

Selectivity-relaxed classical and inverse least squares calibration and selectivity measures with a unified selectivity coefficient

John H. Kalivas¹  | Joan Ferré²  | Alister J. Tencate¹

¹Department of Chemistry, Idaho State University, Pocatello, ID 83209, USA

²Department of Analytical Chemistry and Organic Chemistry, Universitat Rovira i Virgili, 43007 Tarragona, Spain

Correspondence

John H. Kalivas, Department of Chemistry, Idaho State University, Pocatello, ID 83209, USA.

Email: kalijohn@isu.edu

Joan Ferré, Department of Analytical Chemistry and Organic Chemistry, Universitat Rovira i Virgili, 43007 Tarragona, Spain.

Email: joan.ferre@urv.cat

Funding information

European Community's Framework program for Research and Innovation Horizon 2020 (2014-2020), Grant/Award Number: 637232; MPS Chemistry and the OCI Venture Fund, Grant/Award Number: CHE-1506417; National Science Foundation, Grant/Award Number: CHE-1111053

Two popular calibration strategies are classical least squares (CLS) and inverse least squares (ILS). Underlying CLS is that the net analyte signal used for quantitation is orthogonal to signal from other components (interferents). The CLS orthogonality avoids analyte prediction bias from modeled interferents. Although this orthogonality condition ensures full analyte selectivity, it may increase the mean squared error of prediction. Under certain circumstances, it can be beneficial to relax the CLS orthogonality requisite allowing a small interferent bias if, in return, there is a mean squared error of prediction reduction. The bias magnitude introduced by an interferent for a relaxed model depends on analyte and interferent concentrations in conjunction with analyte and interferent model sensitivities. Presented in this paper is relaxed CLS (*r*CLS) allowing flexibility in the CLS orthogonality constraints. While ILS models do not inherently maintain orthogonality, also presented is relaxed ILS. From development of *r*CLS, presented is a significant expansion of the univariate selectivity coefficient definition broadly used in analytical chemistry. The defined selectivity coefficient is applicable to univariate and multivariate CLS and ILS calibrations. As with the univariate selectivity coefficient, the multivariate expression characterizes the bias introduced in a particular sample prediction because of interferent concentrations relative to model sensitivities. Specifically, it answers the question of when can a prediction be made for a sample even though the analyte selectivity is poor? Also introduced are new component-wise selectivity and sensitivity measures. Trends in several *r*CLS figures of merit are characterized for a near infrared data set.

KEYWORDS

classical least squares, inverse least squares, net analyte signal, selectivity, selectivity coefficient, sensitivity

1 | INTRODUCTION

Multivariate calibration is a well-established step in many analytical determinations on the basis of nonselective instrumental measurements. Such measurements are typically spectra although responses from other first-order analytical instruments, or combinations of them, are also used. The property to be determined (measurand) is commonly analyte concentration, although

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2017 The Authors. *Journal of Chemometrics* published by John Wiley & Sons, Ltd.

indexes (eg, motor octane number in gasoline) or physical properties (eg, viscosity in polymers) are also modeled. All discussions henceforth are based on spectroscopic calibration with analyte concentration as the quantity being modeled, but the discussions are applicable to other techniques of instrumental analysis.

From a set of calibration samples with known instrumental responses and quantitative information of the analyte, there are 2 main approaches to stating a multivariate calibration model. These approaches are often referred to as direct (the classical least squares [CLS] model being the most used, but CLS variations have been published¹⁻⁵) and inverse (that include inverse least squares [ILS] methods such as multiple linear regression, principal component regression, partial-least squares [PLS] regression, and ridge regression [RR], among others).^{6,7} To provide accurate predictions from nonselective measurements, each calibration approach sets constraints on the model relative to the non-analyte components (interferents). The understanding of such constraints has led to the concept of net analyte signal (NAS)⁸⁻¹³ and selectivity measures.^{8,14-20}

It is well developed that CLS imposes strict orthogonality constraints while ILS methods do not.²¹ Specifically, depending on the analyte and interferent concentrations in the calibration set as well as the spectral noise structure, different ILS models will be obtained²¹⁻²⁴ while nearly the same orthogonal CLS models are estimated.²¹ In essence, ILS models at appropriate tuning parameter values (eg, latent variables for PLS and ridge parameter for RR) are the models most useful for prediction and not orthogonal to interferents. However, ILS models at correct tuning parameter values are actually commonly nearly orthogonal to the interferents.²⁵ Proposed in this paper is relaxed CLS (*r*CLS), a method that relaxes the CLS orthogonality constraints with relaxation parameters to form a family of models. The process is extended to form relaxed ILS (*r*ILS) models. In fact, the proposed relaxation parameters used to form *r*CLS and *r*ILS models are related to the component-wise selectivity definitions proposed in reference²¹ and this relationship is elaborated in this paper.

Also presented in this paper is a unified definition of the selectivity coefficient. The selectivity coefficient is a measure used in analytical chemistry developed for univariate calibration.^{26,27} It finds most use with ion-selective electrode calibration.^{28,29} The selectivity coefficient (formally defined in Section 3.2 for univariate and Section 3.3 for multivariate) characterizes the sensitivity of a method for an interferent relative to the sensitivity of the method to the analyte. To date, a definition of the selectivity coefficient is nonexistent for multivariate CLS and ILS calibrations. Developed in this paper is a definition that applies to both calibration processes and reduces to the usual univariate definition. With multivariate selectivity coefficients, it is possible to assess the prediction quality expected from a model relative to the sample analyte and interferent concentrations. Thus, even though the selectivity for the analyte can be poor, acceptable concentration predictions may still be possible.³⁰ In addition to the selectivity coefficient, new component-wise selectivity and sensitivity measures are introduced.

This paper begins by discussing the underlying calibration model with respect to CLS and ILS and the orthogonality goals between pure component spectra and model regression vectors. After this, an overview of estimating CLS and ILS models is provided leading into presentations of the relaxed versions of CLS (*r*CLS) and ILS (*r*ILS). With the relaxed multivariate calibration methods established, selectivity coefficients are defined. This section begins with the classical univariate definition and generalizes the definition to multivariate calibration. The *r*CLS process and selectivity coefficient definition are exemplified using a small near infrared (NIR) calibration set of 2 species.

2 | SELECTIVITY-RELAXED CALIBRATION THEORY

For the following mathematical equations and expressions, scalars are represented as nonbold italic letters, column vectors as bold lowercase letters, and matrices as bold uppercase letters. The hat (^) used in the literature over a variable to indicate estimated quantity has been omitted for simplicity except for predicted quantities of the analyte; whether the magnitude is measured or calculated can be deduced from the context. The superscripts “T” and “+” denote transposition and a pseudoinverse, respectively. The matrix **I** represents a properly dimensioned identity matrix. The L₂ norm of a vector is symbolized by $\|\cdot\|$.

2.1 | Underlying model

Let \mathbf{r} ($P \times 1$) be the measured (background corrected) spectrum measured over P wavelengths of a sample with N responding components. Let the subscript a denote the analyte of interest in a sample and i any other sample component, here regarded as interferent for the determination of a . Throughout this paper, it is assumed that a sample spectrum is a linear combination of analyte and interferent pure component spectra following the Beer-Lambert law

$$\begin{aligned}\mathbf{r}^T &= c_a \mathbf{s}_a^T + \sum_i c_i \mathbf{s}_i^T + \mathbf{e}^T \\ &= \mathbf{c}^T \mathbf{S}^T + \mathbf{e}^T\end{aligned}\quad (1)$$

where the columns of \mathbf{S} ($P \times N$) are the pure component spectra of the analyte \mathbf{s}_a and interferents \mathbf{s}_i , \mathbf{c} ($N \times 1$) holds respective concentrations of the analyte (c_a) and interferents (c_i), and \mathbf{e} contains the measurement noise. For simplicity, this paper only considers interfering components, but the developments can also be asserted for other sources of variation in \mathbf{r} , such as measurement temperature and instrument.

A wide range of modeling methods estimate the analyte concentration in a sample by

$$\hat{c}_a = \mathbf{r}^T \mathbf{b}_a, \quad (2)$$

where \mathbf{b}_a contains the model regression coefficients for component a . In some cases, the concentrations of all components in the calibration standards are known and it is possible to derive a vector of regression coefficients for each component. In this case, all components are predicted simultaneously by

$$\hat{\mathbf{c}}^T = \mathbf{r}^T \mathbf{B}, \quad (3)$$

where \mathbf{B} is a matrix of regression coefficients whose a th column is \mathbf{b}_a and $\hat{\mathbf{c}}$ is an estimate of \mathbf{c} in Equation 1.

Because of unselective measurements, a common goal of calibration methods is to provide a model that predicts the analyte concentration without being affected by the interferents. This can be expressed as

$$\frac{\partial \hat{c}_a}{\partial c_k} = \delta_{ka} \quad \delta_{ka} = \begin{cases} 0, & \text{if } k \neq a \\ 1, & \text{if } k = a \end{cases}, \quad (4)$$

where c_k is either the analyte a or an interferent i concentration and δ_{ka} is the Kronecker delta. Figure 1 illustrates the expected behavior according to Equation 4 where the prediction \hat{c}_a is the same as the concentration of a in the sample for any value of c_a , (ie, $\frac{\partial \hat{c}_a}{\partial c_a} = 1$) and, at the same time, the prediction does not change with an increasing amount of interferents in the sample (ie, $\frac{\partial \hat{c}_a}{\partial c_i} = 0 \quad \forall i \neq a$).

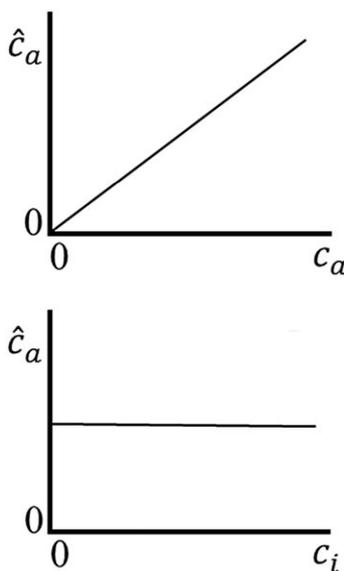


FIGURE 1 Ideal behavior of a fully selective calibration model with c_a and c_i denoting the analyte and interferent concentration. See text for discussion

Inserting Equation 1 into Equation 2 results in

$$\hat{c}_a = c_a \mathbf{s}_a^T \mathbf{b}_a + \sum_i c_i \mathbf{s}_i^T \mathbf{b}_a + \mathbf{e}^T \mathbf{b}_a \quad (5)$$

and applying Equation 4 gives

$$\begin{aligned} \frac{\partial \hat{c}_a}{\partial c_a} &= \mathbf{s}_a^T \mathbf{b}_a = 1 \\ \frac{\partial \hat{c}_a}{\partial c_i} &= \mathbf{s}_i^T \mathbf{b}_a = 0 \quad \forall i \end{aligned} \quad (6)$$

meaning that \mathbf{b}_a must be estimated such that $\mathbf{s}_a^T \mathbf{b}_a = 1$ and at the same time, be orthogonal to the (usually unknown) spectra of all interferences in the sample ($\mathbf{s}_i^T \mathbf{b}_a = 0, \forall i$) and to the noise vector ($\mathbf{e}^T \mathbf{b}_a = 0$). Similar constraints can be stated for other components in the sample acting as the analyte by redefining a and i .

Assuming Equation 1 is valid and taking all components into account simultaneously, then from Equation 3,

$$\hat{\mathbf{c}}^T = \mathbf{c}^T \mathbf{S}^T \mathbf{B} + \mathbf{e}^T \mathbf{B} \quad (7)$$

and

$$\frac{\partial \hat{\mathbf{c}}^T}{\partial \mathbf{c}} = \mathbf{S}^T \mathbf{B}. \quad (8)$$

For \mathbf{B} to fulfill the requirement stated in Equation 4, then

$$\mathbf{S}^T \mathbf{B} = \mathbf{I}, \quad (9)$$

where the identity matrix \mathbf{I} arises from the Kronecker delta in Equation 4.

