Foreword

Since 77 years viscose fibers are produced in Lenzing. Meanwhile almost one million tons of man-made cellulose fibers are produced on our 6 sites worldwide. Recently, we have celebrated the 50th birthday of Lenzing Modal® fiber development. TENCEL® is the third and newest generation of man-made cellulose fibers. It is based on the NMMO process and manufactured at 4 of 6 sites. The required pulp is mainly produced in our two biorefineries based on beech and spruce wood. This pulp production is linked with the production of essential base chemicals like furfural, acetic acid, soda, xylose as well as energy.
These processes are under permanent development and optimization strengthening our international innovation and quality leadership. For successful process innovation, monitoring of process streams and development of new sensors and analytical techniques are essential. Since 2010 we expanded our potential with our partner in the research network PAC- "Process Analytical Chemistry".
Due to the research network PAC we do not only have access to experts at partner universities and institutes. A co-operation with industrial partners outside of our business gives access to new insights and generates new ideas for our own challenges even faster and better.
The topics we investigate include the development of new sensors and analysers as well as the development of evaluation methods and investigation of inhomogeneities in the process. In this issue only a few examples of these activities are described. In future we will follow up on these topics, which contribute to secure our technology leadership.

Dr. Andrea Borgards
Global Director Process Innovation Lenzing AG Lenzing, December 2015

Vorwort

Lenzing produziert an 6 Standorten fast eine Million Tonnen man-made Cellulosefasern. Seit nunmehr 77 Jahren werden in Lenzing Viskosefasern produziert; kürzlich feierten wir den 50sten Geburtstag der Entwicklung der Lenzing Modal®. Als dritte und neueste Generation der man-made Cellulosefasern wird die TENCEL® -Faser, basierend auf dem NMMO-Prozess, mittlerweile an 4 der 6 Standorten erzeugt. In unseren beiden Bioraffinerien wird der überwiegende Teil des dazu benötigten Zellstoffs aus Buche oder Fichte hergestellt. Dies ist gekoppelt mit der Erzeugung von wertvollen Basischemikalien wie Furfural, Essigsäure, Soda, Xylose sowie von Energie.
Die dazu notwendigen Prozesse sind einer ständigen Weiterentwicklung und Optimierung unterworfen, um die internationale Innovations- und Qualitätsführerschaft zu manifestieren. Zu einer erfolgreichen Prozessentwicklung gehören die Überwachung der Prozessströme und die Entwicklung neuer Sensoren und Analysentechniken. Seit 2010 erweitern wir mit unseren Partnern im Forschungsnetzwerk PAC – „Process Analytical Chemistry" unsere diesbezüglichen Möglichkeiten.
Durch das Forschungsnetzwerk PAC haben wir nicht nur Zugang zur Expertise an den Partneruniversitäten und –instituten. Auch die Kooperation mit branchenfremden Firmen trägt dazu bei, neue Sichtweisen kennenzulernen und daraus neue Ideen zu generieren, um eigene Fragestellungen schneller und besser beantworten zu können.
Die Themen reichen von der Entwicklung neuer Messgeräte und Sensoren über die Ausarbeitung verbesserter Auswertemethoden bis hin zur Erforschung der Ursachen von im Prozess auftretenden Inhomogenitäten. Im vorliegenden Band ist nur ein kleiner Auszug dieser Aktivitäten angeführt. Auch in Zukunft werden wir diese Themen weiter verfolgen, denn sie helfen uns, unsere Technologieführerschaft zu sichern.

Dr. Andrea Borgards
Global Director Process Innovation Lenzing AG Lenzing, Dezember 2015

Lenzinger Berichte – PAC-special-issue Autumn 2015

Editorial

Dear reader!
This special issue of the "Lenzinger Berichte" covers an increasingly important topic for chemical industry and shall give you a little insight into the Austrian research network on Process Analytical Chemistry – PAC. We want to make you curious to explore the content of this "PAC-issue" in detail!
At the beginning of this century a new trend was emerging in the fields of chemical and biochemical industry. The trendy topic was "process intensification". In the following years, this topic was picked up in various countries, by different players and in a variety of forms. The Austrian way was, to develop a concept for a research network to bundle the expertise of academic partners and companies.
One of the companies involved in these activities from a very early stage on, was Lenzing AG. They are still a very active and valuable partner in the research network – and this is why we are able to present this special issue today.
Those joint efforts led, in the first stage, to a proposal for a research project: the proposal was submitted in the year 2009 and granted early in 2010. Under the scope of the Austrian funding scheme COMET (Competence Centers for Excellent Technologies) the project PAC (Process Analytical Chemistry – Data Acquisition and Data Processing) represented a perfect fit. During four years (2010 – 2014) a consortium of 17 partners (10 companies and 7 scientific partners) conducted research and development for the improvement of process analytical technologies (PAT) and the utilisation of measurement data to allow for real-time control and optimization of processes.
The involved companies covered a large variety of industrial branches, from base chemicals to pharmaceuticals, from viscose fibers to steel industry, from beer brewing to the production of rubber.
In 2014 new horizons had to be targeted: a follow-up project was initiated by most of the former partners and some new ones. Thus, a consortium of now 19 partners (11 industrial partners and 8 university institutes and research companies) started a cooperation under the project-acronym imPACts: Industrial methods for Process Analytical Chemistry – From measurement technologies to information systems.
The project imPACts is once again a funded project in the COMET-program, administered by the FFG, the Austrian Research Promotion Agency. The scope was greatly extended: new analytical technologies are developed and applied within the project, special attention is paid to understand complex chemical processes, e.g., multi-phase processes and not only is the measured data processed by advanced chemometric algorithms, but also will those chemometric models be processed in a newly developed framework, called CMLCM: Computational Model Live-Cycle Management.
The project imPACts will run until 2018. We have an ambitious research agenda for this time. We invite you, to take your time and browse through some of the topics we are dealing with right here in this issue of "Lenzinger Berichte"!
Beyond imPACts, the acronym PAC has persisted as the name for the Austrian research network for process analytical technologies. It is the clear long-term target for the members of this network to keep on working on the topics of process analytical technologies (PAT) and process intensification.
However, we are convinced, that a sustainable development for the real benefit of the process industry and chemical industry can in the end only be done on an internationally networked level. Neither the big challenges nor the big players are confined to borders of countries.
PAC as the Austrian network teams up with the Austrian GÖCH, the German DECHEMA and the Arbeitskreis Prozessanalytik as well as with UK based CPACT. In 2014, thanks to the SPIRE-PPP in Horizon 2020, a European project called ProPAT was established led by the Spanish company IRIS. ProPAT aims to develop novel sensors and analysers as well as a global control system for closed-loop process control. Its underlying goal of application and outreach has a very similar scope to PAC and it is a perfect partner for PAC on an European level. Together, we will be able to bundle efforts for some topics to achieve increased output and outreach; we will meet, discuss and exchange in order to learn from each other; we will cooperate and aim to issue joint publications!

And what is it all about?

We are working to enable the process industry and especially the chemical industry, to have closer control over their processes.

Closer control – enabled by PAC – means to us:

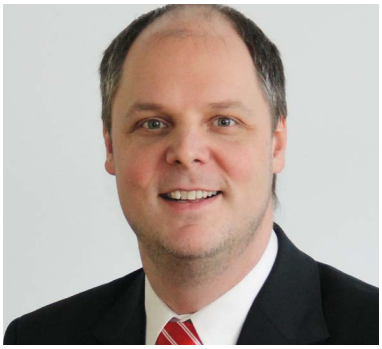
- real-time measurements for all relevant parameters;
- complete models of processes;
- a wide variety of measurement technologies at the industry's fingertips;
- access by soft-sensor technologies to parameters that cannot be measured directly;
- outstanding chemometric modelling concepts, to give reliable chemical information from physical measurements;
- models never get outdated, thanks to new methods from our CMLCM-toolbox;
- processes can always be run on absolute optimum;
- less downtime, more yield, less energy-, time- and raw material consumption;
- lowest production costs;
- and – the one thing always most important at the end of the day: always top quality for your customers.

Go on, read about it – either in this special issue of the "Lenzinger Berichte" or on our website www.k-pac.at.

We hope, you will enjoy to get some insights in work conducted in the PAC-network!

Markus Brandstetter
imPACts – Scientific manager

Thomas Röder
Chairman of the imPACts board
Lenzinger Berichte – Editor in chief



Robert Holzer
imPACts – Consortium manager

In Memoriam Prof. Dr. Burkart Philipp

Shortly after his 90th birthday Prof. Burkhart Philipp passed away in Berlin on the 9th of March 2015. Burkart Philipp was born in Pirna/Saxonia and studied chemistry at the TH (Technical University) Dresden from 1947 to 1950. He received his PhD (Dr. rer. nat.) under the supervision of K. Schwabe in 1952. Since 1953 he had been working at the former Institute of Fiber Research (IFF) of the German Academy of Sciences in Teltow-Seehof, where he started his carrier as an assistant under E. Correns. Later he became a department head and finally the director of the Institute (1969 to 1981 and again 1990 to 1992). At the time the Institute experienced a remarkable broadening of the scientific profile and was renamed Institute for Polymer Chemistry (IPOC) in 1972, belonging to the East German Academy of Sciences (AdW der DDR). After his habilitation (Dr. habil.) in 1956 at the TH Dresden he has been appointed to a professorship of the AdW der DDR and became an adjunct professor for chemistry and polymers, teaching first at Magdeburg and later at the Technical University of Dresden.

Professor B. Philipp has rendered outstanding services to the IPOC-Institute also during the times of the political changes in the 1990s in particular with regard to the re-organization of the institute. He was re-elected as the director in 1990 by the vote of an independent council of IPOC scientists and held this position until 1992. Under his guidance the various research activities of IPOC were screened, bundled and transformed into the appropriate research organizations already existing in West Germany. The Max-Planck-Institute of Colloids and Interfaces as well as the Fraunhofer-Institute for Applied Polymer Research IAP became the largest of the successive Institutes, both nowadays located at Potsdam-Golm. Well into old age, Prof. Philipp has been personally engaged in the work of the Institute for Biomaterials Research of the Helmholtz-Center Geesthacht, the largest Institute remaining in Teltow-Seehof. The model-like successful integration of a large East-German Institute into the scientific landscape of the unified Germany was an extraordinary achievement of B. Philipp which has been recognized by the Federal Cross of Merit (Bundesverdienstkreuz erster Klasse) in 1993.

Cellulose research and polyelectrolytes were the major scientific fields of B. Philipp. Together with his co-workers he published more than 500 papers contributing significantly to the chemistry and technology of the viscose process, the derivatization of cellulose at homogeneous and heterogeneous conditions, the dissolution of cellulose in unconventional solvents, the degradation of cellulose and to the structure and reactivity of this natural polymer. In particular during the late years of his career he worked on regioselective functionalization of cellulose. Furthermore, he dealt with analytics, structure formation, and applications of polyelectrolytes. In a remarkable manner he used and supported interdisciplinary approaches between chemistry, physics, and engineering. His scientific inheritance also covers a number of textbooks which he initiated or promoted as a co-author, including „Grundlagen der makromolekularen Chemie“ (Philipp and Reinisch; Akademie-Verlag, 1976), „Polyelektrolytes“ (Dautzenberg, Jaeger, Kötz, Philipp, Seidel, Stscherbina; Hanser, 1994) as well as the two-volume standard textbook „Comprehensive Cellulose Chemistry“ (Klemm, Philipp, Th. Heinze, U. Heinze, and Wagenknecht; Wiley-VCH, 1998).

In addition to his outstanding scientific achievements he was a demanding and supporting mentor for many young scientists and passed on his experience and the interest in the fascinating and sustainable polymer cellulose. Besides a large number of PhD-students, more than 10 appointments to regular and honorary professorships originated from his school.

Professor B. Philipp was a member of the East-German Academy of Sciences and received, both before and after the German reunification, many awards. He became honorary member of the Kolloidgesellschaft (1992), received the Hermann-Staudinger Award of the GDCH (1994), the International Schwarza-Lyocell-Award (2000), among others. He was an active member of ZELLCHEMING and was one of those who shaped the expert committee of cellulose and cellulose derivatives including the "Cellulosechemiker-Rundgespräch" (Cellulose symposium) for many decades. Already in 1981 he was awarded the Alexander-Mitscherlich-Medal of ZELLCHEMING.

After retirement, despite a serious and progressive disease leading to the loss of his eyesight, he lost neither his mental and intellectual activity nor his open mindedness. In a remarkable way his wife Helga assisted him all the time and helped him to manage his health handicap. Well into his old age, he had been expressing his views kindly and frankly both to the ZELLCHEMING expert group and the Berlin-Brandenburg Association for Polymer Research. For many of his co-workers and students he was a guide and role model with regard to his optimistic and positive way of life, his daringness, efficiency, tenacity, as well as his straightforwardness.

With B. Philipp we lost a great scientist, an outstanding scientific manager, a gifted teacher, and an esteemed colleague.

Prof. Dr. Hans-Peter Fink, Potsdam-Golm
Prof. Dr. Klaus Fischer, Dresden
Prof. Dr. Thomas Heinze, Jena

reprint from: ipw, No. 4-5/2015, p 48, by courtesy of the Association of Chemical Pulp and Paper Chemists and Engineers (ZELLCHEMING Association, Carl-Zeiss-Str. 3, 64331 Weiterstadt, Germany).

A New Sensor System Employing a Mid-infrared Quantum Cascade Laser for Direct Carbon Disulfide Measurement in an Industrial Environment

Johannes Paul Waclawek[1], Roman Schmuck[2], Roland Kroismayr[2], Thomas Röder[2] and Bernhard Lendl[1]

[1] Institute of Chemical Technologies and Analytics, Vienna University of Technology, Getreidemarkt 9/164-UPA, 1060 Vienna, Austria
[2] Lenzing AG, Werkstraße 2, 4860 Lenzing, Austria
e-mail: johannes.waclawek@tuwien.ac.at, bernhard.lendl@tuwien.ac.at

Abstract

A compact gas sensor system based on quartz-enhanced photoacoustic spectroscopy (QEPAS) employing a continuous wave (CW) distributed feedback quantum cascade laser (DFB-QCL) emitting mid-infrared (mid-IR) radiation at 4.59 μm was developed for detection of carbon disulfide (CS_2) at sub-ppmv concentration levels. The work reports the suitability of the sensor system for monitoring CS_2 in process streams at the rayon industry, where this molecule is used in big amounts to produce regenerated man-made cellulose fibers.

Introduction

Quartz-enhanced Photoacoustic Spectroscopy (QEPAS)

Laser-based trace gas spectroscopy is of big interest for a widespread field of applications, including environmental monitoring, medical diagnostics or industrial process control. Strong fundamental absorption features of chemical species in the gas phase can be found in the mid-infrared (mid-IR) region, which can be accessed by quantum cascade lasers (QCLs). QCLs provide a complete coverage of the mid-IR region and have been demonstrated to be very suitable for a large variety of spectroscopic techniques. [1 - 4] A trace gas sensing technique featuring unique properties is called quartz-enhanced photoacoustic spectroscopy (QEPAS). [5, 6] It is based on the principle of photoacoustic spectroscopy (PAS) where the absorption of modulated optical radiation by the analyte leads to a periodic heating of the sample. The periodic heating of the sample in turn causes a periodic thermal expansion, which leads to a periodic pressure change in the media. In this way pressure waves are generated and can be detected by an acoustic transducer. Conventional PAS employs sensitive microphones placed into resonant cells. QEPAS, however, uses a quartz tuning fork (QTF) as a sharply resonant piezoelectric acoustic transducer with an extremely high quality factor, instead of a broadband electric microphone and a relatively low quality-factor resonant photoacoustic cell. The QTF is a commercial mass product used as frequency standard in clocks and watches resonating at 32,768 (= 2^{15}) Hz in vacuum. The QTF is a low-loss piezoelectric element and therefore converts its deformation, caused by generated pressure waves, into separation of electrical charges that can be measured either as voltage or current. Due to the small size of the QTF the QEPAS technique facilitates the measurement of trace gases in an ultra-small acoustic detection module (ADM) with a total effective sample volume of only a few mm^3 in contrast to classical PAS. Further merits of QEPAS are that only the fundamental symmetric vibration of the QTF is piezoelectric active, i.e. when the two prongs bend in opposite directions in the plane of the QTF. That results in excellent environmental noise immunity, because sound waves from distant acoustic sources tend to move the QTF prongs in the same di-

rection, thus resulting in no electrical response. Moreover, the QTF is a rugged element which possesses a very large dynamic range. The detected QEPAS signal is directly proportional to the absorption coefficient per unit concentration of the target species, the concentration of the target species, the laser power, the quality-factor of the acoustic resonator and is inversly proportional to the QTF frequency f_0. [7] An enhancement of the QEPAS signal can be achieved when two metallic tubes acting as a micro-resonator (mR) are added to the QTF sensor architecture. Figure 1 shows a typical QEPAS arrangement where the laser beam is focused between the prongs of the QTF and the acoustic resonator tubes in order to probe the acoustic waves and achieve the highest electric signal. [8] In addition, the detected QEPAS signal is improved for slow relaxing molecules by water vapor, which is known as an efficient catalyst for vibrational-translational (V-T) relaxation processes in the gas phase. [9]

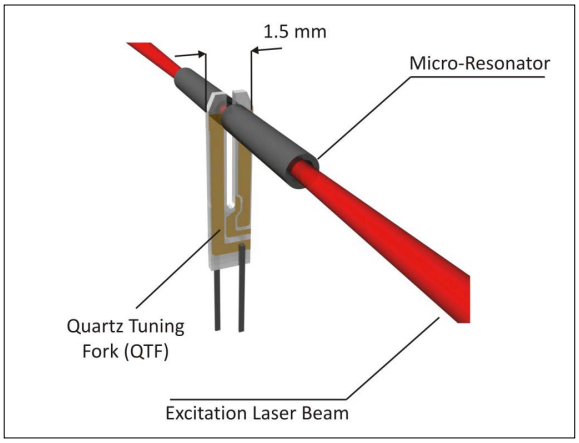


Figure 1. *Typical QEPAS spectrophone configuration.*

The industrial production of fibers out of the natural raw material wood (man-made cellulose fibers) is a chemical-technological process that proceeds in multiple steps. Sulfur containing gases such as carbon disulfide (CS_2) and dihydrogen sulfide (H_2S) are generated during the spinning process. The aim of this work was the development of a portable gas sensor system capable for detection of CS_2 in the process stream of fiber production down to low ppmv concentration rages. The applicability of the CS_2 QEPAS gas sensor for on-line measurements was tested at Lenzing AG (Lenzing, Austria). Furthermore, the results obtained were compared with a method for CS_2 detection based on photoionization (PID).

Wavelength Modulation Spectroscopy

Modulation of the output wavelength of a diode laser can be easily accomplished by modulation of the injection current. This tuning mechanism is very rapid, whereby modulation frequencies up to hundreds of MHz can be applied. The modulation of the injection current results in modulation of the laser frequency and is always accompanied to some extent by modulation of the amplitude, as the injection current also controls the optical output power. In this content amplitude modulation generally acts as an undesirable effect that distorts the signal and is referred to as residual amplitude modulation (RAM). Modulation techniques using frequencies that are much smaller than the half-width of the employed laser source are denoted as wavelength modulation spectroscopy (WMS). [10] The interaction of the target species with the modulated light leads to the generation of signals at the modulation frequency and various harmonics. The analytical signal can then be detected at a suitable harmonic of the modulation frequency, which in most cases is the second harmonic, by use of lock-in amplifiers. [11] One important advantage of modulation techniques is to shift the detection to higher frequencies so that $1/f$ noise is reduced. In case of PAS detection an acoustic signal only evolves when light is being absorbed and no signal is ideally created in the absence of the analyte. Application of WMS thus efficiently allows to get rid of any spurious acoustic signal providing an offset free signal at harmonic detection. [12]

Experimental

CW-DFB-QCL Performance and CS_2 Wavelength Selection

In this work a high heat load (HHL) packaged continuous wave (CW), distributed feedback quantum cascade laser (DFB-QCL) (HHL-14-45, AdTech optics) emitting at ~ 4.59 μm was employed as a compact and powerful light source generating up to more than 75 mW output power. The DFB-QCL operated at a single mode frequency and could be tuned over a few wavenumbers by varying the QCL either by temperature or injection current.

Figure 2 shows a measured absorption spectrum of CS_2 within the spectral region from 6500 to 600 cm^{-1} [13] with a blow-up of the spectral region from 2200 cm^{-1} to 2140 cm^{-1} which was chosen for targeting CS_2. The strongest absorption bands of CS_2 are located in the spectral region between 1550 cm^{-1} and 1500 cm^{-1}, with the strongest line centered around 1541.5 cm^{-1}. Unfortunately, no DFB-QCL covering this line was commercially available back in 2012 when this project was started. In order to perform sensitive CS_2 QEPAS measurements the absorption line centered at 2178.69 cm^{-1} (Figure 2, inset, red arrow) and a QCL operating temperature of 19.45 °C was selected, because of its relative high line intensity and no CO and H_2O interference as well as high laser power.

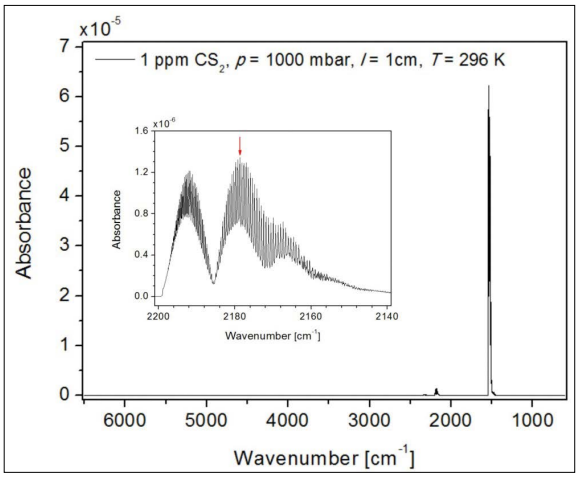


Figure 2. *Measured spectra of CS_2. [13]*

QEPAS Sensor System Architecture and Performance

The optical platform of the QEPAS sensor is schematically depicted in Figure 3a. It employs a high heat load packaged CW-DFB-QCL as light source and the QEPAS gas cell, consisting of the QTF and acoustic resonator tubes. A reference cell filled with 30 % of CS_2 in N_2 at 75 mbar and a MCT-detector located after the ADM were used as a reference channel in order to lock the laser frequency to the center of the selected CS_2 absorption line. The sensor platform is based on 2*f* wavelength-modulation spectroscopy (WMS) and QEPAS detection. The QEPAS based sensor system consists of the optical setup, the gas line with metering valve, QEPAS cell (ADM), pressure controller, micro-vacuum pump and sensors (humidity, pressure, temperature), electronics, main board and display for QCL control, data acquisition and processing (Figure 3b).

In order to implement the 2*f*-WMS technique the DFB-QCL emission wavelength was tuned across the CS_2 absorption line centered at 2178.69 cm^{-1} by applying a ramp to the laser current and modulating it sinusoidally at half of the QTF resonance frequency ($f mod = f_0/2$=16.384 kHz). Once the optical energy is absorbed by the gas, the acoustic wave is generated and detected by the QTF. The induced QTF piezoelectric signal was enhanced by an ultra-low noise trans-impedance amplifier (TA) with a 10 MΩ feedback resistor. The amplified QTF signal was demodulated at f_0, using an internal lock-in amplifier with a time constant set to 1 sec. A CS_2 reference cell and a photodetector signal demodulated at 3*f* were used in order to lock the laser frequency to the center of the selected CS_2 absorption line.

The presence of H_2O vapor influences the QEPAS response to CS_2 by enhancing the V-T energy transfer rate and thus increasing the PA signal. The dependence of H_2O on the response of the QEPAS was acquired as a function of the H_2O concentration. A ~3.2 times improvement of the QEPAS signal was observed when the water content of the analyzed gas mixture was 2.6 %. The sensors sensitivity to the concentration of the trace gas component in a specific gas mixture is also a function of the sample pressure. Moreover, the laser wavelength modulation depth *m* must be optimized for a certain pressure in order to achieve highest 2*f* WMS signal amplitude. Optimum working pressure and modulation depth for a reference gas mixture moisturized with 2.3 % H_2O was found to be p = 75 mbar and m = 0.026 cm^{-1}. Measurements were carried out with continuous gas flow through the ADM. Constant pressure and gas flow within the ADM was accomplished by

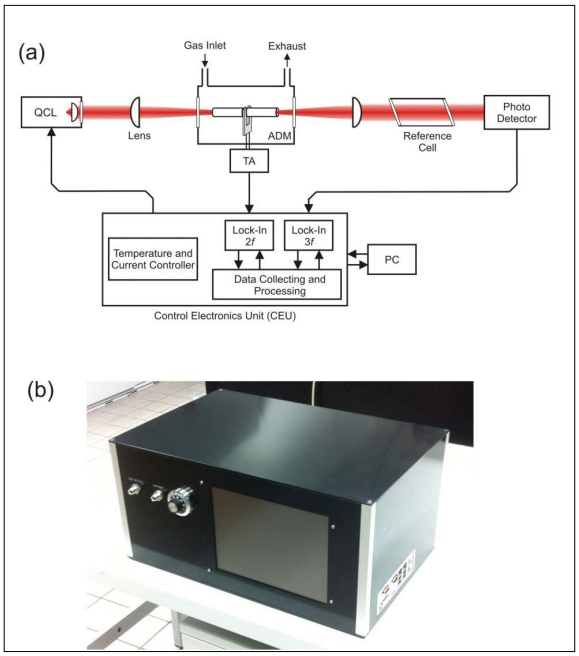


Figure 3. *(a) Schematic diagram of the QEPAS based gas sensor architecture employing a CW-DFB-QCL, (b) QEPAS system suitable for gas monitoring at industrial environment.*

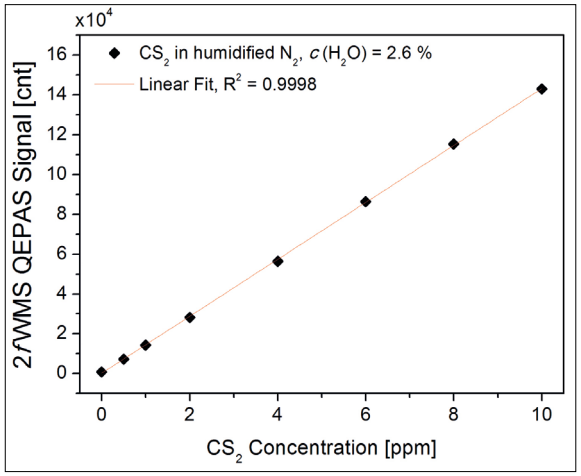


Figure 4. *Dependence of 2f WMS QEPAS signal as a function of CS_2 concentration.*

means of the micro-vacuum pump, pressure controller and metering valve. Sensitivity and linear response of the QEPAS sensor was investigated by performing quantitative measurements of CS_2 in humidified N_2 within the range from 0 to 10 ppmv at optimum operation conditions. Excellent linearity between signal amplitudes and CS_2 concentrations was observed and a sub-ppmv detection limit was determined for the QEPAS based sensor system (Figure 4).

Performance Test at Lenzing AG (Lenzing, Austria)

To demonstrate the suitability of the QEPAS system for continuous detection of CS_2 directly out of the process stream the developed system was installed at Lenzing AG for a few days. The QEPAS sensor does not show any spectroscopic interference at the selected CS_2 absorption line, thus the sample gas could be extracted directly out from the gas stream without any pre-treatment. QEPAS detection was carried out at optimum working conditions in scan and locked mode. In the scan mode, the DC component of the QCL current is slowly tuned so that laser frequency sweeps over the desired spectral range in order to acquire spectral information of the gas sample. Data recorded with the scan mode were used to verify the selectivity of the sensor in the chosen spectral range. In the locked mode, the QCL frequency is locked to the center of the CS_2 absorption line at 2178.69 cm^{-1} by means of a reference cell and a photodetector detector signal demodulated at 3*f* in order to avoid any laser drift. Influence of varying water contents was compensated by simultaneous measurement of the moisture by a humidity sensor. As a reference a photoinization detector (SPID, Analytical Control Instruments Gmbh, Berlin, Germany) was used. The applied PID uses a 10.6 eV cathode lamp to ionize molecules, which are then detected by an electrode. That means all vaporous organic compounds can be detected if they are hit by a photon with higher energy then the minimum ionization energy of the molecule. In this way a broad range of volatile organic compounds can be detected by a PID. Because of the cross-sensitivity to a number of gases, selective measurements of CS_2 can only be performed after a pre-treatment of the sample. The primary interferent with PID detection at the observed process streams was H_2S. In order to perform selective measurements of the target gas the sample was purged through a zinc acetate solution to remove H_2S before PID detection. Disregarding increased maintenance requirements at industry, the pre-treatment raises the sample gas volume, which makes it impractical to detect fast concentration changes of the target molecule.

Figure 5 shows exemplary results of a parallel QEPAS and PID measurement of CS_2 over a few hours. The diagram shows that the concentration trend of both systems coincides. The PID sensor, however, is not as well capable to detect fast concentration changes and does show a retarded detector response compared to the QEPAS system. This behavior is due to the bigger sample gas volume generated by the sample pre-treatment.

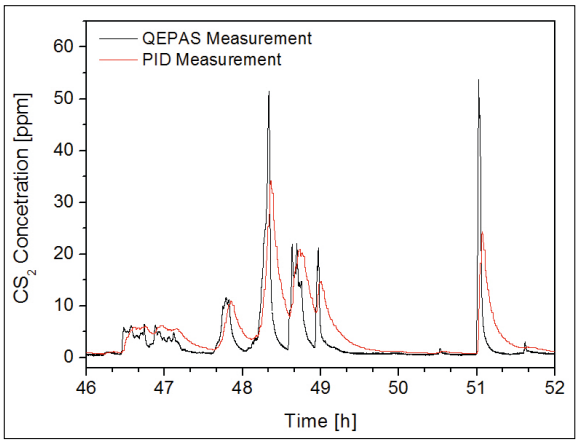


Figure 5. *Comparison PID versus QEPAS, black: QEPAS measurement directly probing sample gas from the process stream, red: PID measurement of pre-treated gas sample.*

Conclusions

The principles of a new sensor system for CS_2 based on mid-infrared quartz-enhanced photoacoustic spectroscopy (QEPAS) have been successfully developed and tested in a research laboratory environment at TU Wien. Subsequently, a rugged and portable sensor system for CS_2, providing sensitive and selective measurements of CS_2 molecules in the industrial gas streams with fast detector response, was developed. This portable sensor system was successfully tested at Lenzing AG during a measurement campaign in 2014. Comparison with the established reference method based on photoionization detection (PID) revealed a clearly faster response of the new QEPAS sensor as well as comparable concentration results. Important advantages of the developed QEPAS sensor system are seen in its high selectivity, which allows continuous direct pre-treatment free measurement of process gases. This not only eliminates additional maintenance, it also guarantees selective concentration readings in addition to a fast response. Because of the generic nature of the proposed sensor system, other analytes can be measured, too always in case a ro-vibrational line of the analyte can be targeted by a CW operated DFB-QCL.

Acknowledgements

JPW and BL acknowledge financial support provided by the Austrian research funding association under the scope of the COMET program within the research network "Process Analytical Chemistry" (contract # 825340) as well as the Carinthian Tech Research. Furthermore the autors would like to thank Lenzing AG for their inkind contributions.

References

[1] L. Gong, R. Lewicki, R. Griffin, F. K. Tittel, C. R. Lonsdale, R. G. Stevens, J. R. Pierce, Q. G. J. Malloy, S. A. Travis, L. M. Bobmanuel, B. L. Lefer, J. H. Flynn: Atmospheric Envi. 77, 893-900 (2013)

[2] S. Solga, T. Schwartz, M. Mudalel, L. Spacek, R. Lewicki, F. K. Tittel, C. Loccioni, T. Risby: J. Breath Research 7, (2013)

[3] A. Varga, Z. Bozóki, M. Szakáll, G. Szabó: Appl. Phys. B 85, 315-321 (2006)

[4] G. Gagliardi, H. P. Loock, Cavity-Enhanced Spectroscopy and Sensing (Springer, London, 2014)

[5] A. A. Kosterev, Y. A. Bakhirkin, R. F. Curl, F. K. Tittel: Opt. Lett. 27, 1902-1904 (2002)

[6] 1. R. F. Curl, F. Capasso, C. Gmachl, A. A. Kosterev, B. McManus, R. Lewicki, M. Pusharsky, G. Wysocki, F. K. Tittel: Chem. Phys. Lett. 487, 1–18 (2010)

[7] A. A. Kosterev, F. K. Tittel, D. Serebryakov, A. L. Malinovsky, I. Morozov: Rev. Sci. Instr. 76, 1-9 (2005)

[8] P. Patimisco, G. Scamarcio, F. K. Tittel, V. Spagnolo: Sensors 14, 6165-6206 (2014)

[9] T.L. Cottrell, J.C. McCoubrey, Molecular Energy Transfer in Gases (Butterworths, London, 1961)

[10] F.K. Tittel, D. Richter, A. Fired, Book chapter "Mid-Infrared Laser Applications in Spectroscopy" in "Solid-State Mid-Infrared Laser Sources" (Springer, 2003)

[11] P. Kluczynski,O. Axner: Appl. Opt. 38, 5803-5815 (1999)

[12] S. Schilt, L. Thévenaz: Infrared Phys. Technol., 48, 154-162 (2006)

[13] PNNL quantitative infrared database for gas-phase sensing, www.nwir.pnl.gov

Bioprocess Monitoring using an Inline Viscosity Sensor

Stefan Clara[1], Aydin Golabgir[2], Martin Heinisch[1], Christoph Herwig[2] and Bernhard Jakoby[1]

[1] Institute for Microelectronics and Microsensors, Johannes Kepler University Linz, Altenberger Str. 69, A-4040 Linz, Austria, e-mail: ime@jku.at, stefan.clara@jku.at
[2] Research Area Biochemical Engineering, Institute of Chemical Engineering, Vienna University of Technology, Gumpendorferstrasse 1a - 166/4, A-1060 Vienna, Austria

Abstract

Inline monitoring of bioprocesses plays an important role for modern manufacturing processes. Higher production rates and better product quality can be reached by tightly monitoring and controlling the progress of the biochemical reactions. The performance of the process control system crucially depends on the quality and on the quantity of (independent) process parameters provided by different types of sensors. New sensing principles increase the knowledge on the biochemical reactions and open up opportunities for specific interventions. In this contribution we present the sensor development process for a resonant inline viscosity measurement system based on three U-shaped wire sensors. We outline the sensor principle, the physical effects, the prototype device and first measurement results obtained in a bio-fermentation process.

