

Monograph

NIR Spectroscopy



NIR Spectroscopy

A guide to near-infrared spectroscopic analysis of industrial manufacturing processes

A guide to near-infrared spectroscopic analysis of industrial manufacturing processes

Table of contents

Table of contents	3
1 Preface.....	5
2 Theory of vibrational spectroscopy.....	7
3 NIRS, MIR, and Raman	11
4 Chemometrics	13
4.1 Data analysis and chemometrics.....	13
4.2 Identification, calibration, and validation of NIR data	13
4.3 Data pretreatment	14
4.4 Reducing the quantity of NIR spectral data	14
4.5 Multivariate calibration methods and validation.....	15
5 NIR spectral region.....	17
6 NIR analysis methods	21
6.1 Qualitative analysis	22
6.2 Quantitative analysis.....	23
7 NIR instrumentation	25
7.1 Laboratory analyzers	25
7.2 Process analyzers	25
8 NIR analysis strategies	27
8.1 Laboratory analysis.....	27
8.2 Atline analysis.....	27
8.3 Online analysis.....	28
8.4 Inline analysis	28
8.5 Data reporting	28
8.6 Maintenance	29
9 Industry sectors and applications	31
10 Standards	35
11 Glossary	37
12 References.....	41

1 Preface

Near-infrared spectroscopy (NIRS) is a simple, quick (< 30 s analysis time), nondestructive technique that provides multiconstituent analysis on virtually any matrix with levels of accuracy and precision that are comparable to primary reference methods. As another great benefit, near-infrared (NIR) analyses require no sample preparation or manipulation with hazardous chemicals, solvents, or reagents: NIRS thus belongs to the reagent-free methodologies, for which reason it is an exemplary analytical contribution to the Environmental Protection Agency's (EPA) green chemistry attempt. The recorded NIR spectra contain a variety of chemical and physical (e.g., particle size) information on the sample and its constituents.

It should be noted, however, that NIR absorption spectra are often complex and normally possess broad overlapping NIR absorption bands that require special mathematical procedures for data analysis. NIR identification and NIR qualification (which are a prerequisite for a quantitative analysis) can be done by comparing a sample spectrum to reference spectra of known materials. Quantification is done by using mathematical models and so-called multivariate data analysis (chemometrics). As the calibration and validation of measured NIR spectral data is correlated through statistical methods to reference data, NIR spectroscopy is a secondary analytical method.

NIR analysis was first used by the United States Department of Agriculture (USDA) in the mid-1960s to detect the internal qualities of apple crops blighted with a devastating condition called «water core». This original application of NIR technology could even predict fruit maturity levels and sugar content, and give indirect measures of taste and texture. From these beginnings, NIR spectroscopy was launched as a popular method to perform rapid, nondestructive analyses on a wide variety of matrices across the agriculture and food industries.

In the pharmaceutical and chemical industries, NIR analysis has been successfully implemented for more than 25 years. Initial applications focused on raw material testing. More recently, attention has turned to analyzing solid and liquid chemical formulations for product quality and in-process monitoring of manufacturing operations. For material inspection, samples are scanned as received and the identity and quality of a material is confirmed using pattern recognition algorithms. As an in-process test and measurement tool, NIR spectroscopy (in conjunction with statistical regression techniques) provides near real-time chemical information for control of chemical production processes and solvent recovery systems as well as drying, blending, and extrusion operations.

Near-infrared spectroscopy has great potential for improving the monitoring and control of industrial processes. To ensure that the most appropriate analysis strategy is implemented and robust NIR methods are developed, the optical characteristics of the sample, the analyte's sensitivity and selectivity, and the production and control requirements must be considered.

This monograph provides a practical overview of the critical factors that are useful and necessary when developing and implementing NIR spectroscopic methods in pharmaceutical and chemical manufacturing operations.

2 Theory of vibrational spectroscopy

Names as «overtone vibrational spectroscopy» and «anharmonic vibrational spectroscopy» illustrate that the understanding of the NIRS theory is by far not easy. Although it is beyond the scope of this monograph to provide an overview of the theoretical background of vibrational spectroscopy, it cannot do without delving into some basic principles that facilitate spectra interpretation.

