DATASET OF FLUORESCENCE SPECTRA AND CHEMICAL PARAMETERS OF OLIVE OILS

AN OPEN SOURCE DATASET

Francesca Venturini *
Institute of Applied Mathematics and Physics
Zurich University of Applied Sciences
Winterthur, Switzerland, vent@zhaw.ch
Artificial Intelligence Research and Development
TOELT LLC, Switzerland

Michela Sperti
PolitoBIOMed Lab
Department of Mechanical and
Aerospace Engineering
Politecnico di Torino, Turin, Italy

Umberto Michelucci
Artificial Intelligence Research and Development
TOELT LLC, Switzerland
umberto.michelucci@toelt.ai
Computer Science Department
Lucerne University of Applied Sciences and Arts
Lucerne, Switzerland

Arnaud Gucciardi
Artificial Intelligence Research and Development
TOELT LLC, Switzerland
arnaud.gucciardi@toelt.ai
Artificial Intelligence Laboratory
University of Ljubljana, Ljubljana, Slovenia

Vanessa M. Martos
Department of Plant Physiology
Faculty of Sciences
Biotechnology Institute
University of Granada, Spain

Marco A. Deriu
PolitoBIOMed Lab
Department of Mechanical and
Aerospace Engineering
Politecnico di Torino, Turin, Italy

January 12, 2023

ABSTRACT

This dataset encompasses fluorescence spectra and chemical parameters of 24 olive oil samples from the 2019–2020 harvest provided by the producer Conde de Benalúa, Granada, Spain. The oils are characterized by different qualities: 10 extra virgin olive oil (EVOO), 8 virgin olive oil (VOO), and 6 lampante olive oil (LOO) samples. For each sample, the dataset includes fluorescence spectra obtained with two excitation wavelengths, oil quality, and five chemical parameters necessary for the quality assessment of olive oil. The fluorescence spectra were obtained by exciting the samples at 365 nm and 395 nm under identical conditions. The dataset includes the values of the following chemical parameters for each olive oil sample: acidity, peroxide value, K_{270}, K_{232}, ethyl esters, and the quality of the samples (EVOO, VOO, or LOO). The dataset offers a unique possibility for researchers in food technology to develop machine learning models based on fluorescence data for the quality assessment of olive oil due to the availability of both spectroscopic and chemical data. The dataset can be used, for example, to predict one or multiple chemical parameters or to classify samples based on their quality from fluorescence spectra.

Keywords Fluorescence · Olive Oil · Chemical Parameters · Quality control

*Contact email: vent@zhaw.ch

1 Summary

The dataset presented is a compilation of measurements of analytical chemistry and fluorescence spectroscopy. The dataset includes fluorescence spectra and chemical parameters of 24 Spanish olive oils from the 2019–2020 harvest. The 24 samples were collected at SCA San Sebastián Puente del Ventorro, Benalua de las Villas, Spain. The data were later measured at the Institute of Applied Mathematics and Physics, Zurich University of Applied Sciences, Technikumstrasse 9, 8401 Winterthur, Switzerland. The fluorescence spectroscopy data was acquired by a miniature spectrometer with a 1024 element CCD array that acquires the entire spectrum in one single measurement. The dataset includes a total of 960 spectra (24 oil samples × 2 excitation wavelengths x 20 repeated measurements). Each of the 960 spectra is an array of 1024 values whose elements are the intensity at the different pixel positions. The chemical parameters were determined by accredited laboratories using the procedures described in the European Commission regulation and its amendment Commission [2013, 1991]. These regulations control the methods for the quality assessment of olive oils and provide a decision tree to verify whether an olive oil class is consistent with the declared quality.

The value of the dataset for research purposes is summarized in the points below.

- The data are useful for studying the link between optical properties (fluorescence and absorption spectroscopy), chemical characteristics (such as oil acidity, peroxide value, and fatty acid content), and olive oil quality (extra virgin, virgin, and lampante olive oil).
- This dataset is the first available that contains fluorescence spectra and chemical analysis obtained by accredited laboratories on samples coming from a single producer.
- Many researchers can benefit from the data: computer scientists can use the data to develop machine learning models that link optical to chemical properties; researchers in food technology that are interested in studying chemical properties of olive oil samples of different qualities; engineers that want to develop new optical analysis techniques alternative to the current expensive and time-consuming analytical chemistry methods.
- This dataset can be used to perform explainability analysis to identify spectral characteristics that are related to different chemical properties (e.g., the acidity of the oil). An example is given in the paper Venturini et al. [2023]. This will further advance the understanding of the complex chemical composition of olive oil and its link to its quality and health benefits.
- This dataset can be used to develop instruments based on fluorescence spectroscopy for the rapid and cost-effective quality assessment of olive oil.

