

# 基于无监督模型转移的苹果跨产地研究

王彬伟

温州大学电气与电子工程学院, 浙江 温州

收稿日期: 2026年4月26日; 录用日期: 2026年5月19日; 发布日期: 2026年5月29日

## 摘要

针对苹果近红外光谱可溶性固形物(SSC)检测中, 不同产地光谱系统性偏移导致模型跨产地泛化性能严重衰减, 传统无监督模型转移方法校正精度不足、易过校正的问题, 本文开展基于滤波学习的无监督模型转移方法(uCTRL)的苹果跨产地无标样品品质检测研究。该方法通过学习线性滤波器投影矩阵, 将源域与目标域光谱映射至共享特征子空间, 在最大化保留SSC预测有效特征的前提下, 最小化产地间光谱分布差异, 实现无监督条件下的跨域模型校正。以北京平谷、新疆阿克苏192个红富士苹果样本为对象开展验证, 结果表明, 经uCTRL校正后, 模型双向跨产地预测RMSEP最大降幅达62.90%, RPD最高提升至2.3432, 满足定量检测要求, 性能显著优于uDOP、di-PLS经典方法, 且无过校正风险, 为苹果品质现场规模化无标样检测提供了可靠技术方案。

## 关键词

苹果, 可溶性固形物, 无监督模型转移, 跨产地检测, uCTRL

# Research on Apple Cross-Producing Area Detection Based on Unsupervised Model Transfer

Binwei Wang

College of Electrical and Electronic Engineering, Wenzhou University, Wenzhou Zhejiang

Received: April 26, 2026; accepted: May 19, 2026; published: May 29, 2026

## Abstract

Aiming at the problems that the systematic offset of spectra from different producing areas leads to serious attenuation of the cross-producing area generalization performance of the model in apple soluble solids content (SSC) detection by near-infrared spectroscopy, and the traditional unsupervised

model transfer methods have insufficient calibration accuracy and are prone to over-calibration, this paper carries out research on label-free cross-producing area quality detection of apples based on Unsupervised Calibration Transfer via Representation Learning (uCTRL). This method maps the spectra of source domain and target domain to a shared feature subspace by learning a linear filter projection matrix, minimizes the spectral distribution difference between producing areas on the premise of maximally retaining the effective features for SSC prediction, and realizes cross-domain model calibration under unsupervised conditions. Verification was carried out with 192 Red Fuji apple samples from Pinggu of Beijing and Aksu of Xinjiang as objects. The results show that after calibration by uCTRL, the maximum reduction of the model's bidirectional cross-producing area prediction RMSEP reaches 62.90%, and the highest RPD is increased to 2.3432, which meets the requirements of quantitative detection. Its performance is significantly better than that of classical methods such as uDOP and di-PLS without over-calibration risk, which provides a reliable technical scheme for on-site large-scale label-free detection of apple quality.

## Keywords

Apple, Soluble Solids Content, Unsupervised Model Transfer, Cross-Producing Area Detection, uCTRL

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

近红外光谱(Near-Infrared Spectrum, NIRS)分析技术凭借快速、无损、无污染的优势,已在农产品品质检测等领域得到广泛应用[1]。可溶性固形物(Soluble Solids Content, SSC)是决定苹果风味品质与商品分级的核心指标, NIRS 技术也已成为苹果 SSC 快速定量检测的主流手段[2]-[5]。但实际应用中,不同产地苹果因种植环境、栽培管理模式差异,光谱会出现系统性分布偏移与基线漂移,基于单一产地建立的校正模型跨域预测性能严重衰减,通用性不足[2] [6]-[8]。传统补充目标域带标签样本重新建模的方式耗时费力、检测成本高,难以适配苹果收购、流通环节的现场规模化检测需求;而模型转移技术可实现校准模型的跨产地复用,有效规避重复建模的资源消耗,目前 di-PLS 等无监督模型转移方法已在农产品、中药等领域的无标样模型适配中得到应用与验证。

