SHIMADZU
Excellence in Science

MALDI-TOF Mass Spectrometry Analysis MALDI-8030

# Application News

# Extra-virgin olive oil authentication: triacylglycerol profiling and machine learning using the Shimadzu MALDI-8020/MALDI-8030 and eMSTAT Solution™

S. Salivo

## User Benefits

- Minimal sample preparation which does not require labor-intensive procedures and excessive solvent use
- Fast, automated sample analysis using a Shimadzu benchtop series linear MALDI-TOF mass spectrometer in positive ion mode
- Easy classification of unknown samples by PCA analysis and machine learning with the Shimadzu eMSTAT Solution software

## ■ Introduction

Extra-virgin olive oil (EVOO) is the highest grade of olive oil due to its unique production process in which olive fruits are crushed solely through mechanical pressing under cold conditions, thus preserving its chemical balance and characteristic properties. EVOO has also been recognised as being nutraceutic, as it provides nutritional and health benefits due to its optimal balance of saturated, monounsaturated and polyunsaturated fatty acids, and minor components such as polyphenols, tocopherols and sterols. These health benefits, and the higher production costs, result in a higher value product which is reflected in the price.

For all these reasons, EVOO is a vulnerable target of adulteration through the addition of alternative, cheaper vegetable oils, in an attempt by producers and traders to fraudulently sell these lower quality adulterated oils as genuine, more expensive EVOO products. Typical vegetable oils which can be found in adulterated EVOOs are palm, canola, hazelnut, pomace, sunflower, corn, soybean, peanut, and others. Besides the economic damage, there are also concerns over the safety of non-genuine EVOOs for human consumption, due to the exposure to potentially hazardous allergens, such those in hazelnut oil.

To combat extra-virgin olive oil frauds, rigorous and efficient quality control assays have been established to protect the identity of this unique product. In the EC Regulation 2568/91, the European Union (EU) provides some analytical methods for the assessment of olive oil quality and authenticity. Among these, those based on triacylglycerols (by high-performance liquid chromatography) and fatty acid and sterol (by gas chromatography) profiles are quite popular. However, these methods often have the drawback of requiring time-consuming extractions, large solvent volumes and long analysis times. Therefore, simplifying and speeding up analysis methods, while preserving the quality and robustness of the data, are necessary steps.

Triacylglycerols (TAGs) account for more than 90% of the chemical content of edible oils. The TAG composition is unique for each type of oil, as the fatty acid expression is plant-specific. Therefore, TAG 'fingerprints' are extremely useful for oil typing, as TAG expression and ratios are a reflection of the fatty acid abundance.

Here, we propose a simple and powerful analytical method to tackle the problem of EVOO fraud. Sunflower oil was used as an example adulterant. The three key steps of the analytical workflow are: 1) minimal sample treatment; 2) fast, automated LDI (matrix-free) analyses of TAG profiles carried out on a Shimadzu benchtop series linear MALDI-TOF mass spectrometer (MALDI-8020/MALDI-8030) in positive ion mode; 3) multivariate analysis and machine learning for the classification of unknown data with the Shimadzu eMSTAT Solution data analysis software (Figure 1). The EVOO adulteration was simulated by spiking sunflower oil into EVOO at 5, 10 and 20% amounts.

