# A chemometric approach: evaluating important parameters in the near infrared in-line calibration

# Vincent A.L. Wortel and Wei G. Hansen

Unichema International, PO Box 2, 2800 AA Gouda, The Netherlands.

# Introduction

### Background

A new ester (ethyl glucoside ester) was at the stage of development in a pilot plant environment. A fatty acid and an alcohol (ethyl glucoside) were reacting to form ester and water (see Figure 1). The process was monitored by the acid value basically, among others.

A NIR spectrometer was installed in the production line as a process analyzer. The advantages of using a process analyzer for a batch process are well known and described in the literature:<sup>1</sup>

- fast selection of reaction endpoint;
- reduction of routine laboratory testing;
- increased process efficiency;
- reduction of off-specification production.

Data used in this paper were collected during four production batches. Three batches were used to build a training set for NIR calibration. The remaining (fourth) batch was used as an independent test set to validate the calibration model.

# Taguchi concept

The Taguchi method for a production process control is described very clearly in the literature.<sup>2,3</sup> The basic idea of the Taguchi concept is to design a process using two important criteria:

(i) process variance;

(ii) average process performance.

Because the influence of all variables on a production process needs to be estimated, a good experimental design is essential. Besides a good spread of experimental points at different levels of parameters, it is also necessary to repeat some of the experimental conditions in order to





calculate the process variance. The process variance is related to the robustness of the process and measured by the signal/noise ratio:

$$S/N \ ratio = 20 \log_{10} \left[ \frac{Mean}{Standard \ Deviation} \right] \dots$$

The first step is to analyze which parameters have significant effects on the process variance, then these parameters will be used to control and reduce the process variance.

The average process performance is calculated by simply averaging repeated experiments. These results are used in the second step to analyze which parameters have significant effects on the average process performance. Only the insignificant parameters of the first step are allowed to be used for the optimisation of the second step, because the stability of a process is assumed to be more important in comparison with the average process performance.

#### Taguchi approach in NIR

Several methods have been suggested in the literature in order to make good calibrations for NIR applications. Chemometric tools such as MLR, PCR and PLS with validation methods, e.g. cross-validation or independent test set validation are used to make reliable NIR calibration models. Various criteria, *RMSEP*, *SEP*, ... are used to evaluate the results of the validation.<sup>4</sup>

In contrast with the Taguchi's optimisation concept, most NIR calibration optimisation methods are only focusing on the average calibration performance. The robustness of the calibration model is underestimated during the calibration. Nevertheless, robustness of the model is a very important indication for the prediction performance of this model in the future. A good evaluation of the calibration model robustness will reduce the number of recalibrations in the future.

In this paper, the Taguchi optimisation concept has been utilized in the NIR calibration to select the optimal calibration parameters. It results in a robust calibration model, that means insensitive to uncontrolled (i.e. unwanted) variance.

# Experimental

The procedures used for this experiment are displayed in a flowchart (Figure 2) to illustrate the complete experiment. The multivariate calibration models are made by using the UNSCRAM-BLER package (Version 5.5). SAS and STATGRAPHICS software are used to evaluate the data and execute the ANOVA calculations. A NIR instrument Prospect IV (UOP, Guided Wave) equipped with a SST probe was installed in-line of the reaction as a process analyser. NIR spectra (66) were automatically measured and collected during four batches, and meanwhile the samples were taken for, among others, an acid value (conform IUPAC 2.201 1979) determined by the wet chemistry analytical methods. The wet chemical analysis results were used as the reference values during the calibration/validation. All spectra were baseline corrected (at 1320 nm) and the set of 66 samples was divided up into a training set for calibration and an independent test set for validation.

In this paper, only three parameters were selected as the main parameters in the experiments to simplify the discussion. These three parameters originated from three different aspects:

- 1. spectrum related settings;
- 2. data treatment related settings;
- 3. regression related settings.

However, to optimise complete applications in real practice, more parameters can/should be used. An overview of the experimental settings is displayed in Table 1.

Experiment	Wavelength selection (nm)	Data pretreatment	Regression method
1	1000–2100	no MSC	PLS-1
2	1000-2100	no MSC	PLS-2
3	1000–2100	MSC	PLS-1
4	1000–2100	MSC	PLS-2
5	1000–1850	no MSC	PLS-1
6	1000–1850	no MSC	PLS-2
7	1000–1850	MSC	PLS-1
8	1000–1850	MSC	PLS-2
9	1300–2100	no MSC	PLS-1
10	1300–2100	no MSC	PLS-2
11	1300–2100	MSC	PLS-1
12	1300–2100	MSC	PLS-2

Table 1. Experimental settings.

