

Reliable Online-Prediction of Characteristic Process Parameters by FTNIR-Spectroscopic Analysis

Wolfram Summerer¹, Marcin Pawliczek¹, Horst Trinker², Peter Hintenaus², Wolfgang Märzinger³,
Thomas Reischer⁴, Martin Nowak⁴, Martin Emsenhuber⁴, Jürgen Kasberger¹
¹ RECENDT GmbH, Linz - Austria,

Corresponding author's e-mail address: wolfram.summerer@recendt.at

² University Salzburg - Austria; ³ i-RED GmbH, Linz - Austria; ⁴ Dynea GmbH, Krems - Austria

Abstract:

Within the industrial research project "Process Analytical Chemistry" (PAC) we are working on FTNIR-spectroscopic measurement systems predicting characteristic parameters of industrial production processes. Those parameters are usually monitored offline or at-line with time consuming and expensive laboratory methods. In this contribution, we present a spectroscopic measurement configuration together with the required chemometric analysis, acting as an online-monitoring system. In order to demonstrate the potential of such a system we use the example of melamine resin production in an industrial process. At company partner Dynea the predicted value of the turbidity point is used as an indicator for the end of the batch reaction (turning off heating). Furthermore, we illustrate a way to verify the chemometric prediction by calculating a confidence interval for each predicted value.

Key words: FTNIR-spectroscopy, process analysis, online-monitoring, confidence interval

Introduction

Environmental and economical challenges of the future lead to increasing efforts on saving energy and resources in chemical industry. In many cases the efficiency and quality of industrial production processes can be improved by applying online analytic technologies. Fourier-transform-near-infrared (FTNIR) spectroscopy is a well-established method in process analytics. In this contribution we use such a measurement system [1] for the online-monitoring of critical process parameters. The use of chemometric algorithms allows the development of mathematical models, which represent the correlation between spectral data and the parameters of interest. This enables the prediction of a process parameter by applying such a model to new spectral data from the process. In the following example we predict the turbidity point of melamine resin at the production plant of company partner Dynea. Moreover we apply two methods to calculate confidence intervals for reviewing the predicted values during online-monitoring.

Measurement Setup

Spectroscopic data from the process is continuously collected by a FTNIR-spectrometer (i-RED) [1]. The measurement is

conducted by a transmission probe (Hellma) with an optical path of 1 mm through the fluid. This immersion probe is connected to a halogen light source on the one side and to the spectrometer on the other side with optical fibers (length: 50 m). The used spectral range in the NIR extends from 1 μm to 2.5 μm . The whole system is installed at the production plant and connected to the process control system by a CANopen fieldbus link and analog input terminals. The particular components of the system are illustrated in Fig. 1.

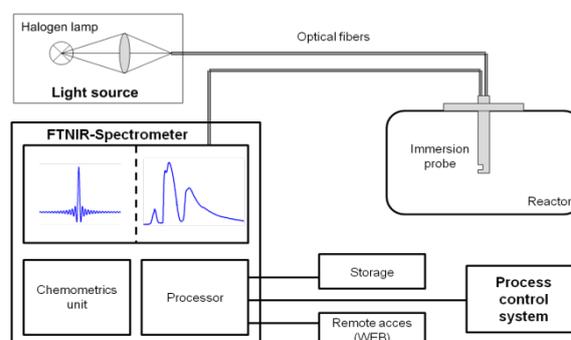


Fig. 1. Sketch of the FTNIR-setup at the production plant. An immersion probe is attached by optical fibers and the spectrometer is integrated in the process control system via a CANopen fieldbus connection.

Thanks to this integration into the process network, additional information like temperature, pressure and pH can be recorded and taken into account for analysis. Parallel to the continuous acquisition of spectroscopic data, the relevant process parameter is measured with a reference method, which is usually based on manual sampling.

Calibration and Modeling

In order to obtain a mathematical model describing the process parameter (reference value) by a NIR spectrum, we used chemometric standard methods like partial least squares regression (PLS). This algorithm builds linear regression models by calculating the principal components of both the independent variables (spectral data) and the dependent variables (reference measurements). This is in contrast to partial component regression (PCR), where only the variance of the spectral data is considered for dimension reduction [2]. The software tool, which was primary used for calibration in this contribution, is the PLS_Toolbox (Eigenvector Research) for use with MATLAB™ [3].

