

近红外光谱通用模型在农产品及食品检测中的研究进展

李明^{1*}, 韩东海^{2*}, 鲁丁强¹, 鲁晓翔¹, 柴春祥¹, 刘文³, 孙柯璇¹

1. 天津商业大学生物技术与食品科学学院, 天津 300134
2. 中国农业大学食品科学与营养工程学院, 北京 100083
3. 湘潭大学化工学院, 湖南 湘潭 411105

摘要 我国人口基数大,农产品和食品的需求量多。农产品和食品的质量与安全与人们的日常生活息息相关。如何实现使用无损、快速、环境友好型、高通量的检测方法对农产品和食品的品质进行检测,是当代社会的发展需要。传统的检测分析方法存在着耗时耗力、检测的样品不能再次出售、产生次品漏检的现象等缺点。近红外光谱分析技术作为一种快速无损的检测手段,逐渐被一些学者以及相关行业人员所重视。然而,近红外光谱分析方法大多数只针对单一物料建立数学模型。对于数量庞大且种类众多的农产品和食品而言,如不同地域、不同年份、不同温度、不同加工方法、不同成分组成甚至是不同品种,这种相对传统的近红外分析方法无疑会增加建模的工作量。随着计算机技术、光谱仪硬件、化学计量学以及互联网技术的发展,相关学者已经开始着手于近红外光谱通用模型的研究与开发来解决这一问题,即建立一个近红外通用模型,能够对多种物料的同一种指标或多种指标进行检测。相比于传统的近红外光谱模型,通用模型具有建模成本低、工作量小等优点,特别使近红外光谱技术在农产品和食品领域中应用以及推广方面具有重要的意义。针对近红外光谱通用模型在农产品和食品检测中的研究进行综述,通过比较传统模型建模方法与通用模型建模方法,分别就建立通用模型过程中样品信息的获取、模型的建立以及样品信息的预测三大建模步骤中使用的方法进行总结,并归纳了近红外光谱通用模型在建模步骤中的要点。近红外光谱通用模型在农产品以及食品品质检测方面的研究尚处于发展阶段。也提出对于通用模型开发与研究的一些建议,并就近红外光谱通用模型预测方法在检测领域的发展趋势进行了进一步展望。

关键词 近红外光谱;通用模型;农产品及食品;无损检测

中图分类号: O657.33 **文献标识码:** R **DOI:** 10.3964/j.issn.1000-0593(2022)11-3355-06

引言

近些年来,随着社会对农产品和食品质量与安全的要求日益增高,传统的检测分析方法已经满足不了目前农产品和食品的发展需要。从生产者角度出发,传统的检测方法不仅耗时,而且破坏性的检测方法导致了检测的样品不能再次出售,成本较大,同时由于大部分为抽样检测,不能对样品进行全面检测,会出现次品漏检的现象,对于商品的出售产生负面效应。对于大多数消费者而言,由于不具备传统检测方法的条件,大部分都以个人经验以及卖家信誉购买商品,单从商品外表难以获知其内部品质,导致购买商品的不确定

性。同时,对于行业专业人员来讲,传统的检测方法对于人员操作要求较高,耗时且耗力,不利于方法的普及以及使用^[1]。因此生产者、消费者还是行业专业人员越来越意识到开发快速、准确、经济的检测手段是十分必要的。

近红外光谱(near infrared spectroscopy, NIR)分析具有快速、无污染、成本低以及无损等技术优势,越来越受到社会的关注。NIR的波长范围是780~2500 nm,该光谱区域主要承载的是含氢基团振动倍频与合频特征信息,如O—H, N—H和C—H等含氢基团。由于大部分农产品以及食品由有机分子构成,含有大量的含氢基团,因此近红外光谱分析技术的应用也是从农业领域起步并慢慢普及。1965年, Norris和Hart^[2]首次采用近红外光谱法对谷物的水分等指

收稿日期: 2022-03-27, 修订日期: 2022-06-15

基金项目: 国家自然科学基金青年科学基金项目(31901782), 天津市农业科技成果转化与推广项目(201901090), 湖南省自然科学基金项目(2020JJ5535), 天津市大学生创新创业训练计划项目(202010069082)资助

