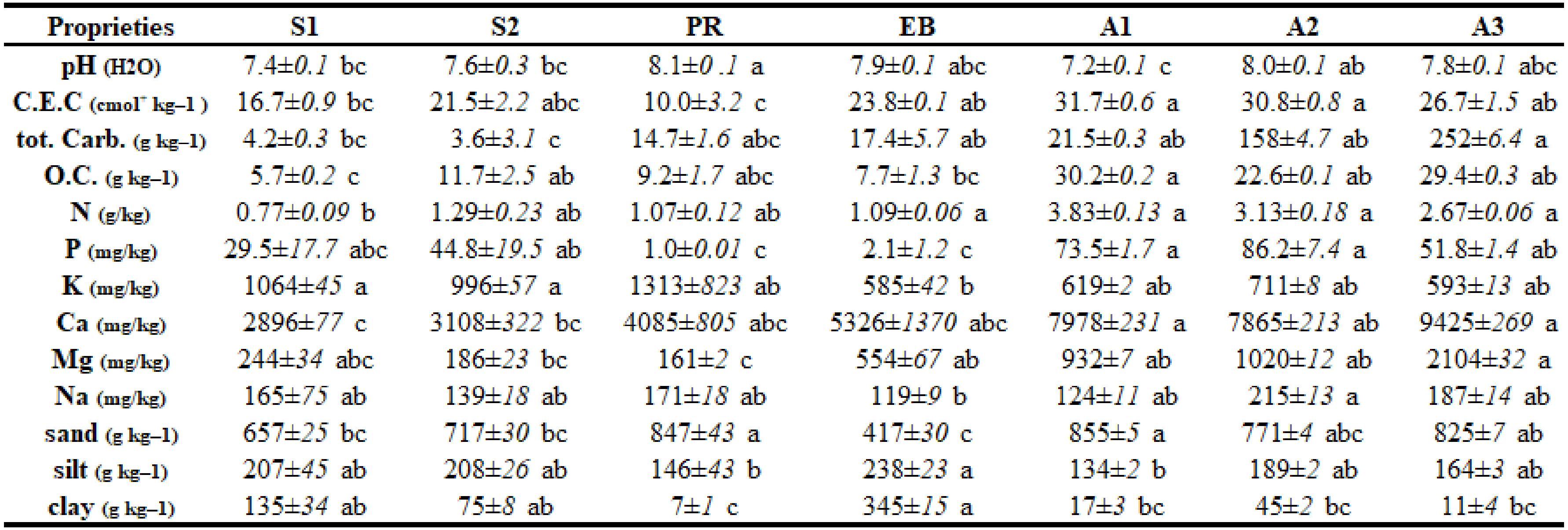


Figure S12. Monthly rainfalls (a) and temperatures (b) in sampling areas for production years 2018 and 2019.

Appendix A. Supplementary data

Table. S1. Chemical and physical characteristics of soil at 40-50 cm depth. Mean values in the same row followed by different letters indicate significant differences ($P < 0.05$) using nonparametric multiple tests ($p<0.05$, Kruskal–Wallis and post-hoc Dunn Test).



Proprieties	S1	S2	PR	EB	A1	A2	A3
pH **(H2O)**	7.4±*0.1* bc	7.6±*0.3* bc	8.1±*0 .1* a	7.9±*0.1* abc	7.2±*0.1* c	8.0±*0.1* ab	7.8±*0.1* abc
C.E.C **(cmol^{+} kg–1)**	16.7±*0.9* bc	21.5±*2.2* abc	10.0±*3.2* c	23.8±*0.1* ab	31.7±*0.6* a	30.8±*0.8* a	26.7±*1.5* ab
tot. Carb. **(g kg–1)**	4.2±*0.3* bc	3.6±*3.1* c	14.7±*1.6* abc	17.4±*5.7* ab	21.5±*0.3* ab	158±*4.7* ab	252±*6.4* a
O.C. **(g kg–1)**	5.7±*0.2* c	11.7±*2.5* ab	9.2±*1.7* abc	7.7±*1.3* bc	30.2±*0.2* a	22.6±*0.1* ab	29.4±*0.3* ab
N **(g/kg)**	0.77±*0.09* b	1.29±*0.23* ab	1.07±*0.12* ab	1.09±*0.06* a	3.83±*0.13* a	3.13±*0.18* a	2.67±*0.06* a
P **(mg/kg)**	29.5±*17.7* abc	44.8±*19.5* ab	1.0±*0.01* c	2.1±*1.2* c	73.5±*1.7* a	86.2±*7.4* a	51.8±*1.4* ab
K **(mg/kg)**	1064±*45* a	996±*57* a	1313±*823* ab	585±*42* b	619±*2* ab	711±*8* ab	593±*13* ab
Ca **(mg/kg)**	2896±*77* c	3108±*322* bc	4085±*805* abc	5326±*1370* abc	7978±*231* a	7865±*213* ab	9425±*269* a
Mg **(mg/kg)**	244±*34* abc	186±*23* bc	161±*2* c	554±*67* ab	932±*7* ab	1020±*12* ab	2104±*32* a
Na **(mg/kg)**	165±*75* ab	139±*18* ab	171±*18* ab	119±*9* b	124±*11* ab	215±*13* a	187±*14* ab
sand **(g kg–1)**	657±*25* bc	717±*30* bc	847±*43* a	417±*30* c	855±*5* a	771±*4* abc	825±*7* ab
silt **(g kg–1)**	207±*45* ab	208±*26* ab	146±*43* b	238±*23* a	134±*2* b	189±*2* ab	164±*3* ab
clay **(g kg–1)**	135±*34* ab	75±*8* ab	7±*1* c	345±*15* a	17±*3* bc	45±*2* bc	11±*4* bc

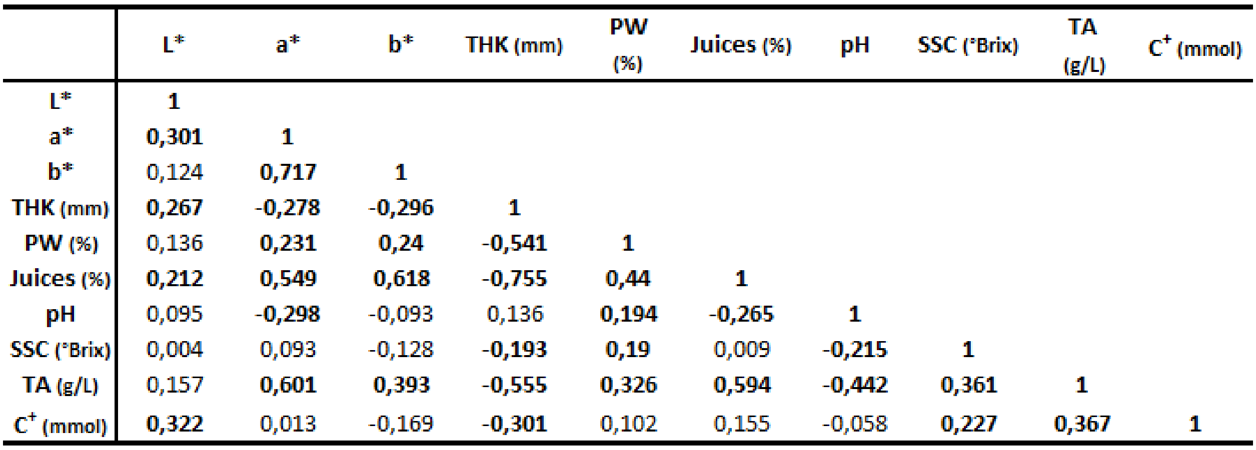
Appendix A. Supplementary data

Table S2. Results of two S-LDA models applied on quality properties according to the production years for EB and among PR and the two lemon orchards of Sorrento PGI area.

