Original Research Article

A new simplex-based approach predicting olive oil blend compositions from fatty acid data

Nabil Semmar [a,*], Jacques Artaud [b]

[a] Institut Supérieur des Sciences Biologiques Appliquées de Tunis, Université de Tunis El Manar, Tunis, Tunisia
[b] Aix-Marseille Université, LISA, EA4672, Equipe METICA, 13397 Marseille Cedex 20, France

ABSTRACT

Olive oils represent complex matrices varying from pure to heterogeneous varietal contents. Quantitative analysis of co-occurring components is fundamental for conformity checking and adulteration alerting (fighting) of commercial oils. Proportions of co-occurring components are governed by additive-dilutive processes which obey to simplex rule. Using simplex rule, we developed an original computational approach to predict proportions of different co-occurring oil varieties from quantitative chemical features of blends. The approach consisted in applying a complete set of N mixtures between different olive oil varieties by gradually varying their proportions. The N simulated mixtures were characterized by N average fatty acid (FA) profiles calculated from N combinations of randomly sampled individual profiles. After k iterations of the mixture design, the k sets of N FA average profiles were used as input in a discriminant analysis to predict proportions of co-occurring olive oil varieties in different blends. Illustrative application concerned blends made by three main French mono-varietal virgin olive oils (*Aglandau*, *Grossane* and *Salonenque*) and benefiting from Protected Designation of Origin label. Predictive model was validated on outside blends and showed prediction errors with an order of 10% susceptible of reduction by applying a larger mixture design.

© 2015 Elsevier Inc. All rights reserved.

1. Introduction

Authentication of agro-alimentary products is a recurrent problem implying several analytical techniques (chromatographic, spectroscopic, DNA markers, etc.) helping to determine specifications and to identify conformity with known labeled products: Protected Designation of Origin (PDO), Registered Designation of Origin (RDO), Protected Geographical Indication (PGI), etc. These labels represent a protection way of product denomination because they guarantee that fabrication was performed according to strict and constraining specifications. For that, production, transformation and development must occur in a well-defined geographical area (climate, soil, etc.) by following a well-recognized and contrasting know-how.

Consequently, products issued from labels' production ways have more enhanced values than other products.

Currently in France, there are eight designations of origins (DO) concerning virgin olive oils (VOOs) (7 PDOs and 1 RDO). Some of these labels are mono-varietal (Nyons, Nice), others are multivarietal (Aix-en-Provence, Vallée des Baux de Provence, etc.) mixing several compositional varieties (e.g. *Aglandau*, *Grossane*, *Salonenque*). These varieties are required by the specifications concerning the trees occurring in the orchards. However, the percentages of olive oil varieties in commercial blends are free under the condition that at least two main varieties co-occur (JORF, 1998, 1999). Moreover, no minimal percentage of compositional oil variety is fixed such that PDO label-benefiting oil blends are commercialized without such information and their organoleptic characteristics are subjected to significant fluctuations.

In previous research works, qualitative analyses were carried out for varietal and geographical origin authentications of VOOs by combining analytical techniques with chemometric models (Aparicio, 2000; Downey et al., 2003; Angerosa et al., 2006; Ollivier et al., 2006; Galtier et al., 2007; Mannina and Segre, 2010;

Casale et al., 2012; Longobardi et al., 2012; Monfreda et al., 2012). Datasets containing fatty acid (FA) profiles were subjected to multivariate analyses (e.g. PCA) for topological differentiation of several olive oil varieties (Amaral et al., 2010; Laroussi-Mezghani et al., 2015). Beyond the multivariate structuration of olive oil varieties into different chemical trends, other works developed pattern recognition models of oil varieties from FA chromatograms using discriminant analysis (DA); FA profiles were statistically analyzed by DA for recognitions of biological and spatio-temporal characteristics of olive oils including cultivar types, geographical origins and harvest years (Sacco et al., 2000; Ollivier et al., 2006; Dıraman et al., 2011).

Beyond multivariate qualitative classification and recognition of separated cultivars, compositional quantification of multivarietal blends represents a complex multi-scale question which has been little studied until now; Marini et al. (2007) previously developed an original chemometric approach to predict proportions of oil varieties from blend profiles by using Self-Organizing Maps (SOM). Although other works were published on the control of multi-species vegetable oils (Maggio et al., 2010; de la Mata et al., 2012; Popescu et al., 2015), to our knowledge, the work of Marini et al. (2007) is the only research strictly made at varietal scale, focusing on the compositional evaluation of multi-varietal olive oil blends. It uses iterative computations to determine geometric locations of blends within a square or hexagonal network (covering all the possible blends). In SOM, blends are geometrically identified by reference to inter-blend weights (i.e. weights between network cells).