Introduction

In modern biochemical production processes, a large number of different process parameters can be measured and controlled [1]. Therefore, many different sensors are integrated into bioreactors to monitor the actual state and transfer the data to the process control system. New sensors provide better data or even new independent data that were not available thus far, allowing further improvement of the production process control and leading to higher efficiency, better process understanding, and therefore higher product quality [2]. Inline viscosity sensors offer different possibilities in biochemical processes. Viscosity is an additional property that is not measurable with other established methods, such as infrared spectroscopy. Special sensors as resonant viscosity sensors offer additional information due to their working principles, providing measurements in new rheological regimes (compared to rheometers used in laboratories) or particle classification [3].

Materials and Methods

Sensor Principle

The used sensor is the so called U-shaped wire sensor [4]. It consists of three parts: a tungsten wire bend in a U-shape, two brass sockets to hold the tungsten wire and a neodymium permanent magnet to provide a magnetic field in the region of the tungsten wire (Fig. 1 & Fig. 2).

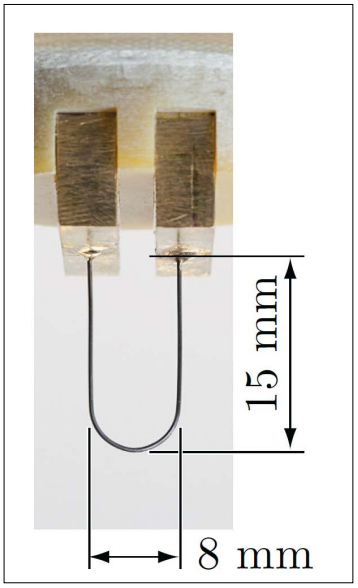


Figure 1. *U-shaped Wire sensor consisting of a tungsten wire mounted on two brass sockets (the permanent magnet is visible in Fig. 2).*

Enforcing an AC current flowing through the wire, an alternating force (with the same frequency) is acting on the tungsten wire (Lorentz force). This force leads to a periodic oscillation of the wire. The motion of the wire in a magnetic field leads to an induced voltage proportional to the velocity of the wire and the magnetic field (motion induction). By performing a frequency sweep around the mechanical resonance frequency associated with this U-shaped beam and measuring the induced voltage, the characteristics of the system and specially the exact resonance frequency f_r and the quality factor Q can be obtained. Immersing the sensor into a viscous fluid, these two values are changed depending on the viscosity-density product of the fluid [4].

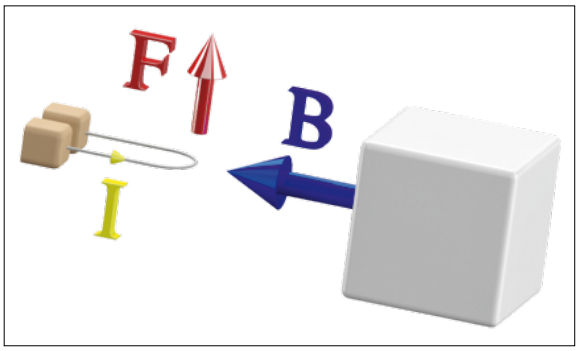


Figure 2. *Schematic for the working principle. The sensor is placed in a B-field B generated by a permanent magnet. An AC current I is driven through the tungsten wire and due to the current in a magnetic field a force F is acting on the wire.*

Measurement Principle

The working principle of resonant viscosity sensors is based on a surface performing sinusoidal (often in-plane) motions in a viscous fluid. Due to the viscosity and the density of the fluid, an attenuated shear wave is generated at the surface which penetrates into the liquid ([5] & [6]).

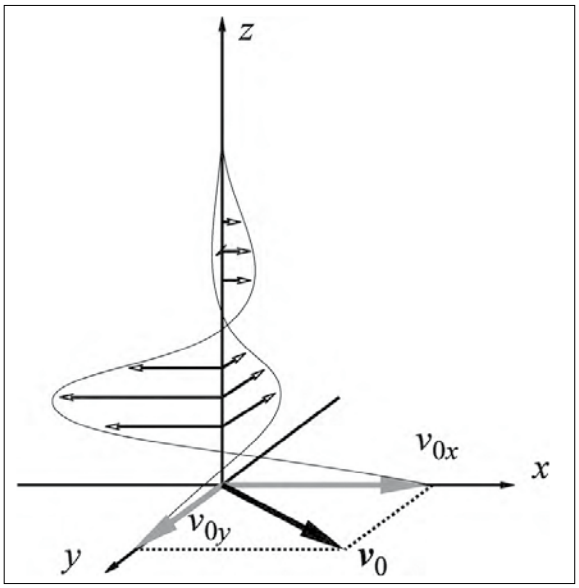


Figure 3. *A surface oscillating in x- or y-direction causes, due to the viscosity and density of the fluid, an exponentially damped shear wave penetrating into the liquid [6].*

The penetration depth δ of the shear wave is dependent on viscosity, density and oscillation frequency [7]:

$$\delta = 2\eta\rho\omega, \qquad (1)$$

where δ, η, ρ and ω denote penetration depth, dynamic viscosity, density, and radian oscillation frequency, respectively.
Fig. 4 shows the penetration depth described in (1) for different oscillation frequencies. This information is especially interesting if particles or contaminations of the measured fluid are present.

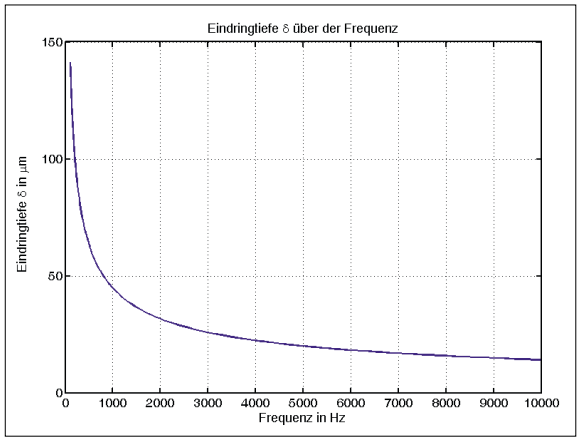


Figure 4. *Illustration graph for the penetration depth over oscillation frequency for constant viscosity and density.*

The standard procedure for measurements is to perform a frequency sweep with the sensor and measure the fre-

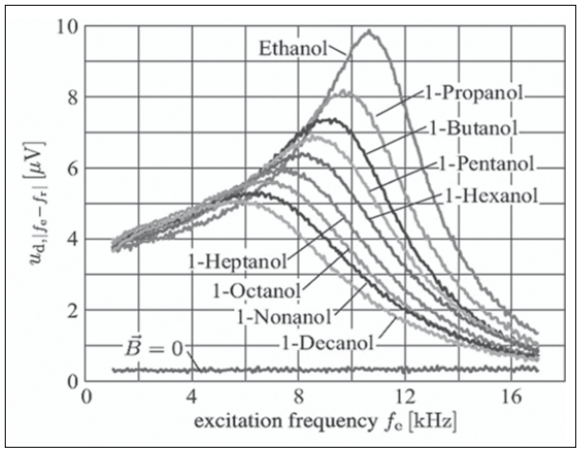


Figure 5. *Frequency response of a resonant (in-plane vibrating plate) viscosity sensor measuring different fluids with different viscosities. The higher the viscosity, the lower is the quality factor and also the resonance frequency [8].*

quency response. Then the resonance frequency f_r and the quality factor Q are extracted from the data. Fig. 5 shows a series of measurements for different fluids (from ethanol, with the lowest viscosity of 1.1 mPa s, to decanol, with 12 mPa s). Note that the resonance frequency f_r and also the quality factor Q are decreasing with increasing viscosity η. As a last step a calibration measurement set is used to calculate back to the viscosity. If only the detection of changes is of interest, the last step is not necessary.

Due to the frequency dependency of the penetration depth of the shear wave, additional measurement applications are feasible. Resonant viscosity sensors probe only a very thin fluid layer atop of their surface. Particles distributed in a liquid influence the macroscopic viscosity, but effectively interact with the shear wave only if the size of the particle is small compared to the penetration depth (Fig. 4 & Fig. 6, [7]). If more than one resonant sensor is used, the gained viscosity values are depending on the particle sizes immersed into the fluid. Sensors with lower oscillation frequencies sense large and small particles however sensors with higher oscillation frequencies are only influenced by small particles.

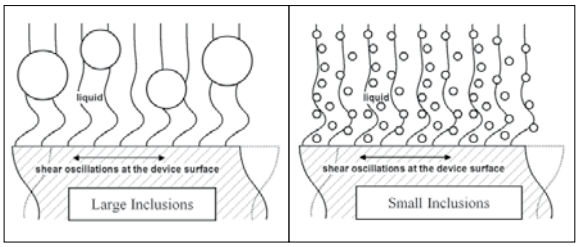


Figure 6. *Interaction of different particle sizes with the penetrating shear wave. Particles which are large compared to the penetration depth of the shear wave do not interact with the shear wave, small particles instead change the resonance characteristics of the sensors and can be sensed [7].*

For particle detection, very accurate measurements are necessary.

Setup

For the probe head itself three U-shaped wire sensors were integrated into a stainless steel casing (see Fig. 7). Each U-shaped wire is soldered on two brass bars and the brass bars are directly connected to the readout cables. A four wire connection reduces the influence of the cables. In front of the U-shaped wires with a distant of 1 mm a cylindrical neodymium permanent magnet is placed to provide the magnetic field required for the excitation and the readout. Due to its low impedance, the sensor is not affected by the conductivity of the surrounding fluid. This qualifies the sensor also for measurements in highly conductive fluids.

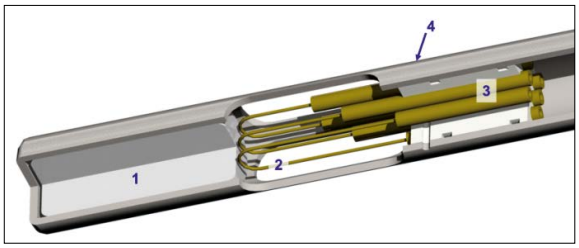


Figure 7. *Probe head containing three U-shaped wire sensors. 1) Permanent magnet, 2) U-shaped tungsten wire, 3) brass sockets and 4) stainless steel housing.*

Figure 8. *Picture of the probe, note the different length of the U-shaped wire sensors leads to different resonance frequencies for each sensor.*

One of the most important parts of the sensor probe is the sealing. During the bioreactor sterilization process pressures up to 2 bar and temperatures over 121 °C are reached. Therefore the holding of the brass bars is made out of PEEK and is sealed to the casing by two special high temperature O-rings made of VITON. To withstand the high pressure the holding is also fixed to the casing by a bolt which runs through the PEEK material and is welded on both sides to the casing. This form-locking manufacturing technique provides the necessary robustness for the intended purpose.

Actuation and Readout

A professional audio interface for PCs was used for the actuation signal generation and also for the readout of the output signal. This is possible due to the frequency range of the mechanical resonators, which lies in the kHz regime. Audio interfaces offer a lot of advantages in this case, e.g., high sampling rate of 192 kHz, very high resolution 24 Bit and easy data transmission to any PC.

For the readout of the sensors, a frequency sweep has to be performed for each sensor separately. In principle, also a simultaneous measurement for all sensors is possible, but leads to higher input amplitudes, due to the crosstalk between the sensors, and was therefore not implemented. A sinusoidal signal with a constant frequency is

applied to the first sensor and simultaneously the output of the sensor is recorded by a line-in channel of the audio interface. As these audio interfaces always provide stereo channels the second output is directly connected to the second input. This input is then used as a reference signal for the phase ϕ and allows the reconstruction of the phase shift between the input and the output signal. In a next step the recorded data is crop to eliminate the settling effects of the signal. A sinus estimator is then used to obtain the amplitude and the phase out of both measured channels. The second direct connected signal is used as a reference for $\phi = 0$. After that the next frequency is processed. When the whole frequency response is measured, a fitting algorithm is used to obtain the resonance frequency f_r and the quality factor Q out of the resonance curve [9]. Both parameters are related to the viscosity density product of the surrounding fluid. The results of an air measurement for all three sensors are shown in Fig. 9, Fig. 10 and Fig. 11.

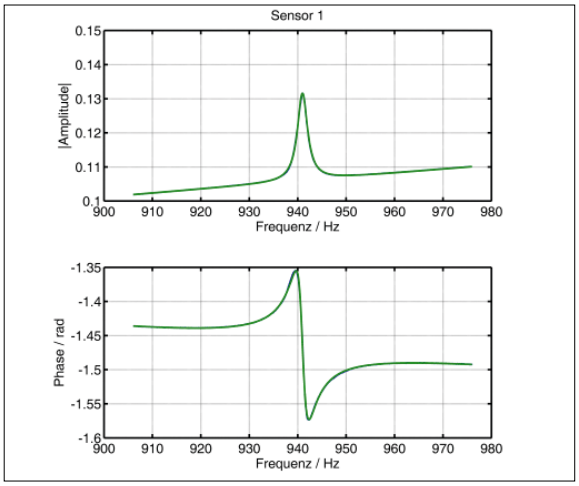


Figure 9. *Frequency response of the largest sensor with the lowest resonance frequency. The measurement was taken in air. The resonance frequency is 941 Hz, the quality factor at 400.*

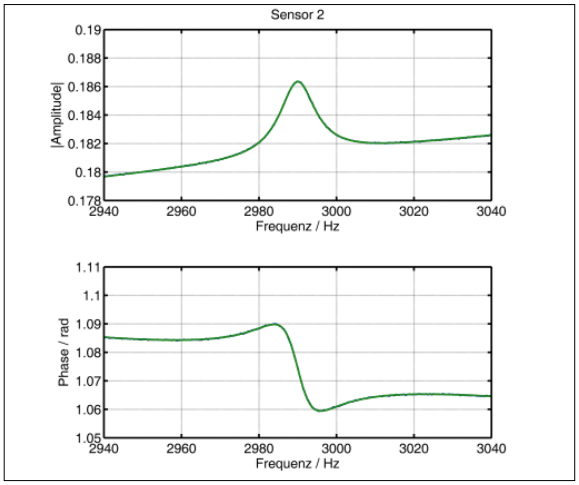


Figure 10. *Frequency response of the second sensor with a resonance frequency of 2990 Hz and a quality factor of 270 measured in air.*

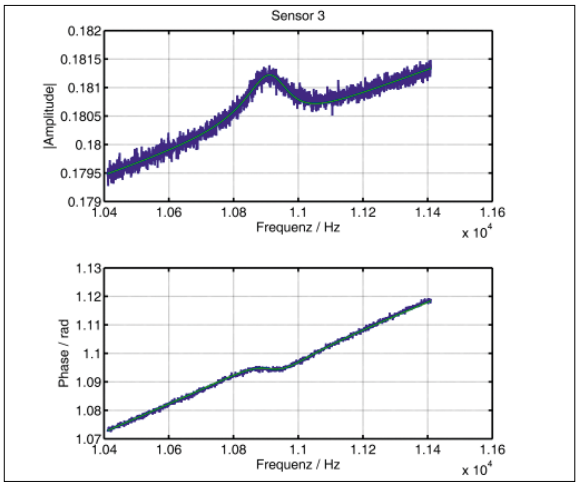


Figure 11. *Frequency response of the smallest sensor with the highest resonance frequency. The resonance frequency is at 11 kHz and the quality factor at 60 measured in air.*

Fig. 12 and Fig. 13 show the stability of the quality factor and the resonance frequency for a longer time period in air.

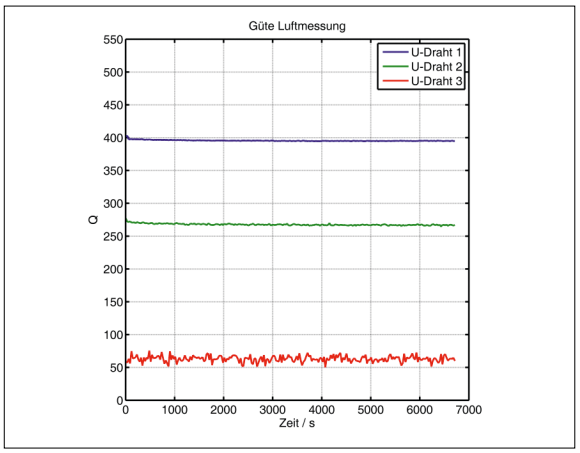


Figure 12. *Quality factors Q for all three sensors in air for a longer time period.*

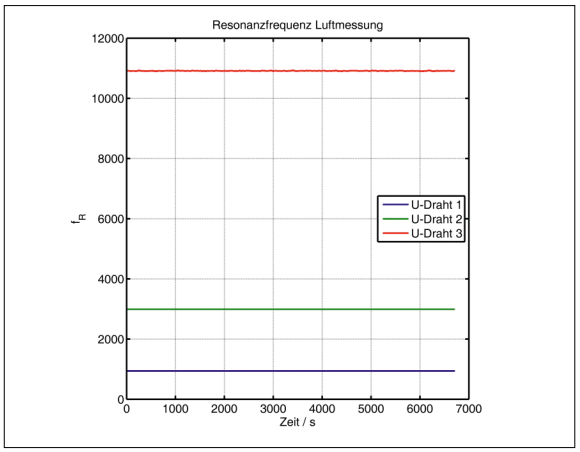


Figure 13. *Resonance frequency fr for all three sensors in air for a longer time period.*

Results and Discussions

Sample Measurements

First some test measurements were performed offline with a series of test samples obtained from a *P*. chrysogenum process [10]. A problem with these measurements is the sedimentation of cells after a short time. This sedimentation process could be observed in the measurements (Fig. 14) as the U-shaped wire sensor was immersed into the sample from above, the tip of the sensor was about 1 mm away from the permanent magnet which defined the bottom of the measurement chamber.

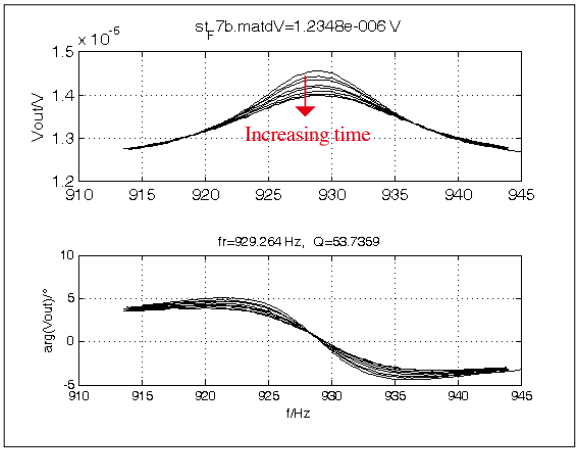


Figure 14. *Shows the resonance curve for test sample measurements over time. The quality factor Q is decreasing, the resonance frequency fr stays very constant.*

The accumulation of the cells on the floor of the measurement chamber leads to a higher viscosity in that region which is also the part of the sensor with the highest sensitivity. Due to this sedimentation effect, at-line measurements where avoided.

Inline Measurements

The first observed production process was a fed-batch culture of *P. chrysogenum*. Due to the filamentary structure of these fungi the probe head was contaminated after a very short time (Fig. 15). After drying and cleaning, the sensor head was again ready for use.

Figure 15. *Probe fouling due to the filamentary structure of the fungi in a P. chrysogenum production process.*

For a second process, we performed measurements in an *E. coli* culture [11], where the cell structure is round and does not show filamentous growth. Here we had no problems with attached cells on the sensor surface. The reference values calculated from the measurements (the measurement results for the change of the quality factor Q and the resonance frequency f_r were scaled and calibrated to the first offline biomass concentration analyses), showed a clear correlation to the biomass concentration determined as dry weight (Fig. 16 & Fig. 17). For both example measurements (f_r and Q), the values at the beginning of the fermentation (lower biomass concentration) correlate better with the concentrations determined by dry weight analyses. The measurements using the resonance frequency f_r (see Fig. 16) as the parameter related to the biomass concentration lead to a mean relative error of 12% and a standard deviation of 0.07. For the measured quality factor Q (Fig. 17) the obtained mean relative error is 18% and the standard deviation is 0.13. Note, however, that the data base for these analyses was fairly limited.

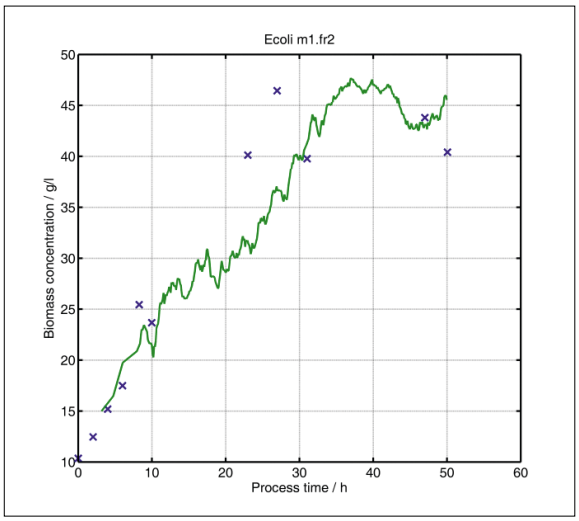


Figure 16. *Measured change of the resonance frequency fr (relating to the resonance frequency at t = 0, solid line) compared to the biomass concentration determined by dry weight analyses (blue crosses).*

Due to the large vibrations of the test reactor the measured signal was very noisy, but as the sampling rate of the sensors is extremely high compared to the offline measurements, a smoothing filter could be applied easily.

Conclusions

We presented a resonant viscosity sensor development for biochemical cell density measurements. The sensor probe consists of three electromagnetically actuated U-shaped wires. The induced voltage is measured and used to reconstruct the resonant frequency response of each sensor consecutively. Out of this measured curves the

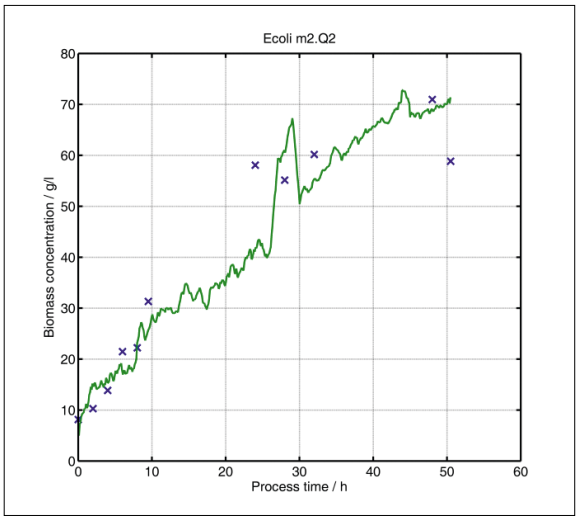


Figure 17. *Measured change of the quality factor Q (relating to the quality factor at t = 0, solid line) compared to the biomass concentration determined by dry weight analyses (blue crosses).*

quality factor Q and the resonance frequency f_r were extracted and used as parameter for the biomass concentration. Both parameters show clear correlations to the concentration. Further improvements e. g. higher robustness against spurious vibrations and a redesign of the sensor housing are necessary to reduce the influence of vibrations on the sensor due to the stirring or the vibrations of the motor.

Acknowledgements

This work was supported by the research network "Process Analytical Chemistry" (PAC, grant 825340), the Austrian Center of Competence of Mechatronics (ACCM), and the Austrian Science Fund (FWF, grant L657-N16).
Financial support was also provided by the Austrian research funding association (FFG) under the scope of the COMET programme within the research project "Industrial Methods for Process Analytical Chemistry - From Measurement Technologies to Information Systems (imPACts)" (contract # 843546). This programme is promoted by BMVIT, BMWFW, the federal state of Upper Austria and the federal state of Lower Austria.

References

[1] P. Biechele, C. Busse, D. Solle, T. Scheper, and K. Reardon, "Sensor systems for bioprocess monitoring," Eng. Life Sci., vol. 15, no. 5, pp. 469–488, Jul. 2015.

[2] P. Kroll, P. Sagmeister, W. Reichelt, L. Neutsch, T. Klein, and C. Herwig, "Ex situ online monitoring: application, challenges and opportunities for biopharmaceuticals processes," Pharmaceutical Bioprocessing, vol. 2, no. 3, pp. 285–300, Jun. 2014.

[3] B. Jakoby, R. Beigelbeck, F. Keplinger, F. Lucklum, A. Niedermayer, E.K. Reichel, C. Riesch, T. Voglhuber-Brunnmaier, B. Weiss, "Miniaturized sensors for the viscosity and density of liquids-performance and issues" Ultrasonics, Ferroelectrics and Frequency Control, IEEE Transactions on, vol.57, no.1, pp.111,120, Jan. 2010, doi: 10.1109/TUFFC.2010.1386

[4] M. Heinisch, E. Reichel, B. Jakoby, "U-Shaped Wire Based Resonators for Viscosity and Mass Density Sensing", Proceedings SENSOR 2013, pp. 52-57, doi: 10.5162/sensor2013/A2.3

[5] S. J. Martin, G. C. Frye, A. J. Ricco and S. D. Senturia, "Effect of surface roughness on the response of thickness-shear mode resonators in liquids", Analytical Chemistry, vol. 65, pp. 2910-2922, 1993, doi: 10.1021/ac00068a033

[6] E. K. Reichel, C. Riesch, F. Keplinger, C. E. A. Kirschhock, B. Jakoby, "Analysis and experimental verification of a metallic suspended plate resonator for viscosity sensing", Sensors and Actuators A: Physical, Volume 162, Issue 2, August 2010, Pages 418-424, ISSN 0924-4247, doi:10.1016/j.sna.2010.02.017.

[7] Jakoby, B., Ecker, A. and Vellekoop, M.J., "Monitoring macro- and microemulsions using physical chemosensors", Sensors and Actuators A: Physical, col. 115, pp.209-214, 2004, doi: 10.1016/j.sna.2004.02.010

[8] C. Riesch, E. K. Reichel, A. Jachimowicz, J. Schalko, P. Hudek, B. Jakoby, and F. Keplinger, "A suspended plate viscosity sensor featuring in-plane vibration and piezoresistive readout", J. Micromech. Microeng., vol. 19, pp. 075010/1–10, 2009. doi: 10.1088/0960-1317/19/7/075010

[9] A.O. Niedermayer et al, "Methods for the robust measurement of the resonant frequency and quality factor of significantly damped resonating devices", 2012 Meas. Sci. Technol. 23 085107, doi:10.1088/0957-0233/23/8/085107

[10] A. E. Posch and C. Herwig, "Physiological description of multivariate interdependencies between process parameters, morphology and physiology during fed-batch penicillin production," Biotechnol Progress, vol. 30, no. 3, pp. 689–699, May 2014.

[11] P. Sagmeister, P. Wechselberger, M. Jazini, A. Meitz, T. Langemann, and C. Herwig, "Soft sensor assisted dynamic bioprocess control: Efficient tools for bioprocess development," Chemical Engineering Science, vol. 96, pp. 190–198, Jun. 2013.

Self-Adaptive Non-linear Methods for Improved Multivariate Calibration in Chemical Processes

Carlos Cernuda[1], Edwin Lughofer[1], Thomas Röder[2], Wolfgang Märzinger[3], Thomas Reischer[4], Marcin Pawliczek[5] and Markus Brandstetter[5]

[1] FLLL, Johannes Kepler University Linz, Austria
[2] Lenzing AG, Werkstraße 2, 4860 Lenzing, Austria
[3] i-Red GmbH, A-4040 Linz, Austria
[4] Metadynea Austria GmbH, Hafenstraße 77, 3500 Krems, Austria
[5] RECENDT – Research Center for Non-Destructive Testing GmbH, 4040 Linz, Austria

This publication gives a comprehensive overview of our joint works during the FFG funded K-Project PAC on the basis of new methodologies for chemometric calibration models and their evaluation in real-world application scenarios.

Abstract

In this paper we present new methods for non-linear multivariate calibration and their application under real-world conditions. The developed non-linear methods, which are applied to FT-NIR absorbance spectra recorded in-line in different industrial production processes (i-Red GmbH and RECENDT GmbH) result in enhanced model quality and robustness. We propose two new concepts for reducing the dimensionality of the calibration problems, which may get severe when several hundreds or thousands of wavelengths are contained in the spectra. One is based on a statistical approach using a modified variant of forward selection, but extending it to extract bands instead of single wavelengths. Thereby, the robustness with respect to noisy recordings is increased. This concept is termed as forward selection with bands; the other one is a wrapper method which is based on a global heuristic search process achieved through genetic algorithms. Internally, they employ a new fuzzified crossover operator in order to weight nearby-lying bands accordingly. The calibration phase is equipped with an own developed non-linear version of PLS (partial least squares) on the basis of Takagi-Sugeno fuzzy inference systems offering the possibility to define piece-wise linear predictors which are combined to a non-linear model through Gaussian kernels, representing fuzzy rules. We will further demonstrate methods how to incrementally adapt the non-linear version of PLS over time with new incoming samples in order to account for significant system dynamics with different outweighing strategies. This is essential to assure high stability and predictive performance of the calibration models in case of dynamic processes over a long timeframe.
The application potential of these methods will be underlined by several results achieved from a viscose fiber production process (at Lenzing AG) and from a melamine resin production process (at Metadynea Austria GmbH). Active learning techniques in single-pass mode will play an essential role to keep target measurements and associated costs on a low economic level.

Introduction

The increasing demand for analytical methods that are capable of in-line process analysis in an industrial environment is accompanied by the need for improved tools for data processing and data analysis. These tools shall guarantee constant method quality and, at best, automatically adapt to varying conditions. Varying conditions can be caused, e.g. by changes in the product composition, cleaning steps, instrumental deficiencies or external influences, such as ambient temperature.

Among the available methods, Near-infrared (NIR) spectroscopy is an attractive approach for the analysis and monitoring of complex industrial production processes. NIR spectroscopy is non-destructive, reagent-

free and provides qualitative and quantitative information on the samples under investigation. The method is well established in process analytical applications and mostly applied in combination with multivariate data analysis tools [1]. The main purpose of these tools is to reduce the high dimensionality of NIR spectra. This reduction step becomes necessary, as NIR absorption bands of liquids and solids are relatively broad. Thus, each band will consists of many spectral data points which all provide more or less the same chemical information. Multivariate data analysis is furthermore capable of establishing quantitative calibration models by linking the chemical information contained in the NIR spectra to the corresponding quantitative information obtained by a reference method. In case of multi-component substances several parameters can be determined out of the same set of spectra, thereby saving time and resources. Most commonly, Partial Least Squares (PLS) regression analysis [2] is applied to establish quantitative multivariate calibration models (=chemometric models).

Considering the practical implementation of multivariate calibration models, it is often challenging to maintain constant model quality over a longer time period. Unexpected changes in the chemical composition or instrument drifts are represented in the acquired spectra and require manual supervision and adaptation of the calibration models. Such re-calibration steps require significant resources, as new calibration sample must be drawn from the process and manually analyzed by reference analysis. Therefore, methods that are able to continuously adapt to gradual or even sudden changes without a need for a re-calibration are highly desired. Another aspect of model robustness concerns the linear nature of conventional multivariate data analysis, which impedes proper modelling of non-linear effects. A practical example for such an effect are the spectral changes induced by temperature variation. Typically, a temperature shift is associated with a change of the absorption properties, including not only the substance of interest, but also the solvent, which is often used for the acquisition of the background spectrum. In the absorption spectrum, a changing background absorption can manifest as spectral shift of the sample spectra. Such shifts are difficult to model by standard linear methods. Furthermore, a drawback of using classical linear regression for model calibration is that a possible high nonlinearity and inhomogeneity among the independent variables (wavebands) directly affect the coefficients of these in the established regression/approximaton mapping/function in a misleading sense.

In this paper we present newly developed methods for multivariate modelling of NIR spectra focusing on self-adaption and non-linear modelling. Practical application of these methods is shown for several parameters in different industrial production processes.

Methods

Data-Driven Dimensionality Reduction in FT-NIR Spectra

FT-NIR spectra recorded at chemical production sites typically contain at least a few hundreds and often up to a few thousands of data points, where each data point represents wavelength dependent spectral information of the sample substance. Figure 1 shows a typical absorbance spectrum on the example of melamine-formaldehyde (MF) resin. Absorbance spectra are usually preferred over an intensity spectrum for model calibration, as potential undesired substances contained in the spectra (such as water content, for instance) are omitted by referencing the sample spectrum to a suitable background spectrum [3].

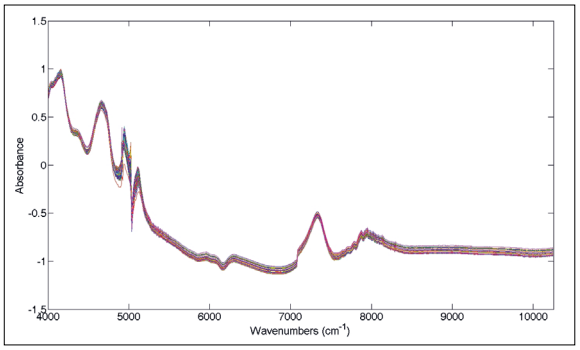


Figure 1. *Typical absorbance spectra measured at various process stages within a melamine resin production system and corresponding to several cloud points --- containing 6000 wavelengths = input dimensions for calibration models.*

When processing this raw data within a batch model calibration step in order to establish a mapping between these spectra and target values (usually concentrations of chemical substances appearing in continuous form (as real numbers)), each of these wavelengths is handled as one single dimension within the modelling technique. In particular, the input to the calibration step is given by a wave-matrix defined as

$$X^{N\times p}$$

where p is the dimensionality and N is the number of calibration samples drawn from the process. Typically, these are costly to gather, as especially their target values are quite time-intensive to be obtained due to high operators' efforts for a detailed analysis of the drawn probes etc. (see also Section 1); hence, $N << p$, which leads to an underdetermined regression modelling problem. Furthermore, usually N should not be only equal

to but much greater than p to guarantee a stable (near-optimal) solution [4]. Regularization is indeed a mechanism to improve and guarantee stability, however it usually deteriorates the quality of the solution in terms of optimality in the least squares sense [5] (quadratic error between observed and predicted target values), even though when the regularization parameter is properly determined (see [6] for a comprehensive survey on parameter choice methods).

A lot of methods have been proposed in literature to handle curse of dimensionality reduction in FT-NIR spectra [7] [8], often achieved through the concept of waveband selection/extraction [9]. We provide a summary of two self-developed strategies for improved optimality and robustness against substantial noise which may be contained in the spectra:

- A statistical and fast approach termed as Forward Selection with Bands (FSB) [10] for extracting wavebands instead of single wavelength using a Greedy-based step-wise selection scheme.
- A slow genetic-based heuristics search approach employing an enhanced crossover operator for retrieving the global optimum in terms of a specific fitness function [11].