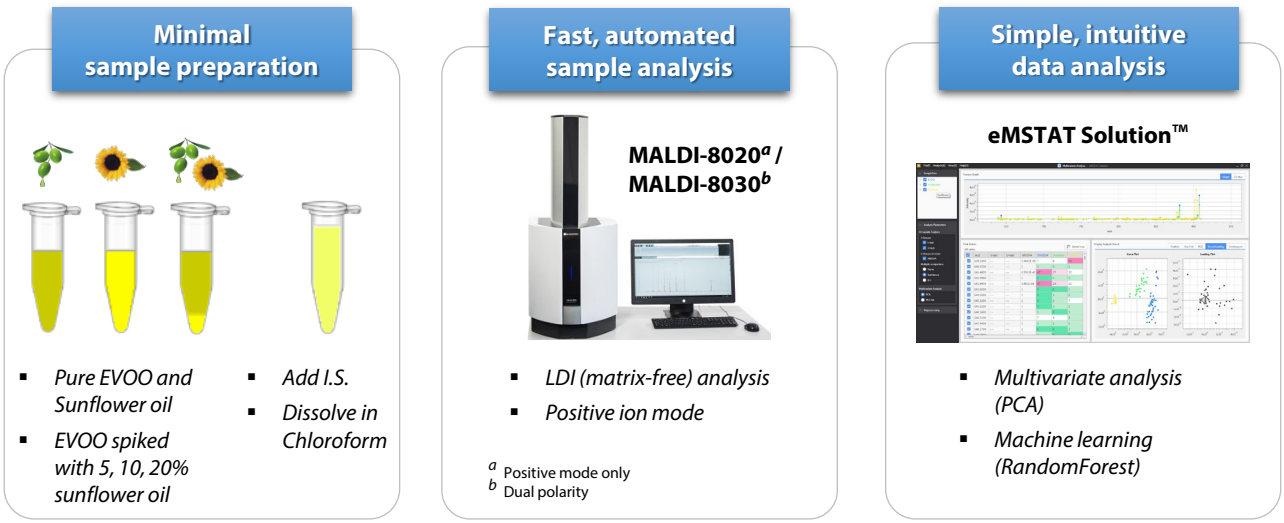


Figure 1. Analytical workflow for the analysis of adulterated extra-virgin olive oils.

## ■ Measurement Conditions and Samples

EVOO and sunflower oils were purchased from local shopping malls. Sample preparation involved dissolution of oil aliquots in chloroform. To simulate the adulteration, mixtures of EVOOs containing 5, 10 and 20% of sunflower oil were prepared. Tricaprin was used as internal standard (I.S.) for mass alignment. To promote the formation of the sodiated TAG species ([M + Na]$^+$), the MALDI target was pre-coated with NaTFA prior to sample spotting. The oil sample solutions were then spotted (0.5 µL) onto the pre-coated MALDI target. LDI (matrix-free) analyses were conducted in positive ion mode on the Shimadzu MALDI-8020 benchtop linear MALDI-TOF mass spectrometer (Shimadzu, Manchester, UK). Note: the MALDI-8020 (positive mode only) and the MALDI-8030 (dual polarity) have identical linear mode performance. All acquired data were processed with the eMSTAT Solutions software (Shimadzu, Japan), and were subjected to multivariate analysis (PCA) and machine learning by means of RandomForest algorithm and KFold cross-validation type.

## ■ Results – Mass spectrometric analysis

TAG profiles were acquired from the pure and adulterated (simulated) oils by LDI (matrix-free) analyses in positive ion mode. Figure 2A shows a comparison between the TAG profiles of EVOO, adulterated EVOO (10%) and sunflower oils (blue, green and yellow traces, respectively). It can be seen how in EVOO, naturally rich in palmitic (P) and oleic (O) acids, the TAGs at *m/z* 881 and 907, i.e., most likely OPO/POO and OOO, are predominant. In sunflower oil, highly rich in linoleic acid (L), the TAGs at *m/z* 877, 901, 903 and 905 (most likely PLL/LPL/POLn, LLL, LLO/LOL and OLO/OOL, respectively) are the most representative. It can be observed how the TAG profile of the adulterated EVOO is highly similar to a genuine EVOO by visual inspection. However, closer examination of the mass spectra reveals how the TAG ratios which are characteristic of a genuine EVOO (e.g., *m/z* 877/907, 881/907, 903/907 and 905/907) are progressively altered as the amount of adulterant oil increases (Figure 2B). Interestingly, the TAG at *m/z* 901 (LLL), characteristic of sunflower oil but not normally expressed in EVOO, is revealed in the EVOO/sunflower mixtures even at the lowest adulteration level tested (Figure 2B).