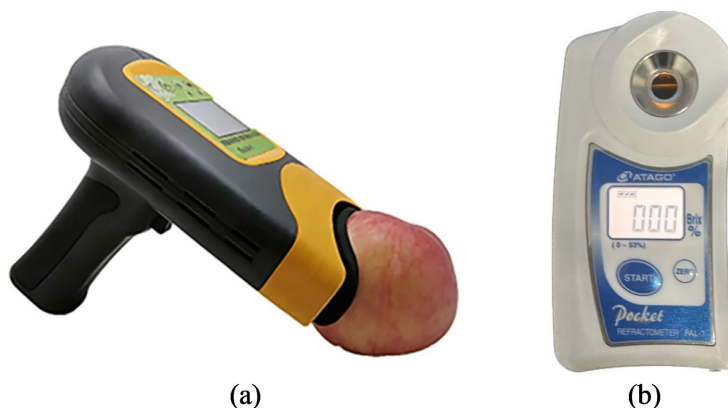
在苹果跨产地无监督模型转移场景中,目标域样本通常仅含光谱信息而无对应 SSC 理化参考值,现有经典无监督方法易受域间大尺度光谱偏移影响,存在校正精度不足、易过校正、鲁棒性差的问题。针对此,本研究基于滤波学习的无监督模型转移方法,通过学习线性滤波器投影矩阵,将源域与目标域光谱映射至共享特征子空间,在最大化保留 SSC 预测有效特征的同时,最小化产地间光谱的系统性分布偏移,实现无监督跨域模型校正。研究以北京平谷、新疆阿克苏产地的 192 个红富士苹果有效样本为对象,采用便携式近红外光谱仪采集样本光谱数据,以偏最小二乘(PLS)为基准校正模型,选取 uDOP、di-PLS 两种经典无监督方法进行对比验证,以预测均方根误差(RMSEP)、残差预测偏差(RPD)为核心指标,系统验证该方法的校正效果与鲁棒性,为苹果品质现场规模化无标样检测提供可靠技术方案。

## 2. 数据采集与方法

### 2.1. 近红外光谱采集仪器

本研究中,苹果样品的近红外光谱数据采用 H100C 型便携式近红外光谱仪完成采集,其可溶性固形

物含量(SSC)采用 PAL-1 型数显式折光仪完成测定, 外观如图 1 所示。



**Figure 1.** (a) Appearance and actual measurement state of H100C portable near-infrared spectrometer; (b) Appearance of PAL-1 digital display refractometer

**图 1.** (a) H100C 型便携式近红外光谱仪外观与实测状态图; (b) PAL-1 型数显式折光仪外观图

子图(a)该仪器的光谱检测范围设定为 650~950 nm, 此波段属于 Vis-NIR, 该区间的光谱信号对苹果果肉中的含氢官能团振动倍频与合频较为敏感, 能够有效反映可溶性固形物、水分等品质指标的成分差异; 同时, 短波近红外的穿透深度适中, 可在避免样品损伤的前提下, 获取苹果表层至浅层果肉的光谱信息, 降低果皮色素对检测结果的干扰。

子图(b)该仪器的核心检测参数为测量范围 0.0~53.0°Brix, 此范围可完全覆盖苹果样品的 SSC 分布区间——常见苹果品种的 SSC 通常在 10.0~15.0°Brix 之间, 即使针对过度成熟(SSC 最高可达 20.0°Brix 以上)或未成熟(SSC 低至 8.0°Brix 以下)的极端样品, 该测量范围仍具备充足的适配性, 可避免因样品 SSC 超出量程导致的检测失效。同时, PAL-1 型 SSC 计具备数显式读数功能, 相较于传统指针式折光仪, 可有效消除人工读数的主观误差。

## 2.2. 实验材料与样本筛选

实验样本的直径严格控制在 85~90 mm 范围内, 单果重量控制在  $200 \pm 10$  g, 该规格为国内市场主流商品苹果的通用标准, 确保实验样本具备良好的市场代表性。两个产地红富士苹果均于 2024 年 10 月采收, 采收后 4℃ 恒温冷藏运输至实验室, 72 h 内完成全部检测, 避免贮藏期品质变化对实验结果的干扰。

两个产地苹果各购 100 个, 为保障实验数据的可靠性与有效性, 购入后首先对样本进行预处理筛选, 剔除标准如下: 表面存在机械损伤、虫眼、霉变的样本, 避免损伤部位对光谱采集与 SSC 检测结果造成干扰; 成熟度不一致、果形畸形的样本, 排除成熟度差异对 SSC 基线的系统性干扰; 重量偏差超过  $\pm 5\%$  的样本, 减少果实个体大小差异带来的光程与散射误差。筛选后, 最终确定北京平谷、新疆阿克苏产地有效样本各 96 个, 合计 192 个有效样本。