Even though the first approach is fast, it should be emphasized that it is a Greedy technique which usually finds the global optimal solution in approximately 76% of the cases [12]. The extraction of bands guarantees a higher robustness against spectral noise, especially when the models are applied to in-line process data coming from the production system. The principal concept of FSB is visualized in Figure 2 below.

First, the most correlated wavelength x^{1xp} to the target vector $\vec{y}$ is sought. The adjacent wavelengths to the left and/or to the right are adjoined to form a complete band as long as one of the following conditions holds:

- The maximal number of wavelengths forming one band is reached (default 50).
- The quality of the original correlation model (measured in terms of R^2 [13]) cannot be further significantly improved by adding wavelengths.

Once one complete band is extracted, a regression model based on the so far selected bands (joined together) is set up using $\vec{y}$ as target. The contribution of the wavelengths to the real target is estimated and substracted from $\vec{y}$ => new $\vec{y}$. With the new $\vec{y}$ (remaining target information), the next most correlated wavelengths is sought from those which are not already included in any selected band. This is repeated until a maximal number of bands is selected, which is the only parameter to be optimized within a cross-validation scheme. For each selected band, several latent variable are extracted using PLS representing the main information about this band. These components (i.e. their scores on

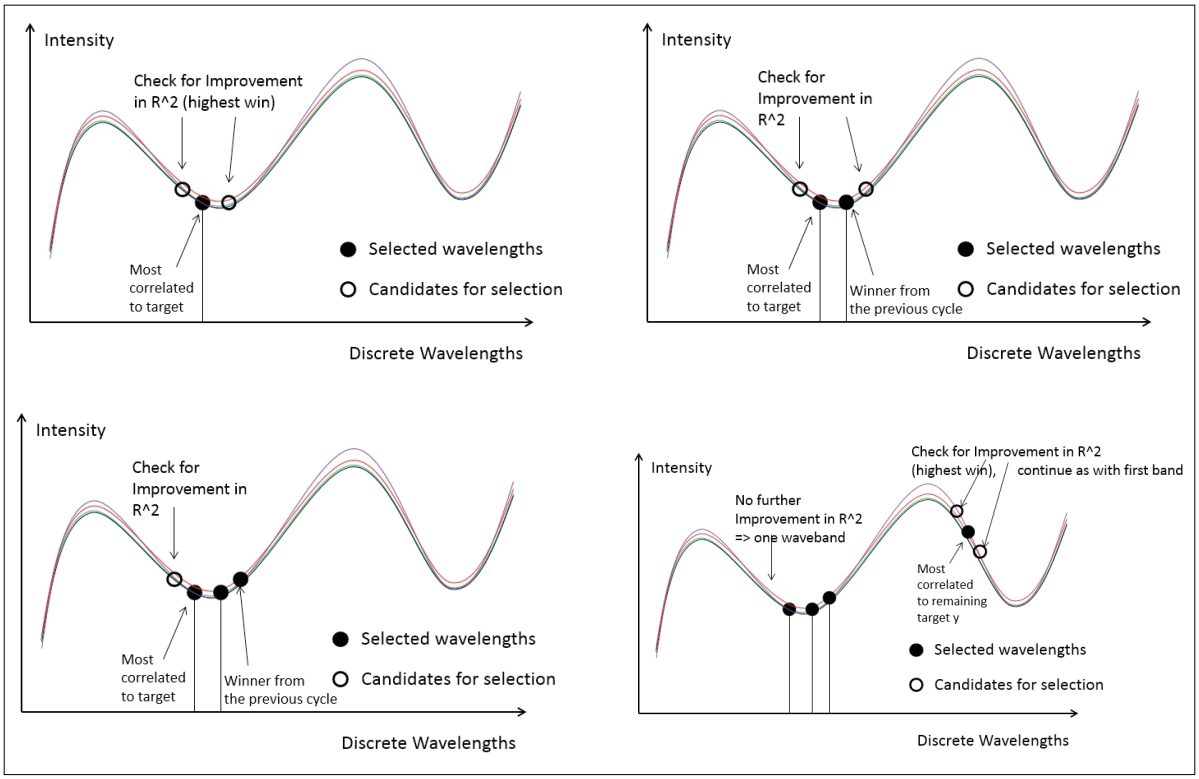


Figure 2. *Stepwise waveband construction according to Forward Selection with Bands (FSB) method (to be read from left upper to right lower corner).*

the original data set) are then used as different input dimensions to the non-linear modelling scheme described in the subsequent section.
For further details about the whole FSB algorithm, please refer to [10], Algorithm 1.
As mentioned above, that FSB finds the optimal solution only in about 76% of the cases due to its "Greedy nature", we tried to improve it towards a higher chance of finding the global optimum using genetic algorithms [14] [15]. In particular, we designed genetic operators on top of FSB integrating fuzzy weights to wavelengths nearby lying a selected one.
The weights assignments thereby follow the design of fuzzy partitions on top of the waves (bands) selected by FSB. Figure 3 provides an example.
This weights assignment finally influences the cross-over operator in a way that wavelengths with higher weights will have a higher chance to be included in the children of the cross-over process. Hence, higher weights are emphasized through the genetic process, but wavelengths with lower weights may still get a chance to become selected. This can happen also in combination with the bands already selected by FSB, thus non-optimal situations may be enhanced while solutions far away from that one achieved by FSB are more unlikely to come true.
The fitness function calculation is achieved through a 10-fold cross-validation error [16]
of a calibration model extracted based on the selected wavebands, as encoded in each individual. This can be a standard PLS model or a more enhanced non-linear model, depending on the selected model architecture (=> wrapper approach for the selected model [17]). This is also a principal difference to the well-known iPLS method [18]
and other state-of-the-art methods such as [19] [20] or [21], which are more embedded and intervened with the concrete calibration model architecture chosen.
For details about the whole genetic-based selection design, please refer to [22].

Non-linear Version of PLS with the Usage of Takagi-Sugeno Fuzzy Systems

Many chemometric modelling techniques exist which operate on a fully linear basis, i.e. which provide regression and prediction models which are linear in their parameters, see [23] [27] [7]. They have been most widely applied since the 80ties. One of the most prominent techniques is PLS regression analysis [2], which serves as foundation for many chemometric modelling software tools (such as the well-known PLS toolbox by Eigenvector Research GmbH, see http://www.eigenvector.com).
However, more and more non-linear effects gain importance in chemical production processes due to increased complexity in terms of mutual interferences of single parameters in multi-component systems, environmental influences, lot sizes, etc.
Therefore, we developed a non-linear version of PLS, which is termed as PLS-FLEXFIS or FLEXFIS+PLS, as building upon a flexible fuzzy inference system (FLEXFIS) [25]. These are flexible in the sense to be able to represent the actual non-linearity degree as requested from the process resp. as implicitly contained in the process. Thereby, PLS acts as a filter to gain the most important principal component directions contained in the data. Afterwards, the fuzzy system is trained based on the scores of the calibration data.
The reason for choosing a fuzzy system as non-linear approximator is four-fold:

- It is able to serve as piecewise linear approximator, which has a funded motivation in statistical research theory, also having some synergies to the well-known and widely used local weighted regression technique (LWR) [26]. The difference is that it provides a global model for the whole calibration set, rather than to re-build a local model

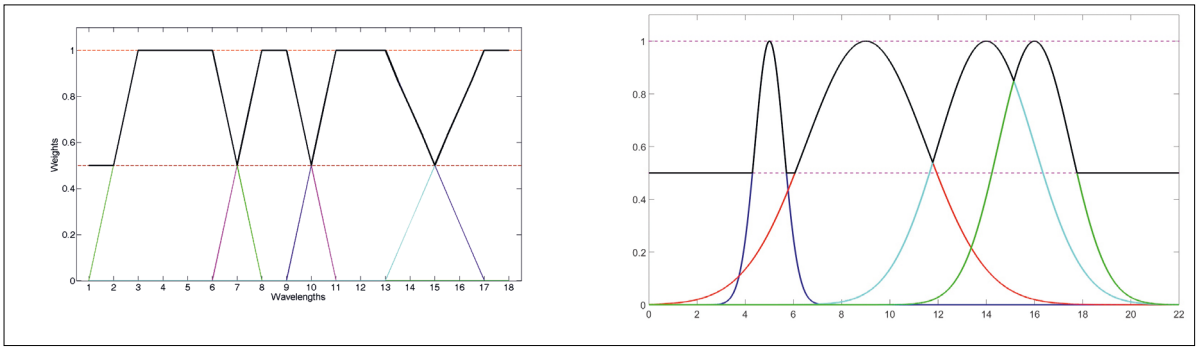


Figure 3. *left: Example for 18 variables and 4 wavebands of widths 4, 2, 3, and 2. Apart from the maximum weight 1 for wavelengths in the bands, i.e. 3, 4, 5, 6, 8, 9, 11, 12, and 13; wavelengths number 1, 7, 10, and 15 get 0.5, and wavelengths number 14 and 16 get 0.75; right: a similar scheme in case when using Gaussian fuzzy sets instead of trapezoidals.*

from scratch, which is more time-consuming and uninterpretable.

- It is known to be an universal approximator, being able to model any non-linear relationship contained in the data with a sufficient accuracy [27].
- It is able to automatically provide a global linear (PLS) model in case when there is no non-linearity contained. Thus, it acts as a real superset method to PLS.
- Finally, it may offer nice interpretability aspects as the components and piece-wise local predictors can be represented in linguistic form in terms of fuzzy rules [28].

In particular, the functional form of a Takagi-Sugeno fuzzy model is given by [29]:

$$\hat{f}(\vec{x}) = \hat{y} = \sum_{i=1}^{C} l_i \Psi_i(\vec{x}) \qquad \Psi_i(\vec{x}) = \frac{\mu_i(\vec{x})}{\sum_{j=1}^{C} \mu_j(\vec{x})} \tag{1}$$

with $\vec{x}$ the input vector (in our case when using PLS as filter, the projected input vector = the scores according to the p most important latent variables), $\mu_i(\vec{x})$ the membership degree of $\vec{x}$ to the ith rule and $l_i(\vec{x}) = w_{i0} + w_{i1}x_1 + \ldots + w_{ip}x_p$ the piecewise local linear predictor for the ith rule, which are combined through the normalized membership function to obtain a non-linear smooth surface.

Once the scores S=X_{proj} are extracted from the calibration set X according to the most important latent variables, i.e.

$$Lat_p = \left[\vec{lat}_1 | \cdots | \vec{lat}_p\right] = X * W_{X,p} \Rightarrow \vec{x}_{proj} = \vec{x} * W_{X,p}$$

our linear algorithm searches for the optimal number of clusters in the data set, whereas one cluster is associated with one rule in the principal component space. This is achieved with an extended version of vector quantization, termed as evolving vector quantization [30], recently extended to be able to model ellipsoidal rules in arbitrary rotational position and equipped with dynamic split-and-merge operations [31] --- see Figures 4 (a) and (b) for a comparison between axis-parallel and arbitrarily rotated rules [32].

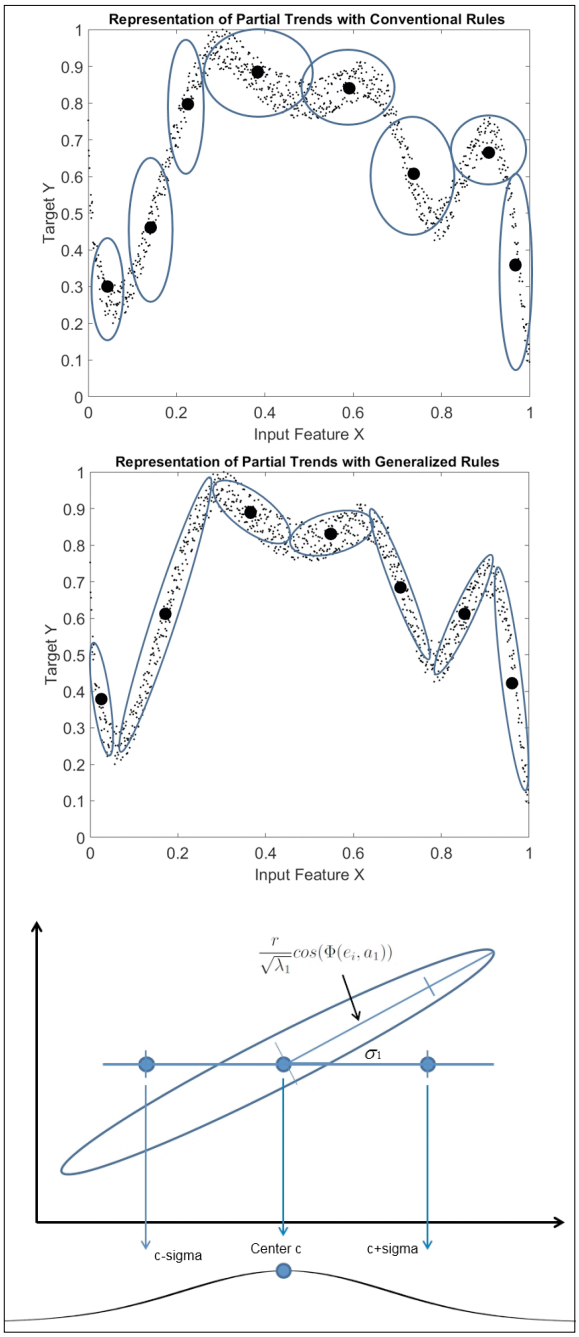


Figure 4. *left: clusters (=rules) extracted in axis-parallel position for modelling the partial linear trends of the overall non-linear (noisy, sinusoidal) relationship between input feature X and target y; middle: clusters (=rules) extracted in rotated position for a more compact and accurate modelling of the same relationship (following closer the partial linear trends); right: projection concept in case of arbitrarily rotated rules respecting its span along the principal components directions.*

This leads to rules defined by multi-dimensional Gaussian kernels:

$$\Psi_i(\vec{x}) = exp(-\frac{1}{2}(\vec{x} - \vec{c}_i)^T \Sigma_i^{-1} (\vec{x} - \vec{c}_i))$$

opposed to conventional ("old-school") axis-parallel rules achieved through the connection of fuzzy sets with a t-norm [33]; $\vec{c}_i$ denotes the center of the rule, $\sum_i^{-1}$ denotes the inverse covariance matrix, defining the shape of the rule. Such rules are called generalized rules, forming a generalized version of Takagi-Sugeno fuzzy systems. These have been recently shown to be able to significantly outperform conventional systems in [32]. The projection concept to form the rules adequately, Figure 4 (right image), again leads to fuzzy sets and finally to linguistically interpretable partitions [28] in the form of the Gaussian fuzzy sets as shown in the right plot of Figure 3.

Once the required rules are extracted, the linear consequent parameters from the piecewise local predictors $l_i(\vec{x})$ are estimated based on a weighted least squares approach, whose solution can be analytically given within a closed form:

$$\hat{\mathbf{w}}_\mathbf{i} = (R_i^T Q_i R_i + \alpha_i I)^{-1} R_i^T Q_i \mathbf{y}$$

This is conducted for each rule separately (termed as local learning), including a regularization parameter α_i for assuring a stable solution in case of matrix rank deficiency. Local learning has some advantages over global learning in terms of robustness and computation speed, as deeply analysed in [28] [34].

Adaptive Calibration Models for Dynamic Processes

In today's chemical processes and production sites, the presence of dynamically changing system behaviours, e.g. due to environmental influences, varying product settings or setups or changing compositions of substances, becomes more and more a necessary challenge to deal with when quantifying the substances using chemometric models [35].

For instance, in the viscose fiber production process (an application scenario which will be under our study in the experiments section below), operating modes and states of the industrial process might vary according to different products (i.e. fiber types) or raw materials processed, hence also the overall composition of the spin bath in detail might vary. In particular, the spin bath might contain additives or degradation products of the cellulose to a varying content [36], what is also conditioned by the dynamics of the industrial continuous operation mode and the recirculation of process media. This usually affects the NIR spectrum in terms of band shape and general appearance, in agreement with data recorded of samples from off-line calibration cycles.

The consequence is that previously calibrated models which worked fine for a batch calibration set and also for preliminary on-line operation modes (yielding new on-line validation data) may become easily outdated once the dynamics become more severe. Then, the model produce wrong quantifications and cannot be reliably maintained and used for condition monitoring and supervision purposes any longer. A real-world example of a significantly deteriorating performance over time of a calibrated static (conventional) PLS model (achieving a very high accuracy on the batch calibration set) is shown in Figure 5, left for sulfuric acid (H_2SO_4) and right for sodium sulfate (Na_2SO_4) contained in the spinbath.

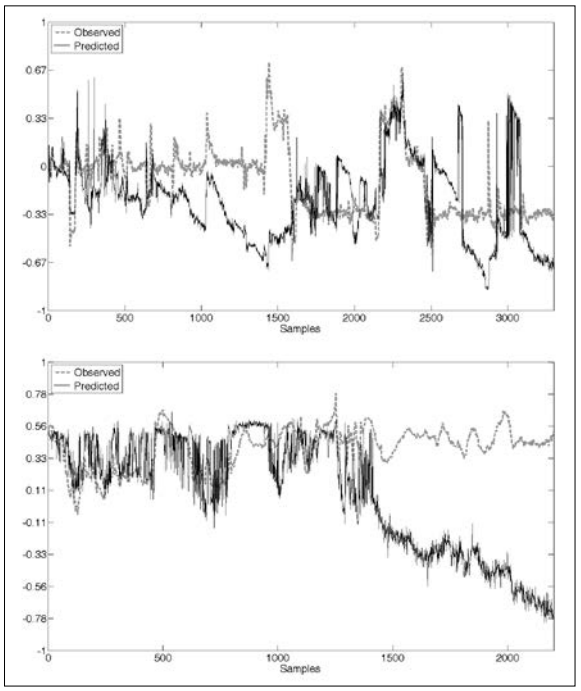


Figure 5. *left: prediction of H2SO4 based on a static calibration model over a timeframe of 2-3 months: the behavior of prediction is only acceptable during the first 300-400 samples, but then runs out of the rudder; right: the same problem for Na2SO4.*

Therefore, we developed a mechanism to adapt calibration models fully automatically 1.) on the fly and 2.) on demand:

- On the fly means that the adaptation is performed in a single-pass block- or sample-wise manner, assuring a fast processing without time-intensive retraining phases.
- On demand means that the model is updated only when it is really required to keep its accuracy/quality --- this is essential in case when target values are costly to measure during the on-line process and thus should be kept on a minimum level (see next subsection).

We developed this specifically for the fuzzy systems model architecture as described in the preliminary section, which resulted in the so-called eChemo learning paradigm [37] (short for evolving Chemometric Modeling). A particular challenge for an appropriate adaptation assuring high quality models is the stability-plasticity tradeoff [38], which on the one hand should assure significant plasticity of the model to change properly to new situations, but on the other hand should be conservative enough to guarantee model stability.

Plasticity is achieved through two concepts during model adaptation:

- Dynamically changing the structure of the fuzzy model according to the current non-linearity degree requested resp. according to the novelty content contained in the current on-line samples
- Outweighing older learned relationships over time to put more emphasis on the current system situation.

The first issue goes hand in hand with the possibility to evolve new rules in case of significant novelty content in current samples (providing the name eChemo) resp. to merge and prune older rules once they are becoming unimportant (e.g. due to redundant information present in the system [39]). The criterion for evolving a new rule is a statistical motivated tolerance region around the multi-dimensional Gaussians. This can be achieved by the prediction interval given by the χ^2 distribution according to [40]. The dimensionality of the learning problem has to be compensated as in case of higher dimensions the distances (to rules) tend to become higher, therefore the tolerance region is violated more easily. In sum, these considerations lead to the following criterion:

$$\min_{i=1,\ldots,C} \left(\sqrt{((\vec{x}-\vec{c}_i)^T \Sigma_i^{-1} (\vec{x}-\vec{c}_i))} \right) > r_i$$

$$r_i = fac * p^{1/\sqrt{2}} * \frac{1.0}{(1-1/(k_i+1))^{m_2}}$$

with p the dimensionality of the feature space, k_i the number of samples falling into rule i so far and fac a tuning parameter controlling the stability-plasticity trade-off (the only sensitive parameter in our method, usually tuned during an initial batch off-line training cycle). If the covariance matrix $\sum_i$ becomes diagonal (e.g. by updating only the standard deviations along each input direction), conventional axis-parallel rules are induced.

If this criterion is not met, the current model is updated, see below.

The second issue is addressed by a forgetting strategy, which is able to either exponentially outweigh older samples (termed as exponential forgetting) in an incremental single-pass manner [37] or to forget them completely (leading to a sliding-window based approach) [22]. The sample weighting strategy with different speeds of forgetting (according to the forgetting factor λ) and with complete forgetting (bold solid line) is visualized in Figure 6.

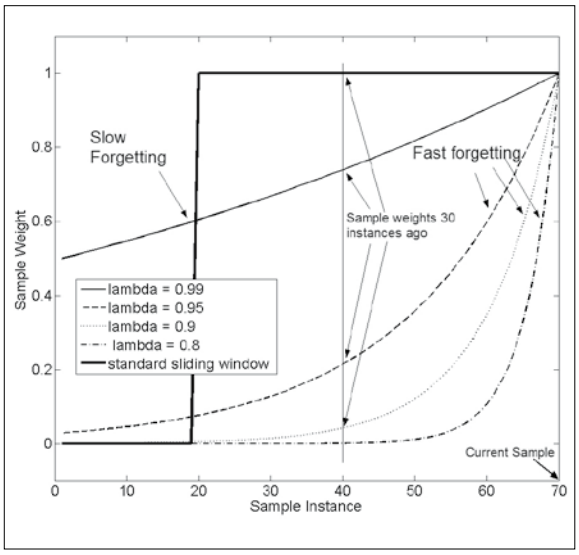


Figure 6. *Smooth forgetting strategies achieving different weights for past samples; compared to a sliding window with fixed width (bold solid line) => complete forgetting of older samples.*

Stability is achieved 1.) by a converging recursive consequent parameters learning, leading to a sub-optimality in the least squares sense, which is close to optimality subject to a constant; and 2.) by a updating the positions and ranges of influences of the rules with a decreasing learning gain over time, fulfilling the Robbins-Monroe conditions.

$$\sum_{N=1}^{\infty} \eta(N) = \infty, \sum_{N=1}^{\infty} \eta^2(N) < \infty$$

The former is a mathematically sophisticated topic as the conventional recursive (weighted) least squares approach (as used in past state-of-the-art adaptive chemometric modelling approaches [41] [42]) for updating the w_i's in the piecewise linear predictors of the ith rule, given by

$$\hat{\vec{w}}_i(k+1) = \hat{\vec{w}}_i(k) + \gamma(k)(y(k+1) - \vec{r}^T(k+1)\hat{\vec{w}}_i(k))$$

$$\gamma(k) = P_i(k+1)\vec{r}(k+1) = \frac{P_i(k)\vec{r}(k+1)}{\frac{\lambda}{\Psi_i(\vec{x}(k+1))} + \vec{r}^T(k+1)P_i(k)\vec{r}(k+1)}$$

$$P_i(k+1) = \frac{1}{\lambda}(I - \gamma(k)\vec{r}^T(k+1))P_i(k)$$

with $P_i(k) = (R_i(k)^T Q_i(k) R_i(k))^{-1}$ the inverse weighted Hessian matrix and $\gamma(k)$ the Kalman gain

[43], only converges when there are no structural changes in the model [44] [45]. However, this is not the case in our model, as rules and thus their membership degrees Ψ_i are updated. By introducing so-called correction terms added to the parameters before the next update cycles, we can balance out this non-optimal situation [25]. This finally achieves sub-optimality subject to a controllable constant according to an expected quasi-monotonic decreasing sequence of correction terms over time.

The latter concerns an appropriate update of the rule centers resp. their inverse covariance matrices, which can be achieved by vector quantization based update concepts (in case of the centers) resp. Neumann series based derivations of recursive (exact) updates in case of inverse covariance matrix, see [LS15] for details.

Some Enhanced Aspects for Improving Robustness and Usability of Calibration

Active Learning Paradigm as a Necessity for Useability in In-line Systems

One central challenge during model adaptation is given by the on demand characteristics. This means that the model should usually not be updated with each single recorded and processed measurement, but only from time to time whenever it is really demanded. This is necessary such that adaptive models can be reliable installed and used within on-line or even in-line production processes where either the target is not available resp. cannot be made available for each measurement (as the goal is to substitute cost-intensive analysis for obtaining target values), e.g. the titration automat in case of viscose fiber production at Lenzing AG.

Hence, it is of utmost importance to have an active learning component embedded, which is able to select the most essential samples for model updates on the fly and in (on-line) single-pass manner [46]. Usually, it is sought for samples

1. for which the model predictions are becoming very uncertain (termed as certainty-based sampling [47]) and
2. which contain significant diversity to the current knowledge already contained in the model.

The former can be measured by so-called model-based confidence intervals [13], indicating a 95% probability that the prediction lies in this interval having a certain width. The width can be seen as uncertainty range of the model output in form of a global (constant width over the whole input space) or local error bar (varying width over the whole input space). We designed such error bars for generalized Takagi-Sugeno fuzzy models, also to be able to produce enhanced output diagnostics, respecting different uncertainty levels in different parts of the input space, see [10] for the concrete formulas. An example of increasing error bars with increasing residuals (observed vs. predicted values) can be visualized in Figure 7.

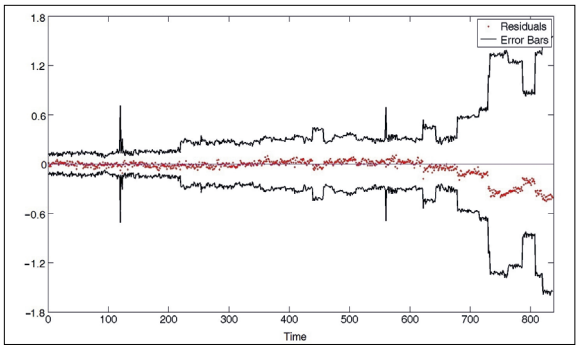


Figure 7. *Residual trends over time (red dots) when not updating the fuzzy model for cloud point prediction in melamin resin production, the black solid lines show the error bar (confidence interval) widths, opening up towards the end of this stream when residuals increase.*

The diversity can be measured in terms of the distance to the current principal component space (spanned by the latent variables extracted from PLS) with the support of T^2 and Q-statistics [48]. The former measures the distance of the sample to the center of the rotated data cloud within the component space, the later measures the projected distance to the principal component space spanned by the latent variables (both indicating novelty) - see Figure 8 for a particular example, where the red dot indicates a new incoming sample.

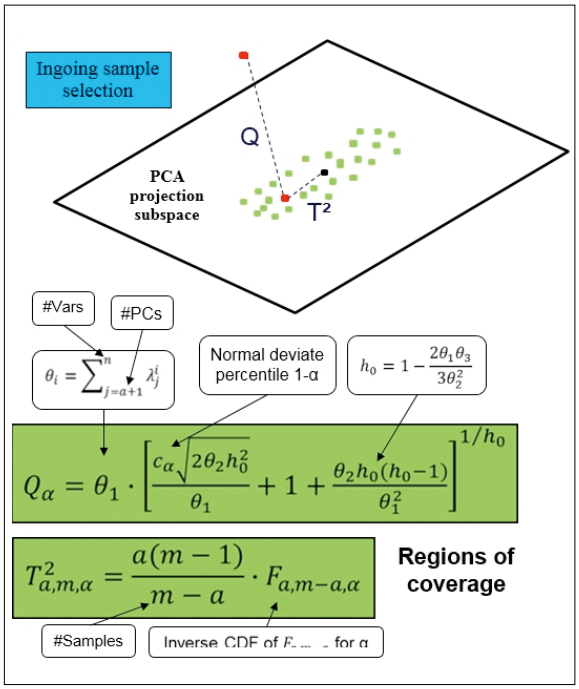


Figure 8. *The geometric interpretation of Hoteling and Q-statistics and the associated thresholds (Q_α and $T^2_{\alpha,m,\alpha}$) for sample selection based on statistical theory [49] [50].*

If one of these two distances exceed a specific threshold, which is statistically motivated as shown in formulas in Figure 8, the sample is used for model update. In case of classical eChemo paradigm with exponential forgetting, the sample is simply sent through the update mechanisms discussed above. In case of sliding window based re-training, the new sample is added to the window, while another sample is deleted from the current window:

- if it is the oldest, it leads to the classical sliding window approach.
- if it is another one selected by a more intelligent strategy using that one with lowest information quality, it may be able to gain some additional model accuracy as shown in [11].

Figure 8b shows the principal process work-flow of the adaptive chemometrics unit as will be embedded in the data acquisition framework shown in Figure 9.

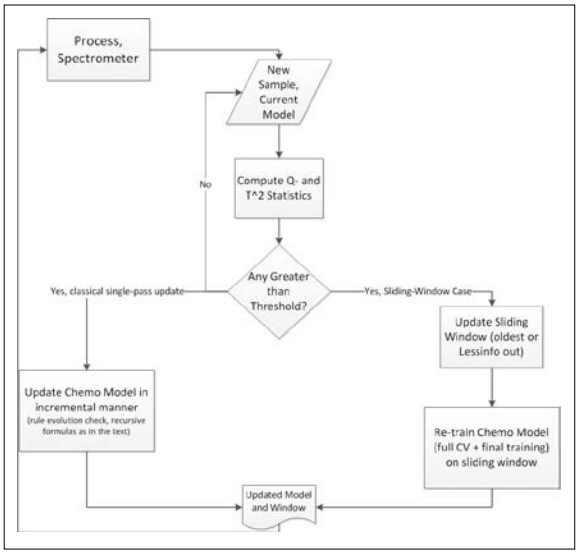


Figure 8b. *Adaptive chemometric model building work-flow in an on-line chemical process.*

Dynamic Principal Component Spaces

So far, only in case when there is a sliding-window based re-training conducted, the principal component space may change according to a full best parameter grid search over different number of latent variables. In the real single-pass adaptation mode, where each sample is used for model update and immediately discarded, afterwards, it has been assumed that the principal component space remains constant throughout the online adaptive modelling phase, once estimated through PLS during an initial batch calibration phase.

However, in reality the position of absorption bands, containing the essential sample information for producing a reliable mapping from the FT-NIR spectra to the targets, may also change [51]. A drift or shift in the process may arise due to a change in the equipment conditions (e.g. a decrease in the lamp intensity) or due to different production (e.g. spin-bath) cycles [36]. A particular example is the determination of sodium chloride in concentrated aqueous solutions where the shift of water absorption bands is the source of analytical information [52].

In order to address such dynamics, it is necessary to update the importance of the wavelengths resp. wavebands over time. This can be directly associated with an update of the principal component space, as the loadings of this space reflect the importance (levels) of the wavelengths in each component, especially in the most influencing components (explaining the largest variation in the data).

For doing so, we proposed an incremental update of the component space in unsupervised manner according to the following strategy [10]:

- Compute the residual vector of a new incoming sample to the component space, similarly as done for calculating the Q-statistics.
- Check whether the norm of the residual vector exceeds a certain threshold, as shown in Figure 8.
- If yes, a new component is added (starting with equal loadings over all wavelengths).
- If no, the component space is incrementally updated by solving a smaller intermediate eigenvalue problem and multiplying the solution matrix (containing eigenvectors as columns) with the current component space matrix, extended with the normalized residual vector having length 1.

This together with the recursive least squares update formulas as mentioned above (for updating the regression coefficients) finally resulted in an incremental Principal Component Regression method (incrPCR), applicable for dynamic linear problems. However, the dynamic expansion of the model structure in fully incremental manner (once the principal component space is expanded) is still a remaining challenge to be investigated (currently, models need to be re-trained when the input structure expands). Another restriction is that the principal component space can be only expanded, but not contracted, neither re-ranked (first principal component always stays the most important one).

Repeated Measurements Handling and Orthogonal Sensor Integration

In many chemical applications, measurements are repeated several times in order to increase the likelihood to receive valid data for particular system states and/or operation modes. This is because spectrometers may fail during their recordings due to several fallout reasons [53]. Once having the data collected, most of the repetitions turn out to be valid as the fallout rate is usually pretty low. Although they are repeated measures,

they may still deviate from each other (having slightly different amplitudes at several wavelengths positions) subject to the noise level contained in the recordings, which also affects the signal-to-noise-ratio of the data [54].

Exploiting the diversity in the different recordings by building a base model for each repetition set finally follows a similar concept as conducted in the well-known bagged ensemble (=bagging) approach [55] (following an exploration of the sample space) --- which is known to usually increase stability and outperform single base models [56].

Assuming to have m repetitions of measurements available, then m (non-linear) regression models $R_1, ..., R_m$ are built with the same concepts as mentioned in the previous sections. An expected model quality on separate validation data can be estimated by strategies such as cross-validation error (batch off-line case) or k-step-ahead prediction errors (incremental on-line case), using measures such as *R-squared(-adjusted)* [13]. These qualities $Q_1, ..., Q_m$ lying in [0,1] can be used as weights

$$w_1, ..., w_m = \frac{Q_1}{\sum_{i=1}^{m} Q_m}, ..., \frac{Q_m}{\sum_{i=1}^{m} Q_m}$$

for combining the predictions $y(R_1), ..., y(R_m)$ of all the regression models $R_1, ..., R_m$, i.e. the final model output $\hat{y}$ is obtained by:

$$\hat{y} = w_1 y(R_1) + w_2 y(R_2) + ... + w_m y(R_m)$$

This strategy has been successfully applied in cloud point prediction (see results section).

In some chemical applications, additional physical system variables (e.g. temperatures, pressures, emissions etc.) may be by-measured additionally to the spectra for enhanced supervision and monitoring purposes. This is usually achieved with so-called orthogonal sensors, additionally installed to the spectrometer. These are delivering numeric values (real numbers) directly indicating the actual state of the system.

Different concepts for combining and amalgamating such numeric measurements with the recorded spectra have been proposed in [10], also respecting different levels of influences of the various sources among each other (higher impact in calibration for spectra and vice versa).