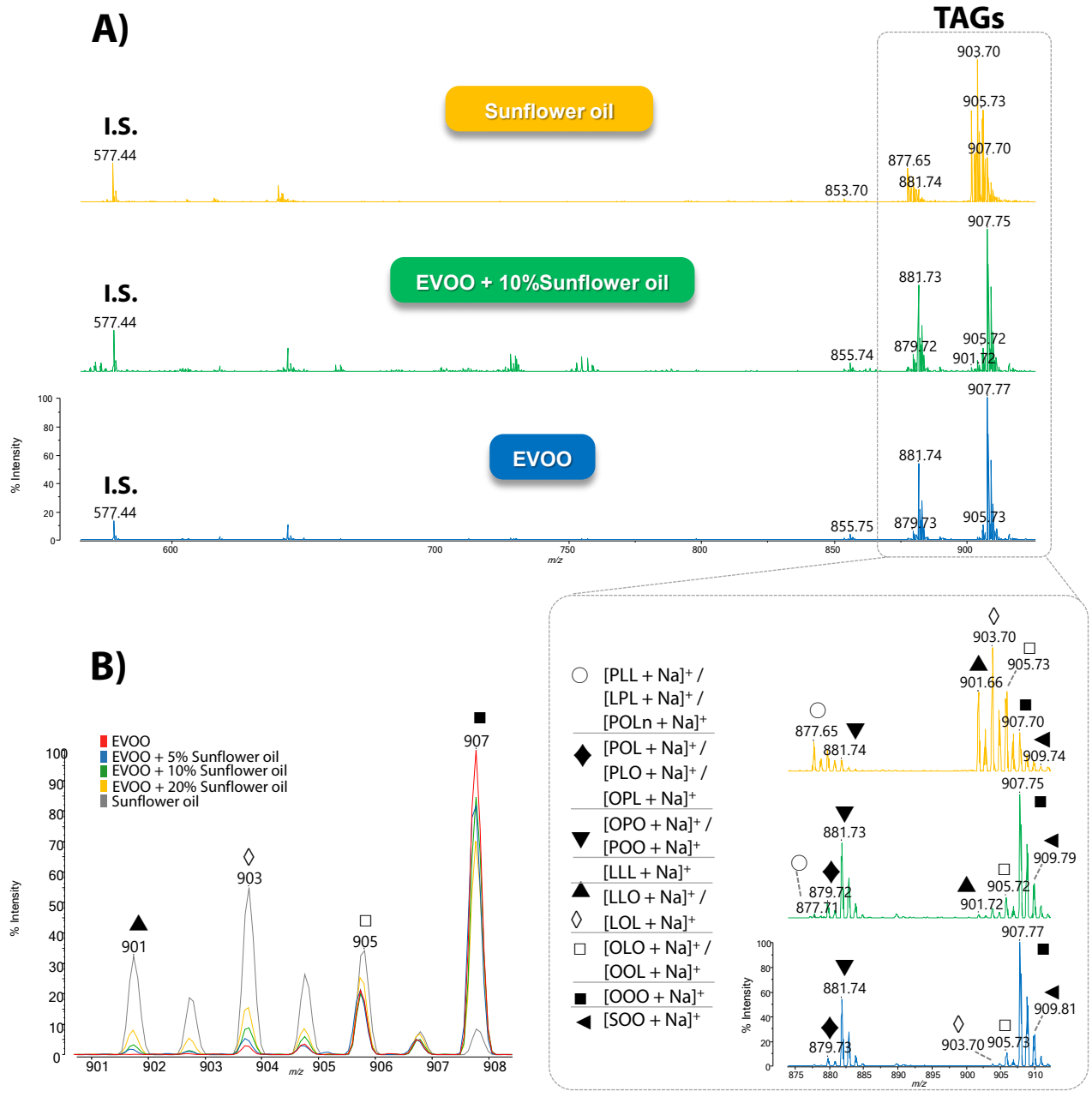


Figure 2. A) TAG profiles of EVOO (blue), adulterated EVOO (10%; green) and Sunflower oil (yellow). Bottom-right inset: expansion on the TAG region showing TAG identity assignment. B) Expansion of the mass spectra in the *m/z* 901-907 region: natural TAG ratios in EVOO are altered by the addition of sunflower oil. The TAG at *m/z* 901, not normally expressed in EVOO, appears in adulterated EVOOs, even at the lowest simulated sunflower oil adulterant level tested (5%).

