Tackling calibration problems in High Throughput Experimentation

ARTICLE · JANUARY 2005

READS
12

4 AUTHORS, INCLUDING:

Susana C. Cruz
Metrohm Applikon
9 PUBLICATIONS **68** CITATIONS
SEE PROFILE

Johan Westerhuis
University of Amsterdam
123 PUBLICATIONS **3,924** CITATIONS
SEE PROFILE

Anal. Chem. **2005,** *77,* 2227−2234

Tackling Calibration Problems of Spectroscopic Analysis in High-Throughput Experimentation

Susana C. Cruz,† Gadi Rothenberg,† Johan A. Westerhuis,*,‡ and Age K. Smilde‡

van t' Hoff Institute for Molecular Sciences and Swammerdam Institute for Life Sciences, University of Amsterdam, Nieuwe Achtergracht 166, 1018 WV Amsterdam, The Netherlands

High-throughput experimentation and screening methods are changing work flows and creating new possibilities in biochemistry, organometallic chemistry, and catalysis. However, many high-throughput systems rely on off-line chromatography methods that shift the bottleneck to the analysis stage. On-line or at-line spectroscopic analysis is an attractive alternative. It is fast, noninvasive, and nondestructive and requires no sample handling. The disadvantage is that spectroscopic calibration is time-consuming and complex. Ideally, the calibration model should give reliable predictions while keeping the number of calibration samples to a minimum. In this paper, we employ the net analyte signal approach to build a calibration model for Fourier transform near-infrared measurements, using a minimum number of calibration samples based on blank samples. This approach fits very well to high-throughput setups. With this approach, we can reduce the number of calibration samples to the number of chemical components in the system. Thus, the question is no longer *how many* but *which type* of calibration samples should one include in the model to obtain reliable predictions. Various calibration models are tested using Monte Carlo simulations, and the results are compared with experimental data for palladium-catalyzed Heck cross-coupling.

The last 150 years have witnessed radical changes in chemistry as far as research areas are concerned, but the basic laboratory work flow is still quite similar to that employed by Perkin in the 19th century. The incorporation of high-throughput experimentation (HTE) work flows in research laboratories, however, is changing this. HTE and combinatorial chemistry tools have surmounted the science barrier and became enabling technologies.[1,2] Robot systems can now perform thousands of experiments per day, yielding mind-boggling amounts of data. This technological revolution is not only about "performing experiments faster". It has important psychological consequences: the value of the basic scientific unit operation (the laboratory experiment) has changed, and chemists must accept this change if they are to make good use of HTE tools. In many cases, the problem has shifted from performing the experiments to monitoring them and analyzing the data. This is especially important when multiple samples must be taken, e.g., in kinetic studies and catalysis research.[3] The coupling of parallel reactors and chromatographic analysis often creates a bottleneck.[4,5] Alternatively, one can use in situ or operando spectroscopy, which features fast and noninvasive reaction sampling.[6] Until now, however, on-line spectroscopic monitoring has been limited to simple reaction mixtures, mainly because of problems associated with spectral overlapping. Deconvolution of the spectra is possible by means of multivariate calibration techniques that enable the efficient extraction of information concerning analytes of interest from multicomponent mixtures.[7] As we recently showed, it is possible to combine in situ spectroscopy with multivariate calibration techniques to monitor the kinetics of complex reactions, such as C−C cross-coupling[8,9] and metal nanocluster formation.[10] However, multivariate calibration requires a large number of calibration samples, the collection of which is often difficult and time-consuming, creating a new bottleneck.

Combinatorial catalysis, for example, is composed of two important stages: catalyst screening/discovery and catalyst optimization. In the screening stage, a large set of catalysts is tested under a few experimental conditions. In this stage, precision is not very important, since the main question here is, "Is the catalyst active?—yes/no". Positive answers to this question will proceed further to the optimization stage, where more experimental parameters are included. If one wants to use spectroscopic analysis in this type of system, the calibration model should be adaptable to both "discovery" and "optimization" modes.

The question is how can one analyze complex reaction mixtures in situ with a minimal calibration effort. The selection of calibration samples is a topic that has received the attention of

* Corresponding auhor. E-mail: j.a.westerhuis@uva.nl.
† van t' Hoff Institute for Molecular Sciences.
‡ Swammerdam Institute for Life Sciences.

(1) Pescarmona, P. P.; van der Waal, J. C.; Maxwell, I. E.; Maschmeyer, T. *Catal. Lett.* **1999**, *63*, 1−11.
(2) Reetz, M. T. *Angew. Chem., Int. Ed.* **2001**, *40*, 284−310.
(3) Lavastre, O. *Actual. Chim.* **2000**, 42−45.
(4) Boelens, H. F. M.; Iron, D.; Westerhuis, J. A.; Rothenberg, G. *Chem. Eur. J.* **2003**, *9*, 3876−3881.
(5) Rothenberg, G.; Boelens, H. F. M.; Iron, D.; Westerhuis, J. A. *Chim. Oggi* **2003**, *21*, 80−83.
(6) Rothenberg, G.; Boelens, H. F. M.; Iron, D.; Westerhuis, J. A. *Catal. Today* **2003**, *81*, 359−367.
(7) Bro, R. *Anal. Chim. Acta* **2003**, *500*, 185−194.
(8) Cruz, S. C.; Aarnoutse, P. J.; Rothenberg, G.; Westerhuis, J. A.; Smilde, A. K.; Bliek, A. *Phys. Chem. Chem. Phys.* **2003**, *5*, 4455−4460.
(9) Rothenberg, G.; Cruz, S. C.; Van Strijdonck, G. P. F.; Hoefsloot, H. C. J. *Adv. Synth. Catal.* **2004**, *346*, 467−473.
(10) Wang, J.; Boelens, H. F. M.; Thathagar, M. B.; Rothenberg, G. *ChemPhysChem* **2004**, *5*, 93−98.