Experimental Setup

Application Scenarios

We investigated three types of chemical applications for evaluation and test purposes of our methodological developments demonstrated in the previous section:

- Quantification of process parameters in polyetheracrylat (PEA) production, which are most important to assure a high quality of the final product, namely hydroxyl (OH) number, viscosity (PL) and acidity (acid number, SZ). Prior to the application of chemometric models, an off-line laboratory analysis by experts and operators has been required to perform a manual supervision of these substances (lasting over 2-3 hours for receiving one concrete value of one substance from a specific probe). The large range of different states in the process during operation leads to totally different compositions of the product. Due to the large variation in these compositions, significant non-linearity in the process is expected, such that a global linear mapping between spectra and target substances is expected to achieve a rather poor performance - an issue which we will verify in the results section when comparing non-linear fuzzy systems with PLS and other standard linear calibration methods.
- On-line prediction of the most important chemical substances for supervising the quality of the final product in viscose fiber production: H_2SO_4, Na_2SO_4 and $ZnSO_4$. The acid and the two salts govern the precipitation and agglomeration of the cellulose from viscose solution and the formation of the viscose fiber. The concentration of those components has a major influence on the fiber properties. Hence, the accurate knowledge and control of the concentration are a pre-requisite for the production of high-quality viscose fibers in the industrial processes. The conventional method to determine those concentrations is titration which is indeed an automated process, but quite time-consuming: it delivers a measurement each 10th minute, whereas the applied NIR spectrometer (see subsequent section) is able to record spectra every 10 seconds, being able to drastically reduce the reaction time onto spin bath recirculations. A complete substitution of the automated titration by chemometric models would be the ideal situation - due to the high system dynamics of the spin bath, self-adaptation of the chemometric models is expected to be required from time to time, omitting a disastrous performance deterioration as shown in Figure 5. Due to the high expense for performing titrations, a requirement of the company is to reduce measurements and thus the update cycles of the chemometric models. Ideally, the goal would be only to measure and update the model 3–4 times a day. This requires an explicit selection of those samples within this timeframe, which are seen as most important to keep the chemometric model on a high quality (or even to improve it).
- On-line prediction of the cloud point in melamine resin production process, which provides information about the progress of the batch condensa-

tion 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. Originally, the cloud point has been monitored manually, which took a high effort for operators, who had to walk to the basin, to draw a sample and analyze it in detail (until the appearance of a cloud in the sample when it cools down). Furthermore, no continuous visualization of the condensation process could be achieved. Initial tests with standard chemometric models (PLS) did not obtain a sufficient accuracy, showing either drifting residuals over a timeframe of several weeks (e.g. Figure 7) or some undesired, unexplainable peaks in the residuals. Non-linearity combined with self-adaptivity introduced in the calibration modeling stages should thus improve this situation. An additional challenge was to integrate measurements from additional sensors into the calibration phase, in order to exploit the fully available information about the state of the process. Furthermore, the noise levels in the three repeated measurements desired a specific treatment for assuring a (more) robust model building.

Data Acquisition

To realize a measurement setup, all components utilized for the process monitoring were installed directly in the chemical plants, which means that a real in-line application has been set up. The whole data acquisition and hardware setup is schematically shown in Figure 9.

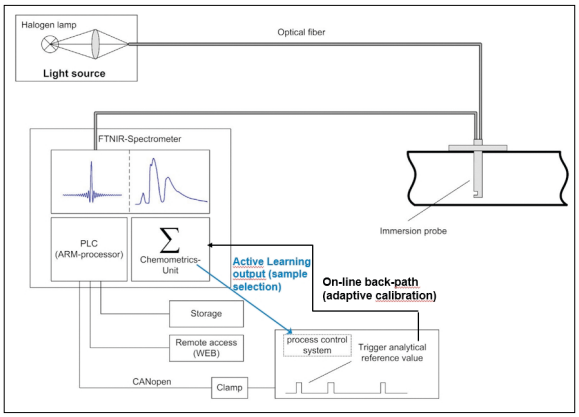


Figure 9. *Schematic view of the data acquisition framework as installed in-line in PEA production, viscose fiber production and melamine resin production, the chemometric unit highlighted by a sum-sign plays an essential role to quantify and predict chemical substances, which are then transferred to the process control system; this transfers back the reference signal to the chemometrics unit for model adaptation.*

The utilized immersion probe (Hellma) is connected to the FT-NIR process spectrometer and to a light source using fiber optics. The probe has an optical path length of 1 mm, with a measurement window processed out of sapphire. The FT-NIR process spectrometer [53] is constructed on the basis of a Michelson interferometer. In order to obtain a robust design, a monolithic interferometer [57] is utilized in the Michelson setup. The implemented programmable logic controller handles time critical signals as well as the communication
to, e.g., a network attached storage server (NAS). Especially for time evolving and adaptive systems, as considered for viscose fiber and melamine resin production, it is important to have an accurate timely synchronization between the analytical reference values and measured IR-spectra. For this task, a digital signal provided by the process automation system was read in via an automation clamp, which was connected to process spectrometer through the CANopen interface. Once the chemical analysis has been carried out, the reference values related to a certain spectral measurement (time of sample withdrawal) are saved automatically and thus fully synchronized with the NIR spectra.
The FT-NIR process spectrometer, whose maximal immersion depth is 690mm, the minimal one 30mm, the outer diameter 25mm and the pipe length of the probe 700mm, was continuously operated in a wavelength range from 3,900 cm^{-1} to 11,000 cm^{-1} with a spectral resolution of 3 cm^{-1}. The spectral region is defined by the extended InGaAs — infrared detector implemented in the system. The measurement rate was 10 measurements per second. The spectra were calculated from an average interferogram calculated from 30 single shot interferograms (measurement time: 3 s) at viscose fiber production, from 100 single shot interferograms (measurement time: 10 s) at PEA production and from 50 single shot interferograms (measurement time: 15 s) at melamin resin production. Under perfect illumination conditions the spectrometer has a signal to noise ratio of 20000:1, at a resolution of three wavenumbers and a measurement rate of one per second.

Data Sets Characteristics

Spectra has been drawn at all three application scenarios over a certain period of time and stored onto harddisc. Table 1 provides a summary of the collected data in all applications.
The data has been pre-processed by standard methods conventional in chemometric modelling such as SNV, mean centering, baseline correction etc. [23] [58], as well as cutting tails and other unimportant information contained in the spectra based on expert knowledge.
Additionally, specific outlier detection routines have

Table 1. *Overview of the data sets characteristics recorded at the various application scenarios.*

Data Set Description	Number of instances	Number of Wavel. (=Dimensionality)	Additional Comment
PEA Production OH visk (PL) SZ	25*20=500 for training 96 for testing	3101	20 consecutive measurements per instance
Viscose H_2SO_4 Na_2SO_4 $ZnSO_4$	 16799 16470 16798	 1762 2550 2192	Outlier removal on initial batch set
Melamine Resin Recipe 1 Recipe 2 Recipe 3 Recipe 4 Alltog (New)	 1608 (3*536) 1269 (3*423) 753 (3*251) 1119 (3*373) 6107	 1249 1249 1249 1249 347 (manual sel.)	Absorbance spectra available, see Figure 1; three orthogonal sensors, three repeated measurements

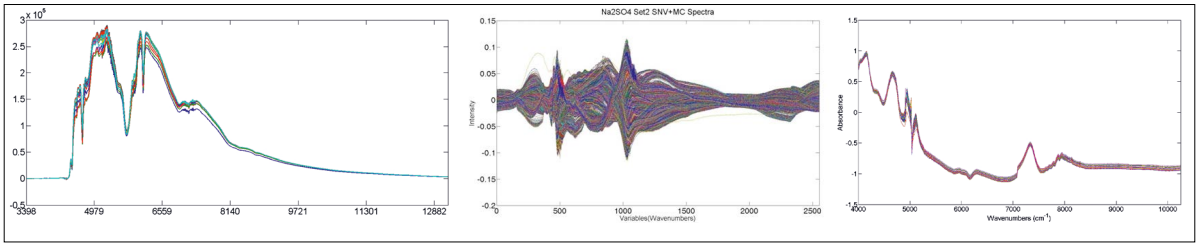


Figure 10. *left: original intensity spectra from PEA production, middle: pre-processed absorbance spectra from viscose production, right: original absorbance spectra from melamine resin production.*

been used to clean up the data for the initial batch off-line calibration phases. One is based on the Mahalanobis distance measure [59], classifiying a sample as outlier when falling out of the tolerance region (statistically motivated prediction interval) [40]. The other is based on statistical approximations in projection methods according to the considerations in [60], leading to χ^2 quantiles with p-values for acceptance regions which can be approximated by Snedecor's F distribution.

For the on-line phase, the uncleaned data has been used in order to mimick the real on-line plug-and-play capability of our methods (a pre-filtering on all available stream data would be an unallowed and invalid fetch-ahead into the future). However, the data has been permanently normalized subject to the wavelengths ranges and characteristics seen so far in the stream (permanently by-updated together with the model). We want to emphasize that a possible incorrect integration of upcoming outlier samples in the stream into the model update may be not that dramatic in case of fuzzy systems, as outlier samples usually trigger a new rule, which would then fire very little for new incoming non-outlier samples and thus contribute very little to the final model output.

Example of spectra with and without pre-processing for the three application scenarios are shown in Figure 10.

Results and discussion

PEA Production

The task is to produce models for the three target parameters viscosity (PL), OH number and acidity (SZ) with highest possible predictive quality. The first evaluation round we have conducted is based on the training data set, on which we performed a 10-fold cross-validation procedure coupled with a best parameter grid search by respecting that consecutive measurements are not in both, training and test folds, in order to gua-

rantee a reliable error which is not too optimistic. The variation range and values of the parameters depend on the concrete calibration modelling method we have used. We compared the most conventional linear methods MLR, PCR and PLS with FLEXFIS combined with FSB, PLS and PCA as filter-based dimension reduction approaches --- for all of these the number of latent variables = the input dimensionality for the calibration models have been varied from 1 to 18. Additionally, we compare with support vector regression (SVR) in two variants (epsilon-SVR and mu-SVR) has having been proposed in various chemical applications to be a fruitful non-linear calibration method [61] [62]. For SVR, the width of the Gaussian kernel as well as the allowed margin C have been varied according to the suggested grid in [HCL06].
The tables below show the results of the cross-validation error for the three targets achieved on the best parameter combination, leading to optimal models in terms of a punished criterion.
This criterion punishes more complex models over lower complex ones such that models with lower complexity receive a higher chance to be selected --- an example of the application of this criterion is provided in Figure 11 below. This is in accordance with the expectation that more complex models more likely overfit on new separate validation data.

Table 2. *cross-validation errors on three target parameters in PEA production when applying various calibration methods onto OH-number (OH-Zahl).*

Results for OH-Zahl (range = 60.74).

Algorithm	Parameters	RMSE	NormRMSE	SD
MLR	–	2.2546	0.0371	1.4684
PCR	$a=5$	2.3044	0.0379	1.5085
PCR	$a=10$	1.9346	0.0319	1.4736
PLSR	$a=6$	**1.7183**	**0.0283**	**0.7143**
ε-SVR	$\varepsilon=0.1$, $C=2^{-1}$, $\gamma=2^{1}$	5.6991	0.0938	10.6569
ν-SVR	$\nu=0.5$, $C=2^{-5}$, $\gamma=2^{3}$	5.6502	0.0930	10.4902
FLEXFIS + PCA	$v=0.9$, $d=7$	3.1730	0.0522	2.6116
FLEXFIS + PLS	$v=0.7$, $d=17$	2.1065	0.0347	1.2061
FLEXFIS + FS	$v=0.9$, $d=3$	**1.4294**	**0.0235**	**0.8478**

Bold values refer to the best methods in terms of minimal RMSE.

Table 3. *cross-validation errors on three target parameters in PEA production when applying various calibration methods onto viscosity (Visk PL/PL).*

Results for Visk PL/PL.

Algorithm	Parameters	RMSE	NormRMSE	SD
MLR	–	15.5784	0.2360	11.9403
PCR	$a=6$	10.1510	0.1538	7.3332
PLSR	$a=3$	10.2541	0.1554	3.6207
ε-SVR	$\varepsilon=0.1$, $C=2^{-3}$, $\gamma=2^{-7}$	13.6281	0.2065	11.3215
ν-SVR	$\nu=0.5$, $C=2^{5}$, $\gamma=2^{-7}$	14.6692	0.2223	6.8904
FLEXFIS + PCA	$v=0.8$, $d=3$	12.0464	0.1825	6.8625
FLEXFIS + PLS	$v=0.9$, $d=3$	11.1374	0.1687	6.8777
FLEXFIS + FS	$v=0.8$, $d=2$	**6.2785**	**0.0951**	**3.5688**

Bold values refer to the best methods in terms of minimal RMSE.

Table 4. *cross-validation errors on three target parameters in PEA production when applying various calibration methods onto acidity (SZ).*

Results for SZ.

Algorithm	Parameters	RMSE	NormRMSE	SD
MLR	–	0.6936	0.0258	0.4496
PCR	$a=4$	1.0564	0.0392	0.3279
PCR	$a=9$	**0.5722**	**0.0213**	**0.3745**
PLSR	$a=4$	0.7373	0.0274	0.2533
PLSR	$a=7$	**0.5241**	**0.0195**	**0.3393**
ε-SVR	$\varepsilon=0.1$, $C=2^{-5}$, $\gamma=2^{-11}$	3.1739	0.1179	5.5629
ν-SVR	$\nu=0.5$, $C=2^{-5}$, $\gamma=2^{-5}$	2.7015	0.1004	5.8859
FLEXFIS + PCA	$v=0.8$, $d=7$	1.3155	0.0489	1.8110
FLEXFIS + PLS	$v=0.8$, $d=7$	1.5978	0.0594	2.4417
FLEXFIS + FS	$v=0.9$, $d=4$	**0.5363**	**0.0199**	**0.3785**

Bold values refer to the best methods in terms of minimal RMSE.

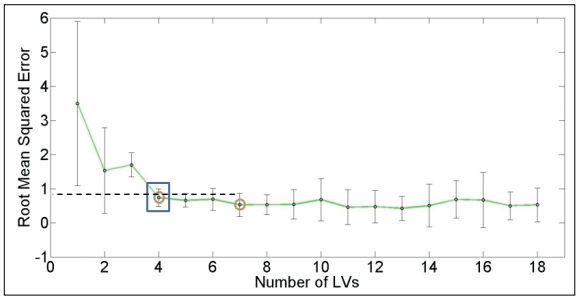


Figure 11. *RMSE CV error over different numbers of latent variables, the intervals representing the standard deviation over the folds; conventional approach based on the minimal CV error would select 7 latent variables (indicated by a circle) whereas our punished criterion selects four latent variables (indicated by a big square) as its induced error is only slightly higher than the minimal one and lies within the standard deviation interval of the minimal one (indicated by a dotted line).*

Obviously, for OH number and viscosity the non-linear models achieved through fuzzy systems (FLEXFIS) could outperform all linear models and also SVR significantly; for SZ the performance of PLS, PCR and FLEXFIS is very similar; SD denotes the standard deviation of the errors over CV folds, NormRMSE can be seen as a kind of percentual error divided by 100 (i.e. 0.01 indicates a 1% error).
Additionally, we studied the performance of the model on a separate validation data set which have been recorded at the same system five months later --- so, this represents really a hard test case how much the system behaviour/dynamics changes in this timeframe resp. how well the actual models still fit into the environment and how reliably they can still quantify the target values. The results for the three targets are shown in the tables below.

Here, the fuzzy systems based non-linear modeling (FLEXFIS) could outperform other non-linear as well as linear models, while the errors became higher than for the cross-validation phase, which is not a big surprise. However, they still stayed in good and for the company acceptable regions for both, OH number and SZ (around 5%), but not for viscosity (errors grow above 20%).

Table 5. *errors on separate validation data for three target parameters in PEA production when applying various calibration methods on the training data with optimal parameter settings achieved through CV --- for OH-number (OH-Zahl, compare with Table 2)*

Validation results for OH (range = 60.74).

Algorithm	RMSE	NormRMSE	Modeling time	Predicting time
MLR	11.3601	0.1870	0.1236	0.0011
PCR5	4.5595	0.0751	0.8014	0.0014
PCR10	**3.4994**	**0.0576**	**0.7604**	0.0061
PLSR	4.1813	0.0688	0.0390	0.0013
ε-SVR	10.9825	0.1808	2.2117	0.0985
ν-SVR	10.5514	0.1737	2.1983	0.0957
FLEXFIS + PCA	**3.6557**	**0.0602**	**0.2280**	0.0013
FLEXFIS + PLS	**3.1520**	**0.0519**	**0.0395**	0.0014
FLEXFIS + FS	**3.5894**	**0.0591**	**0.0336**	0.0014

Bold values refer to the best methods in terms of minimal RMSE.

Table 6. *errors on separate validation data for three target parameters in PEA production when applying various calibration methods on the training data with optimal parameter settings achieved through CV --- for viscosity (VISK PL/PL, compare with Table 3)*

Validation results for Visk PL/PL (range = 66).

Algorithm	RMSE	NormRMSE	Modeling time	Predicting time
MLR	102.2003	1.5485	0.0843	0.00061
PCR	23.9735	0.3632	0.7790	0.00045
PLSR	19.6156	0.2972	0.0136	0.00054
ε-SVR	46.4525	0.7038	2.2429	0.0954
ν-SVR	87.8555	1.3311	2.3352	0.1052
FLEXFIS + PCA	18.5229	0.2806	0.1342	0.00094
FLEXFIS + PLS	19.6156	0.2972	0.0281	0.00096
FLEXFIS + FS	**14.4422**	**0.2188**	**0.0284**	0.00097

Bold values refer to the best methods in terms of minimal RMSE.

Table 7. *errors on separate validation data for three target parameters in PEA production when applying various calibration methods on the training data with optimal parameter settings achieved through CV --- for acidity (SZ, compare with Table 4)*

Results for SZ (range = 26.92).

Algorithm	RMSE	NormRMSE	Modeling time	Predicting time
MLR	11.9130	0.4380	0.0849	0.00054
PCR4	2.8401	0.1044	0.7596	0.00045
PCR9	1.6416	0.0604	0.7702	0.00087
PLSR4	3.0099	0.1107	0.0142	0.00047
PLSR7	**1.5991**	**0.0588**	**0.0205**	0.00058
ε-SVR	7.3331	0.2696	1.7192	0.0791
ν-SVR	7.3331	0.2696	1.7549	0.0788
FLEXFIS + PCA	2.9194	0.1073	0.1409	0.0010
FLEXFIS + PLS	**1.6107**	**0.0592**	**0.0311**	0.0010
FLEXFIS + FS	**1.3344**	**0.0491**	**0.0307**	0.00094

Bold values refer to the best methods in terms of minimal RMSE.

Finally, we tried to improve the waveband selection/extraction used as original input above to FS, PLS and PCA using genetic algorithms including enhanced cross-over operators design.

It finally turned out that on the separate validation set we could achieve the following error reduction [11]:

- For OH number from an RMSE of 3.15 down to around 1 (reduction of around 67%).
- For viscosity from an RMSE of 14.44 down to around 12 (reduction of around 17%).
- For SZ from an RMSE of 1.33 down to around 0.3 (reduction of around 78%).

Hence, we could conclude that a genetic based wavelength selection on original raw spectra really can bring a significant improvement in terms of models' generalization errors.

Viscose Fiber Production

In this application scenario, the main task was to cope with the system dynamics occurring in the spin-bath due to several reasons (see previous section) over a longer timeframe in a sense to keep the accuracy of calibration models initially established during a batch off-line phase at a high level.

Several test runs have been made with various linear and non-linear calibration models, which have been remained static for the whole on-line phase. All these failed showing disastrous deteriorations of model accuracy over time as exemplarily shown in Figure 5 for PLS (but similar occurrences could be observed of SVR, PCR, static fuzzy systems, neural networks etc.). Therefore, we employed the adaptive modelling concepts demonstrated in the methods section. In order to have a fair comparison with the static models, an initial batch model was setup based on a small portion of the whole sample stream. This was achieved by the same procedure (CV with best parameter grid search and model selection) as conducted for the PEA production. Upon the final selected models, we performed a simulation of the real on-line case by loading sample per sample from the data set recorded for evaluation purposes and stored onto hard-disc (its characteristics shown Table 1). For each sample, we performed the following steps (for each sample):

- The target is predicted with the current model (=one-step-ahead prediction), compared with the observed target (as stored in the pre-recorded stream) and the error accumulated and stored (for evaluation purposes).
- Optional: The active learning component is called solely based on the input spectra (as target is not available in the real in-line system) and decides whether the sample should be used for model update. In the classical eChemo paradigm, no active learning is used but each sample always selected for model update.
- If the sample has been selected, the fuzzy model is updated by one of the two variants (which one has to be decided before the modelling process starts):
  - Case 1: by using the dimensionality (number of LVs) and vigilance (most sensitive parameter steering rule evolution versus rule update in FLEXFIS) suggested by the model selection procedures in the off-line phase and conducting a real single-pass, incremental model adaptation.
  - Case 2: performing a re-training on the current window (including the new selected sam-

ple, but deleting the oldest or less informative one) by the same procedure (CV with best parameter grid search and model selection) as described above (case PEA production).

- If the sample has been selected, some help statistics (e.g. the ranges of the wavebands, covariance matrices etc.) are updated.

In case of classical incremental model adaptation, it turned out that the huge model quality deterioration observed for static models could be fully resolved with almost a perfect prediction: Figure 12 shows the observed-versus-predicted plots for the same stream parts and targets as in Figure 5 for static models.

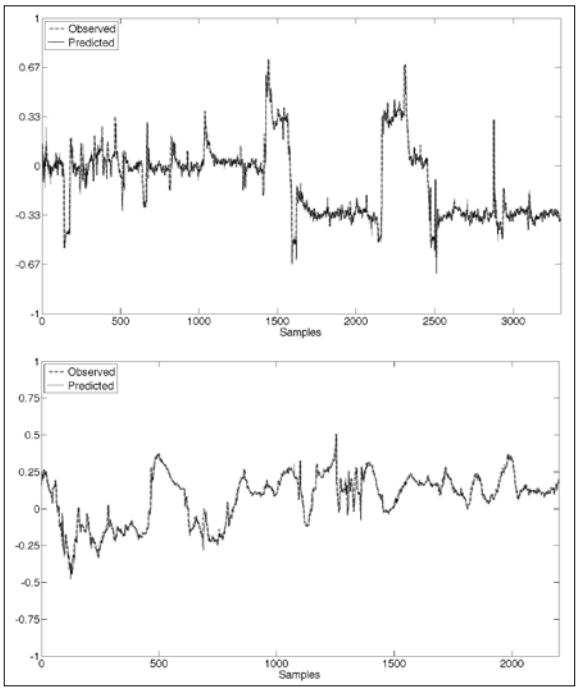


Figure 12. *Observed versus predicted nearly perfectly overlaying each other when updating the calibration model with each sample over a timeframe of around 12-18 days (2000-3000 samples), left: for H_2SO_4, right: for Na_2SO_4 --- compare with the bad results in Figure 5 when using static PLS.*

In summary, very low errors fulfilling the maximal 5% allowance defined by the company could be achieved over the whole streams for each target, as shown in Tables 8-10 (first four rows corresponding to eChemo): around 1.3-1.7% for H_2SO_4 (Table 8), around 1.5-1.8% for Na_2SO_4 (Table 9) and around 1.7-2.5% for ZnSO4 (Table 10), which could be far not achieved with static models (last rows), neither with incremental linear models (incrPCR, last but one rows). An exponential forgetting strategy with lambda = 0.9 (compare with Figure 6) played an essential role to keep the model errors on those low levels. Switching rule pruning on (deletion of redundant rule information [39]) for keeping the model complexity at a low rate and thus ensuring higher computation speed for model updates did not significantly worsen the results.

Table 8. *errors on on-line data streams for H_2SO_4 in viscose fiber production with rule pruning switched on and off ("Pru", "No Pru"), NAvRMSE denotes the percentual deviation which should be maximal 5%.*

On-line validation results for H_2SO_4.

	AvRMSE	NAvRMSE	Corr	Time	AvComp
No Pru-Sel1	0.2495	0.0123	0.9737	0.1433	775.9539
Pru-Sel1	0.2570	0.0127	0.9731	0.2045	599.3145
No Pru-Sel2	0.3447	0.0170	0.9703	0.0031	21.2282
Pru-Sel2	0.3457	0.0171	0.9697	0.0039	12.5538
IncrPCR	2.1124	0.1041	0.4692	0.4077	–
FLEXFIS+FS (static)	5.2633	0.2278	0.1757	0.0001	401

Table 9. *errors on on-line data streams for Na_2SO_4 in viscose fiber production with rule pruning switched on and off ("Pru", "No Pru"), NAvRMSE denotes the percentual deviation which should be maximal 5%.*

On-line validation results for Na_2SO_4.

	AvRMSE	NAvRMSE	Corr	Time	AvComp
No Pru-Sel1	0.6413	0.0154	0.9575	0.0309	260.1313
Pru-Sel1	0.7525	0.0181	0.9074	0.0254	95.4457
No Pru-Sel2	0.8006	0.0192	0.9516	0.0040	29.7643
Pru-Sel2	0.8462	0.0203	0.9150	0.0037	10.7014
IncrPCR	5.6420	0.1115	0.3334	0.1867	–
GLMNet	12.2347	0.2384	0.4134	0.00001	–

Table 10. *errors on on-line data streams for $ZnSO_4$ in viscose fiber production with rule pruning switched on and off ("Pru", "No Pru"), NAvRMSE denotes the percentual deviation which should be maximal 5%.*

On-line validation results for $ZnSO_4$.

	AvRMSE	NAvRMSE	Corr	Time	AvComp
No Pru-Sel1	0.1013	0.0158	0.9796	0.0665	377.3374
Pru-Sel1	0.1119	0.0174	0.9648	0.0845	242.8069
No Pru-Sel2	0.1589	0.0248	0.9720	0.0011	2
Pru-Sel2	0.1589	0.0248	0.9720	0.0012	2
IncrPCR	1.3180	0.2056	0.0545	0.0985	–
StepwiseReg	12.1905	1.6474	0.1466	0.000002	–

However, the problem was that these results have been achieved when using each sample for model update. This was not practicable for the company, as then no titration automat measurement and associated costs can be decreased. In particular, the requirement by the company finally was to measure the targets only 3-4 times a day which leads to an update cycle of each 45th resp. each 60th sample, in order to meet some internal economic goals. Conducting this in a blind equidistant manner without any variability by active sample selection, it lead to huge unacceptable errors of about 30%, 45% resp. 70% percentual errors for the three targets H_2SO_4, Na_2SO_4 and $ZnSO_4$, respectively.

Therefore, we applied our active learning scheme for selecting the most important samples for model updates. We enforced a minimum number of samples that have to be between two selected ones (default 20) in

Table 11. *percentual error rates for the three targets in viscose fiber production when applying equidistant re-training ("Stat"), Euclidean based selection ("Eucl", SoA to be found in [64]) and our active selected scheme ("Dyn") in combination with different sample deletion strategies from the sliding window (random, oldest, lessinfo); the numbers in braces indicate the real average re-train ratio (in case of "Stat" always a fixed value of 60).*

PLSR		Random			Oldest			LessInfo		
Target	Set	Stat	Eucl	Dyn	Stat	Eucl	Dyn	Stat	Eucl	Dyn
H_2SO_4	1	1.32	1.29 (49)	1.30 (52)	1.30	1.29 (50)	1.28 (51)	1.27	1.21 (48)	1.21 (50)
H_2SO_4	2	1.13	1.15 (58)	1.10 (52)	1.09	1.06 (54)	1.04 (54)	1.03	0.98 (50)	0.92 (49)
Na_2SO_4	1	1.71	1.70 (50)	1.71 (52)	1.42	1.39 (54)	1.39 (53)	1.34	1.36 (58)	1.35 (54)
Na_2SO_4	2	1.66	1.66 (52)	1.66 (52)	1.48	1.53 (49)	1.50 (53)	1.60	1.54 (52)	1.49 (51)
$ZnSO_4$	1	**6.03**	**5.98** (51)	**5.90** (54)	**6.10**	**5.05** (53)	**6.01** (52)	**6.13**	**5.97** (51)	**5.99** (51)
$ZnSO_4$	2	**5.93**	**5.91** (53)	**5.85** (55)	**6.03**	**5.92** (53)	**5.92** (53)	**5.91**	**5.91** (52)	**5.89** (54)
FLEXFIS		Random			Oldest			LessInfo		
Target	Set	Stat	Eucl	Dyn	Stat	Eucl	Dyn	Stat	Eucl	Dyn
H_2SO_4	1	1.37	1.34 (53)	1.30 (53)	1.29	1.20 (50)	1.25 (51)	1.26	1.21 (47)	1.19 (49)
H_2SO_4	2	1.19	1.17 (50)	1.06 (51)	1.07	1.08 (54)	1.00 (53)	1.05	1.00 (49)	0.99 (50)
Na_2SO_4	1	1.77	1.73 (54)	1.66 (53)	1.74	1.36 (49)	1.37 (53)	1.66	1.38 (50)	1.37 (51)
Na_2SO_4	2	1.66	1.65 (52)	1.59 (50)	3.01	2.22 (54)	2.00 (56)	1.91	1.64 (53)	1.52 (52)
$ZnSO_4$	1	**6.97**	**6.83** (53)	**6.28** (52)	**5.96**	**5.95** (51)	**5.94** (53)	**5.20**	**5.25** (53)	**5.03** (53)
$ZnSO_4$	2	**6.10**	**5.85** (52)	**5.87** (54)	**5.89**	**5.88** (51)	**5.77** (53)	**6.00**	**5.77** (53)	**5.77** (51)

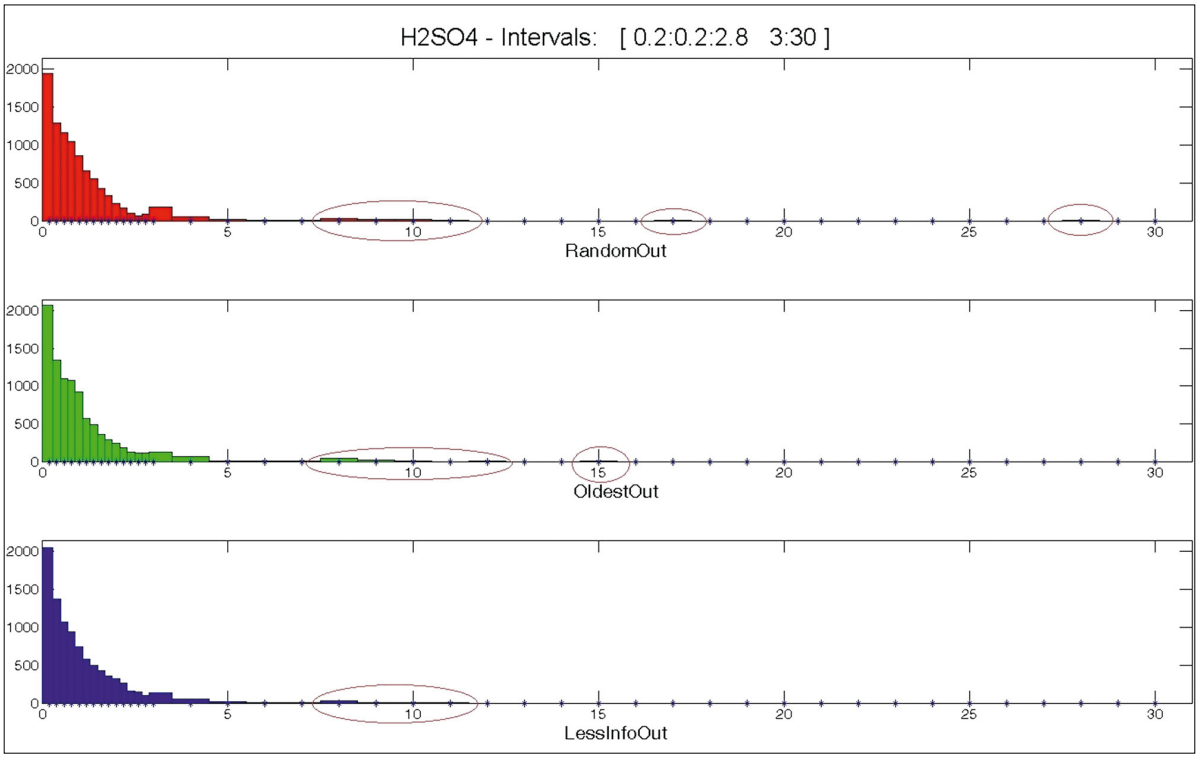


Figure 13. *error histograms for different variants of sample deletion from the window.*

order to be able to meet the requirements regarding the allowed re-train ratio. Indeed, it can happen that 3,4,5 samples are then selected in a row with 20 samples in-between, but on the other hand in lower dynamic phases no sample may be selected at all for a longer time-frame of a few hundreds or even 1000 samples. Finally, it turned out that we could keep the desired re-train ratio of about 45-50 samples in average.

We have been not successful with the classical incremental eChemo paradigm, but could achieve good errors when applying the sliding-window based re-training approach, in a similar manner as when updating the models with each single sample, see Table 11: only for $ZnSO_4$ a slight rise above the desired 5% errors could be observed.

This indicated a change in the importance of the wavelengths (a shift in the spectroscopic information content), as the most essential difference is that waveband selection is carried out newly in the sliding window approach, whereas it always remains fix (obtained from the intial batch phase) for the classical eChemo approach.

