Warsaw, 13/09/2022

dr hab. Małgorzata Wroniak, prof. SGGW
Division of Fat and Oils and Food Concentrates Technology
Department of Food Technology and Assessment
Warsaw University of Life Sciences

Review
of the doctoral dissertation **Qian Ying** MSc ing.
"**The Composition of Triacylglycerols as an Indicator of Edible Oil Adulteration**"
done at Department of Food Technology of Plant Origin
Faculty of Food Science and Nutrition, Poznan University of Life Sciences
under the supervision of prof. dr hab. Magdalena Rudzińska
and academic supervision of dr inż. Arkadiusz Majewski
Poznan University of Medical Sciences

The basis for preparing this review is a letter of July 13, 2022 from prof. dr. hab. Magdalena Rudzinska informing about the resolution of June 30, 2022 of the Scientific Council of the Food and Nutrition Technology Discipline at the University of Life Sciences in Poznan with a request to assess the doctoral dissertation. The review was carried out following the guidelines of the Act from 14 March 2003 on scientific degrees and scientific titles, as well as degrees and titles in the field of art (Journal of Laws of 2017, item 1789).

The doctoral dissertation of Qian Ying MSc ing. focuses on developing a method for detecting adulteration of edible vegetable oils using triacylglycerols as input data for artificial neural networks. Food adulteration is the deliberate substitution and falsification of food, raw materials, or ingredients that are placed on the market to obtain economic benefits. The adulteration of edible oils manifests as blending expensive oil with cheap oils. Such measures are taken to reduce production costs, which affects price competitiveness at the expense of honest producers. Another reason for this could be the desire to conceal the lower than declared quality of an oil. Unfortunately, there is at present no accurate method to identify the adulteration of vegetable oils. While chromatography is still the most popular method of identifying oil fraud, it is not sufficient to differentiate all of the many edible oils. TAGs are the main components of vegetable oils and fats and can be used as indicators for identifying adulteration, given that they have different compositions in different plant oils. Artificial neural networks are widely used in the food field, such as for predicting the origin of raw materials, optimizing processing methods, and obtaining the best quality food. There have only been limited studies on the use of ANNs in detecting adulteration and predicting the quality of plant oils, so this study selected eight commonly used oils as samples for establishing ANN models.

The detection of adulteration is one way of establishing food defense to enhance their food security. Therefore, I consider that the topic and scope of research in this work are accurate and up-to-date, particularly in this era of seeking new, simple, quick methods of food analysis that can be

MW

used to assess its quality and authenticity. It is critical to look for innovative, low-labor, low-cost methods that can be easily applied in laboratory practice to test for adulteration of edible oils.

Evaluation of a doctoral dissertation's layout

The doctoral dissertation submitted for evaluation is an experimental dissertation, with a typical layout, written in English. The dissertation covers 7 chapters, 126 pages in total, in the following layout: abstract (in English and Polish) (4 pages), abbreviations (2 pages), introduction (2 pages), literature review (18 pages)), aims of the thesis, material and methods (7 pages), results and discussion (32 pages), summary (2 pages), conclusions (2 pages), references (13 pages) acknowledgments and a list of tables in the thesis (2 pages), a list of figures in the thesis (2 pages) and an appendix with a list of tables (42 pages). The content of the work contains 20 tables and 15 figures, and additionally, an annex contains 27 tables with detailed results of the analysis. To prepare the work, the Author used 90 literature items. The literature used is in English and from the last decade account for 74%. There are some minor flaws in the list of literature, including missing bibliographic data in Carbone et al. 2001, Fang et al. 2015, Wang from 2015, and Latin names should be in italics. The work is generally written in the correct language, however, there are typos, and minor stylistic and editorial errors. The obtained research results were subjected to statistical analysis and correctly presented in clear and carefully prepared tables and figures. The layout of the work, and division into chapters is correct and typical. Therefore, I conclude that the thesis presented for review meets the formal requirements for dissertations for the doctoral degree.

Evaluation of the literature review

The literature review is presented on 18 pages in 5 subsections. The Author first characterized the main components of edible oils (triacylglycerols, fatty acids, sterols), and discussed the technologies for obtaining them. Then Author focused on the characterization of methods of adulteration of edible oils. Finally, the Author discussed artificial neural networks. To summarize, the literature review was well planned and executed. However, some minor inaccuracies and simplifications were most likely caused by the incorrect/imprecise use of literature sources, but these do not diminish the substantive value of this study. For example, the data on sunflower oil on page 10 differ from those in the table, and the fatty acids in the table do not add up to 100%. On page 13 it is a great simplification to say that oils are obtained by two methods, i.e., by pressing and refining. Refining is an oil refining process, not an oil extraction method. The technological process of obtaining oils presented here is not universal and typical for obtaining oils from seeds - it is rather typical of olive oil extraction. The production data on page 17 is a decade old and out of date.

