Postharvest Biology and Technology 170 (2020) 111326

Contents lists available at ScienceDirect

Postharvest Biology and Technology

journal homepage: www.elsevier.com/locate/postharvbio

Two standard-free approaches to correct for external influences on near-infrared spectra to make models widely applicable

Puneet Mishra [a,*], Jean Michel Roger [b,c], Douglas N. Rutledge [d,e], Ernst Woltering [a,f]

[a] Wageningen Food and Biobased Research, Bornse Weilanden 9, P.O. Box 17, 6700AA, Wageningen, the Netherlands
[b] ITAP, INRAE, Institut Agro, University Montpellier, Montpellier, France
[c] ChemHouse Research Group, Montpellier, France
[d] Université Paris-Saclay, INRAE, AgroParisTech, UMR SayFood, 75005, Paris, France
[e] National Wine and Grape Industry Centre, Charles Sturt University, Wagga Wagga, Australia
[f] Horticulture and Product Physiology Group, Wageningen University, Droevendaalsesteeg 1, P.O. Box 630, 6700AP, Wageningen, the Netherlands

ARTICLE INFO

Keywords:
Model adaption
Multivariate
Scalable spectroscopy
Chemometrics

ABSTRACT

In near-infrared (NIR) spectroscopy of fresh fruit often the external influences due to differences in physical, chemical and environmental conditions lead to model failure. Correction methods are required where standard samples are measured covering all different conditions and then remodeling is performed. However, in the real-world, it is often difficult to measure standard samples. To deal with this, two different approaches to correct for external influences without standard sample measurements i.e., dynamic orthogonalization projection (DOP) and domain adaption (DA), are presented, and for the first time are applied to NIR spectroscopy of fresh fruit. Four different case studies, chosen based on their importance and their frequency of occurrences in the NIR spectroscopy domain, were used for the demonstration. The first case was an adaption to maintain the predictive performance of a model when used on a spectra from a second similar instrument. The second case was the correction of the temperature effects due to sensor heating. The third and fourth cases were about maintaining the model performance for multi-season fruit quality prediction models for mangos and for apples. In all of the cases, the aim was to solve the challenges without resorting to new measurement of standards. The results showed that for all the cases, both DOP and DA improved model performances. Up to 31% increase in R^2_p, and 98% and 66% reduction in prediction bias and root mean squared error (RMSE) of prediction were noted, respectively. The main benefit of the DOP and DA techniques in NIR spectroscopy is the limited need for standard measurements, providing general-purpose tools to complement the NIR spectroscopy and make the models scalable, transferable, and reusable.

1. Introduction

NIR spectroscopy is a rapid, non-destructive sensing technique that is widely used for predicting the physicochemical properties in fresh fruit (Lin and Ying, 2009; Subedi and Walsh, 2009; Wang et al., 2015; Saeys et al., 2019; Walsh et al., 2020;). Applications of NIR spectroscopy can range from agriculture to highly controlled pharmaceutical domains as a process analytical tool (PAT) (Pasquini, 2018). Unlike other analytical techniques such as nuclear magnetic resonance and chromatography which can directly provide the concentration of the analytes, NIR spectroscopy usually requires a multivariate calibration step before it can be used for any application (Saeys et al., 2019; Walsh et al., 2020). The calibration of NIR data is often easy and can be done with latent variable extraction techniques such as partial least-squares regression (PLSR) (Wold et al., 2001). In some more complex cases, non-linear techniques can be used. As well, different pre-processing combinations are always explored in order to have an optimal final model (Rinnan et al., 2009; Engel et al., 2013; Roger et al., 2020).

A major complaint from NIR spectroscopy users is that the developed model often has limitations. Models often need to be updated over time as the quality of the sensor deteriorates, a need for model recalibration when the light source is changed, the perturbing effects of temperature on samples and sensors limiting the usefulness of the model, the model developed on one sensor not being applicable with another sensor, the model not performing well when used on a new batch of materials, bad model performance when tested on multi-season experiments related to

* Corresponding author.
E-mail address: puneet.mishra@wur.nl (P. Mishra).

https://doi.org/10.1016/j.postharvbio.2020.111326
Received 2 June 2020; Received in revised form 6 August 2020; Accepted 6 August 2020
Available online 9 September 2020
0925-5214/© 2020 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

agri-food materials, and many more (Nicolai et al., 2007; Saeys et al., 2019). Many of these problems have been dealt with, mainly by model updating and by calibration transfer between instruments. Model updating is very common practice in the pharmaceutical domain where NIRS is often implemented as a PAT tool. A common practice for model updating is by the measurement of some new extra samples and recalibration of the previously developed model by incorporating the new samples. Calibration transfer is also a widely explored domain in NIR spectroscopy where a model developed on one instrument needs to be transferred to a different instrument (Fearn, 2001; Feudale et al., 2002; Nicolai et al., 2007; Saeys et al., 2019). The aim of calibration transfer is to save the cost and time by reusing the former calibration data. A common approach to do that is by the measurement of standard samples on both instruments and then estimating a transfer function which can be used to adapt the model to the new instrument (Workman, 2017). Several methods for calibration transfer are available such as direct standardization (DS), piece-wise direct standardization (PDS), transfer by orthogonal projection (TOP) (Fearn, 2001; Feudale et al., 2002; Workman, 2017). In summary, several approaches are available in the NIR spectroscopy domain, but they mostly require the measurement of standard samples or new samples, which is a time consuming, costly and appears unattractive to users which may limit the implementation of NIR spectroscopy. The need for standard-free procedures led to the development of the dynamic orthogonal projection (DOP) method which instead of standard samples requires some extra measurements on the new experimental setup (Zeaiter et al., 2006). Standard sample measurements are the measurements under different conditions using the same samples to model the difference between conditions. The extra measurements are the measurements that are performed only in the new condition. DOP removes the detrimental subspace from the calibration dataset by using orthogonal projections within the framework of external parameter orthogonalization (EPO) (Roger et al., 2003). Detrimental subspace can be understood as the subspacesthat capture the differences between two instruments in case of calibration transfer, differences between season in the case of multi-season fruit experiments, and the differences due to the effect of temperature on the NIR spectra for multi-temperature experiments. After removal of the detrimental subspace, the model can be recalibrated and used with improved predictive performance (Zeaiter et al., 2006; Roger, 2016; Roger et al., 2018).

The challenge of sensor calibration and model updating is not limited to the NIR spectroscopy domain as it is frequently encountered in other fields such as computer vision (Kouw and Loog, 2019). Often, object classification and detection algorithms trained on the images acquired with one camera are to be used on images from another. Furthermore, the models need to be adapted to work in different scenes and lighting conditions. To deal with such problems, a family of methods called Domain Adaption (DA) or Transfer Learning has recently emerged (Kouw and Loog, 2018). DA is useful when the distributions of the measurements, coming from different instruments or samples collected under different condition, are not the same. In the NIR spectroscopy domain, this is analogue to the calibration transfer problems when the model made on one instrument need to be used with another instrument and/or the measurements are made under different temperature conditions. One of the commonly used DA techniques is transfer components analysis (TCA) which assumes that if two domains are related to each other then there may exist several common components (or latent variables (LVs)) underlying them but masked due to the differences in the domains (different instruments, temperatures and season conditions) (Pan et al., 2011). TCA aims to learn a set of common components underlying both domains such that the differences in distributions of data in the different domains when projected onto this subspace, can be dramatically reduced (Pan et al., 2011). Then, standard machine learning methods can be used in this subspace to train models that work across measurement conditions.