The error histograms in Figure 13 emphasize the use of more intelligence/efforts in the sample deletion strate-

Table 12. *results from various calibration methods with ("ens") and without ensembles ("orig"), also compared to a pure averaging of the repeated measurements ("aver"), bold values indicate significantly best methods.*

Method	Formulation #1	Formulation #2	Formulation #3	Formulation #4
PLSR (orig)	0.0359 ± 0.0195	0.0439 ± 0.0330	0.0284 ± 0.0090	0.0399 ± 0.0359
PLSR (aver)	0.0358 ± 0.0206	0.0403 ± 0.0284	0.0270 ± 0.0089	0.0397 ± 0.0363
PLSR (ens)	0.0340 ± 0.0263	0.0426 ± 0.0335	0.0264 ± 0.0092	0.0379 ± 0.0355
GLMnet (orig)	0.0326 ± 0.0107	0.0638 ± 0.0692	0.0309 ± 0.0069	0.0476 ± 0.0330
GLMnet (aver)	0.0326 ± 0.0106	0.0679 ± 0.0512	0.0324 ± 0.0084	0.0474 ± 0.0382
GLMnet (ens)	0.0321 ± 0.0110	0.0621 ± 0.0927	0.0310 ± 0.0092	0.0437 ± 0.0349
FLEXFIS + PLS (orig)	0.0297 ± 0.0118	0.0370 ± 0.0277	0.0203 ± 0.0065	0.0428 ± 0.0317
FLEXFIS + PLS (aver)	0.0268 ± 0.0117	0.0343 ± 0.0274	0.0166 ± 0.0047	0.0411 ± 0.0334
FLEXFIS + PLS (ens)	**0.0242 ± 0.0083**	**0.0291 ± 0.0212**	**0.0104 ± 0.0089**	**0.0281 ± 0.0250**
FLEXFIS + FS (orig)	0.1371 ± 0.0179	0.0733 ± 0.0356	0.0786 ± 0.0110	0.0970 ± 0.0414
FLEXFIS + FS (aver)	0.0573 ± 0.0167	0.0620 ± 0.0215	0.0638 ± 0.0143	0.0787 ± 0.0312
FLEXFIS + FS (ens)	0.0482 ± 0.0130	0.0550 ± 0.0215	0.0458 ± 0.0091	0.0695 ± 0.0317
FLEXFIS + FSB (orig)	0.0540 ± 0.0146	0.0626 ± 0.0456	0.0477 ± 0.0082	0.0951 ± 0.0218
FLEXFIS + FSB (aver)	0.0413 ± 0.0171	0.0610 ± 0.0268	0.0464 ± 0.0089	0.0911 ± 0.0240
FLEXFIS + FSB (ens)	0.0395 ± 0.0111	0.0583 ± 0.0398	0.0434 ± 0.0082	0.0601 ± 0.0387

gy, as the larger errors above 10 can be fully omitted with LessInfo approach.

Melamine Resin Production

In this application scenario, we were facing two main challenges:

- Establishing enhanced chemometric calibration models from FT-NIR spectra using orthogonal sensor information and the exploration of repeated measures, to reduce the noise effect. This also includes the search for eventual non-linearities contained in the process to outperform previously applied standard PLS models.
- Performing model adaptation cycles on demand in order to balance out drifts in the process, which may arise due to the degradation of the lamp intensity and changes in the composition of the educt.

Table 12 shows the achieved errors on four different formulations of melamine resin contained within one reactor when applying base and ensemble modelling schemes.

Obviously, fuzzy systems in combination with PLS (FLEXFIS+PLS as a sort of non-linear version of PLS) are able to produce lower errors than the other methods PLS, GLMnet and other dimension reduction schemes including forward selection with bands (FSB). All methods seem to benefit from the ensembling scheme over models extracted from the three repeated measurements. In order to statistically underline these findings from first glance, we performed a Mann-Whitney-Wilcoxon test [65] whose results are shown in Table 13 when comparing ensemble scheme versus sample averaging versus original models and in Table 14 for comparing FLEXFIS+PLS with other methods when applying the ensembling scheme for all.

Table 13. *preference analysis of model ensembles versus sample averaging versus non-ensembles, entries below 0.05 denotes a statistically significant preference (W=first entry worse, S = same, B=first entry better).*

Altogether	Orig Vs Aver	Orig Vs Ens	Aver Vs Ens
PLSR	0.2706 (S)	$3 \cdot 10^{-8}$ (W)	$1 \cdot 10^{-8}$ (W)
GLMNet	0.0556 (S)	$2 \cdot 10^{-4}$ (W)	0.0200 (W)
FLEXFIS + PLS	0.0978 (S)	0.3700 (S)	0.0624 (S)

Table 14. *preference analysis of FLEXFIS+PLS versus all the other methods when applying ensembling scheme, entries below 0.05 denote a statistically significant preference of FLEXFIS+PLS over the method mentioned in the corresponding column heading.*

Ens	PLSR	GLMNet	FLEXFIS + FS	FLEXFIS + FSB
Formulation #1	0.4849 (S)	0.1923 (S)	0.0023 (B)	0.0036 (B)
Formulation #2	0.0036 (B)	0.0320 (B)	0.0070 (B)	0.0106 (B)
Formulation #3	3×10^{-4} (B)	0.2137 (S)	2×10^{-4} (B)	0.0011 (B)
Formulation #4	5×10^{-4} (B)	0.0188 (B)	7×10^{-4} (B)	0.0011 (B)
Altogether	4×10^{-7} (B)	0.0030 (B)	5×10^{-10} (B)	2×10^{-9} (B)

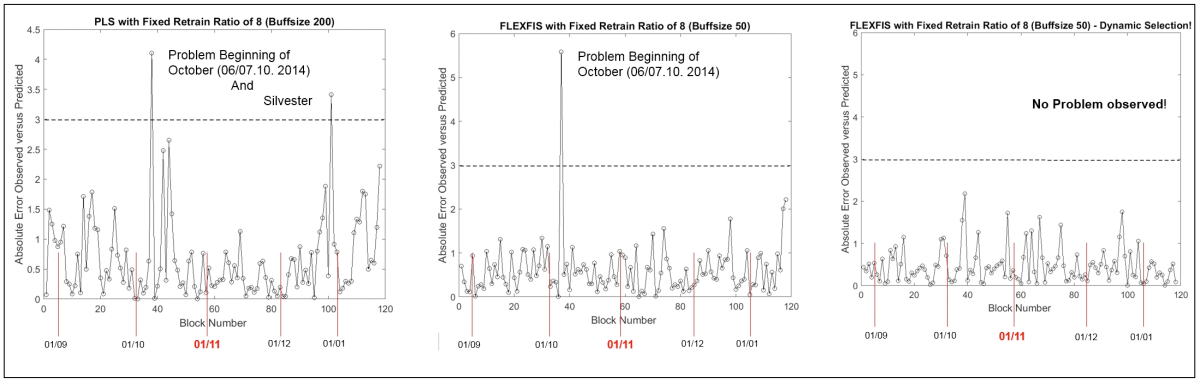


Figure 14. *residuals signals over a timeframe of 5 months when using PLS with a fixed re-train ratio of 8 (left), FLEXFIS with a fixed re-train ratio of 8 (middle) and FLEXFIS with dynamically selected samples for model update (right).*

The entries in both tables represent p-values which when below 0.05 indicate statistical significance with a 5% chance to be wrong. A further note goes to the integration of orthogonal sensors, which could not improve the accuracy of the calibration models. The reason was that the three variables pH value, temperature and pressure were highly correlated with some wavebands in the spectra. Thus, by adding them into the regression modeling phase, no new information content for the target could be gained. Thus, the mapping between spectra and the targets could not be improved.

In order to verify how the off-line built models consecutively behave over time, we applied the finally extracted models (with optimal parameter setting) to a stream recorded over a timeframe of about five months (August 2014 – January 2015). This resulted in drifting residuals after some weeks and undesired peaks in the residual signals with errors significantly above 3, an allowed upper limit during on-line production process. In order to resolve this unpleasant situation, we applied the same concept as for viscose production, namely a sliding window based re-training approach based on actively selected samples. The obtained residuals with the dynamic methods (PLS, FLEXFIS with a fixed retrain ratio and FLEXFIS with a dynamic retrain ratio) over this time period averaged over each data block are shown in Figure 14.

Obviously, the active sample selection can reduce all errors to below the required threshold of 3 as indicated by the horizontal lines in the plots in Figure 14. The re-train ratio of 8 is within the expectations, as usually one batch contains 6 to 8 samples and at most 1 update within one batch should be conducted.

Conclusions

In this paper, we provided a summary of our achievements regarding non-linear calibration modelling from FT-NIR Spectra on the one hand, and their self-adaptivity on the other hand, in order to address dynamic chemical processes which are going beyond having pure linear and static properties and characteristics. Most of the achievements are new developments in terms of combining chemometrics, soft computing, machine learning and genetic algorithms and thus should enrich the scientific community as well as industrial production sites with on-line, dynamic modelling and prediction challenges. The results on three completely different application scenarios show the applicability and feasibility of our new methods, which are statistically compared with state-of-the-art calibration modelling methods (such as MLR, PCR or PLS). In static problems, an error reduction of 30-40% could be achieved in most of the cases, while in dynamic problems it turned out that the usage of classical methods could not deliver any meaningful results at all, whereas our self-adaptive methods could achieve reasonable errors within the upper allowed limits defined by the companies in large parts (with some exceptions). Active learning played an essential role to keep measurement efforts at a minimum level in order to decrease the costs for measurements as much as possible.

Acknowledgements

This main work was funded by the Austrian research funding association (FFG) under the scope of the COMET programme within the research project “Industrial Methods for Process Analytical Chemistry - From Measurement Technologies to Information Systems (imPACts)” (contract # 843546). This publication reflects only the authors' views.

References

[1] M. Blanco, I. Villarroya, NIR spectroscopy: a rapid-response analytical tool. TrAC Trends Anal. Chem. 21, 240–250 (2002).

[2] M. Haenlein and A.M. Kaplan, A beginner's guide to partial least squares (PLS) analysis, Understanding Statistics, vol. 3 (4), pp. 283-297, 2004.

[3] P. Larkin, Infrared and Raman Spectroscopy; Principles and Spectral Interpretation, Elsevier Science Publishing, San Diego, U.S.A., 2011.

[4] T. Hastie, R. Tibshirani and J. Friedman, The Elements of Statistical Learning: Data Mining, Inference and Prediction - Second Edition, Springer, New York Berlin Heidelberg, 2009.

[5] E. Lughofer and S. Kindermann, SparseFIS: Data-driven learning of Fuzzy Systems with Sparsity Constraints, IEEE Transactions on Fuzzy Systems, vol. 18 (2), pp. 396-411, 2010.

[6] F. Bauer and M. Lukas, Comparing parameter choice methods for regularization of ill-posed problems, Mathematics and Computers in Simulation, vol. 81 (9), pp. 1795-1841, 2011.

[7] K. Varmuza and P. Filzmoser, Introduction to Multivariate Statistical Analysis in Chemometrics, CRC Press, Boca Raton, 2009.

[8] J.N. Miller and J.C. Miller, Statistics and Chemometrics for Analytical Chemistry, Prentice Hall, Essex, England, 2009.

[9] H. Chen, T. Pan, J. Chen and Q. Lu, Waveband Selection for NIR Spectroscopy Analysis of Soil Organic Matter based on SG Smoothing and MWPLS method, Chemometric and Intelligent Laboratory Systems, vol. 107 (1), (2011) 139-146.

[10] C. Cernuda, E. Lughofer, P. Hintenaus, W. Märzinger, T. Reischer and M. Pawlicek and J. Kasberger, Hybrid Adaptive Calibration Methods and Ensemble Strategy for Prediction of Cloud Point in Melamine Resin Production, Chemometrics and Intelligent Laboratory Systems, vol. 126, pp. 60-75, 2013.

[11] C. Cernuda and E. Lughofer and P. Hintenaus and W. Märzinger, Enhanced Waveband Selection in NIR Spectra using Enhanced Genetic Operators, Journal of Chemometrics, vol. 28 (3), pp. 123-136, 2014.

[12] B. Korte and J. Vygen, Combinatorial Optimization, Springer, 2005.

[13] F.E. Harrel, Regression Modeling Strategies, Springer, New York, USA, 2001.

[14] J. H. Holland, Adaptation in Natural and Artificial Systems, Univ. Michigan Press, Michigan, U.S.A., 1975.

[15] D. Goldberg, Genetic Algorithms in Search, Optimization, and Machine Learning, Addison-Wesley, Boston, U.S.A., 1989.

[16] M. Stone, Cross-validatory choice and assessment of statistical predictions, Journal of the Royal Statistical Society, vol. 36 (1), pp. 111-147, 1974.

[17] I. Guyon and A. Elisseeff, An Introduction to Variable and Feature Selection, Journal of Machine Learning Research, vol. 3, pp. 1157--1182, 2003.

[18] H. C. Goicoechea, A. C. Olivieri, A new family of genetic algorithms for wavelength interval selection in multivariate analytical spectroscopy, Journal of Chemometrics, vol. 17 (6) (2003) 338–345.

[19] R. Leardi, R. Boggia, M. Terrile, Genetic algorithms as a strategy for feature selection, Journal of Chemometrics, 6 (5) (1992) 267–281.

[20] C. Reyns, S. d. Souza, R. Sabatier, G. Figures, B. Vidal, Selection of discriminant wavelength intervals in NIR spectrometry with genetic algorithms, Journal of Chemometrics, 20 (3-4) (2006) 136–145.

[21] Q. Fei, M. Li, B. Wang, Y. Huan, G. Feng, Y. Ren, Analysis of cefalexin with NIR spectrometry coupled to artificial neural networks with modified genetic algorithm for wavelength selection, Chemometrics and Intelligent Laboratory Systems, 97 (2) (2009) 127 – 131.

[22] C. Cernuda, E. Lughofer, G. Mayr, T. Röder, P. Hintenaus, W. Märzinger and J. Kasberger, Incremental and Decremental Active Learning for Optimized Self-Adaptive Calibration in Viscose Production, Chemometrics and Intelligent Laboratory Systems, vol. 138, pp. 14-29, 2014.

[23] R.G. Brereton, Chemometrics: Data Analysis for the Laboratory and Chemical Plant, John Wiley & Sons, Hoboken, New Jersey, 2003

[24] M. Otto, Chemometrics, 2nd Edition, John Wiley & Sons, Hoboken, New Jersey, 2007.

[25] E. Lughofer, FLEXFIS: A Robust Incremental Learning Approach for Evolving TS Fuzzy Models, IEEE Transactions on Fuzzy Systems, vol. 16 (6), pp. 1393-1410, 2008.

[26] T. Fearn and A.M.C. Davies, Locally-biased regression, Journal of Near Infrared Spectroscopy, vol. 11 (6), pp. 467-478, 2003.

[27] J.L. Castro and M. Delgado, Fuzzy Systems with defuzzification are universal approximators, IEEE Transactions on Systems, Man and Cybernetics, part B: Cybernetics, vol. 26 (1), pp. 149-152, 1996.

[28] E. Lughofer, On-line Assurance of Interpretability Criteria in Evolving Fuzzy Systems --- Achievements, New Concepts and Open Issues, Information Sciences, vol. 251, pp. 22-46, 2013.

[29] T. Takagi and M. Sugeno, Fuzzy Identification of Systems and its Applications to Modeling and Control, IEEE Transactions on Systems, Man and Cybernetics, vol. 15 (1), pp. 116-132, 1985.

[30] E. Lughofer, Extensions of Vector Quantization for Incremental Clustering, Pattern Recognition, vol. 41 (3), pp. 995-1011, 2008.

[31] E. Lughofer and M. Sayed-Mouchaweh, Autonomous Data Stream Clustering implementing Incremental Split-and-Merge Techniques --- Towards a Plug-and-Play Approach, Information Sciences, vol. 204, pp. 54-79, 2015.

[32] E. Lughofer, C. Cernuda, S. Kindermann and M. Pratama, Generalized Smart Evolving Fuzzy Systems, Evolving Systems, on-line and in press, doi: 10.1007/s12530-015-9132-6, 2015.

[33] E.P. Klement, R. Mesiar and E. Pap, Triangular Norms, Kluwer Academic Publishers, Dordrecht Norwell New York London, 2000.

[34] P. Angelov, E. Lughofer and X. Zhou, Evolving Fuzzy Classifiers using Different Model Architectures, Fuzzy Sets and Systems, vol. 159 (23), ppp. 3160-3182, 2008.

[35] J. Einax, Chemometrics in Environmental Chemistry – Applications (2nd edition), Springer, Berlin Heidelberg, 2013.

[36] K. Götze, Chemiefasern nach dem Viscoseverfahren, Springer Verlag, Berlin Heidelberg, New York, 1967.

[37] C. Cernuda, E. Lughofer, L. Suppan, T. Röder, R. Schmuck, P. Hintenaus, W. Märzinger and J. Kasberger, Evolving Chemometric Models for Predicting Dynamic Process Parameters in Viscose Production, Analytica Chimica Acta, vol. 725, pp. 22-38, 2012.

[38] W.C. Abraham and A. Robins, Memory Retention – the Synaptic Stability versus Plasticity Dilemma, Trends in Neurosciences, vol. 28 (2), pp. 73-78, 2005.

[39] E. Lughofer, J.-L. Bouchot and A. Shaker, On-line Elimination of Local Redundancies in Evolving Fuzzy Systems, Evolving Systems, vol. 2 (3), pp. 165-187, 2011.

[40] K. Krishnamoorthy and T. Mathew, Statistical Tolerance Regions: Theory, Applications, and Computation, John Wiley & Sons, Hoboken, New Jersey, 2009.

[41] X. Wu, K.-H. Bellgardt, On-line fault detection of flow-injection analysis systems based on recursive next term parameter estimation, Analytica Chimica Acta, vol. 313 (3) (1995) 161–176.

[42] O. Haavisto, H. Hyotyniemi, Recursive multimodel partial least squares estimation of mineral flotation slurry contents using optical reflectance spectra, Analytica Chimica Acta, vol. 642 (2009) 102–109.

[43] R. E. Kalman, A New Approach to Linear Filtering and Prediction Problems, Journal of Basic Engineering, vol. 82 (35), 1960.

[44] L. Ljung, System Identification: Theory for the User, Prentice Hall PTR, Prentice Hall Inc, Upper Saddle River, New Jersey, 1999.

[45] K.J. Aström and B. Wittenmark, Adaptive Control - Second Edition, Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 1994.

[46] E. Lughofer, Single-Pass Active Learning with Conflict and Ignorance, Evolving Systems, vol. 3 (4), pp. 251-271, 2012.

[47] P. Donmez and J.G. Carbonell, From Active to Proactive Learning Methods, editors: J. Koronacki, Z.W. Ras, S.T. Wierzchon and J. Kacprzyk, Advances in Machine Learning I, Springer, Berlin Heidelberg, (2010), 97-120.

[48] I.T. Jolliffe, Principal Component Analysis, Springer Verlag, Berlin Heidelberg New York, 2002.

[49] L.H. Chiang, E.L. Russell and R.D. Braatz, Fault Detection and Diagnosis in Industrial Systems, Springer, London Berlin Heidelberg, 2001.

[50] P.F. Odgaard, B. Lin and S.B. Jorgensen, Observer and Data-Driven-Model-Based Fault Detection in Power Plant Coal Mills, IEEE Transactions on Energy Conversion, vol. 23 (3), (2008), 659-668.

[51] C. Pasquini, Near Infrared Spectroscopy: fundamentals, practical aspects and analytical applications, Journal of Brazilian Chemical Society, vol. 14 (2), (2003).

[52] J. Lin and C.W. Brown, Spectroscopic Measurement of NaCl and Seawater Salinity in the Near-IR Region of 6801230 nm, Applied Spectroscopy, vol. 47 (2), (1993), 239-241.

[53] P. Hintenaus, G. Kvas and W. Märzinger, An Infrared Spectrometer for Process Monitoring I, Spectroscopy, Proceedings of the IEEE Industria Electronics Society (IECON), Taipei, Taiwan, 2007.

[54] P. Hintenaus, Engineering Embedded Systems, Springer, Berlin Heidelberg, (2015).

[55] L. Breiman, Bagging Predictors, Machine Learning, vol. 24 (2), (1996) 123-140.

[56] P. Brazdil, C. Giraud-Carrier, C. Soares and R. Vilalta, Metalearning, Springer, Berlin Heidelberg, 2009.

[57] Z. Bleier, C. Brouillette, R. Carangelo, A monolithic interferometer for FT-IR spectroscopy, Spectroscopy, vol.14 (10), (1999), 46–49.

[58] A. Rinnan, F. de Berg and S.B. Engelsen, Review of the most common pre-processing techniques for near-infrared spectra, Trends in Analytical Chemistry, vol. 28 (10), 2009.

[59] R. De Maesschalck, D. Jouan-Rimbaud and D.L. Massart, The Mahalanobis distance, Chemometrics and Intelligent Laboratory Systems, vol. 50 (2000), 1-18.

[60] A. Pomerantsev, Acceptance areas for multivariate classification derived by projection methods, Journal of Chemometrics, vol. 22 (2008) 601–609.

[61] N. Hernandez and I. Talavera and R.J. Biscay and D. Porroa and M.M.C. Ferreira, Support vector regression for functional data in multivariate calibration problems, Analytica Chimica Acta, vol. 642 (1-2), (2009) 110-116, 2009.

[62] I.A. Naguib, E.A. Abdelaleem, M.E. Draz and H.E. Zaazaa, Linear support vector regression and partial least squares chemometric models for determination of Hydrochlorothiazide and Benazepril hydrochloride in presence of related impurities: a comparative study, Spectrochim Acta A Mol Biomol Spectrosc, vol. 130, (2014) 350-356.

[63] C.-W. Hsu, C.-C. Chang and C.-J. Lin, A Practical Guide to Support Vector Classification, 2006, Technical Report at the Department of Computer Science and Information Engineering, National Taiwan University, 2006.

[64] F. Douak, F. Melgani, N. Alajlan, E. Pasolli, Y. Bazi, N. Benoudjit, Active learning for spectroscopic data regression, Journal of Chemometrics, vol. 26 (2012) 374–383.

[65] H. Mann and D. Whitney, On a test of whether one of two random variables is stochastically larger than the other, Annals of mathematical Statistics, vol. 18, (1947) 50-60.

Application-oriented Standard-free Methods for Calibration Transfer

Birgit Malli*[1], Thomas Natschläger[1], Marcin Pawliczek[2], Thomas Reischer[3], Wolfgang Kantner[3], Markus Brandstetter[2], Wolfgang Märzinger[4] and Jakub Kowalski[4]

[1] Software Competence Center Hagenberg, 4232 Hagenberg, Austria[1]
[2] RECENDT – Research Center for Non-Destructive Testing GmbH, Altenberger Straße 69, 4040 Linz, Austria
[3] Metadynea Austria GmbH, 3500 Krems, Austria
[4] i-RED Infrarot Systeme GmbH, 4020 Linz, Austria
* Birgit.Malli@scch.at

Abstract

Within the past few decades, the combination of spectroscopic measurement techniques and multivariate calibration methods has become increasingly prominent for the extraction of (bio-) chemical information in various application fields. While the obtained results via such an approach are very satisfying in general, the process of data collection, model calibration and model optimization is a rather time-consuming and cost-intensive one. As such, one certainly aims to prevent the need for a repetition of these steps. In certain cases, though, changes in the environmental conditions, the measurement setup or the measured substance itself may occur and render the calibrated model invalid. In such a situation, either a new model needs to be developed or mathematical operations, referred to as (calibration) transfer methods, can be performed to transfer knowledge from the original to the new setting. Within this contribution, we introduce and discuss a row of application-oriented transfer approaches that are both easy to comprehend and implement. The proposed methods do not assume the availability of transfer standards (i.e. a set of samples measured under the old and new condition) and require the measurement of only a few reference values in the new setting. We evaluate the introduced transfer methods on data from melamine formaldehyde (MF) resin production and show that these techniques can achieve considerable improvement of predictions for data from three forms of changes in the given application context.

Keywords: *Calibration Transfer, Instrument Standardization, Transfer Learning, NIR Spectroscopy, Chemometrics, Melamine Resin*

Introduction

It has been proven within the past few decades that a combination of spectroscopic measurements and multivariate calibration techniques is able to extract (bio-) chemical information in various applications fields. In this work we focus on near-infrared (NIR) spectral data from melamine formaldehyde (MF) resin production, where the objective of this in-line installation comprises in the monitoring of the corresponding polymerization progress and the estimation of the end point of the reaction.

NIR spectroscopy has commonly been applied for on-line / in-line monitoring of (bio-) chemical processes (Du, et al., 2011) quality control of pharmaceutical, petroleum, agriculture and other products, environmental analysis, medical diagnostics and academic research (Kramer, et al., 2008). Some of the benefits of this measurement technique are its non-destructive nature, the possibility for a rapid and continuous process monitoring, the fact that no or only little sample preparation is required in general and in the considered application also the relatively small effort for the implementation within the existing control systems (Kramer, et al., 2008; Fan, et al., 2008; Peng, et al., 2011; Pawliczek, et al., 2015; Du, et al., 2011).

In order to extract chemical information from the measured NIR spectra, the multivariate calibration tech-

nique partial least squares (PLS) (Wold, 1966; Wold, et al., 2001) has been utilized. PLS is a projection method commonly used in chemometrics to handle the mathematically demanding attributes of spectral data.

In general, as well as the studied melamine formaldehyde (MF) resin application, the process of data collection, model calibration and model optimization is rather time-consuming and costly. Therefore, one naturally hopes that the obtained model remains accurate for a very long time. Unfortunately, changes in the environmental conditions, the measurement setup, e.g. due to maintenance operations, or the measured substance itself do usually affect the measured spectra and make the calibration model invalid (Feudale, et al., 2002; Wise & Roginski, 2015; Kramer, et al., 2008; Du, et al., 2011). In such a case, there are two possibilities to deal with this issue: i) calibration of a completely new model and ii) the application of mathematical methods aiming at the transfer of knowledge collected during the calibration process.

Obviously, option ii) causes less costs and effort, if reasonably accurate predictions for the new data setting can be achieved. This is the reason why corresponding methods have been widely studied in literature. They are referred to as calibration transfer or instrument standardization in chemometrics (De Noord, 1994; Bouveresse & Massart, 1996; Feudale, et al., 2002) and are known as transfer learning, domain adaptation or multi-task learning in the machine learning community (Crammer, et al., 2008; Pan, et al., 2008; Mansour, et al., 2008).

In this contribution, we will evaluate a row of transfer approaches which are particularly application-oriented in the sense that they are not based on transfer standards (i.e. a set of samples measured under both the old and new measurement / environmental / sample condition), and do not require any or only a very small amount of reference measurements in the new setting.

Materials and Methods

General Information

Notation and Assumptions

In the following we denote knowledge (data, model,..) belonging to the primary measurement / environmental / sample condition as source (S) knowledge and information coming from the new setting as target (T) information.

The ${}^{(S)}n\times p$ and ${}^{(T)}n\times p$ matrices ${}^{(S)}\boldsymbol{X}$ and ${}^{(T)}\boldsymbol{X}$ contain the spectral measurements (single spectra $\boldsymbol{xi}$ form the rows and columns correspond to wavelengths / wavenumbers). Reference measurements (labels), denoted as ${}^{(S)}\boldsymbol{y}$ and ${}^{(T)}\boldsymbol{y}$, represent column vectors and predictions are generally marked by a hat symbol. If a set of spectra is unlabelled, i.e. no reference value is available for them, then these are written as ${}^{(\cdot)}_{unlab}\boldsymbol{X}$, while labelled data are denoted as ${}^{(\cdot)}_{lab}\boldsymbol{X}$.

Following the setting of the application context, we assume the following data to be available (see Figure 1): A set of labelled S spectra, ${}^{(S)}_{lab}\boldsymbol{X}$, which, together with reference measurements, ${}^{(S)}_{lab}\boldsymbol{y}$, form the basis of the source calibration model, a few labelled target spectra, ${}^{(T)}_{lab}\boldsymbol{X}$, with corresponding ${}^{(T)}_{lab}\boldsymbol{y}$ and a set of unlabelled spectra, ${}^{(T)}_{unlab}\boldsymbol{X}$, in the target domain, where for the sample sizes the relations ${}^{(S)}_{lab}n \gg {}^{(T)}_{lab}n$ and ${}^{(T)}_{unlab}n > {}^{(T)}_{lab}n$ hold.

${}^{(T)}_{unlab}\boldsymbol{X}$ is assumed to be available based on our experience of data collection in practice: Usually, reference values are time-consuming and costly to measure, while spectral data can be obtained very easily or are even recorded automatically. The latter is also the case in the in-line NIR installation considered here and hence it is desirable to make use of these additional spectra, if possible.

In addition to the notation above, we write ${}^{(S)}\beta\in R^n$ and ${}^{(S)}\beta_0\in R$ for the vector of regression coefficients and the intercept term of the source calibration model, respectively and use the notation .∗ for element-wise multiplication.

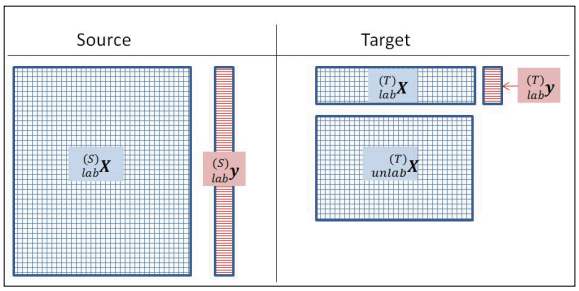


Figure 1. *Picture of the data setting of the investigated application.*

Transfer Standards

Many common chemometric transfer methods are based on the evaluation of differences between the source and target condition (Feudale, et al., 2002). For this, a set of samples, called transfer standards, is assumed to be measured in both the source and target setting. The corresponding spectra are compared in order to learn about the difference between source and target. This information is then utilized to develop appropriate correcting measures.

There are several ways to characterize (Bouveresse & Massart, 1996) and choose (Wang, et al., 1991; Wise & Roginski, 2015; Kennard & Stone, 1969; Wang, et al., 1992) good transfer standards. All of them, though, have naturally in common that the same set of samples can be measured in the source and target domain. Un-

fortunately, this is not the case for the considered industrial application.
Thus, we need to come up with approaches that do not require the existence of transfer standards. Several such methods are described in the following.

Methods

Initial Robust Calibration Model

In a perfect world, we would ask for a source calibration model that is robust enough to deal with changing conditions. For such a model to be built, one can for instance try to incorporate all forms of expected future variation inside the calibration process (De Noord, 1994). This, however, is not feasible in practice as it is generally impossible to foresee all sources and directions of future change. Another approach is based on the idea to build a model only on spectral regions that are not affected by expected changes (Bouveresse & Massart, 1996). This, though, reduces to the same problem that not all forms of change are known beforehand plus the issue that chemically significant areas of the spectra may be removed. Alternatively, ensemble methods as the one proposed in (Ni, et al., 2011), may be pursued. These approaches do certainly represent a valid alternative and are intended to be investigated in future work.
In addition, spectral pre-processing methods, mentioned below, may be able to make the initial calibration model more robust. The investigation of a selection of advanced pre-processing methods shall be addressed in the future.

Model updating

If the direct application of the source model is not accurate enough, a natural extension is the recalibration of the model on an augmented calibration set consisting of the original source data and the labelled target data. This technique, here denoted as T+S, is very common in both literature and practice. In cases where the number of labelled source data exceeds the number of labelled target data considerably, though, the new variance incorporated from the target setting may be "washed out" by the source information, resulting in an unsatisfying model. In order to deal with such a setting, a stronger weighting of the target data seems sensible. In this contribution we investigate the approach to weigh[2] the labelled target data five[3] times before they are added to the source data. This method will be referred to as 5T+S in the following.
Additionally to these simple model updating approaches, automatic model updating methods have been studied (see e.g. (Wise & Roginski, 2015) and reference therein). However, as these methods are currently considered suitable for laboratory settings only (Wise & Roginski, 2015) their applicability in an industrial context shall be investigated in future work.
Obviously, for model updating approaches to be applied, a new model has to be trained and optimized.

Prediction Correction

Another commonly used transfer approach, referred to as LIN hereafter, performs a univariate slope and bias correction of predictions from the original calibration model. In order to learn appropriate values for the slope, k, and the bias, d, target labels, ${}^{(T)}_{lab}\boldsymbol{y}$, and their predictions from the source model, ${}^{(T)}_{lab}\boldsymbol{y}_S = {}^{(T)}_{lab}\boldsymbol{X} * {}^{(S)}\boldsymbol{\beta} + {}^{(S)}\beta_0$, are related in a linear model:
${}^{(T)}_{lab}\boldsymbol{y} = {}^{(T)}_{lab}\boldsymbol{y}_S * k + d \quad k, d \in R$.
The corrected prediction, ${}^{(T)}\hat{y}$, for a new spectrum, ${}^{(T)}_{unlab}\boldsymbol{x}$, is then obtained as:
${}^{(T)}\hat{y} = {}^{(T)}\hat{y}_S * k + d$,
where ${}^{(T)}\hat{y}_S = {}^{(T)}_{unlab}\boldsymbol{x} * {}^{(S)}\boldsymbol{\beta} + {}^{(S)}\beta_0$.
A clear advantage over the model updating approach is that no new calibration model has to be trained and optimized for this method.

Correction of Target Spectra

The idea behind this class of methods is based on the assumption that a correction of the target spectra in a way that makes them more similar to the source measurements may be enough for the source model to be valid again. Typical chemometric transfer approaches following this rationale are based on the availability of transfer standards (Bouveresse & Massart, 1996; Feudale, et al., 2002).
We, however, investigate two basic methods applicable also in standard-free settings which are motivated by the idea that a change in the measurement setting can eventually be modelled by a linear filter which may be approximated as an element-wise multiplication (plain intensity spectra) or element-wise addition (absorption spectra) in the frequency / spectral domain.
Hence, the first approach, named additive mean correction (AMC) by us, subtracts the mean of the available target spectra from a new target spectrum and adds the source mean spectrum. Via this simple correction, the mean over the corrected available target spectra becomes equal to the mean of the source spectra. For a target spectrum, ${}^{(T)}\boldsymbol{x}$, AMC performs
${}^{(T)}\boldsymbol{x}_{amc} = {}^{(T)}\boldsymbol{x} - {}^{(T)}\boldsymbol{\mu} + {}^{(S)}\boldsymbol{\mu}$,
where ${}^{(S)}\boldsymbol{\mu}$ denotes the source mean calibration spectrum and ${}^{(T)}\boldsymbol{\mu}$ denotes the mean over all labelled and unlabelled target spectra.
The second approach assumes a multiplicative relationship ${}^{(S)}\boldsymbol{x} = {}^{(T)}\boldsymbol{x} \; .* \; \boldsymbol{\theta}$, with $\boldsymbol{\theta} \in R^p$, between target and source spectra. We propose to compute the vector $\boldsymbol{\theta}$ as $\boldsymbol{\theta} = \frac{{}^{(S)}\boldsymbol{\mu}}{{}^{(T)}\boldsymbol{\mu}}$, thereby resulting in ${}^{(T)}\boldsymbol{x}_{mmc} = {}^{(T)}\boldsymbol{x} \; .* \frac{{}^{(S)}\boldsymbol{\mu}}{{}^{(T)}\boldsymbol{\mu}}$.

