Early detection of black flesh internal physiological disorder in mango by Vis-NIR spectra and machine learning

[1][2]Jasciane S. Alves; [1]Bruna P. C. Pires; [1]Luana F. Santos; [1]Nilo Ricardo C. M. Júnior; [1]Sandy R. S. Monteiro; [2]Ederson A. Kido; [3]Kerry B. Walsh; [1]Sérgio T. Freitas

[1]Brazilian Agricultural Research Corporation, Embrapa, 56302-970, Petrolina, PE, Brazil ; [2]Federal University of Pernambuco, 1235, Recife, Brazil; [3]Central Queensland University, Rockhampton, Australia.

The objective of this study was to develop a nondestructive method to predict and detect black flesh internal disorder in mango using visible/near-infrared (Vis-NIR) and machine learning. Vis-NIR spectra of healthy and disordered 'Palmer', 'Keitt' and 'Tommy Atkins' mangos were acquired in the wavelength range of 300 to 1100 nm using a portable spectrometer (F-750 Produce Quality Meter, Felix Instruments, WA, USA). Spectra were collected from the equatorial region of both sides of each fruit, with fruit at 24°C (± 1 °C). Spectra were collected from 543 fruit that resulted in 1,086 spectra at harvest (healthy = 350; disordered = 736) and 1,051 after storage at 12 °C (± 1 °C), when the fruit reached the ready-to-eat ripening stage (healthy = 349; disordered = 702). VIS-NIR spectra data were collected at 24 °C (± 1 °C). Spectra data were subjected to the second derivative pre-processing method. Random Forest, Multilayer Perceptron, SMO, LibSVM and J48 algorithms were trialed using the WEKA 3.9 Software. The algorithms performances were evaluated using tenfold cross-validation and model performances were determined by the average accuracy, precision, recall, F-measure, receiver operating characteristics curve (ROC), and Kappa statistics. The models to predict black flesh at harvest showed an average accuracy ranging from 80 to 83.6%, ROC area from 0.80 to 0.91, Kappa from 0.59 to 0.63, precision, recall and f-measure from 0.82 to 0.84 across algorithms triled. The models developed to detect black flesh in ready-to-eat mango showed an average accuracy ranging from 73.2 to 77%, ROC area from 0.63 to 0.86, Kappa from 0.32 to 0.49, precision from 0.74 to 0.77, recall and f-measure from 0.73 to 0.77. The most accurate models to predict at harvest and detect in ready-to-eat mango the incidence of black flesh were developed with the Random Forest algorithm, reaching accuracy of 83.6% and 77%, respectively.

Keywords: Classification models, WEKA, spectroscopy.

Resumo n° 179