10.1021/ac048421c CCC: $30.25
Published on Web 02/25/2005

several researchers.[11−14] The goal is to choose the smallest set of prepared calibration samples that can be used without significantly compromising the model's prediction ability. In many situations, a large number of samples from the system has been obtained, and the exact analyte concentration in the samples is determined using a reference method (often GC). Calibration samples are usually selected from a list of possible candidates using methods such as Kennard−Stone design,[11] simulated annealing,[12] genetic algorithms,[13] successive projections algorithms,[14] or random selection.[15]

In this paper, we investigate a different situation, wherein prior to analysis a number of calibration samples are being prepared by weighing the appropriate amounts of each component, and then the prepared sample is measured. Here we use the net analyte signal (NAS) approach with blank samples, a combination that fits well to HT setups. Two sources of error are important in this situation: the concentration error and the spectral error. We examine strategies to minimize the calibration effort in in situ spectroscopic analysis using both computer simulations and the Heck cross-coupling as an experimental example.

THEORY

NAS Approach. There are several techniques for multivariate calibration, the ones most frequently used being classical least squares, principal component regression, and partial least squares (PLS).[16] Multivariate calibration models are usually built from large sets of samples that include all the possible variability in the data. The collection of the spectra of these samples is usually difficult and time-consuming. Calibration free monitoring[17−23] is appealing, but these methods are not easy to apply since they often give no unique solution and several constraints are required.

Here we chose to use the NAS approach, which is a fairly new technique used in calibration. It was first introduced by Lorber[24] in 1986 and was subsequently applied in several studies.[8,10,25−34] The NAS for a given component is defined as "the part of its spectrum which is orthogonal to the spectra of the other components" and is uniquely related to the concentration of the analyte of interest.[24]

The reason we chose the NAS approach in this study is that applies well to a HT environment. The separation of the interference space from the analyte space (see explanation below) using blank samples makes it easy to extend the calibration model to new situations. One does not need to know in advance the future composition of the system. Rather, by using blank samples (samples that contain a mixture of interferences but no analyte), it is possible to add new interferences sequentially to the calibration model. For example, one can calibrate a system containing one analyte and five interferences (of which, say, three are solvents), using the models shown below. Then, if one wants to test five other solvents in this system, one additional sample of each of these solvents is sufficient to evaluate the model's performance.

The absorbance spectrum of a mixture of K spectroscopic active substances ($k = 1, ..., K$) is measured at J wavenumbers. Assume that the analyte of interest, analyte k, is one component in this mixture. All the remaining components are called the interferences. Each spectrum represents a vector in the J dimensional space, and its length and direction translate, respectively, the spectrum's intensity and shape.

Let $\mathbf{R}$ be the matrix composed of spectra with the analyte of interest, formed by the absorbance spectra of samples containing the analyte and $\mathbf{R}_{-k}$ the matrix composed by the interferences (formed by spectra of samples that do not contain analyte, i.e., blank samples). A spectrum containing analyte k can then be decomposed, by definition, into two orthogonal parts: one part orthogonal to the interference space and one part that lies in the interference space. The latter can be described by a linear combination of the interferences. The unique direction useful for quantification of analyte k, which defines the NAS direction, is therefore orthogonal to $\mathbf{R}_{-k}$.

Temperature is usually considered as an interference, and its variability should be included in the calibration model. However, our previous studies on this chemical system showed that, in this case, temperature has little effect on the spectra.[8]

NAS-based calibration techniques differ in the manner in which the matrix $\mathbf{R}_{-k}$ is defined. Lorber, for example, used pure component spectra.[25] The problem with this approach is that the samples are very far from reaction conditions. Furthermore, the spectrum of a pure component is often different from the spectra at low concentration in a solvent. Also, the Beer−Lambert law is only valid for low/medium concentrations. Instead of pure spectra, Goicoechea and Olivieri used samples with analyte k to define $\mathbf{R}_{-k}$.[30] The analyte part of each spectrum is removed by scaling all samples to equal analyte content and then a centering step removes the analyte contribution.