Table 1. *Properties and requirements for investigated transfer methods.*

Method	Use of orig. S model	S spectra modified	T spectra modified	S model pred. corrected directly	$lab(T)X$ used	$lab(T)y$ used	$unlab(T)X$ used or useable
PLAIN (application of source model)	+	-	-	-	-	-	-
TGT (recalibration on T data)	-	-	-	-	+	+	-
T+S and **5T+S**	-	-	-	-	+	+	-
LIN	+	-	-	+	+	+	-
AMC and **MMC**	+	-	+	-	-	-	+[4]

Due to its definition, we refer to this approach as the multiplicative mean correction (MMC) method.
Similar to LIN, no new calibration model needs to be trained and optimized for these two transfer approaches. Moreover, the possible (not mandatory) use of the set of unlabelled target spectra, which is the generally larger than the one of labelled observations, may be advantageous. Note that no target reference measurements are required / used for AMC and MMC. AMC and MMC are two simple instantiations of the more general idea to make the probability distributions of the observations in the source and target domain more similar. A concept on which many transfer learning approaches from the machine learning community are based on (e.g. Pan, et al., 2011).

Correction of Source and Target Spectra

In some cases, it may help to modify both the source and target spectra to make them more similar. If lacking transfer standards, natural approaches in this group perform standard or more advanced signal pre-processing methods. Standard pre-processing techniques include simple mathematical operations, signal centering or scaling, noise, offset and baseline filtering as well as sample normalization approaches (Feudale, et al., 2002; Woody, et al., 2004; Kramer, et al., 2008; Eigenvector Research Inc., 2015). In our application noise filtering as well as normalization techniques have been investigated. These can, however, not eliminate the need for transfer methods in many of the observed changes.
Examples of more advanced signal pre-processing techniques used in the transfer context are: finite impulse response (FIR) filtering (Blank, et al., 1996; Tan, et al., 2002), an approach strongly related to MSC, generalized moving window MSC (MW-MSC) (Kramer, et al., 2008), wavelet analysis (Feudale, et al., 2002) and a transfer learning version of orthogonal signal correction (OSC) (Wold, et al., 1998; Woody, et al., 2004). Among these techniques, OSC appears as an approach that is both easily implementable and very promising. We therefore plan to investigate this particular approach in future work.
All methods in this group require the training and usually optimization of a new calibration model.

Properties and Requirements of Different Transfer Methods

An overview over the properties and requirements of the transfer methods applied in this contribution can be found in Table 1. The signs + and - indicate that a certain column is true, or false, respectively.

Results and Discussion

Detailed Information on Measurements and Method Application

Data

Data from melamine-formaldehyde (MF) resin production are examined in this contribution. While some information on the measurement setup and model calibration is given here, a more detailed description is provided in (Pawliczek, et al., 2015). Spectroscopic data of several MF resin recipes are continuously collected in-line the production plant. To this end, a commercial FTNIR-spectrometer (i-RED Infrarot Systeme, Austria) combined with a NIR transmission immersion probe with a 1mm measurement gap has been installed. In contrast to the more common approach of studying absorption spectra, the use of single beam spectra, i.e. the transmitted light intensity, has

Table 2. *Investigated transfer problems. Results are only shown for rows in bold font.*

Transfer Problem Number	Source				Target			
	Lamp	Reactor	Recipe	Multipl.	Lamp	Reactor	Recipe	Multipl.
1	**1**	**1**	**1-4**	**-**	**2**	**1**	**1-4**	**-**
3	**1**	**1**	**1**	**-**	**2**	**1**	**1**	**-**
5	**1**	**1**	**2**	**-**	**2**	**1**	**2**	**-**
7	**1**	**1**	**4**	**-**	**2**	**1**	**4**	**-**
9	**2**	**1**	**1-4**	**+**	**3**	**1**	**1-4**	**+**
11	2	1	1	+	3	1	1	+
13	2	1	2	+	3	1	2	+
15	2	1	3	+	3	1	3	+
17	2	1	5	+	3	1	5	+
19	2	1	1-4	+	4	1	1-4	+
21	2	1	1	+	4	1	1	+
23	2	1	2	+	4	1	2	+
25	2	1	3	+	4	1	3	+
27	2	1	5	+	4	1	5	+
29	3	1	1-4	+	4	1	1-4	+
31	3	1	1	+	4	1	1	+
33	3	1	2	+	4	1	2	+
35	3	1	3	+	4	1	3	+
37	3	1	5	+	4	1	5	+
39	**1**	**1**	**1**	**-**	**1**	**2**	**1**	**-**
41	**2**	**1**	**1+3**	**+**	**2**	**2**	**1+3**	**+**
43	**2**	**1**	**1**	**+**	**2**	**2**	**1**	**+**
45	**2**	**1**	**3**	**+**	**2**	**2**	**3**	**+**
47	**3**	**1**	**1+2**	**+**	**3**	**1**	**5**	**+**

been followed, as raw data performed favourably in this industrial application.

In order to monitor the polymerization progress and to estimate the end point of the reaction, samples of the reactor content are taken regularly and analysed manually by a water tolerance method. The obtained values, denoted as turbidity points, depict a measure of reaction completeness and serve as reference measurements for developed chemometric models. The corresponding spectral measurements are assigned based on a signal input actuated manually when a sample of the reactor content is taken. In order to reduce possible measurement errors and small deviations in the manual actuation process, the three closest (w.r.t. time) NIR spectra are associated with each reference value.

The data studied in this contribution cover a row of variances: They were collected using four different light sources (lamps), two different chemical reactors and corresponding fibre lengths, partly incorporated an optical multiplexer in the measurement process and cover five MF resin recipes as well as a time range of about 3.5 years.

Due to this very long time period and the different measurement setups, various chemometric models have been developed and optimized: Some of them calibrated on a single recipe, others on data from different related recipes, some of them using only SNV (data before multiplexer installation) as data pre-processing, others performing certain extents of Savitzky-Golay signal smoothing beforehand, some using only five PLS components, others incorporating up to 8 factors, some using two quite slim spectral regions and others selecting somewhat wider variable intervals.

Even though fairly robust and accurate models could be built, a change in the measurement setup generally required and still requires model updating or recalibration and the corresponding time to collect a significant amount of new data. Furthermore, the introduction of a new MF resin recipe usually demands for a new calibration and model optimization process, particularly if the new recipe is not very closely related to an already modelled one. These issues shall be addressed in our contribution via the investigation of the transfer methods described above which are evaluated on three groups of transfer settings.

Transfer Settings

Three different transfer groups are covered in this contribution:

1) Lamp change: As there are data for four different light sources in various settings available, many of the transfer scenarios fall into this group.

2) Change of Reactor + optical fibre: Two different reactors have been used and hence constitute transfer problems.
3) Recipe change: While four of the five recipes may be modelled together, a fifth recipe does not fit inside such an overarching model, thereby providing another possible transfer setting.

Although the transfer direction is generally fixed within an industrial application, aiming e.g. at the transfer of a model for an earlier light source to the data obtained by the following lamp, we also investigate transfers in the opposite of the natural transfer direction. The reason for this is our intention to study whether or not transfers are generally equally successful in both directions. Furthermore, one may gain insight in the influence of the source sample sizes, modifications of pre-processing and / or variable selection methods and the presence of possible outliers or influential points.

24 transfer problems plus their opposite directions, thereby resulting in a total of 48 transfer scenarios, have been studied. An overview of these transfer problems, not listing scenarios to the opposite direction case[5], is given in Table 2, where lamp changes are marked in orange, reactor changes are coloured in green and the recipe change problem is shown in yellow.

Among these transfer problems, we only discuss results for those settings where the performance of the source model becomes considerably worse in the source to target transfer. We define this situation via the ratio $R=\frac{RMSET_{PLAIN}}{RMSECV_{PLAIN}}$, where $RMSECV_{PLAIN}$ is computed via 5-fold grouped (over sample IDs) cross validation over the source data and $RMSET_{PLAIN}$ denotes the RMSE obtained on all available target data when the source model is used for prediction. If $R \geq 6.5$ for at least one direction, then this transfer problem (both directions) is investigated further. As such, we report results (not all in detail) on the transfer scenarios with numbers 1-10 and 39-48 in Table 2 (bold font).

Sample Splitting and Performance Measures

As shown in Figure 1, we assume that a small set of labelled and a larger set of unlabelled target data is available for calibration transfer. In all investigated transfer settings, a certain number, ${}^{(T)}_{all}n$, of labelled target spectra is available. We split these data into ${}^{(T)}_{lab}\boldsymbol{X}$, ${}^{(T)}_{unlab}\boldsymbol{X}$ and ${}^{(T)}_{test}\boldsymbol{X}$ with their corresponding reference values, where ${}^{(T)}_{unlab}\boldsymbol{y}$ remains unused. The number of labelled samples[6], for which the reference value is permitted to be used during transfer, is varied between 5, 10, 15, 20, 50 and 70[7] samples, $0.5 * {}^{(T)}_{all}n$ of the remaining target data are chosen randomly and used as test set, and all other data form the unlabelled target data set.

Different ways (target sample selections), how to choose / select the l=5, 10, 15, 20, 50 or 70 labelled target samples have been investigated: 1) First denotes the way usually best applicable in an industrial NIR in-line application and uses the first l samples which become available (in chronological order), 2) Random selects l target samples randomly among all available labelled ones and 3) the choice method KennardStone follows the Kennard-Stone algorithm (Kennard & Stone, 1969) to find the subset of l labelled target data accessible for transfer in order to sample the target space optimally; see (Kennard & Stone, 1969) for details.

As was found for various approaches known to obtain representative transfer standards (Bouveresse & Massart, 1996), we also see that different sample selection methods do generally lead to considerably different performances. In order to reduce the influence a particular random data split might have, splitting is repeated five times. Mean values as well as standard deviations[8] over the five obtained $RMSEP=rmseTest\left({}^{(T)}_{Test}\hat{\boldsymbol{y}}, {}^{(T)}_{Test}\boldsymbol{y}\right)$ values serve as performance measures in the graphical illustrations below (different transfer methods were evaluated on the same data splits).

In the following we only present the results for the First and KennardStone target sample selection as we observed that results obtained via a random selection are generally very similar to ones based on a Kennard-Stone selection.

Pre-processing and Model Building

It has been mentioned above, that generally one aims to find a quick and inexpensive way to provide accurate predictions for the target setting. Therefore, we follow the approach to retain the pre-processing methods and variable selection applied in the source calibration model. Moreover, in case of TGT, T+S and 5T+S, where a new PLS model is trained, the number of PLS components is fixed at the same value as used for the source model. AMC and MMC are applied to the pre-processed selected spectral regions.

Software

Results were generated via Matlab code written in-house. All algorithms rely on the software Matlab® release 2012a (Mathworks Inc., Natick, MA, USA) and the Matlab® Statistics Toolbox 2014.

Results for Lamp Change

Among the three transfer groups investigated, the source model usually gives the best results after a lamp change. In numerous cases, though, the accuracy obtained via this plain application of the model trained on source data is far beyond an acceptable threshold

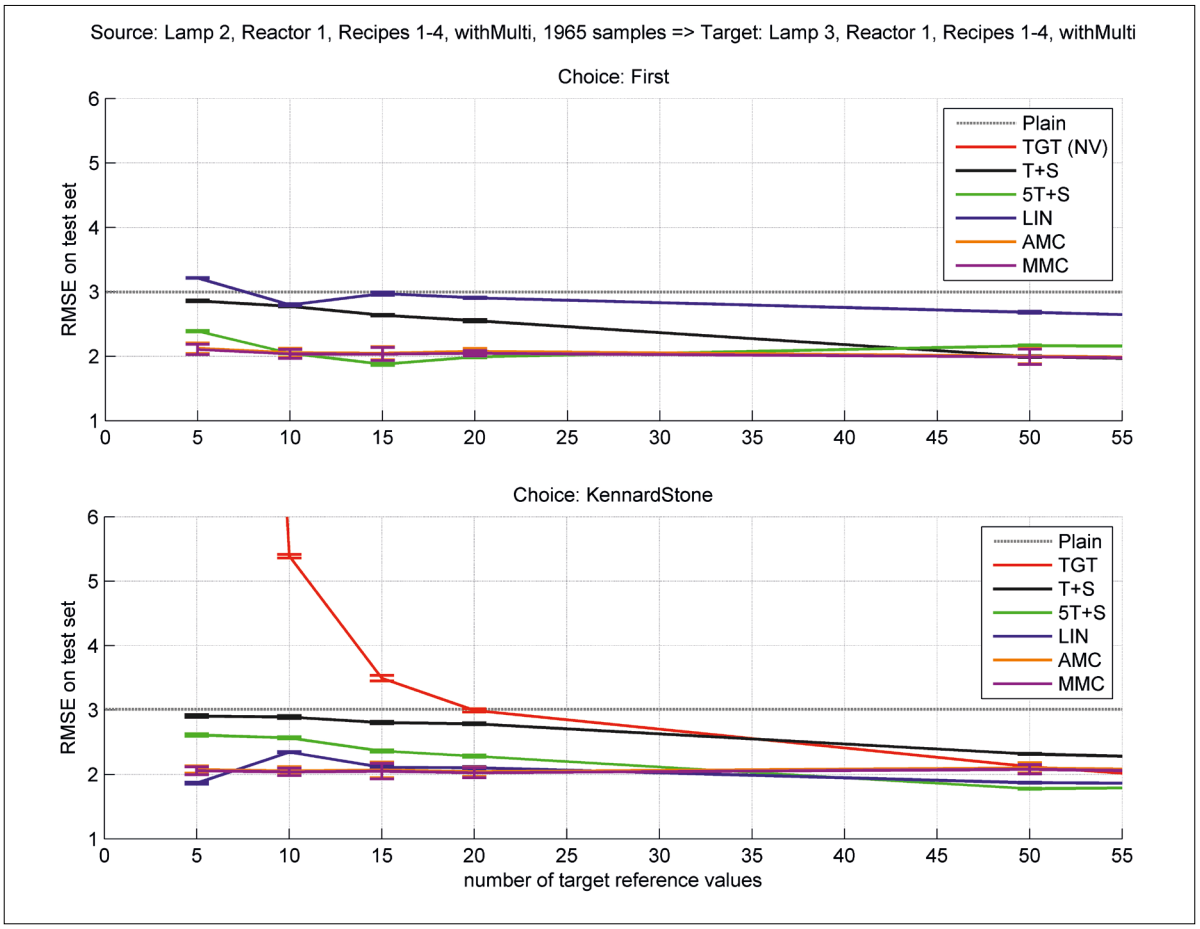


Figure 2. *Prediction performance of different transfer methods for scenario 9 (lamp 2→3). NV indicates that the results of corresponding method are larger than the shown RMSE range.*

and transfer methods shall be applied. Here we discuss results for two different lamp transfers (lamp 2→3; lamp 1→2) and their opposite directions (transfer problems 9/10 and 3/4 in Table 2 where 10 and 4 are the opposite direction).

Results for Lamp Change 2→3 and 3→2 (Scenario 9 and 10)

Figure 2 shows results for choice methods First and KennardStone in case of transfer problem 9. TGT is the worst transfer technique for sample selection method First and too large to be shown on the chosen scale. The corresponding curve also exhibits a high dependence on the chosen samples, resulting in a local minimum (RMSEP ~8) at 50 available target reference measurements. If compared to TGT's results based on a Kennard-Stone selection, it becomes obvious that a simple model calibration based on e.g. 20 labelled target samples can be relatively accurate if the samples are chosen to be representative for the entire target space, which is assumed to be the case for the KennardStone choice. Such an observation has been made for several scenarios.

The plain application of the source model appears to be working fairly well considering that a lamp change has occurred between model calibration and application to the target data. It can however be seen that a considerable improvement of prediction accuracy can be reached by the application of the mean correction methods AMC and MMC (these two methods yield very similar results and are almost indistinguishable in Figure 2) as well as via 5T+S model updating. In case of choice method KennardStone, the same is also true for LIN. While model updating and LIN incorporate the available labelled target data, it is remarkable to see that the methods AMC and MMC which are only based on unlabelled T data can also perform very well in this setting. Although the performance improvement over the source model may not be exceptionally large for this transfer problem, the successful approaches may however be valuable to bridge the time needed to collect enough target data for the calibration of an entirely new target model.

While it may be discussed whether or not the application of transfer methods is really necessary for scenario 9, this is certainly not a question for scenario 10 (opposite lamp change 3→2.) It can be seen in Figure 3 that the plain use of the S model results in an unacceptably large RMSEP. As in practice one is usually

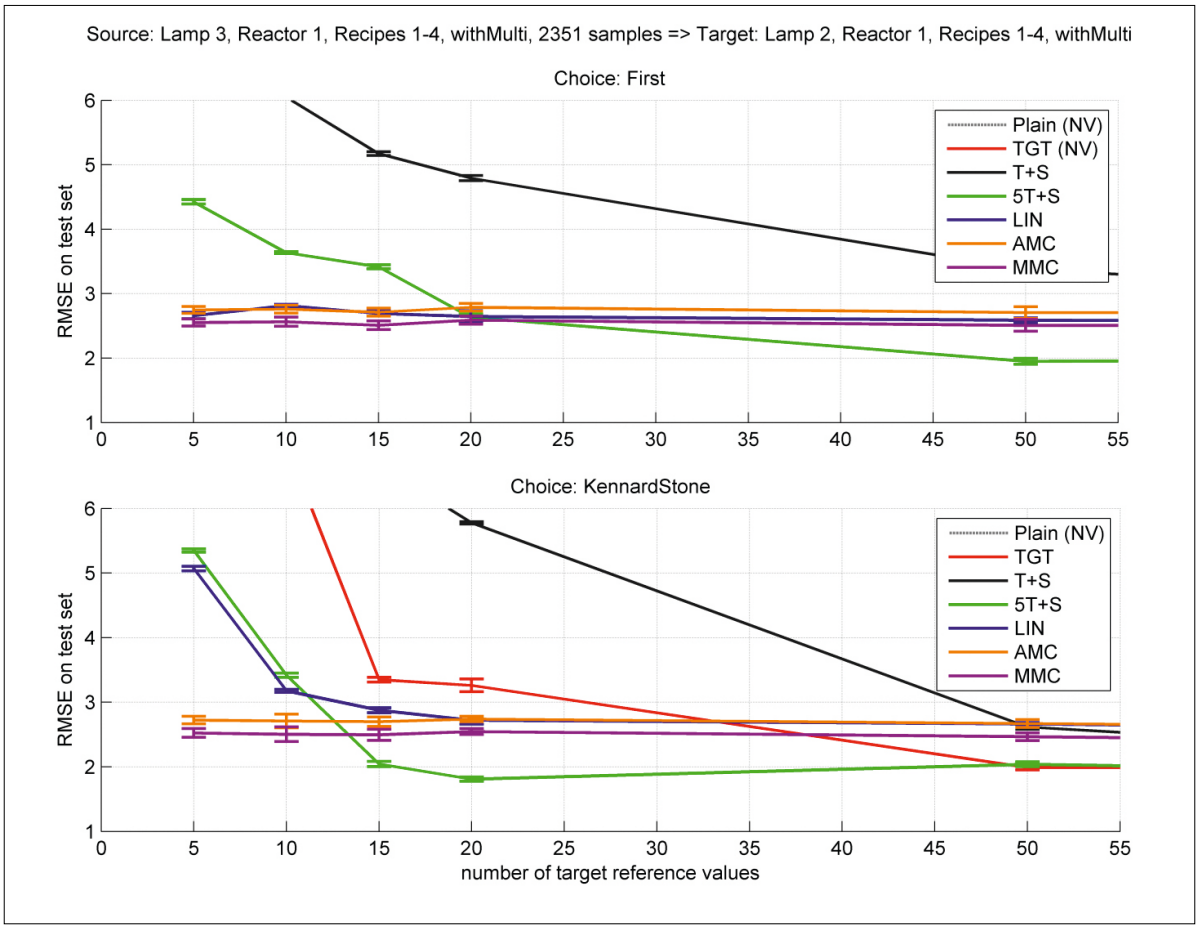


Figure 3. *Prediction performance of different transfer methods for scenario 10 (lamp 3→2). NV indicates that the results of corresponding method are larger than the shown RMSE range.*

stuck to a selection equal or similar to the one from sample selection method First, TGT cannot be seen as a valid alternative either (too large to be shown on the chosen scale). The application of transfer methods LIN, AMC, MMC (LIN and AMC, as well as LIN and MMC, are partly almost indistinguishable in Figure 3) and 5T+S, though, can provide a solution for this transfer scenario.

Results for Lamp Change 1→2 and 2→1 (Scenario 3 and 4)

While performance results and corresponding conclusions for scenario 3 do overly agree with ones described for scenario 10 and are therefore not contained here, a difference can be observed for scenario 4. As shown in Figure 4, RMSEP values obtained for choice method First are considerably larger for all examined transfer methods. Certainly, the model updating, LIN and mean correction methods can achieve a considerable improvement compared to PLAIN and TGT (both not shown in this subplot as too large), the observed performance, though, is still not satisfying for the application context. We believe that our straight-forward approaches may be too simple to capture enough information from the source setting to yield a more pleasing transfer result. Because of this, a row of further, more sophisticated, techniques shall be evaluated in future work. Additionally, the influence of source sample sizes on transfer results shall be investigated.

Results for Reactor + Fibre Change

Compared to the lamp change scenarios, a plain application of the source calibration model after a reactor and fibre change results in extraordinarily large errors from 24 up to 80 times the source RMSECV. As such, there is no question to whether or not the training of a new model (TGT) or some form of transfer methods should be applied. Herein we include results for transfer scenarios 45 (reactor 1→2) and 46 (reactor 2→1).

In Figure 5 it can be seen that all applied forms of transfer learning provide a significant and satisfying improvement over PLAIN (not shown, RMSEP around 70). Only in case of a more spread selection of labelled target samples, as performed by the Kennard-Stone algorithm, TGT becomes an alternative at some point. If the number of available target reference values is lower than 50 or, as usually the case in industry, only samples obtained first are available, TGT presents no

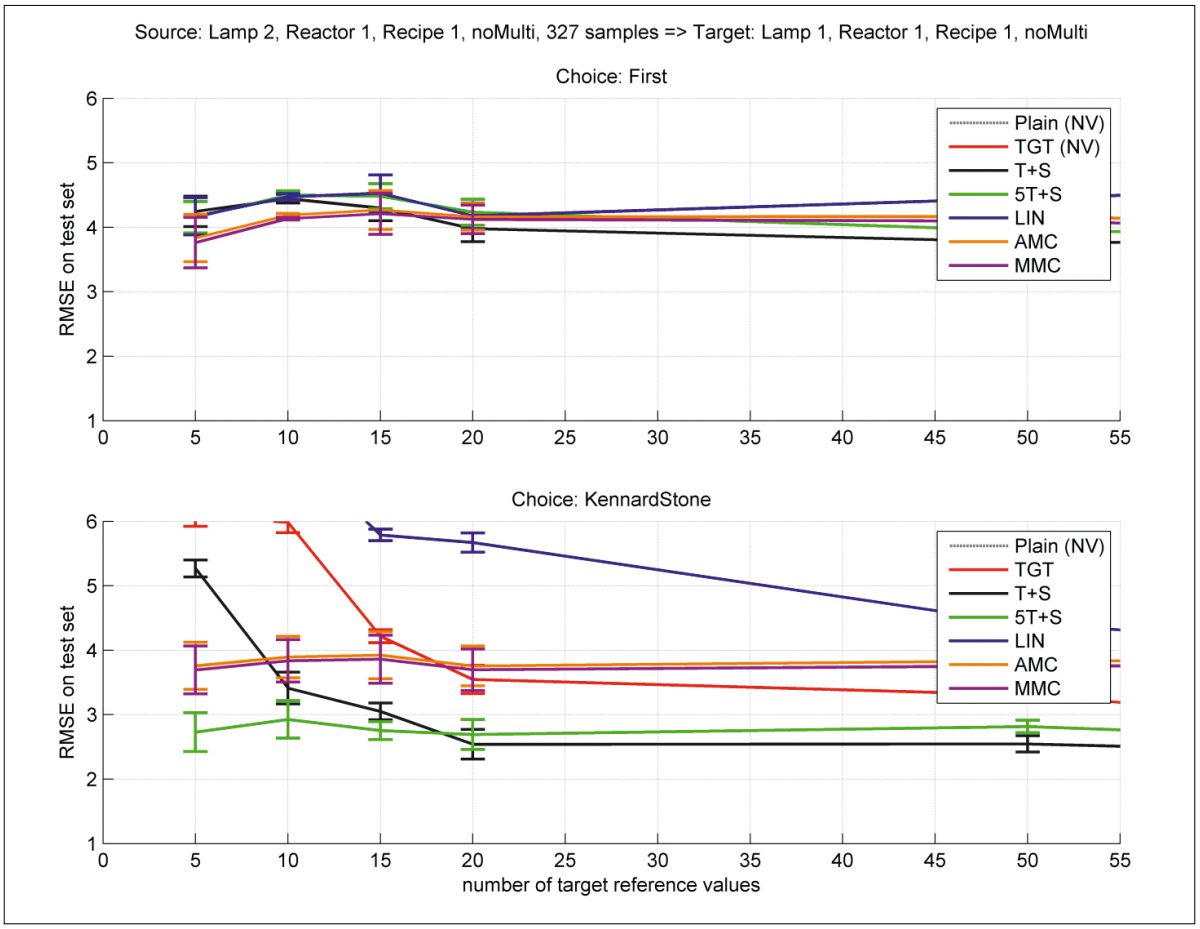


Figure 4. *Prediction performance of different transfer methods for scenario 4 (lamp 2→1). NV indicates that the results of corresponding method are larger than the shown RMSE range.*

valid option. It is obvious again that the adopted choice method plays a significant role for results. Furthermore, it can be seen that LIN, T+S and 5T+S (these methods are partly almost indistinguishable in Figure 5) outperform AMC and MMC in this setting.

In case of scenario 46 (reactor 2→1), a less complex source model has been trained. This may be a reason why RMSEP values for the target setting, displayed in Figure 6 are relatively high for the model updating methods T+S and 5T+S (not visible in choice First, as too large). In contrast, LIN, AMC and MMC perform better in this situation, although only MMC produces relatively satisfying results for choice method First. A large improvement for all methods, particularly for TGT (also not visible in choice First, as too large) and 5T+S, is observed in the KennardStone case. Overall, though, a more satisfying result, especially for First, would be desirable.

Results for Recipe Change

Last but not least, the recipe change scenario 47 is discussed. Again, PLAIN performs rather poorly (RMSEPs around 55) and cannot be used for prediction in the target space. It can be seen in Figure 7 that MMC (not visible in the figure, since RMSEP values lie around 25) is not an option for any choice method in this setting (and also the opposite scenario 48). LIN performs best in all cases and is closely followed by TGT if choice method KennardStone and at least 15 target reference measurements are used. For choice method First, though, TGT jumps up and down seemingly arbitrarily. Thus, once again it can be seen that TGT is not robust with respect to the available target samples.

Discussion

Considerable deviations are observed for results based on different target sample selection methods (or possibilities). While all transfer methods appear to perform better on a wider spread sample selection (Random and KennardStone), TGT appears to be affected most by these changes. This is certainly comprehendible, as TGT trains a new model solely based on the new T samples. If these samples are not representative for the whole T space or in case that an observation does not perfectly fit into this space, the calibrated model will perform poorly.

The results presented above show that transfer problems may be solved or at least considerably reduced via practi-

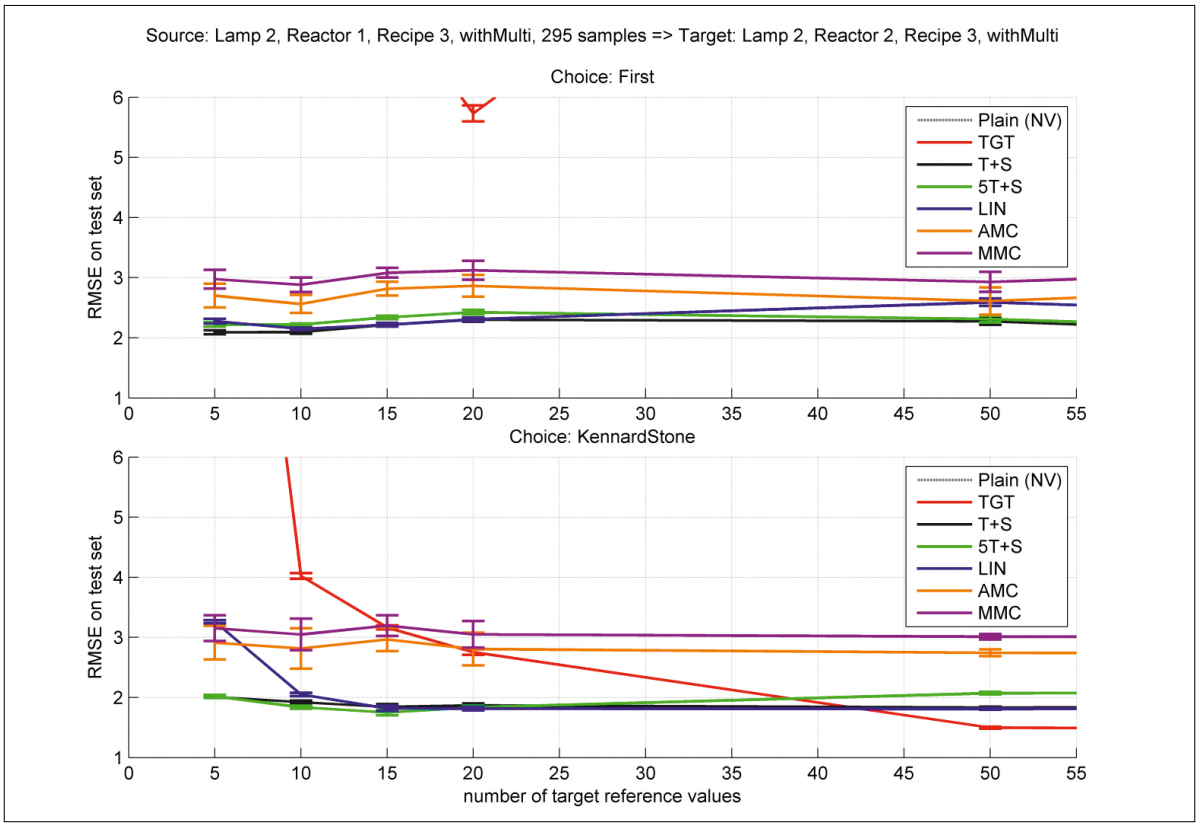


Figure 5. *Prediction performance of different transfer methods for scenario 45 (reactor 1→2). NV indicates that the results of corresponding method are larger than the shown RMSE range.*

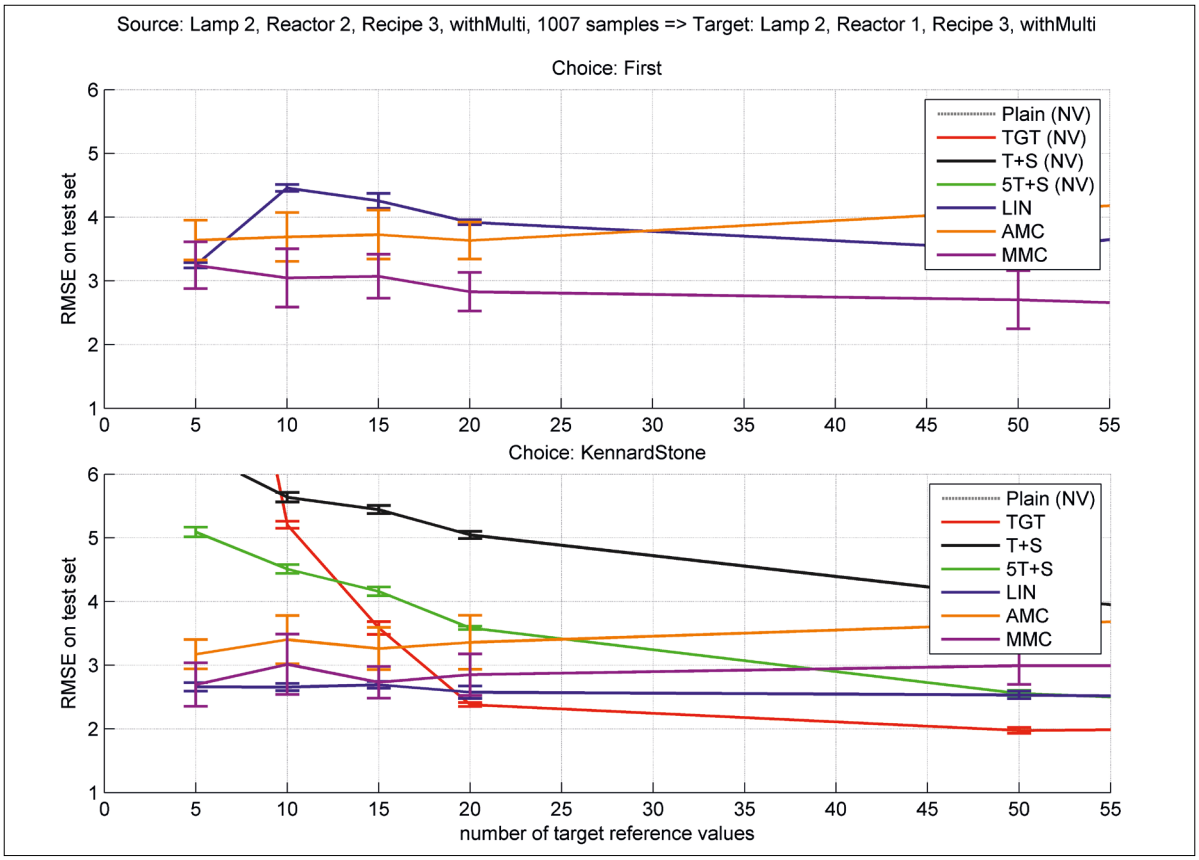


Figure 6. *Prediction performance of different transfer methods for scenario 46 (reactor 2→1). NV indicates that the results of corresponding method are larger than the shown RMSE range.*

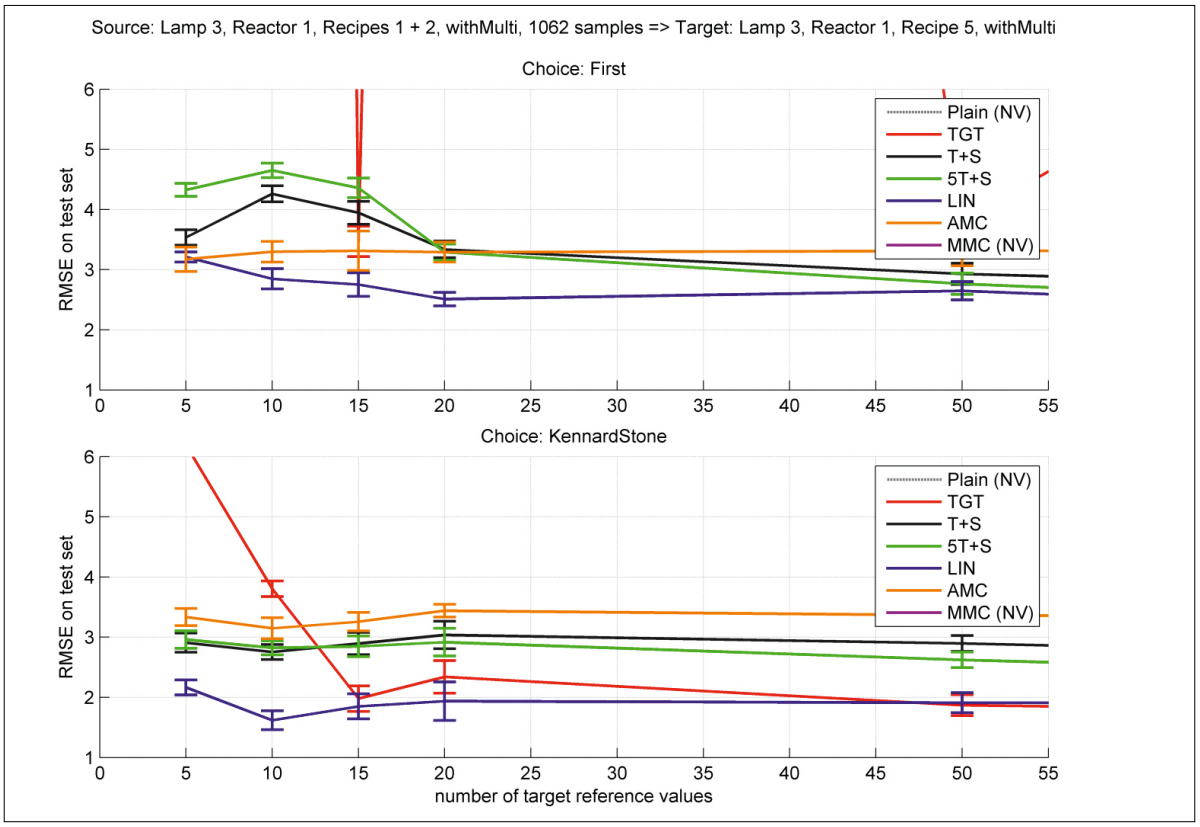


Figure 7. *Prediction performance of different transfer methods for scenario 47. NV indicates that the results of corresponding method are larger than the shown RMSE range.*

cally easily applicable transfer methods. It is natural to raise the question, which of the discussed approaches performs best in a specific transfer problem or in general. While there is no unique answer to this question yet, we summarized the results obtained for scenarios 1-10 and 39-48 in form of average rank maps, found in Figure 8 to Figure 10. In these graphics, we see the average rank of methods over all these scenarios (Figure 8), over the corresponding lamp scenarios, i.e. 1-10, (Figure 9) and over the corresponding reactor+fibre transfer scenarios, i.e. 39-46 (Figure 10), respectively[9]. To this end, we studied which method yielded the lowest (rank 1), second lowest (rank 2), etc RMSEP for a certain scenario, target sample selection method and number of available labelled target samples. Displaying the average ranks, one gets a graphical impression of the methods performing overly well or poorly for a row of transfer problems. As such, it can be seen in Figure 8 that, as assumed, the plain application of the S model presents generally the worst option. The use of TGT performs not much better and is only recommended for a reasonably large number of representative target samples (via choice methods Random or KennardStone). All other methods are usually preferable, where AMC and 5T+S appear best. The higher average ranks of AMC and MMC in the bottom right corner of the KennardStone map can be explained by the fact that these methods do not use any reference values and yield relatively constant RMSEPs over all considered numbers of available labelled target data therefore (variation mostly due to randomized selection of test data). The other methods, except from PLAIN, on the contrary, make use of these labels and can thereby improve their performance with growing labelled target sample size. As such, they obtain lower ranks and displace AMC and MMC.

