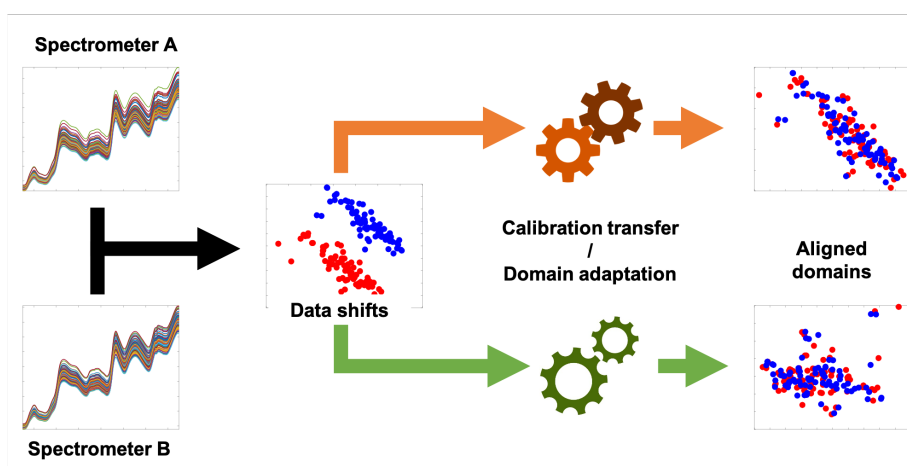


1 Graphical Abstract

2 Cloning instruments, model maintenance and calibration transfer

3 Jean-Michel Roger, Valeria Fonseca Diaz, Ramin Nikzad Langerodi



4 Highlights

5 **Cloning instruments, model maintenance and calibration transfer**

6 Jean-Michel Roger, Valeria Fonseca Diaz, Ramin Nikzad Langerodi

- 7 • A unified view of model maintenance, incorporating chemometrics and
8 machine learning perspectives
- 9 • A review of methods according to their ability to handle different types
10 of data shift
- 11 • A strategy for managing the model maintenance calibration problem

12 Cloning instruments, model maintenance and
13 calibration transfer

14 Jean-Michel Roger^{a,b}, Valeria Fonseca Diaz^c, Ramin Nikzad Langerodi^c

^a*ITAP INRAE, Institut Agro, University Montpellier, Montpellier, 34196, France*

^b*ChemHouse Research Group, Montpellier, 34196, France*

^c*Data Science Group, Software Competence Center Hagenberg (SCCH)
GmbH, Hagenberg, Austria*

15 **Abstract**

Most literature on the application of Non-Destructive Spectral Sensors (NDSS) reports proofs of concept limited to model calculation (calibration) and its application on a so-called independent data set (validation, or test). However, developing NDSS also requires proving that the performance obtained during this first validation remains valid when conditions change. This generic problem is referred to as robustness in chemometrics. When the measurement conditions change, the measured spectrum is subject to a deviation. The reproducibility of the model, and thus of the sensor, with respect to this deviation, defines its robustness. The application of NDSS involves a large number of processes, and thus deviation sources. Instrument cloning, between laboratory instruments or from a benchtop to an online device, is certainly the most concerning issue for deploying NDSS-based applications. This problem has been studied for many years in chemometrics, under the paradigm of calibration transfer, through geometric corrections of spectra, spectral spaces, or calibration model corrections. The same problem has been addressed in the machine learning community under the domain adaptation paradigm. Although all these issues have been addressed separately over the last twenty years, they all fall under the same topic, i.e., model maintenance under dataset shift. This paper aims to provide a vocabulary of concepts for formalizing the calibration model maintenance problem, reviewing recent developments on the subject, and categorizing prior work according to the proposed concepts.

16 *Keywords:* Calibration model maintenance, Calibration transfer, Domain
17 adaptation

Preprint submitted to TrAC

May 7, 2024

18 1. Introduction

19 Public policies on food safety, quality, and traceability rely on a global,
20 unified characterization of the entire food supply chain. Moreover, the food
21 industry's analytical needs go beyond simply complying with current regu-
22 lations; they also encompass the imperative of controlling processes, under-
23 standing the variability of raw materials and final products, and how one
24 influences the other. All this has necessitated the adoption of new analyti-
25 cal technologies to meet the demands of industry, consumers, and regulators
26 alike. Non-destructive spectral sensors (NDSS) - such as near-infrared spec-
27 troscopy, hyperspectral imaging, fluorescence, and Raman spectroscopy - are
28 particularly well suited to answering these questions [1], [2], [3]. NDSS offers
29 a fast, non-destructive, and environmentally friendly method for assessing
30 numerous parameters in a wide range of products and processes. They have
31 proved their worth outdoors, indoors, and online, from farm to supermarket,
32 for a wide range of products [4]. However, NDSS operation relies on cali-
33 bration, which relates the measured signals (a.k.a. spectra) to the desired
34 response, such as analyte concentration or the product class (e.g. fruit vari-
35 ety). Maintaining the predictive accuracy of multivariate calibration models
36 upon matrix, instrumental, or environmental (i.e., measurement condition)
37 changes is one of the key challenges for the practical success of NDSS-based
38 systems. Despite the large number of studies on the subject and the avail-
39 ability of a vast number of algorithms, approaches, and/or workflows for cali-
40 bration model maintenance (CMM), the problem seems to be far from solved
41 as of now. The lack of a theoretical framework formalizing the CMM prob-
42 lem that would allow the categorization of current methodologies and predict
43 under which conditions they are suited for CMM, as well as the spillover of
44 ideas from the machine learning community (coupled with the terminology
45 gap with chemometrics) complicates the alignment of research activities and
46 the efficient application of existing approaches to address CMM in practice.
47 Thus, this review aims at:

- 48 i) Providing a vocabulary of concepts towards formalizing the CMM prob-
49 lem
- 50 ii) Reviewing recent developments on the subject
- 51 iii) Categorizing prior work according to the concepts proposed in i)

52 The rest of this review is organized as follows: Section 2 introduces com-
53 mon terminology related to calibration model maintenance, provides a for-
54 mal problem statement, and reviews some historical background on how the
55 problem has been addressed since the 1980s, highlighting some more recent
56 developments rooted in machine learning research. Section 3 reviews rele-
57 vant literature related to the subject, and section 4 concludes the paper and
58 outlines open problems.

59 2. Problem Statement

60 Statistical learning theory posits that a model learned from some train-
61 ing data - in this case spectra and corresponding quality attributes - will
62 produce accurate predictions on new, test data as long as the underlying
63 distributions of training and test data are equivalent [5]. However, in most
64 real-world applications pertaining to NDSS, the data generating process, and
65 thus the data distribution, tends to change upon internal or external influ-
66 ences that might deteriorate the predictive capability of the corresponding
67 calibration model [6]. Thereby a change in the joint distribution over spectra
68 X and corresponding quality attribute Y (e.g., analyte concentration), i.e.,
69 $P(X, Y)$, can occur due to changes in the so-called *marginal* distribution
70 over the spectra, i.e., $P(X)$, the quality attribute, i.e., $P(Y)$, and/or the
71 *conditional* distribution $P(Y|X)$. The corresponding changes are referred
72 to as *covariate*, *prior* and *conditional* shift, respectively. The dataset shift
73 concept is often associated with the notion of a domain, i.e., the entity where
74 the joint distribution is stationary. Formally, a domain is composed of a pair
75 $\langle P(X), P(Y|X) \rangle$. The domains in which a model is trained and deployed
76 are often referred to as *source domain* ($\mathcal{D}_S := \langle P_S(X), P_S(Y|X) \rangle$) and *target*
77 *domain* ($\mathcal{D}_T := \langle P_T(X), P_T(Y|X) \rangle$), respectively.

78 The implications of dataset shift on model performance are diverse and
79 depend upon which distributions are affected by the changes in the spec-
80 tra and/or quality attributes, and the particular nature of the change (e.g.,
81 whether it is a constant offset affecting the spectra or a more complicated
82 change in the covariance structure of the marginal distribution). Figure 1
83 illustrates the (base) types of prediction drift that must be expected under
84 dataset shift. Often a combination of the prediction drift types is encoun-
85 tered, such as a drift in slope and bias.

86 As illustrated in Figure 2a, a calibration model is robust against prior
87 shift only if the conditional distribution does not change between the source

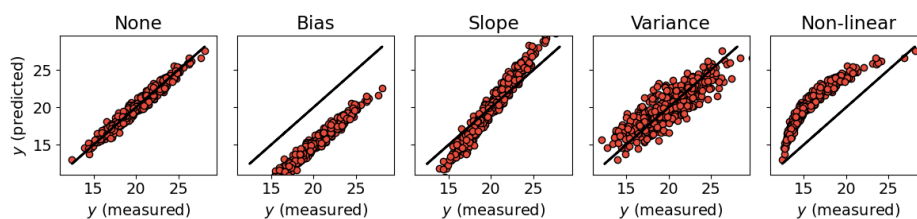


Figure 1: Prediction drift types.