Estimating a regression vector \mathbf{b}_a satisfying Equation 6, or a matrix \mathbf{B} fulfilling Equation 9 is the CLS working framework and the ILS regression targets these equations.²¹ These points are further discussed in the next section.

2.1.1 | Net analyte signal

The requirement $\delta_{ka} = \begin{cases} 0, & \text{if } k \neq a \\ 1, & \text{if } k = a \end{cases}$ keeps interferences from biasing the prediction \hat{c}_a . A model that fulfills this requirement has a null sensitivity for the interferences, or, equivalently, full selectivity for the analyte of interest. Such a model is said to maintain orthogonality and uses for prediction only that part of the analyte signal orthogonal to the interferent space also referred to as the NAS.⁸ This definition pertains to CLS. A more encompassing definition that includes CLS and ILS is to define NAS as that part of a sample signal used for prediction.¹⁰ As such, the NAS plays a fundamental role in the prediction of multivariate models and influences the prediction variance. The NAS for a sample spectrum is commonly given by

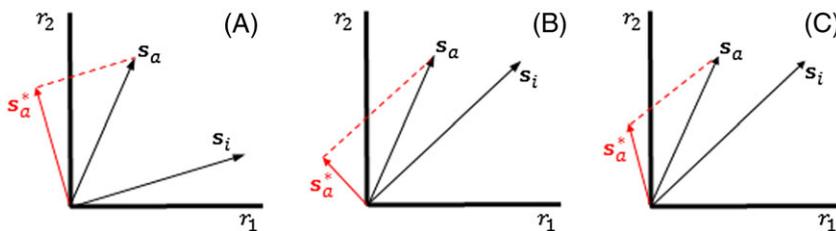
$$\mathbf{r}_a^* = \hat{c}_a \mathbf{s}_a^*, \quad (10)$$

where the net analyte sensitivity vector \mathbf{s}_a^* , can be calculated as

$$\mathbf{s}_a^* = \frac{\mathbf{b}_a}{\|\mathbf{b}_a\|^2}. \quad (11)$$

The constraints in Equation 4 ensure that *modeled* interferences will not bias the prediction of the analyte. To fulfill Equation 4, \mathbf{b}_a must be orthogonal to the spectra of the interferences, the CLS constraint. This is equivalent to saying that the only contribution to the NAS of the unknown sample, \mathbf{r}_a^* , must be generated by only the analyte of interest. The more similar the analyte and interferent spectra become, Figure 2A,B show that the NAS decreases for the orthogonal projection of \mathbf{s}_a against \mathbf{s}_i (the \mathbf{s}_a^* is restricted to be orthogonal to the interferent). Similarly, the NAS decreases as the degree of sample complexity increases. All other things being equal, a small norm of the NAS encompasses a small signal-to-noise ratio, large prediction variance and hence, large overall prediction errors.^{10,30} This realization leads to the awareness that strict adherence to the

FIGURE 2 Example net analyte sensitivity s_a^* effecting the net analyte signal (NAS) (Equation 10) at 2 measured responses r_1 and r_2 . A and B, The NAS becomes smaller for the orthogonal projection of s_a against s_i as the similarity between the analyte spectrum s_a , and the interferent spectrum s_i increases. C, Relaxing the orthogonality constraint in Figure 2B with respect to the interferent can lead to a larger NAS



defining relation in Equation 4 can lead to underperforming models. The underperformance is the price paid by CLS to avoid the interferent from biasing the prediction independent of the interferent concentration.

A remedy for this CLS limitation is realizing that prediction bias caused by an interferent is the result of the model sensitivity for the analyte, the model sensitivity for that interferent, and the amount of interferent in the sample. Hence, when the amount of interferent is expected to be (or known to be) very low or the sensitivity for the interferent is low, a small prediction bias due to the interferent can become acceptable if there is trade-off for another prediction quality such as prediction variance, the same bias/variance tradeoff premise for ILS. Under these circumstances, a null sensitivity for that interferent is not required for CLS, and the orthogonality (ie, full selectivity) constraint of the analyte against all interferences in Equation 4 can be relaxed to increase the NAS and decrease prediction uncertainty. In this case, s_a^* is not constrained to be orthogonal to the interferent. Transitioning from Figure 2B to Figure 2C demonstrates the concept.

2.2 | Classical and inverse calibrations

Using a collection of calibration samples, Equation 1 is expressed as

$$\begin{aligned} \mathbf{R} &= \mathbf{C}\mathbf{S}^T + \mathbf{E} \\ &= \mathbf{c}_a\mathbf{s}_a^T + \sum_i \mathbf{c}_i\mathbf{s}_i^T + \mathbf{E}, \end{aligned} \quad (12)$$

where \mathbf{R} ($M \times P$) holds the measured spectra in rows, \mathbf{C} ($M \times N$) contains the known concentration of the components in the standards, and \mathbf{E} ($M \times P$) is the error matrix associated with this decomposition. The calibration (training) step of CLS amounts to estimating the pure component spectra as

$$\mathbf{S}^T = \mathbf{C}^+\mathbf{R}. \quad (13)$$

In this paper, it is assumed that the calibration standards, and hence \mathbf{S} , properly reflect the composite system of matrix interactions.

Once \mathbf{S} is known, the solution to Equation 9 gives direct access to the CLS regression coefficients satisfying Equations 4 and 6 and is written as

$$\mathbf{B}_{\text{CLS}} = (\mathbf{S}^T)^+. \quad (14)$$

With the matrix of regression vectors, concentrations for calibrated components can be estimated for a sample using Equation 3.

In some analytical problems, such as in the analysis of natural products or very complex mixtures, pure component spectra are not known or cannot be estimated by Equation 13 because not all components concentrations are known even though Equation 12 is the underlying model. In this case, ILS modeling (multiple linear regression, principal component regression, PLS, RR, etc) uses only quantitative information of the analyte of interest for training and does not need pure component spectra. With ILS, the sample prediction for the analyte occurs by Equation 2 by using a vector of regression coefficients for analyte a obtained from the pseudoinverse of \mathbf{R} expressed as

$$\mathbf{b}_{\text{ILS},a} = \mathbf{R}^+\mathbf{c}_a, \quad (15)$$

where \mathbf{R}^+ depends on the ILS method used. Note that if concentrations for more than one responding component are actually known, the ILS model matrix can be computed by

$$\mathbf{B}_{\text{ILS}} = \mathbf{R}^+ \mathbf{C}, \quad (16)$$

where $\mathbf{b}_{\text{ILS},a}$ is the analyte column in \mathbf{B}_{ILS} .

Because of the unique differences between how regression vectors are obtained for CLS and ILS (Equation 14 versus Equation 15 or 16), \mathbf{B}_{CLS} satisfies the orthogonality constraints while $\mathbf{b}_{\text{ILS},a}$ and \mathbf{B}_{ILS} do not. This conclusion is fully derived in Brown.²¹ Thus, ILS processes strive to meet orthogonality and the final model is one that is useful for prediction. This near-orthogonality has been graphically characterized.²⁵ It should be noted that in the hypothetical noise-free measurement situation, the ILS models do fulfill³¹ Equations 6 and 9.

2.3 | Selectivity-*r*CLS

The selectivity requirements in CLS can be relaxed by setting Equation 4 to

$$\frac{\partial \hat{c}_a}{\partial c_k} = l_{kj}, \quad (17)$$

where l is the relaxation parameter relative to k and j that represents either the analyte or an interferent as the situation dictates and is read as the relaxation parameter for the k th component on the j th component with $-\infty < l_{kj} < \infty$. A relaxation parameter represents the change in predicted analyte concentration per unit change in the respective interferent concentration and allows the analyte concentration prediction to be affected by the interferent (ie, introduces prediction bias). Better relaxation parameters are those that slightly increase bias for lower prediction variance. The matrix \mathbf{L} containing these relaxation coefficients l_{kj} replaces the identity matrix \mathbf{I} in Equation 9 resulting in new regression coefficients for relaxed models $\mathbf{B}_{\mathbf{L}}$ expressed as

$$\begin{aligned} \frac{\partial \hat{\mathbf{c}}^T}{\partial \mathbf{c}} &= \mathbf{S}^T \mathbf{B}_{\mathbf{L}} \\ &= \mathbf{L} = \begin{bmatrix} l_{aa} & \cdots & l_{ai} \\ \vdots & \ddots & \vdots \\ l_{ia} & \cdots & l_{ii} \end{bmatrix}. \end{aligned} \quad (18)$$

Solving Equation 18 for the *r*CLS coefficients in $\mathbf{B}_{\mathbf{L}}$ results in

$$\mathbf{B}_{\mathbf{L}} = (\mathbf{S}^T)^+ \mathbf{L} = \mathbf{B}_{\text{CLS}} \mathbf{L}. \quad (19)$$

The $\mathbf{B}_{\mathbf{L}}$ coefficients deviate from \mathbf{B}_{CLS} by the amount defined by \mathbf{L} . It is seen that *r*CLS is a generalization that converges to CLS when full selectivity is imposed by $\mathbf{L} = \mathbf{I}$ and hence, $\mathbf{B}_{\mathbf{L}} = \mathbf{B}_{\text{CLS}}$.

Expanding Equation 19, the regression coefficients in *r*CLS are seen to be

$$\begin{aligned} \mathbf{b}_{\mathbf{L},a} &= \mathbf{b}_{\text{CLS},a} l_{aa} + \sum_i \mathbf{b}_{\text{CLS},i} l_{ia}, \\ \mathbf{b}_{\mathbf{L},i} &= \sum_i \mathbf{b}_{\text{CLS},a} l_{ai} + \mathbf{b}_{\text{CLS},i} l_{ii}, \end{aligned} \quad (20)$$

where $\mathbf{b}_{\mathbf{L},a}$ and $\mathbf{b}_{\mathbf{L},i}$ are respective vectors of regression coefficients in $\mathbf{B}_{\mathbf{L}}$. Using $\mathbf{b}_{\mathbf{L},a}$ to predict the analyte concentration results in

$$\begin{aligned} \hat{c}_a &= \mathbf{r}^T \mathbf{b}_{\mathbf{L},a}, \\ &= (c_a \mathbf{s}_a^T + \sum_i c_i \mathbf{s}_i^T + \mathbf{e}^T) \mathbf{b}_{\mathbf{L},a}, \\ &= (c_a \mathbf{s}_a^T + \sum_i c_i \mathbf{s}_i^T + \mathbf{e}^T) (\mathbf{b}_{\text{CLS},a} l_{aa} + \sum_i \mathbf{b}_{\text{CLS},i} l_{ia}). \end{aligned} \quad (21)$$

Because of the CLS orthogonality constraints, $\mathbf{s}_a^T \mathbf{b}_{\text{CLS},a} = 1$, $\mathbf{s}_a^T \mathbf{b}_{\text{CLS},i} = 0$, $\mathbf{s}_i^T \mathbf{b}_{\text{CLS},a} = 0$, and $\mathbf{s}_i^T \mathbf{b}_{\text{CLS},i} = 1$, Equation 21 reduces to

$$\hat{c}_a = c_a l_{aa} + \sum_i c_i l_{ia} + e \quad (22)$$

Assuming $l_{aa} = 1$ in Equation 22, then the *bias* is 0 when $l_{ia} = 0$ (ie, full model selectivity); otherwise, the $c_i l_{ia}$ terms introduce bias in \hat{c}_a . As shown in the selectivity coefficient sections (Sections 3.2 and 3.3), this bias, however, will be lower if the concentration of the interferent in the sample is low and/or if the instrument sensitivity for the interferent is also low. In turn, \hat{c}_a can benefit from a lower variance if the L_2 norm of $\mathbf{b}_{L,a}$ is smaller. As shown in Section 6, other combinations of l_{aa} and l_{ia} are possible leading to other bias/variance trade-off situations for the relaxed \mathbf{B}_L .