2 Data Description

The dataset consists of one CSV file that contains the columns described in Table 1.

A background file[2] is also provided. The file contains 1024 values that correspond to the intensity measured by the spectrometer without any light (dark counts). This spectrum can be subtracted from the raw fluorescence spectra to remove the effect of the dark counts. The same file can be used for the spectra taken at both 365 nm and 395 nm.

The raw fluorescence spectra of selected oils obtained with excitation at 365 nm and 395 nm are shown in Fig. 1.

3 Materials and methods

3.1 Olive Oil Samples

The dataset contains the fluorescence spectra and the chemical parameters of 24 oils. The oils are characterized by different quality categories: 10 extra virgin olive oil (EVOO), 8 virgin olive oil (VOO), and 6 lampante olive oil (LOO) samples. All samples were provided by Conde de Benalúa, Granada, southern Spain, and were prepared from the 2019–2020 harvest. The properties and values of the chemical parameters of the oil samples are listed in Table 2.

For data acquisition, the samples were placed in commercial 4 ml clear glass vials, taking care that no headspace was present to reduce oxidation. All oils were stored in the dark and at 20 °C during the entire time of the measurements.

3.2 Fluorescence Data Acquisition

The fluorescence spectroscopy data were acquired using the portable sensor described in Venturini et al. [2021]. Since already published, only the most relevant characteristics are reported here. The reader is referred to this publication

[2] Fluorescence_olive_oil_dataset_background.csv

Feature	Datatype	Description
Sample	String	Oil sample name: the values are 'D03','D04','D05', 'D06', 'D07' ,'D08', 'D09', 'D10', 'D 19', 'D20', 'D35', 'D38', 'D45', 'D46', 'D47', 'D49', 'D51', 'D52', 'D53', 'D64', 'D77', 'D81', 'D92','D73'
Repetition	Integer	Repetition number. There are 20 repetition for each oil and led: the iteration number goes from 0 to 19)
Led	Integer	Excitation LED identifier: 1 (395 nm), 2 (365 nm)
Data	Float	The fluorescence spectra. The feature is a string composed of 1024 values given between square brackets and seprated by a comma, as for example [1491.0, 1508.0, ..., 1545.0]. Each value is the raw intensity of the fluorescence signal at the given pixel of the detector of the spectrometer.
Quality	String	Quality of the oil. Possible values are 'EXTRA', 'VIRGIN', 'LAMPANTE'
FAEES	Float	Fatty acid ethyl esters in mg/Kg: content of waxes, fatty acid methyl esters and fatty acid ethyl esters
K232	Float	UV Absorbance at 232 nm (K_{270})
K270	Float	UV Absorbance at 270 nm (K_{232})
Acidity	Float	Acidity: expressed as percentage (%) of oleic acid
Peroxide Index	Float	Quantity of those substances in the sample, expressed in terms of milliequivalents of active oxygen per kilogram (mEqO2/Kg), which oxidize potassium iodide.

Table 1: Information on each feature available in the dataset.

for more details. The schematic design of the spectrometer is sown in Fig. 2. The excitation light was provided by two UV LEDs with emission at 365 nm and 395 nm driven by a current driver (MIC4801, Micrel Inc., San Jose, CA, USA) to adjust the excitation intensity. The fluorescence signal was collected by a miniature spectrometer (STS-Vis, Ocean Optics, Dunedin, FL, USA) with a 1024-element CCD array which acquires the entire spectrum in one single measurement with a resolution of 16 nm. The spectrometer was placed at 90° with respect to the LEDs to avoid the excitation light transmitted by the sample to reach the spectrometer. The sensor has a recess where standard 4 ml clear glass vials with the sample can be inserted.

All spectra of the dataset were acquired on undiluted samples at room temperature under identical conditions (illumination intensity, integration time, and geometry) for a quantitative comparison. The integration time was 1 s. During the measurements, the setup was kept in complete darkness to minimize the effect of stray light.

Each spectrum consists of an array of 1024 values (one for each pixel). The value corresponds to the intensity in counts at the different positions of the pixels. To obtain the wavelength (in nanometers) corresponding to each pixel, the following formula can be used:

$$i = a + b \cdot i + c \cdot i^2 + d \cdot i^3 \tag{1}$$

where i indicates the pixel ($i = 0, ..., 1023$) and

$$\begin{aligned} a &= 337.92288208 \text{ nm} \\ b &= 0.4470772743 \text{ nm} \\ c &= 3.55128 \cdot 10^{-5} \text{ nm} \\ d &= -8.38601 \cdot 10^{-9} \text{ nm} \end{aligned} \tag{2}$$

Calibration parameters were provided by the spectrometer manufacturer. All spectra correspond to the raw data without any data processing (smoothing, background subtraction, or normalization). Since all the measurements were done under identical conditions the intensities are directly comparable.