作者简介: 李明, 1988年生, 天津商业大学生物技术与食品科学学院讲师 e-mail: liming@tjcu.edu.cn

* 通讯作者 e-mail: liming@tjcu.edu.cn; handh@cau.edu.cn

标含量进行了检测。1985 年, Birth^[3] 等开发了一种叫做“Biospect”的近红外光谱仪, 采用其对新鲜洋葱的干物质含量进行了测定。此后, 近红外光谱分析技术在检测领域取得了快速的发展, 能够检测的物料品种以及相关指标也越来越多^[4]。与此同时, 仪器的种类也根据分析对象的种类、运行条件以及特殊需求等得到了进一步的优化, 如离线型、在线型、台式、手持式以及特制型等^[5]。

目前为止, 近红外光谱在农产品和食品品质分析的研究以及应用很多^[6], 但大多数只对于一种物料建立近红外专用数学模型。由于农产品和食品的种类众多, 应对这种情况, 运用传统的近红外分析方法无疑会增加建模的工作量。一般差异较大的物料是难以进行模型通用的。然而近些年来, 一些专家和学者尝试建立近红外通用模型, 即建立一个近红外通用模型, 能够对多种物料的同一种指标或多种指标进行检测。虽说这类通用模型与传统模型相比, 精确度有所下降, 但该类模型不仅具有建模成本低、工作量小等优势, 而且若模型预测结果能满足实际应用要求, 也使近红外光谱技术在农产品和食品领域中的应用具有很大的推广意义。

针对农产品和食品近红外光谱通用模型在建模过程中样品信息的获取、模型的建立以及样品信息的预测三个步骤中的研究进行综述, 对使用的方法进行归纳总结, 提出对于通用模型开发与研究的一些建议, 并就近红外光谱通用模型预测方法在检测领域的发展趋势做进一步展望。

1 通用模型样品信息的获取

近红外光谱分析法的样品信息获取包括采集样品的光谱和测定样品的指标, 这些信息中包含样品的确定信息和不确定信息。确定信息决定了模型的关系信息, 即光谱与指标之间的关系, 不确定信息决定了模型的适用范围, 即关系信息之外的一些影响因素, 这两种信息影响模型的准确性和稳健性^[7]。根据通用模型的目的不同, 获取样品信息的侧重点也不同, 如不同地域、不同批次、不同年份、不同温度、不同加工方法、不同成分组成甚至是不同品种。

王燕妮^[8]等采用近红外光谱分析法, 首先将 8 种饲料原料(菜粕、DDGS、豆粕、麸皮、棉粕、小麦、玉米和杂粕)研磨后并过 0.75 mm 筛, 然后采集粉末状饲料原料的近红外漫反射光谱, 然后测定了这些饲料原料的粗蛋白、粗脂肪和粗纤维 3 种组分的指标含量, 利用偏最小二乘法 (partial least-squares regression, PLS) 算法, 建立了能够对 8 种饲料原料的 3 个指标进行预测的通用检测模型, 通过比较其结果显示, 通用模型的检测结果与单独模型的检测检测结果差异不大, 并且能满足实际生产需要。Piyamartjannok^[9]等利用手持式短波可见近红外光谱仪分别采集了苹果、梨、柿子三种水果的漫透射光谱后, 测得三种水果的可溶性固形物 (soluble solids content, SSC), 同样运用 PLS 算法建立了一种通用模型, 该模型实现对苹果、梨与柿子的 SSC 进行通用预测。通过与单独品种模型的预测标准偏差 (standard error of prediction, SEP) 的范围 (SEP=0.34~0.40°Brix) 对比, 通用模型的 SEP(0.43°Brix) 虽有所下降, 但仍满足检测要求。该

研究表明, 即便品种不同, 但质构相似的薄皮水果是可以建立通用模型来预测相关品质指标的。Fan^[10]等考察了利用近红外光谱仪采集苹果不同部位(赤道处、花萼处、柄附近处)的光谱分别建立 SSC 定量模型, 发现赤道位置采集光谱建立的 SSC 定量模型只能对本位置采集的苹果光谱进行有效的预测 SSC 含量, 其他两个部位建立的 SSC 模型也有相似的结果。为了解决这一问题, 该课题组将三个部位采集的信息合并, 共同建立一个模型, 即“位置通用模型”。结果显示, 位置通用模型对赤道处、花萼处、柄附近处果肉 SSC 的 SEP 分别为 0.386°Brix, 0.486°Brix 和 0.409°Brix, 预测相关系数为 0.977, 0.955 和 0.977。说明该方法能够有效减少由于采集信息的位置不同导致的预测误差。上述研究可知, 通用模型的信息获取可以说是通过一些“校正”与“包容”的思维来攻破模型通用的难题, 即在影响模型准确性和稳健性不大的前提下, 满足检测的要求, 提高模型的适用性。