MW

Unfortunately, this chapter doesn't mention palm oil, which is the world's most important oil in terms of production and consumption, including in Europe.

I have a few questions regarding the literature review: 1) Please describe the technology of obtaining oils from seeds, e.g., refined rapeseed oil. How is crude oil obtained? 2) Please explain, are adulterated and contaminated oils the same? What is the falsification of edible oils with mineral oil, and what amounts are detected in the tests (in %)? 3) Please explain how to distinguish between cold pressed and refined oils and how to detect adulteration.

Evaluation of the aims of the thesis

The purposefulness of the undertaken research was presented and justified by the Author in chapter „Aims of the thesis". The main purpose of the thesis was specified in this chapter, which was to evaluate the ability of an artificial neural network to identify vegetable oil making use of the composition and the content of endogenous components such as triacylglycerols, fatty acids, and sterols. Then, 7 research tasks were presented: evaluating the HTGC-FID method for determining TAGs in vegetable oils; evaluating the HPLC-ELSD method for determining TAGs in vegetable oils; collecting vegetable oil samples; determining TAGs, FAs, and sterols in almost fifty samples from eight kinds of vegetable oil; using PCA to dimensionality reduction inputs for the development of ANN; training ANN for selection the best Models of ANN; testing ANN models and explored the threshold.

No research hypothesis was presented in the paper. However, in my opinion, the research goal and tasks were correctly defined and fully correspond to the topic of the dissertation. Figure 4 was placed here by mistake – it should be in the next chapter.

Evaluation of research methods

The experimental part of the work includes the chapter "Material and methods", which provides the characteristics of the oils and chemical reagents used, and then the analytical and statistical methods are described. The research material consisted of 8 types of vegetable oils: corn, linseed, rapeseed, sunflower, soybean, pumpkin seed, black cumin seeds, and olive oil, which were purchased in a retail network or obtained directly from producers - 35-50 samples each. The content of triacylglycerols (TAGs) in the tested oils was determined using the GC and HPLC method, sterols - using the GC-FID method, and the percentage of fatty acids - using the GC-FID method. The individual instrumental test methods are described in detail. The obtained results were analyzed by PCA, and then neural network models were built using the Statistical program. A supervised learning algorithm was used in this study. The total number of cases in the dataset was 371, with each variable representing a separate case. It was decided to divide the dataset into three subsets

MW

- the teaching set was allocated 70% of the data, and the rest of the data was split evenly between the validation and test sets. A set of empirical data was processed, which was subjected to component analysis, and then three Artificial Neural Networks (ANN) models were developed. In model 1 they were based on a literature review, while in models 2 and 3 they were selected by principal components analysis (PCA). All ANN parameters, such as the number of layers, the number of neurons in the hidden layer, and the learning algorithm were selected arbitrarily. After multiple attempts, the 2 ANN topology was finally established: one input layer, one hidden layer, and one output layer, and BFGS (Broyden Fletcher Goldfarb Shanno) training algorithms for MLP (Multilayer Perceptron) and RBFT (Radial Basis Function Teaching) for RBF (Radial Basis Function) networks. The oil parameters were the input variables, and one of the eight oils was the output variable. In Model 1, the inputs were the results obtained for OOO, OOL, OOS, OOP, C16:0, C18:0, C18:1, C18:2, campesterol, stigmasterol, β-sitosterol in 8 tested oils. To build Model 2, the number of data entered was limited to OOL, OOO, OOP, POL, C18:1, C16:0, C18:0, and campesterol, while Model 3 was based only on six triacylglycerols: OOO, POL, OOP, LOL, OOS, OOL. Two techniques of neural networks were investigated - MLP and RBF.

The methodology, in my opinion, is lacking information on the set of literature data on TAG, fatty acids, and sterols used in model 1. Please specify the number. What publications did the literature data come from? How much of this information was there? Regardless of the above, I believe that the research material and research methods presented in the paper have been adequately and thoroughly described.