We treated this question of blend composition evaluation by using an original simulation approach based on simplex network which has triangular geometry. In simplex networks, geometrical locations of blends are governed by within-blend constraint under which the coordinates (weights) of co-occurring oil varieties obey to unit–sum rule (Fig. 1). This property of conservation turned out to be particularly advantageous for reliable quantitative evaluation of several varieties of olive oils in complex mixtures. Around this goal, the simplex approach was required to:

- be sensitive toward chemical fingerprints of each olive oil variety among all the mixed ones in a same blend;
- overcome identification noise due to mixture between groups showing some similarities in their chemical features; and
- explore inter- and intra-groups variations within blends to analyze links between proportions of co-occurring oil varieties and general feature of mixtures.

Such variations naturally originate from chemical polymorphism occurring within each oil variety; different trees representing a same olive cultivar naturally show some chemical variations among them; it results in blends with similar varietal compositions that do not show exact similarity along their whole chemical profiles. Therefore, repetitions are necessary to construct a model that is able to absorb intra-varietal noise for reliable prediction of blend composition.

In lab, preparation of a wide set of repeated blends represents a very time and money-consuming task. This constraint can be advantageously overcome by the presented simulation simplex approach, which proposes to extensively explore the chemical variability of different olive oil groups in order to prepare a broad library of repeated blends for construction of a high noise-absorbing calibration model.

Simplex is commonly used in chromatography to determine relevant proportions of different miscible solvents for target separations of co-occurring chemical constituents in complex matrices (Glajch et al., 1982; Nyieredy et al., 1985). In this case, a complex matrix chromatogram represents target variables characterized by resolution or selectivity parameters varying in relation to proportions of eluents (input variables). A conceptually inverse approach is presented in this paper, which consisted in using a single chromatogram representing complex matrix (blend) as predictive (input) variable to estimate the proportions of constitutive components (olive oil varieties). This original simplex approach is illustrated here by predictions of proportions of three co-occurring main French olive oil varieties, *Aglandau* (*A*), *Grossane* (*G*), and *Salonenque* (*S*) (Aix-en-Provence and Vallée des Baux de Provence PDOs), from fatty acid methyl esters directly analyzed in blends.

2. Materials and methods

2.1. Virgin olive oil samples

Sampling was carried out during nine successive crops (2002/2003–2010/2011). French VOO samples (*n* = 144) were obtained from the French Inter-Professional Olive Oil Association (AFIDOL), Aix-en-Provence, France (Pinatel et al., 2013). VOOs were produced in laboratory scale by oleodoseur extraction system, from handpicked olive fruits of three French varieties: *Aglandau* (*A*) (*n* = 83), *Grossane* (*G*) (*n* = 26), and *Salonenque* (*S*) (*n* = 35).

All the oil samples were included in the ranges established for EVOOs category (IOC, 2013).

2.2. Fatty acid determination

Olive oil in 2,2,4-trimethylpentane (isooctane) ($\approx$0.06 g mL^{-1}) was trans-methylated with a cold solution of KOH (2 M) (200 μL) according to the European Standard NF EN ISO 5509 Norm (2000). Fatty acid methyl esters (FAME) were analyzed according to the European Standard NF EN ISO 5508 Norm (1995). Analyses were performed using Agilent Technology gas chromatograph 7890A (GC) (Les Ullis, France) equipped with a split/split-less injector (*T* = 250 °C) and flame ionization detector (FID) (*T* = 250 °C). A silica capillary column (60 m × 0.25 mm i.d., 0.25 μm film thickness) coated with polyethylene glycol (Supelco wax, Supelco, France) was used. The carrier gas was hydrogen (column flow 1 mL/min) and the split ratio was 1:60. The oven temperature program was as follows: 20 min at 210 °C, from 210 to 245 °C for 6 °C/min, 20 min at 245 °C. Fatty acid identification was realized in previous work by Ollivier et al. (2003). Fatty acid percentages were determined by internal standardization without taking into account mass response factors. Only fatty acids with a content of more than 0.01% were taken into account.

Repeatability of FAME was evaluated by esterifying 12 times the same VOO sample. Repeatability of injection was also evaluated by injecting 12 times the same FAME sample. The coefficients of variation were lower than 5% for the most important fatty acids and lower than 10% for other minor ones.

2.3. Nomenclature

Fatty acids (FAs): 16:0, palmitic acid (hexadecanoic acid); 16:1ω9, hypogeic acid (7-hexadecenoic acid); 16:1ω7, palmitoleic acid (9-hexadecenoic acid); 17:0, margaric acid (heptadecanoic acid); 17:1ω8, margaroleic acid (9-heptadecenoic acid); 18:1ω9, oleic acid (9-octadecenoic acid); 18:1ω7, Z-vaccenic acid (11-octadecenoic acid); 18:2ω6, linoleic acid (9,12-octadecadienoic acid); 18:3ω3, linolenic acid (9,12,15-octadecatrienoic acid); 20:0, arachidic acid (eicosanoic acid); 20:1ω9, gondoic acid (11-eicosenoic acid); 22:0, behenic acid (docosanoic acid); 24:0, lignoceric acid (tetracosanoic acid).