Interaction of light and physical matter

Spectroscopy uses light – electromagnetic radiation – to analyze materials by describing the energy transfer between light and matter. The energy of a single photon – the light particle – is defined as

$$E = h \cdot \nu = \frac{h \cdot c}{\lambda} = h \cdot c \cdot \tilde{\nu} \quad (\text{Eq. 1})$$

h – Planck constant in $\text{m}^2 \cdot \text{kg} / \text{s}$
 ν – frequency of the light in s^{-1}
 c – speed of light in m / s
 λ – wavelength of the light in m
 $\tilde{\nu}$ – wavenumber of the light in m^{-1}

with $\tilde{\nu} = \frac{1}{\lambda}$

$$\tilde{\nu} [\text{cm}^{-1}] = \frac{10^7}{\lambda [\text{nm}]}$$

Generally, wavenumbers are indicated in cm^{-1} . To convert the wavelength in wavenumbers, the reciprocal of the wavelength in nm is multiplied by 10^7 . Eq. 1 shows that the shorter the wavelength, the higher the wavenumber and the higher the energy of the photon.

The electromagnetic spectrum is divided into several regions. Every region stands for a specific kind of molecular or atomic transition and thus for a specific spectroscopic technique. Gamma rays and X-rays have wavelengths of only a few nm and are very harmful, because they break chemical bonds and ionize molecules. UV radiation covers the region 190–350 nm and visible radiation (VIS) the region 350–780 nm . Photon absorption in the UV/VIS range transfers electrons from low-energy to high energy molecule orbitals and returning to the ground state easily proceeds via fluorescence. Radiation referred to as near-infrared defines the range between 780 and 2500 nm . Molecules can absorb infrared light without later reemission by exciting certain vibrational frequencies. The sample absorbs the frequencies of polychromatic light that corresponds to its molecular vibrational transitions.

Harmonic oscillator

These molecular vibrations can be described using the classical physics model of the harmonic diatomic oscillator as the simplest vibrating system. Two vibrating masses connected with a spring with a given force constant lead to changes in the internuclear distance. Combining Hook's law with Newton's force law, the vibrational frequency corresponds to:

$$\nu_{vib} = \frac{1}{2\pi} \sqrt{\frac{k}{\mu}} \quad (\text{Eq. 2})$$

ν – bond frequency of absorption band
 k – force constant of the bond
 μ – reduced mass of the bonded atoms

The frequency of the vibration and thus the energy are, amongst others, a function of the bond strength. Vibrational frequencies are thus the key that contains information about the structure of the investigated compound. This is the basis for the application of vibrational spectroscopy.

According to Hook's law, the potential energy V of the harmonic oscillator is a quadratic function of the displacement of the vibrating atoms. The parabolic function is symmetrical about the equilibrium bond length r_e .

$$V = \frac{1}{2}k(r - r_e)^2 = \frac{1}{2}kx^2 \quad (\text{Eq. 3})$$

V – potential energy
 r – internuclear distance
 r_e – internuclear distance at equilibrium
 x – displacement of vibrating atoms

Quantum mechanical considerations using the Schrödinger equation show that the vibrational energy for the harmonic oscillator has no continuum for vibrational energy levels, but only certain discrete energy levels.

$$E_{vib} = hv \left(v + \frac{1}{2} \right) \quad (\text{Eq. 4})$$

v – vibrational quantum number (0, 1, 2, ...)

The vibrational quantum number can only have integer values 0, 1, 2, 3... With Eq. 2 the vibrational energy is:

$$E_{vib} = \left(v + \frac{1}{2} \right) \frac{h}{2\pi} \sqrt{\frac{k}{\mu}} \quad (\text{Eq. 5})$$

According to Eq. 5 only transitions between neighboring and equidistant energy levels are allowed in the harmonic oscillator.

The distribution of the molecules among the allowed vibrational levels follows the Boltzmann distribution. At room temperature most molecules reside at the ground vibrational level $v = 0$. Hence, the first allowed transition to the energy level $v = 1$, also called the fundamental transition, dominates the spectrum. The further allowed transitions derive from vibrationally excited levels and the associated bands are much weaker.