## ■ Results – Multivariate analysis

Multivariate analysis was carried out on all acquired samples by means of principal component analysis (PCA) and log10 scaling. Prior to the statistical analysis, mass alignment was performed in the acquisition software (MALDI Solutions, Shimadzu) by lock mass method using the Tricaprin I.S. Figure 3A shows the PCA score plot obtained, demonstrating good separation of the three groups: EVOO (blue), adulterated EVOO (green) and Sunflower oil (yellow). To ensure an adequate separation among the three groups, all processing settings i.e., intensity threshold, mass tolerance and scaling methods were carefully optimised. Intensity alignment was executed in the data analysis software (eMSTAT Solution, Shimadzu) using the TAG at *m/z* 903 (LLO/LOL). The choice of the *m/z* 903 TAG was based on a careful selection from the most statistically relevant markers proposed, and this improved the machine learning model further. Figure 3B shows the PCA loading plot, which reveals a number of markers that contribute to the separation: points which are further away from the origin greatly contribute to the separation, and their position on the plot suggests to which group they are more important (e.g., the *m/z* 901 marker is located in the left-hand side of the plot, being highly representative of a sunflower oil). The markers highlighted in red have the most statistically relevant weighting, based on box plot and ROC curve inspection. For example, in Figures 3C & 3D, the box plot and ROC curve of the *m/z* 907 marker demonstrate the statistical significance of this marker for the group separation, exhibiting sensitivity (Y) and specificity (X) of 97% and 98%, respectively, as well as near-to-one area under the curve (AUC) value for adulterated EVOO vs. EVOO (control). Similarly, the ROC curve of the *m/z* 901 TAG demonstrates the statistical strength of this marker, with values of sensitivity (Y) and specificity (X) of 100% and an AUC = 1, with respect to the EVOO control (Figure 3E).