The aim of the present work is to demonstrate the use of DOP and DA with TCA for dealing with different challenges regarding the implementation of NIRS spectroscopy. Both methods (DOP and TCA) have been used in two very different domains i.e. chemometrics and computer vision. However, they have never been used on fresh fruit analysis for correction of external influences related to different instruments, temperatures and for season corrections. Both DOP and TCA aim to achieve improved model performance, but by different means. The DOP approach aims to remove the effect of the influence factor to make the data as independent as possible of that perturbations. DOP does this by modelling the subspace of the influence factors and then removing it from the data by orthogonal projections (Zeaiter et al., 2006). TCA, on the other hand, aims to retain the common information which is present in the datasets irrespective of the influence factors (Pan et al., 2011).

2. Materials and Methods

2.1. Datasets

In all of the cases, NIR spectroscopy was used for the prediction of dry matter (DM) as reference property in individual fruit. All the spectra were 2nd derivative pre-processed to reveal the underlying peaks related to moisture. Outlying samples were removed based on the use of inner relation plots from PLS decomposition. A summary of datasets used is provided in Table 1. More details about the datasets can be found in Sun et al. (2020a,2020b), Anderson et al. (2020), and Teh et al. (2020).

Olive fruit instrument transfer. The olive fruit dataset consists of olive fruit measured with two identical portable Felix spectrometers. The original dataset consisted of NIR spectroscopy measurement on two sides of each olive fruit and a single DM value. The two spectra from each fruit were averaged in this analysis. The dataset was used to demonstrate the model transfer from one instrument to the other.

Mango fruit temperature correction. The mango fruit temperature dataset consisted of mango fruit measured at two different sensor temperature levels of the Felix NIR spectroscopy instrument. The original dataset has three temperature levels, low (~15 °C), medium (~25 °C) and high (~30 °C). However, for this demonstration, only the low and medium temperature levels were used.

Mango fruit season correction. Mango fruit season data consisted of multi-year mango fruit NIR spectroscopy data (2015–2019) measured with the Felix NIR spectroscopy instrument. For this demonstration, the data from 2016 and 2019 were used. As well, due to the large number of samples, 1/3 of the measurements were used. The samples were randomly selected.

Apple fruit season correction. The apple fruit multi-season dataset consisted of NIR spectroscopy measurements on individual apple acquired using the Felix NIR spectroscopy instrument. The two seasons were from 2015 and 2016.

2.2. Data analysis

Two approaches, i.e., dynamic orthogonal projection (DOP) and transfer component analysis (TCA) were used to correct the effects of perturbations on the NIR spectroscopy data. All the data analyses were performed using MATLAB 2017b, Natick, MA, USA.

2.2.1. Dynamic orthogonal projections

DOP is a model maintenance method developed to deal with the physical, chemical and environmental affects in the spectroscopic modelling (Zeaiter et al., 2006). The approach is based on the correction of the calibration dataset based on the new reference measurements performed in different physical, chemical and environmental conditions. The correction is performed using orthogonal projections based on the subspace defined by the difference between the calibration spectra and the spectra under new conditions. Let R be a set of samples measured in the new conditions. Let **Yr** be the reference values and **Xr**

Table 1
Summary of datasets used.

Data set	Spectral range (nm)	Training (samples × wavelengths)	Optimizing DOP (samples × wavelengths)	Testing DOP/TCA (samples × wavelengths)	Reference measurement	Literature
Olive fruit instrument transfer	705–1115	186 × 135	31 × 135	65 × 135	DM (%)	(Sun et al., 2020a)
Mango fruit temperature correction	705–1128	10003 × 142	196 × 142	800 × 142	DM (%)	(Anderson et al., 2020; Sun et al., 2020a,2020b)
Mango fruit season correction	705–1115	455 × 135	133 × 135	350 × 135	DM (%)	(Anderson et al., 2020; Sun et al., 2020a,2020b)
Apple fruit season correction	729–975	1219 × 83	207 × 83	800 × 83	DM (%)	(Teh et al., 2020)

the measured spectra of this samples. The DOP method starts by estimating virtual standards, i.e. the spectra **Xr*** that should have been measured in correspondence with **Yr**, if the calibration conditions were maintained. This is done by means of linear combinations of the original calibration data matrix, using kernels centered on **Yr** values. Once the virtual standards are prepared then the difference spectra between **Xr** and **Xr*** are calculated. The orthogonal basis for the difference spectra is estimated by principal components analysis, and finally, the calibration spectra are projected orthogonally to that basis. Models calibrated on these data become insensitive to the differences (physical, chemical and environmental conditions).

To explain the case for a two-batch scenario related to NIR spectroscopy of fruit, let us consider that the spectral data acquired for Batch1 is given by $\mathbf{X_1}$ and the corresponding reference values by $\mathbf{Y_1}$. The spectral data acquired as Batch2 (under different physical, chemical and environmental conditions) be given by $\mathbf{X_2}$, without reference measurements. Consider a small set of measurements performed on batch 2 for which spectra can be given as $\mathbf{X_r}$ and the corresponding reference measurements as $\mathbf{Y_r}$.

The step of DOP includes (Zeaiter et al., 2006):

1 Virtual spectral standards are created by estimating $\mathbf{X_r}$ as a linear combination of $\mathbf{X_1}$. This linear combination is provided by the kernel functions centered on the elements of $\mathbf{Y_r}$ and applied to Y_1.

$$\hat{X_r} = AX_1 \quad (1)$$

Where $\mathbf{A = F(Y_1)}$, where **F** is the kernel function centered on $\mathbf{Y_r}$

2 The difference spectra **D** are computed as the difference between $\mathbf{\hat{X}_r}$ and $\mathbf{X_r}$

$$\mathbf{D = \hat{X}_r - X_r} \quad (2)$$

3 An orthonormal basis **P** of the space spanned by **D** is estimated by a principal component analysis (PCA)

$$\mathbf{D = TP^T + E} \quad (3)$$

where **T** and **P** are the first *k* scores and loadings of the PCA, and **E** is the residuals

4 The spectra from Batch 1 ($\mathbf{X_1}$) are corrected by orthogonal projection as

$$\mathbf{X^*_1 = X_1\,(I - PP^T)} \quad (4)$$

where I is the identity matrix and PP^T is the product of P by P^T. Doing such an operation removes the differences between two batches and the new calibration can be done with $\mathbf{X^*_1}$ and $\mathbf{Y_1}$ and then the model can be applied directly to the $\mathbf{X_2}$. The number of dimensions removed (*k*) must be tuned, e.g. by examining how the proximity between the two batches improves with increasing *k*. The external parameter orthogonalization (EPO) components were optimized using the validation procedure by choosing the number of components corresponding to lowest error for a tuning set.

2.2.2. Domain adaption with Transfer Component Analysis

TCA is a common DA technique but new to the chemometrics domain. TCA applies when there are differences between data sets which are expected to have identical variations. In (Pan et al., 2011) the term distribution is used to characterize these variations. TCA assumes that there exist differences in the data distributions for data from different domains. In the case of NIRS, it can be assumed that the variance – covariance matrix between two instruments, two temperature conditions and two seasons is different. Some of the latent space captures the intrinsic structure underlying the data while the rest just carries non useful information (Pan et al., 2011). The latent space used by a model, e.g. the one spanned by the loadings of a PLS, may correspond to the differences between domains, or to the similarities between domains. TCA finds a latent feature space that minimizes the distributional difference of the domains by the calculation of transfer components across domains in a Reproducing Kernel Hilbert Space (RKHS) using Maximum Mean Discrepancy (MMD) (Pan et al., 2011). MMD is a non-parametric distance measure for data distributions in RKHS. RKHS is the Hilbert space of functions in which point evaluation is a continuous linear functional.