(11) Wu, W.; Walczak, B.; Massart, D. L.; Heuerding, S.; Erni, F.; Last, I. R.; Prebble, K. A. *Chemom. Intell. Lab. Syst.* **1996**, *33*, 35−46.
(12) Kalivas, J. H. *J. Chemom.* **1991**, *5*, 37−48.
(13) Broudiscou, A.; Leardi, R.; PhanTanLuu, R. *Chemom. Intell. Lab. Syst.* **1996**, *35*, 105−116.
(14) Filho, H. A. D.; Galvão, R. K. H.; Araújo, M. C. U.; Silva, E. C.; Saldanha, T. C. B.; José, G. E.; Pasquini, C.; Raimundo, I. M., Jr.; Rohwedder, J. R. *Chemom. Intell. Lab. Syst.* **2004**, *72*, 83−91.
(15) Miller, J. N.; Miller, J. C. *Statistics and Chemometrics for Analytical Chemistry*, 4th ed.; Prentice Hall: Upper Saddle River, NJ, 2000.
(16) Martens, H.; Næs, T. *Multivariate Calibration*; John Wiley & Sons Ltd.: New York, 1989.
(17) Sans, D.; Nomen, R.; Sempere, J. *Comput. Chem. Eng.* **1997**, *21*, S631−S636.
(18) Jiang, J. H.; Ozaki, Y. *Appl. Spectrosc. Rev.* **2002**, *37*, 321−345.
(19) Ozaki, Y.; Sasic, S.; Jiang, J. H.; Siesler, H. W. *Macromol. Symp.* **2002**, *184*, 229−247.
(20) Sasic, S.; Kita, Y.; Furukawa, T.; Watari, M.; Siesler, H. W.; Ozaki, Y. *Analyst* **2000**, *125*, 2315−2321.
(21) Windig, W.; Antalek, B.; Sorriero, L. J.; Bijlsma, S.; Louwerse, D. J.; Smilde, A. K. *J. Chemom.* **1999**, *13*, 95−110.
(22) Bijlsma, S.; Smilde, A. K. *Anal. Chim. Acta* **1999**, *396*, 231−240.
(23) Bijlsma, S.; Boelens, H. F. M.; Hoefsloot, H. C. J.; Smilde, A. K. *Anal. Chim. Acta* **2000**, *419*, 197−207.
(24) Lorber, A. *Anal. Chem.* **1986**, *58*, 1167.
(25) Lorber, A.; Faber, K.; Kowalski, B. R. *Anal. Chem.* **1997**, *69*, 1620−1626.
(26) Faber, N. M. *J. Chemom.* **1998**, *12*, 405−409.
(27) Faber, N. M. *Anal. Chem.* **1999**, *71*, 557−565.
(28) Faber, K.; Lorber, A.; Kowalski, B. R. *J. Chemom.* **1997**, *11*, 419−461.
(29) Boque, R.; Larrechi, M. S.; Rius, F. X. *Chemom. Intell. Lab. Syst.* **1999**, *45*, 397−408.
(30) Goicoechea, H. C.; Olivieri, A. C. *TrAC−Trends Anal. Chem.* **2000**, *19*, 599−605.
(31) Olesberg, J. T.; Arnold, M. A.; Hu, S. Y. B.; Wiencek, J. M. *Anal. Chem.* **2000**, *72*, 4985−4990.
(32) Ferre, J.; Faber, N. M. *Chemom. Intell. Lab. Syst.* **2003**, *69*, 123−136.
(33) Skibsted, E. T. S.; Boelens, H. F. M.; Westerhuis, J. A.; Witte, D. T.; Smilde, A. K. *Appl. Spectrosc.* **2004**, *58*, 264−271.
(34) Boelens, H. F. M.; Kok, W. T.; de Noord, O. E.; Smilde, A. K. *Anal. Chem.* **2004**, *76*, 2656−2663.

The approach employed here makes use of spectra of blank samples. In this case, the interference space can be directly determined from $\mathbf{R}_{-k}$.[8,34] This approach allows the calibration effort to be reduced, and even more important, with only a few samples it is possible to assess whether the calibration is feasible at all. A detailed explanation of the NAS approach as used here was published elsewhere.[8] It is important to realize the consequences this approach has on the design of the calibration samples used. The samples used to define the interference space are free of analyte, and thus, the analyte contribution appears only in the samples used to define the analyte space.

After finding the unique direction related with the concentration of analyte, NAS_k, the projections of new spectra on this direction are calibrated to concentration according to eq 1, where c_k is the concentration of analyte k.

$$\mathrm{NAS}_k = bc_k + e \tag{1}$$

The slope b of this calibration line is fitted using calibration samples in a least-squares way and e is the residual. The NAS direction is directly proportional to the concentration of the analyte, so the multivariate problem reduces to a univariate one.

Even in the case of spectral overlap, it is possible to minimize the number of calibration samples to the number of chemical components in the system.[24,25] The samples are divided in two groups: $K - 1$ samples with no analyte, to define the interference space (the blank samples), and one sample containing analyte to define the analyte space. The fundamental question is no longer *how many* samples but rather *which type* of samples: Which $K - 1$ samples should be selected to optimally describe the $K - 1$-dimensional interference space, and which sample should be selected to define the NAS direction in order to have the minimum prediction error for new samples.

Single-Component versus Multicomponent Samples. The reaction takes place in a solvent, and therefore, the calibration samples are conveniently prepared in solvent. One can prepare samples with only one component mixed with the solvent (called "single-component samples") or samples with many or all the components in the solvent ("multicomponent samples"). Figure 1 shows this concept for a system with four components (A, B, C, and solvent). The concept of single-component samples is simpler, as each component (except for the solvent) is present in only one sample. On the other hand, multicomponent samples are more difficult to visualize once each component is present in more samples.

Which type of samples should be included in the calibration model? In high-throughput laboratories, the samples are usually prepared by robots, and therefore, any type of sample is easy to prepare. However, this is not the case for standard analytical laboratories where multicomponent samples take more effort and time to prepare. Also, multicomponent samples contain more sources of error as one weighs more components. Here we will examine which type of samples should be included in the calibration model to obtain reliable concentration predictions.

The most informative way to evaluate which samples to include in the calibration model is to make a feasibility study with experimental data, but that is exhaustive and time-consuming, especially when the experimental design leads to a large number of samples. Another way of solving the problem, demonstrated in this paper, is to assess the solution by use of computer simulations. The simulated calibration samples should be significantly realistic and include error values. Each sample should be generated as a linear combination of the underlying pure component spectra. When these are not available, they can be calculated from measured spectra of mixtures of the components.

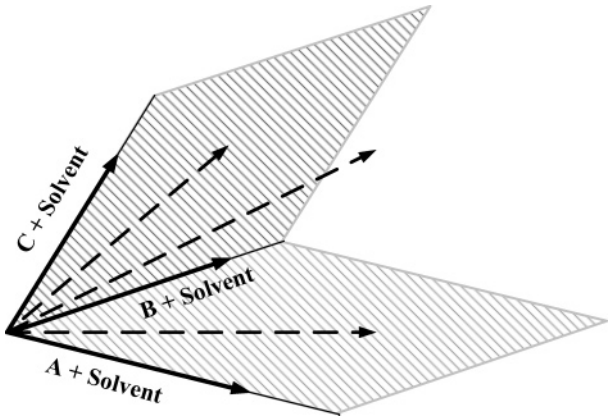


Figure 1. Single-component samples vs multicomponent samples. Single-component samples are formed by one component mixed in solvent (solid lines) and multicomponent samples by two or more components in solvent (broken lines). For simplicity, the figure relates to a system with three components (A, B, C) in a solvent. The solid lines represent single-component samples, e.g. {A + solvent}, and the broken lines represent multicomponent samples, e.g. {A + B + C + solvent}.