为消除苹果球形结构导致的内部 SSC 分布不均问题, 同时保障光谱采集与 SSC 测定的空间对应性, 对每个有效样本沿赤道面均匀划分 4 个采样方位, 每个方位对应 1 个独立采样点, 后续光谱采集与 SSC 测定均以采样点为基本单元完成, 最终取 4 个采样点的光谱均值与 SSC 均值作为该苹果样本的有效数据。对筛选后的北京平谷、新疆阿克苏产地有效苹果样本各 96 个依次进行唯一编号, 结合苹果的球形结构与前文设计的均匀采样方案, 将每个苹果沿中心轴线均匀划分的 4 个方位赤道面四周作为独立数据采集区域, 每个区域对应 1 个采样点如图 2 所示[4] [9]-[12]。



**Figure 2.** Schematic diagram of experimental apple samples and sampling point division

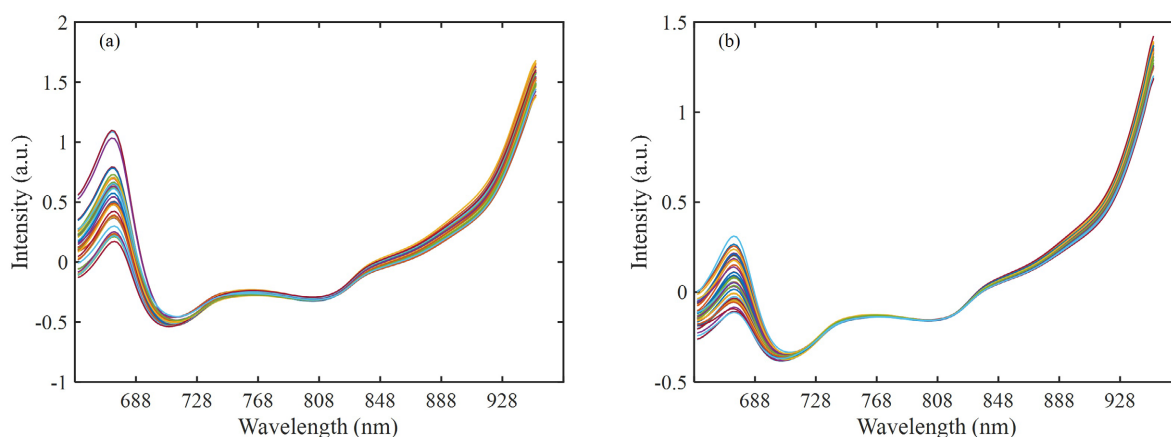
**图 2.** 苹果样本与采样点划分示意图

对筛选后的北京平谷、新疆阿克苏产地有效苹果样本各 96 个依次进行唯一编号，结合苹果的球形结构与前文设计的均匀采样方案，将每个苹果沿中心轴线均匀划分的 4 个方位赤道面四周作为独立数据采集区域，每个区域对应 1 个采样点。具体光谱采集操作流程如下：

1) 仪器开机与校准：将专用保护罩装配于检测探头处，长按开机键启动 H100C 光谱仪，待仪器完成内置基线校准与系统自检后，进入待机工作状态；

2) 设备连接：通过 USB 数据线建立仪器与计算机的物理连接，打开 H100C 配套的光谱管理软件，在软件端完成仪器设备的匹配与连接，确认软件界面显示“设备就绪”；

3) 光谱采集操作：取下探头保护罩，将苹果目标赤道位置紧贴探头检测窗口，确保采样区域完全覆盖检测窗口且无漏光，扣动仪器手柄扳机，卤素灯光源启动并照射苹果样本，同步完成单次光谱数据采集；待软件界面显示光谱曲线绘制完成且光谱计数递增后，完成该采样点的 1 次采集；每个采样点重复采集 3 次，取 3 次采集的光谱平均值作为该采样点的有效光谱数据。两个产地苹果样本的原始近红外光谱曲线如图 3 所示。