As before, the aim is to predict the $\mathbf{Y_2}$ from $\mathbf{X_2}$ with the calibration developed on $\mathbf{X_1}$ and $\mathbf{Y_1}$. The TCA presented in (Pan et al., 2011) assumes that the marginal distribution of $\mathbf{X_1}$ and $\mathbf{X_2}$ are not equal ($P(\mathbf{X_1}) \approx P(\mathbf{X_2})$), but that there exists however a transformation φ such that $P(\varphi(\mathbf{X_1})) \approx P(\varphi(\mathbf{X_2}))$ and $P(\mathbf{Y_1}/\varphi(\mathbf{X_1})) \approx P(\mathbf{Y_2}/\varphi(\mathbf{X_2}))$, where P denotes the marginal distribution. Once the transformation is done with the φ, models on $\varphi(\mathbf{X_1})$ and $\mathbf{Y_1}$ can be developed and applied to $\varphi(\mathbf{X_2})$. However, the key challenge is to find the φ as there is no $\mathbf{Y_2}$ available, so the φ cannot be learned by minimizing the distance between $P(\mathbf{Y_1}/\varphi(\mathbf{X_1})) \approx P(\mathbf{Y_2}/\varphi(\mathbf{X_2}))$. The TCA as presented in (Pan et al., 2011) proposes to learn φ such that the distance between the marginal distributions $P(\varphi(\mathbf{X_1}))$ and $P(\varphi(\mathbf{X_2}))$ is small, and, the $\varphi(\mathbf{X_1})$ and $\varphi(\mathbf{X_2})$ preserve important properties of $\mathbf{X_1}$ and $\mathbf{X_2}$.

The steps of TCA include:

1 Construction of kernel matrix **K** from $\mathbf{X_1}$ and $\mathbf{X_2}$ as defined by the Gram matrices for Batch1($\mathbf{K}_{X1,1}$), Batch2 ($\mathbf{K}_{X2,2}$) and cross-batch ($\mathbf{K}_{X1,X2}$)

$$\mathbf{K} = \begin{bmatrix} K_{X_{1,1}} & K_{X_{1,2}} \\ K_{X_{2,1}} & K_{X_{2,2}} \end{bmatrix}$$

where the kernel matrix is defined on all the data by minimizing the distance (MMD) between the Batch1 and Batch2 while maximizing the embedded data variance. With the kernel trick, the MMD distance can be estimated as $tr(KL)$, where $\boldsymbol{K} = [\varphi(x_i)^T \varphi(x_j)]$, and $L_{ij} = {}^1/_{n_1{}^2}$ if x_i, $x_j \in \boldsymbol{X}_1$, else $L_{ij} = {}^1/_{n_2{}^2}$ if $x_i, x_j \in \boldsymbol{X}_2$, otherwise, $L_{ij} = -({}^1/_{n_1 n_2})$

2 The objective function thus becomes:

$\max_{K \geq 0} \{ trace(KL) - \lambda\, trace(K) \}$ subject to contraints on $\boldsymbol{K}$

where the first term in the objective minimizes the distance between distributions, while the second term maximizes the variance in the feature space, and $\lambda \geq 0$ is a trade-off parameter

3 After dealing with the problem as a trace optimization problem as presented in (Pan et al., 2011), the final step leads to an eigen decomposition problem and selection of m leading eigenvectors as

$$W = \text{eig}\left((KLK + \mu I)^{-1} KHK\right)$$

where, I is the identity matrix, $\mu > 0$ is a trade-off parameter, **H** is the centering matrix $\boldsymbol{H} = I_{n_1+n_2} - ({}^{1}/_{n_1+n_2})11^{T}$, where 1 is the column vector of all 1's and $I_{n_1+n_2}$ is the identity matrix and **W** are the leading eigenvectors which can be used for the data transformation

4 Finally, both the Batch1 and Batch2 data can be transformed as

$X_1^* = X_1 W$ and $X_2^* = X_2 W$

5 The new calibration on X_1^* and Y_1 can be developed and directly applied on X_2^*

2.2.3. Partial least-squares regression

To compare the improvements using DOP and TCA, PLSR was used. PLSR is a common chemometric technique used for the calibration of NIR spectroscopy data of fresh fruit (Saeys et al., 2019; Walsh et al., 2020). PLSR works by extracting the LVs which explain the variance in both the explanatory and response variables (Wold et al., 2001). The LVs are used for estimating the scores for the original data matrix and these scores are used for performing the multi-linear regression. By doing so, PLSR avoids the collinearity problem in the regression and leads to optimal models. In the present work, the number of LVs (selected corresponding to error stabilization point) for the PLSR was optimized with a 10-fold venetian-blind cross-validation. The model performance was quantified as root mean squared error of prediction (RMSEP), prediction bias and prediction R^2.

3. Results and discussion

3.1. Data distributions

Fig. 1 presents the distribution of DM (%) for the four data sets presented in Table 1. For all four cases the distribution can be assumed to be normally distributed. Such a distribution is commonly encountered when estimating properties related to biological samples such as fruit.

3.2. Olive fruit dry matter copy instrument predictions

Fig. 2 presents the results of PLSR modelling on the non-transformed data (Fig. 2A), the DOP- (Fig. 2B) and the TCA- transformed data (Fig. 2C). The samples from instrument 1 are in green and from