	Number of samples		λ Wilks	p-value	% Class	%Cross V	selected variables [a]		
	2018	2019							
EB	11	11	0,072	< 0.0001	100	100	**b****	**TA****	**a***
S1; S2; PR	34	44	0,220	< 0.0001	97	96	**TA****	**THK****	**pH***

a: Column from left to right by decreasing F values; ** p-value < 0.0001; * p-value < 0.001

Table S3. Spearman correlation analysis among quality proprieties in lemon juices. Peel chromaticity (L*, a*, b*); thickness of pericarp (THK); peel water (PW); juice yield (Juices%); soluble solid content (SSC), titratable acidity (TA), the molar concentration of positive charge from the cations (C+). Values in bold (R) are different from 0 at significance level α = 0.05.



	L*	**a***	**b***	**THK (mm)**	**PW (%)**	**Juices (%)**	**pH**	**SSC (°Brix)**	**TA (g/L)**	**C^+ (mmol)**
L*	**1**									
a*	**0,301**	**1**								
b*	0,124	**0,717**	**1**							
THK (mm)	**0,267**	**-0,278**	**-0,296**	**1**						
PW (%)	0,136	**0,231**	**0,24**	**-0,541**	**1**					
Juices (%)	**0,212**	**0,549**	**0,618**	**-0,755**	**0,44**	**1**				
pH	0,095	**-0,298**	-0,093	0,136	**0,194**	**-0,265**	**1**			
SSC (°Brix)	0,004	0,093	-0,128	**-0,193**	**0,19**	0,009	**-0,215**	**1**		
TA (g/L)	0,157	**0,601**	**0,393**	**-0,555**	**0,326**	**0,594**	**-0,442**	**0,361**	**1**	
C^+ (mmol)	**0,322**	0,013	-0,169	**-0,301**	0,102	0,155	-0,058	**0,227**	**0,367**	**1**

Chapter 4: Feasibility study of authentication PGI Italian lemon fruits by near-infrared (NIR) spectroscopy: the case study of Limone di Sorrento PGI.

1 Introduction

In Mediterranean countries, citrus production is particularly developed due to the favorable climatic conditions. The peculiarity of Mediterranean citrus is its prevalent propensity to be consumed fresh, unlike citrus production in other countries, such as Brazil and the United States, where the majority of the fruits are processed for the production of concentrated juice (Giare & Giuca, 2008). The cultivation of lemons is well-known and praised for its quality in southern Italy, mainly in Sicily, Calabria and Campania. In Campania, two high-quality products have been certified and registered as Protected Geographical Indication (PGI) products, i.e. *Limone di Sorrento* (Reg. CE n. 2446/2000) and *Limone Costa D'Amalfi* (Reg. CE n. 1356/2001). These PGI products are specifically marketed for fresh consumption (intact lemon) or consumption of the processed product (fresh juice). *Limone di Sorrento* refers to the Citrus lemon (L.) *Burm.f. Femminello Ovale di Sorrento*, also known as *Limone di Massa Lubrense* or *Massese*, and is produced in the area of the Sorrento Peninsula and Capri Island. However, this genotype is not the only genotype that grows in the mentioned PGI area. As reported on the product specification, 15% cultivation of other cultivars and genotypes are also allowed. In Sorrento Peninsula, the lemon orchards are covered with "pergolato sorrentino", a system made by vertical wooden canes, preferably chestnut, covering the whole lemon grove with reeds known as "pagliarelle". This traditional cultivation system is characteristic and unique to lemons produced in the Sorrento and Amalfi Penisula regions. Apart from the higher costs associated with the traditional production of lemons, the name also carries a story and tradition that is valued by consumers who are willing to pay a high price/premium for the product. Due to this higher selling point, there is a great economic incentive linked to the misuse and mislabelling of PGI lemons, where fraudsters mislead consumers through the sale of generic/cheaper lemons as PGI products. Given the difficulty of subjectively distinguish different valued lemons, it is vital to develop a suitable fraud detection and screening technique. Several analytical techniques have been used to authenticate citrus fruit. One of the most prevalent is the determination of the mineral composition of juices from orange fruits (Simpkins et al., 2000), clementine fruits (Benabdelkamel et al., 2012) and several lemon juices (Gaiad et al., 2016; Pellerano et al., 2008). Other methods have also been used for the authentication of citrus fruits, such as: the analysis of aroma-active compounds (volatiles) by Headspace Solid-Phase

Micro-Extraction coupled with Gas Chromatography-Mass Spectrometry (HS-SPME-GC-MS) (Zhang et al., 2019); metabolic profiling by Nuclear Magnetic Resonance (NMR) spectroscopy (Salazar et al., 2018); 3D-front-face fluorescence (Ammari et al., 2015) and the use of rapid and non-destructive near-infrared (NIR) spectroscopy combined with other techniques as reported in Amenta et al. (2016). NIR spectroscopy can be used as rapid screening technique owing to its easy use and low cost, but above all for its on-line application in the food industry (Jha et al., 2014). The use of NIR spectroscopy to determine a food product's geographical origin, has been successfully applied on many products such as honey (Latorre et al., 2013)(Woodcock et al., 2009); *dong quai* or female ginseng (Woo et al., 2005); wheat (Zhao et al., 2013); wine (Cozzolino et al., 2011), (Liu et al., 2006); olive oils (Jiménez-carvelo et al., 2019), (Hirri et al., 2016); sea cucumber (Guo et al., 2018). The success of this technique lies in the fact that the NIR spectra of each sample is unique and related to the vibrational frequency of functional groups present in the product. Ultimately, this vibrational frequency of all functionals groups produces a unique molecular 'fingerprint' that can be used to confirm the identity of a sample (Luykx & van Ruth, 2008). Typical NIR spectra (4000–10000 cm^{-1}) are characterized by absorption bands that are combination bands (overtone) originated, in the infrared region, of vibrational and rotational transitions of molecular bonds (Nicolaï et al., 2007). The introduction of the Fourier-transform near-infrared technique (FT-IR) has increased the use of IR, due to measurements all of the infrared frequencies simultaneously, the technique allows very rapid screening of samples (Luykx & van Ruth, 2008). Usually, the spectra are very similar, they are dominated by the water (overtone bands of the OH-bonds at 760, 970 and 1450nm) (Polesello, Giangiacomo, & Dull, 1981). For these reasons, the NIR spectroscopy spectra require multivariate statistical analyses to extract useful information from spectra (Blanco & Villarroya, 2002).This study aims to apply NIR spectroscopy as a rapid method to authenticate *Limone di Sorrento* (cultivar *Ovale di Sorrento*) intended for fresh consumption (intact lemon) and consumption as processed product (fresh juice). The study will seek to verify if the spectroscopic profile can be used to discriminate PGI lemons from Non-PGI lemons grown in areas geographically close to each other (Campania Region). It will thereby develop a tool that is fast and economic to verify the PGI origin of lemons and juices for the small farms within the large PGI area, and to distinguish PGI lemons from lemons of other origins.