Depending on the sample and degree of relaxation from orthogonality, the selectivity- r CLS model can have an improved prediction performance compared to the CLS model. Values for l_{kj} must be optimized for the dataset at hand, since the interferent spectral characteristics (interferent wavelength sensitivities and similarities with the analyte spectrum) and concentrations differ from one dataset to another. Also, note that for a given analyte, only the corresponding column of \mathbf{L} must be optimized. One approach to solve for \mathbf{L} (or the \mathbf{I}_a column) is to use a cross-validation strategy while systematically varying \mathbf{L} . Depending on the number of components and hence values in \mathbf{L} , this process is probably not feasible and an optimization algorithm such as simulated annealing or genetic algorithm could be used. For both situations, the optimal \mathbf{L} is sought that minimizes prediction error. Neither approaches were studied in this paper and only trends in measures of model quality are reported as relaxation parameters varied.

2.3.1 | Analysis of r CLS as a basis change from CLS

To sustain Equation 12 ($\mathbf{R} = \mathbf{CS}^T + \mathbf{E}$), the corresponding relaxed expression is expressed as

$$\begin{aligned} \mathbf{R} &= \mathbf{CLL}^{-1}\mathbf{S}^T + \mathbf{E} \\ &= \mathbf{C}_L\mathbf{S}_L^T + \mathbf{E} \end{aligned} \quad (23)$$

where $\mathbf{C}_L = \mathbf{CL}$ and $\mathbf{S}_L^T = \mathbf{L}^{-1}\mathbf{S}^T$. By inserting the invertible matrix \mathbf{L} , the solutions are not unique (ie, multiple solutions with the same fit \mathbf{E} can be obtained). Algebraically, the measured sample spectra (\mathbf{R}) can be expressed as their coordinates on a given vector base. The column vectors in \mathbf{S} represent N independent basis vectors that span the space of all possible mixtures of the N components where the linear model is valid. Hence, by setting the base to the pure component spectra at unit concentration, the coordinates of each sample in \mathbf{R} for this pure component base are the concentrations of each component in the respective mixture spectra. These concentrations are contained in \mathbf{C} and mathematically expressed by Equation 1.

The N column vectors in \mathbf{S}_L also span a vector space for spectra in \mathbf{R} (apart from noise). The coordinates of the spectra \mathbf{R} in this basis set are now given by \mathbf{C}_L . Thus, \mathbf{L}^{-1} is the transformation matrix from the \mathbf{S} basis set to an \mathbf{S}_L basis set. Depending on \mathbf{L} (degree of relaxation or transformation), different basis vectors can be formed and, respectively, the corresponding r CLS models in \mathbf{B}_L . The components of \mathbf{S}_L can be thought of as pseudo-components where a pseudo-component is a basis vector made up of mixtures of pure components that serves the same purpose as the pure component basis vectors.

The creation of pseudo-components from pure components amounts to relaxing the selectivity constraints of the CLS model. The concentration estimate of the analyte in a sample by r CLS ($\hat{c}_{L,a}$) is an unbiased estimate of the amount of the pseudo-component $\mathbf{s}_{L,a}$ but a biased estimate of the amount of the pure component \mathbf{s}_a . As an example, it is possible to tune the relaxation (transformation) parameters in order for the pseudo-component to be almost the same as the pure component spectrum (eg, a two-component situation with $\mathbf{s}_{L,a} = 0.99\mathbf{s}_a + 0.01\mathbf{s}_i$). In this situation, the solution $\hat{c}_{L,a}$ is as a biased estimate of c_a due to a small contribution from the interferent.

The transformation of the basis vectors in \mathbf{S} by \mathbf{L}^{-1} is equally matched by transformation of the \mathbf{B}_{CLS} basis vectors (the N column vectors) to the new basis set \mathbf{B}_L with \mathbf{L} being the transformation matrix. These basis vector changes are illustrated in Figure 3 (patterned after the general situation in Vainchtein³²). From Figure 3, it is recognized that \mathbf{S} and \mathbf{B}_{CLS} have the special property of being reciprocal basis vectors ($\mathbf{S}^T\mathbf{B}_{CLS} = \mathbf{I}$) as do \mathbf{S}_L and \mathbf{B}_L ($\mathbf{S}_L^T\mathbf{B}_L = \mathbf{I}$). Because of this reciprocal nature within CLS and r CLS, the respective calibrations and predictions can be expressed as covariant and contravariant components with additional special relationships.³²⁻³⁴

Shown in Figure 4 is a geometric depiction (a rendition from the generic portrayal in Vainchtein³²) for a two-component CLS system (analyte and interferent) for a single spectrum from \mathbf{R} . The sample spectrum \mathbf{r} is not only represented as linear combination of the basis vectors in \mathbf{S} but also as a linear combination of the reciprocal basis vectors from \mathbf{B}_{CLS} . The components for the linear combination of the \mathbf{S} basis vectors are respective concentrations. Using Equation 1 with only one interferent, the relationship is written as

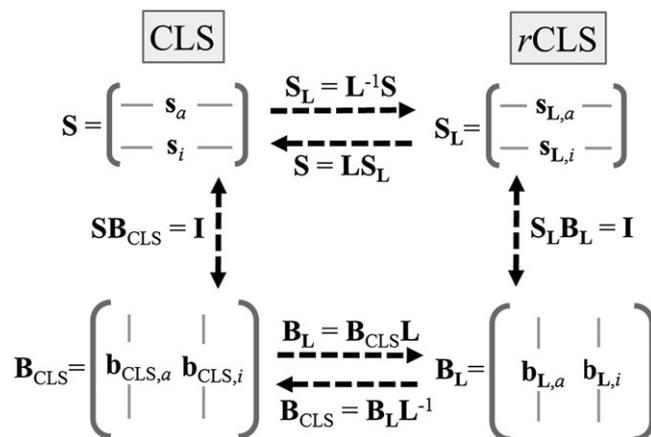


FIGURE 3 The classical least squares (CLS) covariant (\mathbf{S}) and reciprocal contravariant (\mathbf{B}_{CLS}) basis vectors with the respective changes to the relaxed classical least squares ($r\text{CLS}$) basis vectors. Example depicts a two-component situation for the analyte a and the interferent i

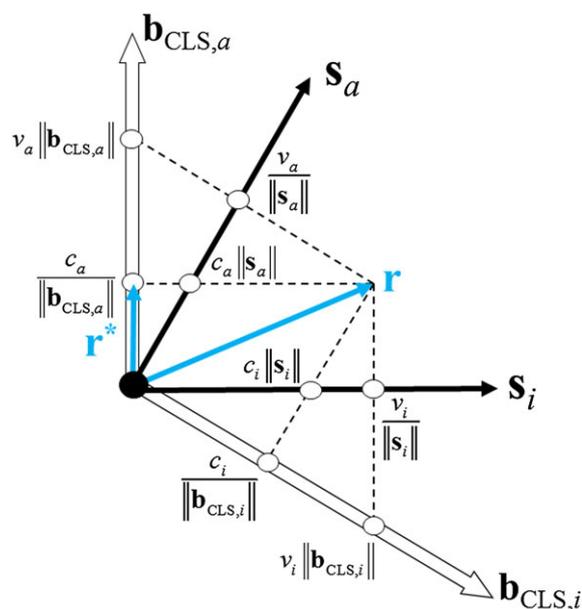


FIGURE 4 Relationship between classical least squares (CLS) covariant and contravariant basis vectors and the net analyte signal. A similar relationship can be characterized for the relaxed classical least squares ($r\text{CLS}$) basis vectors. The NAS (\mathbf{r}^*) and corresponding L_2 norm of \mathbf{r}^* ($c_a/\|\mathbf{b}_{\text{CLS},a}\|$) are shown

$$\mathbf{r}^T = c_a \mathbf{s}_a^T + c_i \mathbf{s}_i^T. \quad (24)$$

The concentrations are known as the contravariant components of \mathbf{r} in the covariant basis vectors \mathbf{s}_a and \mathbf{s}_i . These components are estimated individually by $\hat{c}_a = \mathbf{r}^T \mathbf{b}_{\text{CLS},a}$ and $\hat{c}_i = \mathbf{r}^T \mathbf{b}_{\text{CLS},i}$ or simultaneously by $\hat{\mathbf{c}}^T = \mathbf{r}^T \mathbf{B}_{\text{CLS}}$. The CLS regression vectors are sometimes referred to as the contravariant vector (or contravariant spectrum) because when the inner product (dot product) is taken between \mathbf{r} and a regression vector \mathbf{b}_{CLS} , the corresponding contravariant component (concentration) is obtained. The contravariant vectors are reciprocal basis vectors to the covariant basis vectors in \mathbf{S} .

The components for the linear combination of the \mathbf{B}_{CLS} basis vectors are known as the covariant components. A sample spectrum in this basis set is expressed by

$$\mathbf{r}^T = v_a \mathbf{b}_{\text{CLS},a}^T + v_i \mathbf{b}_{\text{CLS},i}^T \quad (25)$$

where v_a and v_i represent the covariant components. These components are computed individually by $v_a = \mathbf{r}^T \mathbf{s}_a$ and $v_i = \mathbf{r}^T \mathbf{s}_i$ or simultaneously by $\mathbf{v}^T = \mathbf{r}^T \mathbf{S}$. The vectors in \mathbf{S} can be labeled the covariant vectors because when the dot product is taken between \mathbf{r} and a pure component spectrum \mathbf{s} , the corresponding covariant component is obtained.

A similar depiction in Figure 4 can be made for the transformed basis vectors \mathbf{S}_L and \mathbf{B}_L by substituting the respective transformed basis vectors for \mathbf{S} and \mathbf{B}_{CLS} . Equally, the following equations represent the transformed contravariant and covariant situation

$$\begin{aligned} \mathbf{r}^T &= c_{L,a} \mathbf{s}_{L,a}^T + c_{L,i} \mathbf{s}_{L,i}^T \\ &= v_{L,a} \mathbf{b}_{L,a}^T + v_{L,i} \mathbf{b}_{L,i}^T \end{aligned} \quad (26)$$

From Figure 4, the relationship between the contravariant basis set and the NAS is clarified. Specifically, the contravariant component c_a can be found from either the component $c_a \|\mathbf{s}_a\|$ shown in Figure 4 or the projection of \mathbf{r} onto the contravariant vector $\mathbf{b}_{\text{CLS},a}$. This projection is the NAS labeled \mathbf{r}^* with the L_2 norm $\|\mathbf{r}^*\| = c_a / \|\mathbf{b}_{\text{CLS},a}\|$. Not as useful from an analytical chemist's perspective, Figure 4 shows the covariant components.

The ambiguity in the decomposition $\mathbf{R} = \mathbf{C}_L \mathbf{S}_L^T + \mathbf{E}$ results from the lack of constraints on \mathbf{L} in Equation 23. Thus, any invertible \mathbf{L} is a solution as demonstrated in Section 6 and many solutions (regression vectors) with no physical interpretation are possible that form effective predictions,²²⁻²⁴ ie, the models just work. Similarly, without constraints, \mathbf{C}_L and \mathbf{S}_L are not easy to chemically interpret. Regardless, the pseudo-components, by definition, meet the CLS hard constraints of $\mathbf{S}_L^T \mathbf{B}_L = \mathbf{I}$ giving $\hat{\mathbf{C}}_L = \mathbf{C}_L + \mathbf{E} \mathbf{B}_L$.

Because ILS models are not reciprocal basis vectors with pure component spectra in \mathbf{S} (the ILS regressions are not unique^{21,22}), an ILS regression vector is not a true contravariant vector. Thus, a sample spectrum cannot be represented as a linear combination of all respective ILS regression vectors (one for the analyte and one for each response forming interferent, \mathbf{B}_{ILS}) and simultaneously, a linear combination of \mathbf{S} . This is also true for the basis vectors transformed by \mathbf{L}^{-1} and \mathbf{L} . However, it should be noted that while the reciprocal basis set description (the covariant, contravariant, and orthogonality strictness) does not hold for ILS, many of the computations (NAS, net sensitivity vector) are still relevant. These points are further elaborated in the next section.

2.4 | Selectivity-*r*ILS

By expressing ILS in a relaxed format (*r*ILS), insight can be gained on how close an ILS solution comes to satisfying the CLS selectivity constraints. Even though all sample components are not known when using inverse models, it can still be assumed that the Beer-Lambert law (Equations 1 and 12) characterizes the underlying model generating the measured spectra.