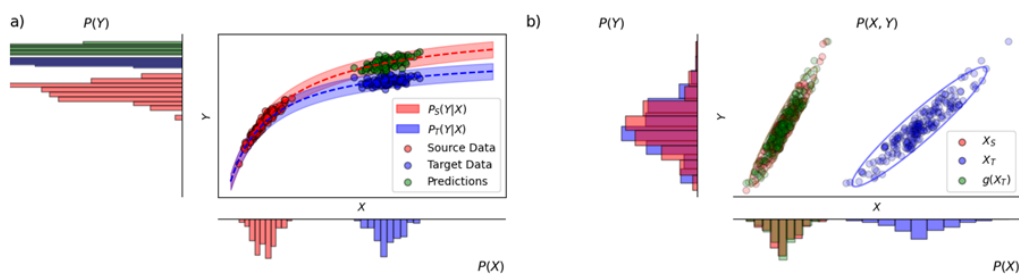


Figure 2: Dataset shifts. a) Prior shift with (blue area) and without (red area) conditional shift. The dashed lines indicate (optimal) source and target domain models, and the green data points indicate predictions of the source model on the target domain patterns X . b) Covariate shifts given some source (X_S) and target (X_T) domain data. The function $g(\cdot)$ re-scales X_T to match the distribution of X_S leading to an alignment of the joint distribution $P(X, Y)$.

88 and target domain, i.e., if $P_S(Y|X) = P_T(Y|X)$. Under conditional shift,
 89 however, there is a risk that the source domain model predictions on the
 90 target domain spectra (green points) systematically deviate from the true
 91 values of Y (blue points). In contrast to prior shift, which is always associ-
 92 ated either with a covariate (as in Figure 2a) or conditional shift¹, covariate
 93 shift exclusively affects the marginal distribution over the spectra ($P(X)$).
 94 Thus, it can be corrected by proper alignment of the corresponding marginal
 95 distributions (Figure 2b).

96 The goal of instrument cloning and model maintenance in NDSS is to use
 97 data sampled from either \mathcal{D}_S and/or \mathcal{D}_T to obtain a model that performs
 98 well in the latter – usually in terms of the mean squared error (MSE) of
 99 prediction. On the other hand, the goal of robust calibration is to derive a

¹Note that if $P_S(X)$ would be equivalent to $P_T(X)$ in Figure 2A, while $P_S(Y) \neq P_T(Y)$, implies conditional shift.

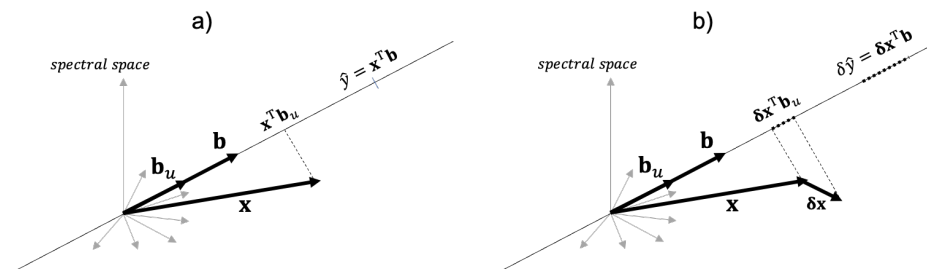


Figure 3: Illustration of the effect of a spectral perturbation on the model output. \mathbf{x} is the measured spectrum, $\delta \mathbf{x}$ is the spectral perturbation, \hat{y} is the predicted response and $\delta \hat{y}$ is the resulting error. Inspired from [13]

100 model that is invariant with respect to the domain in which it is deployed,
 101 such that on average it has the same performance in every (possible) domain.

102 *2.1. How the robustness problem has been handled in chemometrics*

103 Any quality control regulatory body has not officially defined the concept
 104 of robustness. However, it is acknowledged as a critical quality of the NDSS
 105 calibration models. The robustness has been extensively discussed in the
 106 literature, with various definitions tailored to each specific field of application
 107 [7, 8, 9]. In chemometrics applied to NDSS, robustness can be defined as the
 108 reproducibility of the calibration model in the face of spectral perturbations
 109 [10].

110 When the *measurement conditions* of the spectrum change, a spectral
 111 disturbance $\delta \mathbf{x}$ is added to the spectrum \mathbf{x} which should have been measured
 112 in calibration conditions [11]. As illustrated in Figure 3, the addition of $\delta \mathbf{x}$
 113 provokes an error $\delta \hat{y}$ in the y estimation [12, 13]. In Figure 3-a, the measured
 114 spectrum is projected onto the model direction, yielding an estimated
 115 response \hat{y} . In Figure 3-b, it can be noticed that following the same mech-
 116 anism, the spectral perturbation provokes an error $\delta \hat{y}$. Thus, improving the
 117 robustness of a calibration relies on managing $\delta \mathbf{x}$ and its influence on \hat{y} .

118 Changes in measurement conditions, which cause spectral disturbances,
 119 may correspond to changes in a continuous variable, such as product tem-
 120 perature, particle size, or spectrometer temperature. Likewise, they may be
 121 linked to changes in a discrete variable, such as the instrument, the season,
 122 or the product origin. Each specific value of this variable that influences
 123 the measurements establishes a measurement condition. Thus, maintaining

124 the calibration's robustness relies on keeping the model running under differ-
125 ent measurement conditions. In the chemometrics community, consideration
126 of this problem dates back to the era of personal computers, i.e., to the
127 1980s. The concept of measurement condition initially concerned the change
128 of spectrometer, to develop calibrations that are robust across instruments
129 of laboratory networks. In 1985, Shenk discussed the possibility of not re-
130 calibrating a model when changing spectrometers. He suggested performing
131 a *calibration transfer* instead, following his patented method [14, 15]. The
132 change of spectrometer then became an emblematic cause of the robustness
133 problem, and the vocabulary *master* and *slave* became established to des-
134 ignate two measurement conditions between which the calibration transfer
135 takes place [14].

136 Shenk's method acts by *geometrically correcting* slave spectra to resemble
137 master spectra (or vice versa). Other geometric methods have been proposed,
138 introducing the term *optical standardization* [16, 17, 18, 19, 20, 21]. All these
139 methods try to suppress the source of disturbances, i.e., to lower $\delta\mathbf{x}$ in Figure
140 3.

141 Unlike previous methods, which treat spectra as signals, methods for
142 *correcting the measurement vector space* emerged in the 2000s. In 2001,
143 Wise *et al.* proposed to correct the differences between master and slave
144 covariance structures using a General Least Squares (GLS) [22]. Next came
145 orthogonal projection-based methods, which remove from the spectral vector
146 space the subspace generated by variations in measurement conditions [23,
147 24, 25, 26]. All these methods try to make the model space independent from
148 the disturbances, i.e. to lower the projection of $\delta\mathbf{x}$ on \mathbf{b} in Figure 3, and thus
149 making the model as close as possible to the *Net Analyte Signal* (NAS) [27].

150 2.2. How machine learning concepts can help to improve the robustness

151 From a statistical point of view, the "lack of robustness under data set
152 shift" problem is a so-called *sample selection bias* problem. It arises due to
153 the fact that, in practice, it is usually impossible to calibrate using a com-
154 prehensive sample covering all possible sources of variability (either mea-
155 surement and sample related) affecting the spectral signals. Inferring the
156 parameters of a statistical model under a biased sample is a long-standing
157 problem in statistics and related disciplines. In the 1970s Heckmann, in his
158 seminal work on estimating statistical models under sample selection bias,
159 proposed to estimate the selection probability for each (calibration) sample
160 and use this probability to derive an unbiased estimate under the ordinary

161 least squares (OLS) model [28]. However, the Heckmann correction requires
162 the observation of some (instrument) variable that constitutes the sample
163 selection mechanism (e.g., when it is known that only samples with some
164 interferent concentration above a certain threshold are included in the cali-
165 bration set). In the mid-2000s, with the advent of (open-source) deep learning
166 models, the sample selection bias problem was re-discovered by the machine
167 learning community. The importance weighting approach by Zadrozny rep-
168 represents an important development from this period that has constituted the
169 fields of *domain adaptation* (DA) and *transfer learning* (TL) [29]. Impor-
170 tance weighting assigns a weight to each source domain sample (\mathbf{x}, y) that is
171 proportional to $\frac{P_T(X=\mathbf{x})}{P_S(X=\mathbf{x})}$, i.e., assigning a high weight to those samples that
172 have a high probability under the target domain (marginal) distribution.