2 Materials and methods

2.1 Collection of fruit samples and study area

Fifty-six fully ripened lemon fruits of *Citrus limon* (L.) *Burm.f. Femminello Ovale di Sorrento* from five different sampling sites in southern Italy were collected. Forty-four of the collected lemons originate from three different farms of the *Limone di Sorrento* PGI area: 13 lemons from the municipality of Massa Lubrense (MS-PGI farm) (40°36’28’’ N; 14°20’36’’E); 26 lemons from the municipality of Piano di Sorrento (PZ-PGI farm) (40°37’51’’N; 14°24’22’’ E); and 6 lemons from the municipality of Sorrento (SR-PGI farm) (40°36’59’’N; 14°23’10’’E). The remaining 12 lemons originated from two Non-PGI farms located in the *Limone di Sorrento* PGI area: 6 lemons from the germplasm field of the Department of Agricultural Sciences, University of Naples Federico II (PR-NOPGI farm) (40°48’54’’N; 14°20’49’’ E), and 6 lemons from the germplasm field of The Regional experimental agricultural farm "Improsta" at Eboli (EB-NOPGI farm), (40°33’21’’ N; 14°58’53’’E). All lemons (samples) and trees, in each farm, were randomly selected, to ensure the greatest variability among the farms (inside). Within each farm, 5-7 lemons were collected per tree. After sampling, all lemon fruits were first cleaned in tap-water and then washed in deionized water. The fruits were stored for one week at 4°C during which the NIR spectra of the surface (i.e. peel) was obtained. The lemon fruits were then cut and squeezed to extract the juice using a domestic juicer. The extracted juices were centrifugated at 2000 *g* for 20 min to separate the liquid from the pulp to minimize the effect of light scatter caused by the suspended particles during the NIR spectra acquisition step (Rinnan et al., 2009).

2.2 NIR spectrum acquisition

Near-infrared (NIR) spectra were acquired by a PerkinElmer Fourier-transform near-infrared, FT-IR/NIR spectrometer (FRONTIER) equipped with a NIRA accessory in reflectance mode, over the absorbance range 10000 to 4000 cm^{-1} (1000 to 2500 nm) with an interval of 4 cm^{-1} with a scan time of 64 scans (samples and background). The NIR spectra were acquired within one week after sample collection. During analysis, the fruits were stored in a refrigerator at a temperature of 4°C. For each lemon, along its equatorial circumference, at each eighth of a turn, a total of eight NIR spectra were acquired. The fruits were placed in direct contact with the detector. After the squeeze and centrifuge step, 2 ml lemon juice were transferred into a glass petri dish (diameter of 30 mm and depth of 10 mm) after which two spectra were acquired for each lemon juice sample. Samples were analysed in duplicate. All spectra were acquired at room temperature (20-25°C).

2.3 Chemometric analysis

The eight spectra of each intact lemon were randomly divided into two groups and mediated. For each lemon, two spectra (per intact lemon/peel and juice, respectively) were used for data analysis taking into account the analytical variability that is needed to minimize tissue heterogeneities, instrumental noise, ambient effects, and other sources of variability as reported in (Nicolaï et al., 2007). The data analysis, spectra acquisition, and averaging was carried out by PerkinElmer Spectrum software (Version 10.03.06, 2011 PerkinElmer, Inc.). The Unscrambler X10.4 software (Camo Software AS., Oslo, Norway) was used to pre-process the spectral data and to perform multivariate data analysis. All spectra (n = 112) were divided into two groups: calibration set (n = 58) and validation set (n = 54). The spectra of the PGI lemons (total number = 88), were randomly divided into two groups: 34 as the calibration set and 54 as the validation set. Hence, the complete calibration set was composed of 24 Non-PGI lemons spectra (12 PR-NOPGI, 12 EB-NOPGI) and 34 PGI lemons spectra (10 MS-PGI, 18 PZ-PGI, 6 SR-PGI), while the validation set contained only the spectra of 54 PGI lemons (16 MS-PGI, 32PZ-PGI, 6 SR-PGI). The pre-processing of NIR spectra is the necessary step before the chemometric analysis (Rinnan et al., 2009). Multiplicative Scatter Correction (MSC), Standard Normal Variate (SNV) and derivate correction (first and second derivate) are the three methods most widely used (Roggo et al., 2007; Blanco & Villarroya, 2002). The derivate methods broadly applied are Norris-Williams derivatives and Savitzky-Golay (SG) polynomial derivates (Rinnan et al., 2009). The pre-processing tecniques were applied to decrease the detrimental effect on the signal-to-noise ratio and improve both the explorative analysis and the classification models. For screening the natural grouping of samples, principal component analysis (PCA) was carried out, while partial least square discriminant analysis (PLS-DA) (±0.5 cutoff criteria), and three different methods of discriminant analysis: Linear Discriminant Analysis (LDA), Quadratic discriminant analysis (QDA) and Mahalanobis discriminant analysis MDA methods were used for classification of the selected samples.

3 Results and discussion

3.1 Description of NIR spectra

All the acquired absorbance spectra for the lemon juice samples are shown in Fig.1a (raw spectra). The two highest absorptions were observed around 7200-6600 cm^{-1} (second overtone region) and 5300-4800 cm^{-1} (first overtone region). The presence of the highest absorptions in these regions could justify not only the presence of water, showing strong adsorption and known as the major constituent of lemon juice but also of phenolic and aromatic compounds and carbohydrates. The regions with high absorptions for water is expected to dominate since fruit contains 80–90% water (Gòmez, He, & Annia Garcìa Pereira, 2006; (Miyamoto & Kitano, 2009). In the same way, the spectra of the intact lemons (where the surfaces/peels were scanned) (Fig.1b) are very similar to those shown in (Magwaza et al., 2014) for 'Nules Clementine' citrus. For both the juice and intact lemon, high absorptions were observed in the last part of the spectra, around 4000 cm^{-1} (combination band region), that coincides with the high presence of carbohydrates. As shown in Fig.1, the spectral dataset of intact lemons display an improved division (parting of spectra from different classes) for the spectra of PGI lemons (blue line) and Non-PGI lemons (red line) than that observed for the lemon juices. This lack of division of the juice samples could be due to their higher water content, and subsequent higher absorption bands, than the surface or peel of the intact lemons.

3.2 Exploratory analysis

To explore the grouping of the spectra in the calibration set, principal component analysis (PCA) was performed using the raw spectra and two different pre-processing techniques (SNV and SNV+ SG, first derivative, 2nd order polynomial, 15 points). The samples were visualized based on the selection of two different groups for the PCA to verify if the trend in sample grouping was due to the geographical origin of samples. The first group contained two groups: PGI and Non-PGI samples; and the second group three sub-groups: (MS+PZ+SR)-PGI, PR-NOPGI and EB-NOPGI. In Fig.2, the PCA score plots with raw spectra (Fig.2a), SNV (Fig.2b), SNV+ SG (Fig.2c) of lemon juices are shown. The score plots of the raw spectra (Fig.2a and 2a^1) showed no distinct grouping of the samples for groups. The slight grouping could be related to undesired scatter effects, most probably for the strong presence of water, and that it could have caused a negative influence on the extracted information. The first four PCs explained 99% of the total variance (PC1 76%; PC2 22%; PC3 1%; PC4 1%). The score plots of SNV processed spectra (Fig.2b and Fig.2b^1) showed an improved grouping than the raw spectra plot, better with two groups than three groups because there is no clear difference between the Non-PGI groups. The first four PCs explained 99% of the total variance (PC1 88%; PC2 9%; PC3 1%; PC4 1%). Fig.2c and 2c^1 show the PCA score plots of the SNV+SG processed spectra. The first three PCs explained 20% of the total variance (PC1 10%; PC2 6%; PC3 4%) and showed a clear grouping between the three different areas, PR-NOPGI; EB-NOPGI and (MS+PZ+SR)-PGI, but not between the PGI groups (data not shown) (Fig. S1, provided in Appendix A. Supplementary data). This lack of grouping, for single farms, could be due to the similar conditions of growing and soils, of the PGI samples (see Chapter 1). For the NIR spectra of the intact lemons, the same statistical analyses were applied. The first three PCs, of the PCA using the raw spectra, explained 100% of the total variance (PC1 95%; PC2 4%; PC3 1%). The score plots (Fig.3a and 3a^1) showed a slight grouping trend according to the origin, better than the grouping of the score plots (four PCs with 99% of total variance) of the SNV processed spectra (Fig.3b and 3b^1) and the score plots (five PCs with 94% of total variance) of the SNV+SG processed spectra; where distinct groupings were not apparent. The loadings line plots are shown in Fig. 2 (a^2, b^2, c^2) for the lemon juice samples, and in Fig. 3 (a^2, b^2, c^2) for the whole lemons. The loadings show the importance of variables (wavenumbers) on the natural grouping of samples. Overall the main variables correlated with the natural separation of samples (loadings $\neq 0$) as absorption bands for both whole lemons and lemon juice samples were observed around 8800-8600 cm^{-1} (second overtone of C-H), 7300-7000 cm^{-1} (first overtone C-H combination bands), 5300-5100 cm^{-1} (first overtone C-H combination bands) and 4000 cm^{-1} (combination bands region) for the first PC (the highest percentage value of variability explained).