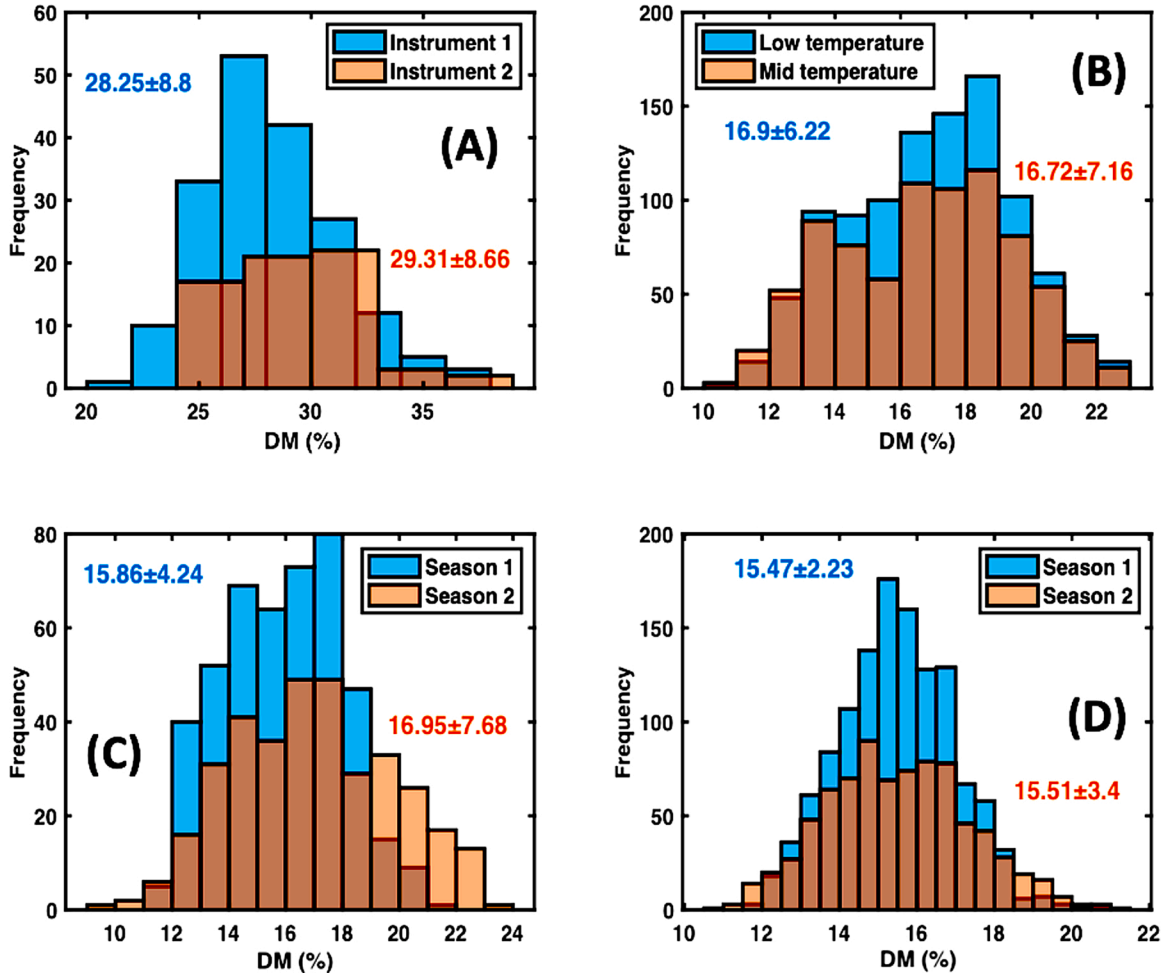

Fig. 1. Distributions of dry matter (DM) in % for calibration and test set. (A) instrument transfer (olive fruit), (B) different temperature (mango fruit), (C) different season correction (mango fruit) and (D) different season (apple fruit). The numerical values represent the mean ± 2 standard deviation.

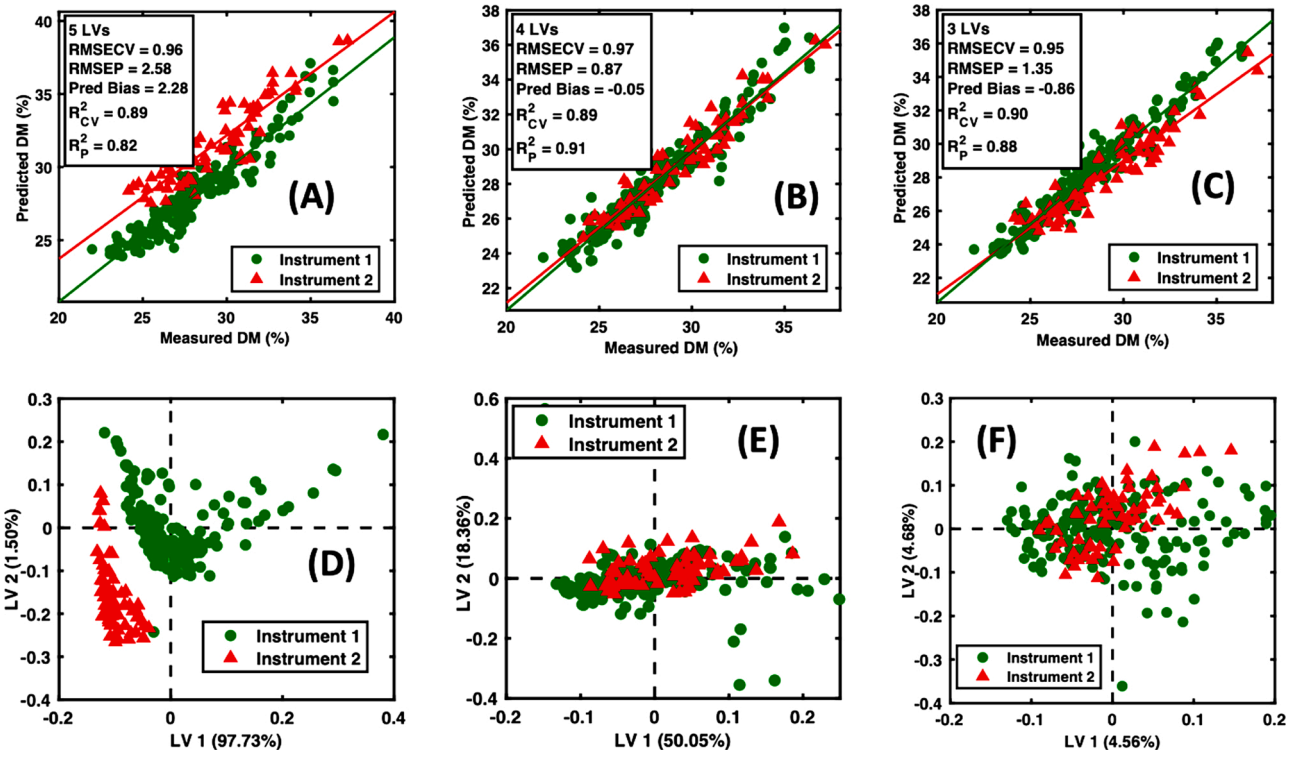

Fig. 2. Model made on data from instrument 1 and tested data from instrument 2 to predict DM (%) in olive fruit. Data from instrument 1 is presented in green circles and instrument 2 in red triangles. (A) Partial least-squares regression (PLSR) calibrated on instrument 1 tested on instrument 2, (B) PLSR after dynamic orthogonal projection (DOP) calibrated on instrument 1 tested on instrument 2, (C) PLSR after transfer component analysis (TCA) calibrated on instrument 1 tested on instrument 2, (D) scores distribution for instrument 1 and instrument 2 measurements after PLSR, (E) scores distribution for instrument 1 and instrument 2 measurements from PLSR after DOP, and (F) scores distribution for instrument 1 and instrument 2 measurements from PLSR after TCA. LV in the score plots stands for latent variables. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

instrument 2 in red. In all the cases, the model was calibrated on instrument 1 and tested on instrument 2. It can be noted that both DOP and TCA improved the model performance, i.e. high R^2_p and low RMSEP compared to no data transformation. The scores on the first two latent variables (LVs) are shown in Fig. 2D-F. The scores plots show that the raw data from multiple instruments have differences in distributions (Fig. 2D), can be seen for the red (instrument 2) and green (instrument 1) clusters, which are due to differences between the instruments. However, these differences were removed/reduced by the transformations using DOP and TCA. The R^2_p was highest for DOP with a total 11% increase compared to no transformation. DOP also reduced the prediction bias by 97% compared to no transformation. The RMSEP with the use of DOP was reduced by 66%. The number of LVs were lowest for TCA (3 LVs), followed by DOP (4 LVs), then by no transformation (5 LVs). A summary of all the models is shown in Table 2.

3.3. Temperature correction for improved prediction

Fig. 3 presents the results of PLSR modelling on the non-transformed data (Fig. 3A), the DOP- (Fig. 3B) and the TCA- transformed data (Fig. 3C) for correcting he temperature effect in the NIRS. The samples from low temperature are in green and from medium temperature are in red. In all the cases, the model was calibrated in low temperature and tested on medium temperature. It can be noted that both DOP and TCA improved the model performance, i.e. high R^2_p and low RMSEP compared to no data transformation. The scores on the first two latent variables (LVs) are shown in Fig. 3D–F. The R^2_p was the highest for TCA with a total 12% increase compared to no transformation. TCA also reduced the prediction bias by 85% compared to no transformation. The RMSEP with the use of TCA was reduced by 29%. The number of LVs were the lowest for TCA (3 LVs), followed by DOP (5 LVs), and then by no transformation (8 LVs).

3.4. Correction of seasonal effect in mango fruit

Fig. 4 presents the results of PLSR modelling on the non-transformed data (Fig. 4A), the DOP- (Fig. 4B) and the TCA- transformed data (Fig. 4C). The samples from season 1 are in green and from season 2 are

Table 2
Summary of results from partial least-squares regression (PLSR), dynamic orthogonal projection (DOP) and transfer component analysis (TCA) for dry matter (DM) prediction. The best improvements are highlighted in red.