MSC = multiple signal correction.<sup>6</sup>

Noise (1% relative to the signal) is added to the spectra to represent future instability in the NIR spectrometer and to make the S/N ratio possible. This has resulted in 12 calibration models. Each calibration model is used to make a prediction for the test set: once with the normal test set and once with the noise added test set. The *RMSEP* (root mean squares error of prediction) of the independent test set is used as representative parameter for the prediction performance and is evaluated using a multiple ANOVA<sup>5</sup> (analysis of variance) calculation. ANOVA is an indication of whether a data treatment has a significant effect on the response value. ANOVA is also a useful tool to analyse the cause and effect of the relationship between independent variables (in this case wavelength selection, data treatment and regression method) and response value, which could be a clear index for the performance of the calibration. To use the concepts of Taguchi, the *RMSEP* with and without adding noise are transformed into the average S/N ratio. The main effects and interactions<sup>3</sup> are both included in the calculation, the insignificant effects are eliminated to decrease the number of points of freedom.

# Results and discussion

#### Parameter selection

The first parameter is the wavelength selection for the calibration. The wavelength range of the spectrometer is 1000–2100 nm. The use of the complete wavelength range for calibration is the first option.

The second option is defined by excluding the low wavelength part of the spectrum (from 1000–1300 nm), which contains mainly absorptions of second overtones of CH bond vibrations.

The third option of wavelength selection is the elimination of the high wavelength part of the spectrum. This part of the spectrum is normally more noisy and unstable in comparison with the rest of the spectrum, because of the limited range of Tungsten lamp illumination and the InGaAs detector sensitivity, which varies the spectra the most, as the lamp and detector deteriorate with time. As a consequence, the calibration will be inaccurate for a long time prediction model. On the other hand, the absence of information of the water (1900–1940 nm) and acid (1880–1910 nm) vibration bands could cause problems during calibration as well.

The second parameter is the data pretreatment, two options are defined:

- a MSC treated data set;
- a non-treated data set.

Full multiple signal correction (additive and multiplicative) is done on the complete selected spectral range.

The third parameter is the regression method, with two options:

- PLS-1;
- PLS-2.

Although the acid value and water content are not the only parameters to monitor the process, these data are randomly selected as examples for PLS-2 calibration and for explaining the chemometric concepts.

### Analysis of S/N ratio RMSEP

It has been found that the type of calibration (PLS-1 or PLS-2) has no significant influence on the robustness of the calibration. The ANOVA of the RMSEP S/N ratio resulted in three significant effects:

- wavelength selection;
- data pretreatment;
- wavelength selection \* data pretreatment interaction.

The wavelength selection definitely has the most significant effect. The results are visualized in an interaction plot (Figure 3), which shows clear differences between the calibration robustness, in the S/N ratio scale!

The calibration model with the highest S/N ratio is the model with the best robustness. Three combinations have a S/N ratio of approximately arbitrary units value of 50 or more:

- wavelength selection 1300–2100 nm without MSC;
- wavelength selection 1000–2100 nm without MSC;
- wavelength selection 1000–2100 nm with MSC.

The wavelength selection of 1000–1850 nm reduces the robustness in comparison with the other wavelength selections. The combination of the range 1000–1850 nm with MSC resulted in a very dramatic reduction of the S/N ratio. The spectral information between 1850 and 2100 seems to be very important for the acid value calibration especially in combination with MSC. The lack of the second overtone C=O vibration (~1918 nm) has a drastic effect on the calibration. In contrast the combination of MSC with the wavelength selection 1000–2100 nm resulted in the highest S/N ratio.

# Analysis of average RMSEP

The second step of the optimisation is the minimisation of the *RMSEP* value. However the ANOVA of the average *RMSEP* resulted in the same significant parameters as the S/N ratio *RMSEP* analysis. The F-values are much bigger in comparison with the ANOVA of the S/N ratio.

The interaction plot (Figure 4) shows a clear difference between the different combinations. If Figure 4 was the only information available, the choice would have been the wavelength



Figure 2. Flowchart of experiment.

selection 1000–1850 nm. But according to the robustness analysis this selection results in a less robust calibration. The minimisation of the average *RMSEP* value is limited by three options from the maximisation of the S/N ratio:

- wavelength selection 1300–2100 nm without MSC (avg. *RMSEP* = 1.89);
- wavelength selection 1000–2100 nm without MSC (avg. *RMSEP* = 1.89);
- wavelength selection 1000-2100 nm with MSC (avg. *RMSEP* = 2.44).

The difference between the wavelength selection of 1000–2100 nm and 1300–2100 nm, both without MSC, is trivial. In this case, the second overtone C–H vibrations in the wavelength range 1000–1300 nm does not seem crucial for the calibration model.