3.3 PGI and Non-PGI discrimination

Based on the results of the exploratory analysis, the two groups of spectra, PGI and Non-PGI (intact fruits and lemon juices), were classified using three different methods of LDA, i.e. linear, quadratic, and Mahalanobis. The models were built as described in section 2.3. The validation test was performed only for PGI lemons, and these samples were randomly selected from the three different farms in the PGI areas. The same pre-processing techniques used before (sections 3.2) were applied. The results of the discriminant models are shown in Table 1. The LDA models were carried out using three principal components. The best LDA model for lemon juices was obtained with the quadratic method using SNV+SG pre-processed spectra (1st derivative, 2nd order polynomial, 15 points) (Table 1). The results show 98% accuracy, where 100% of the PGI samples and 96% of Non-PGI samples were correctly classified, while 98% of the validation set was correctly classified. Moreover, the best LDA model for intact lemon was obtained with the linear method using the raw spectra. The results show 83% correct classification for Non-PGI samples and 88% for PGI samples (86% of total accuracy). The result of PGI validation was 89% correctly classified. Fig. S1 (provided in Appendix A. Supplementary data) shows the main variables correlated with the discrimination of samples. The water absorption bands, around 5150 cm^{-1} and 7000 cm^{-1} (Chen et al., 2009), seemed not to have a high influence on discrimination because it correlated with the 3rd PC in the set of lemon juices (Fig. S1a), and 2nd and 3rd PCs in the set of intact lemons (Fig.S1b). The most intensive bands, in both cases, correspond to overtones of the carbonyl group ($-CH_2$ and $–CH$) possibly as a result of the presence of polyphenols, volatile acid and some aroma compounds (Chen et al., 2009), which are dominant compounds in lemons. These differences could also be explained by different climatic and cultivation conditions where traditional cultivation is practiced in the PGI area and not in the Non-PGI area.

3.4 Classification of geographical origin

To understand the unique characteristics of lemons grown in the PGI area, it is necessary to evaluate not only the differences between PGI and Non-PGI samples (section 3.3) but also the differences between the different production farms inside the PGI areas: MS-PGI, PZ-PGI and Sr-PGI and the two farms Non-PGI: PR-PGI and EB-PGI taken separately. Although the farms investigated in this study are geographically close to one another, there could still be differences due to different growth conditions (soil and climatic condition) that could lead to unique profiles in lemon composition, i.e aroma as reported in Zhang et al., (2019). The PLS-DA was performed to discriminate the samples that came from different farms. The best PLS-DA was chosen with high R^2 (Pearson determination coefficient), low SEC (standard error of calibration), low SEP (standard error of prediction), a high percentage of spectra correctly classified. The true number of regression factors were determined according to the minimum SEP as too many numbers of regression factors would result in overfitted models. Also, large differences between SEC and SEP would lead to modelling noise by affecting prediction (Gòmez et al., 2006; Shao et al., 2007). Modelling using PLS-DA was first evaluated with three groups: PGI (MS+PZ+SR-PGI), PR-NOPGI, EB-NOPGI and then with five groups (MS-PGI, PZ-PGI, SR-PGI, PR-NOPGI, EB-NOPGI). The PLS-DA results are shown in Table 2. Among all models tested, the best PLS-DA model was obtained with SNV+SG pre-processed spectra (1st derivative, 2nd order polynomial, 15 points) for both lemon juices and intact lemons. The models with three groups showed better results than the models with five groups. The best model for lemon juices (with three groups) produced 100% correct classification of samples in the calibration set and 91% correct classification of samples in the validation set. In the same way, the best model of intact lemons produced 74% correct classification of samples in the calibration set and 65% correct classification of samples in the validation set. Upon closer investigation, these models were able to discriminate the PGI lemons from the two Non-PGI lemons, but it could not successfully separate the PGI lemon samples that originated from the three PGI farms. The major problem for the modelling was the overlap in the three PGI farms (Table S1 and S2, provided in Appendix A. Supplementary data). For lemon juices from MS-PGI, only 63% was correctly classified, PZ-PGI only 53%, and for SR-PGI no samples were correctly classified. Furthermore, a few samples for MS-PGI (7%) were classified as PR-NOPGI. In the same way, for intact lemons from MS-PGI, only 19%, in the PZ-PGI only 28% and for SR-PGI only 17% were correctly classified. Few samples for MS-PGI (31%) and PZ-PGI (9%) were classified as PR-NOPGI. The high percentage of misclassification could be due to the different cultivation and climatic condition of PGI lemons which make the product profile unique, but also complex. In the end, similar results, in terms of percent of correct classification, were found for black tea (Ren et al., 2013) with percent of accuracy to 94.3%; were found in Australian and Spanish Tempranillo wine (Liu et al., 2006) by PLS-DA (93.5%) and LDA (91.6%); and were found in Spanish wine vinegar (>90%) by PLS-DA (Ríos-Reina et al., 2018).

4 Conclusions

The use of the NIR spectral fingerprinting technique has shown satisfactory results to discriminate the PGI lemon samples from the Non-PGI lemon samples produced in two areas geographically close to one another (Campania Region). The results indicate that the LDA and PLS-DA classification, using the NIR spectra of the lemon juice samples, performed better (higher level of discrimination) than the models developed using the NIR spectra of intact lemon samples. The discrimination models, so built, could be subjected to misclassifications when they will be validated with samples of different years (different growing conditions) or with samples that came from other farms also in the PGI area. For this, it is necessary that these discrimination models must be implemented year to year with samples not only from the same orchards but also with samples that come from different orchards inside the PGI area. The results have shown that there are common characteristics among the farms in the *Limone di Sorrento* PGI area, different from the Non-PGI areas that could be related to the unique traditional cultivation methods of the PGI area. Future application with other techniques could provide more evidence to support this relationship (cultivation environment/products). Extending the database with further representative samples, and farms, in all PGI area is necessary to improve the statistical models to provide lower error rates for the authentication of lemon fruits. Considering the low price of the device, the non-destructive, fast and easy application (samples without or minimal preparation), ease of measurement, and flexibility of the measurement set-up, NIR spectroscopy could serve as a viable tool to distinguish the lemon fruits according to origin, and thereby also act as a technical control measure to help detect food fraud.