**Figure 3.** (a) Apples from Pinggu, Beijing; (b) Apples from Aksu, Xinjiang

**图 3.** (a) 北京平谷苹果；(b) 新疆阿克苏苹果

### 2.3. 无监督模型转移方法

无监督动态正交投影<sup>[13]</sup> (Unsupervised Dynamic Orthogonal Projection, uDOP)模型转移方法是动态正交投影(DOP)结合光谱空间变换(SST)的无监督版本，主要用于没有标签数据的场景。该方法依赖源域数

据和响应值以及目标域数据中单独的光谱测量值,通过对源域和目标域的奇异值分解(Singular Value Decomposition, SVD)得到负载和投影,投影并结合源域负载重新构建目标域中的目标样本。

$$\hat{X}_t = X_t P_t P_s^T P_s^T \quad (1)$$

其中,  $\hat{X}_t$  为目标域虚拟标准光谱矩阵,  $X_t$  为目标域原始光谱矩阵,  $P_s$ 、 $P_t$  分别为源域、目标域光谱经 PCA 分解得到的载荷矩阵。

域不变偏最小二乘[1](Domain-invariant Partial Least Squares, di-PLS)方法是在普通 PLS 的基础上加入了域正则化,在最大化潜在变量投影域目标属性的协方差以保证模型的预测能力的前提下,通过惩罚源域和目标域数据在潜在空间中投影的方差差异,实现两域数据在潜在空间的分布对齐,从而提取对域间系统差异不敏感的域不变特征。

$$\max_w \text{Cov}(Xw, y)^2 - \lambda \cdot \|\text{Cov}(X_S w) - \text{Cov}(X_T w)\|_2^2 \quad (2)$$

其中,  $w$  为潜变量空间权向量,  $\lambda \in [0, 1]$  为正则化参数,  $X_S$ 、 $Y_S$  分别为源域输入光谱矩阵与输出参考值矩阵,  $C_S$ 、 $C_T$  分别为源域和目标域输入数据的协方差矩阵。

基于滤波学习的无监督模型转移方法[14](Unsupervised Calibration Transfer via Representation Learning, uCTRL)是基于滤波器学习框架的无监督模型转移方法,核心面向无目标域标样的跨域光谱检测场景。该方法通过学习线性滤波器投影矩阵,将源域与目标域光谱数据映射至共享特征子空间,在最大化保留苹果可溶性固形物含量(SSC)预测有效特征的前提下,最小化不同产地间光谱的系统性分布偏移与基线差异,实现源域校正模型向目标域的无标样高效校准,解决了传统模型转移方法在无标样场景下校正精度不足、易出现过校正的应用瓶颈。

$$\min_b \|\bar{x}_S - \bar{x}_T * b\|_2^2 + \lambda \|b\|_2^2 \quad (3)$$

其中,  $b$  为域间转换滤波器,  $\bar{x}_S$ 、 $\bar{x}_T$  分别为源域、目标域平均光谱,  $*$  为卷积运算,  $\lambda$  为正则化参数。

$$X_{T,corr} = X_T * b - c \quad (4)$$

其中,  $X_{T,corr}$  为校正后目标域光谱矩阵,  $X_T$  为目标域原始光谱矩阵,  $c$  为背景校正项。

## 2.4. 评价指标

为了全面评价所建立模型的性能,采用预测均方根误差(Root Mean Square Error of Prediction, RMSEP)和残差预测偏差(Residual Predictive Deviation, RPD) [15]作为评价指标。预测均方根误差是通过计算预测值与真实值之间的误差来衡量模型的预测误差,数值越小,表示模型的预测能力越强;其计算公式为:

$$\text{RMSEP} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (5)$$

其中,  $y_i$  是第  $i$  个观测值,  $\hat{y}_i$  是模型预测的第  $i$  个值,  $n$  是观察值的总数。

残差预测偏差是一种用于评估校准模型在预测数据集上性能的统计指标,其定义为总体样本标准差( $\delta$ )与预测均方根误差。RPD 值越大,表示实际值与预测值之间的一致性越好,反映出模型具有更强的泛化能力,能够在未知样本上进行准确的预测。其计算公式为:

$$\text{RPD} = \frac{\delta}{\text{RMSEP}} \quad (6)$$

## 3. 结果分析

为系统验证 uCTRL 方法在苹果 SSC 无标样跨产地检测场景中的校正效果与泛化性能,本研究分别

以北京平谷、新疆阿克苏产地红富士苹果样本为源域构建 PLS 基准校正模型，以另一产地样本为独立目标域测试集，选取 uDOP、di-PLS 两种经典无监督模型转移方法作为对比，以预测均方根误差(RMSEP)、残差预测偏差(RPD)为核心定量评价指标，不同方法的跨产地预测性能如表 1 所示。