Dataset	PLSR				DOP				TCA			
	LVs	R^2_p	RMSEP (%)	Bias (%)	LVs	R^2_p	RMSEP (% DM)	Bias (% DM)	LVs	R^2_p	RMSEP (% DM)	Bias (% DM)
Olive fruit instrument transfer	5	0.82	2.58	2.28	4	0.91	0.87	−0.05	3	0.88	1.35	−0.86
Mango fruit temperature correction	8	0.76	1.56	0.86	5	0.78	1.26	0.10	3	0.85	1.11	0.39
Mango fruit season correction	6	0.61	2.25	1.40	7	0.80	1.28	−0.26	4	0.75	1.49	−0.54
Apple fruit season correction	6	0.90	0.74	−0.44	5	0.91	0.58	0.008	3	0.93	0.59	−0.29

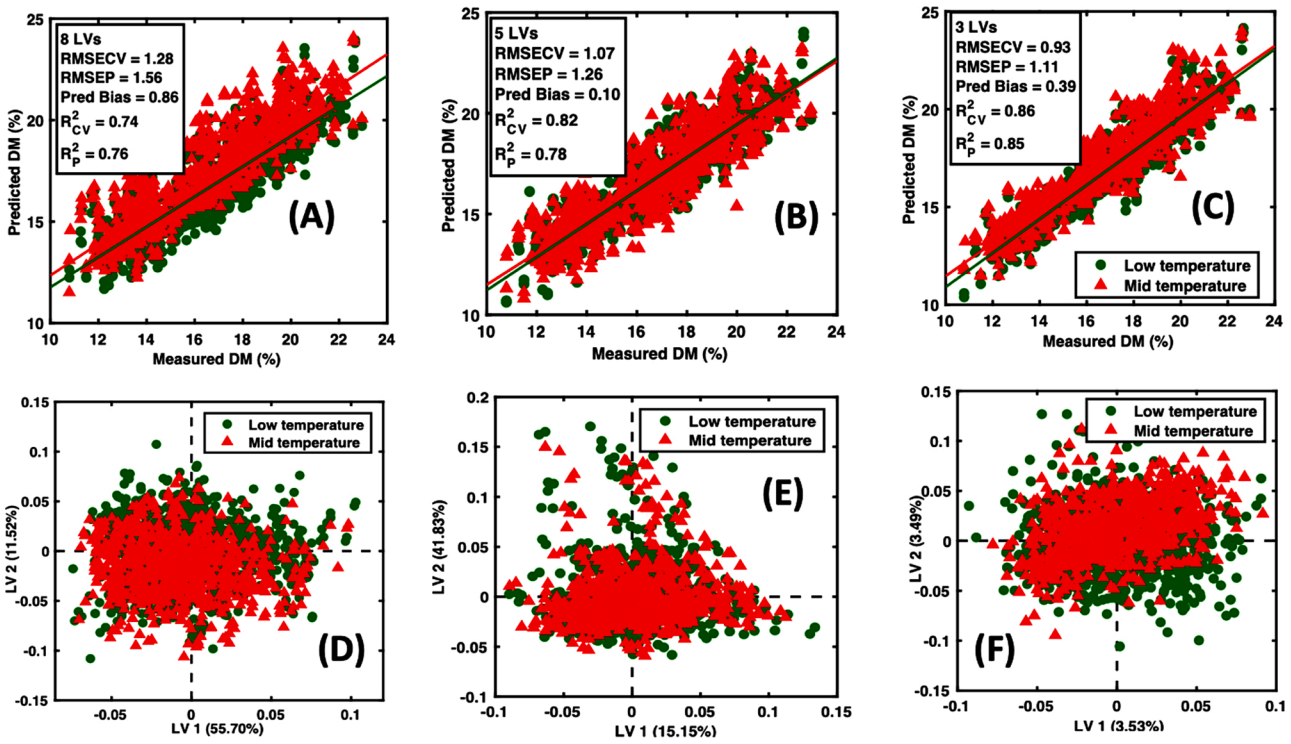

Fig. 3. Temperature correction for prediction of DM (%) in mango fruit. Data from low temperature is presented in green circles and medium temperature in red triangles. (A) Partial least-squares regression (PLSR) calibrated on low temperature tested on medium temperature measurements, (B) PLSR after dynamic orthogonal projection (DOP) calibrated on low temperature tested on medium temperature measurements, (C) PLSR after transfer component analysis (TCA) calibrated on low temperature tested on medium temperature measurements, (D) scores distribution for low and medium temperature measurements after PLSR, (E) scores distribution for low and medium temperature measurements from PLSR after DOP, and (F) scores distribution for low and medium temperature measurements for PLSR after TCA. LV in the score plots stands for latent variables. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

in red. In all the cases, the model was calibrated on season 1 and tested on season 2. It can be noted that both DOP and TCA improved the model performance, i.e. high R^2_p and low RMSEP compared to no data transformation. The scores on the first two latent variables (LVs) are shown in Fig. 4D–F. The scores plots show that the raw data from multiple seasons have differences in distributions (Fig. 4D), as can be seen from the red (season 2) and green (season 1) clusters, which reflect the seasonal effect. The differences were removed/reduced by the transformation performed by DOP and TCA. The R^2_p was highest for DOP with a total 31% increase compared to no transformation. DOP also reduced the prediction bias by 80% compared to no transformation. The RMSEP with the use of DOP was reduced by 43%. The number of LVs was lowest for TCA (4 LVs), followed by no transformation (6 LVs), which was almost the same as DOP (7 LVs).

3.5. Apple fruit multi-season dataset

Fig. 5 presents the results of PLSR modelling on the non-transformed data (Fig. 5A), the DOP (Fig. 5B) and the TCA transformed data (Fig. 5C). The samples from season 1 are in green and from season 2 are in red. In all the cases, the model was calibrated on season 1 and tested on season 2. It can be noted that both DOP and TCA improved the model performance, i.e. high R^2_p and lower RMSEP compared to no data transformation. The scores on the first two latent variables (LVs) are shown in Fig. 5D–F. The scores plots show that the raw data from multiple seasons have differences in distribution (Fig. 5D), as the red (season 2) and green (season 1) clusters show differences due to the seasonal effect. The differences were removed/reduced by the transformation performed by DOP and TCA; as well, the distributions are more compact compared to no data transformation. The R^2_p was highest for TCA with a total 1% increase compared to no transformation. DOP reduced the prediction bias by 98% compared to no transformation. The RMSEP with the use of DOP was reduced by 21% compared to no transformation. The number of LVs was the lowest for TCA (3 LVs), followed by DOP (5 LVs), and then by no transformation (6 LVs).

3.6. Effects of changing response distribution on DOP and TCA performance

The apple fruit multi-season data was used for comparison of the performances in the situation where the test dataset is not like the calibration dataset. Such a problem can be encountered in the real-case scenario when first batch has homogenous normal distribution of fruit properties, while in the second batch, the fruit are either too raw or too ripe i.e. the properties are not normally distributed (Fig. 6A and 6E). Such a distribution difference is of critical importance to the performance of methods like TCA, which relies on distribution matching. The DOP performed better in terms of low prediction bias and RMSEP (Fig. 6C and G) compared to both TCA (Fig. 6D and H) and standard PLSR (Fig. 6B and F). TCA does not perform as well as like DOP as it relies on the distribution matching and resulted in increased bias when data distribution was different. However, the prediction R^2 was highest for TCA as compared to both standard PLSR and DOP with a lower number of LVs (only 3 in the case of TCA).