References

Amenta, M., S.,Fabroni, Costa, C., Rapisarda, P., Fabroni, S., Costa, C., & Rapisarda, P. (2016). Traceability of “Limone di Siracusa PGI” by a multidisciplinary analytical and chemometric approach. *Food Chemistry*, *211*, 734–740. https://doi.org/10.1016/j.foodchem.2016.05.119

Ammari, F., Redjdal, L., & Rutledge, D. N. (2015). Detection of orange juice frauds using front-face fluorescence spectroscopy and Independent Components Analysis. *Food Chemistry*, *168*, 211–217. https://doi.org/10.1016/j.foodchem.2014.06.110

Benabdelkamel, H., Di Donna, L., Mazzotti, F., Naccarato, A., Sindona, G., Tagarelli, A., & Taverna, D. (2012). Authenticity of PGI “clementine of Calabria” by multielement fingerprint. *Journal of Agricultural and Food Chemistry*, *60*(14), 3717–3726. https://doi.org/10.1021/jf2050075

Blanco, M., & Villarroya, I. (2002). NIR spectroscopy: A rapid-response analytical tool. *TrAC - Trends in Analytical Chemistry*, *21*(4), 240–250. https://doi.org/10.1016/S0165-9936(02)00404-1

Chen, Q., Zhao, J., & Lin, H. (2009). Study on discrimination of Roast green tea (Camellia sinensis L.) according to geographical origin by FT-NIR spectroscopy and supervised pattern recognition. *Spectrochimica Acta - Part A: Molecular and Biomolecular Spectroscopy*, *72*(4), 845–850. https://doi.org/10.1016/j.saa.2008.12.002

Cozzolino, D., Cynkar, W. U., Shah, N., & Smith, P. A. (2011). Can spectroscopy geographically classify Sauvignon Blanc wines from Australia and New Zealand? *Food Chemistry*, *126*(2), 673–678. https://doi.org/10.1016/j.foodchem.2010.11.005

Gaiad, J. E., Hidalgo, M. J., Villafañe, R. N., Marchevsky, E. J., & Pellerano, R. G. (2016). Tracing the geographical origin of Argentinean lemon juices based on trace element profiles using advanced chemometric techniques. *Microchemical Journal*, *129*, 243–248. https://doi.org/10.1016/j.microc.2016.07.002

Gòmez, A. H., He, Y., & Annia Garcìa Pereira. (2006). *Non-destructive measurement of acidity , soluble solids and firmness of Satsuma mandarin using Vis / NIR-spectroscopy techniques*. *77*, 313–319. https://doi.org/10.1016/j.jfoodeng.2005.06.036

Guo, X., Cai, R., Wang, S., Tang, B., Li, Y., & Zhao, W. (2018). Non-destructive geographical traceability of sea cucumber (*Apostichopus japonicus*) using near infrared spectroscopy combined with chemometric methods. *Royal Society Open Science*, *5*(1), 170714. https://doi.org/10.1098/rsos.170714

Hirri, A., Bassbasi, M., & Platikanov, S. (2016). FTIR Spectroscopy and PLS-DA Classification and Prediction of Four FTIR Spectroscopy and PLS-DA Classification and Prediction of Four Commercial Grade Virgin Olive Oils from Morocco. *Food Analytical Methods*, *9*(August), 974–981. https://doi.org/10.1007/s12161-015-0255-y

Jha, S. N., Narsaiah, K., Jaiswal, P., Bhardwaj, R., Gupta, M., Kumar, R., & Sharma, R. (2014). Nondestructive prediction of maturity of mango using near infrared spectroscopy. *Journal of Food Engineering*, *124*, 152–157. https://doi.org/10.1016/j.jfoodeng.2013.10.012

Jiménez-carvelo, A. M., Lozano, V. A., & Olivieri, A. C. (2019). Comparative chemometric analysis of fluorescence and near infrared spectroscopies for authenticity con firmation and geographical origin of Argentinean extra virgin olive oils. *Food Control*, *96*(August 2018), 22–28. https://doi.org/10.1016/j.foodcont.2018.08.024

Latorre, C. H., Crecente, R. M. P., Martín, S. G., & García, J. B. (2013). A fast chemometric procedure based on NIR data for authentication of honey with protected geographical indication. *Food Chemistry*, *141*(4), 3559–3565. https://doi.org/10.1016/j.foodchem.2013.06.022

Liu, L., Cozzolino, D., Cynkar, W. U., Gishen, M., & Colby, C. B. (2006). Geographic Classification of Spanish and Australian Tempranillo Red Wines by Visible and Near-Infrared Spectroscopy Combined with Multivariate Analysis. *Journal of Agricultural and Food Chemistry*, *54*, 6754–6759.

Luykx, D. M. A. M., & van Ruth, S. M. (2008). An overview of analytical methods for determining the geographical origin of food products. *Food Chemistry*, *107*(2), 897–911. https://doi.org/10.1016/j.foodchem.2007.09.038

Magwaza, L. S., Landahl, S., Cronje, P. J. R., Nieuwoudt, H. H., Mouazen, A. M., Nicolaï, B. M., Terry, L. A., & Opara, U. L. (2014). The use of Vis/NIRS and chemometric analysis to predict fruit defects and postharvest behaviour of "Nules Clementine" mandarin fruit. *Food Chemistry*, *163*, 267–274. https://doi.org/10.1016/j.foodchem.2014.04.085

Miyamoto, K., & Kitano, Y. (2009). Non-Destructive Determination of Sugar Content in Satsuma Mandarin Fruit by near Infrared Transmittance Spectroscopy. *Journal of Near Infrared Spectroscopy*, *3*(4), 227–237. https://doi.org/10.1255/jnirs.73

Nicolaï, B. M., Beullens, K., Bobelyn, E., Peirs, A., Saeys, W., Theron, K. I., & Lammertyn, J. (2007). Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: A review. *Postharvest Biology and Technology*, *46*(2), 99–118. https://doi.org/10.1016/j.postharvbio.2007.06.024

Pellerano, R. G., Mazza, S. S., Marigliano, R. A., & Marchevsky, E. J. (2008). Multielement analysis of argentinean lemon juices by instrumental neutronic activation analysis and their classification according to geographical origin. *Journal of Agricultural and Food Chemistry*, *56*(13), 5222–5225. https://doi.org/10.1021/jf073555n

Ren, G., Wang, S., Ning, J., Xu, R., Wang, Y., Xing, Z., Wan, X., & Zhang, Z. (2013). Quantitative analysis and geographical traceability of black tea using Fourier transform near-infrared spectroscopy (FT-NIRS). *Food Research International*, *53*(2), 822–826. https://doi.org/10.1016/j.foodres.2012.10.032

Rinnan, Å., Berg, F. van den, & Engelsen, S. B. (2009). Review of the most common pre-processing techniques for near-infrared spectra. *TrAC - Trends in Analytical Chemistry*, *28*(10), 1201–1222. https://doi.org/10.1016/j.trac.2009.07.007

Ríos-Reina, R., García-González, D. L., Callejón, R. M., & Amigo, J. M. (2018). NIR spectroscopy and chemometrics for the typification of Spanish wine vinegars with a protected designation of origin. *Food Control*, *89*, 108–116. https://doi.org/10.1016/j.foodcont.2018.01.031

Roggo, Y., Chalus, P., Maurer, L., Lema-martinez, C., & Jent, N. (2007). *A review of near infrared spectroscopy and chemometrics in pharmaceutical technologies*. *44*, 683–700. https://doi.org/10.1016/j.jpba.2007.03.023

Salazar, M. O., Pisano, P. L., González Sierra, M., & Furlan, R. L. E. (2018). NMR and multivariate data analysis to assess traceability of argentine citrus. *Microchemical Journal*, *141*(May), 264–270. https://doi.org/10.1016/j.microc.2018.05.037

Shao, Y., He, Y., Gómez, A. H., Pereir, A. G., Qiu, Z., & Zhang, Y. (2007). Visible/near infrared spectrometric technique for nondestructive assessment of tomato “Heatwave” (Lycopersicum esculentum) quality characteristics. *Journal of Food Engineering*, *81*(4), 672–678. https://doi.org/10.1016/j.jfoodeng.2006.12.026

Simpkins, W. A., Louie, H., Wu, M., Harrison, M., & Goldberg, D. (2000). Trace elements in Australian orange juice and other products. *Food Chemistry*, *71*(4), 423–433.

https://doi.org/10.1016/S0308-8146(00)00150-3

Woo, Y. A., Kim, H. J., Ze, K. R., & Chung, H. (2005). Near-infrared (NIR) spectroscopy for the non-destructive and fast determination of geographical origin of Angelicae gigantis Radix. *Journal of Pharmaceutical and Biomedical Analysis*, *36*(5), 955–959. https://doi.org/10.1016/j.jpba.2004.08.037