Statistical modeling was performed with a calibration dataset, where the process parameter (turbidity point) was measured by the operating personnel at the plant. Predicting now the process parameter by analyzing a spectrum, means in simple terms to multiply the regression vector obtained from the model by the corresponding spectrum. The predictive power of the resulting mathematical model strongly depends on the utilized preprocessing methods as well as on the suitable selection of relevant spectral regions (see Fig. 2).

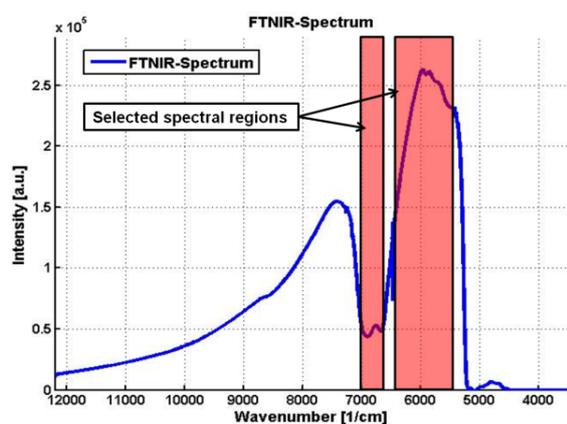


Fig. 2. FTNIR-Spectrum of melamine resin with the relevant spectral regions (selected).

The best results were achieved by applying standard normal variate scaling (SNV) as preprocessing procedure to spectral data. The most appropriate wavebands, which could be

found during the modeling period, are highlighted in Fig. 2.

After a period of data acquisition the measured reference values were correlated with the corresponding spectra. This was carried out successfully for a number of different recipes of the melamine resin. As the results and the calibration process are similar, we show the results of just one exemplary recipe.

Every calibration result is cross-validated (CV) and reviewed (see e.g. Fig. 3) before a specific model is implemented online. Cross-validation was performed with the “leave one out” method, where every sample is consecutively left out and predicted by a model containing the remaining samples. Statistical variations of the process and the measurement conditions require the usage of data from a comparatively large number of batches for calibration. The model shown in Fig. 3 uses spectra and reference data from over 120 batches produced within 9 weeks, while the number of reference measurements per batch was between 6 and 9. This corresponds to around 1000 samples.

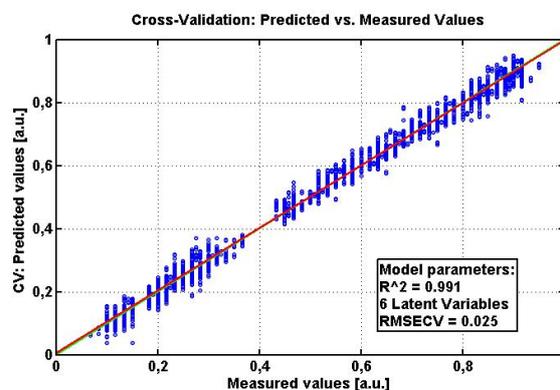


Fig. 3. Cross-Validation result of the calibration model used for online-prediction.

On the way to an online-implementation of a model there is often a lot of re-calibrating and validating necessary. Thereby in addition to cross-validation it is very useful to validate the model by using a test data set. Comparing predicted and measured values of the modeled process parameter is the preferable way to validate a model for online-application.

Process Implementation / Online-Monitoring

A robust and validated chemometric model enables the prediction of a process parameter by evaluation of the currently measured spectrum. This evaluation is computed online on the process spectrometer system, while the batch process is running. In order to demonstrate the capability of the system, we monitor the turbidity point for each individual

batch of the melamine resin production at company partner Dynea (see Fig. 4). In this application the turbidity point is a parameter, which provides information about the progress of the condensation process in melamine resin production. Monitoring the value of that parameter indicates the best point of time to turn off heating in order to stop the condensation.