173 Transfer learning and domain adaptation are machine learning problems
174 where a model developed respectively for one task, or domain, should be
175 reused to improve learning on a related task, or in a related domain [30].
176 For example, TL is used to fine-tune a model, initially developed to dis-
177 criminate between images of cats and dogs, to enable discrimination between
178 images of cars and boats. DA is a special case of TL, where the task does
179 not change across domains [31]. The first generally accepted formalization
180 of TL/DA was proposed in 2006 by Ben-David *et al.*, which provided impor-
181 tant insights into the necessary conditions for generalization under data set
182 shift [32, 33] prompting a research boom on the subject in machine learning
183 [34] and later also in chemometrics community [6]. By applying the domain
184 adaptation paradigm to analytical chemistry, we can notice that the concepts
185 of initial and secondary conditions, or master and slave instruments, com-
186 monly used in chemometrics correspond to the ones of *source domain* and
187 the *target domain* used in DA. Furthermore, different problems encountered
188 in analytical chemistry can be characterized by a particular type of dataset
189 shift. Calibration transfer problems, for instance, often amount to correcting
190 for covariate shift while a change in sample matrix often leads to conditional
191 shift [6]. Finally, the robustness concept can be defined as the reproducibility
192 of a calibration model in the face of domain changes.

193 2.3. How to manage the robustness

194 Managing the robustness of calibration models benefits from following a
195 strategy that depends on the characteristics of the influencing variables, i.e.
196 the domains [7, 35, 36, 37]. Figure 4 shows a decision tree that can help select

197 a domain transfer method, depending on the characteristics of the domain
198 in which the model is applied.

- 199 • The first step is to decide how damaging the domain change really is.
200 Answering this question is not so simple, and has been the subject of
201 much discussion in statistical journals, under the heading of sensitivity
202 analysis [38], but also in chemometrics [10, 39].
- 203 • If the answer is yes, the second step is to consider the controllability
204 of the influencing variables responsible for the domain change. For
205 example, if the domain change is due to the spectrometer's temperature
206 variation, placing it in a temperature-controlled environment may solve
207 the problem.
- 208 • If this control is impossible or not desired, the third step is to ask
209 whether the domain is known when the model is applied. For exam-
210 ple, is the spectrometer temperature measured simultaneously with the
211 spectrum?
- 212 • If the domain is known while applying the model, three solutions exist:
 - 213 – *a priori correction*: The measured spectrum is modified using the
214 information about \mathcal{D} . For example, a spectrum measured on a
215 slave spectrometer can be geometrically modified to match that
216 which should have been measured on the master spectrometer.
 - 217 – *model correction*: The model is modified according to the infor-
218 mation about \mathcal{D} . For example, the model can be chosen according
219 to the variety of the fruit measured.
 - 220 – *a posteriori correction*: The model output is modified according
221 to the information about \mathcal{D} . For example, the estimated sugar
222 content of a fruit is modified using the law relating the estimation
223 error to its temperature, which will have been learned experimen-
224 tally.
- 225 • If the domain is unknown when the model is applied, a robust calibra-
226 tion must be performed, for example by calibrating a model on a data
227 set comprising a diversity of fruit origins.

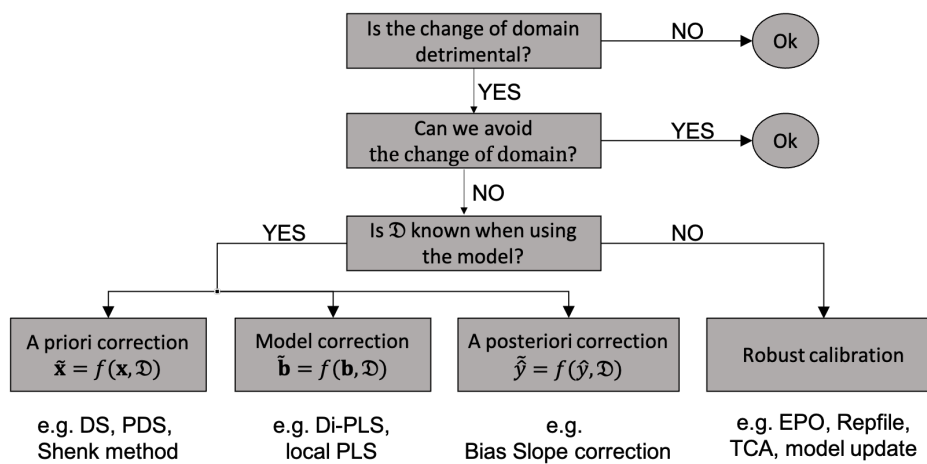


Figure 4: A strategy for managing the model robustness. \mathcal{D} is the current domain. Cited methods are described in section 3

228 3. Review on calibration transfer / domain adaptation methods

229 Table 1 summarizes the categories of methods, indicating the types of
 230 samples required and whether or not they produce a model that works in
 231 both domains. Categories are based on Figure 4. We will describe here each
 232 category's characteristics and generalize the methods' trends.

233 3.1. a priori correction

234 Linear standardization methods have been the most classical and predom-
 235 inant solutions for calibration transfer. In the context of transfer between
 236 instruments, where they are called optical standardization, the objective is
 237 to standardize the signal response between source and target instruments.
 238 These methods can be used in two different ways: forward or backward [71].
 239 Forward standardization transforms the source spectra so that they resemble
 240 the target spectra; this transformation must then be applied to all the spectra
 241 in the calibration database, and a new model recalculated, which can then be
 242 applied directly to the target spectra. Backward standardization transforms
 243 target spectra so that they resemble source spectra; this transformation must
 244 then be applied to each measured spectrum before the model calibrated in
 245 the source domain is applied to it.

246 The two most predominant techniques in this category are called Direct
 247 Standardization (DS) and Piecewise Direct Standardization (PDS) [40, 41].

Table 1: Compilation of methods for calibration transfer/domain adaptation. Std: (1) Standard samples, (2) Standard-free with reference values, (3) Standard-free without reference values. \mathcal{D} : (0) Domain independent model, (1) Domain dependent model. Type of sift managed: (P) prior, (M) marginal, (C) conditional.

Category (Cf Fig. 4)	Method name	Std	\mathcal{D}	Shift	Ref.	
A priori correction	Shenk's patented method	1	1	M	[14, 15]	
	(Piecewise) Direct Standardization ((P)DS)	1	1	M	[40, 41]	
	Spectral Space Transformation (SST)	1	1	M	[42]	
	Finite Impulse Response (FIR) Filter	3	1	M	[43]	
	Principal Components Canonical Correlation Analysis (PC-CCA) ¹	1	1	M	[44]	
	Calibration transfer based on the weight matrix of PLS (CTWM)	1,2	1	M,C,P	[45]	
	Calibration transfer via filter learning	1	1	M	[46]	
	(Kernel) Joint PCA	1,2	1	M	[47]	
	A posteriori correction	Slope and Bias Correction (SBC)	2	1	M,C,P	[48]
		Calibration transfer with affine invariance	3	1	M	[49]
Model correction (Domain Adaptation)	Transfer Component Analysis (TCA)	3	0	M,C,P*	[50, 51]	
	Trimmed Scores Regression (TSR)	1	1	M	[52, 53]	
	Maximum Likelihood PCA (MLPCA)	1	1	M	[52, 53, 54]	
	Joint and Unique Multiblock Analysis (JUMBA)	1	1	M	[55]	
	Domain Invariant Partial Least Squares (di-PLS)	3	1	M,C,P*	[56, 57, 58]	
	Graph-based calibration transfer (GCT-PLS) ²	1	1	M,C,P*	[59]	
	Domain Adaptive Partial Least Squares (da-PLS)	3	1	M,C,P*	[60]	
Robust Modeling (Variable Selection)	Correlation-Analysis-Based Wavelength Selection (CAWS)	1	0	M	[61]	
	Screening Wavelengths (SWCSS)	1,3	0	M	[62]	
	Domain-invariant Covariate Selection (di-CovSel)	3	0	M	[63]	
Robust Modeling (Concatenation)	Model Update	1,2	0,1	M,C,P	[64, 65]	
	RepFile	1	0	M	[66]	
	Joint Y PLS (JYPLS)	1	0,1 ³	M,C,P	[53, 67]	
Robust Modeling (Orthogonalization)	External Parameter Orthogonalization (EPO)	1	0	M	[23]	
	Transfer by Orthogonal Projection (TOP)	1,3	0	M	[24]	
	Dynamic Orthogonal Projection (DOP)	2	0	M,C,P	[25]	
	Tikhonov Regularization (TR)	1,2	0	M,C,P	[65]	
	Orthogonal Space Regression (OSR)	2	1	M,C	[68]	
	Dual-Domain Calibration Transfer Orthogonal Projection (DDTOP)	1,3	0	M	[69]	
Unsupervised Dynamic Orthogonal Projection (uDOP)	3	0	M	[70]		

* These methods can accommodate labeled data from the target domain and are thus capable of accounting for prior and conditional shifts. ¹ Source and target domain can have different numbers of variables. ² Can cope with CT problems where standards and calibration samples have different physicochemical properties. ³ Note that JYPLS-inv is domain-dependent while JYPLS-noninv is domain-independent.