Using Equation 2, the fitted concentrations for a sample using an ILS calibration is

$$\begin{aligned} \hat{c}_a &= \mathbf{r}^T \mathbf{b}_{\text{ILS},a} \\ &= c_a \mathbf{s}_a^T \mathbf{b}_{\text{ILS},a} + \sum_i c_i \mathbf{s}_i^T \mathbf{b}_{\text{ILS},a} + e \end{aligned} \quad (27)$$

Applying the relaxed parameter definition to Equation 27 results in

$$\begin{aligned} \frac{\partial \hat{c}_a}{\partial c_a} &= \mathbf{s}_a^T \mathbf{b}_{\text{ILS},a} = l_{aa} \\ \frac{\partial \hat{c}_a}{\partial c_i} &= \mathbf{s}_i^T \mathbf{b}_{\text{ILS},a} = l_{ia} \quad \forall i \neq a \end{aligned} \quad (28)$$

where the derivative in the second equation is taken for every interferent i . As already noted, ILS models implicitly relax the orthogonality constraint to maintain a bias/variance trade-off relative to respective minimization criteria,¹⁶ ie, $l_{aa} \neq 1$ and $l_{ia} \neq 0$. Variations of $\mathbf{s}_a^T \mathbf{b}_{\text{ILS},a}$ and $\mathbf{s}_i^T \mathbf{b}_{\text{ILS},a}$ (now termed l_{aa} and l_{ia}) as ILS tuning parameters vary has been graphically characterized.²⁵ Specifically, these terms were used to characterize the bias/variance tradeoff in selecting tuning parameters.

An r CLS model is formed from a linear combination of all analyte and interferent CLS models. Similarly, an r ILS model is formed from a linear combination of all ILS models. This combination to form the r ILS models \mathbf{B}_Q is expressed as

$$\mathbf{B}_Q = \mathbf{B}_{\text{ILS}}\mathbf{Q}, \quad (29)$$

where \mathbf{Q} denotes a matrix of coefficients to form a linear combination of \mathbf{B}_{ILS} . Note that \mathbf{Q} is different from the \mathbf{L} used in r CLS. With CLS, there is an inverse relationship between \mathbf{S}^T and \mathbf{B}_{CLS} and hence, the \mathbf{L} matrix used in Equation 19 to form the relaxed models \mathbf{B}_L as the linear combinations of \mathbf{B}_{CLS} is the same \mathbf{L} matrix obtained in Equation 18 for the inner products of the pure component spectra with \mathbf{B}_L (Figure 3 is a diagram of the inverse basis relationship). With ILS, the inverse relationship between \mathbf{S}^T and \mathbf{B}_{ILS} does not exist. For example, given the generic situation of Equation 9 ($\mathbf{S}^T\mathbf{B} = \mathbf{I}$), the only \mathbf{B} satisfying this equality is \mathbf{B}_{CLS} .

As with r CLS, the degree of deviations of \mathbf{L} from \mathbf{I} for the r ILS models in \mathbf{B}_Q depend on the coefficients in \mathbf{L} shown by

$$\begin{aligned} \frac{\partial \hat{\mathbf{c}}^T}{\partial \mathbf{c}} &= \mathbf{S}^T\mathbf{B}_Q \\ &= \mathbf{L} \end{aligned} \quad (30)$$

relative to prediction by $\hat{\mathbf{c}}^T = \mathbf{c}^T\mathbf{S}^T\mathbf{B}_Q + \mathbf{e}^T\mathbf{B}_Q$, similar to Equation 7 for CLS. The more \mathbf{L} for a particular modeling process deviates from \mathbf{I} , the greater the deviation from obtaining an orthogonally constrained model. When \mathbf{Q} is the identity matrix, then the r ILS models are the ILS models and \mathbf{L} has the relaxation coefficients expressing the amount of orthogonality obtained for the ILS models. This statement agrees with previous work where the matrix $\mathbf{W} = \mathbf{S}^T\mathbf{B}_{\text{ILS}} - \mathbf{I}$ is defined¹⁶ and can be rearranged to $\mathbf{W} = \mathbf{L} - \mathbf{I}$. Note that this second equation is true for both r CLS and r ILS. The \mathbf{W} can be interpreted as a measure of departure from the CLS.

Prediction of the analyte in a sample by a particular r ILS model is expressed by

$$\begin{aligned} \hat{c}_a &= \mathbf{r}^T\mathbf{b}_{Q,a} \\ &= (c_a\mathbf{s}_a^T + \sum_i c_i\mathbf{s}_i^T)\mathbf{b}_{Q,a} + e, \\ &= c_a l_{aa} + \sum_i c_i l_{ia} + e. \end{aligned} \quad (31)$$

The more of \mathbf{S} that is known, the better the assessment of the interferent effects defined by \mathbf{L} for particular r ILS model. Interpretation of the respective r ILS l_{aa} and l_{ia} values is the same as for r CLS; l_{ia} values denote the change in predicted analyte concentration per unit change in the respective interferent concentrations relative to the model used for prediction. This aspect is further commented on in the Selectivity section.

Values in \mathbf{L} are not as simple for r ILS as for r CLS. This complexity can be seen by expanding Equation 29 for a simple situation with an analyte and one interferent to

$$\begin{aligned} \mathbf{b}_{Q,a} &= \mathbf{b}_{\text{ILS},a}q_{aa} + \mathbf{b}_{\text{ILS},i}q_{ia}, \\ \mathbf{b}_{Q,i} &= \mathbf{b}_{\text{ILS},a}q_{ai} + \mathbf{b}_{\text{ILS},i}q_{ii}, \end{aligned} \quad (32)$$

where $\mathbf{b}_{Q,a}$ and $\mathbf{b}_{Q,i}$ are the vectors of regression coefficients in \mathbf{B}_Q . Using $\mathbf{b}_{Q,a}$ expressed in Equation 32 to predict the analyte concentration results in

$$\begin{aligned} \hat{c}_a &= \mathbf{r}^T\mathbf{b}_{Q,a}, \\ &= c_a\mathbf{s}_a^T\mathbf{b}_{Q,a} + c_i\mathbf{s}_i^T\mathbf{b}_{Q,a} + e, \\ &= c_a(\mathbf{s}_a^T\mathbf{b}_{\text{ILS},a}q_{aa} + \mathbf{s}_a^T\mathbf{b}_{\text{ILS},i}q_{ia}) + c_i(\mathbf{s}_i^T\mathbf{b}_{\text{ILS},a}q_{aa} + \mathbf{s}_i^T\mathbf{b}_{\text{ILS},i}q_{ia}) + e. \end{aligned} \quad (33)$$

From comparing Equation 33 to Equation 31, the l_{aa} and l_{ia} values are written as

$$\begin{aligned} l_{aa} &= \mathbf{s}_a^T\mathbf{b}_{\text{ILS},a}q_{aa} + \mathbf{s}_a^T\mathbf{b}_{\text{ILS},i}q_{ia}, \\ l_{ia} &= \mathbf{s}_i^T\mathbf{b}_{\text{ILS},a}q_{aa} + \mathbf{s}_i^T\mathbf{b}_{\text{ILS},i}q_{ia}. \end{aligned} \quad (34)$$

Equation 34 shows that the final l_{aa} and l_{ia} values depend on the relationship between the base ILS models being used and amount (or weight) of the base ILS models combined to form the final analyte model vector $\mathbf{b}_{\mathbf{Q},a}$. If $\mathbf{Q} = \mathbf{I}$, then $l_{aa} = \mathbf{s}_a^T \mathbf{b}_{\text{ILS},a}$ and $l_{ia} = \mathbf{s}_i^T \mathbf{b}_{\text{ILS},a}$.

As with r CLS, a systematic variation of \mathbf{Q} or optimization algorithm can be used to determine appropriate \mathbf{Q} values (or the \mathbf{q}_a column) for r ILS. However, because all the ILS models are rarely known for the analyte and each interferent, then the practicality of using r ILS is nearly nonexistent. It is interesting to note that Equation 29 can be thought of as a more general regression strategy based on stacked regression³⁵⁻³⁷ where a collection of analyte regression models are used. For example, using a set of analyte models for a \mathbf{B} , a simple stacked regression gives $\mathbf{c}_a = \mathbf{R}\mathbf{b}_{\mathbf{q}_a} = \mathbf{R}\mathbf{B}\mathbf{q}_a = \hat{\mathbf{C}}\mathbf{q}_a$. Least squares can be used to solve for \mathbf{q}_a .

3 | SELECTIVITY

While selectivity is the focus in this section, analyte and interferent sensitivities are key figures of merit^{20,38} that selectivity depends on. In particular, the accuracy of analytical determinations can be adversely affected by other species present in the sample and/or variations of analytical conditions (eg, changes in temperature in NIR spectra) that alter the instrument response for the sample. Sensitivity and selectivity are 2 figures of merit balanced in forming calibration models.^{18,25} Approaches to measure and quantify sensitivity and selectivity have been provided (previous works¹⁸⁻²⁰ and references therein). Regarding selectivity, Ridder and Brown¹⁷ state “a method is fully selective against a particular interferent if the result it renders is independent of the concentration of the interferent. If a method is not fully selective, then the results it produces will depend on the concentration of the interferent.” In this work, a selectivity definition is proposed based on r CLS and is applicable to the usual multivariate CLS and ILS calibration models and reduces to the standard univariate definition. Because sensitivity is key to defining selectivity, a brief overview of univariate sensitivity is given next.

3.1 | Sensitivity in univariate calibration

Sensors are designed to respond to the amount of a component of interest (the “analyte” a) in a sample. The “sensitivity” of the sensor for analyte a , s_a , is defined³⁹ as the change in the measured response r per unit change of concentration of a (c_a) expressed as

$$\frac{dr}{dc_a} = s_a. \quad (35)$$

In univariate CLS calibration with narrow concentration ranges, the relationship between response and concentration should be describable by a straight line. In this situation, sensitivity is constant over the concentration range and is the slope of the regression line.³⁹ If the relationship is not linear, then the sensitivity changes depending on the value of c_a . The sensitivity can also depend on the sample matrix, in which case the standard addition method is used. In addition to the analyte, the sensor can also respond to the presence of other components in the sample (interferents) and hence, the measured response is a mix of responses generated by the analyte and interferents. The degree of contribution from an interferent signal depends on how sensitive the sensor is to the interferent (s_i). This sensitivity can be obtained by measuring samples with increasing known amounts of interferent when available.

3.2 | Selectivity and the selectivity coefficient in univariate CLS

In univariate calibration, the (background corrected) measured response is assumed to follow the linear model

$$r = c_a s_a + e \quad (36)$$

with prediction by

$$\begin{aligned} \hat{c}_a &= \frac{r}{s_a}, \\ &= r b_a, \\ &= c_a s_a b_a + e b_a, \end{aligned} \quad (37)$$

where $b_a = 1/s_a$ is defined for clarity. Imposing the condition in Equation 4 on Equation 37 gives

$$\begin{aligned}\frac{\partial \hat{c}_a}{\partial c_a} &= s_a b_a \\ &= 1\end{aligned}\quad (38)$$

If unexpected interferences contribute to the measured response of the unknown sample, Equation 36 is now written as

$$r = c_a s_a + \sum_i c_i s_i + e \quad (39)$$

and prediction with Equation 37 is given by
$$\hat{c}_a = c_a + \sum_i c_i s_i b_a + e b_a. \quad (40)$$

A nonzero sensitivity of the sensor for interferent i ($s_i \neq 0$) introduces a bias in the prediction \hat{c}_a in the amount $c_i s_i b_a$. This bias depends on the amount of the interferent i (c_i) and on the sensitivity for the analyte ($b_a = 1/s_a$) and for the interferent (s_i).

Figure 5 depicts the situation for one interferent. Mathematically, $\frac{\partial \hat{c}_a}{\partial c_a} = s_a b_a = 1$ is true, but the interferent conditions in Equation 4 are not followed and $\frac{\partial \hat{c}_a}{\partial c_i} = s_i b_a \neq 0$ for each interferent.