**Table 1.** Comparison of data results for different unsupervised model transfer methods

**表 1.** 不同无监督模型转移方法的数据结果对比

数据集	Algorithm	北京平谷		新疆阿克苏	
		RMSEP	RPD	RMSEP	RPD
北京平谷	None	0.2777	2.6014	0.5544	1.3031
北京平谷	uCTRL	0.2777	2.6014	0.3083	2.3432
北京平谷	uDOP	0.2777	2.6014	0.9551	0.7564
北京平谷	di-PLS	0.2777	2.6014	0.3428	2.1075
新疆阿克苏	None	1.7111	0.7134	0.1418	8.6093
新疆阿克苏	uCTRL	0.6348	1.9229	0.1418	8.6093
新疆阿克苏	uDOP	1.3358	0.9139	0.1418	8.6093
新疆阿克苏	di-PLS	0.8798	1.3876	0.1418	8.6093

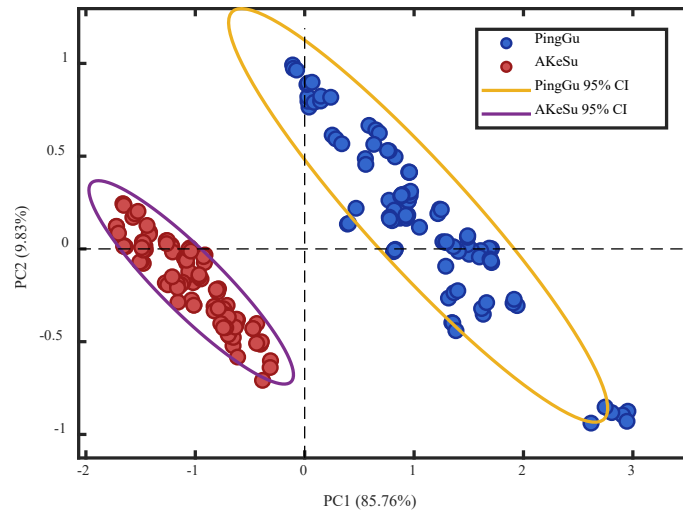
当以北京平谷产地为源域时，无模型转移的基准模型直接应用于新疆阿克苏目标域时性能严重衰减，RMSEP 达 0.5544°Brix，RDP 仅为 1.3031，低于定量检测阈值，无法实现苹果 SSC 的有效跨产地预测；经模型转移方法校正后，uCTRL 与 di-PLS 均实现了模型跨域性能的显著提升，其中 uCTRL 方法校正效果最优，目标域预测 RMSEP 降至 0.3083°Brix，较无模型转移组降低 44.39%，RPD 提升至 2.3432，满足可靠定量检测要求，其预测精度显著优于 di-PLS 方法；而 uDOP 方法未实现有效域间校正，甚至出现预测性能的进一步劣化，其核心原因在于该方法对源域与目标域的光谱分布相似度要求较高，而不同产地苹果的光谱存在显著的系统性偏移，导致无监督场景下虚拟标准样本构建偏差，最终造成模型预测失效。

在反向跨产地验证中，以新疆阿克苏产地为源域、北京平谷产地为目标域时，无模型转移的基准模型完全丧失定量预测能力，RMSEP 高达 1.7111°Brix，RPD 仅为 0.7134；经 uCTRL 方法校正后，模型跨域预测 RMSEP 降至 0.6348°Brix，较基准模型降低 62.90%，RPD 提升至 1.9229，实现了有效定量检测，其性能同样显著优于 uDOP 与 di-PLS 方法。值得注意的是，所有模型转移方法均未改变源域本地样本的预测性能，源域本地预测 RMSEP 与 RPD 均与基准模型保持一致，验证了 uCTRL 方法在实现域间光谱分布对齐的同时，可完整保留源域模型的核心预测能力，无过校正风险。