2 通用模型的建立

当物料的光谱信息以及待测指采集之后, 需要建立两者之间的数学关系, 即建立模型。相对于传统的近红外分析模型, 通用模型的建立在样本集筛选、光谱预处理、变量选择以及建模方法上都需要考虑由于不同地域、年份、温度、成分组成以及品种等因素的复杂性产生的影响。

2.1 通用模型样本集的筛选

样本筛选是将采集的样本信息分为校正集与验证集。当样本集中的样品数足够多且具有代表性、参数分布均匀以及性质范围应包含待测样品等时, 预测模型的准确性和稳健性才能满足一定的检测需求。对于通用模型而言, 更需要保证模型的适用性, 因此通用模型样本集的筛选相对复杂。

Kennard-Stone (KS)^[11] 方法是样本集筛选中最为常用的一种分类方法, 该方法根据光谱特征进行选取。Wang^[12]等首先采集五个品种 (Red Comice, Wujiuxiang, Abate, Cascade, Conference) 洋梨的漫透射光谱以及光谱采集对应位置的 SSC 与果肉硬度值, 每个品种约 94~96 个样本信息, 通过 KS 方法把各个品种样本信息按照三比一的比例分为校正集与验证集, 随后将分离后的各个品种的校正集和验证集合并组成一个新的校正集 (358 个) 和验证集 (118 个)。经统计, SSC 与果肉硬度的新校正集的指标范围分别为 11.3%~18.5%, 1.9~71.2 N, 验证集的指标范围分别为 11.6%~18.0%, 3.5~66.5 N。最后运用多元线性回归 (multiple linear regression, MLR) 算法建立了洋梨的 SSC 通用预测模型, 用 PLS 法建立果肉硬度的通用检测模型。结果显示, SSC 与硬度通用模型的预测决定系数分别为 0.87 与 0.66, SEP 分别为 0.45% 和 8.11 N, 该通用模型具有一定的应用潜力。Set partitioning based on joint x-y distance (SPXY) 法^[13] 是在 KS 方法的基础上演变而来的一种新的样本筛选方法, 该方法考虑了样本的光谱特征以及对应的指标两种因素, 进而对样本信息进行筛选。马卉^[14]等采集了五种药材饮片 (黄芩、栀子、板蓝根、黄柏、胖大海) 的近红外光谱, 并测定了五种药材的水分含量, 同样用 Wang 的方法, 运用 SPXY 距离划

分法将五种药材的样本信息划分为校正集与验证集。经统计显示,该方法划分的校正集含水量范围涵盖验证集,证明划分的校正集具有良好的代表性,有利于建立这五种药材饮片水分含量的通用检测模型。常冬^[15]等通过建立四种不同饲料原料的淀粉含量通用定量检测模型时发现,通过KS方法筛选的通用模型样本集比单个品种的分布范围更加广泛,通用模型的评价参数甚至要好于独立模型。通过以上研究可以推断,无论使用哪种样本筛选方法,应先将独立对象的样本按照一定比例进行校正集和验证集的预筛选,然后将各自筛选出来的校正集与验证集进行整合后,校正集和样本集的分布数据更佳合理,建模效果较好。

2.2 通用模型的光谱预处理方式

采集后的光谱信息除了含有样品自身的有效信息外,还包含其他无效信息,以及对建立模型产生负面效应的信息,如噪声、背景以及杂散光等。在建立通用模型时,由于样本具有地域、时间、大小以及品种的多样性的光谱多样性也会对建立通用模型带来不利。因此如何使用有效的方法对光谱进行处理,是建立通用模型时需要考虑的问题。常见的光谱预处理方法有均值中心化、标准化、归一化、平滑、导数、标准正态变量变换、多元线性校正、傅里叶变换、小波变换、正交信号校正以及净分析信号等^[16]。