248 These two techniques depend on the availability of standard samples which
249 are used to fit a linear model to transform the target instrument signal into
250 the source instrument signal. The difference between DS and PDS is the cor-
251 respondence that is assumed between the wavelengths of each instrument. In
252 DS, the transformation to a wavelength of one instrument is considered a lin-
253 ear combination of all the wavelengths of the other instrument. In the PDS,
254 the transformation to a wavelength of one instrument corresponds to a linear
255 combination of wavelengths within a small window around the wavelength of
256 the other instrument (i.e. piecewise).

257 Another predominant method in this category is Spectral Space Trans-
258 formation (SST) [42]. In SST, the same type of standardization as in DS
259 and PDS takes place but is fit under a different approach. SST uses the for-
260 mulation of Beer Lambert's law on each instrument and couples it with the
261 Singular Value Decomposition (SVD) of a horizontally concatenated matrix
262 of standard samples. This allows to standardize the spectra via the loadings
263 of each domain [42].

264 3.2. *a posteriori* correction

265 The second category of calibration transfer methods involves the transfor-
266 mation or adjustment of predictions when using a source calibration model
267 directly on target spectra. The most predominant classical technique in this
268 category is Slope and Bias Correction (SBC) where, as the name indicates,
269 the slope and bias of the predictions are corrected [48]. The slope and bias
270 terms are estimated with least squares linear regression using samples mea-
271 sured in the target domain whose reference values are available. One of the
272 major advantages of SBC is that it can cope with either prior, covariate and
273 conditional shifts without the need for transfer standards. However, its main
274 drawback is that multivariate outlier statistics like e.g., Hotelling's T^2 , Q -
275 Residuals [72], or cumulative sum statistics [73] are not available to assess
276 if predictions for new samples (in the target domain) are reliable. Further-
277 more, SBC requires spectra and corresponding reference measurements (i.e.,
278 labeled samples) from the target domain to learn the correction function.
279 Calibration transfer with affine invariance, on the other hand, does not use
280 standard samples nor target reference values but obtains scores and predic-
281 tions for the source samples and the target samples based on a separate set
282 of samples in the target domain. Then, regression coefficients between the
283 scores and the predicted values are calculated for each set. To align these
284 regression coefficients, angle and bias terms are calculated. These are then

285 used to correct the predictions for the target instrument [49]. On the down-
286 side, the method can handle only covariate shift problems and might thus
287 have limited applicability in practice.

288 3.3. Model correction: Variable selection

289 As a step in the pipeline of model building, variable selection started
290 being considered to achieve calibration transfer. While variable selection
291 has been greatly explored for model building, only three approaches were
292 found to tackle the difference between instruments. The first approach cor-
293 responds to wavelength selection based on correlation analysis (CAWS). In
294 CAWS, standard samples are used to calculate Pearson correlations for each
295 wavelength between the two instruments. The wavelengths with the highest
296 absolute correlation values are selected for further recalibration, so the re-
297 sulting model can be directly applied to spectral measurements in the target
298 instrument [61].

299 Another solution was recently proposed to relax the requirement for the
300 availability of standard samples. In this solution, a screening of wavelengths
301 is done by calculating the consistency and stability of the wavelength in each
302 instrument (SWCSS²). Such consistency is quantified by the similarity in
303 standard deviation and mean between the two sets of samples. As in CAWS,
304 the wavelength selection is followed by a recalibration of the model. A major
305 drawback of both approaches is that despite their capability to select wave-
306 lengths that are consistent across domains with respect to their (marginal)
307 distributions, it is not guaranteed that these variables will carry (sufficient)
308 Y -predictive information. To address this shortcoming, di-CovSel was re-
309 cently proposed as an extension to the original CovSel algorithm by Roger
310 *et al.* [74] aiming at identifying variables that are both consistent across
311 domains and predictive with respect to Y [63] without requiring transfer
312 standards nor labeled target domain samples. Conceptually, di-CovSel in-
313 volves a DA-inspired domain regularization approach first proposed in [56].
314 We note that all variable selection approaches discussed here aim at obtain-
315 ing robust models that are domain-independent, whereas they are capable of
316 handling covariate shifts only.

²The abbreviation comes from the title of the original publication [62]

317 3.4. Model correction: Domain adaptation methods

318 One of the first approaches within the framework of domain adaptation
319 was Transfer Component Analysis (TCA) which has recently been tested
320 for calibration transfer [51, 13]. In unsupervised TCA, a kernel matrix is
321 calculated to be indicative of the discrepancy of the samples within each in-
322 strument and between the two instruments. This optimization is a trade-off
323 between finding maximum variance and finding common variance between
324 the two sample sets [50]. The components can then be used for further cal-
325 ibration. A semi-supervised TCA uses the kernel matrix and the available
326 reference values to extend the optimization criterion to include the depen-
327 dence between the kernel matrix of spectral measurements and the kernel
328 matrix of reference values [50].

329 The main drawback of TCA is that when used with a linear kernel, which
330 for NDSS applications is often the optimal choice due to (approximately) lin-
331 ear relationships between the predictors and the response variable, it can only
332 account for differences in the means related to $P_S(X)$ and $P_T(X)$. When the
333 shift involves more complicated changes related to higher order distributional
334 properties like e.g., variance, a universal kernel (like the Gaussian kernel) is
335 required which comes at the risk of over-fitting the corresponding (non-linear)
336 calibration model – in particular for small calibration sets. To address this
337 problem, Nikzad-Langerodi *et al.* in [56] proposed a domain-regularization
338 approach for partial least squares regression that allows to model the response
339 Y using a linear model, while (implicitly) accounting for domain differences
340 in terms of mean and variance, i.e. di-PLS regression. Despite the flexibility
341 of the corresponding model in the sense that it can accommodate labeled
342 and/or unlabeled data from the target domain, and work with calibration
343 standards [59] or in standards-free settings, it assumes normal distribution
344 of the domains. In domain adaptative PLS (da-PLS), proposed later, a simi-
345 lar optimization problem that employs the (non-parametric) Hilbert-Schmidt
346 independence criterion (HSIC) between the domains, as a constraint, is for-
347 mulated [75], which does not make assumptions about the underlying distri-
348 butions. In terms of the classification presented in Table 1, apart from TCA,
349 all domain adaptation approaches proposed/employed so far (in the realm
350 of analytical chemistry) yield domain-dependent models, i.e., when used to
351 make predictions it must be known which domain the data originates from.
352 When reviewing existing literature we also noted that some approaches, while
353 not explicitly referring to DA, are related to this category of methods, e.g.,
354 trimmed scores regression (TSR) and maximum likelihood PCA (MLPCA)

355 proposed in [52, 54, 53]. The basic idea of these approaches is to model a
356 common latent structure underlying the source and target domains, and sub-
357 sequently use this structure to infer how the source domain spectra would
358 look like in the target domain. A similar (multi-block) approach was pro-
359 posed by Skotare *et al.* in [55], where the authors used On-PLS to extract
360 common latent variables across transfer standards measured on multiple NIR
361 spectrometers to stabilize the transfer between some source and target do-
362 main instrument. In a similar line of research, Mikulasek *et al.* recently
363 extended the original di-PLS method to accommodate Mango samples from
364 multiple domains and found that the corresponding model not only per-
365 formed better in the target domains but was also more robust when applied
366 in an out-of-domain context [58].

367 3.5. Robust modeling: Concatenation

368 Concatenation refers to the methods that merge datasets obtained from
369 the source and target domain to achieve an adaptation of the model that
370 takes into account the variability of both domains.

371 The easiest way to do this is simply to supplement the source calibra-
372 tion database with target spectra and corresponding reference values [76].
373 Another way consists in augmenting the database with a *repeatability file*,
374 which consists in the collection of differences in measured spectra between
375 two instruments or two conditions, i.e. $\delta\mathbf{x}$ from Figure 3, associated with
376 null Y values [66]. Two natural questions arise with concatenation-based
377 modeling approaches: Are the domains "compatible" at all and if so, what
378 emphasis should be put on the the source and target domain data when fit-
379 ting the model? The first question is not easy to answer (theoretically) and
380 is an area of active research. However, if the domains are not compatible,
381 so-called *negative transfer* might be the consequence, i.e., including source
382 domain data harms learning in, and generalization to, the target domain [77].
383 As to the weighting problem, Stork & Kowalski showed that the variance in
384 the predicted concentration on a test sample is related to its outlyingness
385 (a.k.a. *leverage*) with respect to the augmented calibration set [64], and that
386 the leverage is the only quantity affected by changing the sample weighting.
387 This allows us to derive an optimal weighting of (labeled) source and target
388 domain data that minimizes the prediction variance for a (unlabeled) test
389 set. One decade later Kalivas *et al.* re-addressed the weighting problem
390 and showed that it can be cast as a least-squares regression problem with a
391 two-norm penalty [78]. The corresponding regularization parameter controls

392 the weight that is assigned to the target domain samples. Concatenation ap-
393 proaches have also been disseminated for calibration transfer. Joint-Y PLS
394 is a technique that attempts to unify a calibration model across several in-
395 struments by concatenating the scores of two instruments and modifying the
396 regression step in PLSR to acquire a unique regression vector [67, 53]. This
397 method is suitable when standard samples are available, but also when sep-
398 arate sets with reference values are available for the different instruments.
399 In terms of dataset shift, concatenation-based techniques that involve la-
400 beled data from all domains are, in principle, capable of correcting for all
401 three types of shifts while those that only use unlabeled data from the target
402 domain(s) account for covariate shift only.