The absolute sensitivity of the sensor for an interferent is not as important as the relative sensor sensitivity for the interferent compared to the analyte sensitivity. In this sense, “selectivity” for an analyte means lack of sensitivity (low enough sensitivity) for a potential interferent compared to that for the analyte. Accordingly, the univariate “selectivity coefficient” for analyte a in the presence of interferent i is defined as

$$\begin{aligned}K_{a,i} &= \frac{s_i b_a}{s_a b_a} \\ &= \frac{s_i}{s_a}\end{aligned}\quad (41)$$

with $-\infty < K_{a,i} < \infty$. In univariate calibration with spectroscopic data, the sensitivities in Equation 41 are usually positive, but for other systems of analysis, negative values are possible. The selectivity coefficient is used to quantitate the model sensitivity towards the interferent relative to the analyte model sensitivity.^{18,26,27} With the selectivity coefficient now defined, then dividing Equation 39 by s_a and Equation 40 by $s_a b_a$ with rearrangement gives, respectively,

$$\hat{c}_a = c_a + \sum_i c_i K_{a,i} + e b_a \quad (42)$$

The analyte prediction error is obtained by subtracting the true analyte concentration value from Equation 42 producing

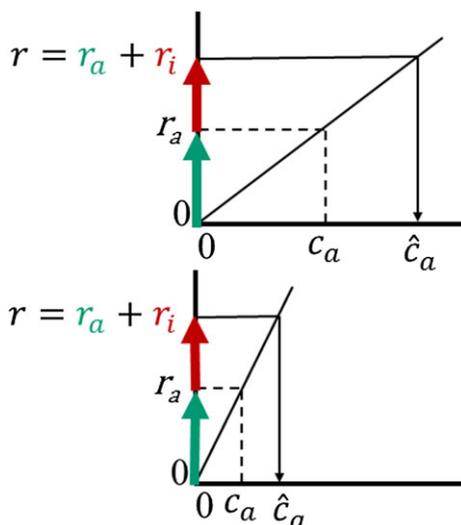


FIGURE 5 Bias introduced by an interferent in a prediction not only depends on the sensitivity of the sensor for the interferent but also the sensitivity of the sensor for the analyte (through the slope of the calibration line)

$$\hat{c}_a - c_a = \sum_i c_i K_{a,i} + eb_a \quad (43)$$

and expressed as a relative error by

$$\frac{\hat{c}_a - c_a}{c_a} = \sum_i \frac{c_i}{c_a} K_{a,i} + \frac{eb_a}{c_a} \quad (44)$$

Equations 43 to 44 show the crucial significance of selectivity coefficients for quantitating the effect each interferent has on a particular analysis. For example, a $K_{a,i} = 0.001$ means that the sensor is 1000 times more sensitive to the analyte than the interferent. Also shown in Equations 42 and 43 is the direct effect analyte sensitivity has on the prediction variation due to random error e . Specifically, the greater the analyte sensitivity, the smaller the effect. Selectivity, like sensitivity, is continuous and when it is stated that “a sensor is selective for the analyte of interest,” actually what is meant is that the sensor is selective enough such that errors caused by the sensitivity to the interferent become irrelevant. A sensor can be checked for a number of potential interferents followed by calculating the selectivity coefficient $K_{a,i}$ for each interferent. The greater the number of low $K_{a,i}$ values, the more the sensor can be considered globally selective for the analyte.

Applying notation previously developed for selectivity assessment in multivariate ILS¹⁶ to the univariate situation and starting with Equation 40, the prediction error is written as

$$\begin{aligned} \hat{c}_a - c_a &= c_a(s_a b_a - 1) + \sum_i c_i s_i b_a + eb_a \\ &= c_a w_a + \sum_i c_i w_i + eb_a \\ &= \sum_i c_i w_i + eb_a \end{aligned} \quad (45)$$

where $w_a = s_a b_a - 1$ and $w_i = s_i b_a$ and because in this univariate situation, $b_a = 1/s_a$, then $w_i = K_{a,i}$. The corresponding univariate selectivity indicators developed for multivariate ILS¹⁶ (presented in the next section) are for the analyte

$$\text{SEL}_a^{\text{IND}} = 1/(1 + w_a^2) \quad (46)$$

and each interferent

$$\text{SEL}_i^{\text{IND}} = 1/(1 + w_i^2). \quad (47)$$

The interferent prediction problem shown in Equations 42 to 44 does not exist with multivariate CLS. In this situation, the regression vector cancels the effect of the interferent ($\mathbf{s}_i^T \mathbf{b}_a = 0$) to fulfill the orthogonality constraint.

3.3 | Selectivity, sensitivity, and the selectivity coefficient in multivariate *r*CLS and ILS

3.3.1 | Selectivity coefficient

For congruence with univariate calibration, a definition for the selectivity coefficient in multivariate calibration should consider how nonmodeled interferents affect the prediction from the multivariate calibration model. Moreover, the selectivity coefficient should also be defined for modeled interferents since they also affect the model's prediction performance by decreasing the NAS and increasing the prediction variance.

Applying Equation 41 for the univariate situation to the multivariate case gives the selectivity coefficient as

$$\begin{aligned} K_{a,i} &= \frac{\mathbf{s}_i^T \mathbf{b}_a}{\mathbf{s}_a^T \mathbf{b}_a}, \\ &= \frac{l_{ia}}{l_{aa}}, \end{aligned} \quad (48)$$

where the generic model vector \mathbf{b}_a notation indicates any analyte model vector via CLS, *r*CLS, ILS, *r*ILS, or other method and $-\infty < K_{a,i} < \infty$ with values approaching zero denoting an optimal selectivity situation as with the univariate definition. Using

the generic model, Equation 22 (*r*CLS) and Equation 32 (*r*ILS) are generalized to

$$\hat{c}_a = c_a l_{aa} + \sum_i c_i l_{i,a} + e \quad (49)$$

and incorporating Equation 48 for the selectivity coefficient definition into Equation 49 produces

$$\hat{c}_a = l_{aa} (c_a + \sum_i c_i K_{a,i}) + e, \quad (50)$$

which is the multivariate representation of univariate situation in Equation 42. If a model is accurately predicting the analyte, the model is doing a good job of making the contributions sum to the analyte concentration in Equation 49. Specifically, while regression vectors can have uniquely different shapes and magnitudes, the analyte concentration predictions are similar.²¹⁻²⁴

The corresponding multivariate prediction error equations to univariate Equations 43 and 44 are, respectively,

$$\hat{c}_a - c_a = c_a (l_{aa} - 1) + l_{aa} \sum_i c_i K_{a,i} + e, \quad (51)$$

$$\frac{\hat{c}_a - c_a}{c_a} = (l_{aa} - 1) + l_{aa} \sum_i \frac{c_i}{c_a} K_{a,i} + \frac{e}{c_a}. \quad (52)$$

With selectivity coefficients, Equations 50 to 52 show the effect an interferent will have on the analyte prediction (or prediction error) for a sample. Different from the univariate situation, the prediction does depend on the relaxation parameter for the analyte (l_{aa}) where for the univariate case, $l_{aa} = 1$.

With respect to the approach taken in Brown and Ridder,¹⁶ Equation 47 for univariate calibration becomes for multivariate calibration

$$\begin{aligned} \hat{c}_a - c_a &= c_a w_a + \sum_i c_i w_i + e \\ &= c_a (l_{aa} - 1) + \sum_i c_i l_{i,a} + e \end{aligned} \quad (53)$$

Unlike the univariate situation, $w_i \neq K_{a,i}$ for multivariate calibration and Equation 53 written to include the selectivity coefficient is

$$\hat{c}_a - c_a = c_a w_a + l_{aa} \sum_i c_i K_{a,i} + e. \quad (54)$$

The selectivity coefficient for a particular interferent and model can also be represented by the ratio of the L_2 norms of the net sensitivity interferent \mathbf{s}_i^* and analyte \mathbf{s}_a^* vectors that are the projections of the respective pure component spectra onto the model vector. With this notation, the selectivity coefficient is written as

$$K_{a,i} = \frac{\|\mathbf{s}_i^*\|}{\|\mathbf{s}_a^*\|}. \quad (55)$$

For comparison to Equation 48, the absolute value of Equation 48 can be used if a negative value is used or results for one of the relaxation parameters.

3.3.2 | Selectivity and sensitivity

The CLS component-wise net sensitivity vector was presented and discussed using Equation 11 can be modified for *r*CLS and others as

$$\mathbf{s}_a^* = \frac{\mathbf{b}_a}{\|\mathbf{b}_a\|^2} l_{aa}, \quad (56)$$

where again the \mathbf{b}_a notation indicates any analyte model vector via CLS, *r*CLS, ILS, *r*ILS, or other method. The scalar selectivity is

$$\begin{aligned}
 SEL_a &= \frac{\|\mathbf{s}_a^*\|}{\|\mathbf{s}_a\|} \\
 &= \frac{l_{aa}}{\|\mathbf{s}_a\| \|\mathbf{b}_a\|}
 \end{aligned} \tag{57}$$

with the corresponding sensitivity written as

$$\begin{aligned}
 SEN_a &= \|\mathbf{s}_a^*\|, \\
 &= \frac{\mathbf{s}_a^T \mathbf{b}_a}{\|\mathbf{b}_a\|}, \\
 &= \frac{l_{aa}}{\|\mathbf{b}_a\|}.
 \end{aligned} \tag{58}$$

Respective interferent selectivity and sensitivity measures SEL_i and SEN_i relative to the analyte regression vector are obtained by using l_{ia} and \mathbf{s}_i instead of l_{aa} and \mathbf{s}_a in Equations 57 and 58. The selectivity measures proposed in Equations 46 and 47 only depend on respective relaxation parameters (deviation from orthogonality for a model). The selectivity measures proposed here depend on both the relaxation parameters and the full sensitivity of the analyte or interferent relative to the particular analyte model. Equivalently, the selectivity defined in Equation 57 for the analyte and the corresponding definition for the interferent are respective $\cos\theta$ values for the angular relationships between pure component analyte and interferent spectra and the model. In the case of r CLS, the corresponding $\cos\theta$ values are computed by $\cos\theta_a = \mathbf{s}_a^T \mathbf{b}_{L,a} / (\|\mathbf{s}_a\| \|\mathbf{b}_{L,a}\|)$ and $\cos\theta_i = \mathbf{s}_i^T \mathbf{b}_{L,a} / (\|\mathbf{s}_i\| \|\mathbf{b}_{L,a}\|)$. To avoid negative SEL and SEN values for the analyte and interferent when a negative value for a relaxation parameter is obtained, absolute values can be used.

In many calibration situations, it is unlikely to know \mathbf{S} or the full \mathbf{C} and ILS is commonly used in these situations. To characterize the model analyte sensitivity, the best approximation is probably the usual $SEN_a = \frac{1}{\|\mathbf{b}_{ILS,a}\|}$. Even though \mathbf{S} may not be known, the interferent or non-analyte space (\mathbf{N}) space can be estimated and used to span the interferent space. In this case, for ILS, it has been suggested and evaluated²⁵ to project the ILS model vector onto the space spanned by \mathbf{N} to form \mathbf{b}_N and define the model selectivity as the cosine of the angle between these 2 vectors computed by $\cos\beta = \frac{\|\mathbf{b}_N\|}{\|\mathbf{b}_{ILS,a}\|}$. A value close to 0 indicates good model selectivity towards the analyte. Equivalently, the ILS model vector could be projected orthogonally to \mathbf{N} to form \mathbf{b}_N^\perp and the model selectivity becomes $\sin\beta = \|\mathbf{b}_N^\perp\| / \|\mathbf{b}_{ILS,a}\|$ with a value close to one denoting good model selectivity for the analyte relative to the \mathbf{N} used.