Liu^[17]等运用可见/近红外光谱便携仪分别采集了苹果、梨、桃三种水果的光谱信息,积分时间分别为100,90和60 ms,然后测得SSC值,并建立三种水果的SSC通用定量检测模型。研究发现,首先通过标准正态变量变换处理,能够在一定程度上消除水果大小以及内部果肉颗粒造成的散射影响,Savitzky-Golay平滑处理能够消除一些高频噪声,二阶导数处理能够消除由于采集光谱积分时间不同导致的基线漂移现象,经过上述三种光谱预处理之后,三种水果的光谱在840~920 nm波长范围内表现出良好的光谱一致性,在这个波段范围尝试PLS以及MLR两种建模方法进行通用模型的建立。结果显示,PLS通用模型的预测决定系数和SEP分别为0.97和0.45°Brix,MLR通用模型的预测决定系数和SEP分别为0.96和0.46°Brix,模型均达到了一定的预测要求。彭丹等^[18]采集了不同原料(大豆和菜籽)的不同等级油的近红外光谱,共建立了四种油的过氧化值通用检测模型,即不同等级的大豆油、不同等级的菜籽油以及两种同一等级但原料不同的食用油。结果显示,通过二阶导数处理后的光谱建立的一级油的通用模型效果最好,其推测二阶导数处理能够消除基线以及背景干扰,提高有效信息的表达,有利于建立通用模型。Saranwong等^[19]建立近红外光谱通用模型,考察该模型对两种芒果是否被果蝇的虫卵幼虫侵染进行无损检测。结果表明,采集后的光谱进行二阶导数处理,能够有效的表征两种芒果正常储藏和被果蝇虫卵及幼虫侵染后的果实品质的变化,建立的偏最小二乘法判别通用模型能够对两种芒果是否被果蝇虫卵及幼虫侵染进行正确判别,即芒果被蝇虫卵及幼虫侵染后48 h,校正集判别假负率和判别假正率均为1.1%,验证集判别假负率为6.2%,判别假正率为1.2%。Li等^[20]使用Antaris II近红外分光光度计半透反射模式扫描三种不同等级蜂蜜近红外光谱,并通过高效液相色谱法测

定蜂蜜的水分、5-羟甲基糠醛、果糖和葡萄糖含量,建立三种等级蜂蜜的指标定量预测的通用模型。结果显示,在建立水分、5-羟甲基糠醛和果糖三种指标的通用定量模型时,采用一阶导数、Savitzky-Golay平滑与多元散射校正的光谱预处理组合方式,模型效果最好。对于葡萄糖模型,仅使用一阶导数的光谱预处理方式效果最佳。通过上述研究推断,建立近红外光谱模型时,对采集的光谱信息运用预处理手段可以消除噪声以及背景的信息,提高有效信息。同时在建立通用模型时,适当选取某些波段再进行光谱预处理,还能消除由于样品多样性产生的光谱不一致性,利于建立通用模型。

2.3 通用模型的变量选择

光谱分析法建立的模型在对样品进行分析时,由于光谱的数据量越来越大,也会产生一些该技术在农产品及食品品质高通量无损快速检测需求的不利因素,其主要表现为增加计算成本、对仪器硬件要求较高以及工作效率低等方面^[21]。因此,近些年来,大量的研究工作专注于变量选择,筛选后的有效信息用于建立变量少的数学模型。由于采集光谱使用的仪器、采集方式等不同,导致了光谱的表现形式不同,光谱预处理也具有放大信号的作用,即放大了由于品种、产地、加工方式等产生的光谱差异,这些因素对于建立近红外通用模型是不利的。因此相比于传统的模型,通用模型具有更加多样化的样本信息,选择有效、少且全的变量信息是一个相对复杂的研究课题。