However, based on both analysis of the S/N ratio and the average *RMSEP* value, the best selection of wavelength range should be 1000–2100 nm. If the average *RMSEP* (2.44) is acceptable, this calibration gives the highest S/N ratio (most robust model). The average *RMSEP* can be reduced by selecting the same wavelength range but without MSC, the disadvantage of this alternative model is a reduction of the robustness.







Figure 4. Interaction plot average RMSEP.

#### Spectroscopic explanation of result

It is critical for chemists to understand and explain results in chemistry from the previous ANOVA approach. The focus is on experiment 1 and 3 (1000–2100 nm) and 9 and 11 (1300–2100 nm). The difference between the models is present in the influence plots (leverage against residual) and loading plots of the calibration models.

The loading plot shows that in calibration 1 (Figure 5) there is one important peak (at 1918 nm) and in calibration 3 (Figure 6) there are two important peaks (1918 nm and 1726 nm). The peak at 1918 nm is associated with the concentration of the fatty acid and the peak at 1726 nm is



Figure 5. Loading plot of calibration 1.

assigned to the second overtone  $CH_2$  vibration. For the calibration 1 the peak at 1726 nm and 1918 nm have the same loading sign (both positive), while for calibration 3 the loadings at 1726 nm and 1918 nm have an opposite signs. The absolute value of the loading at 1726 nm in calibration 1 is much smaller than in calibration 3.

Nine consecutive spectra from the original test set are compared with the equivalent MSC spectra to interpret these differences. Special attention is payed to the peak area at 1726 and 1918 nm.

The variance of the spectra with a wavelength selection of 1000–2100 nm is reduced in the wavelength range of 1500–1650 nm and 1850–2000 nm by MSC, which is the desired effect. The



Figure 6. Loading plot of calibration 3.



Figure 7. MSC effect on the test set. Wavelength range 1880–1980 nm.

variance at the peak at 1950 nm (Figure 7) for the MSC corrected set (test set 3) is slightly reduced in comparison with the untreated set (test set 1).

In contrast, the multiple signal correction has an unexpected peak at 1726 nm (Figure 8). The first effect to be noticed is the change in the sequence of the spectra for calibration 1 and 3. The absorptions of the original spectra are increasing from spectrum 1 to 9. In contrast, the maximum absorption at 1726 nm in the corrected spectra decreases from spectrum 1 to 9.

Second, the observed variance of the peak at 1726 nm is increased due to MSC. The comparison of the test sets for calibration 9 and 11 shows also a rearranging in the sequence of the spectra. In contrast with test set 3, the variance does not increase significantly.

This opposite loading direction of the  $CH_2$  and C=O peaks is indeed explained by the extraction of water during the reaction, which causes a relativly higher  $CH_2$  concentration during the process. The  $CH_2$  is inversely related to the acid value during this specific process. Because the variation of  $CH_2$  concentration in the spectra is very subtle without proper data treatment, the instrument deviation can easily cover the real variation from  $CH_2$  concentration change. Naturally, after MSC data treatment (to reduce the spectral variation), the spectra absorption at ~1726 nm will be only related to  $CH_2$  concentration. In other words, the order of the spectra is changed.

Summary of this data analysis

The Taguchi method of analysis performed well for the optimisation of the calibration parameters. If using the results of the analysis of the *RMSEP* only, it would have lead to a wavelength selection of 1000–1850 nm with or without MSC, leading to a poor robustness calibration model.

In addition, the results of chemometrics approach correspond very well with the chemistry concepts.

# Conclusions

To conclude our discussion, evaluating important parameters in the NIR calibration should not only be based on the average performance but also on the robustness. The analysis of the *RMSEP* 



Figure 8. MSC effect on the test set. Wavelength range 1720–1730 nm.

only would lead to a false wavelength range selection, due to the poor robustness of this calibration model. This calibration model could promise accurate results during calibration and validation, despite the use of an independent test set. However, after some time the accuracy of the calibration would gradually decrease.

Whereas, the S/N ratio *RMSEP* analysis of the calibration model is a powerful tool to evaluate the robustness of the calibration model. The combination of S/N ratio and average *RMSEP* ANOVA will insure a much better robust calibration model with a good *RMSEP*.

# References

- 1. K.J. Clevett, Process and Quality Control 6, 203 (1995).
- Khosrow Dehnad, *Quality Control, Robust Design and the Taguchi Method*. Wadsworth and Brooks, pp. 3–50 (1989).
- 3. R.A. Stone and A. Veevers, Journal of Chemometrics 8, 103 (1994).
- 4. H. Martens and T. Næs, *Multivariate Calibration*. John Wiley & Sons, Chichester (1989).
- E. Morgan, K.W. Burton and P.A. Church, *Chemometrics and Intelligent Laboratory Systems* 5, 283 (1989).
- 6. O.E. de Noord, Chemometrics and Intelligent Laboratory Systems 23, 65 (1994).