Classification of Soybean Paste Using Laser-Induced Breakdown Spectroscopy and *k*-Nearest Neighbors Analysis

Sang-Ho Nam[1,*] and Yonghoon Lee[1]

[1] Department of Chemistry, Mokpo National University, 1666 Yeongsan-ro, Jeonnam 58554, Republic of Korea
[*]Corresponding author email: shnam@mokpo.ac.kr

Laser-induced breakdown spectroscopy (LIBS) is one of the simple elemental analysis method using laser-ablation sampling [1]. In typical LIBS analysis, the sample preparation can be simply done without acid digestion and dilution process and thus allows a rapid analysis of multiple elements. In this study, we conducted elemental discrimination of soybean paste according to its geographical origin using LIBS. Soybean paste is a seasoning widely used in East Asian countries. It is characterized by its high protein and amino acid content, excellent storage stability, and distinctive flavor and aroma, making it widely favored in culinary practices. In South Korea, numerous food companies distribute soybean paste products, with a significant portion being produced in China. Distinguishing the origin of soybean paste is crucial for food distribution management. A total of 167 soybean paste samples were used in this study, with 101 being domestically produced and 66 from China. LIBS equipment measurements revealed the presence of elements such as C, Mg, Na, K, H, Ca, Cl, and P in soybean paste samples. The discrimination capabilities of the emission lines of these elements were investigated by using the concept of interclass distance. To evaluate the interclass distance, the difference between average of specific emission intensities measured for Chinese and Korean soybean paste samples was taken first and scaled by the common pooled standard deviation. The emission lines of Mg, P, and C showed relatively larger interclass distances (> 0.5). Thus, the three emission line intensities were considered in the k-NN modeling. Also, the classification performances of three 1-variable, three 2-variable, and one 3-variable models were evaluated following the leave-one-out validation process. Among the possible combinations of variables, the usage of both Mg and C emission intensities provided the best model. The highest classification accuracy (86.2%) was achieved by the Mg-C model at k = 13. Our research suggests that the classification methodology combining LIBS and k-NN can be an effective and practical choice for screening particular soybean paste products.

Acknowledgements: This work was supported by the Korea Basic Science Institute (KBSI) National Research Facilities and Equipment Center (NFEC) grant funded by the Korea government (Ministry of Education) (2019R1A6C1010005). The authors thank Mr. Heesu Chae for his assistance in data analysis.

References: [1] R. E. Russo, X. Mao, J. J. Gonzalez, V. Zorba and J. Yoo, "Laser Ablation in Analytical Chemistry", Anal. Chem. 85(13) 6162-6177 (2012).