Precondition for the absorption of a light photon is that the frequency of the light photon equals the energy difference between two vibrational states of the bond. But that's not all that is required for energy absorption.

Energy absorption according to the resonance theory

According to the resonance theory, a way has to be found to successfully transfer the energy to the molecule (e.g., by exerting physical contact or, in case of magnetic materials, by applying a magnetic field). For molecular vibrations, this transfer occurs via the molecular polarity, that is, that interaction of infrared radiation with a vibrating molecule only occurs when the vibration is accompanied by a change of the dipole moment μ .

$$\frac{\delta\mu}{\partial x} \neq 0 \quad (\text{Eq. 6})$$

The degree of this dipole moment change within the vibrational transition determines to a strong extent the intensity of light absorption. According to Eq. 6, only heteronuclear, diatomic molecules show (vibrational-spectral) transitions between the photons of the light and the molecule vibrations. The hydrogen atom is the lightest atom. Therefore, bonds with hydrogen exhibit the largest vibrations (C–H, N–H, O–H, and S–H).

Most molecules have more than two atoms which can perform more than one type of vibration per bond (stretching, bending). Furthermore, not all bonds vibrate independently of each other; many real vibrations are coupled.

Anharmonic oscillator

The potential energy curve of the harmonic oscillator (Eq. 3) is parabolic and only transitions between consecutive energy levels that additionally cause a change in dipole moment are possible. According to this model, infinite energies can be stored in the molecule without bond breaking. Experience teaches, however, that all molecules dissociate, when only enough energy is applied and the vibrating bond is extended. Moreover, strong repelling forces are observed, when the atoms are pressed together. Due to this mechanical so-called anharmonicity – bond dissociation and Coulombic repulsion –, the classical ball and spring model has to be modified. Therefore, Eq. 3 which describes the potential energy as a quadratic function of the displacement is extended by adding higher-order terms of displacement:

$$V = \frac{1}{2}kx^2 + \frac{1}{2}k_ax^3 + k_bx^3 + \dots \quad (\text{Eq. 7}) \quad \text{with } k_a, k_b \ll k$$

With the corrections in Eq. 7, the Schrödinger equation provides the following term for the vibrational energy of the molecule

$$E_{vib} = hv \left(v + \frac{1}{2} \right) - y \left(v + \frac{1}{2} \right)^2 \quad (\text{Eq. 8}) \quad y - \text{anharmonicity constant}$$

The anharmonicity now allows transitions between noncontiguous energy states were $\Delta v = \pm 2, 3, \dots$. Additionally, the vibrational energy levels are no longer equidistant and energy differences decrease with increasing quantum number v . Instead of the parabolic shape, the potential energy curve is represented by an asymmetric Morse function. Besides this mechanical anharmonicity, the electronic anharmonicity should be mentioned. It is generally much smaller than the mechanical anharmonicity and it arises from the nonlinear relationship between dipole moment and atomic displacement.

Implications for NIRS

To sum up, the occurrence and the spectral properties of NIR absorption bands are – besides the dipole moment change – due to the large mechanical anharmonicity of the vibrating atoms. This is reflected in

- overtone transitions that correspond to quantum numbers greater than one and that appear as multiples of the fundamental vibrational frequency. Overtone transitions appear between 780 and 2000 nm. First and higher overtones are much less likely than the fundamental vibrational frequency. That's the reason why the bands are weaker.
- combination modes that appear in polyatomic molecules, where multiple vibrational modes interact. They are the sum of multiples of each interacting frequency. Their absorption bands emerge between 1900 and 2500 nm.
- the nonequidistant energy states of a vibration. That means that allowed transitions become smaller in energy.

3 NIRS, MIR, and Raman

Spectroscopic techniques excel by their possibility to gain rapid and accurate information from the high-resolution spectra of solid and liquid samples without prior (NIRS) or reduced (Raman) sample preparation. The advantages of NIRS, mid-infrared (MIR), and Raman are numerous (Table 1).