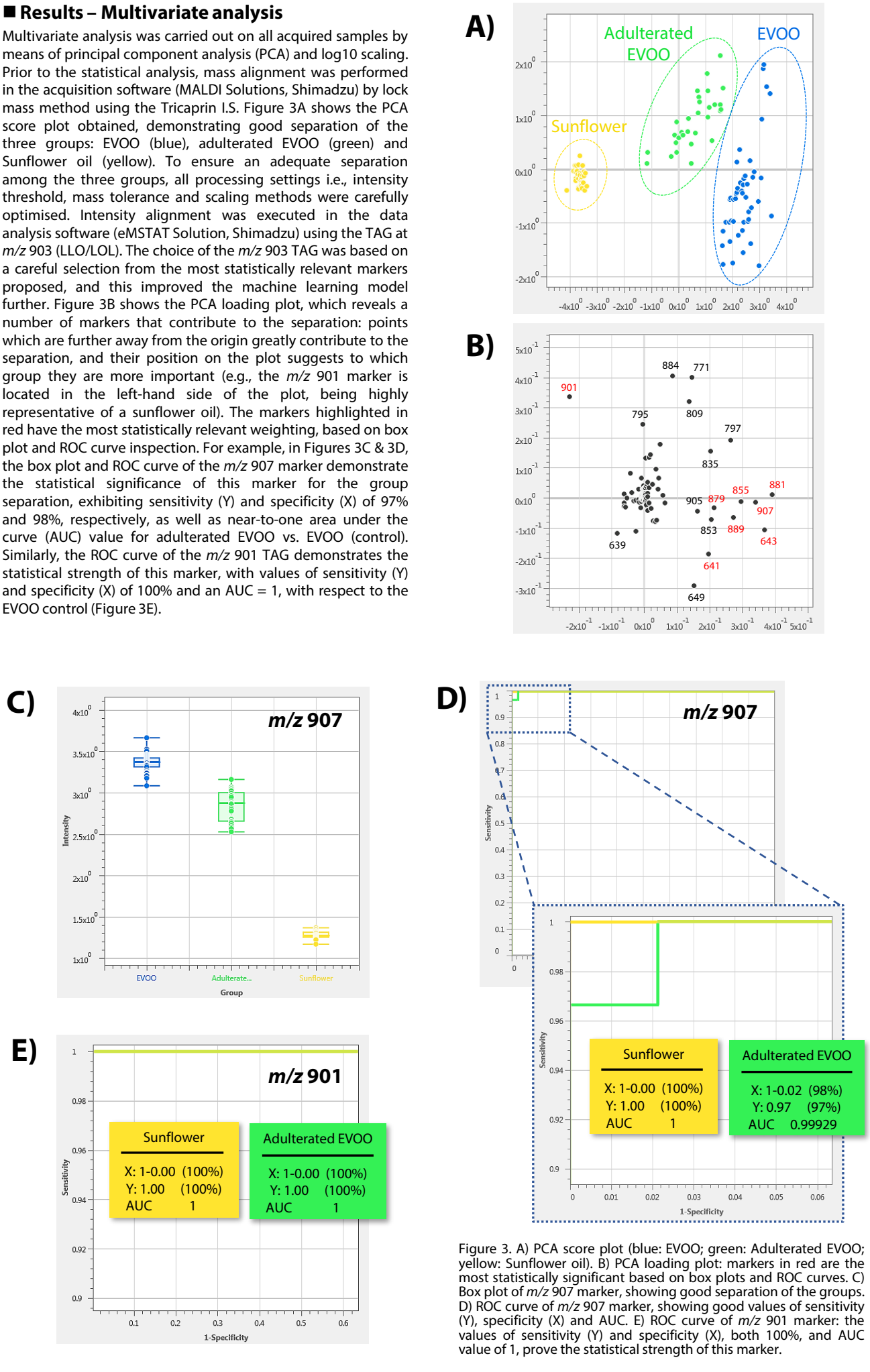


Figure 3. A) PCA score plot (blue: EVOO; green: Adulterated EVOO; yellow: Sunflower oil). B) PCA loading plot: markers in red are the most statistically significant based on box plots and ROC curves. C) Box plot of *m/z* 907 marker, showing good separation of the groups. D) ROC curve of *m/z* 907 marker, showing good values of sensitivity (Y), specificity (X) and AUC. E) ROC curve of *m/z* 901 marker: the values of sensitivity (Y) and specificity (X), both 100%, and AUC value of 1, prove the statistical strength of this marker.

## ■ Results – Machine learning

Discriminant analysis by means of machine learning was carried out based on the PCA classification analysis performed previously. The algorithm used was RandomForest, with KFold cross-validation. A total of 113 data (47 EVOO, 30 adulterated EVOO and 36 sunflower oil) were used to train the model, and 78 'unknown' (blinded) datasets (26 EVOO and 52 adulterated EVOO) were used to validate the model accuracy. All settings were optimised to provide 0% estimated error, which gives an indication of the model accuracy. Figure 4 shows the results of the classification of the 78 'unknown' datasets. While a correct classification of pure EVOOs and EVOOs with higher percentages of sunflower oil is easier to achieve, the biggest challenge arises from the lowest amounts of adulterant oil, e.g., 5%. As can be observed in Figure 4, all 'unknown' data were correctly classified, thus demonstrating the model is 100% accurate. It can be even noted how the distribution of the 'unknown' adulterated data in the plot follows the logic of the amount of adulterant oil, with the 5% adulterated oils being closer to the genuine counterpart, hence reflecting the high similarity of the two types. Conversely, EVOOs adulterated with the highest amount of sunflower oil (20%) are further from the genuine EVOOs and closer to their adulterant counterpart.

## ■ Conclusion

In this application, we present a simple and robust analytical method to pinpoint a very common fraud which affects the olive oil industry, i.e., extra virgin olive oil adulteration with cheaper vegetable oils. We propose a very simple and minimal sample preparation method, which combines with the fast, automated analysis capability of the Shimadzu benchtop series MALDI-TOF mass spectrometers (MALDI-8020/MALDI-8030), and the power of multivariate analysis and machine learning of the Shimadzu eMSTAT Solution software.