Woodcock, T., Downey, G., & O'Donnell, C. P. (2009). Near infrared spectral fingerprinting for confirmation of claimed PDO provenance of honey. *Food Chemistry*, *114*(2), 742–746. https://doi.org/10.1016/j.foodchem.2008.10.034

Zhang, W., Chen, T., Tang, J., Sundararajan, B., & Zhou, Z. (2019). Tracing the production area of citrus fruits using aroma-active compounds and their quality evaluation modles. *Journal of the Science of Food and Agriculture*, *September*. https://doi.org/10.1002/jsfa.10026

Zhao, H., Guo, B., Wei, Y., & Zhang, B. (2013). Near infrared reflectance spectroscopy for determination of the geographical origin of wheat. *Food Chemistry*, *138*(2–3), 1902–1907. https://doi.org/10.1016/j.foodchem.2012.11.037

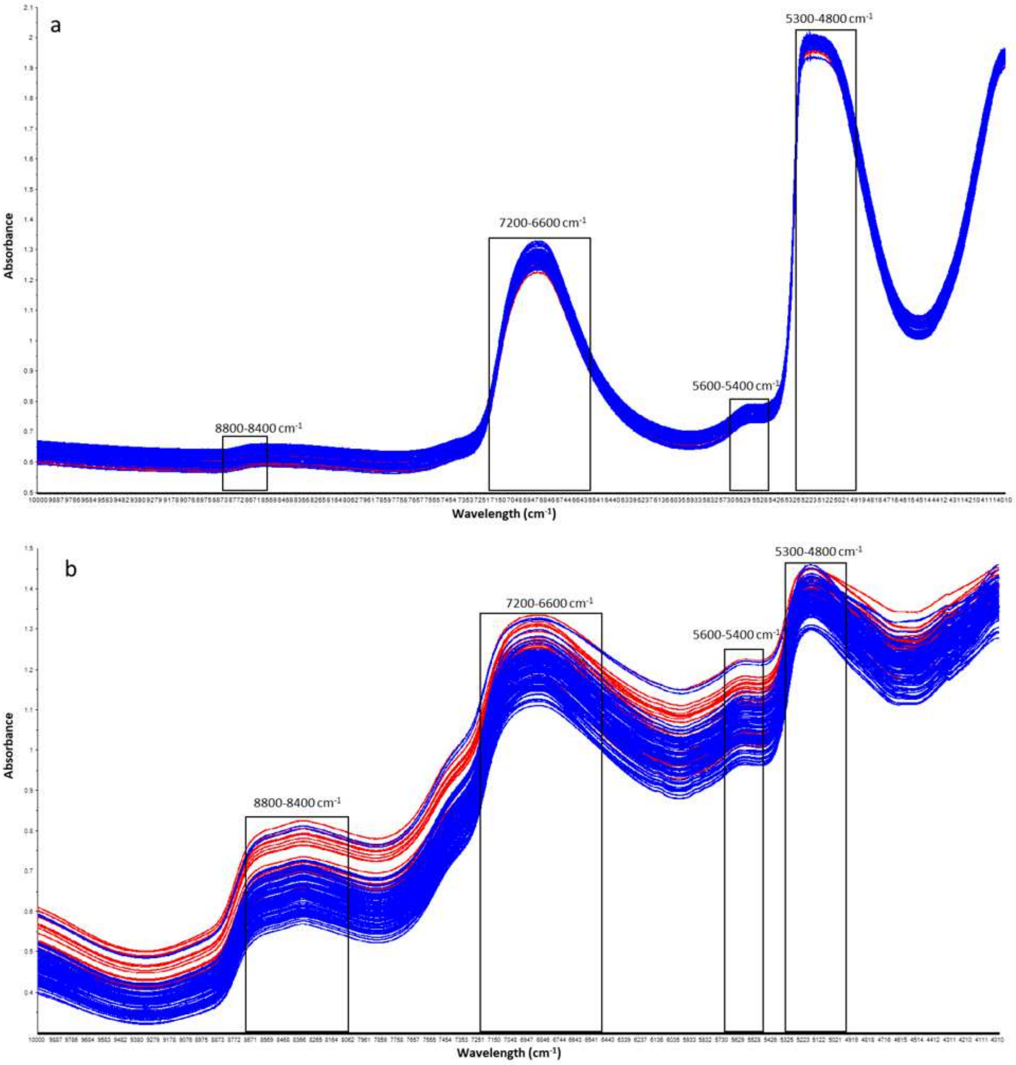


Figure 1 The raw NIR spectra (4000-10000 cm-1) for lemon juice (a) and intact lemon/peel samples (b), PGI: blue; NO-PGI: red.

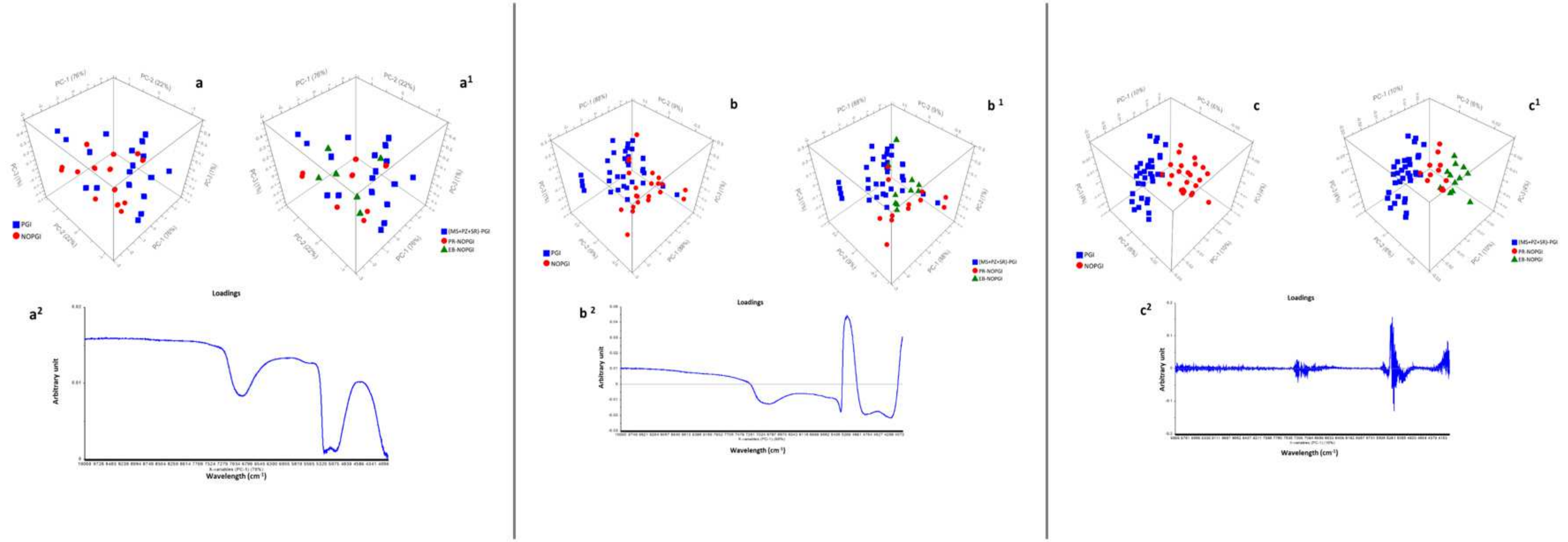


Figure 2 Principal component analysis (PCA) scores plots of the (a) raw spectral data, (b) standard normal variate (SNV) pre-processed spectral data, and (c) SNV with Savitzky-Golay (SG) first derivative (2nd order polynomial, 15 points) pre-processed spectral data for the different lemon juices of two groups (a; b; c), three groups (a^1; b^1; c^1), and the loadings line plots of three groups (a^2; b^2; c^2). PGI: blue; NO-PGI: red for two groups, and PGI: blue; PR-NOPGI: red; EB-NOPGI: green for three groups.