4. Discussion

The NIR spectroscopy data modelling of fruit suffers from robustness problems (Nicolai et al., 2007; Saeys et al., 2019). Often the calibration models fail to perform well when tested on new season data or measurements performed under different conditions (Zeaiter et al., 2006). There are technical challenges from the point of view of NIR

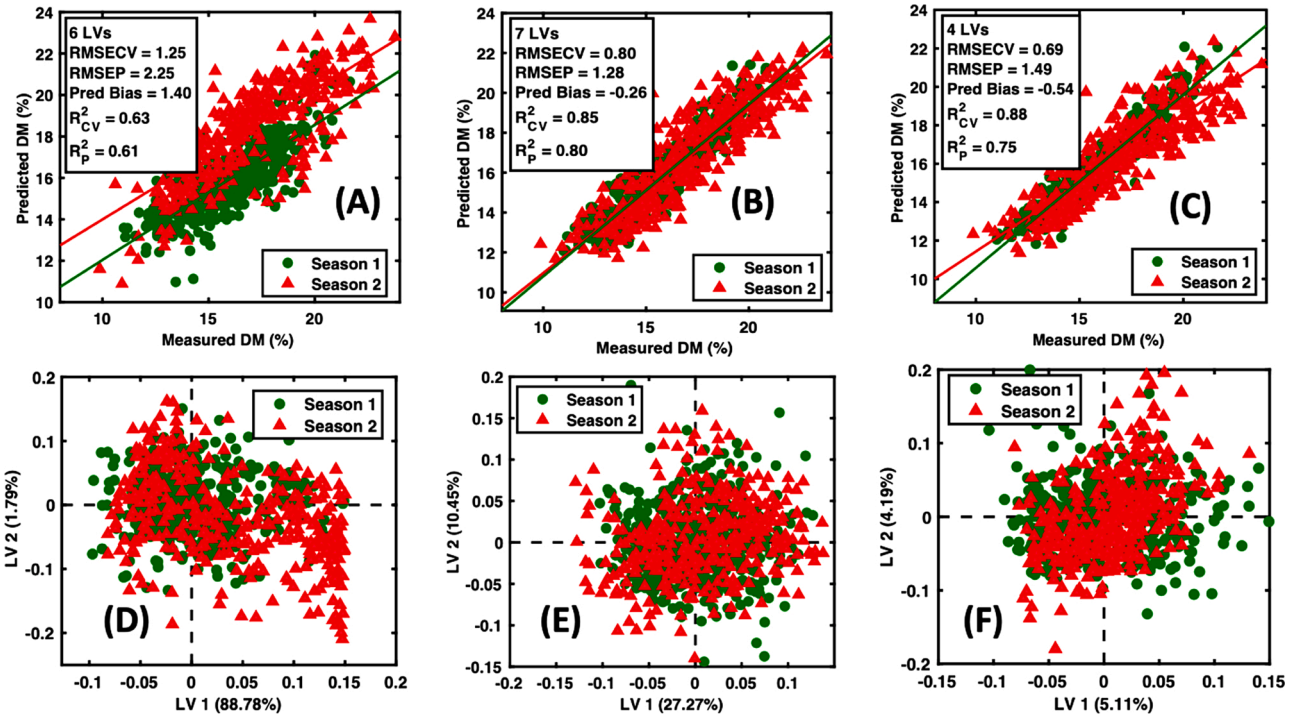

Fig. 4. Season correction for prediction of DM (%) in mango fruit. Data from season 1 is presented in green circles and from season 2 in red triangles. (A) Partial least-squares regression (PLSR) calibrated on season 1 tested on season 2 measurements, (B) PLSR after dynamic orthogonal projection (DOP) calibrated on season 1 tested on season 2 measurements, (C) PLSR after transfer component analysis (TCA) calibrated on season 1 tested on season 2 measurements, (D) scores distribution for season 1 and season 2 measurements after PLSR, (E) scores distribution for season 1 and season 2 measurements from PLSR after DOP, and (F) scores distribution for season 1 and season 2 measurements from PLSR after TCA. LV in the score plots stands for latent variables. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

spectroscopy such as detector sensitivities, differences in light source illuminations, temperature of the sensors, which all induce perturbance in the signals. However, the modelling approaches used to process NIR spectroscopy data do not always consider these perturbations and focus mainly on the development of simple PLSR with a combination of either exploration of various pre-processing techniques or variable selection. Methods such as DOP and TCA to model and remove the influences and to identify the underlying common subspace are still not adapted to the NIR data analysis of fruit properties. In the present work, both DOP and TCA data transformation methods improved model performances (high R^2_p and low RMSEP and bias) without performing any extra standard measurements, when confronted by challenges such as instrument transfer, temperature correction and seasonal variations.

In this study,DOP outperformed all the original results of the instrument transfer methods for the olive data (Sun et al., 2020a,2020b) by obtaining a $R^2_p = 0.91$ and RMSEP and bias of 0.87% and −0.05% respectively. Furthermore, the TCA obtained results comparable to all the classical standard-based instrument transfer methods in an unsupervised approach. In the original work, the results obtained with the classical standard-based instrument transfer methods (direct standardization, model updating, spectral space transformation, piecewise direct standardization) showed as lowest RMSEP 1.32% (corresponding to spectral space transformation) and as bias 0.37% (corresponding to piecewise direct standardization), respectively, with R^2_p ranging from 0.73 to 0.85. In the case of the mango fruit dataset used in this study, the original work gave as the best $R^2_p = 0.82$ and the lowest RMSEP and bias of 1.05% and 0.10%, respectively. In this study, the TCA resulted in a higher $R^2_p = 0.85$ without the need of any extra measurements compared to the extra samples used to perform EPO in the original study. The DOP performed similarly to the original study in terms of similar bias. In comparison to the multi-season data set no direct comparison can be made with the original studies for mango fruit (Anderson et al., 2020), as in this study, we only used data from two seasons data to compare the performance of the model developed in one season with data taken in a different season. In the original study, global modelling was performed with data from multiple seasons for mango fruit. Similarly for apples, (Teh et al., 2020) no direct comparison can be made with the study as the original study used samples from both seasons to calibrate the original models. However, like the original study both the TCA and DOP obtained $R^2_p = 0.91$–0.93 with RMSEP = 0.58–0.59% but with the model being based on data from the only one season, 2015.

DOP and TCA does not require any standard measurements which fits well into the framework of portable spectroscopy which require sharable models, that are widely deployable among users. However, DOP requires a small number of extra measurements performed under the new conditions (physical, chemical and environmental condition) to design the influence factor subspace for orthogonal projection (Zeaiter et al., 2006). TCA is completely unsupervised and does not require any reference measurements (Pan et al., 2011). But TCA assumes that the datasets of the two batches are similar. If the Y distribution differs between the two batches, this will decrease the performance of TCA. For example, in the case of inter-season transfer for mango fruit, the mean value of Y in season 2 is more than 1% higher than in season 1 (see Fig. 1C). Therefore, there is by nature, a difference in the mean spectra of the two seasons, which is contrary to the assumptions of TCA. The correction made by TCA therefore has the effect of removing this mean spectrum from the latent space, thus eroding the useful subspace. Thus, the resulting model becomes poorer compared to DOP. On the contrary, DOP, using reference values of Y, can take this difference into account. Therefore, DOP works better on this example, as shown in Fig. 4(B) and (C) and therefore, in all examples, the bias was the lower for the DOP transformation compared to TCA.

