ScienceDirect

Analytica Chimica Acta
Volume 1225, 8 September 2022, 340154

Unsupervised dynamic orthogonal projection. An efficient approach to calibration transfer without standard samples

Valeria Fonseca Diaz [a], Jean-Michel Roger [b c], Wouter Saeys [a]

Show more

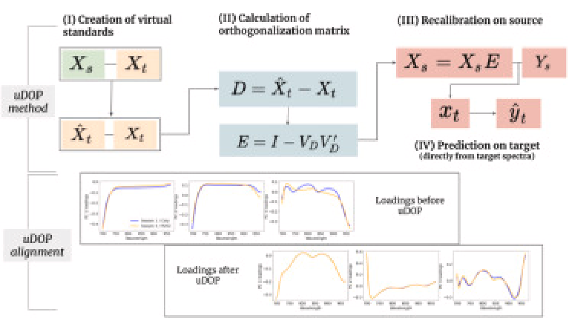
Share Cite

https://doi.org/10.1016/j.aca.2022.340154
Get rights and content

Abstract

Calibration transfer has been traditionally performed in the context of transferring models between instruments using standard samples. Recently, new methodologies and applications have shown that transfer techniques can be adopted to achieve calibration transfer between other types of domains, such as product form, variant or seasonality. In addition, to achieving a higher efficiency for calibration transfer, it is desirable to perform the transfer without the need for standard samples or new reference analyses. Therefore, we propose a method for unsupervised calibration transfer based on the orthogonalization for structural differences between domains. The method has been successfully applied to one simulated dataset and two real datasets. In the studied cases, the proposed methodology allowed to achieve a successful transfer of calibration models and enabled the interpretation of the interferences responsible for the degradation of the original calibration models when transferred to the new domain.

Graphical abstract



Download : Download high-res image (229KB)
Download : Download full-size image

Introduction

The quantification of chemical constituents is increasingly done in an indirect way using multivariate sensors (e.g. spectroscopy) which require calibration models. As these models are built on a limited set of training data, methodological approaches have been proposed to keep calibration models valid in the long term [[1], [2], [3], [4], [5]]. Thanks to the rapid evolution of technology, there are many opportunities to make processes more efficient through the implementation of new instrumentation. Complementary, there is an expansion of the domains in which the same products are manufactured or harvested, such as new plants, new cultivars or new compositions. The implementation of the calibration models in these new domains poses specific challenges. To fully exploit the potential of this technological progress, the calibration models should be easily adapted to the sources of variability that new measurements can bring to avoid that their prediction performance degrades [6].

As manufacturers struggled in producing instruments which were completely identical and practitioners were reluctant to repeat the entire calibration phase for every new instrument, calibration transfer methodologies have been elaborated which allow to transfer a calibration model from one source or master instrument to target or secondary instruments with minimal degradation of the model performance [1,7,8]. To this end, linear standardization methods have been proposed which match the instrument responses using so-called standard samples. These samples are units that are measured simultaneously with the source and target instruments. The most popular methods that were designed for this purpose are (Piecewise) Direct Standardization (DS, PDS) [9,10] and Spectral Space Transformation (SST) [2]. These methods achieve the standardization of the instrument signal by calculating a transformation matrix that can be used to adapt the response of the secondary or target instrument as if it was measured in the primary or source instrument, so that the existing calibration model based on the latter can be used. While calibration transfer has been traditionally understood as the transfer between instruments, the general definition of model transfer goes beyond the domain of the instrumentation, enabling to achieve calibration transfer between other types of domains, such as product form, variant or seasonality [3,11,12]. Standardization methods are suitable for transfer problems in which standard samples are available or can be measured. However, acquiring standards may not be possible when the model needs to be transferred between seasons, product variants, instruments that are located far away from each other or even instrumentation that is no longer available. This more general definition of calibration transfer as the adaptation of a calibration model to be valid in a new target domain, requires methods which can cope with cases where such standard samples are not available [1].

One way to cope with the unavailability of standard samples involves adaptation of an existing calibration based on new reference analyses acquired in the target domain. A popular approach in this category is the so-called Slope and Bias Correction (SBC) [7] which allows to adjust the predictions in the target domain. Other methods have been designed to enhance the transfer when predictions need more than a slope and bias adjustment, such as joint modelling [5,13] and supervised orthogonalization [14]. Dynamic orthogonal projection (DOP) is one of the most effective methods to transfer models between any type of domains. It involves an estimation of the chemical, physical, or environmental sources of variation that are responsible for the differences between the signals acquired in the source and target domains through association of samples via their reference values [14].

As calibration transfer without the need for standard samples is highly desirable, the newest contributions in this field have focused on the design of methodologies that do not require standard samples or new reference analyses [3,11,12,15]. The most predominant methodology has been categorized as domain invariant or domain adaptation [16,17]. The main characteristic of these methods is to model the bilinear relationship between spectral measurements and reference values with a constraint for domain invariability between a source and a target domain. The two most recent developments in this context are Domain Invariant Partial Least Squares (di-PLS) [3] and Domain Adaptive Partial Least Squares (da-PLS) [18]. Such a constraint for domain invariability has also been adopted to update models using the differences in the covariance matrices of the domains as predictors submatrices [15].

The domain invariant methods have proven to be successful in building bilinear models that account for the differences in variability observed in the source and target domains [12,18,19]. One of the main challenges of these methods that are still under study is the establishment of suitable approaches to tune the regularization parameter of these models which increases their complexity compared to classical Partial Least Squares Regression (PLSR) [15,19]. Therefore, we investigated a method that can deliver a satisfactory transfer of models without the need for standard samples or new reference analyses with a lower complexity of training.