Evaluation of the results and discussion

The most extensive part of the work is the chapter "Results and discussion", which is divided into sub-chapters. The research carried out and described so far allowed the Author to gather 59 parameters relating to FAs, TAGs, and sterols for the eight kinds of oil collected. C16:0, C18:1, C18:2 OOP, POL, OOL, LOL, campesterol, campestanol, stigmasterol, β-sitosterol, and sitostanol were detected in all eight oils. Interestingly, some oils contained characteristic components: for example, β-amyrin and α-spinasterol were found only in sunflower oil and pumpkin oil, respectively. Principle component analysis was carried out and three ANN models were established. The results showed that MLP had more satisfactory results - a valuable finding that will contribute to future experiments. The developed method with the use of ANN and TAG as input data allowed for the identification of oil adulteration at the level of 20% -40%. The optimal topology had a better threshold than the other models, fewer parameters, and was a fast and efficient method of TAG determination. Massive TAG species caused the unique results seen in Model 3. Due to the different importance of the parameters selected for different oils, the detection threshold will vary greatly. However,

when looked at comprehensively, Mode 3 has great advantages. Compared with the complicated extraction process of sterols and fatty acid methylation experiments, the HTGC method can rapidly analyze TAG components. Moreover, Model 3 has only six input variables, which is convenient and reduces the workload of ANN experiments. Model 3 thus has potential. In future studies, it would be better to select the most important TAGs in each oil.

I conclude that the research findings have been accurately presented in a variety of tables and figures. The extensive range of tests and analyses performed enabled the Author to obtain a diverse set of results, providing a solid foundation for the detailed characteristics of the oils. It should be noted that there have been few reports on the examination of the authenticity of oils using artificial neural networks, so the obtained results contribute significantly to the scientific literature in this field. In this chapter, some reservations are raised by the careful preparation of the content (mistaken units of mg/kg or mg / g, no reference to the figures, typos, e.g., C16 instead of C18: 2, repeated content), some inaccuracies and simplifications have also crept here, and the discussion of the obtained results it is often too sketchy, sometimes tables and graphs precede the issues discussed. However, this does not diminish the substantive value of the work and I believe that the obtained results have been sufficiently discussed.

Please answer the following questions about this section of the work: 1) Why was the analysis and identification of minor fatty acids in most oils, such as C20 and C22, abandoned? When examining adulteration in olive oil under EU Regulation, the analysis of trace amounts of fatty acids is crucial. 2) No squalene was found in olive oil in this study; how can this be explained? According to the literature, it ranges from 0.7 to 12 g/kg. 3) I haven't found any information on what hidden layers/neurons are. For example, what are the 9 hidden layers in MLP 11-9-1? Please elaborate.

Evaluation of the summary and conclusions

The Author presented ten very brief statements and conclusions resulting directly from the research in the two-page chapter "Conclusions." The Author demonstrated that an Artificial neural network (ANN) is a useful tool to identify adulteration of vegetable oil and Principal Component Analysis (PCA) can be used to decrease input data for ANN. The unquestionable novelty of this work is the proposal for the application of Model 3, which showed similar performance for training, validation, and testing to Models 1 and 2. This reduces needed input data into the TAG, without the need for FA and sterols determination. It has been demonstrated that the MLP M3.4 *3 artificial neural network model with a 6-15-1 structure based on 6 TAGs' properly identifies adulteration of 8 kinds of vegetable oils with a threshold 20%. The presented conclusions are related to the purpose of the work and research tasks, and they indicate the completion of the planned research.

To summarize, the extensive research planned by the Ph.D. student was consistently carried out, the aim of the work was achieved, and the research tasks were completed. The numerous results of instrumental analyses obtained were subjected to appropriate statistical analysis, carefully presented, adequately discussed, and confronted with the available literature. Correct conclusions were drawn as a result of the findings. The study provides valuable theoretical information as well as ready-made practical solutions in the field of analyzing the quality and authenticity of edible oils. It should be emphasized that the critical remarks and observations presented in the review were intended to improve the Author's scientific and research knowledge.

Conclusion

The doctoral dissertation of Qian Ying MSc ing. is an interesting and innovative application-oriented research paper. It concerns the development method for detecting adulteration of edible vegetable oils using triacylglycerols as input data for the artificial neural network. The work contributes to the advancement of knowledge in the field of food sciences and nutrition, and it takes an innovative approach to the problem at hand. The work was completed using very well chosen, modern, instrumental research techniques, demonstrating that the Author has mastered the research technique, can pose, and solve research problems, allowing for independent research, and indicates extensive theoretical knowledge on the topic represented.

To summarize, Mrs. Qian Ying's work, titled "The composition of triacylglycerols as an indicator of edible oil adulteration," meets the requirements for doctoral dissertations under the Act of March 14, 2003, on academic degrees and scientific titles, as well as degrees and titles in the field of art (consolidated text: Journal of Laws of 2003, No. of 2014, item 1852, as amended), and I am applying to the Scientific Council of the Food and Nutrition Technology Discipline, University of Life Sciences in Poznan for allowing the admission of Qian Ying MSc ing. to the next stages of the doctoral dissertation.

Małgorzata Wroniak