As well, TCA assumes that the distributions of the spectral data are the same for the batches. This will not be the case if the spectra measured on the batches do not refer to the same Y distributions, i.e. at least with similar means and standard deviations. The example of the simulated

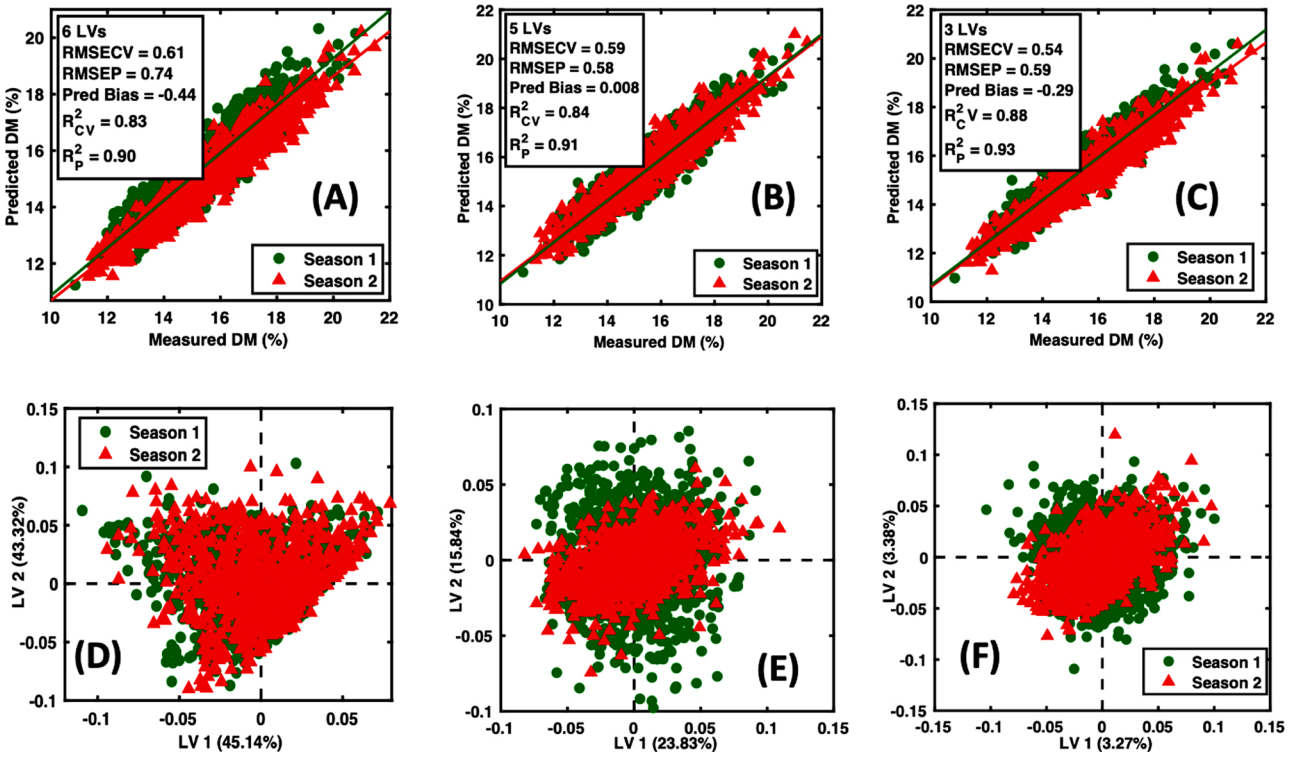

Fig. 5. Season correction for prediction of DM (%) in apple fruit. Data from season 1 is presented in green circles and from season 2 in red triangles. (A) Partial least-squares regression (PLSR) calibrated on season 1 tested on season 2 measurements, (B) PLSR after dynamic orthogonal projection (DOP) calibrated on season 1 tested on season 2 measurements, (C) PLSR after transfer component analysis (TCA) calibrated on season 1 tested on season 2 measurements, (D) scores distribution for season 1 and season 2 measurements after PLSR, (E) scores distribution for season 1 and season 2 measurements from PLSR after DOP, and (F) scores distribution for season 1 and season 2 measurements from PLSR after TCA. LV in the score plots stands for latent variables. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

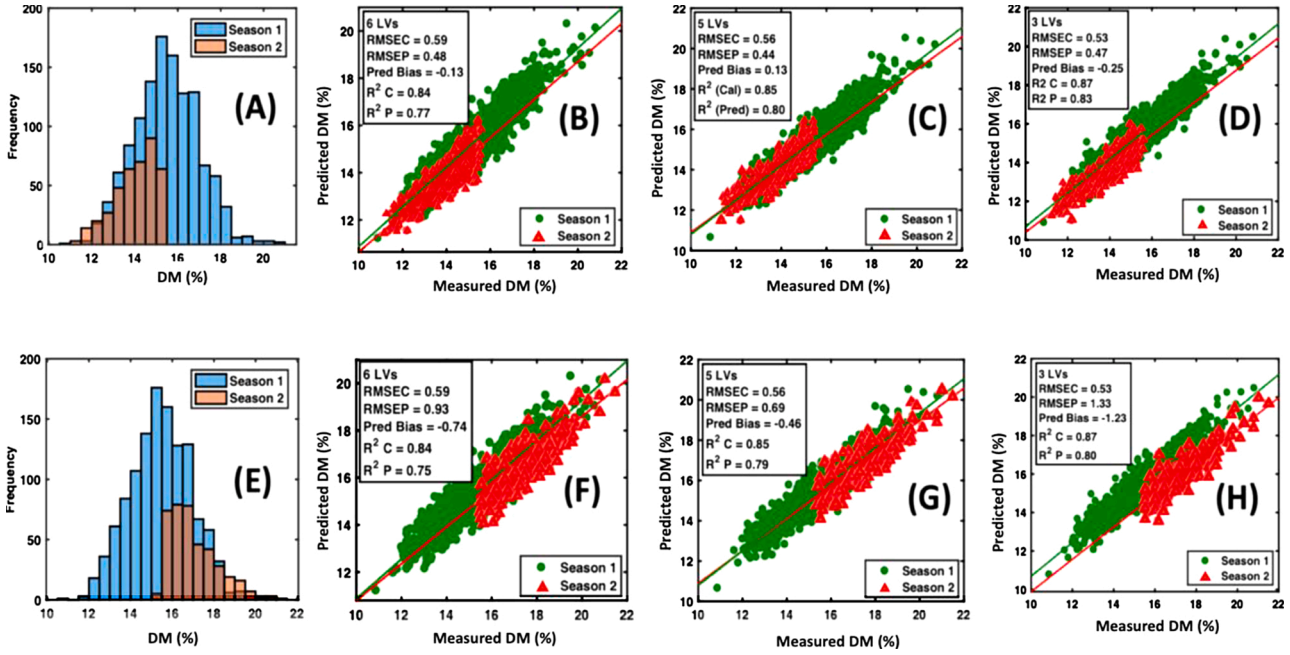

Fig. 6. Comparison of partial least-squares regression (PLSR), dynamic orthogonal projection (DOP) and transfer component analysis (TCA) for different test distributions. The dataset used was the apple fruit multi-season dataset. In regression analysis plots, data from season 1 is presented in green circles and from season 2 in red triangles. (A) Histogram explaining season 1 and reduced season 2 data (low dry matter), (B) PLSR calibrated on season 1 tested on season 2 measurements, (C) PLSR after DOP calibrated on season 1 tested on season 2 measurements, (D) PLSR after TCA calibrated on season 1 tested on season 2 measurements, (E) histogram explaining season 1 and reduced season 2 data (high dry matter), (F) PLSR calibrated on season 1 tested on season 2 measurements, (G) PLSR after DOP PLSR calibrated on season 1 tested on season 2 measurements, and (H) PLSR after TCA PLSR calibrated on season 1 tested on season 2 measurements. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

differences in the apple fruit multi-season data showed that when the test set distributions are very different from the calibration set, the TCA performance is poorer compared to that of DOP (Fig. 6). TCA will perform better in those cases where the chemical properties are identically distributed for the different situations. This favorable situation seems to be encountered often in the case of fruit quality predictions, as in all four real-life cases presented here, the distributions from the different batches were similar.

Another important point is the selection of optimal EPO dimensions for the DOP correction. There are many ways of doing it (Roger and Boulet, 2018). In the present work, due to the availability of the test set, the validation approach, based on tuning samples was used. In those cases where no extra test set is available, non-objective criteria explaining the effect of EPO on the decrease of the detrimental effect can be used. One such approach is to visualize the evolution of the variance of the corrected batch and the second batch, where the inflection point may indicate the optimal number of EPO dimensions (Roger and Boulet, 2018). TCA does not require any such tuning.