In contrast to the overall maps shown in Figure 8, average ranks for the lamp specific scenarios, displayed in Figure 9, show a certain shortcoming of method LIN. As such, it may be beneficial to investigate other methods initially. Above this observation, the maps are similar to the overall case.

Figure 10, on the other hand, may lead to the conclusion that LIN appears a valid approach to apply first.

In all settings though, 5T+S appears to be relatively high ranked. It can also be observed in these maps as well as from the actual RMSEP values that the weighted model updating (5T+S) performs generally better than normal model updating (T+S), particularly for the more interesting areas with small numbers of available labelled T data. This makes us believe that an even stronger weighting may be beneficial; thus, such a stronger weighting shall be investigated in future work.

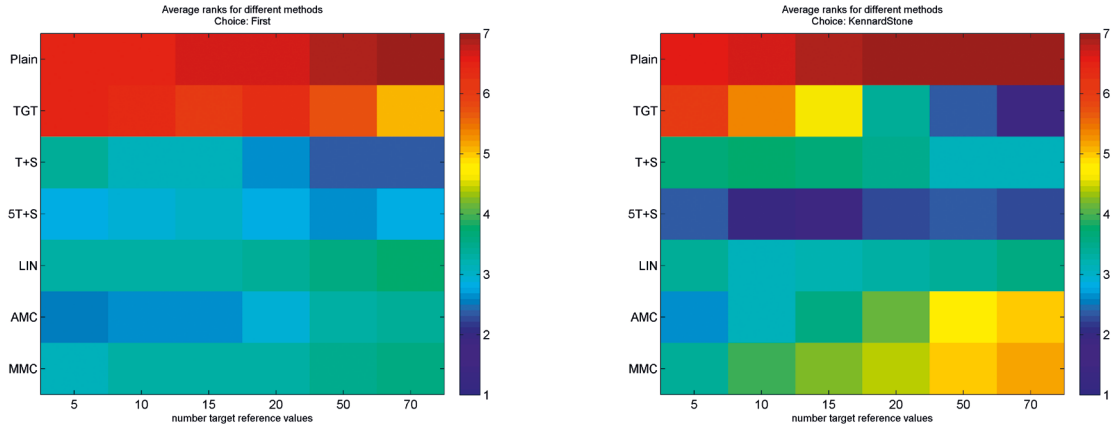


Figure 8. *Average rank for evaluated methods computed over all scenarios.*

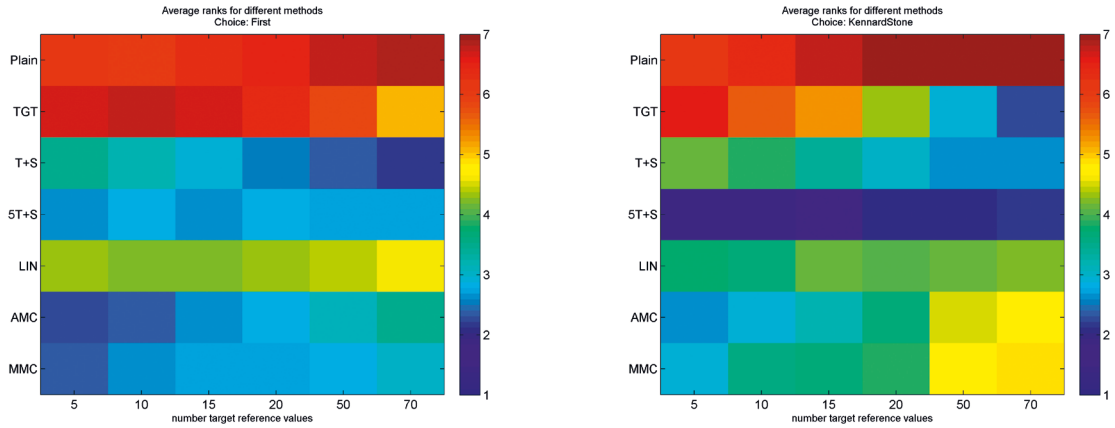


Figure 9. *Average rank of evaluated methods computed over the lamp change scenarios.*

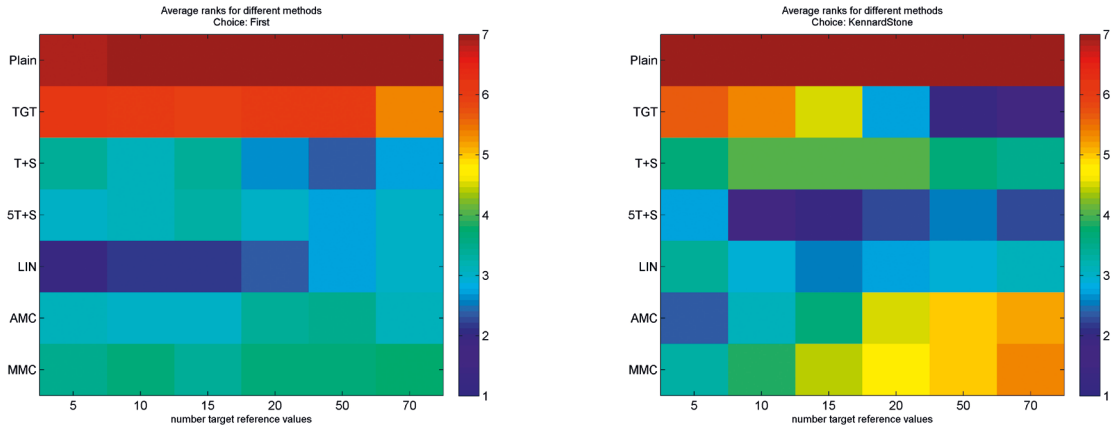


Figure 10. *Average rank of evaluated methods computed over the reactor+fibre scenarios.*

Conclusions

Within this contribution we show that a row of rather simple transfer approaches can achieve considerable improvement of predictions for data from a changed setting. The methods introduced and discussed here are both easy to comprehend and implement as well as application-oriented in the sense that they do not assume the availability of transfer standards and require only a few labelled data in the target space. Regarding a suitable choice among the proposed methods, average rank maps may provide guidance.

Although the results presented here prove the suitability of the investigated straight-forward methods, there is still potential for improvement with respect to both, prediction accuracy and robustness (regarding the choice method as well as the number of available target samples). Because of this, we intend to investigate

further and more sophisticated transfer methods from both chemometrics and machine learning in future work. To this end, we are currently analysing the transfer OSC technique (Woody, et al., 2004), as well as an approach based on a combination of transfer component analysis (Pan, et al., 2011) and linear regression. Additionally, a more advanced weighting instead of the 5T+S approach and a possible idea to develop artificial transfer standards shall be studied.
Last, but not least, it shall be mentioned that it is not yet clear at this point, what role different source sample sizes and calibration settings may play for subsequent transfer results. Thus, this question shall also be investigated in the future.

Acknowledgements

This work was funded by the Austrian research funding association (FFG) under the scope of the COMET programme within the research project "Industrial Methods for Process Analytical Chemistry - From Measurement Technologies to Information Systems (imPACts)" (contract # 843546). This programme is promoted by BMVIT, BMWFW, the federal state of Upper Austria and the federal state of Lower Austria.

References

Blank, T. B., Sum, S. T. & Brown, S. D., 1996. Transfer of Near-Infrared Multivariate Calibrations without Standards. Anal. Chem., Volume 68, pp. 2987-2995.

Bouveresse, E. & Massart, D., 1996. Improvement of the piecewise direct standardisation procedure for the transfer of NIR spectra for multivariate calibration. Chemometrics and Intelligent Laboratory Systems, Volume 32, pp. 201-213.

Bouveresse, E. & Massart, D., 1996. Standardisation of near-infrared spectrometric instruments: A review. Vibrational Spectroscopy, Volume 11, pp. 3-15.

Crammer, K., Michael, K. & Wortman, J., 2008. Learning from Multiple Sources. The Journal of Machine Learning Research, Volume 9, pp. 1757-1774.

De Noord, O. E., 1994. Multivariate calibration standardization. Chemometrics and Intelligent Laboratory Systems, 25(2), pp. 85-97.

Du, W. et al., 2011. Maintaining the predictive abilities of multivariate calibration models by spectral space transformation. Analytica Chimica Acta, Volume 690, pp. 64-70.

Eigenvector Research Inc., 2015. Eigenvector Wiki - Preprocessing. [Online]

Available at: http://wiki.eigenvector.com/index.php?title=Preprocessing [Accessed 11 09 2015].

Fan, W., Liang, Y., Yuan, D. & Wang, J., 2008. Calibration model transfer for near-infrared spectra based on canonical correlation analysis. Analytica chimica acta, Volume 623, pp. 22-29.

Feudale, R. N. et al., 2002. Transfer of multivariate calibration models: a review. Chemometrics and Intelligent Laboratory Systems, Volume 64, pp. 181-192.

Kennard, R. W. & Stone, L. A., 1969. Computer Aided Design of Experiments. Technometrics, Volume 11, pp. 137-148.

Kramer, K. E., Morris, R. E. & Rose-Pehrsson, S. L., 2008. Comparison of two multiplicative signal correction strategies for calibration transfer without standards. Chemometrics and Intelligent Laboratory Systems, pp. 33-43.

Mansour, Y., Mohri, M. & Rostamizadeh, A., 2008. Domain Adaptation with Multiple Sources. s.l., Neural Information Processing Systems.

Ni, W., Brown, S. D. & Man, R., 2011. Stacked PLS for calibration transfer without standards. Journal of Chemometrics, 25(3), pp. 130-137.

Pan, S. J., Kwok, J. T. & Yang, Q., 2008. Transfer Learning via Dimensionality Reduction. s.l., Association for the Advancement of Artificial Intelligence.

Pan, S. J., Tsang, I., Kwok, J. T. & Yang, Q., 2011. Domain adaptation via transfer component Analyisis. Trans. on Neural Networks, pp. 199-210.

Pawliczek, M. et al., 2015. Long-term in-line monitoring of melamine formaldehyde resins using NIR spectroscopy. Lenzinger Berichte, p. in print.

Peng, J., Peng, S., Jiang, A. & Tan, J., 2011. Near-infrared calibration transfer based on spectral regression. Spectrochimica Acta, Volume 78, pp. 1315-1320.

Tan, H., Sum, S. T. & Brown, S. D., 2002. Improvement of a Standard-Free Method for Near-Infrared Calibration Transfer. Appl. Spectrosc., Volume 56, pp. 1098-1106.

Wang, Y., Lysaght, M. J. & Kowalski, B. R., 1992. Improvement of multivariate calibration through instrument standardization. Anal. Chem, 64(5), p. 562–564.

Wang, Y., Veltkamp, D. & Kowalski, B. R., 1991. Multivariate instrument standardization. Anal. Chem., 63(23), pp. 2750-2756.

Wise, B. M. & Roginski, R. T., 2015. A Calibration Model Maintenance Roadmap. IFAC-PapersOnLine, 48(8), pp. 260-265.

Wold, H., 1966. Estimation of principal components and related models by iterative least squares. In: Multivariate Analysis. New York: Academic Press, pp. 391-420.

Wold, S., Antti, H., Lindgren , F. & Öhman, J., 1998. Orthogonal signal correction of near-infrared spectra. Chemometrics and Intelligent Laboratory Systems, 44(1-2), pp. 175-185.

Wold, S., Sjöström, M. & Eriks, L., 2001. PLS-regression: a basic tool of chemometrics. Chemometrics and Intelligent Laboratory Systems, 58(2), pp. 109-130.

Woody, N. A., Feudale, R. N., Myles, A. J. & Brown, S. D., 2004. Transfer of Multivariate Calibrations between Four Near-Infrared Spectrometers Using Orthogonal Signal Correction. Anal. Chem, pp. 2595-2600.

[1] The research reported in this article has been partly supported by the Austrian Ministry for Transport, Innovation and Technology, the Federal Ministry of Science, Research and Economy, and the Province of Upper Austria in the frame of the COMET center SCCH.

[2] Weighting is achieved by fully resampling / copying the target data five times without adding any noise.

[3] We choose this value, in order to guarantee, that the weighting of the target data does only exceed the weight of the source data in isolated investigated cases (namely those cases where the number of source samples is low and the number of incorporated labelled target samples is high). A higher weighting would result in a performance rather similar to full recalibration for certain transfer scenarios, while a lower weighting would provide results barely distinguishable from T+S.

[4] Alternatively, ${}^{(T)}_{lab}\boldsymbol{X}$ can be used.

[5] This is the reason why only impair transfer problem numbers are shown in Table 2. Transfer problem number 2 denotes the opposite direction of transfer scenario 1, transfer problem 4 denotes the opposite direction of transfer scenario 3, etc.

[6] The term samples refers to a measured reference value and its corresponding three assigned spectra.

[7] The figures below do only cover the region up to 55 samples.

[8] Note that standard errors can also be obtained for choice method First. This is due to the fact that the test set is drawn randomly and can therefore vary.

[9] We did not include a map for the recipe change transfer, since there were only two scenarios (47 and 48) investigated for this case.

Long-term In-line Monitoring of Melamine Formaldehyde Resins using NIR Spectroscopy

Marcin Pawliczek[1], Thomas Reischer[2], Jakub Kowalski[3], Wolfgang Märzinger[3], Wolfgang Kantner[2] and Markus Brandstetter[1]

[1] RECENDT – Research Center for Non-Destructive Testing GmbH, 4040 Linz, Austria
[2] Metadynea Austria GmbH, 3500, Krems, Austria
[3] i-RED Infrarot Systeme GmbH, 4020, Linz, Austria

Abstract

Within the K-project "Process Analytical Chemistry - PAC" key parameters of various industrial production processes were monitored by using FTNIR-spectroscopy and chemometrics. In case of melamine-formaldehyde (MF) resin production one of the key parameters is the turbidity point during MF resin condensation reaction. This parameter is typically determined off-line. In this work a method for in-line monitoring the turbidity point was developed and a long-term study was performed. For this purpose, in-line FTNIR spectral and off-line analytical measurements were performed for a set of calibration data and used for developing robust and reliable chemometric models for the turbidity point. Specific models were developed for different resin recipes and optimized continuously. The obtained results confirm that, despite recipe-specific spectral differences, a common model for the turbidity point is applicable for various recipes. The developed in-line method was capable of determining the turbidity point with a validated accuracy between 2.7% and 4.4% depending on the recipe. Thereby, the stopping point of the production process could be determined with a maximum time resolution of 8 seconds. Furthermore, the number of manual sampling steps could then be reduced significantly depending on the recipe. The chemometric models have to be recalibrated rarely, mostly after maintenance operation (e.g. change of the light source).

Keywords: *NIR spectroscopy, Partial Least Squares, melamine resin*

Introduction

Melamine formaldehyde (MF) resins are among the most widely used industrial thermosetting plastic materials, e.g. in high-pressure laminates for furniture or flooring (1). In general, industrially produced MF resins are the result of condensation of melamine with formaldehyde in water at a certain temperature, pH and solids content (2). As described in (3) the final product is prepared in two separated condensation stages. During the first stage formaldehyde reacts with melamine in aqueous solution. The obtained pre-condensate mixture consists of the different monomeric as well as short linear and branched oligomeric methylolmelamine compounds (Figure 1). The most important determining factors for the composition and quality of the MF resin are temperature, pH and duration of the condensation step as well as the order and time course of heating together with reagent addition. In a second stage the resin solutions are cured by the application of heat and pressure to give an insoluble highly cross-linked MF resin (Figure 1) (3).

To control the polycondensation progress of melamine and formaldehyde and to estimate the end point of the reaction samples are regularly taken and analysed by a water tolerance method. For that purpose a certain amount of resin solution is placed in a test tube and mixed with a certain amount of water and cooled till the sample becomes turbid. This method is a manual procedure based on visual inspection, thus influenced by the human factor. The obtained values are denoted as turbidity points and serve as reference for process control. In order to ensure strict control of the resin condensation during the batch process an operator is fully occupied with the acquisition

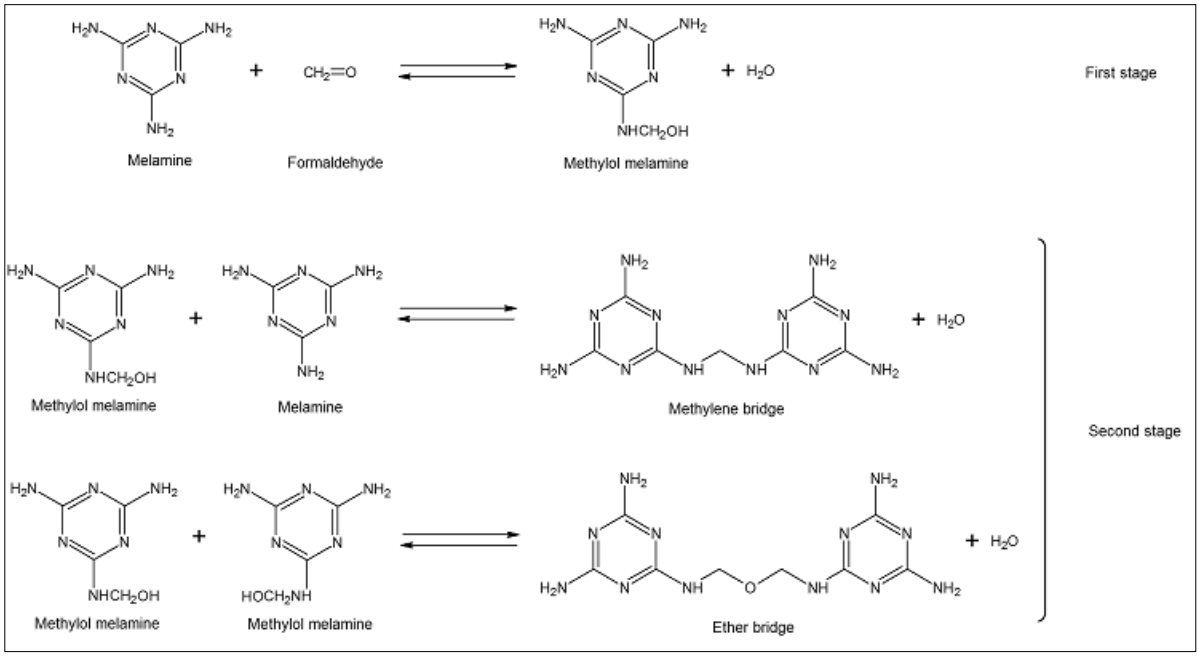


Figure 1. *Reaction scheme of MF resin condensation.*

of turbidity points. In order to make better use of human resources and to reduce sampling errors an in-line analytical method was developed.

There are several reasons for choosing Fourier-transform near-infrared (FTNIR) spectroscopy in combination with multivariate data analysis for this application. This method combination offers a fast quantitative and qualitative, as well as non-destructive and reagent-free analysis of liquids. Furthermore, NIR spectroscopy can be implemented with relatively small effort in existing control systems for process monitoring and quality control (4).

Materials and Methods

Measurement Equipment

At the production plant a commercial FTNIR-spectrometer (i-RED Infrarot Systeme, Austria) (5) and two NIR transmission immersion probes with a 1 mm measurement gap connected via fiber optics (32 m distance) and an optical multiplexer (Leoni, Germany) have been installed (Figure 2). Along with spectral data additional parameters such as process temperature and pressure are recorded via a CANopen® fieldbus connection to the process control system.

Spectral Data and Reference Values

Spectroscopic data of several recipes of MF resin are continuously collected in-line at the production plant (Figure 3). At the same time, the corresponding reference values (i.e. turbidity point) were collected by several qualified operators – up to nine values per batch. Assignment of reference values acquired by

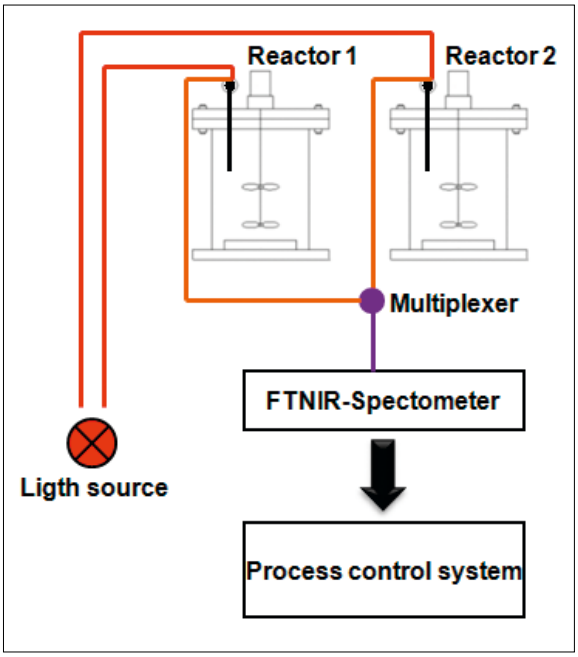


Figure 2. *Schematic diagram of the implemented FTNIR in-line measurement.*

off-line analysis to the measured spectral data is achieved by a signal input which is manually actuated by the operator. For each reference point three corresponding spectra are assigned which are collected within a time period of 24 seconds.

During these measurements the single beam spectrum, i.e. the transmitted light intensity was used as raw data instead of the absorption spectrum. Thus, obtained data contains characteristics of the light source and MF resin. This approach was found to deliver best results in subsequent data analysis.

Data Analysis

The analysis of spectra and chemometric modeling on calibration data was performed using Partial Least Squares (PLS) regression analysis (Matlab (Mathworks) and PLS-Toolbox (Eigenvector Research)). Spectral pre-processing included Savitzky-Golay (SG) filtering and standard normal variate (SNV) scaling. This pre-processing approach was defined after extensive investigations of several MF-resin recipes and production batches. The accuracy of spectroscopic/chemometric prediction of the turbidity point strongly depended on the selection of relevant spectral regions which were identified and optimized during model calibration. Finally, the calibration results were validated by cross-validation (CV) before a specific model was implemented in-line. The chemometric evaluation of the measured spectral data was performed on the spectrometer by the software chemo engine (6). The results were provided in real time to process control system.

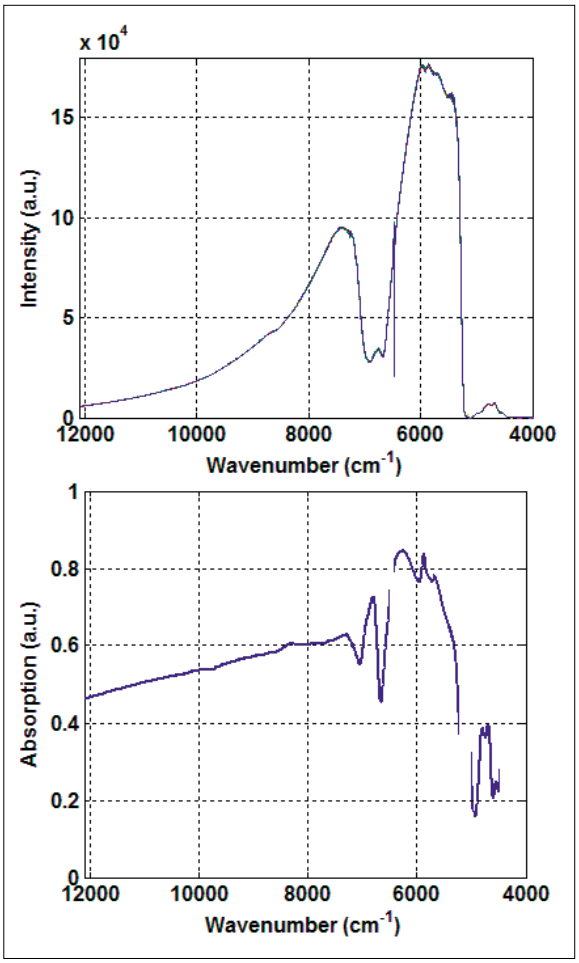


Figure 3. *Typical NIR single beam intensity (top) and absorption spectra (bottom) of MF resin collected at the production plant.*

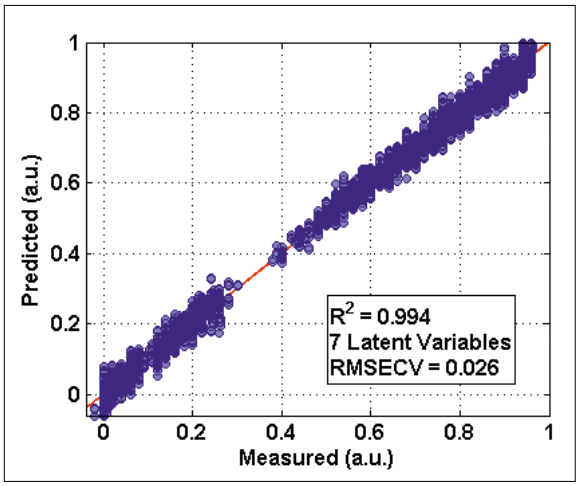


Figure 4. *Result of cross-validation for the overarching model for four similar products (recipes).*

Results and Discussion

Chemometric Models

Satisfying correlations (R2) between manually determined reference values (turbidity points) and spectral features of MF resin could be confirmed by PLS-analysis (Figure 4). Application of defined pre-processing methods and choosing of relevant spectral regions turned out crucial to achieve a well-functioning model. A significant amount of spectral data from several batches was necessary to get reasonably stable calibration models. This was owed to the rough industrial measurement environment at the production site. To minimize time consuming re-calibration, the focus of modelling was set on the long-term stability of the chemometric models. Concerning the different MF-resin recipes it was found that the more rarely produced recipes show better accuracy when applying an overarching model created using data from several similar recipes instead a recipe-specific model. This effect was ascribed to varying measurement conditions, involving cleaning steps of the production reactor. Such effects are modelled best, if a larger calibration set is available. An additional advantage of an overarching model is the shorter down-time of the method, required for collection of spectral data and reference values. For frequently produced MF resins there is no relevant difference in accuracy between recipe-specific models and the overarching model (Table 1). The achieved model prediction accuracy was validated by manually taken reference values and is either given by the root mean square error of cross-validation (RMSECV) or by the RMSE of prediction (RMSEP). The latter is obtained by an independent validation data set. Typical values of RMSECV and RMSEP were found to be in the range of 2.7% - 4.4%, which is highly satisfactory for the considered pro-

Table 1. *Calibration und prediction results of chemometric modelling for several recipes. Comparison includes both model types: overarching and recipe-specific.*

Recipe	Calibration Quality parameters of cross validation: R^2, RMSECV [%]		Validation (14.07.2014 - 12.11.2014) Quality parameter of prediction: RMSEP [%]		
	Overarching model (including four similar recipes)	Recipe-specific models	Overarching model (including four similar recipes)		Recipe-specific models
1	0.994, 2.6	0.995, 2.3	3.6	3.4	3.6
2		0.995, 2.5		3.6	4.2
3		0.995, 2.2		3.6	4.4
4		0.994, 2.8		3.9	3.9
5	---	0.993, 2.6	---		2.7

cess. Therefore, the method was found to be suitable for practical use in in-line process monitoring. Concerning the RMSECV and RMSEP values it has to be noted, that the chemometric models were validated with the reference method. In the present case the method error of the reference method (water tolerance test) was approximately 2%.

Long-term Stability

The established models have been in use for routine process monitoring over a period of six months. The so achieved turbidity points were provided to the process control system. The long-term accuracy was generally satisfying, except for short periods in which not yet known disturbances occurred (Figure 5). The amount of manual samplings per batch could be significantly reduced per batch. In some periods even a single manual measurement per batch was sufficient. As long as the obtained prediction error was within the confidence interval the end point of condensation was determined by in-line values only. In certain production batches the condensation process was fully controlled based on in-line values of the turbidity point. No negative influence on the quality of MF resins was observed.

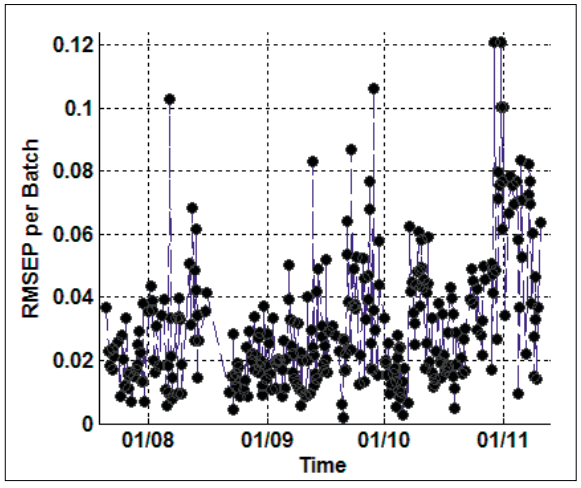


Figure 5. *Trend of the prediction error of the turbidity point (calculated as RMSEP per batch) between 14.07.2014 and 12.11.2014.*

Robustness

For industrial routine monitoring high robustness of the chemometric model is essential. Several effects on the in-line measurement, which have an influence on the robustness of the prediction model, were observed. These influences include, e.g., a slow continuous decrease of the light source intensity of the spectra (calculated as integral of the spectrum, see Figure 6). However, by using a larger calibration data set, which included the decrease of light intensity, adverse ef-

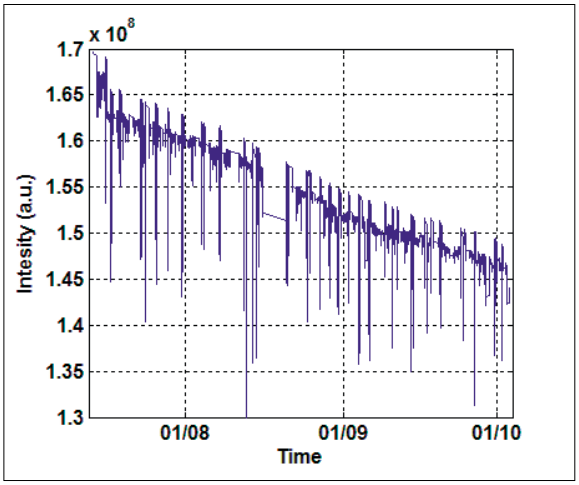


Figure 6. *Typical trend of decreasing intensity of light source over time. The negative spikes have no relevance for the measurement. They appear before and after the condensation, mainly due to temperature variation.*

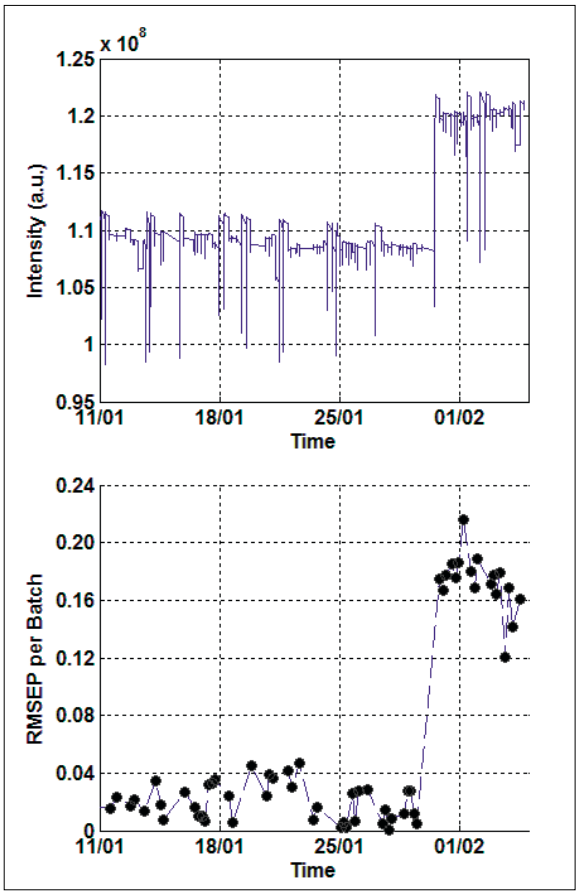


Figure 7. *Effect of light source replacement on the integral intensity of single beam spectra (top) and prediction accuracy of chemometric model (bottom).*

fects on model robustness can be very well covered. Even if the light intensity during validation is significantly lower than in calibration data. Intensity losses on a much shorter time scale were also observed, which were caused by undissolved particles blocked in the measuring gap of the in-line probe. In most cases of such events the models are vulnerable. Typically, there occurs an offset in the prediction. The same effect is observed after the change of the light source during annual maintenance (Figure 7). This can be corrected very quickly by using a simple offset correction, but for the longer term it is needed to recalibrate the model. This can be done later after a sufficient amount of new spectral data and related reference values were acquired. Very high sensitivity towards changes in the light intensity is related to low signal-to-noise ratio and small changes of the spectral features due to degree of condensation. In addition, permanently changing interference effects arising from the optical multiplexer when switching between measurement channels reduce the accuracy of the model. Although the latter effect has a rather low influence on model accuracy, it is difficult to correct.