We included two sources of error in the simulated calibration samples: a spectral error, which is dependent on the spectrometer used, and a concentration error. The latter deserves special attention: In this study, the calibration samples are prepared before analysis by weighing the appropriate amounts of each component. Let C_k be the concentration, in weight fraction, of component k to be calculated:

$$C_k = m_k / \sum_{k=1}^{K} m_k \tag{2}$$

It is assumed that the masses of the components (m_1, m_2, ..., m_K) are independent. The variance in the concentration, σ_K^2, is then found by error propagation theory as the weighed sum of the variances of the different sources of error in the preparation of the sample (see Appendix).

The calibration samples are simulated as a linear combination of the underlying pure component spectra. The errors are included according to eq 3 (for derivation see Appendix), where N denotes all the components in the mixture except for the solvent ($N = 1$, ..., $K - 1$), and the subscript *sol* stands for the solvent component.

$$\mathbf{S_a} = (\sum_{n=1}^{N}(\mathrm{con}_n + wgEr_n)\cdot\mathbf{P}_n + (\mathrm{con}_{\mathrm{sol}} - \sum_{n=1}^{N} wgEr_n)\cdot\mathbf{P}_{\mathrm{sol}}) + \mathbf{spEr} \tag{3}$$

The vector $\mathbf{S_a}$ is the simulated spectrum of the calibration sample to prepare, con_n is the concentration ratio for component

n (calculated relative to the concentration of a pure sample of this component), $\mathbf{P}_n$ is the pure component spectra of component n estimated from samples at reaction level, and *wgEr* and **spEr** are normal distributed random errors, with zero mean and variances σ_R^2 and SE, respectively. SE is the variance of the spectral error and depends only on the spectrometer used.

A calibration model based on the NAS approach is then developed for each set of K calibration samples and studied with the aid of simulations. The predictive ability of the models can be evaluated by examining the root-mean-square of the error in prediction, RMSEP.

$$\mathrm{RMSEP} = \left(\frac{\sum_{v=1}^{V} (c_v - \hat{c}_v)^2}{V} \right)^{1/2} \tag{4}$$

Here, c_v is the known analyte concentration in sample v, $\hat{c}_v$ is the predicted concentration from the calibration model, and V is the number of validation samples. The spectra from which the concentrations are predicted are a function of the concentration error and the spectral error, as are the RMSEP values. A single simulation will not necessarily provide a representative result, and hence, one runs many simulations to get more realistic results. Each calibration model will, therefore, generate a large number of RMSEP values. To compare the different models, the RMSEP values for each model are condensed into an average value, the average error of prediction (AEP, eq 5), where NR is the number of simulations per model ($nr = 1, \ldots, \mathrm{NR}$).

$$\mathrm{AEP} = \left(\frac{\sum_{nr=1}^{\mathrm{NR}} \mathrm{RMSEP}_{nr}}{\mathrm{NR}} \right) \tag{5}$$

EXPERIMENTAL SECTION

Materials and Instrumentation. Near-infrared spectra were recorded on a Perkin-Elmer Spectrum GX FT-IR instrument from 15 000 to 2700 cm^{-1}. The spectra were recorded as single scans (2.4 s/scan), 2-cm^{-1} resolution and 0.4-cm path length. Unless noted otherwise, all chemicals were purchased from commercial firms and used as received. *n*-Butyl acrylate (NBA) and *n*-methyl pyrrolidinone (NMP) were purified prior to use by filtration over a plug of basic alumina (~150 mesh, 58 Å). The palladium catalyst was synthesized in situ by adding 1 mol % $\mathrm{Pd(OAc)_2}$ and 6 mol % $\mathrm{P(Ph)_3}$ ligand.

Reaction progress was monitored using fiber optics with starting concentrations of ~0.20 M of *n*-butyl acrylate (NBA), an excess of triethylamine ($\mathrm{Et_3N}$), and an excess of iodobenzene (PhI). Each reaction was carried out in a batch reactor at 60 °C (thermoregulated by a water bath). The reaction volume was set to ~40 mL. All liquid components were introduced in the stirred reactor. The ligand was dissolved in NMP in a glass vial, prior to adding the $\mathrm{Pd(OAc)_2}$. This mixture was preheated to 60 °C and added to the reactor to start the reaction.

Data Preprocessing and Wavenumber Selection. FT-NIR measurements contain not only variation related to composition change in the reaction mixture but also variation that is caused by changing physical properties in the mixture, for example, density, viscosity, and temperature. Some of these phenomena introduce offsets or slope changes to the spectra that can be corrected for by spectral preprocessing methods.[27] The simulated data were offset corrected prior to application of the calibration model.

Although multivariate techniques are full spectral methods, the performance of the methods usually improves when a suitable wavelength region is selected.[35] NIR measurements are disturbed, for example, by water vapor[36] in the following areas: 5102−5618 and 6667−7463 cm^{-1}. As the NMP solvent is highly hygroscopic, the spectral range used for quantification[8] was 5950−6300 cm^{-1}.

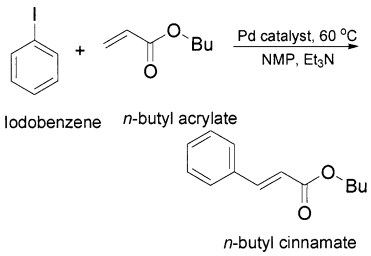
In the following sections, PhI is considered as the analyte, but the same approach can be applied to any of the other components.