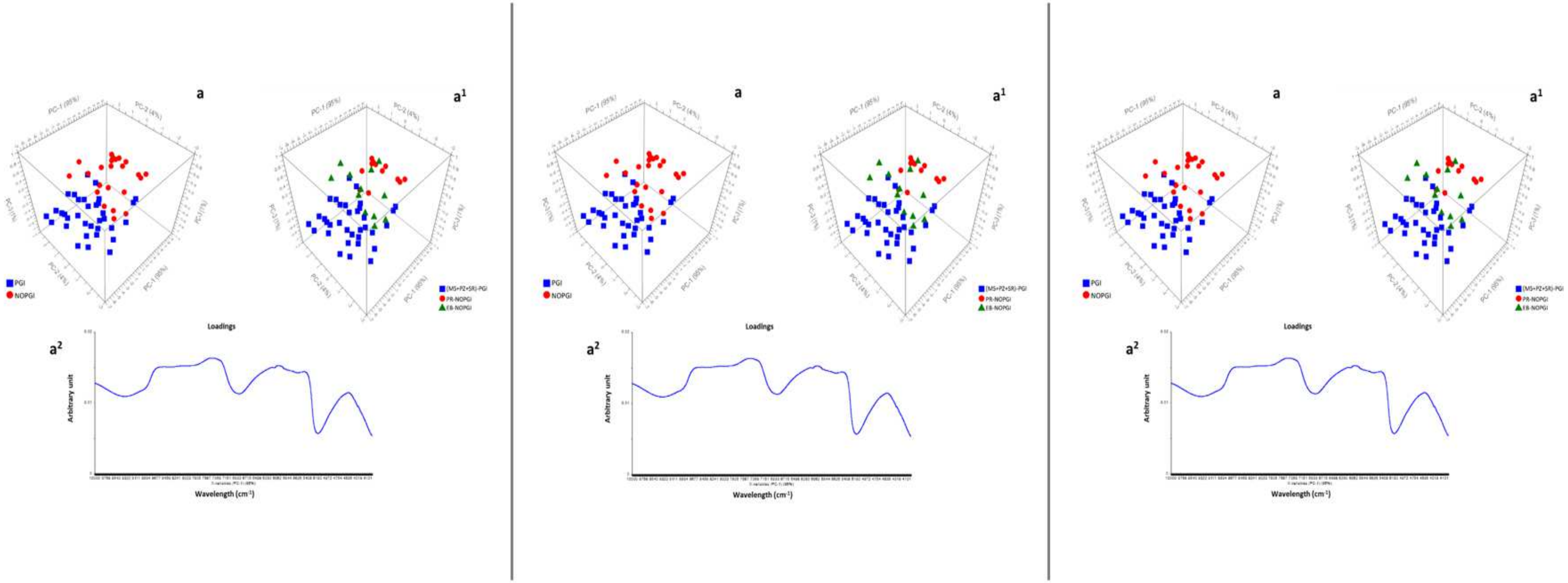


Figure 3 Principal component analysis (PCA) scores plots of the (a) raw spectral data, (b) standard normal variate (SNV) pre-processed spectral data, and (c) SNV with Savitzky-Golay (SG) first derivative (2nd order polynomial, 15 points) pre-processed spectral data for the different intact lemons/peels of two groups (a; b; c), three groups (a^1; b^1; c^1), and the PC-1 loadings line plots of three groups (a^2; b^2; c^2). PGI: blue; NOPGI: red for two groups, and PGI: blue; PR-NOPGI: red; EB-NOPGI: green for three groups. For interpretation of the colour, please refer to the online version of the paper.

Table 1. Results of Discriminant Analyses using three different methods (Calibration set): linear (LDA), quadratic (QDA) and Mahalanobis (MDA).

	Pre-processing	**Methods**	**% Total Accuracy**	**% Accuracy per groups**	
				NO-PGI	**PGI**
Lemon Juices					
	Raw spectra				
		LDA	76	75	77
		QDA	79	75	82
		MDA	71	54	82
	SNV				
		LDA	74	75	74
		QDA	81	79	82
		MDA	77	66	85
	SNV+SG				
		LDA	97	91	100
		QDA	98	95	100
		MDA	98	91	100
Intact Lemons					
	Raw spectra				
		LDA	86	83	88
		QDA	84	83	85
		MDA	86	79	91
	SNV				
		LDA	69	66	71
		QDA	78	75	79
		MDA	74	66	79
	SNV+SG				
		LDA	71	66	74
		QDA	74	66	79
		MDA	74	66	79

Table 2. Correct classification rates of the external validation set by Partial Least Squares Discriminant Analysis (PLS-DA).

Samples	GROUP	FACTOR	R^2 CAL	SEC	SEP	% CAL	% VAL
Lemon juice	3	2	0.95	0.18	0.24	100%	91%
	5	2	0.88	0.45	0.67	65%	46%
Intact lemon	3	4	0.69	0.44	0.51	74%	65%
	5	2	0.59	0.84	0.86	24%	23%

(Group 3) PGI, PR-NOPGI, EB-NOPGI; (Group 5) MS-PGI, PZ-PGI, SR-PGI, PR-NOPGI, EB-NOPGI; (FACTOR) Number of factors used for the generation of the models; (R^2 CAL) Pearson determination coefficient of the Calibration model; (SEC and SEP) Standard Error of the Calibration and Prediction (Validation) models; (%CAL and %VAL) Percentage of samples correctly classified in the Calibration and Validation set.