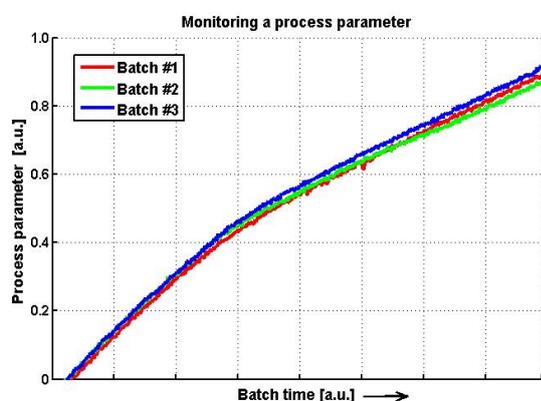


Fig. 4. Representative examples for the online-monitoring of a process parameter (3 consecutive batches).

After a period of testing and re-training, the obtained chemometric models for predicting the turbidity point turned out to be very accurate and robust. But the maximum predictive power of the applied model is naturally limited by the accuracy of the reference measurement. During online monitoring comparative reference measurements were performed in order to check the capability calibrated model. These have shown good accordance to the predicted values of the FTNIR-system, as depicted in Fig. 5 for a single exemplary batch.

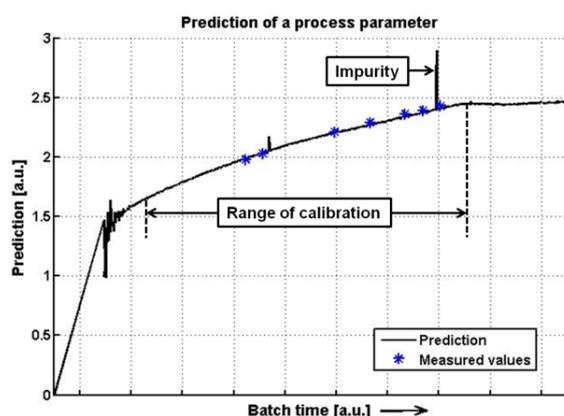


Fig. 5. Predicted process parameter compared with reference measurements for one representative batch.

The long term stability of the model in use was tested over several weeks. During this test the RMSEP (root mean square error of prediction)

normalized to the data range of the process parameter was around 4%, which is approximately the same magnitude as the estimated error of the reference method.

The range of calibration is the part of the batch, where condensation takes place and reference measurements are available. The periods before and after the reaction (filling of the reactor and cool-down phase) are not important for process control and therefore not covered by calibration. Due to the continuous spectral measurements one gets a predicted value of the process parameter for every corresponding spectrum, even if the process state is out of calibration range or if other disturbances occur. To get a measure for the reliability of the displayed value we compared two closely related ways to calculate confidence intervals.

Confidence Intervals

Even though a lot of effort has been spent in the recent years to predict process parameters by spectroscopic methods, there is only limited work done to estimate the confidence level of predicted values. Such a confidence interval would be beneficial, because in some cases one has small impurities or unforeseen disturbances (e.g. bubbles), which can strongly influence the prediction. To detect such errors and to verify the predicted values we used confidence intervals as described in [4] and [5]. The principle of calculating the confidence intervals (CI) is expressed by eq. (1) in a simplified form:

$$CI = \pm t_{\alpha} (MSE)^{1/2} (1 + D_M)^{1/2} \quad (1)$$

where MSE is the mean squared error of the training set and D_M is a Mahalanobis distance measure of the current spectrum. The factor t_{α} represents a value of Student's t distribution, which uses the number of batches and the number of principal components from the PLS-model. The value of $1-\alpha$ defines the level of confidence (e.g. 95%) and is also included in the factor t_{α} . The main difference between [4] and [5] in calculating the confidence intervals is in the different measure for Mahalanobis distances. The effect of these differences on the confidence interval appears in the varying slope in case of disturbances. The method used by PLS_Toolbox (Eigenvector Research) can be found in [4].

Figure 6 shows the confidence intervals calculated for the batch depicted in Fig. 5 using both methods. The level of confidence was set to 95%, which corresponds to an α -value of 0.05. Small impurities in the probed liquid, like the one marked in Fig. 5 and 6, can be detected

and are indicated by these confidence intervals. Outside the range of calibration the confidence intervals strongly arise, resulting from high Mahalanobis distances of these spectra.

