



HAL
open science

An iterative selection algorithm: A decision aid to select the best extra virgin olive oils competing in an international contest

Fabien Girard, Jacques Artaud, Christian Pinatel, Magalie Claeys-Bruno, Catherine Rébufa

► To cite this version:

Fabien Girard, Jacques Artaud, Christian Pinatel, Magalie Claeys-Bruno, Catherine Rébufa. An iterative selection algorithm: A decision aid to select the best extra virgin olive oils competing in an international contest. *Food Control*, 2023, 151, pp.109776. 10.1016/j.foodcont.2023.109776. hal-04074757

HAL Id: hal-04074757

<https://amu.hal.science/hal-04074757>

Submitted on 19 Apr 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

An iterative selection algorithm: A decision aid to select the best extra virgin olive oils competing in an international contest

Fabien Girard^a, Jacques Artaud^a, Christian Pinatel^b, Magalie Claeys-Bruno^a, Catherine Rébufa^{a,*}

^a Aix Marseille Univ, Avignon Université, CNRS, IRD, IMBE, Marseille, France

^b Centre Technique de l'Olivier (CTO), Maison des Agriculteurs, 13100, Aix-en-Provence, France

ARTICLE INFO

Keywords:

Selection algorithm
Extra virgin olive oil
Descriptors
International competition

ABSTRACT

Extra virgin olive oils (EVOO) are typically categorized in accordance with the nature of their fruitiness (green or ripe) and the intensity of their fruitiness (delicate, medium, or robust) in accordance with the proposal of the competitor (producers) or a chemical analysis certificate established by an official panel. Depending on the category in which EVOO participates, its categorization may either disadvantage or benefit its rating. To eliminate this ambiguity and maximize an EVOO's probability of winning, the jury of the International "Word Edible Oils" competition use an innovative method to assess the EVOOs using a restricted amount of sensory descriptors (aromatic maturity, structure and fruitiness). Independent of category, the best EVOOs with comparable organoleptic properties have been identified using a statistical processing of the scores of the three descriptors. This is an iterative version of the technique developed by Wootton, Sergent, and Phan-Tan-Luu (iWSP), which generates subspaces that enable a local selection of the best EVOOs in a 2D aromatic maturity vs structure plan. In each subspace, their ranking is determined by their fruitiness score. An iterative approach that takes into consideration the changing sequence of subspace formation and the varying size of the subspace enables the selection of the best EVOOs. The development of the iWSP algorithm enabled the elimination of category constraints, (ii) the selection of the best EVOOs among those with comparable organoleptic features based on a large number of simulations, and (iii) the creation of a ranking to aid the jury's final judgment.

1. Introduction

Due to its high selling price, virgin olive oil is one of the most regulated and monitored agri-food items, which might lead to fraudulent activities. Consequently, virgin olive oil is subject to precise and continually evolving European regulations and international standards that define the physicochemical and organoleptic characteristics of the different categories of virgin olive oils: extra virgin olive oil (EVOO), virgin olive oil (VOO), ordinary virgin olive oil (OVOO), and lampante virgin olive oil (LOO) (Conte et al., 2020; IOC, 2018; EU regulations 2022/2104). VOOs are the only agro-food items that need an organoleptic evaluation prior to being deemed suitable for human consumption (EVOO and VOO in EU area and OVOO in certain countries outside). Many regional, national, and international competitions are held worldwide to recognize the finest olive oils (Rébufa et al., 2021). The

rewarded oils recognize the top producers, who then increase their output, advertise their oils abroad, and boost the trade of high-quality goods. By generating attention, competitions assist customers in identifying factors that might lead their selection and in discovering the variety of EVOO flavors. Each competition has its unique rules, but all include an organoleptic examination by knowledgeable tasting jurors who evaluate the absence of negative attributes (EVOO) and the strength (robust, moderate, or delicate) of fruitiness (IOC, 2018; EU standards 2022/2104). Often, taster jurors employ analogous sensory descriptors experienced by the ortho-nasal and retro-nasal methods to better characterize monovarietal olive oils, Protected Design Origin (PDO) oils, terroir oils, and others (Bendini et al., 2012; Bongartz & Oberg, 2011). Most competitions suggest a rating of EVOOs according to the type of their fruitiness (green or ripe) and the level of their fruitiness (strong, medium, or delicate), resulting in a total of six EVOO categories

* Corresponding author. Aix Marseille Univ, Univ Avignon, CNRS, IRD, IMBE UMR 7263. Pôle de l'Etoile, Équipe Biotechnologie et Chimiométrie, Case 451, 13397 Marseille cedex 20, France.

E-mail addresses: fabien.girard@univ-amu.fr (F. Girard), jacques.artaud@univ-amu.fr (J. Artaud), c.pinatel@ctolivier.org (C. Pinatel), m.claeys-bruno@univ-amu.fr (M. Claeys-Bruno), c.rebufa@univ-amu.fr (C. Rébufa).

(IOC, 2018). Hence, EVOOs may participate in the category selected by the producer during registration or in the category determined by a certificate of sensory analysis required by the contest rules (the last one being delivered by an official panel holding IOC recognition). But in the two cases, this categorization can penalize or benefit a producer for the following reasons: the competitor placed his EVOO in the incorrect category because he lacked a reference, or his EVOO is close to the categories' limits as defined by the chemical and organoleptic analysis certificate. This product may thus compete in a category other than its own. Furthermore, EVOOs in the boundaries of two categories may be far from samples in the same category, but close to samples in the next category. This circumstance may or may not enable the producer to get an award for its EVOOs. Yet, the granting of a reward also relies on the performance of specialists (the jurors in the competition or the official panels issuing sensory analysis certificates). To strengthen the use of official approach, several studies have compared the findings of different trained panels in an effort to enhance their performance, training, measurement methods, and the necessity for robust, repeatable reference materials (Amelio, 2019; Barbieri et al., 2020). Recent research has paid particular attention to the classification of panel performance tools according to sensory method types, the statistical methods and software used to obtain a good agreement between instrumental measurements and human sensory results, and the control of panel performance drift (Sipos et al., 2021). The correlation between the chemical composition of extra-virgin olive oils (EVOOs) and their sensory profile as judged by a panel was still the subject of publications. Target components were volatile and phenolic ones, measured from several analytical methods (MS, HS-MS, HS-SPME-GC-MS, HPLC-DAD) in order to distinguish the grade of olive oils (EVOO, VOO or LOO) or to group monovarietal EVOOs with comparable sensory and chemical characteristics (Cecchi et al., 2022; Gerhardt et al., 2019; Quintanilla-Casas et al., 2020). Instrumental data fusion of an electronic nose based on headspace mass spectrometry, an electronic tongue based on Middle Infrared spectroscopy, and an electronic eye based on UV-vis spectrophotometry has been a valuable tool for olive oil classification based on their category and the presence of certain sensory defects (Borràs et al., 2016). Statistical processing of these data was conducted using a hierarchical clustering approach to identify oil groups with similar intensities for the perceived olfactory-gustatory attributes, or by applying unsupervised and supervised chemometric tools to develop predictive models for olive oil classification (Borràs et al., 2016; Cecchi et al., 2022; Gerhardt et al., 2019). Their purpose was to approximate human panel judgments as closely as possible. These experimental measures, often based on the search for flaws, enhanced the categorization of EVOO panels, given that the search for defects is simpler than the classification of EVOOs devoid of defects. As far as we are aware, no scientific studies have yet published a strategy for selecting virgin olive oils to be rewarded in a competition as equally as feasible. In the majority of olive oil contests, no information is given on the final sample selection technique; it is unknown how the scores of the (sometimes very many) organoleptic descriptors are utilized to rank the samples.