Simulations. A computer simulation program for generation of multivariate data, calibration, and validation of the results was developed in Matlab.[37] For each simulation, a number of samples were generated as a linear combination of the underlying pure component spectra (eq 3 above). The spectra were superimposed with normally distributed noise, with zero mean and variances of 5×10^{-5} (au) and 7×10^{-6} (au) for the spectral error and the concentration error, respectively. These spectra were analyzed with the NAS software toolbox developed in our group.[38] Each simulation was run 200 times to ensure sufficiently precise sampling distribution (increasing the number of simulations from 200 to 10 000 gave results within ±6%, confirming the robustness of the models).

Model Validation. The validation of each model is based on 30 samples, designed to mimic reaction conditions. For example, a sample with high concentration of reactants has a low concentration of product. Sample concentrations were in the range of 0.0−0.3 M for all components except for the solvent, which had concentrations of 8.7−10 M.

RESULTS AND DISCUSSION

The model system considered in this study is the Heck reaction between PhI with NBA to give *n*-butyl cinnamate (NBC) in NMP (eq 6). $\mathrm{Et_3N}$ was used as a base.



(6)

Samples of pure NMP and of mixtures of known concentrations of the components were measured using Fourier transform near-infrared (FT-NIR) spectroscopy. Each measured sample can be expressed as a linear combination of the underlying pure component spectra (concentration of each component multiplied by its pure spectrum). The pure component spectra (Figure 2) were then obtained from these equations in a least-squares way.

(35) Xu, L. A.; Schechter, I. *Anal. Chem.* **1997**, *69*, 3722−3730.
(36) Davies, A. M. C. *NIR News* **1992**, *3*, 8−9.
(37) Release 13 version 6.5, MathWorks Inc.
(38) This toolbox can be downloaded from http://www-its.chem.uva.nl/research/pac/.

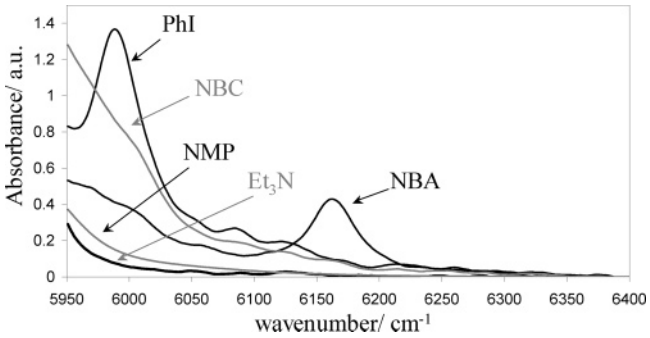


Figure 2. Pure component spectra in the Heck reaction between PhI and NBA estimated from mixture spectra.

Table 1. Design for Studying Sample Composition on the Calibration Model[a]

	analyte space	
interference space	single comoponent	multicomponent
single component	model A	model C
multicomponent	model B	model D

[a] The composition of the samples in the analyte and the interference spaces are varied systematically in models A–D.

Table 2. Sample Composition in the Calibration Sets for Each Model (PhI is the Analyte)

		concentration/M				
entry	model	NBA	PhI	NBC	Et_3N	NMP
1	A	0	0.3	0	0	10.0
2		0.3	0	0	0	10.0
3		0	0	0.3	0	9.8
4		0	0	0	0.3	10.0
5		0	0	0	0	10.4
6	B	0	0.3	0	0	10.0
7		0	0	0.3	0.3	8.7
8		0.3	0	0	0.3	8.7
9		0.3	0	0.3	0	8.7
10		0	0	0	0	10.4
11	C	0.3	0.3	0.1	0.3	8.7
12		0.3	0	0	0	10.0
13		0	0	0.3	0	9.8
14		0	0	0	0.3	10.0
15		0	0	0	0	10.4
16	D	0.3	0.3	0.1	0.3	8.7
17		0	0	0.3	0.3	8.7
18		0.3	0	0	0.3	8.7
19		0.3	0	0.3	0	8.7
20		0	0	0	0	10.4

Which Samples Should One Choose? There are five chemical components in this system (NBA, PhI, NBC, Et_3N, and the solvent, NMP) that have overlapping absorbencies in the 5950–6300-cm^{-1} region. Therefore, the minimum number of calibration samples is five. What type of samples yields the lowest prediction error? Single-component or multicomponent samples? To study the effect of the chosen type of samples in the model, the sample's composition was varied in analyte and interference space as shown in Table 1. When the spectra of blank samples are available, the minimum number of samples defining the analyte space is one and the analyte concentration is measured only once. As in this paper, we use blank samples to define the interference space, four blank samples are used to define this space, and the fifth sample is used to define the NAS direction.

The calibration model must span the concentration of each component. This includes the solvent (which has a higher concentration than any other component). To observe the contribution of each component, the concentration level of the components in the calibration samples should be varied. It is then crucial to include at least two levels in concentration of each component. The solvent concentration is nearly constant. A high concentration sample of solvent is only introduced in the calibration model when adding a pure solvent sample to the interference space. Several models, with and without a sample of pure solvent in the interference space, were tested (not shown here), and it was found that the lowest AEP values are obtained when including a "solvent sample". The chosen combinations for each model include samples with low and high concentration of all components rather than also having intermediate concentration levels. The models below are the combinations with the lowest AEP for each of the groups presented in Table 1.

The composition of the samples in each of the calibration sets is shown in Table 2. Again, note the importance of using only one analyte concentration contribution. Each sample is simulated according to eq 3, using normally distributed random errors with zero mean and variances of 5×10^{-5} (au) and 7×10^{-6} (au) for the spectral and the concentration error variances, respectively. For each model, the first samples (entries 1, 6, 11, and 16) are the analyte samples, and the other samples are used to define the interference space. Also note that models A and B have the same analyte sample as have models C and D. Furthermore, models A and C have the same interference samples as have models B and D. For calibration, the samples are divided into analyte and interference spaces. The validation samples, on the other hand, always contain analyte and are made to mimic reaction conditions. The analyte samples are as close as possible from reaction conditions. In reaction conditions, a high concentration of product is never present in a sample with high concentration of reactant. This is why the analyte multicomponent sample has a lower concentration of product with a high concentration of analyte.