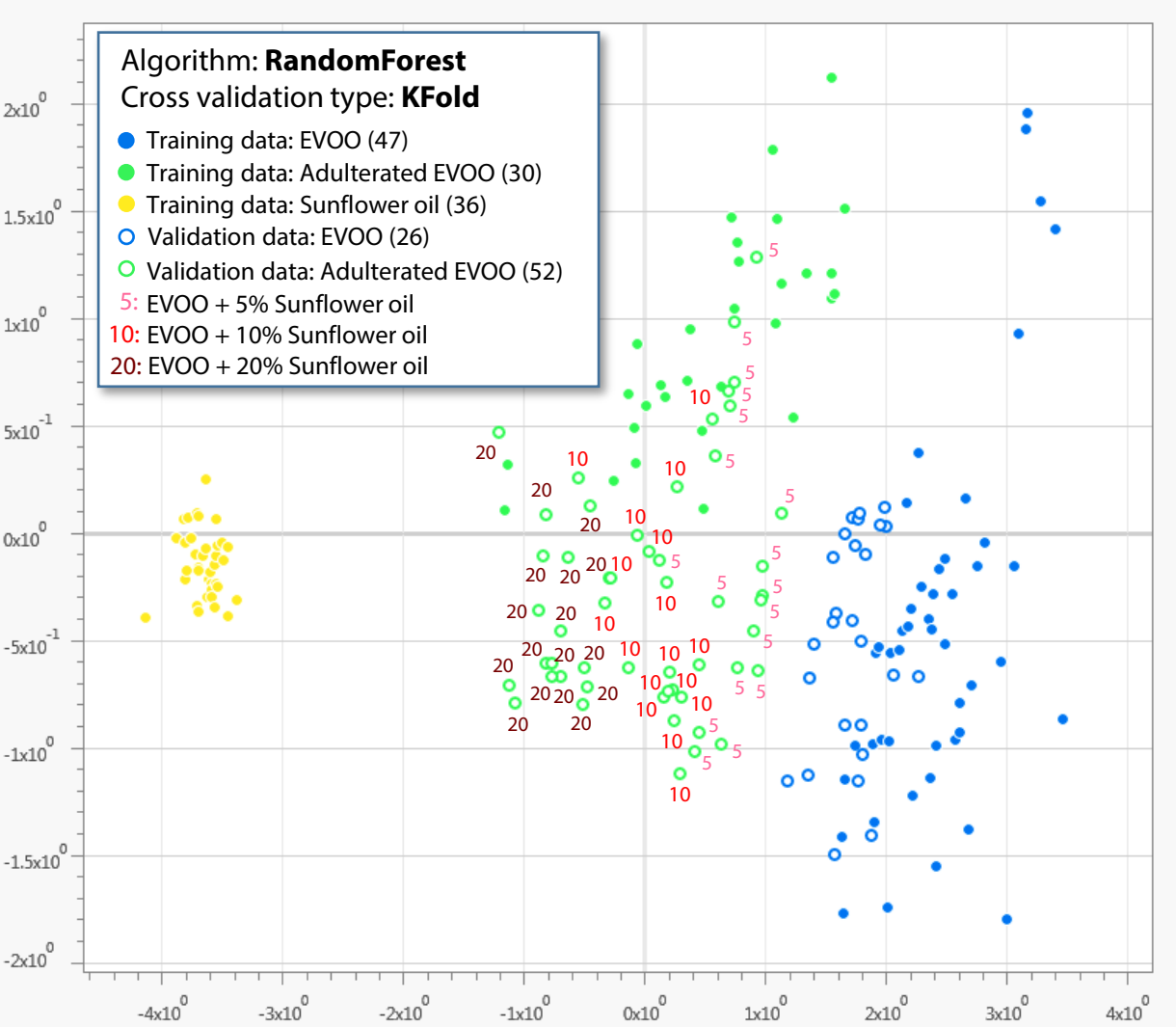


Figure 4. Score plot showing the classification results of the 'unknown' (blinded) data with machine learning (RandomForest; Kfold). 26 EVOOs and 52 adulterated EVOOs 'unknown' data were used to validate the model and its accuracy (100%). The distribution of the 'unknown' adulterated EVOO data in the plot reflects the amount of sunflower oil spiked in EVOO.



Shimadzu Corporation
Analytical & Measuring Instruments Division
Global Application Development Center

www.shimadzu.com/an/

12-MO-485-EN First Edition: Feb. 2022