In addition, this research provides a mathematical method as a decision-making tool for using the findings of the several tasting phases of the worldwide "Word Edible Oils" competition hosted in France by the Agency for the Valorization of Agricultural Products (AVPA). In contrast to other EVOO competitions, the "World Edible oils" (formerly known as AVPA) contest evaluates the organoleptic qualities of registered virgin olive oils using just three distinguishing descriptors. There are (i) the "aromatic maturity," which describes, in a continuous manner, the nature of oil fruitiness (from green to ripe), (ii) the "structure," which describes the intensity of oil bitterness and pungency (delicate, medium, or robust), and (iii) the "fruitiness," which accounts for the overall aromatic value of oil, i.e., the harmony and intensity of fruitiness (Rébufa et al., 2021). As opposed to other competitions, the AVPA event does not require that the fruitiness category of samples be known at the time of registration. Here, the mathematical technique

used to process AVPA findings is based on an innovative application of Wootton, Sergent, and Phan-Tan-(WSP) Luu's algorithm. Earlier, the WSP method was created for producing space-filling designs constructed from a collection of N candidate points obtained using an interesting sampling approach. Hence, the WSP method allowed for the gradual elimination of points, ensuring that all points in the final design were evenly distributed and separated by a minimal distance d_{min} (Santiago et al., 2012; Sergent, Luu, & Elguero, 1997; Sergent, Luu, Faure, & Elguero, 1997). An adaptive version (aWSP) was later designed for considering experimental constraints or the increase of the point density in a zone of particular interest to solve the problem of missing points (Beal et al., 2014) and was subsequently adapted for the reduction of unsupervised variables (Ballabio et al., 2014). Here, an iterative selection variant of this algorithm (now dubbed iWSP) was developed in order to identify the finest olive oil samples throughout the contest's selection phases. Using the basic WSP theory of the nearest neighbor, the iWSP algorithm created subspaces that permitted a local selection of the EVOOs in the 2D plan (aromatic maturity score versus structure score). The innovation was to employ a third dimension, the fruitiness score, to choose the top EVOOs in each 2D subspace already identified. The purpose of this study was to demonstrate how the iWSP methodology is constructed (i) to select the best olive oils among those with similar organoleptic properties (ii) to allow olive oils placed near category limits (green/ripe and robust/medium/delicate) to have the best chance of being well evaluated and subsequently awarded (iii) to enable a safe ranking with the aid of iterations on the building order and size of subspaces. Nevertheless, the jury always has the option to amend the categorization of EVOOs offered by the iWSP algorithm in order to award specific EVOOs based on their appreciations in addition to the descriptor scores.

2. Material and method

2.1. Specificity of the extra virgin olive oil selection by "Word Edible Oils" contest

AVPA has been organizing an international competition for virgin olive oil in partnership with the Centre Technique de l'Olivier (France Olive, Aix-en-Provence, France) for the past two decades. The "Word Edible Oils" contest attracts over 200 oils from 20 countries annually and is open to growers, millers, and their groups (cooperatives, denominations of origin). The competition employs three non-redundant and unambiguous descriptors to classify the registered oils, with the aim of showcasing the diversity of extra virgin olive oils (EVOOs) from around the Mediterranean and the world. Objective descriptors such as aromatic maturity and structure are used, as they are easier to evaluate for jurors and more noticeable to consumers. However, fruitiness, a more subjective descriptor, presents a challenge in evaluation.

Aromatic maturity is a specific AVPA contest descriptor that signifies the fruitiness's maturity (from green to ripe). The rating ranges from zero (green fruitiness) to one (ripe fruitiness).

On a scale from 0 to 10, AVPA judges provide ratings for bitterness and pungency, as described by IOC rule (IOC/T.20/Doc. No. 15/Rev. 10). These scores, weighted by empirical factors derived through experimentation, are used to compute the structural descriptor (Eq.1). These coefficients account for the length of perception of these traits and the perception of customers for whom bitterness is more irritating than pungency (Pinatel, personal results). The structure score is also scored on a scale of 0–10.

$$Structure = [(0.62 \times bitterness + 0.38 \times pungency) + \text{Max}(bitterness; pungency)] / 2 \quad (\text{Eq.1})$$

Fruitiness is the third specific AVPA description that explains the overall aromatic smells of the EVOOs. It is independent of bitterness, pungency, and the placement of aromatic maturity. On a scale from 0 to

10, judges rate the harmony and intensity of EVOO's fruitiness for its quality. On a scale ranging from 0 to 20, the fruitiness score is calculated by adding the scores of the last two characteristics.

The process for selecting oils has been reported before (Rébufa et al., 2021). Basically, it comprises of three oil selection processes by an expert jury, harmonising the scores and removing the lowest-ranked virgin olive oils in each round using statistical processing of attribute scores. At each selection phase, three jury groups taste distinct oil samples. For successive choices, the same jurors are selected. Olive oils with defects are removed after the first round of selection, and only EVOOs are permitted to continue in the competition. At the conclusion of the third round, a mathematical ranking of the EVOOs is offered to the jury as a tool for recognizing the top oils (about 20 percent of the total number of initially registered oils). Hence, the jury remains sovereign in every instance. This competition offers the publishing of results through a 2D map on which awarded EVOOs are positioned according to the typical fruitiness categories, informing producers of the type and strength of the fruitiness of their oil samples, as is customary in other competitions.

The results of the various AVPA contest selection procedures for the years 2020, 2021, and 2022 were made available for this research. These were included in Table S1 of the supplementary material. For graphical representations, just the results of the first selection were used.

2.2. Iterative WSP model (iWSP) for determination of EVOOs with similar organoleptic profile and selecting the best ones

During a previous study conducted in collaboration with the AVPA organizer (unpublished works), the scores of the three parameters (aromatic maturity, structure, and fruitiness) were used to establish a spatial graphical representation of the various EVOOs tasted by defining a three-dimensional scatter plot. A mathematical model enabled the selection of EVOOs based on the Euclidean distance between the location of the EVOOs and the barycenter. This model, however, tended to favor the selection of EVOOs with a green and strong fruitiness. To circumvent this issue, a new strategy was developed with the iWSP algorithm by utilizing initially two of the three organoleptic parameters (aromatic maturity and structure) to define a 2D plan in which the coordinates (x,y) of EVOOs are positioned according to the score values of the corresponding descriptors, without data transformation. Aromatic ripeness and structure have been selected as characteristics for the 2D plan because they are readily sensed by customers and more objectively assessed by AVPA judges. Within the point scatterplot (where each point represents an EVOO), the iterative mode of the iWSP algorithm enables the delineation of subspaces comprised of EVOOs with comparable aromatic maturity and structural scores. The number of subspaces generated by the iWSP algorithm is such that all EVOOs are members of at least one subspace. Furthermore, the iWSP method allows adjusting the size of the subspaces while taking the totality of the samples into account. The EVOOs with the highest fruitiness score (0–20) within each subspace are chosen as the best. A further feature of the iWSP approach is that it allows a configurable proportion of the best EVOOs to be picked in each subspace. The head of the competition jury determines this proportion.

The different steps of the iWSP algorithm are described below.