4 | MEAN SQUARE ERROR OF PREDICTION

Using the approach set out in other works,^{16,21} the expected mean square error of prediction ($\langle MSEP \rangle$) for an arbitrary \mathbf{b}_a model is expressed as

$$\langle MSEP \rangle = \mathbf{w}^T \boldsymbol{\Psi} \mathbf{w} + \mathbf{b}_a^T \boldsymbol{\Sigma} \mathbf{b}_a, \tag{59}$$

where $\boldsymbol{\Psi}$ denotes the expected concentration covariance for the population of samples being considered, $\boldsymbol{\Sigma}$ symbolizes the expected measurement error covariance, and \mathbf{w} is the $(N \times 1)$ vector with analyte value $w_a = l_{aa} - 1$ and respective values for each interferent are $w_i = l_{ia}$. The first term in Equation 59 is termed bias, and the second term is the variance contribution. With this derivation of $\langle MSEP \rangle$, the influence of the relaxation parameters in the bias contribution are directly obtainable. The influence of the relaxation parameters on variance for r CLS can be distinguished from Equation 59 by replacing the generic \mathbf{b}_a with $\mathbf{b}_{L,a}$ calculated from Equation 19.

The bias term in Equation 59 depends on the selectivity and sensitivity of the calibration model and features of the population samples being predicted, eg, relative amounts of the interferent compared to the analyte. However, in the CLS case with full orthogonality, \mathbf{w}_G is a zero vector causing removal of the bias term. Calculated in this paper is the MSEP for the analyte based on observed values for $\boldsymbol{\Psi} = \mathbf{C}^T \mathbf{C} / (M)$ and $\mathbf{w}(\mathbf{l}_a)$.

5 | EXPERIMENTAL

5.1 | Algorithms and data preprocessing

All algorithms were written by the authors using MATLAB 8.1 (The Math Works, Natick, Massachusetts). Calibration samples are mean-centered relative to the calibration set. New samples are mean-centered to the calibration set mean prior to prediction. For simplicity, $\Sigma = \sigma^2 \mathbf{I}$ and $\sigma^2 = 1$ are assumed when computing MSEF.

5.2 | The r CLS relaxation values

Graphical results presented are based on varying the l_{aa} relaxation parameter value from 2×10^{-17} to 1.98 with the first value being 2×10^{-17} and the second starting at 0.02 incrementing to 1.98 by 0.02 for 100 values. Simultaneously, l_{ia} values range from -0.975 to 1.000 in increments of 0.025 for 80 values. For example, when computing the r CLS models for an analyte by Equation 20, 8000 $\mathbf{b}_{L,a}$ relaxed regression vectors are obtained. The CLS regression vector is calculated when $l_{aa} = 1$ and $l_{ia} = 0$. Similarly, the corresponding 8000 $\mathbf{b}_{L,i}$ interferent relaxed regression vectors are determined by Equation 20. The analyte and interferent relaxed regression vectors are then used to compute respective measures of model quality for the graphical assessments. The \mathbf{B}_{CLS} regression vectors used in Equation 20 are obtained from Equation 14 using estimated matrix effected pure component spectra from Equation 13.

5.3 | Near infrared data

In addition to pure component methanol and water samples, 9 mixtures of methanol and water were prepared with methanol concentrations in increments of 10% (by volume) ending⁴⁰ at 90%. The water concentrations were such that summed sample concentrations are 100%. Spectra of the 9 mixtures plus pure water and methanol were measured in a 0.5-mm flow cell using an NIRSystem model 6500 NIR spectrophotometer. Spectra were recorded from 1100 to 2500 nm in 2-nm increments for 700 points per spectrum. All samples were measured over 1 hour. The spectra are plotted in Figure 6. Methanol is the analyte, and due to the sample set size, leave-one-out cross-validation (LOOCV) was used for CLS and r CLS modeling. For all studies, only the 9 mixtures were used. As noted previously, estimated matrix effected pure component were calculated from Equation 13 using the 9 mixtures.

6 | RESULTS AND DISCUSSION

To graphically assess the trade-offs between different measures of model quality (bias, variance, MSEF, $K_{a,i}$, sample-wise prediction errors, etc) for r CLS, a simple two-component system (an analyte and interferent) was studied. In this case, only 2 relaxation parameters are needed (l_{aa} and l_{ia}) and observations from images of model quality measures are possible as the 2 relaxation parameters vary across a range of values. See Section 5 for ranges. With more than 2 analytes, such images are not possible.

6.1 | Prediction error, bias, and variance

Shown in Figure 7 are the mean values from LOOCV for the root mean square error of calibration (RMSEC), RMSE of cross-validation (RMSECV), MSEF, bias, and the model vector L_2 norms ($\|\mathbf{b}_{L,a}\|$). The RMSEC, RMSECV, and bias images indicate

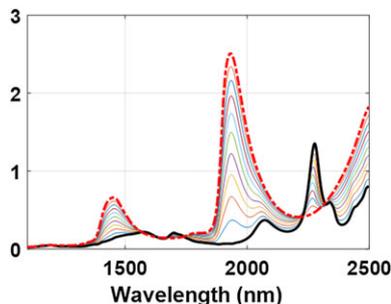


FIGURE 6 Plot of sample mixtures and pure components methanol (black solid) and water (red dash)

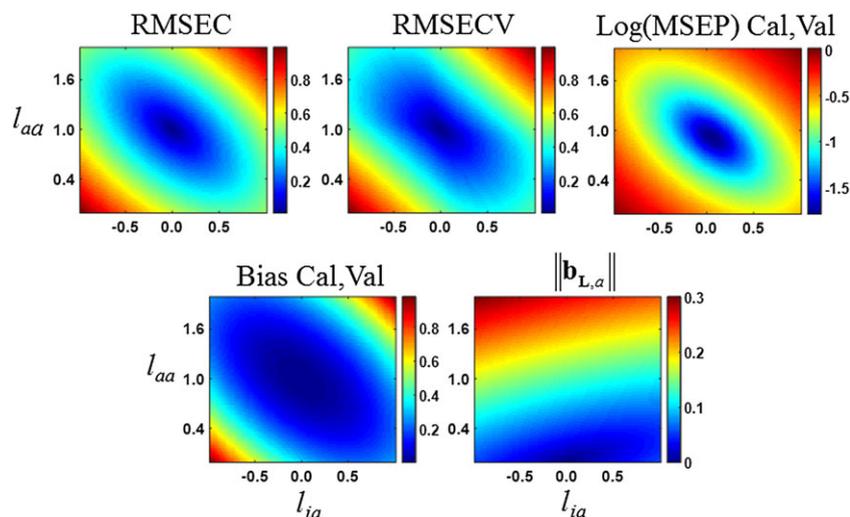


FIGURE 7 Mean images of analyte methanol root mean square error of calibration (RMSEC), RMSE of cross-validation (RMSECV), log(MSEP), bias, and the model vector L_2 norm $\|\mathbf{b}_{L,a}\|$. Respective calibration and validation bias and log (MSEP) mean images are the same and only one is shown to represent calibration (Cal) and validation (Val)

that no relaxation from CLS ($l_{aa} = 1$ and $l_{ia} = 0$) can provide smaller prediction errors and bias over all the samples. However, the MSEP image (log values displayed) shows the minimum MSEP occurs for a slightly r CLS model. In this case, the relaxation parameters for the r CLS model $\mathbf{b}_{L,a} = \mathbf{b}_{CLS,a} \mathbf{l}_a$ at the minimum MSEP are $l_{aa} = 0.920$ and $l_{ia} = 0.050$. Lower MSEP values occur for r CLS because the minimum model vector L_2 norm shown in Figure 7 is at an r CLS model with low l_{aa} and l_{ia} values. It should be noted that mean MSEP images over the LOOCV images are the same for the calibration and validation sets and hence, only one is shown in Figure 7. This statement is also true for the respective mean bias images. However, while the individual sample-wise RMSEC and $\|\mathbf{b}_{L,a}\|$ images are essentially the same across the LOOCV, sample-wise RMSECV, bias, and MSEP images do vary depending on the sample left out.

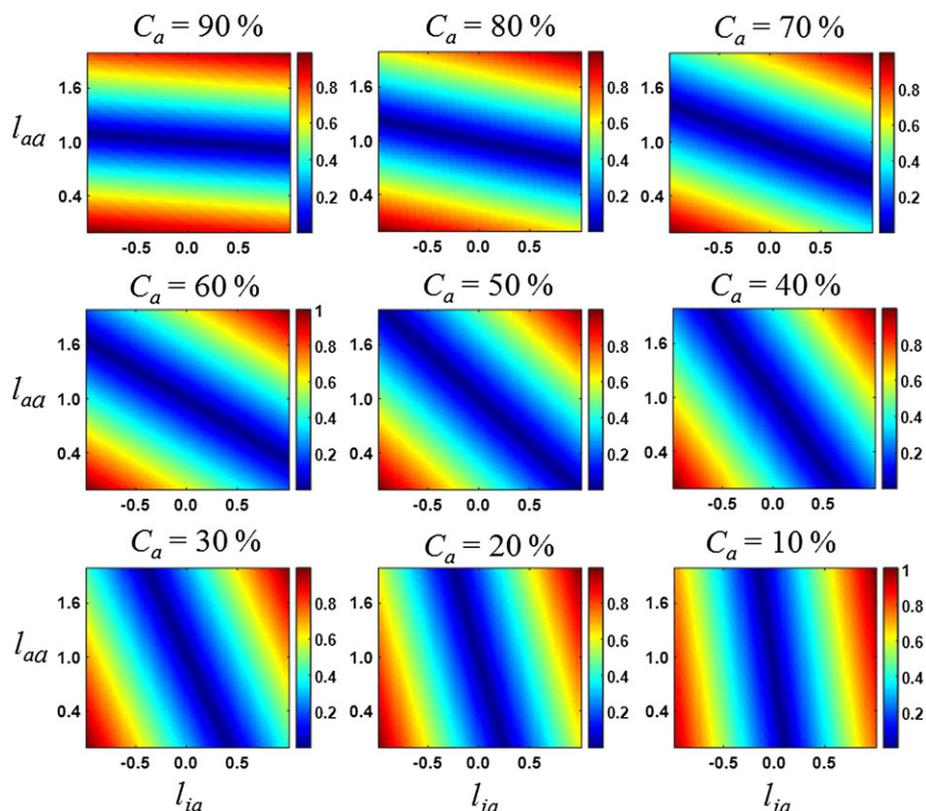


FIGURE 8 Sample-wise images of analyte methanol root mean square error of cross-validation for each sample left out. The c_a values correspond to the sample left out

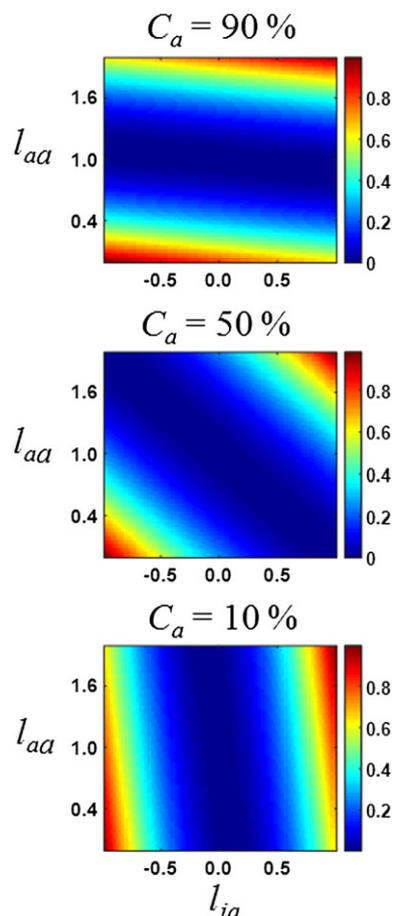


FIGURE 9 Three of the 9 sample-wise analyte methanol validation bias images when the respective sample (c_a value) is left out as the validation sample

Inspection of the sample-wise RMSECV and bias images shown, respectively, in Figures 8 and 9 reveals that RMSECV and bias values track similarly as the relaxation parameter values vary. Because of this similarity, only 3 of the 9 bias images are represented. Also noteworthy is that for a given sample, many models will predict equivalently agreeing with previous work.^{22,24} For the situations presented, the $\mathbf{b}_{L,a}$ models in the deep blue zones primarily differ by respective L_2 norm values (magnitude), not by changes in shape (direction). Figures 8 and 9 show that when the interferent to analyte ratio is low, acceptable r CLS predictions are possible when the l_{aa} values are nearly 1 regardless of the value of l_{ia} . Conversely, when the analyte to interferent ratio is high, the l_{aa} values can now range while the value of l_{ia} must be close to 0 for an acceptable prediction error.