403 3.6. Robust modeling: Orthogonalization

404 Orthogonalization techniques have been developed to modify the space in
405 which the model is constructed, to make \mathbf{b} orthogonal to $\delta\mathbf{x}$ (Cf Figure 3). In
406 short, these techniques project the vector space spanned by \mathbf{X} , $Row(\mathbf{X})$ or-
407 thogonally to the vector subspace containing the $\delta\mathbf{x}$, and as \mathbf{b} lies in $Row(\mathbf{X})$,
408 it naturally becomes orthogonal to any $\delta\mathbf{x}$.

409 This mechanism was initially developed in External Parameter Orthog-
410 onalization (EPO), which has proved effective in making a Brix prediction
411 model insensitive to the temperature of the fruit [23]. Then, orthogonal-
412 ization started to be used as a solution to make calibration models robust
413 against the differences between source and target instruments, with Transfer
414 by Orthogonal Projection (TOP) [24].

415 With the availability of standard samples, the most direct solution for
416 transfer using orthogonalization is to estimate the interference information
417 from the difference between the two measurement conditions. A new calibra-
418 tion model is fitted with the orthogonalized signals. The resulting calibration
419 model is robust against the difference between the measurement conditions.

420 Other orthogonalization techniques estimate the interference matrix under
421 different criteria. One of the first methods that proposed orthogonaliza-
422 tion solutions for multiple instruments simultaneously is TOP in which the
423 interference matrix is based on the mean spectra for all instruments [24]. In
424 this way, orthogonalization allows handling the between-instrument variabil-
425 ity among different spectrometers.

426 Inspired by TOP, other methods such as Dual-Domain Calibration Trans-
427 fer (DDTOP) use orthogonalization by estimating the difference between

428 instruments from wavelet decomposition of the spectral signals in both in-
429 struments [69]. Because these methods are based on mean or structural dif-
430 ferences between the instruments, they can be performed without standard
431 samples.

432 Another solution from process control monitoring suitable for calibra-
433 tion transfer is Dynamic Orthogonal Projection (DOP), which estimates the
434 interference matrix from the spectral differences of samples with similar ref-
435 erence values. This method uses reference values in the source and the target
436 domain to create a set of virtual standard samples. Such virtual standard
437 samples are identified using a nearest neighbors approach which serves to
438 calculate the spectral signals of the target samples as if they were measured
439 on the source instrument. This allows us to obtain the set of target samples
440 and its virtual standards which are used to calculate the interference matrix
441 and proceed with orthogonalization [25]. This method has been applied to
442 different types of corrections, such as process control conditions, temperature
443 drifts, and calibration transfer [25, 51, 79].

444 An unsupervised version of DOP has recently been published [70]. It pro-
445 poses the creation of virtual standards based solely on two spectra matrices of
446 the source and target domains. An SVD provides the scores and loadings of
447 these two matrices. A source version of the target spectra is then estimated
448 by projecting the target scores onto the source loadings. This method has
449 been successfully applied to the transfer of a PLS model predicting the acid
450 detergent fiber fraction of sugarcane between two portable spectrometers. It
451 was also used to establish a PLS model for predicting the dry matter value
452 of mangoes, robust over two varieties and two harvest seasons.

453 Orthogonalization approaches may be too aggressive in the sense that
454 they can remove domain differences at the expense of the NAS. It's worth
455 noting, however, that this risk can be easily detected since it translates into
456 model degradation, visible in calibration and cross-validation performances.
457 Some approaches though, such as the Tikhonov regularization approach by
458 Kalivas *et al.* [65], or the GLS approach by Wise *et al.* [22] relax the orthog-
459 onality constraint and try to balance better the trade-off between making \mathbf{b}
460 orthogonal to $\delta\mathbf{x}$ (Cf Figure 3) and preserving the NAS.

461 4. Discussion and open problems

462 This review shows that the causes of robustness problems in calibration
463 models can be formalized as dataset shifts, whereas covariate shift is the most

464 common one. It corresponds to cases where the measurement condition is
465 affected, for example when a spectrometer is changed, leading to a change in
466 the marginal distribution over the spectra. This is the easiest case and has
467 been addressed by historical correction methods developed by the chemomet-
468 rics community. Conditional shifts correspond to a change in the relationship
469 between the measured spectrum and the response to be predicted. They can
470 occur frequently in NDSS when calibrations are based on secondary cor-
471 relations. For example, the estimation of a fruit's sugar content by NIRS
472 spectroscopy is based on the spectral responses of all the compounds in the
473 fruit that are correlated with the sugar content. However, this internal fruit
474 correlation can change with the variety or the origin of the fruit. Finally,
475 prior shifts correspond to changes in the distribution of the response to be
476 predicted, which can occur when the NDSS needs to extrapolate beyond the
477 concentration range of the calibration samples.

478 Numerous methods have been developed to improve the robustness of
479 calibration models, both in the field of chemometrics and machine learning.
480 They differ according to whether or not they need paired samples (i.e., trans-
481 fer standards) or reference values from both domains, whether or not they
482 require knowledge of the domain when applying the model, and whether or
483 not they handle prior, conditional, and/or covariate shifts. A careful analysis
484 of the situation is therefore imperative before implementing a CMM.

485 As far as covariate shift is concerned, domain adaptation techniques can
486 be considered the most promising methodology since they provably minimize
487 an upper bound on the generalization error in the target domain, and neither
488 require transfer standards nor labeled samples from the target domain [57].
489 Methods dedicated to prior and conditional shifts are less numerous and
490 require the use of labeled target domain samples. A largely open challenge
491 here is the assessment of the "compatibility" of the domains to avoid negative
492 transfer.

493 Almost all published methods are aimed at the simple case of a source
494 and a target domain. In reality, however, the user is often confronted with
495 the case of multiple domain changes. For example, a change of spectrometer
496 will often occur simultaneously with a change of harvest year. This raises the
497 question of how to combine several transfers. Some methods, such as those
498 using data augmentation or orthogonalization, can deal with this problem
499 globally, while others, such as those based on model correction, will have to be
500 used successively for each domain change. New methodological developments
501 are expected in this field.

502 A similar problem arises for instrument networks. What is the best strat-
503 egy to employ in this case? Is it possible to have a single model for all
504 instruments? This would offer the advantage of not having to manage sev-
505 eral models. Should we decide that there is only one master instrument?
506 This multi-domain problem, as yet little treated as such in the literature,
507 should be the subject of specific research, that will address the problem of
508 multiple and combined shifts. With the trend towards the availability of
509 larger and open-source data, we also foresee further developments related
510 to the application of deep learning to address the robustness problem. The
511 first promising results in this direction comprise, e.g., the recently proposed
512 Deep Tutti Frutti model by Passos & Mishra capable of fruit-independent,
513 NIR-based dry matter prediction [80].

514 The methods described in this review show that progress has been made
515 in solving the problem of calibration transfer in chemometrics. Nevertheless,
516 industrial practitioners are often still unaware of methodological advances
517 that offer the efficient solutions required to transfer calibration models. To
518 date, most commercial chemometrics software only offers solutions based on
519 linear normalization and orthogonalization, with none of the many methods
520 identified in this review available. Efforts must therefore be made to develop
521 and enrich open software, such as ChemFlow [81], FACT [82] or Saisir [83].

522 This study aims to explain the reasons for robustness problems and to list
523 the main methods for improving it. In other words, it aims to answer the why
524 and the how. However, the question of when to carry out a CMM is also prac-
525 tically very relevant. The most reliable solution is still to acquire reference
526 measurements periodically and implement a conventional, threshold-based
527 process diagnostic tool. Little methodological research has been carried out
528 to provide more elaborate tools. When the products being measured are
529 stable, such as flour, sealed cups containing a range of samples can be mea-
530 sured periodically. The predicted values for these cups can be used to replace
531 comparisons with reference analyses [84]. The evolution of measured spectra
532 can be examined using multivariate techniques [85]. In short, new spectra
533 are projected into the space of a PCA or a PLS, where the quality of their
534 projection is evaluated with the so-called Hotelling statistic and residuals (T^2
535 and Q). A new method has recently been proposed for estimating a regres-
536 sion model's prediction bias and covariate shift in real-time, without using
537 target reference values [37]. A set of current spectra, acquired during process
538 monitoring is used to re-estimate the intercept of the calibration model. The
539 covariance structure of the current spectra is also calculated as loadings. The

540 calibration set is then projected onto these loadings, giving scores that are
541 compared with the scores obtained directly by a PCA on the calibration set.
542 Despite this original contribution, the issue of unsupervised diagnosis of the
543 calibration models remains largely open.