1. Construction of an orthonormal 2D plan defined by the scores of the structure and aromatic maturity attributes. *
2. Projection of the N EVOOs in this 2D plan.
3. Initialization of the number of selected EVOOs (SEL variable equal to a certain percentage of N).
4. Initialization of the number of clusters to pave the plan (CLU variable).
5. Initialization of the starting coordinates for the construction of the clusters. **

- a. Determination of the Euclidean distance between each EVOO and the starting coordinates.
 - b. Determination of the first master point (MPI) from the smallest distance.
 - c. Determination of the Euclidean distance between each EVOO and MPI .
 - d. Iterative calculation of the minimum diameter (d_{min}) of the clusters to obtain CLU clusters including all the EVOOs. The centre of a cluster is necessarily an EVOO.
 - e. Identification of the EVOOs at the centre of the clusters as master points (MPI).
 - f. Iterative calculation of the percentage of EVOOs in each cluster to obtain the final selection (SEL). This selection is made by comparison with the 3rd attribute, the fruitiness score.
6. Initialization of the new starting coordinates for the construction of clusters. Then go to the step 5.
 7. Initialization of the new number of clusters to pave this plan (CLU). Then go to the step 4.

*In this study, the scores of the structure and aromatic maturity descriptors can be varied between 0 and 10.

** Nine starting coordinates ((0,0), (0,5), (0,10), (5,0), (5,5), (5,10), (10,0), (10,5) and (10,10)) were chosen to ensure that the best selected EVOOs were independent of the first master point.

The different steps of iWSP algorithm were presented in Fig. 1. Fig. 1a described the main organizational units allowing the iterative calculation of the number of selected EVOOs and the number of clusters in which EVOO organoleptic properties are compared. A detailed description of the algorithm itself was given in Fig. 2a.

In the following parts, the word “simulation” will mean the fact to realize the step number 5 described above.

2.3. Software

The iWSP model was developed with *MATLAB*, version 2020a.

3. Results and discussion

The first section illustrates why EVOOs' location at the category limit may be beneficial or disadvantageous during the contest selection procedure. Next, a thorough description of the iWSP algorithm selection mode is provided to illustrate that it is an effective method for avoiding the category issue and proposing a ranking of the EVOOs, while allowing the jurors free to pick the EVOOs to be rewarded.

3.1. Problem of EVOOs at the limits of categories

The majority of competitions use the IOC technique for evaluating the organoleptic qualities of EVOO and classifying them (IOC, 2018). So, the strength of perception of positive attributes (fruity, bitter, and pungent) was classified as robust (when the median of the attribute is more than 6.0), medium (when the median is between 3.0 and 6.0), and delicate (when the median of the attribute is less than 3.0). In the several phases of the “World Edible Oils” competition, bitterness and pungency are rated by judges, and then utilized to determine the structure descriptor. The aromatic maturity and fruitiness are also measured as two more distinguishing descriptors. By similarity with the IOC rule, the results of the “World Edible Oils” competition are exhibited on a map, a two-dimensional plan based on the aromatic maturity and structure scores of all EVOOs participated in the competition (Fig. 2). The awarded oils (with gold, silver, and bronze medals) are highlighted by always appearing with their anonymity number. To assist producers in evaluating the nature and intensity of the fruitiness of their EVOOs (i.e., a correspondence with the aromatic maturity and structure descriptors), it was specified (i) horizontally the zones bounding the green and ripe fruitiness and (ii) vertically the regions corresponding at the different

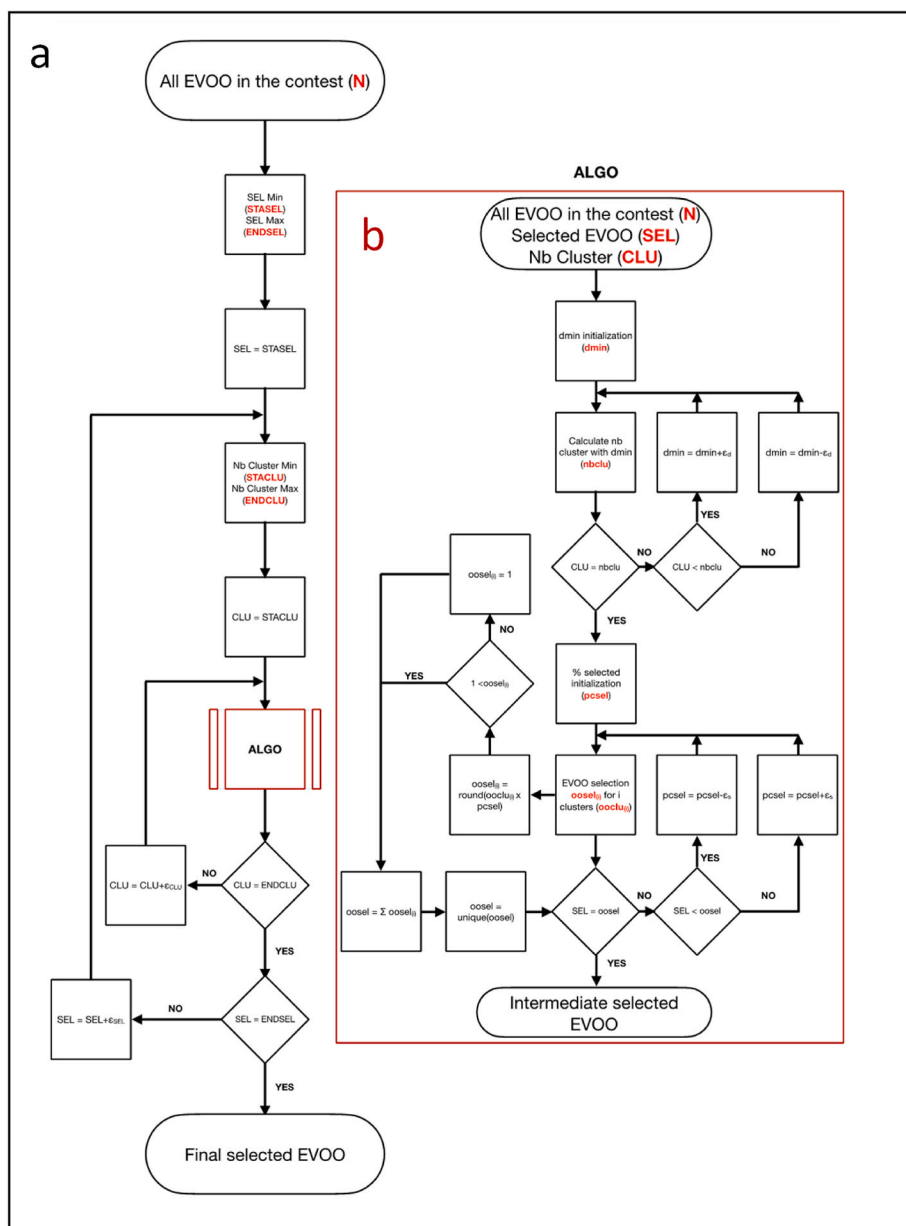


Fig. 1. Flow chart of iWSP algorithm for selecting the best EVOOs among virgin olive oils registered to “World Edible Oils” contest: main organizational units (a) and detailed units (b). CLU: cluster number; d_{min} : minimum diameter of selection area; ENDCLU: maximum cluster usually 20); ENDSEL: maximum selected EVOO (usually between 120 and 200); EVOO: Extra Virgin Olive Oil; N: registered olive oil samples; SEL: selected EVOO; STACLU: minimum cluster (usually 10); STASEL: minimum selected EVOO (usually 20); ϵ_{CLU} : loop increment for the number of clusters (usually 1); ϵ_{SEL} : loop increment for selected EVOO (usually 10). All numbers are integers except d_{min} , $pcsel$, ϵ_d and ϵ_s .

structure domains by analogy with the fruity intensity described by “sweet”, “bitter and pungent”, or “very bitter and/or very pungent.”