3.3 Chemical Analysis

For each olive oil sample, the dataset includes the values of the following chemical parameters: acidity, peroxide value, K_{270}, K_{232}, ethyl esters concentration and the samples quality class (EVOO, VOO, or LOO) (see Tab. 2).

The chemical parameters were determined by accredited laboratories using the procedures described in the European Commission regulation and its amendment (Commission [2013, 1991]).

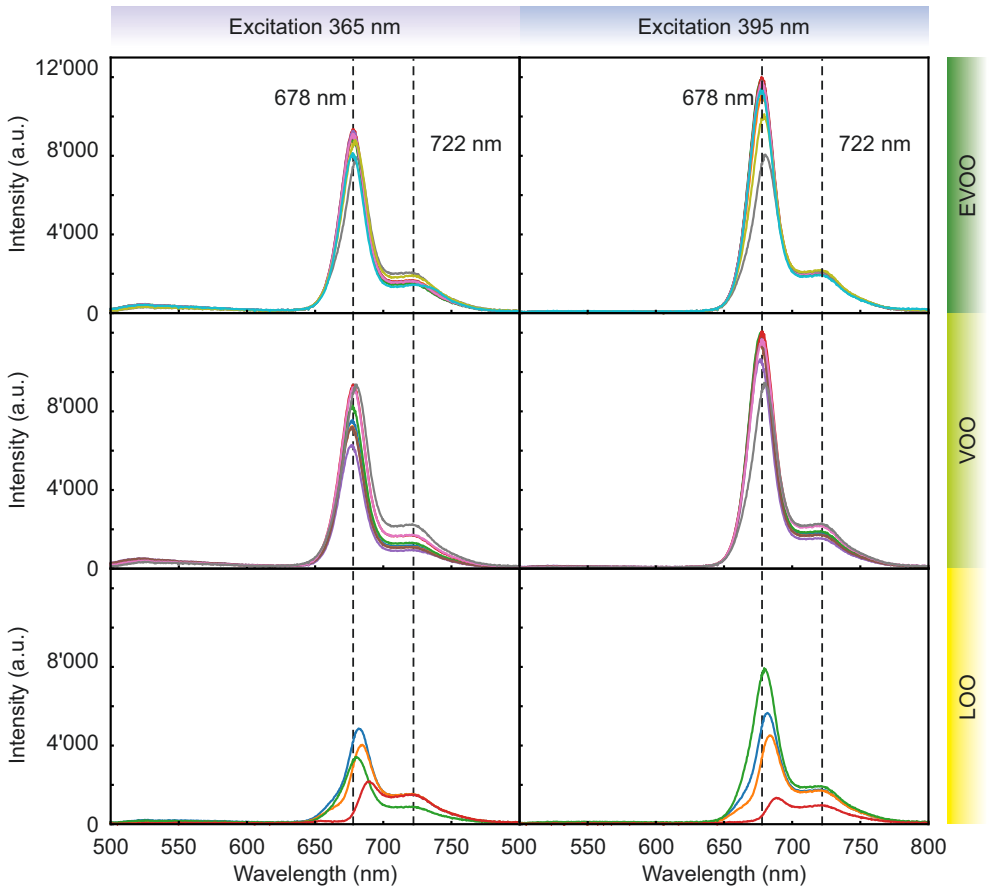

Figure 1: Fluorescence emission spectra of selected olive oils divided in the quality classes EVOO, VOO and LOO. On the left: spectra obtained with excitation at 365 nm; on the right: spectra obtained with excitation at 395 nm. Each curve shows a single spectrum without averaging or smoothing after the background subtraction. Reproduced from Venturini et al. [2023].

4 Funding

This research was supported by the projects: “VIRTUOUS” funded by the European Union’s Horizon 2020 Project H2020-MSCA-RISE-2019 Grant No. 872181; “SUSTAINABLE” funded by the European Union’s Horizon 2020 Project H2020-MSCA-RISE-2020 Grant No. 101007702; “Project of Excellence” from Junta de Andalucia-FEDER-Fondo de Desarrollo Europeo 2018. Ref. P18–H0-4700.

5 Author Contributions

Conceptualization: Francesca Venturini and Umberto Michelucci; methodology: Francesca Venturini and Umberto Michelucci; software, Michela Sperti and Arnaud Gucciardi; validation, Francesca Venturini and Umberto Michelucci; formal analysis, Francesca Venturini and Umberto Michelucci; investigation, Francesca Venturini and Umberto Michelucci; resources, Vanessa M. Martos; data curation, Michela Sperti and Arnaud Gucciardi; writing, original draft preparation, Francesca Venturini and Umberto Michelucci; writing, review and editing, Francesca Venturini, Umberto Michelucci, Arnaud Gucciardi and Marco A. Deriu; funding acquisition, Vanessa M. Martos and Marco A. Deriu. All authors have read and agreed to the published version of the manuscript.