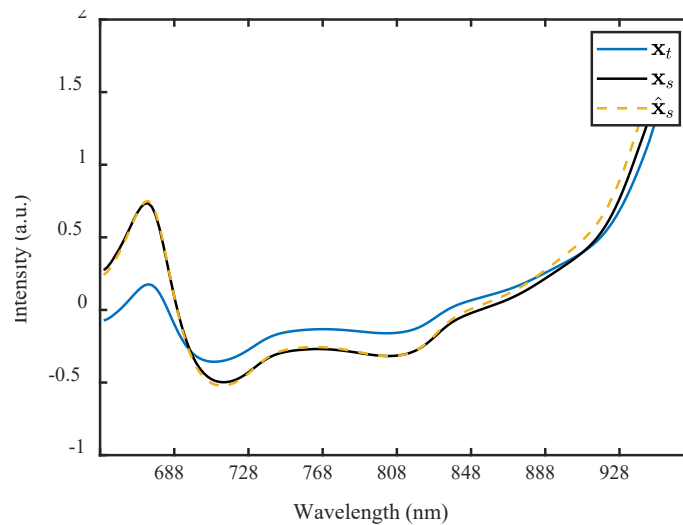
为直观表征北京平谷与新疆阿克苏苹果的光谱分布差异，验证跨产地模型转移的必要性与可行性采用 PCA 对预处理后的全光谱数据进行降维，提取前两个主成分 PC1、PC2 绘制得分空间分布图，结果如图 4 所示。

为验证 uCTRL 无监督模型转移方法的光谱校正与域适配效果，绘制源域原始光谱、目标域原始光谱与经 uCTRL 校正后的源域光谱的对比曲线如图 5 所示。从图中可以看出，校正后与原始源域光谱在全检测波段内的特征峰谷位置、光谱变化趋势与响应强度几乎完全重合，无明显的特征信息丢失与畸变，证明 uCTRL 方法可完整保留源域光谱中与样本属性相关的核心特征。同时，校正后的源域光谱与目标域光谱的光谱轮廓与变化规律高度匹配，二者仅在 688 nm 附近的特征峰处存在小幅强度差异，其余波段均实现了良好的光谱对齐，表明 uCTRL 方法可有效消减源域与目标域之间的光谱分布偏移，具备优异的跨域

模型转移性能。



**Figure 4.** Spatial distribution of principal component scores of apple spectra from different producing areas  
**图 4.** 不同产地苹果光谱的主成分得分空间分布图



**Figure 5.** Comparison of average spectral correction effects  
**图 5.** 平均光谱校正效果

#### 4. 总结

本研究针对苹果可溶性固形物近红外光谱检测中，不同产地光谱系统性偏移导致模型跨域泛化性能严重衰减、传统有监督模型转移方法依赖目标域带标签标样的行业痛点，引入 uCTRL 无监督模型转移方法开展苹果跨产地无标样品质检测研究，以北京平谷、新疆阿克苏产地的 192 个红富士苹果有效样本为研究对象，以 PLS 为基准校正模型，与 uDOP、di-PLS 经典无监督方法开展对比验证，结果表明经 uCTRL 校正后的模型双向跨产地预测 RMSEP 最大降幅达 62.90%，RPD 最高提升至 2.3432，可实现苹果可溶性固形物的可靠定量检测，校正性能显著优于对比方法，且能完整保留源域模型核心预测能力、无过校正风险，该方法大幅降低了近红外光谱技术的应用门槛与检测成本，为苹果品质现场规模化无标样检测提