The distribution of EVOOs chosen in the first selection stage of the “World Edible oils” competition during the previous three years (from 2020 to 2022) in a 2D plan based on their aromatic maturity and structure scores was shown using the same graphical approach (Fig. 3). To assess the impact of harvest year on the orientation of the scatter plot, a regression line has been generated. There was a little fluctuation in the slope of regression lines, which corresponded to the shift in scores between green and ripe EVOOs. The value of the ordinate at the origin was lower for the 2022 data set compared to the 2021 and 2020 data sets, whose structure scores reached values close to 9.5. These variances may be attributable to yearly fluctuations in the chemical composition of the EVOOs (phenolic compounds, volatile organic compounds, and others) (Angerosa & Campestre, 2013), as well as the organoleptic evaluation of the jury experts (Barbieri et al., 2020).

For each year’s final presentation of AVPA findings, a little alteration was made to the limit values of aromatic maturity and structure scores that define the six normal categories (green/ripe and robust/medium/

delicate) to account for these EVOO classification modifications. In fact, as shown in Fig. 2, the majority of EVOOs are located in the center of the maturity scale, approximately 50 percent. For a standard presentation of results to the public, it is required to position EVOOs in the standard categories, while adhering to the fundamental concept of this work (categories are created arbitrarily, and the influence of their constraints on the awarding of awards must be minimized). In order to maintain a fair quantity of EVOOs in each group, the green-ripe threshold was modified to between 49% and 50%. The light/medium and medium/intense boundaries were then changed in the same manner to create six groups of equal size. In order to have a sense of the EVOO distribution in these categories, the findings for 2021 and 2022 were shown on Fig. 4 according to the harvest year. In addition to the issue of category limitations, this 2D scatter plot reveals an inequality in the density of the EVOOs. In contrast to certain isolated EVOOs, which are surrounded by few oils with comparable organoleptic qualities, there are EVOOs with intense competition surrounding them. Additionally, EVOOs located close or on the boundaries of these categories (red spots on Fig. 4) may be fined if they do not participate in the correct category according to

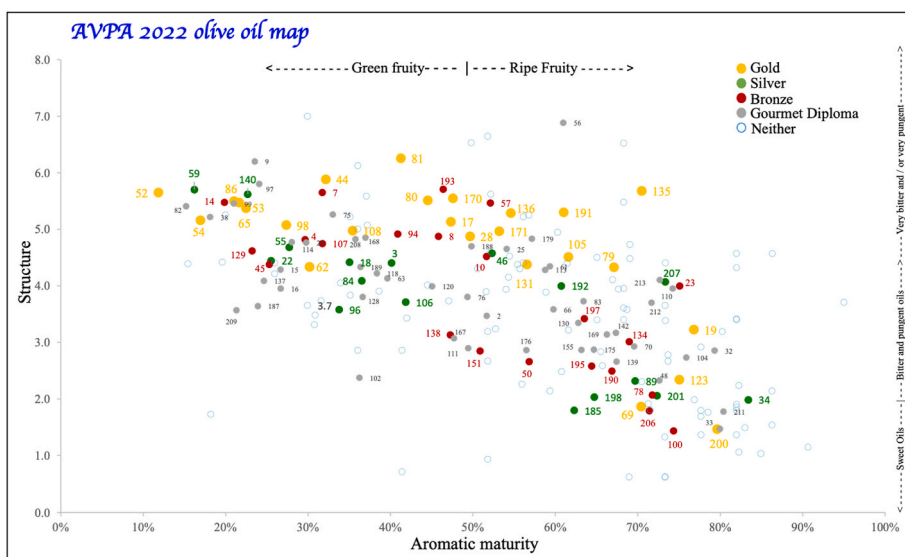


Fig. 2. Result map of « World Edible oil » contest for 2022 year.

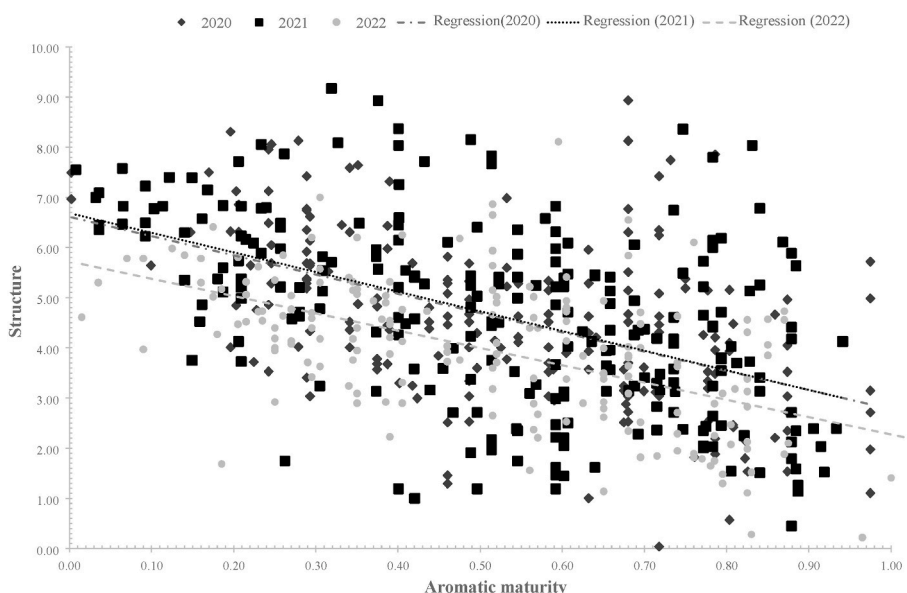


Fig. 3. EVOO distribution after the first selection step of “World Edible Oils” contest, in the 2D plan (aromatic maturity score versus structure score) for the 2020–2022 years.

the rules of specific competitions. A sample that, for instance, would be categorized in the lower portion of the medium green category has little chance of being awarded since it is competing with medium green oils in the higher portion of the category. If, on the other hand, it had been placed towards the top of the delicate green category, it would have a great likelihood of being awarded. However, no competition, with the exception of the one organized by the AVPA, presents a way to avoid penalizing producers whose EVOOs are at the category’s restrictions. Section 3.2 presents a mathematical technique for picking the best EVOOs entered in the “World Edible Oils” competition, regardless of categorization, by constructing a local selection system to evaluate only EVOOs with comparable organoleptic features.

3.2. iWSP algorithm: a construction model of selecting spaces to solve categorization problems

Due to the overabundance of data delivered by measuring devices

that facilitate data acquisition, different algorithms have been developed for selecting a subset of points, as representative as possible of the initial set. Often these methods are used to obtain a reduced amount of data in order to keep the information of good quality. Some are based on distance calculations while others use clusters of points in space (Beal, 2015). Among these methods, WSP algorithm was chosen as the basic algorithm for establishing iWSP for the following reasons.

The iWSP algorithm allowed grouping in restricted areas, EVOOs with organoleptic characteristics locally similar (very little difference in scores of aromatic maturity and structure) and selecting, in these restricted areas, the best EVOOs from their fruitiness’s score.

Unlike k-means algorithm (Lei et al., 2008) in which the number of clusters was fixed initially by the user and OptiSim approach (Clark, 1997) needing to define initially the number of points to be selected, iWSP algorithm iterates over a variable number of clusters and a variable size for these clusters (d_{min}) to obtain the number of clusters including all the EVOOs.