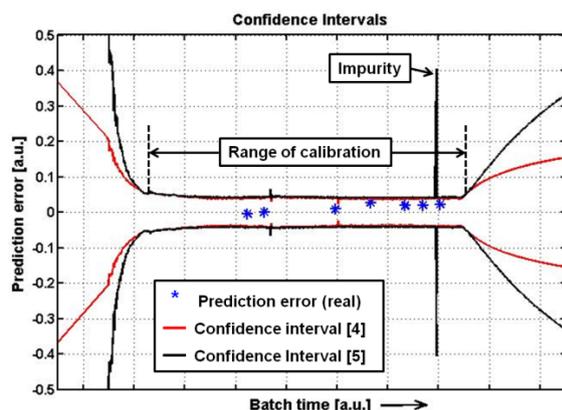


Fig. 6. Confidence intervals (confidence level: 95%) calculated for every predicted value over time with actual prediction errors.

The actual prediction error has been evaluated by test measurements in the industrial process and it should remain between the upper and lower boundary of the confidence interval. Aside the indication of wrong predictions due to impurities, the confidence interval also strongly arises when the measured spectra are outside the data space of calibration (cf. beginning and end of batch in Fig. 6). Large values of the confidence interval suggest that one should not trust the predicted values at that state of the process.

As the calculation of the reviewed confidence intervals is done offline up to now, the next step is the online-implementation and online-validation of the shown methods. Due to good progress in programming the spectrometer system is already prepared to do the required calculations.

Conclusion

We have shown the development of a method to online-monitor characteristic process parameters in industrial production processes within the research project PAC. The potential and reliability of combining FTNIR-spectroscopic measurements and chemometric analysis, was presented exemplary by monitoring a crucial process parameter in melamine resin production.

Future prospects in this project are further improvement and development of the online-analytic system and the extension of the shown methods to other plants and new processes. Developing advanced chemometric methods to further improve accuracy and long term stability of the existing system plays an important role.

This includes research and development in the field of output diagnostics, like the presented confidence intervals. The integration of additional, orthogonal parameters like temperature or pressure in the modeling process is currently work in progress.

Analyzing and monitoring industrial production processes with such systems enables a deeper understanding of these processes. Therefore the usage of a reliable monitoring system in process control often leads to improvements in product quality and process efficiency.

Acknowledgements

This work was funded by the Austrian research funding association (FFG) under the scope of the COMET program within the research network "Process Analytical Chemistry (PAC)" (contract # 825340). This program is promoted by BMVIT and BMWFJ.

The work has also been supported by the European Regional Development Fund (EFRE) in the framework of the EU-programme Regio 13, and the federal state of Upper Austria.

References

- [1] P. Hintenaus, G. Kvas, W. Märzinger, An Infrared Spectrometer for Process Monitoring I, *Spectroscopy, Proc. IECON (33rd)*, 2576-2579 (2007); doi: 10.1109/IECON.2007.4459989
- [2] S. Wold, A. Ruhe, H. Wold, W. J. Dunn, The Collinearity Problem in Linear Regression. The Partial Least Squares (PLS) Approach to Generalized Inverses, *SIAM J. Sci. and Stat. Comput.* 5, 735-743 (1984); doi: 10.1137/0905052
- [3] B. M. Wise, N. B. Gallagher, R. Bro, J. M. Shaver, W. Windig, R. S. Koch, Chemometrics Tutorial for PLS_Toolbox and Solo, *Eigenvector Research, Inc.*, 1-420 (2006)
- [4] R. Bro, A. Rinnan, N. Faber, Standard error of prediction for multilinear PLS 2. Practical implementation in fluorescence spectroscopy, *Chemometrics and Intelligent Laboratory Systems* 75, 69-76 (2005); doi: 10.1016/j.chemolab.2004.04.014
- [5] R. Reiss, W. Wojsznis, R. Wojewodka, Partial least squares confidence interval calculation for industrial end-of-batch quality prediction, *Chemometrics and Intelligent Laboratory Systems* 100, 75-82 (2010); doi: 10.1016/j.chemolab.2009.11.003