5. Conclusions

Multiple physical, chemical and environmental factors affect the NIR spectroscopy models related to fruit quality prediction. In the present work, DOP and TCA were successfully implemented to correct for these effects leading to higher R^2_p and lower bias and RMSEP. The main benefit of DOP and the TCA is limiting the need of new standard measurements for model updating and calibration transfer. The bias was lower in the case of DOP compared to the TCA. Up to 31% increase in $R^2_{p,}$ and 98% and 66% reductions in prediction bias and root mean squared error of prediction, respectively were noted. In addition, the number of LVs was lower for TCA in all the datasets compared to DOP, indicating that TCA captures the underlying subspaces more efficiently than does DOP. This study also shows that the use of TCA requires precautions to be taken concerning the distributions of the responses in the different datasets.

CRediT authorship contribution statement

Puneet Mishra: Conceptualization, Data curation, Investigation. **Jean Michel Roger:** Formal analysis, Software, Visualization. **Douglas N. Rutledge:** Formal analysis, Methodology, Software. **Ernst Woltering:** Writing - review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

Prof. Kerry B. Walsh and Dr. Phul Subedi from Central Queensland University, Australia for sharing the Olive fruit dataset used in the study.

Dr. Soon Li The and Prof. Kate Evans from Washington State University for sharing the Apple fruit multi-season dataset used in the study.

References

Anderson, N.T., Walsh, K.B., Subedi, P.P., Hayes, C.H., 2020. Achieving robustness across season, location and cultivar for a NIRS model for intact mango fruit dry matter content. Postharvest Biol. Technol. 168, 111202 https://doi.org/10.1016/j.postharvbio.2020.111202.

Engel, J., Gerretzen, J., Szymańska, E., Jansen, J.J., Downey, G., Blanchet, L., Buydens, L. M.C., 2013. Breaking with trends in pre-processing? TrAC Trends Anal. Chem. 50, 96–106. https://doi.org/10.1016/j.trac.2013.04.015.

Fearn, T., 2001. Standardisation and calibration transfer for near infrared instruments: a review. J. Near Infrared Spectrosc. 9 (4), 229–244. https://doi.org/10.1255/jnirs.309.

Feudale, R.N., Woody, N.A., Tan, H., Myles, A.J., Brown, S.D., Ferré, J., 2002. Transfer of multivariate calibration models: a review. Chemom. Intell. Lab. Syst. 64 (2), 181–192. https://doi.org/10.1016/S0169-7439(02)00085-0.

Kouw, W.M., Loog, M., 2018. An introduction to domain adaptation and transfer learning. arXiv preprint arXiv 1812, 11806.

Kouw, W.M., Loog, M., 2019. A review of domain adaptation without target labels. IEEE Trans. Pattern Anal. Mach. Intell. arXiv:1901.05335.

Lin, H., Ying, Y., 2009. Theory and application of near infrared spectroscopy in assessment of fruit quality: a review. Sens. Instrum. Food Qual. Saf. 3 (2), 130–141. https://doi.org/10.1007/s11694-009-9079-z.

Nicolai, B.M., Beullens, K., Bobelyn, E., Peirs, A., Saeys, W., Theron, K.I., Lammertyn, J., 2007. Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: a review. Postharvest Biol. Technol. 46 (2), 99–118. https://doi.org/10.1016/j.postharvbio.2007.06.024.

Pan, S.J., Tsang, I.W., Kwok, J.T., Yang, Q., 2011. Domain adaptation via transfer component analysis. IEEE Trans. Neural Netw. 22 (2), 199–210. https://doi.org/10.1109/tnn.2010.2091281.

Pasquini, C., 2018. Near infrared spectroscopy: A mature analytical technique with new perspectives – a review. Anal. Chim. Acta 1026, 8–36. https://doi.org/10.1016/j.aca.2018.04.004.

Rinnan, Å., Berg, F. v. d., Engelsen, S.B., 2009. Review of the most common pre-processing techniques for near-infrared spectra. TrAC Trends Anal. Chem. 28 (10), 1201–1222. https://doi.org/10.1016/j.trac.2009.07.007.

Roger, J.-M., Boulet, J.-C., 2018. A review of orthogonal projections for calibration. J. Chemom. 32 (9), e3045. https://doi.org/10.1002/cem.3045.

Roger, J.-M., Boulet, J.-C., Zeaiter, M., Rutledge, D.N., 2020. Pre-processing Methods★. Reference Module in Chemistry, Molecular Sciences and Chemical Engineering. Elsevier. https://doi.org/10.1016/B978-0-12-409547-2.14878-4.

Roger, J.-M., Chauchard, F., Bellon-Maurel, V., 2003. EPO–PLS external parameter orthogonalisation of PLS application to temperature-independent measurement of sugar content of intact fruits. Chemom. Intell. Lab. Syst. 66 (2), 191–204. https://doi.org/10.1016/S0169-7439(03)00051-0.

Roger, J.M., 2016. Orthogonal projections in the row and the column spaces. NIR News 27 (7), 15–20. https://doi.org/10.1255/nirn.1640.

Saeys, W., Do Trong, N.N., Van Beers, R., Nicolai, B.M., 2019. Multivariate calibration of spectroscopic sensors for postharvest quality evaluation: a review. Postharvest Biol. Technol. 158 doi:UNSP 11098110.1016/j.postharvbio.2019.110981.

Subedi, P.P., Walsh, K.B., 2009. Non-invasive techniques for measurement of fresh fruit firmness. Postharvest Biol. Technol. 51 (3), 297–304. https://doi.org/10.1016/j.postharvbio.2008.03.004.

Sun, X., Subedi, P., Walker, R., Walsh, K.B., 2020a. NIRS prediction of dry matter content of single olive fruit with consideration of variable sorting for normalisation pre-treatment. Postharvest Biol. Technol. 163, 111140 https://doi.org/10.1016/j.postharvbio.2020.111140.

Sun, X.D., Subedi, P., Walsh, K.B., 2020b. Achieving robustness to temperature change of a NIRS-PLSR model for intact mango fruit dry matter content. Postharvest Biol. Technol. 162 https://doi.org/10.1016/j.postharvbio.2019.111117. ARTN 111117.

Teh, S.L., Coggins, J.L., Kostick, S.A., Evans, K.M., 2020. Location, year, and tree age impact NIR-based postharvest prediction of dry matter concentration for 58 apple accessions. Postharvest Biol. Technol. 166, 111125 https://doi.org/10.1016/j.postharvbio.2020.111125.

Walsh, K.B., McGlone, V.A., Han, D.H., 2020. The uses of near infra-red spectroscopy in postharvest decision support: A review. Postharvest Biol. Technol. 163, 111139 https://doi.org/10.1016/j.postharvbio.2020.111139.

Wang, H.L., Peng, J.Y., Xie, C.Q., Bao, Y.D., He, Y., 2015. Fruit quality evaluation using spectroscopy technology: a review. Sensors 15 (5), 11889–11927. https://doi.org/10.3390/s150511889.

Wold, S., Sjostrom, M., Eriksson, L., 2001. PLS-regression: a basic tool of chemometrics. Chemom. Intell. Lab. Syst. 58 (2), 109–130. https://doi.org/10.1016/S0169-7439(01)00155-1.

Workman, J.J., 2017. A review of calibration transfer practices and instrument differences in spectroscopy. Appl. Spectrosc. 72 (3), 340–365. https://doi.org/10.1177/0003702817736064.

Zeaiter, M., Roger, J.M., 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. Chemom. Intell. Lab. Syst. 80 (2), 227–235. https://doi.org/10.1016/j.chemolab.2005.06.011.