Li等^[22]采用便携式可见近红外光谱仪分别采集了三种甜瓜(玛瑙、金红宝、西州蜜)两个部位(赤道部位、花萼部位)处的可见近红外漫透射光谱,光谱范围是550~1100 nm,检测对应部位的SSC值,建立三种甜瓜SSC的通用模型。结果显示,采集的三种甜瓜的光谱,由于品种以及采集部位不同,原始光谱的特征差异较大,为了消除基线以及甜瓜表面纹路的影响,采用的多元散射校正、Savitzky-Golay平滑以及一阶导数的光谱预处理方式,处理后光谱范围在550~749 nm的光谱差异较大,这是由于甜瓜表面颜色不同导致的,而在750~950 nm处的光谱一致性较好。因此在750~950 nm范围内运用竞争性自适应重加权算法(competitive adaptive reweighted sampling, CARS)仅选取24个变量,建立的PLS通用模型能够具备一定的应用潜力(RMSEP=0.73°Brix, $R_p^2=0.83$)。有研究先将三种粮食样品经过粉碎处理,再通过18目筛后进行光谱扫描,通过物理方法消除了由于颗粒大小造成的影响,光谱与水分含量测定后,分别运用蒙特卡洛无信息变量消除算法(Monte carlo uninformative variable elimination, MCUVE)与蒙特卡洛无信息变量消除算法与连续投影算法联用算法MCUVE-SPA(Monte carlo uninformative variable elimination successive projections algorithm, MCUVE-SPA)对光谱的变量进行筛选后,建立具有对粮食水分进行预测的通用定量模型。相比于原始变量(127个),MCUVE与筛选的变量分别为18和7个,经历了变量筛选后的建立的模型仍可满足需求。周婷等^[23]采集了7个不同批次的温州蜜桔的光谱信息(200~1100 nm),建立SSC模型,光谱经过预处理后,选取230~950 nm波段范围建立模型,同时采用无信息消除变量法(uninformative variable

elimination, UVE)、CARS 法以及 UVE-CARS 组合的方法建立 SSC 的 PLS 通用定量模型。结果表明,通过 UVE-CARS 方法选取的变量数目最少,仅 34 个变量,模型的预测效果最佳,即 $R_{cv}=0.704$, $RMSECV=0.949\%$, $R_p=0.732$, $RMSEP=0.873\%$ 。经上述研究可推测,通用模型建立时,变量筛选需先通过物理、化学以及化学计量学方法消除其他因素干扰,选取合适的建模波段范围作为基础,再通过算法选出变量进行建模,效果较好。

2.4 通用模型的建模方法

近红外光谱分析中常用的定量方法有多元线性回归法、主成分回归法、偏最小二乘法等一些线性校正方法,同时还有人工神经网络以及支持向量机等非线性校正方法^[24]。目前来说,PLS 方法在近红外光谱分析中应用最为广泛,逐渐成为一种标准的常用方法,而一些非线性方法的应用也逐渐增多^[25]。

Pocwiardowski 等^[26]运用近红外光谱仪采集了五种蔬菜种子的近红外光谱,使用 PLS 法建立了五种蔬菜种子的水分含量的通用预测模型,结果显示,校正模型的相关系数为 0.996 8, Q 系数^[27]达到 0.890 4 (Q 系数大于 0.7, 表明该模型预测性能较好)。Torres 等^[28]使用微型光谱仪采集了两个柑橘品种 (Oranges 和 Mandarins) 光谱,测得两种柑橘的颜色指数 (L^* , a^* , b^* , C^* , h^* , 比色指数)、硬度、果皮、纵径、横径、重量、厚度以及果汁质量,比较使用线性回归方程 (PLS 法) 与非线性方程 (non-linear regression techniques, LOCAL 算法^[29]) 建立两种柑橘水果模型的结果,其显示使用 LOCAL 算法建立的各个指标通用模型效果均比 PLS 法更佳。Huang 等^[30]使用可见近红外光谱仪采集了三种茶叶、六个部位的光谱信息,并测得四种主要儿茶素和咖啡因含量,分别用不同的建模方法建立了通用定量模型。结果显示,使用 MLR 方法建立的四种种儿茶素的通用定量模型效果最好,模型的相关系数分别为 0.949, 0.893, 0.968 和 0.931, 使用 PLS 法建立的咖啡因的通用定量模型效果最好,模型的相关系数为 0.917。经上述研究可知,无论是线性校正方法还是非线性校正方法,都能建立通用模型。