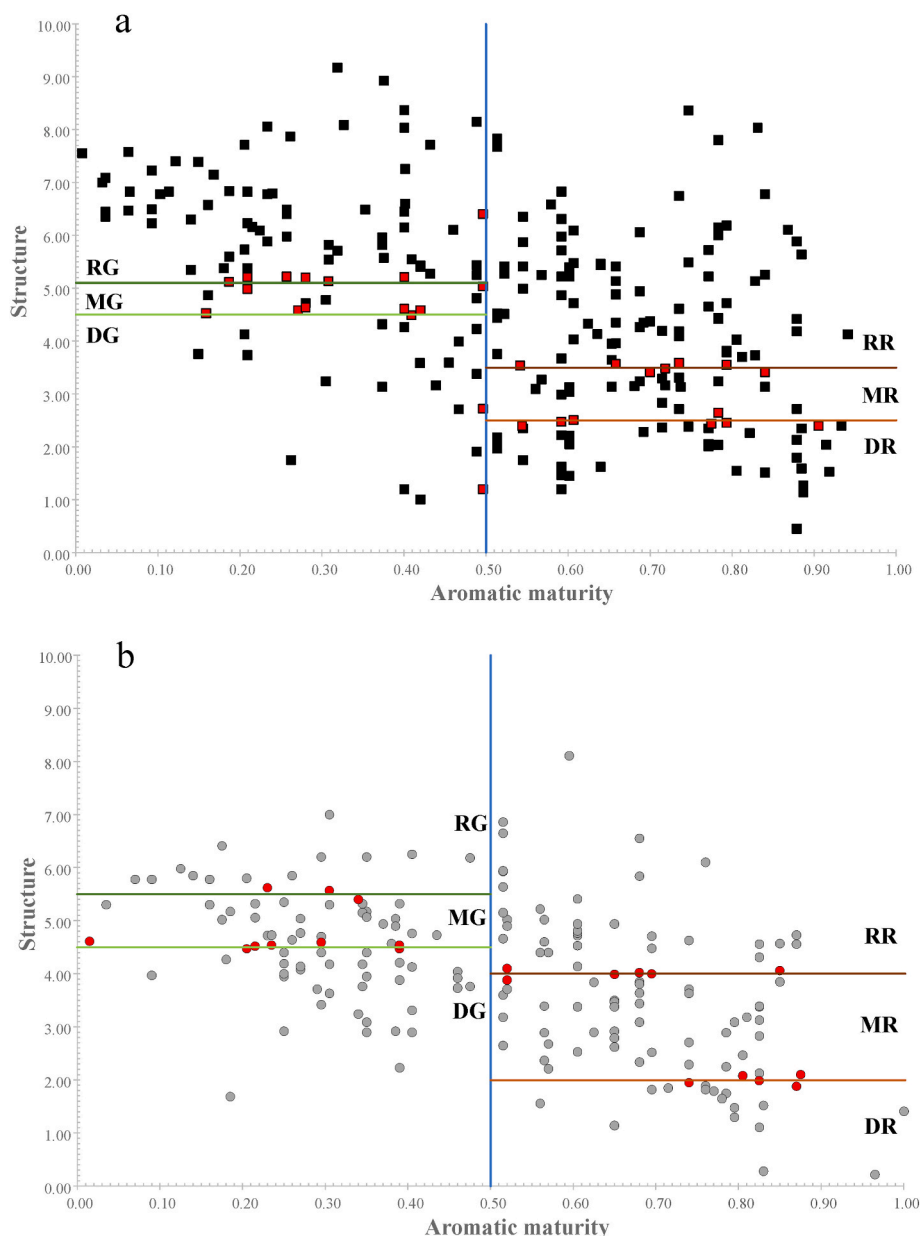


Fig. 4. EVOO distribution after the first selection step of “World Edible Oils” contest, in the 2D plan (aromatic maturity score *versus* structure score) for the 2021 (a) and 2022 (b) years for the visualization of EVOOs placed at the limit of categories. DG: delicate green fruitiness; MG: medium green fruitiness; RG: robust green fruitiness; DR: delicate ripe fruitiness; MR: medium ripe fruitiness; RR: robust ripe fruitiness; horizontal lines: fictive frontiers between the different intensities of fruitiness; vertical line: fictive frontier between green and ripe fruitiness; red points: EVOOs in limit of frontiers.

To consider the influence of the choice of starting point on the final selection of the best EVOOs, different starting coordinates of iWSP algorithm have been considered to define the first candidate EVOO (also called master point) around which the whole selection process took place. A combination of different starting coordinates of the model was used to pave the 2D space (aromatic maturity score *versus* structure score), which has been transformed in an orthonormal coordinate system (scores values between 0 and 10) to simplify the calculations.

In Distance-Based Optimal Design (DBOD) and OptiSim algorithms, the points were selected by sequential substitution or randomly respectively (Clark, 1997; Marengo & Todeschini, 1992). The modified algorithm of Kennard and Stone (KS) also considered different starting coordinates by searching for the closest candidate point to the centre of the domain and then for its furthest point (Daszykowski et al., 2002). But the distribution of points in space was not necessarily representative and uniform in the case of OptiSim, KS and k-means algorithms. In k-means clustering, the datasets with elliptical shapes, for example, posed a problem and the data needed to be reworked before using this algorithm. In iWSP algorithm, the used paving has made it possible to avoid

these disadvantages. As describing previously in the iWSP process (§3.2), different starting points such as the coordinates of each EVOO and nine starting coordinates, comprised between 0 and 10 with an increment of 0.5, were considering to optimize the selection of the best EVOOs.

The only parameter fixed by the user in the iWSP algorithm has been the percentage of EVOOs to be selected, to let the tasting jury adjust it, as he wished, according to the number of registered EVOOs.

Thus, this methodology has showed different advantages because it allowed that (i) all EVOOs belonged to at least one cluster (ii) the percentage of the best selected EVOOs was the same for each cluster (iii) at least one oil was selected by default in the clusters if their number not allowed satisfying the % of EVOOs to be selected (30% for example); this EVOO has been the master point even if only one oil was present, (iv) if an EVOO has belonged to several clusters, it could be selected several times as the best one, but in the final stage of the algorithm (for the choice of EVOOs verifying $30\% \times N = 60$ if $N = 200$), it was only counted once.

For a better understanding, the results of the main steps of this

approach were presented in Fig. 5 for the 2020-year data of the first selection round of “World Edible Oils” contest. In this example, the starting coordinate was the centre of the scatter plot (a calculated point without correspondence with an EVOO and not shown). After the calculation of Euclidian distances between the centre and all EVOOs, the centre-nearest EVOO was considered as the first master point. From this last EVOO, the building of subspaces started. After a projection of aromatic maturity and structure scores in an orthonormal plan (Fig. 5a), the process determined the candidate EVOOs using the centre of the scatter plot as starting point (Fig. 5b). This first step showed already a uniform repartition of the candidate EVOOs among the registered EVOOs. The iterations on the number of the clusters and on their dimension (d_{min} value) allowed to create subspaces recovering the EVOO sample set and ensured that all EVOOs, locally in competition, had similar aromatic maturity and structure scores (Fig. 5c). The final selection of the best

EVOO samples was done in each cluster (or subspace) by keeping a certain percentage of samples with the higher fruitiness score (3D graph of Fig. 5d). Finally, a uniform selection of the best EVOOs (blue points) was obtained among the recorded samples (Fig. 5d) whereas this is not the case with different point selection methods except for the adaptive WSP algorithm (Beal, 2015; Beal et al., 2014). When an EVOO is alone in a cluster, it is not due to the mathematical model but only to the fact that there is no competing EVOO with similar organoleptic properties. Then, this EVOO will be automatically selected by the algorithm even if its fruitiness score is low because it is distinguished from other EVOOs by its aromatic maturity and structure scores; but only the contest jury will approve or not this selection. The fact that an EVOO stands alone may be due to the nature of the variety or a blend of varieties that make up the oil, a terroir effect, an original extraction technology ... but it is not to be dismissed. Thus, the iWSP algorithm gives the chance to all oils to be