Appendix A. Supplementary data

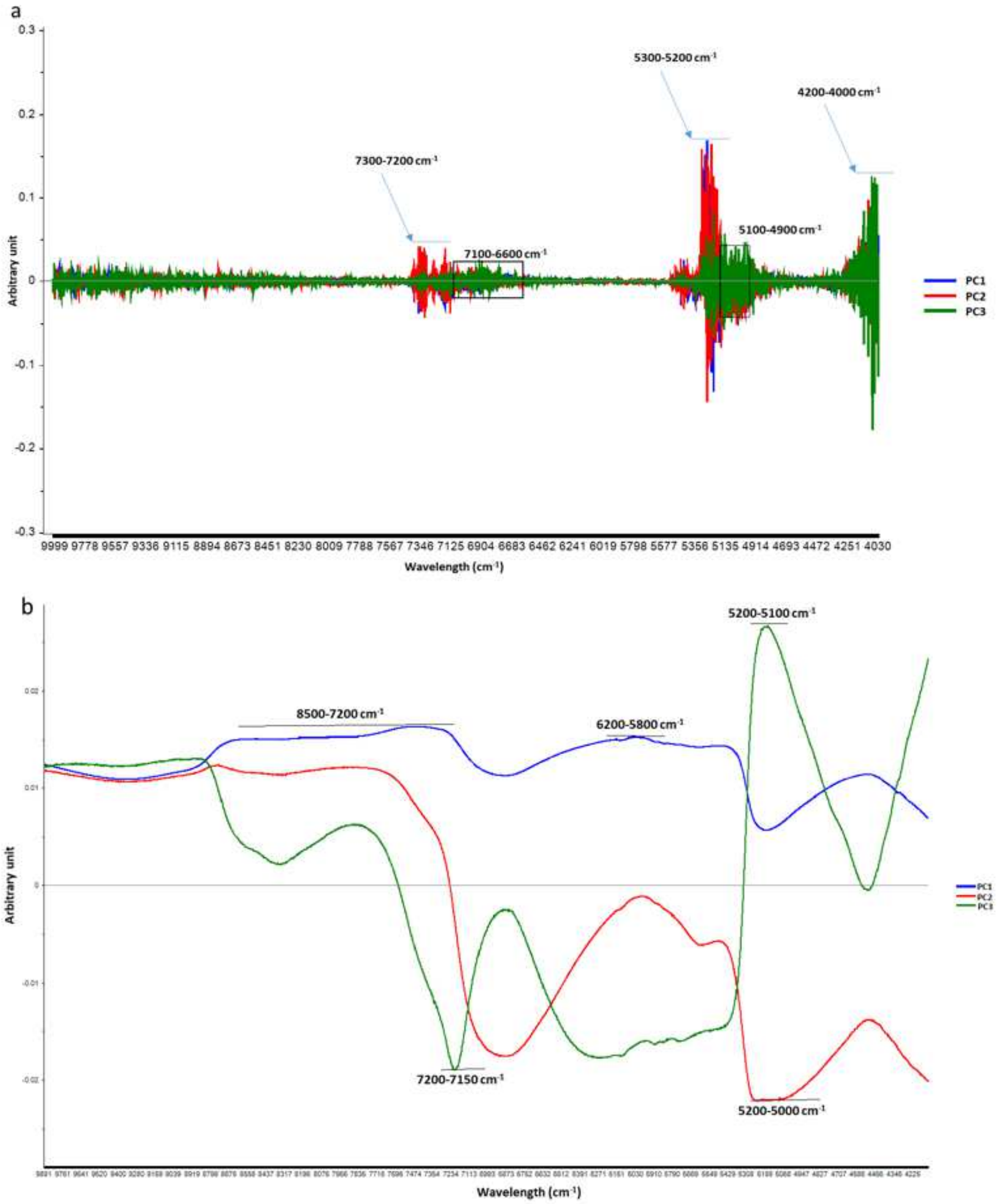


Fig. S1. Loading line plots for Principal Component (PC)1, PC2, and PC3 of the (a) lemon juices and (b) intact lemons/peels generated by Discriminant Analysis of the spectral data showing the wavelengths with the highest absorptions.

Table S1 Confusion matrix of validation (% correctly validated) for the Partial Least Squares Discriminant Analysis (PLS-DA) model based on spectra of intact lemons pre-processed using Standard Normal Variate (SNV)and Savitzky-Golay (SG) polynomial derivate in the PGI area.

% Correct	**MS-PGI**	**PZ-PGI**	**SR-PGI**
MS-PGI*	**19%**	50%	0%
PZ-PGI°	59%	**28%**	3%
SR-PGI	0%	83%	**17%**

*31% of samples were classified as PR-NOPGI; °9% of samples were classified as PR-NOPGI

Table S2 Confusion matrix of validation (% correctly validated) for the Partial Least Squares Discriminant Analysis (PLS-DA) model based on spectra of lemon juices pre-processed using Standard Normal Variate (SNV) and Savitzky-Golay (SG) polynomial derivate in the PGI area.

% Correct	**MS-PGI**	**PZ-PGI**	**SR-PGI**
MS-PGI*	**63%**	31%	0%
PZ-PGI	47%	**53%**	0%
SR-PGI	33%	67%	**0%**

*7% of samples were classified as PR-NOPGI.

GENERAL CONCUSION

The main aim of my thesis work was to build different discrimination models able to authenticate and track high-quality lemon fruits such as the Limone di Sorrento PGI (cv. Femminello Ovale di Sorrento) and Limone Costa d'Amalfi PGI (cv. Femminello Sfusato Amalfitano) produced in the Campania region. For this purpose, lemons collected from orchards in different PGI and not PGI areas and in two consecutive production years (2018 and 2019) were analysed for their geochemical and spectroscopical fingerprinting. A special emphasis was given to the interaction of environmental cultivation conditions (pedoclimatic conditions) in defining the specific profile of lemon fruits. Among the most widely applied techniques for authentication and traceability of agri-food products, we selected soil-related indicators such as multielement fingerprinting and Sr isotope signature, along with spectroscopic fingerprinting by Near-Infrared Spectroscopy (NIR). Our results demonstrated that lemon fruit profiles (geochemical and spectroscopical profiles) were strongly related to the cultivation soil as well as to the climatic conditions of cultivation area.

With regard to geochemical indicators, we built different chemometric discriminant models based on mineral element profile and ^{87}Sr/^{86}Sr isotope ratio of lemon juices both at small (regional) and large geographical scale. The latter was achieved by comparing Campania lemons with samples from Calabria region and from southern California. In both cases, foreign lemons were gathered in local markets.

On a small geographical scale and taking into consideration the Limone di Sorrento PGI, the Stepwise Linear Discriminant model (S-LDA) classified correctly 97.7% of juice samples and validated 93.8% of juice samples (external to the model) according to provenance sites. In this case, for understanding the relationship between soil and lemon juices, Pearson correlation analysis and S-LDA were applied to different production years. Our conclusion was that the use of non-essential elements (Ba, Rb, and Sr) was more reliable than essential ones in discriminating the different provenance sites.

The comparison of mineral element profiles and ^{87}Sr/^{86}Sr isotope ratio of lemon juices from different provenances (Limone di Sorrento PGI and Limone Costa d'Amalfi PGI both on a small and a large geographical scale) confirmed previous results giving for all elements 98.4% and 90.9% of juice samples correctly classified and validated, respectively, 86.3% and 72.7% for essential elements and 92.4% and 86.3% for not-essential elements. It is worth noting that although the performance of the whole elemental profile gave a better result than the non-essential elements profile, the reliability of the two models, calculated as the ratio between the percentage of correctly validated and classified samples, was similar. The Sr isotope ratio showed a clear differentiation among the lemon fruits from the Campania region and extra-regional samples. In addition, by analysing the $^{86}Sr/^{87}Sr$ ratio in cultivation soils, it was clear that the Sr isotope ratio of lemon juices was more closely related to that of the bioavailable fractions than to that of the total Sr in soil. Finally, combining on an explorative level (Principal Component Analysis, PCA) isotopic signature and mineral element profile of lemon juices by a low-level data fusion approach, the results showed a better clustering according to geographical origin than the two determinations taken separately.

Application of NIR spectroscopy fingerprinting to intact lemons showed the strong influence of environmental growing conditions on the quality of lemon fruits. The explorative analysis based on raw NIR spectra showed natural clustering of samples according to production years. When the two years were considered separately, data were grouping by sites of origin better then by different cultivars. These results were confirmed by the application of six LDA models, all suggesting to build discrimination models according to origins (PGI and not PGI productions) on the base on one production year. The Multiple Linear Regression (MLR) between NIR spectra and lemon quality parameters highlighted a good relation between internal and external quality and the possibility to build predictive models. However, the results of MLR suggested to use NIR spectroscopy as quality predictive models on both cultivars only for single years rather than to implement datasets with fruits from different production year. In addition, on one part of the samples collected in 2019 (intact lemons and juices), were applied two different chemometrics models i.e., LDA and Partial Least Square Discriminant Analysis (PLS-DA). The results showed better provenance discrimination using the lemon juices than the intact lemons.

The comparison between the results obtained applying the geochemical fingerprinting approach with those of NIR spectroscopy, highlighted a higher stability of the former in discriminating lemon fruits collected in two different production years, especially when not essential elements are considered. However, the numerous vantages of NIR spectroscopy application (non-destructive, rapid, and cheap) as well as the results obtained, make the technique useful to be used for rapid screening in order to verify the quality and geographical origin of lemon fruits during the year. In both approaches, the study of the pedoclimatic features was fundamental to understand the nature of discriminating variables.

Discriminant models based on geochemical and spectroscopical profiles of lemon fruits could substantially contribute to implementing a blockchain system for Campanian lemon traceability, providing real-time information not only to the final consumers but also to manufacturers, distributors, and retailers. To improve their robustness and reliability in detecting frauds, additional research should be conducted to include a greater number of lemon farms (or sampling points) in the

PGI areas and to enlarge the existing database including lemon samples from other regions and validate the models built.