介邓飞,谢丽娟,饶秀勤,应义斌.近红外光谱变量筛选提高西瓜糖度预测模型精度[J].农业工程学报,2013,29(12):264-270

近红外光谱变量筛选提高西瓜糖度预测模型精度

Improving accuracy of prediction model for soluble solids content of watermelon by variable selection based on near-infrared spectroscopy

投稿时间： 2012-12-12 最后修改时间： 2013-05-15

中文关键词：近红外光谱,模型,无损检测,西瓜,可溶性固形物,变量筛选,漫透射

英文关键词:near-infrared spectroscopy models nondestructive determination watermelon soluble solids content variable selection diffuse transmittance

基金项目:农业部公益性项目（201003008-5）

作者	单位
介邓飞	1. 浙江大学生物系统工程与食品科学学院，杭州3100582. 农业部设施农业装备与信息化重点实验室，杭州310058
谢丽娟	1. 浙江大学生物系统工程与食品科学学院，杭州3100582. 农业部设施农业装备与信息化重点实验室，杭州310058
饶秀勤	1. 浙江大学生物系统工程与食品科学学院，杭州3100582. 农业部设施农业装备与信息化重点实验室，杭州310058
应义斌	1. 浙江大学生物系统工程与食品科学学院，杭州3100582. 农业部设施农业装备与信息化重点实验室，杭州310058

中文摘要:

水果的内部品质是水果分级、保鲜及存储的一项重要指标，利用近红外光谱技术对西瓜内部品质进行快速无损检测研究有着非常重要的意义。为了研究变量筛选方法对西瓜糖度预测模型精度的影响，该文以麒麟瓜为研究对象，利用近红外漫透射光谱技术对麒麟瓜可溶性固形物含量（SSC）进行检测，采用偏最小二乘回归（PLSR），多元线性回归（MLR）和主成分回归（PCR）建立麒麟瓜可溶性固形物数学模型，并探讨等间隔平均光谱和等间隔抽取光谱变量筛选结合连续投影算法（SPA）对预测模型精度的影响。研究结果表明：光谱经等间隔抽取（间隔5，115个变量）经归一化预处理，结合SPA优选出6个波长建立的PLSR预测模型的相关系数（rpre）为0.828、校正均方根误差（RMSEC）为0.589、预测均方根误差（RMSEP）为0.611。该模型预测效果相对较优，建模时间短，提高了模型的预测能力和预测精度。该研究为西瓜内部品质的在线无损检测提供研究基础。

英文摘要:

Abstract: Fruit internal quality is an important indicator in the fruit grading, preservation and storage stage; rapid non-destructive detection of fruit internal quality can improve the marketable value of watermelon. Near infrared spectroscopy (NIRS) is a powerful technology with the virtue of the convenience and accuracy. This work focused on soluble solids content (SSC) determination of the Qilin watermelon. On basis of the NIRS, we adopted a near-infrared diffuse transmittance technique. We used the home-built measurement system with a fiber optic spectrometer to acquire the spectra. Partial least squares regression (PLSR), multiple linear regression (MLR) and principal component regression (PCR) were used to establish mathematical models. In order to improve the predictive models, firstly, variables were reduced by the interval spectral average method and the interval spectral extraction method, respectively.Further model optimization was carried out by a successive projections algorithm (SPA). The results showed that the PLSR models with 115 variables obtained by an interval of 5 in 574 variables were the best one in first step. Different pretreatments were employed using the 115 variables. By comparison of the results of the PLSR models with different pretreatment, we adopted the normalization pretreatment as the input of SPA algorithm, six optimal wavelengths (702.32, 713.68, 732.58, 770.23, 863.53 and 904.21 nm) were picked out. In a comparison of the predictive results of PLSR, MLR and PCR, the performance of the PLSR model for SSC prediction was better, the correlation coefficient of prediction (rpre) was 0.828, root mean square error of calibration (RMSEC) was 0.589, and root mean square error of prediction (RMSEP) was 0.611. The results revealed that this model took less time for modeling and had reliable predictive ability. This study showed that the home-built measurement system was stable, and provided a theoretical basis for on-line nondestructive detection of the internal quality of watermelon.

查看全文 下载PDF阅读器

关闭