Location, year, and tree age impact NIR-based postharvest prediction of dry matter concentration for 58 apple accessions

Soon LiTeh, Jamie L. Coggins, Sarah A. Kostick and Kate M. Evans

Postharvest Biology and Technology, Volume 166, August 2020, 111125

Abstract

Conventional assessments of postharvest traits, like dry matter concentration (DMC) require destructive sampling that can often be limited by fruit availability of apple trees in early years. Alternatively, nondestructive prediction of postharvest traits using NIR spectroscopy has been demonstrated to be rapid and robust in various fruits and vegetables, including apples. With most prediction models built using one to few cultivars, it remains largely unknown if NIR-based prediction can be applied in an apple breeding program with numerous selections that are added, replicated, and evaluated annually. In this study, 2252 fruit from 58 accessions grown at three orchard sites were destructively measured for DMC, and nondestructively measured for NIR spectra. Nondestructive prediction of DMC exhibited high accuracies in most analyses. In characterizing DMC predictive performance of within- versus between-years, both models were highly predictive and comparable, albeit slightly higher for the former. Further analysis of location × year effects revealed that location was a more important factor than year in influencing predictive performance. Finally, in cultivar-specific models, prediction made using fruit from more established trees as a calibration set consistently yielded higher prediction accuracy. Overall, the use of three statistical metrics enhanced our understanding of NIR prediction accuracy, while a control calibration set in all analyses provided a statistical baseline for comparing accuracy outputs. This study provides a framework for understanding the impacts of location, year and tree age on NIR prediction accuracy of DMC in diverse apple breeding accessions. In addition, this study demonstrates the importance of assessing predictive performance using multiple statistical metrics.