3 通用模型样品信息的预测

随着近红外技术使用环境的多样化,比如环境的温度、湿度、仪器台间差、待测样本状态等,使得采集的光谱产生差异等因素,给使用建立好的通用模型对未知样品信息的预测带来困难。为了解决这一问题,模型传递的研究也是近期研究的热点^[31],同时该方法的研究也为实现模型的通用开辟了新的道路。Wongsapun 等^[32]首先建立了泰国茉莉香米掺假的偏最小二乘判别模型,随后采用正交投影算法^[33]提取泰国茉莉香米的变量,再通过模型传递对未知掺假样品的近红外光谱进行校正,使其在模型的预测范围之内,相比于之前未进行模型传递的模型,预测值 RMESP 值降低, Q^2 值提高。这种方法不仅实现了对泰国茉莉香米是否掺假的量化检测,同时也降低了建立模型的成本,实现了模型的通用。李小昱等^[34]应用模型传递的方法,实现了不同品种猪肉 pH

值的近红外定量预测模型的通用。该课题组先建立了山黑猪肉 pH 值的近红外光谱 PLS 定量模型,随后使用光谱值校正传递算法消除了山黑猪肉与零号土猪肉之间的光谱差异,进而对零号土猪肉的 pH 值进行预测,模型传递后的预测相关系数为 0.837, 相比没有传递之前提高了 101.69%, 模型传递后的预测均方根误差为 0.085 6, 相比没有传递之前降低了 52.55%, 研究表明校正后模型的预测效果均达到应用水平。

近些年来,互联网、大数据、云服务等技术不断兴起,通过这些新兴技术与近红外光谱技术快速、便捷、无损的特点相结合,为现代农业发展提供了思路。郭志明等^[35]通过研发低成本、低能耗、小型的苹果手持便携仪器采集光谱信息,通过物联网云端数据库实现对苹果品质的数据传输以及远程维护,建立了 17 个产区富士苹果的 SSC、硬度以及维生素 C 含量的通用定量检测模型,并对两个地区的苹果 (陕西洛川和甘肃天水) 的进行预测,其独立预测相关系数 R_i 平均值分别为 0.931, 0.794 和 0.755, 独立预测均方根误差平均值分别为 0.596%, 1.563 kg · cm⁻² 和 0.942 mg · (100 g)⁻¹, 结果显示该系统可以将通用模型与云服务技术相结合,对苹果的品质进行预测。

综上所述,随着新技术的发展,某些通用模型对样品信息的预测已经突破了传统形式上的模型预测方式,这种将无损快速检测技术与其他新兴技术有机地结合起来运用到农产品与食品的研究成果,也为未来通用模型的开发与利用提供了新的探索方向。

4 结 论

农产品和食品的质量与安全与人们的日常生活息息相关,需求量大且种类多给农产品和食品品质检测带来了巨大的工作量。近红外光谱由于其速度快、灵敏度高、采样方式灵活多变以及无损检测等技术优势,已经在农产品以及食品品质检测领域占有一席之地。目前,大多数研究是针对于某种特定情况下对某一种物料进行检测。但在实际生活中,面对不同地域、不同年份、不同温度、不同加工方法、不同成分组成甚至是不同品种的农产品以及食品,传统的近红外光谱“一对一”式模型已经满足不了该项技术在农产品与食品领域的产业化应用。

随着计算机技术、光谱仪硬件、化学计量学以及互联网技术的发展,一部分学者已经开始着手于近红外光谱通用型模型的研究与开发。然而,近红外光谱通用型模型在农产品以及食品品质检测方面的研究尚处于发展阶段,还存在着一些问题,主要体现在以下方面。(1)目前通用型模型研究的农产品以及食品种类不多,至今报道的文献仍然相对较少,主要集中在果蔬与粮油的品质检测,而且该项技术应用还不广泛。(2)目前通用型模型主要应用在定量检测,对于用于定性检测的通用模型研究相对比较较少。(3)建立通用型模型时,建模方法主要以多元线性回归以及偏最小二乘法居多,其他的建模方法相对较少,应用也不多。(4)通用型模型建立的机理研究尚不深入。