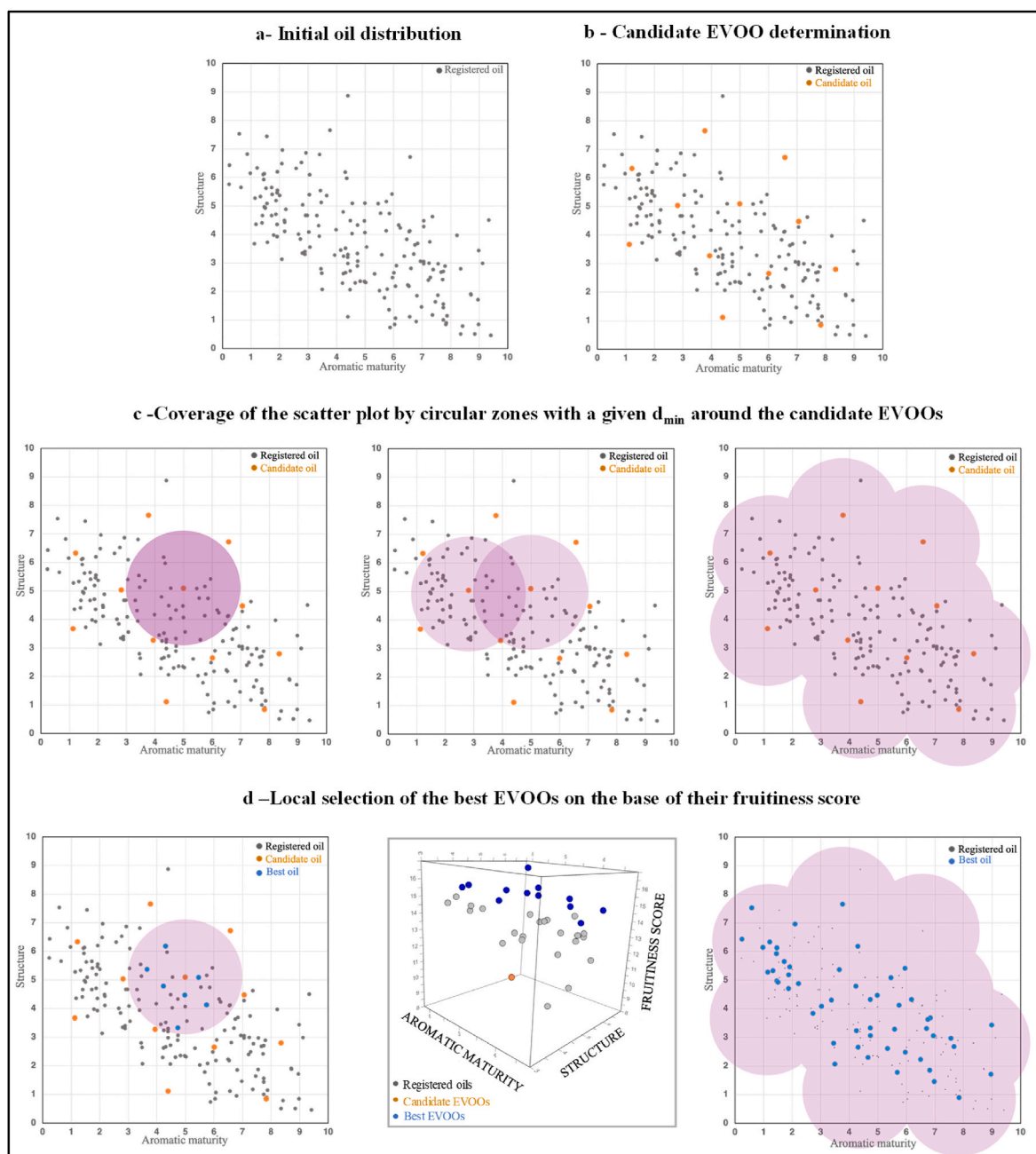


Fig. 5. Schematic visualization of iWSP approach to select the best EVOOs (registered in 2020-year “World Edible Oils” contest) using the scatter plot centre as the starting point.

selected and does away with any initial categorization of the EVOOs registered in the contest. This is no longer a constraint for participants who are no longer obliged to propose a category for their EVOO, (a category which is not necessarily the right one, especially for EVOOs placed at the limit of the categories).

As is common practice in the awarding of prizes in many olive oil contests, it was also possible to present the results of the iWSP selection in a classic form known to the participants. The best selected EVOOs are positioned in the aromatic maturity *versus* structure plan according to the six categories previously defined: green/ripe and robust/medium/delicate (Fig. 6). This figure shows that EVOOs that might not have been selected in some olive oil contests because of their position at the limit of a fruitiness category, are here among the best EVOOs selected by the iWSP approach. This figure illustrates well the fact that the iWSP algorithm leads to a uniform distribution of the best EVOOs to be awarded, independently of one categorization which may induce a bias when selection is done inside a category.

3.3. EVOO classification based on their selection frequencies, according to the different iWSP starting coordinates

The last step of iWSP process also allowed providing an additional contribution for the classification of the best EVOOs for awarding the prizes. To be sure that one EVOO was considered as the best one, numerous simulations have been done by using various starting points and subspace dimensions; then the results obtained for all simulations were compiled. For each considered starting coordinates, the final step of iWSP process conducted to a calculation of a selection frequency for each EVOO. The sum of the selection frequencies of a given EVOO was used to calculate its selection percentage and thus to define its final

ranking: higher the sum, the better the ranking. It was noted that the selection percentage depended on the number of simulations that have converged. However, it was found that some values of the *subspace number* variable could not allow the convergence of the algorithm because a minimum of clusters was required to cover the entire scatter plot. An extract of the calculations with four specific values of *SEL* variable was presented in Table 1. These results were performed with 99 simulations (9 starting coordinates x 11 cluster sizes (d_{min})). The nine starting coordinates corresponded to the following pairs of values (aromatic maturity, structure): (0,0), (0,5), (0,10), (5,0), (5,5), (5,10), (10,0), (10,5) and (10,10). The values of d_{min} varied from 10 to 20 with a step of 1. The selection percentage was the ratio between the number of times the oil is selected as “best” and the number of simulations that have converged. In the whole process with 99 simulations, the average was calculated for 11 *SEL* values while in Table 1 (given as an example), the average was for 4 *SEL* values only. Some EVOOs were never selected whatever the *SEL* value (samples 1 and 24 for example), because they always had close neighbours with a higher fruitiness score whatever the simulation configuration. In the case of EVOO number 14, if only 30 percent of the oils are selected as the best ($SEL = 30$), this EVOO was selected only twice on 83 convergent simulations and if the percent of the oils was increased to 50, this EVOO was selected 27 times on 66 convergent simulations. Thus, this method allowed to consider all the possibilities of selection, by varying the final selection percentage (*SEL*), to obtain a classification having considered all possible situations for a given sample.