544 This review shows that CMM is a common concern in both machine
545 learning and chemometrics, but conceptual approaches differ. Chemomet-
546 ric approaches are essentially based on signal processing and linear algebra,
547 while machine learning approaches are based on the notions of conditional
548 distribution and probability. It's a safe bet that the combination of these
549 two paradigms will pave the way for new methods capable of resolving the
550 questions still open to us.

551 This review does not cover all methodological developments on CMM
552 in chemometrics since this would be prohibitive in the current format. We
553 thus want to point the interested reader to the excellent literature surveys by
554 Workman [86] and Mishra *et al.* [87] on calibration transfer and standards-
555 free calibration transfer, respectively, and to Feudale *et al.* [88] for a general
556 primer on the more general CMM problem.

557 Acknowledgements

558 The authors acknowledge the SensorFINT project (COST Action No.
559 CA19145). Valeria Fonseca Diaz and Ramin Nikzad-Langerodi acknowledge
560 funding from the Federal Ministry for Climate Action, Environment, Energy,
561 Mobility, Innovation and Technology (BMK), the Federal Ministry for Dig-
562 ital and Economic Affairs (BMDW), and the Province of Upper Austria in
563 the frame of the COMET - Competence Centers for Excellent Technologies
564 programme managed by the Austrian Research Promotion Agency FFG, and
565 the FFG project Interpretable and Interactive Transfer Learning in Process
566 Analytical Technology (Grant No. 883856).

567 References

- 568 [1] H. Huang, H. Yu, H. Xu, Y. Ying, Near infrared spectroscopy for on/in-
569 line monitoring of quality in foods and beverages: A review, *Journal of*
570 *food engineering* 87 (3) (2008) 303–313.
- 571 [2] M. H. Ahmad, A. Sahar, B. Hitzmann, Fluorescence spectroscopy for the
572 monitoring of food processes, *Measurement, modeling and automation*
573 *in advanced food processing* (2017) 121–151.

- 574 [3] Y. Sun, H. Tang, X. Zou, G. Meng, N. Wu, Raman spectroscopy for food
575 quality assurance and safety monitoring: A review, *Current Opinion in*
576 *Food Science* 47 (2022) 100910.
- 577 [4] V. Baeten, H. Rogez, J. Fernández Pierna, P. Vermeulen, P. Dardenne,
578 Vibrational spectroscopy methods for the rapid control of agro-food
579 products, *Handbook of food analysis 2* (2015) 591–614.
- 580 [5] V. Vapnik, *The nature of statistical learning theory*, Springer science &
581 business media, 2013.
- 582 [6] R. Nikzad-Langerodi, E. Andries, A chemometrician’s guide to
583 transfer learning, *Journal of Chemometrics* 35 (11) (2021) 1–23.
584 doi:10.1002/cem.3373.
- 585 [7] H. Swierenga, A. De Weijer, R. Van Wijk, L. Buydens, Strategy for
586 constructing robust multivariate calibration models, *Chemometrics and*
587 *Intelligent laboratory systems* 49 (1) (1999) 1–17.
- 588 [8] M. Mulholland, Ruggedness testing in analytical chemistry, *TrAC*
589 *Trends in Analytical Chemistry* 7 (10) (1988) 383–389.
- 590 [9] S. Pharma, *Stp pharma pratiques methodology*, Rapport de Commission
591 2 (4) (1992) 205–206.
- 592 [10] M. Zeaiter, J.-M. Roger, V. Bellon-Maurel, D. Rutledge, Robustness of
593 models developed by multivariate calibration. part i: The assessment of
594 robustness, *TrAC Trends in Analytical Chemistry* 23 (2) (2004) 157–170.
- 595 [11] F. Chauchard, J. Roger, V. Bellon-Maurel, Correction of the temper-
596 ature effect on near infrared calibration—application to soluble solid
597 content prediction, *Journal of near infrared spectroscopy* 12 (3) (2004)
598 199–205.
- 599 [12] J. M. Roger, F. Chauchard, P. Williams, Removing the block effects in
600 calibration by means of dynamic orthogonal projection. Application to
601 the year effect correction for wheat protein prediction, *Journal of Near*
602 *Infrared Spectroscopy* 16 (3) (2008) 311–315. doi:10.1255/jnirs.793.

- 603 [13] N. I. Pedge, M. Papillaud, J.-M. Roger, Investigation of long-term stabil-
604 ity of a transmission raman calibration model using orthogonal projec-
605 tion methods, *Chemometrics and Intelligent Laboratory Systems* (2024)
606 105115.
- 607 [14] J. Shenk, M. Westerhaus, W. Templeton Jr, Calibration transfer be-
608 tween near infrared reflectance spectrophotometers 1, *Crop science*
609 25 (1) (1985) 159–161.
- 610 [15] J. Shenk, M. Westerhaus, U.s. patent 48666644 (1989).
- 611 [16] T. Fearn, C. Eddison, R. Withey, I. A. Cowe, A method for wave-
612 length standardisation in filter instruments, *Journal of Near Infrared*
613 *Spectroscopy* 4 (1) (1996) 111–118.
- 614 [17] Y. Wang, D. J. Veltkamp, B. R. Kowalski, Multivariate instrument stan-
615 dardization, *Analytical chemistry* 63 (23) (1991) 2750–2756.
- 616 [18] E. Bouveresse, D. Massart, Improvement of the piecewise direct stan-
617 dardisation procedure for the transfer of nir spectra for multivariate cal-
618 ibration, *Chemometrics and intelligent laboratory systems* 32 (2) (1996)
619 201–213.
- 620 [19] Y. Xie, P. K. Hopke, Calibration transfer as a data reconstruction prob-
621 lem, *Analytica chimica acta* 384 (2) (1999) 193–205.
- 622 [20] Y. Xie, P. K. Hopke, P. Paatero, Erratum to calibration transfer as a
623 data reconstruction problem-(*analytica chimica acta* 384 (1999) 193-
624 205), *Analytica Chimica Acta-Including Cumulative Indexes* 392 (2)
625 (1999) 347–348.
- 626 [21] C. E. Anderson, J. H. Kalivas, Fundamentals of calibration transfer
627 through procrustes analysis, *Applied Spectroscopy* 53 (10) (1999) 1268–
628 1276.
- 629 [22] B. M. Wise, H. Martens, M. Høy, R. Bro, P. B. Brockhoff, Calibration
630 transfer by generalized least squares, in: *Proceedings of the Seventh*
631 *Scandinavian Symposium on chemometrics (SSC7)*, 2001, pp. 19–23.

- 632 [23] J. M. Roger, F. Chauchard, V. Bellon-Maurel, EPO-PLS external
633 parameter orthogonalisation of PLS application to temperature-
634 independent measurement of sugar content of intact fruits, *Chemometrics and Intelligent Laboratory Systems* 66 (2) (2003) 191–204.
635 doi:10.1016/S0169-7439(03)00051-0.
636
- 637 [24] A. Andrew, T. Fearn, Transfer by orthogonal projection: Making near-
638 infrared calibrations robust to between-instrument variation, *Chemometrics and Intelligent Laboratory Systems* 72 (1) (2004) 51–56.
639 doi:10.1016/j.chemolab.2004.02.004.
640
- 641 [25] M. Zeaiter, J. M. Roger, V. Bellon-Maurel, Dynamic orthogonal projec-
642 tion. A new method to maintain the on-line robustness of multivariate
643 calibrations. Application to NIR-based monitoring of wine fermenta-
644 tions, *Chemometrics and Intelligent Laboratory Systems* 80 (2) (2006)
645 227–235. doi:10.1016/j.chemolab.2005.06.011.
- 646 [26] Y. Zhu, T. Fearn, D. Samuel, A. Dhar, O. Hameed, S. G. Bown, L. B.
647 Lovat, Error removal by orthogonal subtraction (eros): a customised pre-
648 treatment for spectroscopic data, *Journal of Chemometrics: A Journal*
649 *of the Chemometrics Society* 22 (2) (2008) 130–134.
- 650 [27] A. Lorber, K. Faber, B. R. Kowalski, Net analyte signal calculation in
651 multivariate calibration, *Analytical Chemistry* 69 (8) (1997) 1620–1626.
- 652 [28] J. J. Heckman, The common structure of statistical models of trun-
653 cation, sample selection and limited dependent variables and a simple
654 estimator for such models, in: *Annals of economic and social measure-*
655 *ment*, volume 5, number 4, NBER, 1976, pp. 475–492.
- 656 [29] B. Zadrozny, Learning and evaluating classifiers under sample selection
657 bias, in: *Proceedings of the twenty-first international conference on Ma-*
658 *chine learning*, 2004, p. 114.
- 659 [30] S. J. Pan, Q. Yang, A Survey on Transfer Learning, *IEEE Transac-*
660 *tions on Knowledge and Data Engineering* 22 (10) (2010) 781–789.
661 doi:10.1007/978-981-15-5971-6_83.
- 662 [31] J. Jiang, A literature survey on domain adaptation of statistical classi-
663 fiers, URL: <http://sifaka.cs.uiuc.edu/jiang4/domainadaptation/survey>
664 3 (1-12) (2008) 3.