From Figure 8, it is observed for all samples that the orthogonality constraints can be relaxed and increases in prediction errors are obtained because of adding bias to the prediction relative to the CLS prediction at $l_{aa} = 1$ and l_{ia} with a variance reduction (the $\|\mathbf{b}_{L,a}\|$ image in Figure 7). However, most notable, it is also possible to achieve an equivalent prediction error with a reduction in the variance. In this case, l_{ia} increases in value relative to 0 with a corresponding decrease in the l_{aa} value from 1. Conversely, equivalent predictions errors are possible with increases in variance (undesirable solutions) where now the l_{ia} values are more negative with corresponding increases in the l_{aa} value above 1.

The images in Figures 9 and 7 reveal that the sample-wise bias/variance trade-off has several potential combinations depending on the interferent to analyte concentration ratio. Possible are increased bias for a decrease in variance (a goal of r CLS, but the tradeoff should not be significant), a bias increase for an increase in variance (undesirable models), and lastly, apparently, no change in bias accompanied by decrease in variance (a desirable model).

Displayed in Figure 10 are the sample-wise MSEF images. The trends are like those seen in Figure 8 for the RMSECV images with the exception that each r CLS model is unique relative to the MSEF. For all samples, the orthogonality constraints can be relaxed to obtain reduced MSEF values. Specifically, reducing l_{aa} values from 1 while increasing the l_{ia} values from 0. Thus, while each sample can be predicted with low errors and bias by several respective r CLS model, only one sample-wise r CLS model has the smallest MSEF.

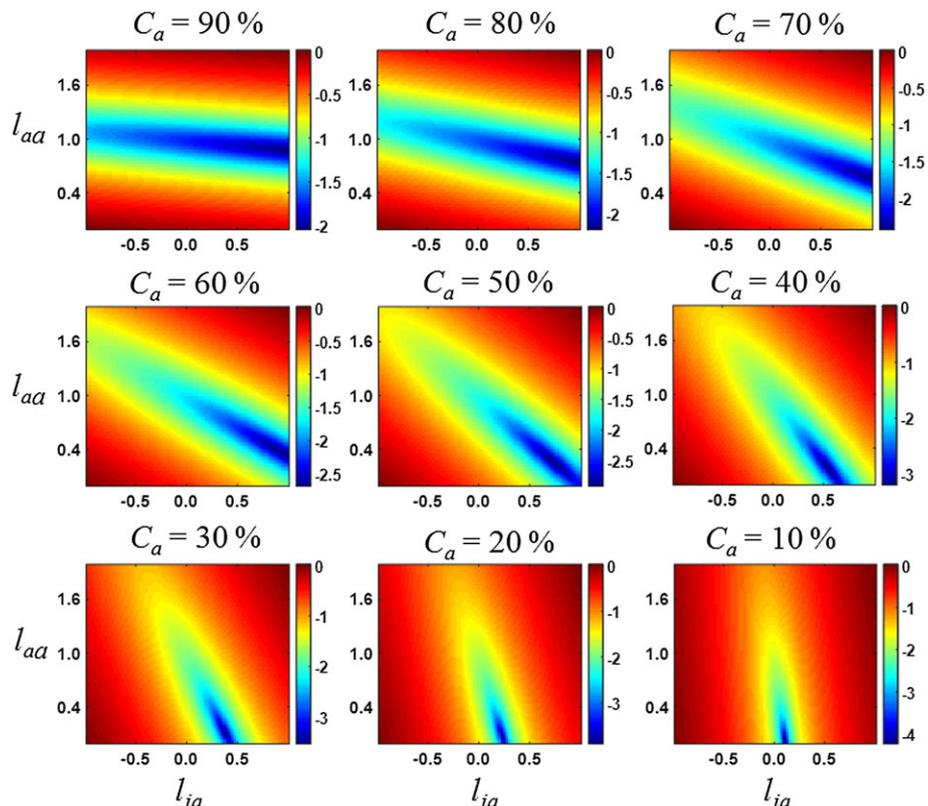


FIGURE 10 Sample-wise images of analyte methanol log(MSEP) for each validation sample left out. The c_a values correspond to the sample left out

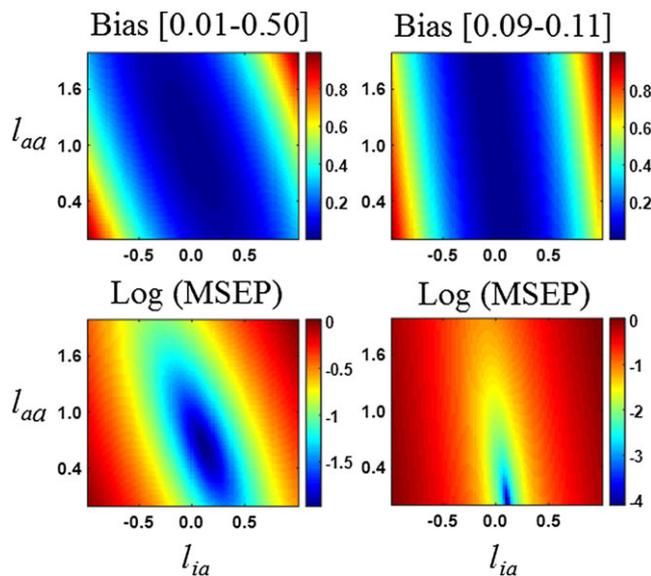


FIGURE 11 Mean calibration analyte methanol bias and log(MSEP) images for two simulated calibration sets with reduced analyte concentration ranges (indicated on top row of images) for the validation sample with methanol concentration at 10% and water at 90%

Because of the sample-wise variations, 2 simulated calibration ranges were made for the validation sample with the greatest interferent to analyte concentration ratio (with the interferent at 90% and the analyte at 10%). Shown on the left in Figure 11 is one calibration situation where the 8 samples in **C** have the analyte concentration equally incremented from 1% to 50% instead of from 20% to 90% as in Figures 9 and 10 to more closely span the analyte concentration for this particular validation sample. As with the full calibration range, the interferent concentration varies to maintain a concentration sum of 100% in the 8 simulated calibration samples. The bias and MSEP images are obtained using Equation 59 with the same $\mathbf{b}_{L,a}$ and l_{aa} and l_{ia} ranges

used with full calibration range. The only change is the analyte and interferent reduced C ranges and hence, a concentration covariance matrix more closely centered around the validation sample. The second simulated calibration case is shown on the right of Figure 11 where now the analyte concentration range is further reduced to more closely span the analyte concentration in the same validation sample. In this case, the analyte concentration varies from 1% to 11% in equal increments and the interferent concentration, respectively, varies to maintain the concentration sum at 100% in the eight calibration samples. In comparing the bias and MSEP images in Figures 9 and 10 for the full calibration range to those in Figure 11, it is observed that as the calibration concentration range decreases to more closely bracket the validation sample situation, the calibration bias and MSEP images converge to those of the specific validation sample. Using local calibration ranges for specific samples allows for better matrix matched calibration samples to the specific samples, thereby leveraging the advantages of r CLS. Using the mean MSEP of the local calibration set now allows proper selection of the relaxation terms for the validation sample where the MSEP are small. Thus, the more a calibration set is descriptive of a new sample for quantitation, the more likely the selection of the best l_a by the MSEP for the calibration set will match the optimal l_a for the new sample.

6.2 | Sensitivity and selectivity

By using the relaxed sensitivity measures in Equation 58, the trade-off between maximizing the model sensitivity towards the analyte and minimizing the model sensitivity to the interferent is possible. This trade-off is demonstrated in Figure 12 where it is also observed from the top image that the traditional sensitivity measure does not provide this trade-off information. In conjunction with prediction errors in Figure 8 and MSEP values in Figure 10, r CLS models associated with greater sensitivity have lower prediction errors and MSEP values. It is also seen that if the sensitivity is low for the analyte, predictions of new samples with acceptable prediction error are still possible. However, if the interferent to analyte ratio is high, then the MSEP can be poor

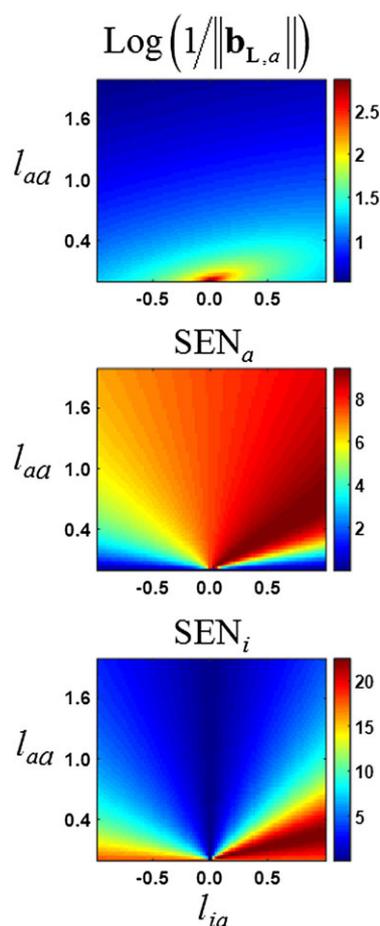


FIGURE 12 Mean sensitivity images relative to the analyte methanol relaxed classical least squares (r CLS) calibration models across the leave-one-out cross-validation (LOOCV) for the general measure $1/\|\mathbf{b}_{L,a}\|$ (shown as $\log(1/\|\mathbf{b}_{L,a}\|)$) and the relaxed version (Equation 58) relative to respective relaxation parameter values (absolute for SEN_i due to the allowed negative values for l_{ia})

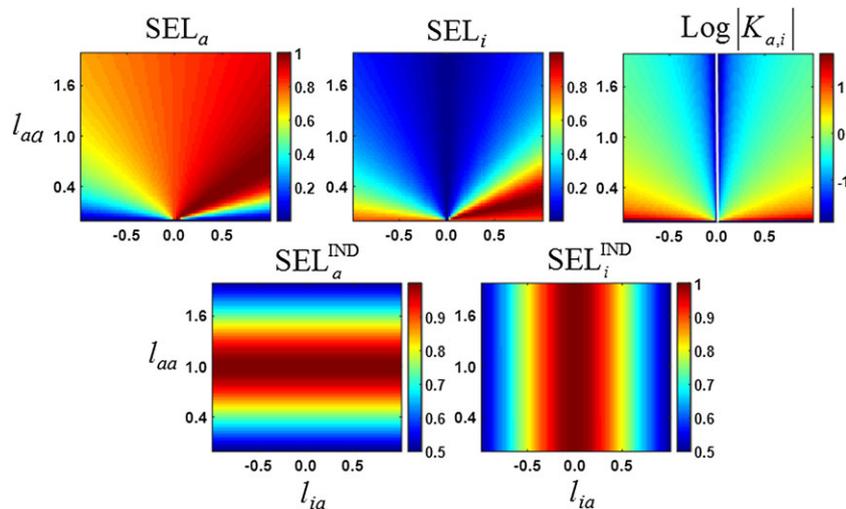


FIGURE 13 Mean selectivity (Equation 57) images relative to the analyte methanol relaxed classical least squares (r CLS) calibration models across the leave-one-out cross-validation (LOOCV) (absolute for SEL_i due to the allowed negative values for l_{ia}). Also shown is the mean $\log(|K_{a,i}|)$ with white at $l_{ia} = 0.0$. The selectivity measures from Brown and Ridder¹⁶ are the 2 images in the bottom row (Equations 46 and 47, respectively)

compared to when the interferent to analyte concentration ratio is low. Because the mean sensitivity values (absolute for SEN_i due to the allowed negative values for l_{ia}) presented in Figure 12 are essentially the same as the sample-wise images, only the mean image is shown.