采用近红外光谱技术建立通用型模型对农产品以及食品品质进行检测尚处于科学研究阶段, 应用相对较少。结合通用型模型的优点, 改进的预测方法必然是未来检测领域的趋

势。同时, 近红外光谱通用性模型的建立也为其他方式无损检测通用型模型的研究与开发提供新的思路。

References

- [1] WANG Jian-wei, TAO Fei(王建伟, 陶 飞). *Anhui Agricultural Science Bulletin(安徽农学通报)*, 2021, 27(17): 155.
- [2] Norris K H, Hart J R. *Principles & Methods of Measuring Moisture Content in Liquides & Solids*, 1997, 4(1): 19.
- [3] Birth G S, Dull G G, Renfroe W T, et al. *JAmerSocHortSci*, 1985, 110(2): 297.
- [4] TIAN Hua(田 华). *Science and Technology of Food Industry(食品工业科技)*, 2021, 42(18): 41.
- [5] LIN Fang, WU Li-hua, LIU Man-man, et al(林 房, 吴丽华, 刘曼曼, 等). *Liquor Making(酿酒)*, 2020, 47(1): 12.
- [6] LI Chang-bin, XIAO Zhong-shan(李长滨, 肖忠闪). *Modern Animal Husbandry(现代牧业)*, 2020, 4(2): 51.
- [7] YAN Yan-lu, CHEN Bin, ZHU Da-zhou(严衍禄, 陈 斌, 朱大洲). *Near Infrared Spectroscopy-Principles, Technologys and Applications(近红外光谱分析的原理、技术与应用)*. Beijing: China Light Industry Press(北京: 中国轻工业出版社), 2013.
- [8] WANG Yan-ni, SUN Qi, XIAO Zhi-ming, et al(王燕妮, 孙 琦, 肖志明, 等). *Quality and Safety of Agro-Products(农产品质量与安全)*, 2016, (6): 55.
- [9] Kawano, Sumio. *Journal of Near Infrared Spectroscopy*, 2014, 22(5): 367.
- [10] Fan S, Zhang B, Li J, et al. *Biosystems Engineering*, 2016, 143: 9.
- [11] Stone R W K A. *Technometrics*, 1969, 11(1): 137.
- [12] Wang J, Wang J, Chen Z, et al. *Postharvest Biology and Technology*, 2017, 129: 143.
- [13] Galvão R K H, Araujo M C U, José G E, et al. *Talanta*, 2005, 67(4): 736.
- [14] MA Hui, FENG Xue-jing, CHEN Ming, et al(马 卉, 冯雪静, 陈 明, 等). *Chinese Journal of Modern Applied Pharmacy(中国现代应用药学)*, 2021, 38(23): 2932.
- [15] CHANG Dong, ZHANG Sen, WANG Feng-xiang, et al(常 冬, 张 森, 王凤香, 等). *Journal of the Chinese Cereals and Oils Association(中国粮油学报)*, 2017, 32(8): 136.
- [16] CHU Xiao-li(褚小立). *Molecular Spectroscopy Analytical Technology Combined with Chemometrics and its Applications(化学计量学方法与分子光谱分析技术)*. Beijing: Chemical Industry Press (北京: 化学工业出版社), 2011.
- [17] Liu R, Qi S, Han D. *Journal of Near Infrared Spectroscopy*, 2015, 23(5): 301.
- [18] PENG Dan, LI Lin-qing, LIU Ya-li, et al(彭 丹, 李林青, 刘亚丽, 等). *Spectroscopy and Spectral Analysis(光谱学与光谱分析)*, 2020, 40(6): 1828.
- [19] Saranwong S, Thanapase W, Suttiwijitpukdee N, et al. *Journal of Near Infrared Spectroscopy*, 2010, 18(4): 271.
- [20] Li X, Wu Z, Xin F, et al. *Pharmacognosy Magazine*, 2017, 13(49): 193.
- [21] Zhang D, Xu L, Wang Q, et al. *Food Analytical Methods*, 2018, 12: 1.
- [22] Li M, Han D, Liu W. *Biosystems Engineering*, 2019, 188: 31.
- [23] ZHOU Ting, LIU Miao-miao, MAO Fei, et al(周 婷, 刘苗苗, 毛 飞, 等). *Journal of Food Safety & Quality(食品安全质量检测学报)*, 2020, 11(11): 3460.
- [24] ZHANG De-hu(张德虎). *Jiangsu Agricultural Sciences(江苏农业科学)*, 2020, 48(16): 235.
- [25] CHU Xiao-li(褚小立). *Practical Manual of Near Infrared Spectral Analysis Techniques(近红外光谱分析技术实用手册)*. Beijing: China Machine Press(北京: 机械工业出版社), 2016.
- [26] Powiadowski W, Szulc J, Gozdecka G. *Czech Journal of Food Sciences*, 2020, 38(2): 131.
- [27] Elfadl E, Reinbrechta C, Claupein W, et al. *International Journal of Plant Production*, 2010, 4: 259.
- [28] Torres I, Pérez-Marín D, María-José D, et al. *Biosystems Engineering*, 2017, 153: 140.
- [29] Pérez-Marín D, Garrido-Varo A, Guerrero J E. *Talanta*, 2007, 72(1): 28.
- [30] Huang Y, Dong W, Alireza S, et al. *Computers and Electronics in Agriculture*, 2020, 173: 1.
- [31] ZHANG Jin, CAI Wen-sheng, SHAO Xue-guang(张 进, 蔡文生, 邵学广). *Progress in Chemistry(化学进展)*, 2017, 29(8): 902.
- [32] Wongsaijun S, Theanjumol P, Kittiwachana S. *Food Analytical Methods*, 2021, 14(5): 997.
- [33] Boulet J C, Roger J M. *Chemometrics & Intelligent Laboratory Systems*, 2012, 117: 61.
- [34] LI Xiao-yu, ZHONG Xiong-bin, LIU Shan-mei, et al(李小昱, 钟雄斌, 刘善梅, 等). *Transactions of the Chinese Society of Agricultural Machinery(农业机械学报)*, 2014, (9): 216.
- [35] GUO Zhi-ming, WANG Jun-yi, SONG Ye, et al(郭志明, 王郡艺, 宋 烨, 等). *Transactions of the Chinese Society of Agricultural Engineering(农业工程学报)*, 2021, 37(22): 271.