Figure 3 shows the AEP for each model in millimolar. The analyte space in models A and B is defined using a single-component sample, and that in models C and D using a multicomponent sample. The interference space in models A and C is formed by single-component samples, while that in B and D is formed by multicomponent samples. If we compare model A with B and model C with D, we observe the effect of changing the composition of the samples in the interference space. Clearly, there is an improvement of the models' predictive ability when the interference space is formed using multicomponent samples.

To analyze the effect of the samples' composition in the analyte space, we should compare models A with C, and models B with D. These models are grouped by the same interference space. Comparing models A with C, we see an improvement when using single-component samples in the analyte space (though the differences are rather small). However, if we compare B with D we see no significant change in the AEP. This indicates that multicomponent samples are preferred in the interference space, and in this case, the composition of the sample in the analyte space has no effect on the prediction ability of the models.

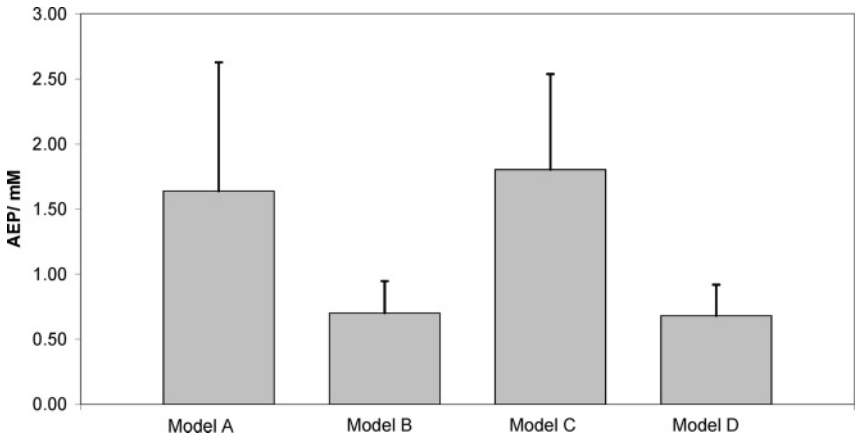


Figure 3. Values of averaged error in prediction, AEP, averaged out from 200 runs, for the models shown in Table 2. The error bars are the standard deviation of the values over the 200 runs.

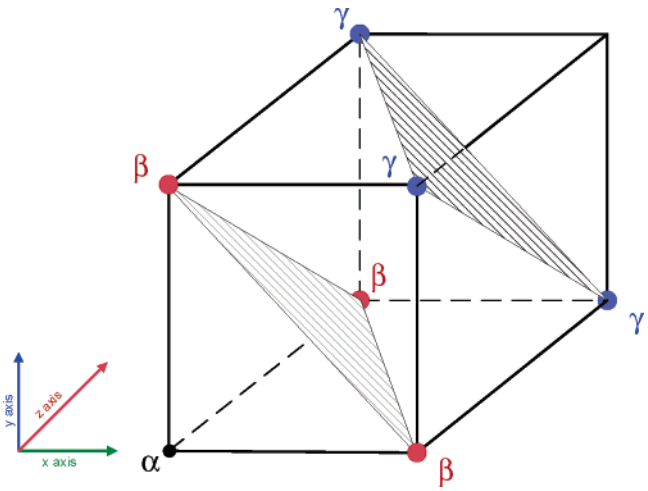


Figure 4. Schematic representation of the interference space. The axes *x*, *y*, and *z* represent NBA in solvent, Et_3N in solvent, and NBC in solvent, respectively. Red points (β) pertain to single-component samples in the interference space of models A and C, while the blue points (γ) correspond to multicomponent samples forming the interference space in models B and D. The bottom left point α indicates a solvent sample.

Why are multicomponent samples better for the interference space than single-component samples? Figure 4 gives a schematic representation of the {NBA, Et_3N, NBC, solvent} system. The cube represents the entire interference space, and the axes represent NBA in solvent, Et_3N in solvent, and NBC in solvent, respectively. The bottom left point α denotes pure solvent. Note that the further away one goes from α, the lower the solvent concentration. The red points β denote the single-component samples used to construct the interference space in models A and C, and the blue points γ denote the multicomponent samples used to construct the interference space in B and D. Point α is present in all models. Thus, the subspace $\{\beta, \beta, \beta, \alpha\}$ pertains to the range of the interference space covered by models A and C, and the subspace $\{\gamma, \gamma, \gamma, \alpha\}$ pertains to the range of the interference space covered by models B and D.

If all interference samples were error-free, both subspaces could be used to construct exactly the same interference space (the cube). Measurement errors (both concentration and spectral), however, will lead to differences. As the single-component samples cover only a small volume of the cube, measurement errors will affect the single-component case much more than the multicomponent case. To understand why this happens, let us consider the effects of each type of error on the interference space separately.

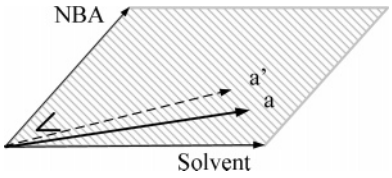


Figure 5. Effect of the concentration error on the interference space. The two dimensions represented by orthogonal arrows indicate the solvent and the NBA directions. The vectors a and a′ are two samples of NBA in solvent.

