

## **Application of FT-MIR spectroscopy for the discrimination of geographical origin of thyme honeys**

Dafni Dimakopoulou-Papazoglou<sup>\*a</sup>, Kyriaki Chatzikyriakidou<sup>a</sup>, Nikolaos Ploskas<sup>b</sup>, Konstantinos Koutsoumanis<sup>a</sup>, Eugenios Katsanidis<sup>a</sup>

<sup>a</sup> Department of Food Science and Technology, School of Agriculture, Faculty of Agriculture, Forestry and Natural Environment, Aristotle University of Thessaloniki, Greece

<sup>b</sup> Department of Electrical and Computer Engineering, Faculty of Engineering, University of Western Macedonia, Greece

### **Abstract**

Honey has been characterized as a functional food with high commercial value. In recent years, many consumers prefer monofloral honeys as they have specific nutritional characteristics and unique aroma and flavor. The specific characteristics of honey are provided both by the botanical and the geographical origin. The aim of the present study was to differentiate thyme honey samples with regards to their geographical origin, using FT-MIR spectroscopy, combined with multivariate statistical analysis. A total of 74 thyme honey samples were collected from different geographical regions of Mediterranean countries, namely Spain (17), Greece (26) and Tunisia (31). For all samples, FT-MIR spectra were acquired using a Jasco FTIR 6700 spectrophotometer equipped with an 3-reflection ATR diamond in the region of 4,000 – 400 cm<sup>-1</sup>. Regarding spectra pre-processing, different methods were evaluated, such as smoothing with the Savitzky-Golay algorithm, standard normal variate (SNV) and first and second derivatives. Principal Component Analysis was used for dimensionality reduction and then the Random Forest classifier was utilized for this multi-class classification problem with a 10-fold cross validation, using Python. The experimental results indicate that our predictive model can identify the geographical origin of thyme honeys at an accuracy of 91.33%, F1-score of 99.33%, precision of 99.33%, recall of 99.33%, on average for the test set. In conclusion, predictive models trained on FT-MIR spectral data could be successfully used to predict the geographical origin of honey samples.

**Keywords:** Thyme honey; Mediterranean area; FT-MIR; Geographical origin; Random Forest classifier

**Acknowledgments:** This research is part of the PRIMA programme. Grant agreement No 1932, project MEDIFIT (Call 2019 Sec. 1 Agrofood IA).