When the relaxed selectivity measures defined in Equation 57 are used, the corresponding images in Figure 13 (absolute for SEL_i due to the allowed negative values for l_{ia}) are similar to the respective sensitivity trends shown in Figure 12 except the selectivity values range between 1 and 0. Similar interpretations to the trends made for RMSECV and MSEP and sensitivity can be stated for the selectivity values. Conversely, the selectivity measures defined by Equations 46 and 47 are shown in the bottom row of Figure 13 and the trends with RMSECV and MSEP are not obvious. These selectivity definitions only depend on the respective relaxation parameters (deviation from orthogonality). The selectivity measures proposed here not only depend on the relaxation parameter values but the selectivity values are also dictated by sensitivity of the analyte relative to the model. Depending on the interferent to analyte concentration ratio, the selectivity may be low for the analyte, but predictions of samples with acceptable prediction errors are still possible. This trade-off is well characterized by the selectivity coefficient proposed in Equations 48 and 55 and imaged in Figure 13. As with the sensitivity images in Figure 12, the sample-wise selectivity values shown in Figure 13 are essentially the same as those for the mean images.

The selectivity coefficient values displayed in Figure 13 in conjunction with the RMSECV images in Figure 8 show that if the interferent concentration is large relative to the analyte concentration, acceptable predictions are possible when the selectivity coefficient is low. It is possible to also have low prediction errors where the selectivity coefficients are large. Low prediction errors can occur because the l_{aa} values are small thereby compensating for the large interferent prediction bias. Again, the sample-wise selectivity images are essentially the same as the mean image in Figure 13 and only the mean image is shown.

7 | CONCLUSION

With r CLS, a family of CLS models can be formed. Some r CLS models leverage interferent concentrations into the prediction to increase bias with an improvement in variance. Other combinations of relaxation parameters are possible that vary the degree of the bias/variance trade-off and hence, the MSEP, ie, numerous models can be formed that predict well, but only one has the lowest MSEP relative to the variance. It was shown that the more local the calibration set (concentration ranges closely bracket the analyte and interferent concentrations), the greater opportunity for improvement of r CLS over CLS. An r ILS strategy was presented to form a family of ILS models. However, for the main reason ILS is used (all concentrations of spectrally responding species are not known), this approach is not practical. Regardless, a selectivity coefficient is defined that is applicable to univariate and multivariate calibration. With such a coefficient, it is possible to assess the predictability of a sample given knowledge of expected analyte and interferent concentrations.

If the goal is to obtain a model that predicts well and targets the CLS orthogonality constraints, then a penalty regression approach can assist in this. Specifically, the regression vector satisfying the minimization expression

$$\min \left(\|\mathbf{R}\mathbf{b}_a - \mathbf{c}\|^2 + \lambda^2 \|\mathbf{b}_a\|^2 + \eta^2 \|\mathbf{S}\mathbf{b}_a - \mathbf{i}\|^2 \right) \quad (60)$$

is sought where the first 2 terms are the usual RR terms and the third term is an extra orthogonality constraint weighted by the η tuning parameter with \mathbf{i} denoting a column vector of zeros with a 1 at the row position for the pure component analyte spectrum in \mathbf{S} . Alternatively, the \mathbf{i} in Equation 60 could be a vector of target relaxation parameter values. If \mathbf{S} is not known, then non-analyte samples can be used instead giving

$$\min\left(\|\mathbf{R}\mathbf{b}_a - \mathbf{c}\|^2 + \lambda^2\|\mathbf{b}_a\|^2 + \eta^2\|\mathbf{N}\mathbf{b}_a\|^2\right), \quad (61)$$

where \mathbf{N} is a collection of samples without the analyte.⁴¹ For Equations 60 and 61, sought from the regression is a $\mathbf{b}_a = \mathbf{b}^\perp + \mathbf{b}_i$, where \mathbf{b}^\perp denotes the part of \mathbf{b}_a orthogonal to the \mathbf{S} and/or \mathbf{N} and \mathbf{b}_i is the contribution from the interferences.⁴¹ In both Equations 60 and 61, the MSEP is not directly minimized, but with an MSEP penalty included, the solution is no longer directly computable. As demonstrated in Section 6, the more closely the calibration set mimics a particular sample regarding the full sample matrix, the more likely the solutions of Equations 60 or 61 will correspond to a minimum MSEP.

ACKNOWLEDGEMENTS

This material is based upon work supported by the National Science Foundation under grant no. CHE-1111053 (co-funded by MPS Chemistry and the OCI Venture Fund) and under grant no. CHE-1506417 (co-funded by CDS&E) and is gratefully acknowledged by the authors. Funding from the European Community's Framework program for Research and Innovation Horizon 2020 (2014-2020) under grant agreement no. 637232 is also acknowledged. Paul Gemperline is thanked for allowing the use of the NR data set.

REFERENCES

1. Haaland DM, Chambers WB, Keenan MR, Melgaard DK. Multi-window classical least-squares multivariate calibration methods for quantitative ICP-AES analyses. *Appl Spectrosc.* 2000;54:1291-1302.
2. Haaland DM, Melgaard DK. New prediction-augmented classical least squares (PACLS) methods: Application to unmodeled interferences. *Appl Spectrosc.* 2000;54:1303-1312.
3. Haaland DM, Melgaard DK. New classical least-squares/partial least squares hybrid algorithm for spectral analyses. *Appl Spectrosc.* 2001;55:1-8.
4. Haaland DM, Melgaard DK. New augmented classical least squares methods for improved quantitative spectral analyses. *Vib Spectrosc.* 2002;29:171-175.
5. Melgaard DK, Haaland DM, Wehlburg CM. Concentration residual augmented classical least squares (CRACLS): A multivariate calibration method with advantages over partial least squares. *Appl Spectrosc.* 2002;56:615-624.
6. Næs T, Isaksson T, Fern T, Davies T. *A User Friendly Guide to Multivariate Calibration and Classification*. Chichester, UK: NIR Publications; 2002.
7. Hansen PC. *Rank-Deficient and Discrete Ill-Posed Problems: Numerical Aspects of Linear Inversion*. Philadelphia, PA: SIAM; 1998.
8. Lorber A. Error propagation and figures of merit for quantification by solving matrix equations. *Anal Chem.* 1986;58:1167-1172.
9. Lorber A, Faber K, Kowalski BR. Net analyte signal calculation in multivariate calibration. *Anal Chem.* 1997;69:1620-1626.
10. Ferré J, Faber NM. Net analyte signal calculation for multivariate calibration. *Chemom Intell Lab Sys.* 2003;69:123-136.
11. Ferré J, Brown SD, Rius FX. Improved calculation of the net analyte signal in inverse multivariate calibration. *J Chemom.* 2001;15:537-553.
12. Faber NKM. Efficient computation of net analyte signal vector in inverse multivariate calibration models. *Anal Chem.* 1998;70:5108-5110.
13. Bro R, Andersen CM. Theory of net analyte signal vectors in inverse regression. *J Chemom.* 2003;17:646-652.
14. Kalivas JH, Lang PM. *Mathematical Analysis of Spectral Orthogonality*. New York, NY: Marcel Dekker; 1994.
15. Faber NM, Ferré J, Boqué R, Kalivas JH. Quantifying selectivity in spectrophotometric multicomponent analysis. *Trends Anal Chem.* 2004;22:352-361.
16. Brown CD, Ridder TD. Framework for multivariate selectivity analysis, part I: Theoretical and practical merits. *Appl Spectrosc.* 2005;59:787-803.
17. Ridder TD, Brown CD. Framework for multivariate selectivity analysis, part II: Experimental applications. *Appl Spectrosc.* 2005;59:804-815.
18. Dorkó Z, Verbić T, Horvai G. Selectivity in analytical chemistry: Two interpretations for univariate methods. *Talanta.* 2015;132:680-684.
19. Dorkó Z, Verbić T, Horvai G. Comparison of the single channel and multichannel (multivariate) concepts of selectivity in analytical chemistry. *Talanta.* 2015;139:40-49.
20. Olivieri AC. Analytical figures of merit: From univariate to multiway calibration. *Chem Rev.* 2014;114:5358-5378.

21. Brown CD. Discordance between net analyte signal theory and practical multivariate calibration. *Anal Chem.* 2004;76:4364-4373.
22. Brown CD, Green RL. Critical factors limiting the interpretation of regression vectors in multivariate calibration. *Trends Anal Chem.* 2009;28:506-514.
23. Seasholtz MB, Kowalski BR. Qualitative information from multivariate calibration models. *Appl Spectrosc.* 1990;44:1337-1348.
24. Mosteller F, Tukey JW. *Data Analysis and Regression: A Second Course in Statistics.* Reading, MA: Addison-Wesley; 1977.
25. Kalivas JH, Palmer J. Characterizing multivariate calibration tradeoffs (bias, variance, selectivity, and sensitivity) to select model tuning parameters. *J Chemom.* 2014;28:347-357.
26. Harvey D. *Modern Analytical Chemistry.* Boston, MA: McGraw Hill; 2000.
27. Skoog DA, Holler FJ, Crouch SR. *Principles of Instrumental Analysis.* Seventh ed. Pacific Grove, CA: Thomson Brooks/Cole; 2018.
28. Pungor E, Tóth K, Hrabéczy-Páll A. Selectivity coefficients of ion-selective electrodes. *Pure Appl Chem.* 1979;51:1913-1980.
29. Umezawa Y, Umezawa K, Sato H. Selectivity coefficients for ion-selective electrodes: Recommended methods for reporting $K_{A,B}^{\text{pot}}$ values. *Pure Appl Chem.* 1979;67:507-518.
30. Bennett KH, Cook KD, Seebach GL. Simultaneous analysis of butane isomer mixtures using process mass spectrometry. *J Am Soc Mass Spectrom.* 2000;11:1079-1085.
31. Nadler B, Coifman RR. Partial least squares, Beer's law and the net analyte signal: Statistical modeling and analysis. *J Chemom.* 2005;19:45-54.
32. Vainchtein A. Change of basis, reciprocal basis vectors, covariant and contravariant components of a vector and metric tensor. Math 1550 Lecture Notes, <http://physastro-msci.tripod.com/webonmediacontents/notes1.pdf> Accessed November 23, 2016.
33. Sanchez E, Kowalski BR. Tensorial calibration: I. First-order calibration. *J. Chemom.* 1988;2:247-263.
34. Borisenko AI, Tarapov IE. In: Silverman RA, ed. *Vector and Tensor Analysis with Applications.* New York, NY: Dover Publications; 1968.
35. Xu L, Jiang JH, Zhou YP, Wu HL, Shen GL, Yu RQ. MCCV stacked regression for model combination and fast spectral interval selection in multivariate calibration. *Chemom Intel Lab Syst.* 2007;876:226-230.
36. Ni W, Brown SD, Man R. Stacked partial least squares regression analysis for spectral calibration and prediction. *J Chemom.* 2009;23:505-515.
37. Li Z, Lv J, Si G, Zhang Y, Wang Q, Liu S. An improved ensemble model for the quantitative analysis of infrared spectra. *Chemom Intel Lab Syst.* 2015;146:211-220.
38. Fragoso W, Allegrini F, Olivieri AC. A new consistent parameter for measuring the quality of a multivariate analytical methods: Generalized analytical sensitivity. *Anal Chim Acta.* 2016;933:43-49.
39. Currie LA. Nomenclature in evaluation of analytical methods including detection and quantification capabilities. *Pure Appl Chem.* 1995;67:1699-1723.
40. Kalivas JH, Gemperline PJ. Calibration. In: Gemperline PJ, ed. *Practical Guide to Chemometrics.* Second ed. Boca Raton, FL: CRC Press Taylor & Francis Group; 2006.
41. Andries E, Kalivas JH. Interrelationships between generalized Tikhonov regularization, generalized net analyte signal, and generalized least squares for desensitizing a multivariate calibration to interferences. *J Chemom.* 2013;27:126-140.

How to cite this article: Kalivas JH, Ferré J, Tencate AJ. Selectivity-relaxed classical and inverse least squares calibration and selectivity measures with a unified selectivity coefficient. *Journal of Chemometrics.* 2017;31:e2925. <https://doi.org/10.1002/cem.2925>