4. Conclusions

The iWSP algorithm looks to be an ideal method for picking the top

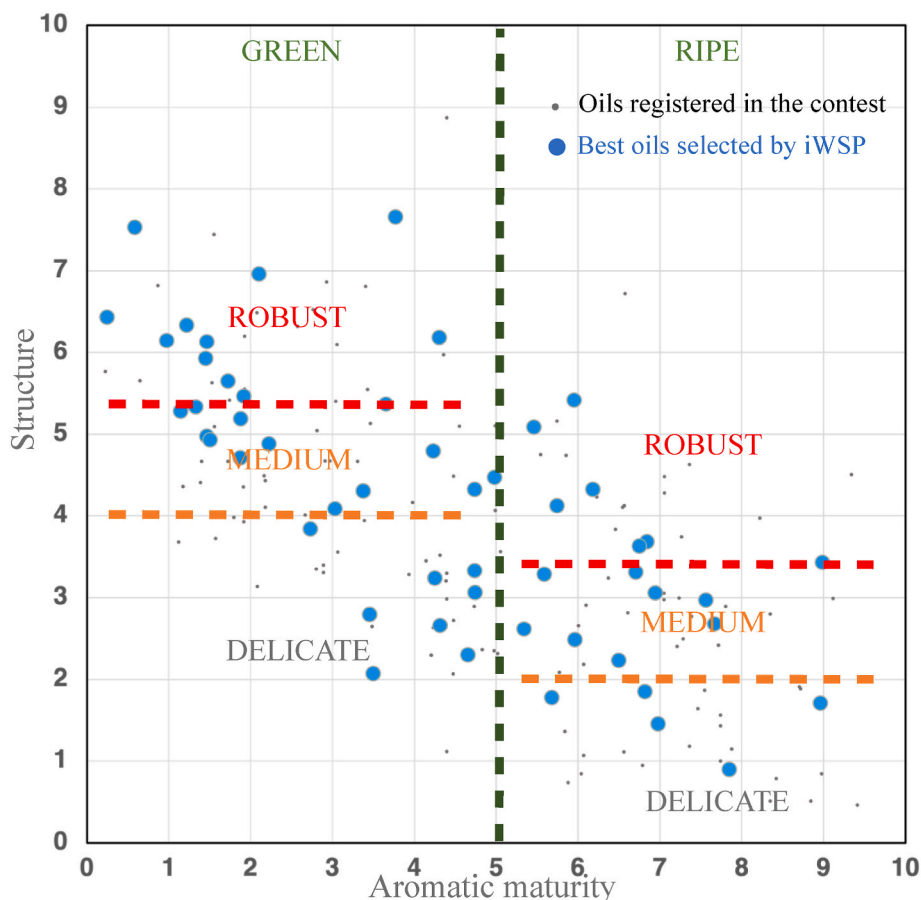


Fig. 6. iWSP’s selection of the best EVOOs to reward and visualization of their repartition in regard to the classical system of categorization.

Table 1

Result examples obtained for the ranking of the first 24 olives oils samples (2nd step of “World Edible Oils” contest, 2022-year, 147 registered oils).

Sample	Scores given by tasting jury after the first selection step			Oil selection FREQUENCY for four values of SEL parameter (30–90) and 99 simulations				Oil selection PERCENTAGE for four values of SEL parameter (30–90) and 99 simulations				Selection PERCENTAGE AVERAGE	Rank out of the oils initially registered	
	Maturity	Structure	Fruitiness	SEL = 30	SEL = 50	SEL = 70	SEL = 90	SEL = 30	SEL = 50	SEL = 70	SEL = 90			
1	0.37	5.07	15.30	0	0	0	0	0	0	0	0	0	0	114
2	0.54	4.15	16.00	0	2	37	63	0	3	45	90	35	35	72
3	0.49	4.88	16.60	5	21	72	70	6	32	88	100	56	56	54
4	0.29	5.02	17.00	9	31	74	70	11	47	90	100	62	62	50
7	0.33	6.13	17.10	0	15	55	55	0	23	67	79	42	42	67
8	0.41	4.45	16.60	3	11	38	57	4	17	46	81	37	37	71
9	0.24	6.19	17.20	0	0	8	32	0	0	10	46	14	14	91
10	0.52	4.52	16.50	7	23	57	70	8	35	70	100	53	53	57
13	0.67	3.35	15.30	0	0	0	7	0	0	0	10	3	3	106
14	0.21	5.17	17.20	2	27	53	59	2	41	65	84	48	48	61
15	0.26	4.19	16.80	0	1	11	55	0	2	13	79	23	23	84
16	0.26	3.97	16.44	0	0	0	4	0	0	0	6	1	1	108
17	0.43	5.58	17.30	50	65	82	70	60	98	100	100	90	90	25
18	0.45	4.01	16.60	35	40	73	70	42	61	89	100	73	73	36
19	0.74	3.36	18.00	83	66	82	70	100	100	100	100	100	100	1
20	0.50	6.54	16.20	0	0	0	0	0	0	0	0	0	0	114
21	0.47	3.75	15.60	0	0	6	31	0	0	7	44	13	13	94
22	0.25	3.93	17.60	83	66	82	70	100	100	100	100	100	100	1
23	0.82	4.32	17.10	21	45	79	70	25	68	96	100	72	72	40
24	0.54	4.55	15.50	0	0	0	0	0	0	0	0	0	0	114
S–Nb				83	66	82	70							

Non-converged simulations are excluded from the percentage calculation. The average of the selection percentage has been calculated for the four SEL values. S–Nb: number of simulations that converged on 99 in total.

EVOOs entered into the “World Edible Oils” competition. This French international competition uses just three organoleptic descriptors to evaluate the sensory qualities of EVOOs. It does not employ particular EVOO categories and a single EVOO selection mode within them, unlike other competitions, so as not to develop a bias at the category boundary. Currently, it was the only competition using mathematical methods for its many participant selection processes. The iterative selection algorithm developed here is novel in that it (i) compares EVOOs with similar organoleptic characteristics; (ii) selects locally the EVOOs with the highest scores; and (iii) proposes to a final sovereign jury, the EVOOs with the highest scores to assist it in rewarding the EVOOs that it considers to be the best. This study is the product of a partnership between taste experts and modelers who intended to ensure the fairness of EVOO prizes in an international competition. In development is a future version of iWSP that will expand this mathematical method for rewarding EVOOs with n descriptors. The subspaces will be constructed using the $(n-1)$ goal descriptors submitted in the different competitions. Hence, Euclidian distances will be computed in $(n-1)$ dimensions. The ultimate decision will be made utilizing the latter dimension. Hence, this methodology may be used to the selection and comparison of certain features of label-food goods (wine, coffee, cheese, and others) in order to promote their quality, enhance their value, and differentiate the places from which they originate. However, the iWSP technique may also be useful for the labeling of technological items based on their technical, environmental, or social features, after these products’ attributes have been measured.

Funding sources

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

CRedit authorship contribution statement

Fabien Girard: Conceptualization, Methodology, Software, Formal analysis, Writing. **Jacques Artaud:** Conceptualization, Methodology, Writing. **Christian Pinatel:** Resources, Reviewing. **Magalie Claeys-Bruno:** Conceptualization, Reviewing. **Catherine Rébufa:**

Conceptualization, Methodology, Formal analysis, Writing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

The authors would like to acknowledge Jean-Emmanuel Jourde, general secretary of AVPA for agreeing to share anonymously the results of the “World Edible Oils” contest.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.foodcont.2023.109776>.