- 665 [32] S. Ben-David, J. Blitzer, K. Crammer, F. Pereira, Analysis of representa-
666 tions for domain adaptation, *Advances in neural information processing*
667 *systems* 19 (2006).
- 668 [33] S. Ben-David, J. Blitzer, K. Crammer, A. Kulesza, F. Pereira, J. W.
669 Vaughan, A theory of learning from different domains, *Machine Learning*
670 79 (1-2) (2010) 151–175. doi:10.1007/s10994-009-5152-4.
- 671 [34] F. Zhuang, Z. Qi, K. Duan, D. Xi, Y. Zhu, H. Zhu, H. Xiong, Q. He,
672 A comprehensive survey on transfer learning, *Proceedings of the IEEE*
673 109 (1) (2020) 43–76.
- 674 [35] J.-M. Roger, Développements chimiométriques pour améliorer la ro-
675 bustesse des mesures spectrométriques appliquées aux agro-procédés,
676 Ph.D. thesis, Université Montpellier II-Sciences et Techniques du
677 Languedoc (2005).
- 678 [36] V. Fonseca Diaz, A. van Nuenen, J. DÃaz Olivares, et al., A com-
679 plete scheme for unsupervised multivariate calibration monitoring, in:
680 NIR2023, Date: 2023/08/20-2023/08/24, Location: Innsbruck, Austria,
681 Austrian Society of Analytical Chemistry (ASAC), 2023, pp. 187–187.
- 682 [37] V. F. Diaz, B. Aernouts, E. Bobelyn, B. Nicolai, W. Saeys, Monitoring
683 of multivariate calibration models in the absence of new reference values:
684 the regression case, *Chemometrics and Intelligent Laboratory Systems*
685 240 (2023) 104884.
- 686 [38] S. Iyengar, J. Greenhouse, Sensitivity analysis and diagnostics, *Hand-*
687 *book of research synthesis and meta-analysis* (2009) 417–433.
- 688 [39] Y. Vander Heyden, A. Nijhuis, J. Smeyers-Verbeke, B. Vandeginste,
689 D. Massart, Guidance for robustness/ruggedness tests in method valida-
690 tion, *Journal of pharmaceutical and biomedical analysis* 24 (5-6) (2001)
691 723–753.
- 692 [40] Y. Wang, D. J. Veltkamp, B. R. Kowalski, Multivariate Instru-
693 ment Standardization, *Analytical Chemistry* 63 (23) (1991) 2750–2756.
694 doi:10.1021/ac00023a016.
- 695 [41] E. Bouveresse, D. L. Massart, Improvement of the piecewise direct stan-
696 dardisation procedure for the transfer of NIR spectra for multivariate

- 697 calibration, *Chemometrics and Intelligent Laboratory Systems* 32 (2)
698 (1996) 201–213. doi:10.1016/0169-7439(95)00074-7.
- 699 [42] W. Du, Z. P. Chen, L. J. Zhong, S. X. Wang, R. Q. Yu, A. Nordon,
700 D. Littlejohn, M. Holden, Maintaining the predictive abilities of multi-
701 variate calibration models by spectral space transformation, *Analytica*
702 *Chimica Acta* 690 (1) (2011) 64–70. doi:10.1016/j.aca.2011.02.014.
703 URL <http://dx.doi.org/10.1016/j.aca.2011.02.014>
- 704 [43] H. Tan, S. T. Sum, S. D. Brown, Improvement of a standard-free method
705 for near-infrared calibration transfer, *Applied spectroscopy* 56 (8) (2002)
706 1098–1106.
- 707 [44] X. Fan, H. Lu, Z. Zhang, Direct calibration transfer to prin-
708 cipal components via canonical correlation analysis, *Chemomet-*
709 *rics and Intelligent Laboratory Systems* 181 (June) (2018) 21–28.
710 doi:10.1016/j.chemolab.2018.08.006.
711 URL <https://doi.org/10.1016/j.chemolab.2018.08.006>
- 712 [45] F. Zhang, R. Zhang, J. Ge, W. Chen, W. Yang, Y. Du, Calibration
713 transfer based on the weight matrix (CTWM) of PLS for near infrared
714 (NIR) spectral analysis, *Analytical Methods* 10 (18) (2018) 2169–2179.
715 doi:10.1039/c8ay00248g.
- 716 [46] Z. Xie, X. Chen, J.-M. Roger, S. Ali, G. Huang, W. Shi, Calibration
717 transfer via filter learning, *Analytica Chimica Acta* (2024) 342404.
- 718 [47] P. Shan, Y. Zhao, Q. Wang, Y. Ying, S. Peng, Principal compo-
719 nent analysis or kernel principal component analysis based joint spec-
720 tral subspace method for calibration transfer, *Spectrochimica Acta -*
721 *Part A: Molecular and Biomolecular Spectroscopy* 227 (2020) 117653.
722 doi:10.1016/j.saa.2019.117653.
723 URL <https://doi.org/10.1016/j.saa.2019.117653>
- 724 [48] T. Fearn, Standardisation and calibration transfer for near infrared in-
725 struments: A review, *Journal of Near Infrared Spectroscopy* 9 (4) (2001)
726 229–244. doi:10.1255/jnirs.309.
- 727 [49] Y. Zhao, Z. Zhao, P. Shan, S. Peng, J. Yu, S. Gao, Calibration transfer
728 based on affine invariance for NIR without transfer standards, *Molecules*
729 24 (9) (2019). doi:10.3390/molecules24091802.

- 730 [50] S. J. Pan, I. W. Tsang, J. T. Kwok, Q. Yang, Domain adaptation via
731 transfer component analysis, *IEEE Transactions on Neural Networks*
732 22 (2) (2011) 199–210. doi:10.1109/TNN.2010.2091281.
- 733 [51] P. Mishra, J. M. Roger, D. N. Rutledge, E. Woltering, Two standard-free
734 approaches to correct for external influences on near-infrared spectra to
735 make models widely applicable, *Postharvest Biology and Technology* 170
736 (2020) 111326. doi:10.1016/j.postharvbio.2020.111326.
737 URL <https://doi.org/10.1016/j.postharvbio.2020.111326>
- 738 [52] F. Arteaga, A. Ferrer, Dealing with missing data in MSPC: Several
739 methods, different interpretations, some examples, *Journal of Chemometrics* 16 (8-10) (2002) 408–418. doi:10.1002/cem.750.
- 741 [53] A. Folch-Fortuny, R. Vitale, O. E. de Noord, A. Ferrer, Calibration
742 transfer between NIR spectrometers: New proposals and a
743 comparative study, *Journal of Chemometrics* 31 (3) (2017) 1–11.
744 doi:10.1002/cem.2874.
- 745 [54] F. Arteaga, A. Ferrer, Framework for regression-based missing data im-
746 putation methods in on-line MSPC, *Journal of Chemometrics* 19 (8)
747 (2005) 439–447. doi:10.1002/cem.946.
- 748 [55] T. Skotare, D. Nilsson, S. Xiong, P. Geladi, J. Trygg, Joint and unique
749 multiblock analysis for integration and calibration transfer of nir instru-
750 ments, *Analytical chemistry* 91 (5) (2019) 3516–3524.
- 751 [56] R. Nikzad-Langerodi, W. Zellinger, E. Lughofer, S. Saminger-Platz,
752 Domain-Invariant Partial-Least-Squares Regression, *Analytical Chem-*
753 *istry* 90 (11) (2018) 6693–6701. doi:10.1021/acs.analchem.8b00498.
- 754 [57] R. Nikzad-Langerodi, W. Zellinger, S. Saminger-Platz, B. A. Moser, Do-
755 main adaptation for regression under Beer–Lambert’s law, *Knowledge-*
756 *Based Systems* 210 (2020) 106447. doi:10.1016/j.knosys.2020.106447.
757 URL <https://doi.org/10.1016/j.knosys.2020.106447>
- 758 [58] B. Mikulasek, V. Fonseca Diaz, D. Gabauer, C. Herwig, R. Nikzad-
759 Langerodi, Partial least squares regression with multiple domains, *Jour-*
760 *nal of Chemometrics* 37 (5) (2023) e3477.

