

Title Variable selection in visible and near-infrared spectra: Application to on-line determination of sugar content in pears

Author Huirong Xu, Bing Qi, Tong Sun, Xiaping Fu and Yibin Ying

Citation Journal of Food Engineering, Volume 109, Issue 1, March 2012, Pages 142-147

Keywords Variable selection; Vis–NIR spectroscopy; Sugar content; On-line determination; Pear

Abstract

Informative variable (or wavelength) selection plays an important role in quantitative analysis by visible and near-infrared (Vis–NIR) spectroscopy. Four different variable selection methods, namely, stepwise multiple linear regression (SMLR), genetic algorithm-partial least squares regression (GA-PLS), interval PLS (iPLS), and successive projection algorithm-multiple linear regression combined with GA (GA-SPA-MLR), were studied to determine the sugar content of pears. The results derived by these techniques were then compared. The calibration model built using GA-SPA-MLR on 18 selected wavelengths (2% of the total number of variables) exhibited higher coefficient of determination (R^2) = 0.880 and root mean square error of prediction (RMSEP) = 0.459°Brix for the validation set. Results show that the accuracy of the quantitative analysis conducted by Vis–NIR spectroscopy can be improved through appropriate wavelength selection. Despite the RMSEP value of GA-SPA-MLR was a slightly higher than that of GA-PLS, considering that this model was simpler and easier to interpret, GA-SPA-MLR can be used for industrial applications.