6 Data Availability

The data presented in this study are openly available in Dataset of Fluorescence Spectra and Chemical Parameters of Olive Oils at `https://data.mendeley.com/datasets/thkcz3h6n6/6`, DOI: 10.17632/thkcz3h6n6.6.

Label	Acidity (%)	Peroxide value (mEq O_2/kg)	K_{270}	K_{232}	FAEES (mg/Kg)	Quality
D03	0.35	8.4	0.123	1.435	26	VOO
D04	0.34	8.6	0.108	1.403	40	VOO
D05	0.36	10.3	0.112	1.44	18	VOO
D06	0.31	9.2	0.151	1.484	18	VOO
D07	0.50	8.9	0.150	1.537	47	VOO
D08	0.40	8.5	0.158	1.546	25	VOO
D09	-	-	-	-	-	LOO
D10	-	-	-	-	-	LOO
D19	0.25	4.9	0.13	1.540	10	EVOO
D20	0.26	4.6	0.14	1.540	10	EVOO
D35	0.17	6.4	0.12	1.63	8	EVOO
D38	0.16	6.4	0.12	1.63	9	EVOO
D45	0.17	4.9	0.12	1.63	7	EVOO
D46	0.18	5.0	0.13	1.63	8	EVOO
D47	0.18	5.2	0.13	1.64	16	EVOO
D49	0.9	9.9	-	-	-	LOO
D51	2.16	-	-	-	-	LOO
D52	1.78	22	-	-	-	LOO
D53	0.7	8.7	-	-	-	LOO
D64	0.2	7.1	0.13	1.63	29	VOO
D73	0.2	8.9	0.14	1.66	15	EVOO
D77	0.24	10.4	0.13	1.74	26	VOO
D81	0.16	4.9	0.12	1.63	9	EVOO
D92	0.18	5	0.17	1.91	15	EVOO

Table 2: List of olive oil samples and their physicochemical characteristics. FAEES: fatty acid ethyl esters, EVOO: extra virgin olive oil, VOO: virgin olive oil, LOO: lampante olive oil.

7 Ackowledgments

The authors would like to thank Michael Baumgartner and Ivo Herzig (Institute of Applied Mathematics and Physics, Zurich University of Applied Sciences, Winterthur, Switzerland) for help for the realization of the sensor, and Josep Palau Caballero and Arturo Jimenez (SCA San Sebastián Puente del Ventorro, s/n, 18566 Benalua de las Villas, Spain) for providing the oil samples.

8 Conflicts of Interest

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

9 Abbreviations

The following abbreviations are used in this manuscript:

LOO	Lampante Olive Oil
EVOO	Extra Vigrin Olive Oil
VOO	Virgin Olive Oil
CCD	Charge-Coupled Device
LED	Light Emitting Diode
UV	Ultraviolet
FAEES	Fatty Acid Ethyl Ester

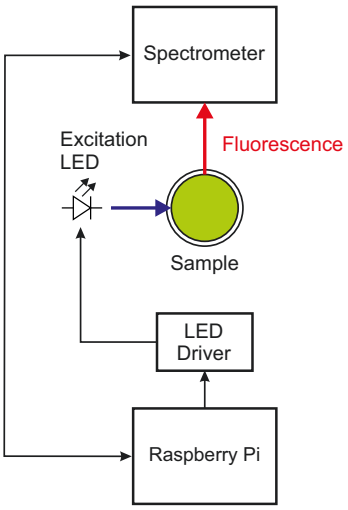

Figure 2: Schematics of the portable fluorescence sensor. Blue: excitation light, red: fluorescence light. From Venturini et al. [2021].

References

European Commission. Commission implementing regulation no 1348/2013 of december 17 2013. *Official Journal of the European Union*, 338:31–67, 2013.

European Commission. Commission regulation (eec) no. 2568/91 of 11 july 1991 on the characteristics of olive oil and olive-residue oil and on the relevant methods of analysis official journal l 248, 5 september 1991. *Offic. JL*, 248:1–83, 1991.

Francesca Venturini, Michela Sperti, Umberto Michelucci, Arnaud Gucciardi, Vanessa M Martos, and Marco A Deriu. Extraction of physicochemical properties from the fluorescence spectrum with 1d convolutional neural networks: Application to olive oil. *Journal of Food Engineering*, 336:111198, 2023.

Francesca Venturini, Michela Sperti, Umberto Michelucci, Ivo Herzig, Michael Baumgartner, Josep Palau Caballero, Arturo Jimenez, and Marco Agostino Deriu. Exploration of spanish olive oil quality with a miniaturized low-cost fluorescence sensor and machine learning techniques. *Foods*, 10(5):1010, 2021.