- 761 [59] R. Nikzad-Langerodi, F. Sobieczky, Graph-based calibration trans-
762 fer, *Journal of Chemometrics* 35 (4) (2021) 1–12. arXiv:2006.00089,
763 doi:10.1002/cem.3319.
- 764 [60] G. Huang, X. Chen, L. Li, X. Chen, L. Yuan, W. Shi,
765 Domain adaptive partial least squares regression, *Chemometrics*
766 *and Intelligent Laboratory Systems* 201 (March) (2020) 103986.
767 doi:10.1016/j.chemolab.2020.103986.
768 URL <https://doi.org/10.1016/j.chemolab.2020.103986>
- 769 [61] Z. Xu, S. Fan, W. Cheng, J. Liu, P. Zhang, Y. Yang, C. Xu,
770 B. Liu, J. Liu, Q. Wang, Y. Wu, A correlation-analysis-based wave-
771 length selection method for calibration transfer, *Spectrochimica Acta -*
772 *Part A: Molecular and Biomolecular Spectroscopy* 230 (2020) 118053.
773 doi:10.1016/j.saa.2020.118053.
774 URL <https://doi.org/10.1016/j.saa.2020.118053>
- 775 [62] L. Ni, M. Han, S. Luan, L. Zhang, Screening wavelengths with consis-
776 tent and stable signals to realize calibration model transfer of near in-
777 frared spectra, *Spectrochimica Acta - Part A: Molecular and Biomolec-*
778 *ular Spectroscopy* 206 (2019) 350–358. doi:10.1016/j.saa.2018.08.027.
779 URL <https://doi.org/10.1016/j.saa.2018.08.027>
- 780 [63] V. F. Diaz, P. Mishra, J.-M. Roger, W. Saeys, Domain invariant co-
781 variate selection (Di-CovSel) for selecting generalized features across
782 domains, *Chemometrics and Intelligent Laboratory Systems* 222 (Jan-
783 uary) (2022) 104499. doi:10.1016/j.chemolab.2022.104499.
784 URL <https://doi.org/10.1016/j.chemolab.2022.104499>
- 785 [64] C. L. Stork, B. R. Kowalski, Weighting schemes for updating regression
786 models—a theoretical approach, *Chemometrics and Intelligent Labora-*
787 *tory Systems* 48 (2) (1999) 151–166.
- 788 [65] J. H. Kalivas, G. G. Siano, E. Andries, H. C. Goicoechea, Calibration
789 maintenance and transfer using tikhonov regularization approaches, *Ap-*
790 *plied Spectroscopy* 63 (7) (2009) 800–809.
- 791 [66] J. S. Shenk, M. O. Westerhaus, New standardization and calibration
792 procedures for nirs analytical systems, *Crop Science* 31 (6) (1991) 1694–
793 1696.

- 794 [67] S. García Muñoz, J. F. MacGregor, T. Kourti, Product transfer between
795 sites using Joint-Y PLS, *Chemometrics and Intelligent Laboratory Sys-*
796 *tems* 79 (1-2) (2005) 101–114. doi:10.1016/j.chemolab.2005.04.009.
- 797 [68] Z. Lin, B. Xu, Y. Li, X. Shi, Y. Qiao, Application of orthogonal space
798 regression to calibration transfer without standards, *Journal of Chemo-*
799 *metrics* 27 (11) (2013) 406–413. doi:10.1002/cem.2536.
- 800 [69] D. V. Poerio, S. D. Brown, Dual-Domain Calibration Transfer Using
801 Orthogonal Projection, *Applied Spectroscopy* 72 (3) (2018) 378–391.
802 doi:10.1177/0003702817724164.
- 803 [70] V. Fonseca Diaz, J.-M. Roger, W. Saeys, Unsupervised dynamic
804 orthogonal projection. an efficient approach to calibration transfer
805 without standard samples, *Analytica Chimica Acta* 1225 (2022) 340154.
806 doi:<https://doi.org/10.1016/j.aca.2022.340154>.
807 URL <https://www.sciencedirect.com/science/article/pii/S0003267022007255>
- 808 [71] B. H. Campbell, When to use chemometric near infrared transfer steps,
809 *Near Infrared Spectroscopy: The Future Waves* (1996) 218–220.
- 810 [72] R. J. Rainer, M. Mayr, J. Himmelbauer, R. Nikzad-Langerodi, Opening
811 the black-box of neighbor embeddings with hotelling’s t2 statistic and q-
812 residuals, *Chemometrics and Intelligent Laboratory Systems* 238 (2023)
813 104840.
- 814 [73] R. Nikzad-Langerodi, E. Lughofer, C. Cernuda, T. Reischer, W. Kant-
815 ner, M. Pawliczek, M. Brandstetter, Calibration model maintenance in
816 melamine resin production: Integrating drift detection, smart sample
817 selection and model adaptation, *Analytica chimica acta* 1013 (2018) 1–
818 12.
- 819 [74] J. Roger, B. Palagos, D. Bertrand, E. Fernandez-Ahumada, Covsel: vari-
820 able selection for highly multivariate and multi-response calibration: ap-
821 plication to ir spectroscopy, *Chemometrics and Intelligent Laboratory*
822 *Systems* 106 (2) (2011) 216–223.
- 823 [75] J. Huang, S. Romero-Torres, M. Moshgbar, Practical Considerations
824 in Data Pre-treatment for NIR and Raman Spectroscopy, *American*
825 *Pharmaceutical Review* (2010).
826 URL <http://www.americanpharmaceuticalreview.com/Featured-Articles/116330-Prac>

- 827 [76] P. Gujral, M. Amrhein, D. Bonvin, Drift correction in multivariate cali-
828 bration models using on-line reference measurements, *Analytica Chimica*
829 *Acta* 642 (1-2) (2009) 27–36. doi:10.1016/j.aca.2008.12.044.
- 830 [77] Z. Wang, Z. Dai, B. Póczos, J. Carbonell, Characterizing and avoid-
831 ing negative transfer, in: *Proceedings of the IEEE/CVF conference on*
832 *computer vision and pattern recognition*, 2019, pp. 11293–11302.
- 833 [78] J. H. Kalivas, G. G. Siano, E. Andries, H. C. Goicoechea,
834 Calibration maintenance and transfer using Tikhonov regulariza-
835 tion approaches, *Applied Spectroscopy* 63 (7) (2009) 800–809.
836 doi:10.1366/000370209788701206.
- 837 [79] M. Dabros, M. Amrhein, P. Gujral, U. Von Stockar, On-line recal-
838 ibration of spectral measurements using metabolite injections and dy-
839 namic orthogonal projection, *Applied Spectroscopy* 61 (5) (2007) 507–
840 513. doi:10.1366/000370207780807830.
- 841 [80] D. Passos, P. Mishra, Deep tutti frutti: Exploring cnn architectures
842 for dry matter prediction in fruit from multi-fruit near-infrared spectra,
843 *Chemometrics and Intelligent Laboratory Systems* 243 (2023) 105023.
- 844 [81] J.-C. Boulet, M. Brandolini-Bunlon, G. Chaix, B. Jaillais, E. Latrille,
845 M. Lesnoff, A. Mallet, S. M. Garcia, M. Metz, J.-M. Roger, et al., *Chem-*
846 *house: A research and development centre for chemometrics*, *NIR news*
847 32 (7-8) (2021) 36–38.
- 848 [82] S. D. Brown, The chemometrics revolution re-examined, *Journal of*
849 *Chemometrics* 31 (1) (2017) e2856.
- 850 [83] C. B. Cordella, D. Bertrand, Saisir: A new general chemometric toolbox,
851 *TrAC Trends in Analytical Chemistry* 54 (2014) 75–82.
- 852 [84] P. Dardenne, Calibration transfer in near infrared spectroscopy, *NIR*
853 *news* 13 (4) (2002) 3–7.
- 854 [85] B. De Ketelaere, M. Hubert, E. Schmitt, Overview of pca-based statis-
855 tical process-monitoring methods for time-dependent, high-dimensional
856 data, *Journal of Quality Technology* 47 (4) (2015) 318–335.

- 857 [86] J. J. Workman, A Review of Calibration Transfer Practices and Instru-
858 ment Differences in Spectroscopy, *Applied Spectroscopy* 72 (3) (2018)
859 340–365. doi:10.1177/0003702817736064.
- 860 [87] P. Mishra, R. Nikzad-Langerodi, F. Marini, J. M. Roger, A. Biancolillo,
861 D. N. Rutledge, S. Lohumi, Are standard sample measurements still
862 needed to transfer multivariate calibration models between near-infrared
863 spectrometers? the answer is not always, *TrAC Trends in Analytical*
864 *Chemistry* 143 (2021) 116331.
- 865 [88] R. N. Feudale, N. A. Woody, H. Tan, A. J. Myles, S. D. Brown, J. Ferré,
866 Transfer of multivariate calibration models: a review, *Chemometrics and*
867 *intelligent laboratory systems* 64 (2) (2002) 181–192.