Motivated by the concept of dynamic orthogonal projection (DOP) to create virtual standard samples that would provide information of the interference for transfer, the aim of this study was to develop a variant which does not require standard samples or new reference values. The proposed method relies on existing source calibration data and a separate set of spectral measurements from the target domain. Starting from these data, it models the differences between the domains based on the loadings structure of the spectral spaces to calculate components to be filtered out from the calibration spectra. This enables the training of a model suitable for use on the target domain.

This article is organized as follows: First, the development of the method is presented with a positioning of its equivalence with the state of the art including transfer methods involving orthogonalization. Second, the potential of the method is demonstrated with one simulation dataset and two real case studies. The latter correspond to applications in the agrofood industry related to transfer between instruments for sugarcane samples and transfer between cultivars and seasons for mango samples. The obtained performance results are compared to those obtained with the reported methods. Finally, conclusions and recommendations are presented.

Section snippets

Theory

The proposed method is inspired by DOP in which virtual standards are calculated in order to obtain an orthogonalization matrix that can filter the differences between the involved domains or instruments. The fundamental difference relies on how these virtual standards are calculated. For that, only the spectral measurements of two separate sets of samples in each domain are used. Therefore, the proposed method has been named Unsupervised Dynamic Orthogonal Projection (uDOP). The virtual…

Data sets

Three datasets were used to evaluate the uDOP method: one simulation set and two real case studies. The first dataset corresponds to a simulation of high-dimensional data using the principles of bilinear modeling for the spectral data and linear regression for the relationship with the response variable [21]. In order to create a dataset with a low signal-to-noise ratio (SNR) so that the model training and the domain adaptation would be more similar to real case studies, particularly in the…

Results

The mean spectra of each dataset are illustrated in Fig. 1 for the corresponding source and target domains. The difference between domains in the simulation dataset can be observed as a shift in the spectral values between variables 600 and 800. In the case of the sugarcane dataset, the two instruments show an offset from the source to the target domain whose magnitude does not appear to be constant. The mean structure in the mango dataset for both domains seems to change as the difference…

Discussion

In the case of DMC prediction in mango samples, the transfer by uDOP was not so successful as that obtained with DOP. The degradation of the source calibration model when applied to samples from a different cultivar measured three seasons later combined a large dispersion with nonlinearity. In this case, the model adaptation by TOP was not successful, which suggests that the degradations were caused by strong structural and variability differences in the domains. Although DOP proved to be the…

Conclusions

As the need for standard samples or new reference analyses limits the possibilities for calibration transfer, we proposed a novel method for model transfer based on orthogonalization which does not require standard samples or reference analyses for the target domain. We demonstrated the applicability and suitability of the uDOP method for calibration transfer problems regarding transfer between instruments as well as transfer or adaptation for domains related to seasonality and cultivar groups. …

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.…

Acknowledgement

Valeria Fonseca Diaz is funded as aspirant to doctoral fellow of the Research Foundation-Flanders (FWO Brussels, Belgium). We thank Dr. Ramin Nikzad-Langerodi for the important discussions about domain invariability methodology for calibration transfer and Dr. Puneet Mishra for his support on literature review and sharing open data sources.…

Recommended articles

References (31)

P. Mishra *et al.*
Are standard sample measurements still needed to transfer multivariate calibration models between near-infrared spectrometers? The answer is not always
TrAC, Trends Anal. Chem. (2021)

W. Du *et al.*
Maintaining the predictive abilities of multivariate calibration models by spectral space transformation
Anal. Chim. Acta (2011)

R. Nikzad-Langerodi *et al.*
Domain adaptation for regression under Beer–Lambert's law
Knowl. Base Syst. (2020)

P. Shan *et al.*
Principal component analysis or kernel principal component analysis based joint spectral subspace method for calibration transfer
Spectrochim. Acta, Part A (2020)

P. Mishra *et al.*
Handling batch-to-batch variability in portable spectroscopy of fresh fruit with minimal parameter adjustment
Anal. Chim. Acta (2021)

E. Bouveresse *et al.*
Improvement of the piecewise direct standardisation procedure for the transfer of NIR spectra for multivariate calibration
Chemometr. Intell. Lab. Syst. (1996)

V.F. Diaz *et al.*
Domain invariant covariate selection (Di-CovSel) for selecting generalized features across domains
Chemometr. Intell. Lab. Syst. (2022)

M. Zeaiter *et al.*

Dynamic orthogonal projection. A new method to maintain the on-line robustness of multivariate calibrations. Application to NIR-based monitoring of wine fermentations
Chemometr. Intell. Lab. Syst. (2006)

G. Huang *et al.*
Domain adaptive partial least squares regression
Chemometr. Intell. Lab. Syst. (2020)

Z. Huang *et al.*
Improved generalization of spectral models associated with Vis-NIR spectroscopy for determining the moisture content of different tea leaves
J. Food Eng. (2021)

View more references

Cited by (7)

Standard-free sample model transfer of visible/near -infrared spectral model of apple ripeness under seasonal variation
2024, Journal of Food Composition and Analysis

Show abstract

Application of SWSRA-DS algorithm in improving the model transfer for near infrared analysis of pulpwood holocellulose
2023, Infrared Physics and Technology

Show abstract

Monitoring of multivariate calibration models in the absence of new reference values: The regression case
2023, Chemometrics and Intelligent Laboratory Systems

Show abstract

Predicting components of pulpwood feedstock for different physical forms and tree species using NIR spectroscopy and transfer learning ↗
2024, Cellulose

Standard-Free Sample Calibration Transfer of Nir Spectral Model of Apple Ripeness Under Seasonal Variation ↗
2023, SSRN

A Federated Classification Approach of Waste Lubricant Oils in Geographically Distributed Laboratories ↗
2022, Industrial and Engineering Chemistry Research

View all citing articles on Scopus ↗

View full text