Research Progress of Universal Model of Near-Infrared Spectroscopy in Agricultural Products and Foods Detection

LI Ming^{1*}, HAN Dong-hai^{2*}, LU Ding-qiang¹, LU Xiao-xiang¹, CHAI Chun-xiang¹, LIU Wen³, SUN Ke-xuan¹

1. School of Biotechnology and Food Science, Tianjin University of Commerce, Tianjin 300134, China

2. College of Food Science and Nutritional Engineering, China Agricultural University, Beijing 100083, China

3. School of Chemical Engineering, Xiangtan University, Xiangtan 411105, China

Abstract China has a large population, and the demand for agricultural products and food is great and diverse. Moreover, the quality and safety of agricultural products and foods are closely related to people's daily life. Therefore, it is the development needs of contemporary society to use nondestructive, rapid, environmentally friendly and high-throughput testing methods to detect the quality of agricultural products and foods. The traditional detection and analysis methods have some disadvantages, such as time and labor consumption, the tested samples cannot be sold again after testing and the defective products missing detection. As a rapid and nondestructive detection method, near-infrared spectroscopy (NIR) has been paid more and more attention by some scholars and related industry personnel. However, most NIR analysis methods only build mathematical models for single material. For a large number and variety of agricultural products and foods, such as different regions, different years, different temperatures, different processing methods, different components and even different varieties, this relatively traditional NIR analysis method will undoubtedly increase the workload of preliminary modeling. With the development of computer technology, spectrometer hardware, stoichiometry and internet technology, relevant scholars have begun researching and developing a universal NIR model to solve this problem. That is to establish a near-infrared universal model, which can detect the same index or multiple indexes of various materials. Compared with the traditional NIR model, the universal model has the advantages of low modeling cost and small workload, which makes the application and promotion of NIR spectroscopy technology in agricultural products and foods field of great significance. This paper reviews the research on the universal model of NIR in detecting agricultural products and foods. By comparing the traditional model modeling method with the universal model modeling method, the methods used in the three modeling steps of sample information acquisition, model establishment and sample information prediction in building the universal model are summarized. At the same time, the main points of modeling of NIR universal model in agricultural products and foods detection are summarized. Currently, the research of the NIR universal model in agricultural products and food quality detection is still in the development stage. In this paper, some suggestions on the development and research of the universal model are proposed, and the development trend of the NIR universal model in agricultural products and foods detection has further prospected.

Keywords Near-infrared spectroscopy; Universal model; Agricultural products and foods; Non-destructive detection

(Received Mar. 27, 2022; accepted Jun. 15, 2022)

* Corresponding authors