The concentration error has no effect on the interference space. This bold statement is best illustrated by way of an example: Let us assume that we want to construct an interference space that covers NBA and solvent (Figure 5). A solvent sample and sample a are used to span this space, but because of a concentration error, too much NBA is weighed in. The resulting sample, a′, contains therefore a slightly higher concentration of NBA and a slightly lower concentration of solvent. Nevertheless, the {solvent, a} samples still define exactly the same space as the {solvent, a′} samples.

If the concentration error does not affect the estimation of the interference space, it is the spectral error that is causing the differences between the single-component and multicomponent interference spaces. Let us look again at Figure 4, where we now assign the origin point as α. Assuming homoscedastic spectral errors, an estimation of the interference space that uses the β points (single-component sample vectors) will be more affected by spectral error than one that uses the γ points (multicomponent sample vectors), as the latter are further away from the origin.

The concentration error on the analyte space has a direct effect on the calibration line and thus on the prediction error. When too much analyte is weighed, the NAS value is overestimated. In this case, a one-point calibration model will result in too low concentration predictions for all samples. The effect of the spectral error on the analyte space is partly reduced by the interference space; i.e. the spectral error in the same direction as the interferences space is removed.

Effect of Changing Error Levels. Computer simulations can be considered as a link between theory and experimental work.

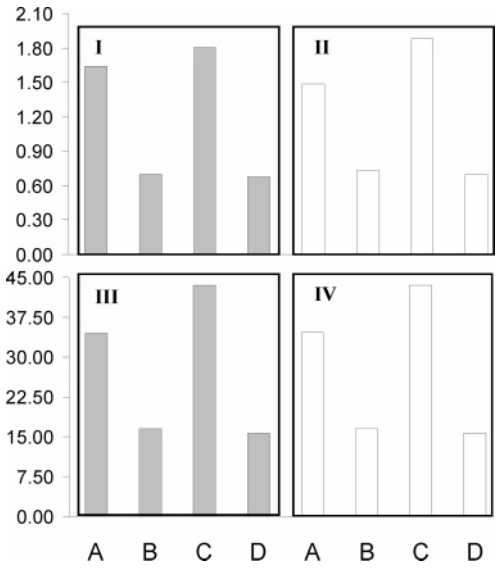


Figure 6. Effect of increasing the variances of the concentration error, σ_R^2 (au), and spectral error, SE (au), on the AEP of the models. (I) $\sigma_R^2 = 7 \times 10^{-6}$ and SE $= 5 \times 10^{-5}$ (original data); (II) $\sigma_R^2 = 7 \times 10^{-5}$ and SE $\times 10^{-5}$; (III) $\sigma_R^2 = 7 \times 10^{-6}$ and SE $= 5 \times 10^{-4}$; (IV) $\sigma_R^2 = 7 \times 10^{-5}$ and SE $\times 10^{-4}$. Note the difference in scale between the top and bottom graphs.

They have the advantage that the experimental parameters, such as error levels and number of components, can easily be changed and results are obtained quickly. To study the effect of changing the error levels in the models, the variances used to simulate the spectra were varied over the ranges 5×10^{-5}–5×10^{-4} and 7×10^{-6}–7×10^{-5} for the spectral and concentration errors, respectively.

Figure 6 shows the results when increasing the concentration error and the spectral error by a factor of 10. The original data are plotted in I. Plot II shows the results obtained when the concentration error is increased by a factor of 10 from the original value. Plot III shows the results when the spectral error is increased by 10 times from the original value. Finally, plot IV shows the results when the variances of both the errors are changed by a factor of 10.

The increase of the concentration error has almost no effect on the models' prediction ability. Conversely, increasing the spectral error has a dramatic effect (note the different scales). Thus, for this application, the contribution of the spectral error is much larger than the contribution of the concentration error. From Figure 2 it can be seen that PhI is almost completely overlapped from the solvent NMP. Therefore, quantification becomes very difficult if the spectral noise increases. As discussed before, the concentration error affects only the analyte space. However, the spectral error affects both spaces leading to the observed results. We can then conclude that the investment in a HT laboratory should be on the spectrometer rather than on the weighing robot.

Experimental Example. The Heck reaction was monitored at 60 °C using FT-NIR spectroscopy.[8] A PLS calibration model (for a detailed explanation on PLS theory, see ref 39) for PhI quantification based on GC results was built using 32 samples. For comparison, the necessary samples to build models A and B

(39) Wold, S.; Sjostrom, M.; Eriksson, L. *Chemom. Intell. Lab. Syst.* **2001**, *58*, 109–130.

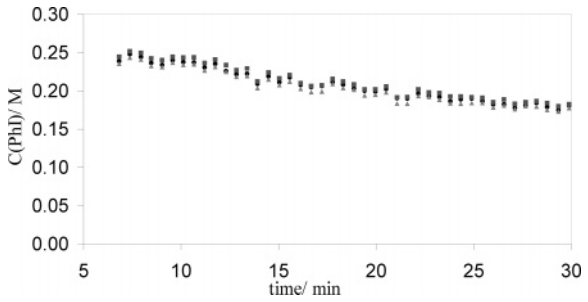


Figure 7. Concentration profiles for the Heck reaction at 60 °C obtained from experimental data. Legend: ●, PLS model based on 32 samples; ■, model A; ◆, model B.

were prepared and measured using the same spectrometer. Figure 7 shows the concentration predictions obtained with these three models: A and B built from experimental data and the PLS model. For convenience, each point is averaged out of 10 time points.

Figure 7 gives us three important points: First, the difference between the predictions of models A and B are close to the PLS model predictions (the root-mean-square differences between the PLS model and models A and B are 6.30 and 5.54 mM, respectively). Thus, the two models, with only 5 samples, are comparable to the PLS model that uses 32 samples. Second, the predictions of models A and B have a small systematic deviation from the PLS model (A is always higher, and B always lower). This is probably due to a concentration error in the analyte samples in these models (each model is based on a single analyte sample). Third, the deviation from the expected monotonic decreasing concentration curve for PhI is due to sampling and spectral errors (less obvious here because of averaging in time). The fact that the same spectrum (and thus the same spectral error realization) is applied to all three models results in a similar deviation in the concentration prediction, since the models are similar. Repeating the reaction under the same conditions confirmed these results.