供了可靠技术支撑, 也为其他农产品的无监督跨域光谱检测提供了方法参考。

## 参考文献

- [1] Poerio, D.V. and Brown, S.D. (2018) Dual-domain Calibration Transfer Using Orthogonal Projection. *Applied Spectroscopy*, **72**, 378-391. <https://doi.org/10.1177/0003702817724164>
- [2] Grabska, J., Beć, K., Ueno, N. and Huck, C. (2023) Analyzing the Quality Parameters of Apples by Spectroscopy from Vis/NIR to NIR Region: A Comprehensive Review. *Foods*, **12**, Article 1946. <https://doi.org/10.3390/foods12101946>
- [3] Asma, U., Morozova, K., Ferrentino, G. and Scampicchio, M. (2023) Apples and Apple By-Products: Antioxidant Properties and Food Applications. *Antioxidants*, **12**, Article 1456. <https://doi.org/10.3390/antiox12071456>
- [4] Li, L., Huang, W., Wang, Z., Liu, S., He, X. and Fan, S. (2022) Calibration Transfer between Developed Portable Vis/NIR Devices for Detection of Soluble Solids Contents in Apple. *Postharvest Biology and Technology*, **183**, Article ID: 111720. <https://doi.org/10.1016/j.postharvbio.2021.111720>
- [5] Malherbe, W., Biggs, R. and Sitas, N. (2024) Comparing Apples and Pears: Linking Capitals and Capacities to Assess the Resilience of Commercial Farming Operations. *Agricultural Systems*, **217**, Article ID: 103934. <https://doi.org/10.1016/j.agsy.2024.103934>
- [6] Francini, A., Romeo, S., Cifelli, M., Gori, D., Domenici, V. and Sebastiani, L. (2017) <sup>1</sup>H NMR and PCA-Based Analysis Revealed Variety Dependent Changes in Phenolic Contents of Apple Fruit after Drying. *Food Chemistry*, **221**, 1206-1213. <https://doi.org/10.1016/j.foodchem.2016.11.038>
- [7] Tang, Y., Gao, S., Zhuang, J., Hou, C., He, Y., Chu, X., *et al.* (2020) Apple Bruise Grading Using Piecewise Nonlinear Curve Fitting for Hyperspectral Imaging Data. *IEEE Access*, **8**, 147494-147506. <https://doi.org/10.1109/access.2020.3015808>
- [8] Li, X. and Zhu, W. (2011) Apple Grading Method Based on Features Fusion of Size, Shape and Color. *Procedia Engineering*, **15**, 2885-2891. <https://doi.org/10.1016/j.proeng.2011.08.543>
- [9] Diels, E., van Dael, M., Keresztes, J., Vanmaercke, S., Verboven, P., Nicolai, B., *et al.* (2017) Assessment of Bruise Volumes in Apples Using X-Ray Computed Tomography. *Postharvest Biology and Technology*, **128**, 24-32. <https://doi.org/10.1016/j.postharvbio.2017.01.013>
- [10] Khatiwada, B.P., Subedi, P.P., Hayes, C., Carlos, L.C.C. and Walsh, K.B. (2016) Assessment of Internal Flesh Browning in Intact Apple Using Visible-Short Wave near Infrared Spectroscopy. *Postharvest Biology and Technology*, **120**, 103-111. <https://doi.org/10.1016/j.postharvbio.2016.06.001>
- [11] Tao, Y. and Zhou, J. (2017) Automatic Apple Recognition Based on the Fusion of Color and 3D Feature for Robotic Fruit Picking. *Computers and Electronics in Agriculture*, **142**, 388-396. <https://doi.org/10.1016/j.compag.2017.09.019>
- [12] Jiang, X., Zhu, M., Yao, J., Zhang, Y. and Liu, Y. (2022) Calibration of near Infrared Spectroscopy of Apples with Different Fruit Sizes to Improve Soluble Solids Content Model Performance. *Foods*, **11**, Article 1923. <https://doi.org/10.3390/foods11131923>
- [13] Zeaiter, M., Roger, J.M. and Bellon-Maurel, V. (2006) Dynamic Orthogonal Projection. A New Method to Maintain the On-Line Robustness of Multivariate Calibrations. Application to Nir-Based Monitoring of Wine Fermentations. *Chemometrics and Intelligent Laboratory Systems*, **80**, 227-235. <https://doi.org/10.1016/j.chemolab.2005.06.011>
- [14] Xie, Z., Chen, X., Roger, J., Ali, S., Huang, G. and Shi, W. (2024) Calibration Transfer via Filter Learning. *Analytica Chimica Acta*, **1298**, Article ID: 342404. <https://doi.org/10.1016/j.aca.2024.342404>
- [15] Chang, C., Laird, D.A., Mausbach, M.J. and Hurburgh, C.R. (2001) Near-Infrared Reflectance Spectroscopy-Principal Components Regression Analyses of Soil Properties. *Soil Science Society of America Journal*, **65**, 480-490. <https://doi.org/10.2136/sssaj2001.652480x>