Influence of Temperature

Temperature is one of the main factors influencing spectral data. The MF-resin condensation process is characterized by a well-defined temperature profile that has to be followed to achieve the desired product parameters. In the investigated process two different temperature profiles were applied – a standard profile and a customized profile (Figure 8). Figure 8During the time of model evaluation, it was confirmed that spectra as well as model predictions for a certain batch are highly dependent on the temperature profile used during the process. As a consequence a temperature correction method, that should compensate the difference between the profiles, was developed. A disadvantage associated with the variation in the temperature profile was a higher risk in end point determination of the condensation process. However, it was possible to minimize this risk due to the much tighter monitoring of the turbidity point by in-line FTNIR measurements.

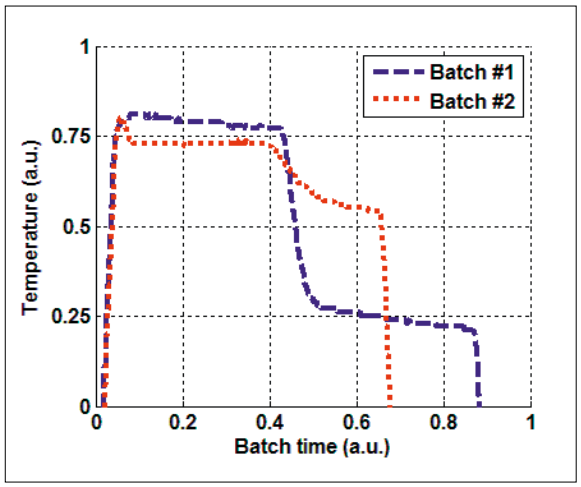


Figure 8. *Differences between temperature profiles of MF resin production.*

Conclusions

We presented various aspects of developing and implementing FTNIR measurements in combination with multivariate data modelling in a chemical batch process for MF resin production. Challenges, such as cleaning steps, decreasing light source intensity or temperature influence on the modelling results were identified and corrected. In conclusion it was proven that the developed chemometric models are capable of achieving satisfactory predictions of the turbidity point – a measure for the end point of the production process - over several months. The applied in-line method offers much tighter process control and reduces manual sampling steps which are particularly prone to errors. The operators, which are usually fully occupi-

ed with manual turbidity point determination, could perform other monitoring activities during condensation process, thereby enabling a better use of human resources. Furthermore, the method allowed to monitor the stop-condition automatically in an objective way without interaction of personnel.

Acknowledgements

This work was funded by the Austrian research funding association (FFG) under the scope of the COMET programme within the research project "Industrial Methods for Process Analytical Chemistry - From Measurement Technologies to Information Systems (imPACts)" (contract # 843546). Additionally we acknowledge support from the European Regional Development Fund (EFRE) in the framework of the EU-program REGIO 13, and the Federal State of Upper Austria.

References

1. W. H. Binder, M. Dunky, Encyclopedia of Polymer Science and Technology (John Wiley & Sons, Inc., Hoboken, NJ, USA, 2002).
2. S. Jahromi, Storage stability of melamine-formaldehyde resin solutions, 1. The mechanism of instability. Macromol. Chem. Phys. 200, 2230–2239 (1999).
3. J. Mijatovic, W. H. Binder, F. Kubel, W. Kantner, Studies on the stability of MF resin solutions: investigations on network formation. Macromol. Symp. 181, 373–382 (2002).
4. Y. Roggo et al., A review of near infrared spectroscopy and chemometrics in pharmaceutical technologies. A Rev. near infrared Spectrosc. Chemom. Pharm. Technol. 44, 683–700 (2007).
5. P. Hintenaus, G. Kvas, W. Märzinger, in IECON Proceedings (2007), pp. 2576–2579.
6. P. Hintenaus, W. Märzinger, H. Pöll, P. M. Hintenaus, M. Wolfgang, "An Infrared Spectrometer for Process Monitoring II , Chemometry and Automatization" (2010).

The Reactions of Cellulose and Hemicellulose Degradation Products in the Viscose Fibre Spin Bath

Georg Mayr[1*], Franz Zeppetzauer[1], Thomas Zweckmair[2], Daniela Bauer[3], Sabine Hild[4], Antje Potthast[2], Thomas Rosenau[2] and Thomas Röder[5]

[1] Competence Centre of Wood Composites and Wood Chemistry K-Plus, Altenberger Str. 69, 4040 Linz, Austria
[2] University of Natural Resources Department of Chemistry, Muthgasse 18, 1190 Vienna, Austria
[3] Kelheim Fibres GmbH, Regensburger Straße 109, 93309 Kelheim, Germany
[4] Johannes Kepler University Linz, Institute of Polymer Science, Altenberger Str. 69, 4040 Linz, Austria
[5] Lenzing AG, Werkstraße 2, 4860 Lenzing, Austria
[*] contact: g.mayr@kplus-wood.at

Abstract

In viscose fibre production degradation reactions of cellulose and hemicelluloses do occur. Degradation products from the alkaline degradation steps, i.e. steeping and pre-ageing, are apparently diverted into the acidic spin bath, where they undergo further reactions. Although the primary degradation products are well characterised, further reactions of these degradation products are hardly investigated. We tested the reactivity of degradation products under conditions, which prevail in the acidic spinning media of the viscose process. The studied chemical species were xylose, glucose, isosaccharinic acid, lactic acid and glycolic acid. We found strong indications that intermediate products, such as furfural and derivatives thereof, are formed, which readily polymerize and yield solid deposits. We showed that xylose, glucose and isosaccharinic acid trigger the subsequent formation of precipitants. Among the model compounds studied, xylose formed solid deposits most rapidly. The promoting effect of hydrogen sulfide on the formation of deposits was verified as well. Suspended solid particles act as crystallization points.

Keywords: *Viscose, deposits, cellulose degradation, glucose, xylose, isosaccharinic acid, lactic acid, glycolic acid, furfural*

Introduction

In viscose fibre production cellulose is subject to a sequence of chemical conversion processes. Initially the cellulose is steeped with sodium hydroxide lye, which activates cellulose and affords the so-called alkali cellulose. The alkali cellulose is pre-aged with the aim to adjust the molar mass of the cellulose. The aged alkali cellulose is then xanthogenated by reaction with carbon disulfide. The xanthogenate groups facilitate the dissolution of the cellulose in caustic lye. A viscous mass is obtained, the so called viscose dope. This dope is spun into the acidic spin bath. The xanthogenate groups are decomposed, carbon disulfide is released and the cellulose precipitated in the form of viscose fibres. Then the fibre is thoroughly washed, cut, after-treated and dried.

These harsh chemical conditions cellulose is exposed to cause some degradation of the polysaccharidic material. The degradation reactions which occur under alkaline conditions are well known[1-5]. Degradation reactions in acidic media are reported in literature as well[6-8]. However, the chemical behaviour of the degradation products in the viscose process have hardly been an issue up to now although the degradation products accumulate in the closed and recycled process streams having an important impact on the viscose fiber product quality.

Degradation reactions occur mostly during alkaline steeping and the subsequent pre-ageing of the alkali cellulose. In principle three different cellulose degradation mechanisms can be discerned under alkaline conditions[9,10]:

- End-wise degradation (peeling)
- Alkaline chain scission
- Aerobic (oxidative) alkaline degradation

These mechanisms can be differentiated by the conditions, under which they occur, and the resulting degradation products. The end-wise degradation (peeling) (Figure 1) does not involve air (oxygen) and typically occurs at moderate temperatures. In the peeling reaction, the anhydroglucose unit is split off (peeled off) one by one starting from the reducing sugar end. Initially in the open-ring structure the carbonyl group moves from position 1 to position 2 of the glucose unit in an Alberda-van-Eckenstein-Lobry-de-Bruin transformation (EBT)[11]. After deprotonation of the enol hydroxyl group an alkoxy-group can be easily eliminated from the beta-position of the enolate. The polysaccharide chain is eliminated from the C4-atom by this beta-alkoxy elimination (see Figure 1). A vicinal dicarbonyl compound, 4-deoxy-D-glycero-2,3-hexodiulose[9, 10] is yielded.

Under alkaline conditions the obtained dicarbonyl compound undergoes further reactions. Most important is a benzylic acid rearrangement yielding iso- and meta-saccharinic acid[9, 10]. The meta-saccharinic acid is obtained via a benzylic acid rearrangement of 4-deoxy-D-glycero-1,2-hexodiulose, which is formed from 4-deoxy-D-glycero-2,3-hexodiulose by an EBT. The formation of glucoisosaccharinic acid is depicted in Figure 2. A retro-aldol reaction yields products of lower mass (Figure 3)[9, 12]. A number of degradation products was found, and not all formation mechanisms of them are clear. Beside meta- and iso-saccharinic acids, lactic acid and 2,3-dihydroxypropanoic acid are typical degradation products[9].

The alkaline scission occurs only at elevated temperatures, above 170°C[9, 10]. It is a hydrolytic scission of glycosidic bonds, catalysed by high hydroxyl ion concentrations. It requires conformational changes from the $^4C_1$ into the $^1C_4$ chair conformation, which is only possible at higher temperatures[10]. In principle every glycosidic bond in the polysaccharide chain can be cleaved this way. As a consequence, fragments of larger molecular mass are formed, from oligomeric to polymeric size.

The aerobic (oxidative) alkaline degradation mechanism takes place in the presence of air and oxygen. By oxidation, a carbonyl group is formed on an anhydroglucose-unit, which in turn immediately leads to chain scission under alkaline conditions. This happens again by beta-alkoxy elimination (compare Figure 1), which in contrast to the endwise peeling might occur anywhere along the chain, wherever an oxidized groups has been introduced before. As a result fragments of oligomeric to polymeric character are formed.

The degradation mechanisms described are not only affecting cellulose but also hemicellulose, for instance xylan from hardwood.

The acidic degradation of cellulose proceeds accor-

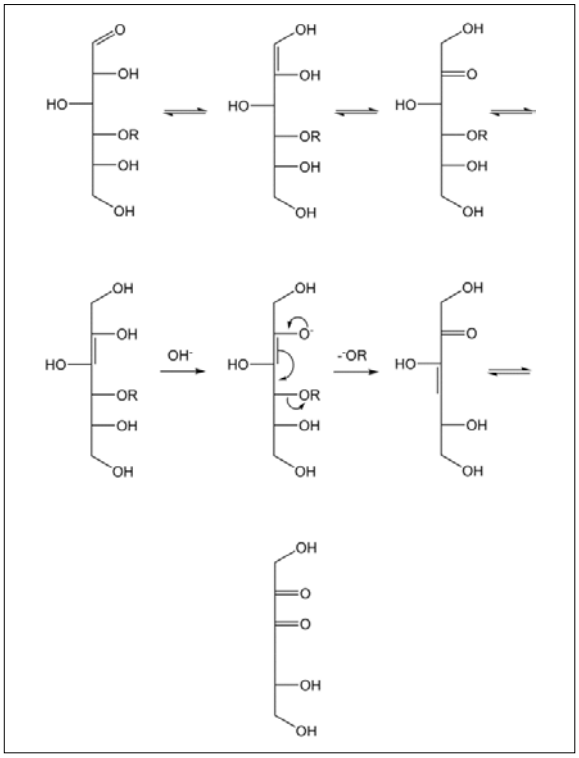


Figure 1. *The splitting off of a reducing sugar by the Peeling reaction. Via a beta-alkoxy elimination the 4-deoxy-D-glycero-2,3-hexodiulose is formed.*

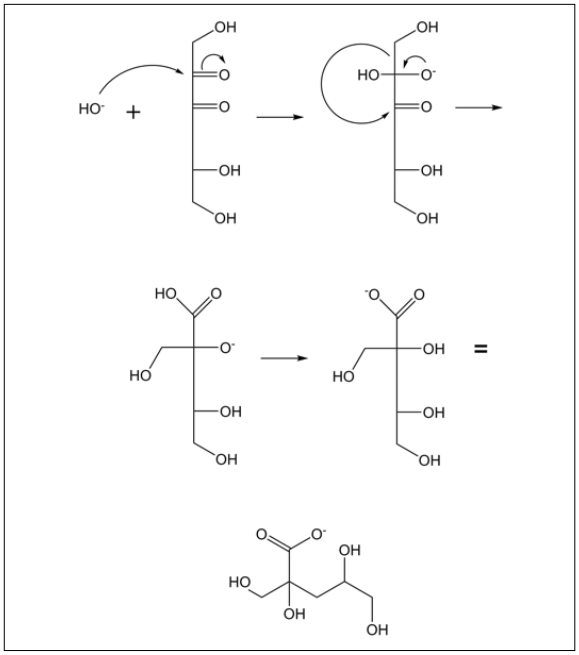


Figure 2. *Reaction of 4-deoxy-D-glycero-2,3-hexodiulose to glucoisosaccharinic acid via benzylic acid rearrangement (BAR). Glucometasaccharinic is generated from 4-deoxy-D-glycero-1,2-hexodiulose by an analogous benzylic acid rearrangement.*

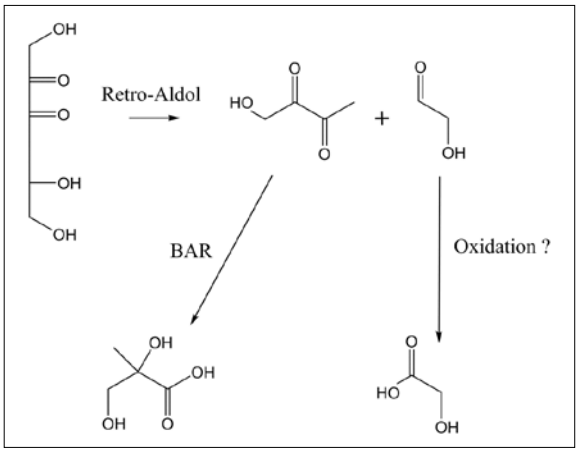


Figure 3. *Decomposition of 4-deoxy-D-glycero-2,3-hexodiulose (on the left-top) via Retro-Aldol reaction. BAR…Benzylic acid rearrangement.*

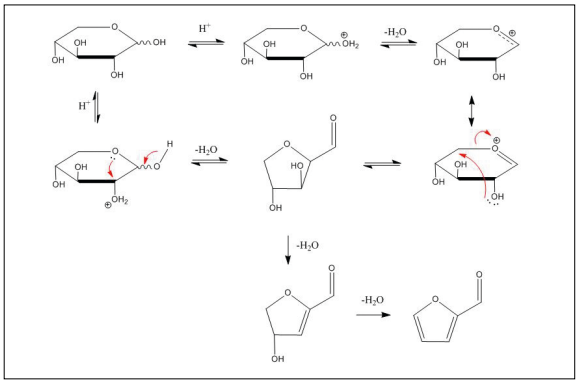


Figure 4. *Reaction mechanism for the formation of furfural from xylose by proton catalysed dehydration.*[14]

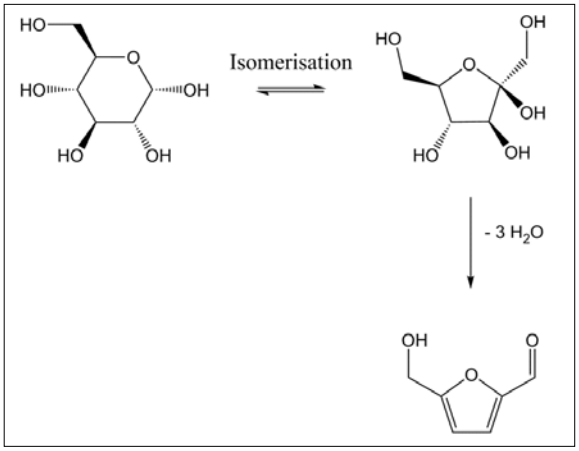


Figure 5. *Formation of 4-hydroxymethylfurfural from glucose via isomerisation to fructose.*[15]

ding to a typical hydrolytic reaction mechanism. At first the oxygen atom of the glycosidic bond is protonated. The bond to the adjacent C1 atom breaks, forming a carbenium ion which is attacked by a water molecule in the next step. After the release of a proton a neutral glucose unit and a neutral residual cellulose chain is formed[12]. Monomeric glucose units are predominantly formed. Oligomeric fragments scarcely emerge due to the kinetics of bond scission[13]. However, it is not clear, which role acidic degradation plays in the viscose process, as the cellulose and the fibre, respectively, remain only a few seconds in the acidic spin bath during the spinning step. The extent of degradation depends, besides the chemical matrix (pH, oxygen access) and the temperature also on the supermolecular structure of cellulose. The accessibility of the cellulose chains with regard to degrading agents, for instance the OH- anion, is essential. Thus the packing and crystallinity of cellulose plays a key role for the extent of degradation[9]. The more accessible the cellulose packing is for the degrading reagent, the faster the degradation is.

The alkaline degradation mechanisms play a role – as mentioned – predominantly during the steeping of pulp, the aging of alkali cellulose and within the corresponding lyes. The steeped pulp is pressed to yield the alkali cellulose. The lye, separated from the pulp slurry, is recycled. One part is used for repeated steeping, another part of this lye is used for the dissolution of the xanthogenated cellulose, which yields the viscose dope. Since the viscose dope is spun into the spinning bath, the alkaline degradation products end up in the acidic spinning bath. This change of the chemical properties of the matrix is able to induce chemical reactions and conversions of the degradation products.

It is well known that sugars dehydrate to furan derivatives under acidic conditions at higher temperatures. Xylose reacts to furfural[14] (Figure 4) whereas glucose is converted to hydroxymethylfurfural[15] (Figure 5). Both reactions might occur in the acidic spin bath of the viscose process. The monomeric sugars can be formed from acidic degradation reactions of hemicellulose and cellulose fragments, which are soluble in the acidic spin bath. Polysaccharides, which are not soluble, cannot act as source. They reside in the spin bath only for a few seconds during the spinning step. However, fragments from polymeric polysaccharides, which are formed for instance by alkaline oxidations during ageing, might be of such low molecular mass that they dissolve in the acidic spinning bath. Due to recycling of the spin bath longer residence times result, upon which those soluble oligomers are likely to be hydrolysed to monomers.

It is well-established that furfural and its derivatives readily polymerize to insoluble molecular species[16-18], which precipitate and form deposits. The occurrence of deposits in the spin bath circuits of industrial facilities can be observed indeed. These deposits are objectionable for production since in the staple fibre they are detrimental for product quality. The formation of furfural and its derivatives from cellulose degradation products is one possible reason for the formation of those depo-

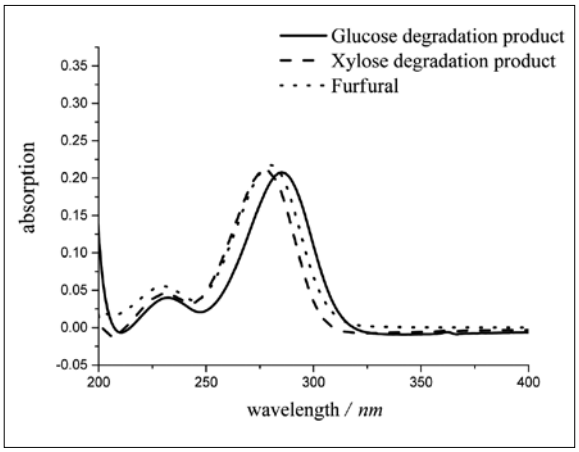


Figure 6. *Comparison of UV-spectra of degradation products of xylose and glucose with the spectrum of furfural.*

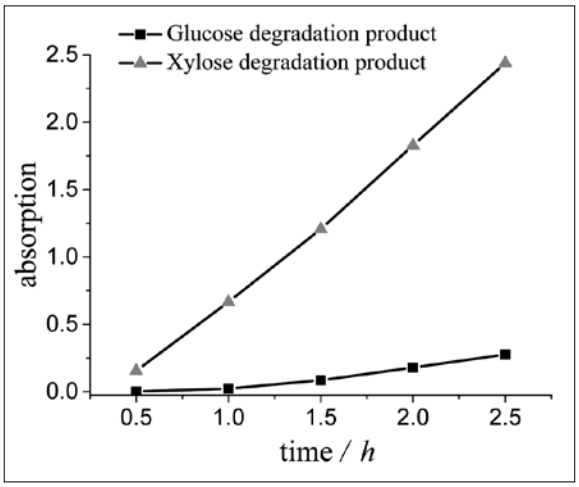


Figure 7. *Increase of the UV-absorption band at ~280 nm after different times at the elevated temperature of 90°C in synthetic viscose spin bath.*

sits. It is unknown, which degradation products form deposits. Furthermore it is not clear, if degradation products of cellulose or hemicellulose can form precipitates under the conditions prevailing in the acidic spin bath. In deliberate furfural production, for instance, temperatures above 150°C are used[14, 16]; temperatures which are never reached in the viscose spin bath. In this paper we present results, which indicate that the formation of furfural and its derivatives plays indeed a key role for the occurrence of deposits. We further limited the bandwidth of degradation products leading to precipitates. As product quality builds the centre of interest, a number of factors i.e. the influence of hydrogen sulfide and suspended solids, were studied with regard to deposit formation in the viscose process.

Materials and Methods

In the first step we checked, if and in particular which hemicellulose degradation products form solid contaminations under industrial process conditions. Therefore the conditions in the acidic viscose spin bath have been simulated in the laboratory. A solution, analogous to the spin bath, has been heated for a defined time. 100 mL of a laboratory prepared spin bath, i.e. an aqueous solution containing H_2SO_4, $ZnSO_4$ and Na_2SO_4, was heated to 90°C. The chemical species xylose, glucose, glucoisosaccharinic acid, lactic acid and glycolic acid, which are known to be cellulose or hemicellulose degradation products[9], were added at defined concentrations. If not stated otherwise, 500 mg were dissolved in 100 mL, resulting in a concentration of 5 g/L. Samples were taken for analysis by UV-Vis spectroscopy with an Agilent 5483 instrument. After heating at the elevated temperature of 90°C for 72 h the solid precipitants formed were separated, sucked off via a suction filter, dried and weighted.

Results and Discussion

During the thermal impact experiments with addition of xylose, glucose and glucoisosaccharinic acid to a simulated spinning bath the colour of the reaction solution turned yellow. In case of lactic acid and glycolic acid no such colouring was observed. We noticed a distinct absorption band around 280 nm. For xylose the absorption band was exactly located at 278 nm and for glucose at 285 nm. In the visible region a broad, undefined absorption over the whole wavelength area evolved. The colouring of the reaction solution as well as the intensity of the band increased with continued heating. Therefore we assigned the band around 280 nm to a primary degradation product. Furfural absorbs at 280 nm as well. This indicates that the first degradation product of xylose, which is formed at a temperature of 90°C, is furfural, which is in accordance with the hypothesis derived from literature[14, 16, 19-21]. The slight difference of 2 nm between the absorption maximum of the xylose conversion product and the absorption maximum of furfural might be caused by side products. The absorption of the glucose degradation product at 285 nm is probably mainly derived from 4-hydroxymethylfurfural. Figure 6 shows absorption bands of the glucose degradation product, the xylose degradation product and furfural in comparison.

Furthermore the progress of the formation of furfural or 4-hydroxymethylfurfural was tracked by the increase of the absorption intensity with time. The increase of this band (~280 nm) within the first 2.5 hours of the reaction is depicted in Figure 7. It is obvious, that the increase for xylose is much faster than for glucose. The analogous spectroscopic investigations for glucoisosaccharinic acid will be performed in a subsequent study.

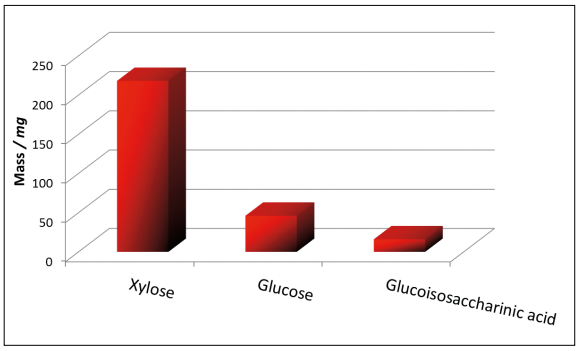


Figure 8. *Mass of precipitates formed after the heating to 90° C for 72 h. Data are presented for xylose, glucose and glucoisosaccharinic acid. The experiments were conducted with equal concentrations and reaction volumes. The experiments were conducted with equal concentrations and reaction volumes.*

The thermal impact experiments applying xylose, glucose and glucoisosaccharinic acid resulted in the formation of solid dark precipitants. In case of lactic acid and glycolic acid no formation of solids was detected by thermal impact experiments. This is in agreement with the hypothesis, that the solid deposits are formed via furfural and its derivatives. The increase of the UV-band at ~280 nm indicates that xylose degrades faster than glucose. In connection with the hypothesis, that the solids are formed by polymerization reactions from furfural (and its derivatives) it could be hypothesized, that xylose also forms solid deposits faster than glucose. In order to verify this hypothesis we carried out thermal impact experiments with an initial concentration of 20 g/L of xylose, glucose or glucoisosaccharinic acid and a total volume of 100 mL. After heating to 90°C for 72 h we measured 218 mg deposits for xylose, 46 mg for glucose and 16 mg for glucoisosaccharinic acid (Figure 8). This supports our hypothesis that xylose degrades and forms solid deposits faster than glucose and glucoisosaccharinic acid. We assume that the initial isomerization step required for furan formation is faster in the case of xylose as compared to glucose.

Besides the chemical degradation of polysaccharides discussed so far, also hydrogen disulfide present in the spin bath of the industrial process might play an important role. Hydrogen disulfide is released via a side reaction during the de-xanthogenation in the spinning step. Considering furfural formation as formulated in Figure 4 and the nucleophilic character of the sulfide ion the question arises, if the sulfide ion could participate in the reaction. In order to answer that question we carried out thermal impact experiments with xylose as test species (5 g/L or 33.3 mmol/L, 100 mL batch) with and without the addition of hydrogen sulfide. 0.64 mol of gaseous H_2S were discharged into the reaction mixture (of 100 mL) during the first two hours. The experiment without hydrogen disulfide yielded 80 mg solid deposits after 72 hours whereas the experiment with hydrogen disulfide yielded 118 mg. This represents a significant increase by 38 mg and proves that hydrogen disulfide promotes the precipitate formation. The participation of sulfide ions in the reactions of the organic species is only one possibility, the sulfide could affect the formation of solids. Raman spectra of the formed precipitate proved, that elemental sulphur is formed under the test conditions as well.

In a further set of experiments we hypothesized if the presence of solids accelerates the formation of solid precipitants. In order to verify this hypothesis, we performed three separate thermal impact experiments, with xylose as source material, and with the addition of 100 mg of fine particles of silicon dioxide and activated carbon in two separated experiments. In order to determine the mass of the formed precipitates, from the mass of total solids filtered off after the thermal impact experiment the mass of initially added particle material was subtracted. The yield of solid precipitates of 80 mg was increased to 95 mg by silicon dioxide and to 131 mg by activated carbon. Since an increase is obtained with both, very different particle materials, we concluded that particles in the solution somehow act as seed nuclei. The distinctly higher increase by activated carbon probably is owed to its high effective surface area and surface structure. The determined mass of precipitates in the case of addition of activated carbon as well might be increased by material, absorbed by the activated carbon.

Conclusions

The concept of thermal impact experiments in viscose spin bathes proved to be a valuable tool for the investigation of reactions of dissolved cellulose degradation products under process conditions. The situation in the acidic spin bath of industrial production was simulated. We tested the degradation products xylose, glucose, glucoisosaccharinic acid, lactic acid and glycolic acid for their reactivity in acidic media. UV spectra of initially soluble degradation products of xylose and glucose supported the hypothesis that furfural or a derivative thereof is formed. Extended reaction time of 72 h at 90°C lead to solid precipitates from xylose, glucose and glucoisoccharinic acid. Upon applying xylose the highest amount of precipitates was formed. Lactic acid and glycolic acid do not undergo any further reaction. This is in accordance with the hypothesis that further reactions proceed via furfural (and its derivatives).

We were also able to prove, that hydrogen sulfide promotes the formation of solid deposits. This is of relevance, as hydrogen sulfide is released by a side reaction during the spinning step of the viscose process. Solids, suspended in an acidic reaction medium, promote the formation of solid contaminants as well.

Acknowledgements

This work was funded by the Austrian research funding association (FFG) under the scope of the COMET programme within the research project “Industrial Methods for Process Analytical Chemistry - From Measurement Technologies to Information Systems (imPACts)” (contract # 843546). This programme is promoted by BMVIT, BMWFW, the federal state of Upper and Lower Austria. We also thank Kelheim Fibres GmbH and Lenzing AG for their in-kind contributions.

References

[1] D. Mozdyniewicz and H. Sixta, "Carbohydrated degradation reactions during alkaline steeping of dissolving pulp - infuelnce of air exclusion", Lenzinger Berichte 90, 103 (2012).

[2] G. Kraft, M. Kraft, F. Zeppetzauer, H. Weber, D. Mozdyniewicz and T. Röder, "Cellulose degradation during viscose processing", Lenzinger Berichte 91, (2013).

[3] P. Barthel and B. Philipp, "Untersuchungen zum Abbauverlauf bei der Alkalicellulose-Vorreife verschiedenartier Zellstoffe", Faserforschung und Textiltechnik 18, 266 (1967).

[4] M. A. Glaus and L. R. V. Loon, "Degradation of Cellulose under Alkaline Conditions", Environmental Science & Technology 42, 2906 (2008).

[5] J. Sartori, A. Potthast, A. Ecker, H. Sixta and P. Kosma, "Alkaline degradation kinetics and CE-separation of cello- uand xylooligomers. Part I", Carbohydrate Research 338 (11), 1209 (2003).

[6] D. Klemm, B. Philipp, T. Heinze, U. Heinze and W. Wagenknecht, Comprehensive Cellulose Chemistry - Fundamentals and Analytical Methods. p. 85, WILEY VCH Verlag, Weinheim, (1998).

[7] J. Gütsch, T. Nousiainen and H. Sixta, "Comparative evaluation of autohydrolysis and acid-catalyzed hydrolysis of Eucalyptus globulus wood", Bioresource Technology 109, 77 (2012).

[8] O. Battista, S. Coppicic, J. Howsmoon, F. Morehead and W. Sisson, "Level-Off Degree of Polymerization", Industrial & Engineering Chemistry 48 (2), 333 (156).

[9] C. J. Knill and J. F. Kennedy, "Degradation of cellulose under alkaline conditions", Carbohydrate Polymers 51, 281 (2003).

[10] A. Potthast, "Chemistry of Kraft Cooking" In Handbook of pulp, Ed by H. Sixta. Wiley-VCH, Weinheim, p. 175 (2006).

[11] Y. J. Topper and S. D. W. Jr, "The alkali-catalyzed conversion of glucose into fructose and mannose", Journal of Biological Chemistry 189, 191 (1951).

[12] T. P. Nevell, "Degradation of cellulose by acids, alkalis and mechanical means", Cellulose chemistry and its applications, 223 (1985).

[13] G. Joksimovic and Z. Markovic, "Investigation of Mechansim of Acidic Hydrolysis of Cellulose", Acta Agriculturae Serbica 12 (24), 51 (2007).

[14] M. J. Antal, T. Leesomboon and W. S. Mok, "Mechanism of formation of 2-furaldehyde from D-xylose", Carbohydrate Research 217, 71 (1991).

[15] R. J. H. G. Joseph Zakzeski, Arjan T. Smit, and Bert M. Weckhuysen, "Solid Acid-Catalyzed Cellulose Hydrolysis", Chemsuschem 5, 430 (2012).

[16] K. J. Zeitsch, The chemistry and technology of furfural and its many by-products. p. 3, 17, 19, 28, Amsterdam, Elsevier, (2000).

[17] R. Sanchez, J. Rieumont and M. I. B. Tavares, "Structural Characterization of the Polymeric Lattice of Furfuraldehyde Resins", Polymer Testing 17, 398 (1998).

[18] A. Gandini and M. N. Belgacem, "Furan derivatives and furan chemistry at the service or macromolecular materials" In Monomers, Polymers and Composites from Renewable Resources, Ed by M. N. Belgacem. Elsevier, Oxford, Amsterdam, p. 123 (2008).

[19] A. P. Dunlop, "Furfural formation and behavior", Industrial and Engineering Chemistry 40 (2), (1948).

[20] A. Takagaki, M. Ohara, S. Nishimura and K. Ebitani, "One-pot Formation of Furfural from Xylose via Isomerization and Successive Dehydration Reactions over Heterogeneous Acid and Base Catalysts", Chemical Letters 39, (2010).

[21] C. K. Lin, "Furfural formation by acid hydrolysis of the pentosan containing materials", PhD-thesis, School of Forest Resources, NC State University (1977).