Cost/Information Tradeoff. Models with a minimum number of samples are convenient, especially for HTE. Such models may give accurate predictions as long as the assumptions on which the models are based are valid. One may choose to test these assumptions, but this "costs" more calibration samples. The user must decide on this tradeoff.

Linearity in the analyte space can easily be checked by including samples at different levels of concentration and forcing the line through these points. This will show whether the points all fall on a straight line or not. Until now the calibration lines for analyte k considered are defined by one sample and therefore forced through zero. This is a valid assumption, since a concentration of zero should give a NAS value of zero. Not forcing the line through zero would require an additional sample to set the calibration line. However, this extra point can also be considered a validation step. Obtaining an offset close to zero would ensure that the NAS direction is well chosen.

Using the NAS approach, one has the possibility of tuning the calibration for the application at hand. For the discovery stage in HTE (a "go"/"no go" analysis), the calibration model does not need to be very accurate. The main question at this stage is "Is there any product or not?" On the other hand, the optimization stage needs an accurate answer about the amounts of product,

costing more analysis time. The approach applied here is flexible and can be tuned to one situation or the other.

CONCLUSIONS

By using the NAS approach for calibration, one can reduce the minimum number of samples to the number of chemical components in the system. In the present work, we have looked at how these few calibration samples should be chosen to obtain good predictions. Using blank samples (i.e., samples that contain only interferences and no analyte) facilitates the sequential addition of new interferences to the model, which fits well to high-throughput environments. Our simulation results and experimental data show that the samples used to build the interference space should be composed of several components mixed in a solvent (i.e., multicomponent samples). Choosing the right samples improves the predictive ability of the model. Also, extra interferences can systematically be introduced in the model by simply including additional calibration samples. Very high concentration errors will of course deteriorate the analysis, and any type of sample will yield the same results. However, the effect of the spectral error on the prediction ability of the models is higher than that of the concentration error. Further studies in our laboratory will include extending the calibration model to new situations and developing calibration-free monitoring techniques.

APPENDIX

Derivation of the Variance in the Concentration, σ_R^2. The concentration in the weight fraction of component k (C_k) is given by eq 2. It is assumed that m_1, m_2, ..., m_K are independent. The variance in the concentration, σ_R^2, is found by error propagation theory as the weighted sum of the variances of the different sources of error (masses of the different components) in the preparation of the sample. These variances are all equal and depend only on the analytical scale, ${\sigma_{m_1}}^2 = {\sigma_{m_2}}^2 = ... = \sigma^2$:

$$\sigma_{C_k}^2 = \left(\frac{\partial C_k}{\partial m_1}\right)^2 \sigma_{m_1}^2 + \left(\frac{\partial C_k}{\partial m_2}\right)^2 \sigma_{m_2}^2 + ... + \left(\frac{\partial C_k}{\partial m_K}\right)^2 \sigma_{m_K}^2 = \sigma^2\left[\left(\frac{\partial C_k}{\partial m_1}\right)^2 + \left(\frac{\partial C_k}{\partial m_2}\right)^2 + ... + \left(\frac{\partial C_k}{\partial m_K}\right)^2\right]$$

The partial derivatives in the previous equation lead to

$$\left(\frac{\partial C_k}{\partial m_k}\right) = \frac{\partial}{\partial m_k}\left(\frac{m_k}{\sum_{k=1}^{K} m_k}\right) = \frac{1}{\sum_{k=1}^{K} m_k} - \frac{m_k}{(\sum_{k=1}^{K} m_k)^2} \quad (8)$$

$$\left(\frac{\partial C_k}{\partial m_{j\neq k}}\right) = \frac{\partial}{\partial m_{j\neq k}}\left(\frac{m_k}{\sum_{k=1}^{K} m_k}\right) = -\frac{m_k}{(\sum_{k=1}^{K} m_k)^2} \quad (9)$$

where j stands for one of the components except for component k. The average number used in this study for the variance in the concentration, σ_R^2, was found to be 7×10^{-6}.

Derivation of Eq 3. Let $wgEr_k$ be the normal distributed random error in concentration of component k, and ϵ_k the variance of the error in the mass for each component. Including these errors in eq 2 gives

$$C_k + wgEr_k = \frac{m_k + \epsilon_k}{\sum_{k=1}^{K}(m_k + \epsilon_k)} = \left(C_k + \frac{\epsilon_k}{\sum_{k=1}^{K} m_k}\right)\cdot\left(1 - \frac{\sum_{k=1}^{K}\epsilon_k}{\sum_{k=1}^{K} m_k}\right) \quad (10)$$

The error in concentration of component k is then given by eq 11:

$$wgEr_k = \frac{\epsilon_k}{\sum_{k=1}^{K} m_k} - \frac{C_k\sum_{k=1}^{K}\epsilon_k}{\sum_{k=1}^{K} m_k} \quad (11)$$

Since the concentration is calculated as a weigh fraction, $\sum_{k=1}^{K} C_k = 1$, the sum of all the errors in concentration is

$$\sum_{k=1}^{K} wgEr_k = \frac{\sum_{k=1}^{K}\epsilon_k}{\sum_{k=1}^{K} m_k} - (\sum_{k=1}^{K} C_k)\frac{\sum_{k=1}^{K}\epsilon_k}{\sum_{k=1}^{K} m_k} = 0 \quad (12)$$

In other words, the error in the solvents' concentration, for example, is the sum of the error of all the other components. Therefore, the errors were included in the simulated samples as presented in eq 3.

Received for review October 26, 2004. Accepted January 19, 2005.

AC048421C