References

- Agency for the Valorization of Agricultural Products (AVPA). <https://www.avpa.fr/avpa>. (Accessed 3 September 2022) Accessed.
- Amelio, M. (2019). Olive oil sensory evaluation: An alternative to the robust coefficient of variation (Cv_r %) for measuring panel group performance in official tasting sessions. *Trends in Food Science & Technology*, 88, 567–570. <https://doi.org/10.1016/j.tifs.2019.02.044>
- Angerosa, F., & Campestre, C. (2013). Sensory quality: Methodologies and application. In *Handbook of olive oil, analysis and properties* (pp. 523–560). Aparicio R. and Harwood J.
- Ballabio, D., Consonni, V., Mauri, A., Claeys-Bruno, M., Sergent, M., & Todeschini, R. (2014). A novel variable reduction method adapted from space-filling designs. *Chemometrics and Intelligent Laboratory Systems*, 136, 147–154. <https://doi.org/10.1016/j.chemolab.2014.05.010>
- Barbieri, S., Brkić Bubola, K., Bendini, A., Bučar-Miklavčič, M., Lacoste, F., Tibet, U., Winkelman, O., García-González, D. L., & Toschi, T. G. (2020). Alignment and proficiency of virgin olive oil sensory panels: The OLEUM approach. *Foods*, 9, 355. <https://doi.org/10.3390/foods9030355>

- Beal, A. (2015). *Description et sélection de données en grande dimension* [Doctoral dissertation]. Aix-Marseille University.
- Beal, A., Claeys-Bruno, M., & Sergent, M. (2014). Constructing space-filling designs using an adaptive WSP algorithm for spaces with constraints. *Chemometrics and Intelligent Laboratory Systems*, 133, 84–91. <https://doi.org/10.1016/j.chemolab.2013.11.009>
- Bendini, A., Valli, E., Barbieri, S., & Toschi, T. G. (2012). Sensory analysis of virgin olive oil. In B. Dimitrios (Ed.), *Olive oil - constituents, quality, health properties and bioconversions* [internet]. London: IntechOpen. <https://www.intechopen.com/chapters/27029>. <https://doi.org/10.5772/29650>.
- Bongartz, A., & Oberg, D. G. (2011). Sensory evaluation of extra virgin olive oil (EVOO). Extended to include the quality factor “harmony”. *Journal of Agricultural Science and Technology A*, 1, 422–435, 2011 [http://olivenoel.ingds.de/pdf/Sensory Evaluation of EVOO extended to include the quality factor Harmony def.pdf](http://olivenoel.ingds.de/pdf/Sensory%20Evaluation%20of%20EVOO%20extended%20to%20include%20the%20quality%20factor%20Harmony%20def.pdf).
- Borràs, E., Ferré, J., Boqué, R., Mestres, M., Aceda, L., Calvo, A., & Busto, O. (2016). Olive oil sensory defects classification with data fusion of instrumental techniques and multivariate analysis (PLS-DA). *Food Chemistry*, 203, 314–322. <https://doi.org/10.1016/j.talanta.2016.04.040>
- Cecchi, L., Parenti, A., Bellumori, M., Migliorini, M., Mulinacci, N., & Guerrini, L. (2022). Clustering monovarietal extra virgin olive oil according to sensory profile, volatile compounds, and k-mean algorithm. *European Journal of Lipid Science and Technology*, 124, Article 2200038. <https://doi.org/10.1002/ejlt.202200038>
- Clark, R. (1997). OptiSim: An extended dissimilarity selection method for finding diverse representative subsets. *Journal of Chemical Information and Computer Sciences*, 37, 1181–1188. <https://doi.org/10.1021/ci970282v>
- Conte, L., Bendini, A., Valli, E., Lucci, P., Moret, S., Maquet, A., Lacoste, F., Brereton, P., García-González, D. L., Moreda, W., & Toschi, T. G. (2020). Olive oil quality and authenticity: A review of current EU legislation, standards, relevant methods of analyses, their drawbacks and recommendations for the future. *Trends in Food Science & Technology*, 105, 483–493. <https://doi.org/10.1016/j.tifs.2019.02.025>
- Daszykowski, M., Walczak, B., & Massart, D. L. (2002). Representative subset selection. *Analytica Chimica Acta*, 468, 91–103. [https://doi.org/10.1016/S0003-2670\(02\)00651-7](https://doi.org/10.1016/S0003-2670(02)00651-7)
- 29 July European Commission delegated Regulation (EU) 2022/2104. (2022). *Supplementing regulation (EU) No 1308/2013 of the European parliament and of the council as regards marketing standards for olive oil, and repealing commission regulation (EEC) No 2568/91 and commission implementing regulation (EU) No 29/2012.*
- Gerhardt, N., Schwolow, S., Rohn, S., Pérez-Cacho, P. R., Galán-Soldevilla, H., Arce, L., & Weller, P. (2019). Quality assessment of olive oils based on temperature-ramped HS-GC-IMS and sensory evaluation: Comparison of different processing approaches by LDA, kNN, and SVM. *Food Chemistry*, 25. <https://doi.org/10.1016/j.foodchem.2018.11.095>, 278–720–728.
- International Olive Oil Council (IOC). (2018). IOC/T.20/Doc. No 15/Rev.10. *Sensory analysis of olive oil, method for the organoleptic assessment of virgin olive oil* <https://www.internationaloliveoil.org/wp-content/uploads/2019/11/COI-T20-Doc.-15-REV-10-2018-Eng.pdf>. (Accessed 3 September 2021) Accessed.
- Lei, X. F., Xie, K. Q., Lin, F., & Xia, Z. Y. (2008). An efficient clustering algorithm based on local optimality of K-means. *Journal of Software*, 19, 1683–1692. [https://doi.org/10.1016/S0031-3203\(02\)00060-2](https://doi.org/10.1016/S0031-3203(02)00060-2)
- Marengo, E., & Todeschportini, R. (1992). A new algorithm for optimal, distance-based experimental design. *Chemometrics and Intelligent Laboratory Systems*, 16, 37–44. [https://doi.org/10.1016/0169-7439\(92\)80076-G](https://doi.org/10.1016/0169-7439(92)80076-G)
- Quintanilla-Casas, B., Julen Bustamante, J., Guardiola, F., García-González, D. L., Barbieri, S., Bendini, A., Toschi, T. G., Vichi, S., & Tres, A. (2020). Virgin olive oil volatile fingerprint and chemometrics: Towards an instrumental screening tool to grade the sensory quality. *LWT - Food Science and Technology*, 121, Article 108936. <https://doi.org/10.1016/j.lwt.2019.108936>
- Rébufa, C., Pinatel, C., Artaud, J., & Girard, F. (2021). A comparative study of the main international extra virgin olive oil competitions: Their impact on producers and consumers. *Trends in Food Science & Technology*, 110, 591–593. <https://doi.org/10.1016/j.tifs.2021.02.019>
- Santiago, J., Claeys-Bruno, M., & Sergent, M. (2012). Construction of space-filling designs using WSP algorithm for high dimensional spaces. *Chemometrics and Intelligent Laboratory Systems*, 113, 26–31. <https://doi.org/10.1016/j.chemolab.2011.06.003>
- Sergent, M., Luu, R. P. T., & Elguero, J. (1997). Statistical analysis of solvent scales. Part 1. *Anales de Química*, 93, 71–75.
- Sergent, M., Luu, R. P. T., Faure, R., & Elguero, J. (1997). Statistical analysis of solvent scales. Part 2. *Anales de Química*, 93, 295–300.
- Sipos, L., Nyitrai, A., Hitka, G., Friedrich, L. F., & Kókai, Z. (2021). Sensory panel performance evaluation-comprehensive review of practical approaches. *Applied Sciences*, 11, Article 11977. <https://doi.org/10.3390/app112411977>