All three spectroscopic techniques are economic and facilitate qualitative and quantitative as well as noninvasive and nondestructive analysis. Further advantages include that they are reagent- and waste-free and require no additional auxiliary chemicals. For all these reasons, spectroscopic techniques are ideally suited for industrial quality control and process monitoring. For nearly every application, there is the right technique.

The most flexible technique, however, is NIRS. Especially the availability of efficient chemometric evaluation tools and software as well as light-fiber optics has made NIRS to an invaluable tool for academic research and industrial quality control. NIRS allows the determination of multiple values in a single determination. In the process environment, NIRS stands out by the possibility of on-site measurements and remote sampling – not least because the spectrometer can be placed far away from the sampling point.

Table 1: Some characteristics of NIR, MIR, and Raman spectroscopies

	Raman	MIR	NIR
Wavenumber	50–4000 cm ⁻¹	200–4000 cm ⁻¹	4000–12500 cm ⁻¹
Bonds	homonuclear bonds such as C–C, C=C, S–S	polar bonds such as C=O, C–O, C–F	H-containing bonds such as C–H, O–H, N–H, S–H
Absorption bands due to	scattered radiation	absorbed radiation (basic vibration)	absorbed radiation (overtones and combination)
Absorption	strong	weak	weak
Absorption bands	well-resolved, assignable to specific chemical groups	well-resolved, assignable to specific chemical groups	series of overlapping bands
Signal intensity	poor	good	good
Quantification	intensity (I) ~ concentration	log I ₀ /I ~ concentration (Lambert-Beer law)	log I ₀ /I ~ concentration (Lambert-Beer law)
Excitation conditions	change of polarizability α	change of dipole moment μ	change of dipole moment μ
Selectivity	high	high	low, requires calibration and chemometrics
Interference	broad fluorescence baseline	water	Water, physical attributes (e.g., sample size, shape, and hardness)
Particle size	independent	dependent	dependent
Applicability for atline, online, inline	good	poor	good
Radiation source	Monochromatic (laser VIS/NIR region)	Polychromatic by global tungsten	Polychromatic by global tungsten
Sample preparation	none	reduced (except ATR*)	none

ATR = attenuated total reflection

NIRS can analyze through some translucent packaging materials and is thus ideally suited for quality control of raw materials. The penetration depth also allows to analyze bulk material. Intermolecular hydrogen bondings and dipol interactions alter vibrational energy and thus lead to shifts in existing absorption bands or even the formation of new absorption bands. The fact that the spectra are strongly impacted by physical parameters (e.g., particle size, density, and moisture content) is on the other hand the reason, why NIRS is ideally suited for the capture of these characteristics.

NIRS – more than MIR and Raman – belong to indirect or secondary methods. To obtain information, their spectra have to be compared to primary or reference methods. Therefore, mathematical means are applied to the spectra to establish a NIR «predictive model» to prognosticate the composition of the investigated sample. The next chapter deals with this issue.

4 Chemometrics

This monograph gives an impression of some of the chemometric procedures on a basic level. If necessary, a deeper insight into chemometrics can be gained by studying the literature listed in chapter 12. Whether or not a specific technique can or should be applied depends on the particular problem and the software that is used for data analysis.

4.1 Data analysis and chemometrics

A good definition of chemometrics stems from Massart (2003): «Chemometrics is a chemical discipline that uses mathematics, statistics and formal logic (a) to design or select optimal experimental procedures; (b) to provide maximum relevant chemical information by analyzing chemical data; and (c) to obtain knowledge about chemical systems».

NIR absorption spectra are often complex and normally possess broad overlapping NIR absorption bands. Moreover, chemical, physical, and structural properties of all species present in a sample influence the measured spectra. Also small sample-to-sample differences of a sample series can cause very small spectral differences. This means that the NIR absorption data obtained (measured spectra) is depending on more than one variable simultaneously and thus is multivariate. Another way to put it is that more than one variable (absorption at a single wavelength) is recorded for a sample: a measured NIR spectra of a sample can be seen as a huge number of measurements that have to be performed (absorptions measured at hundreds of wavelengths) and thus this data is multivariate. Chemometrics uses mathematical and statistical procedures for multivariate data analysis to filter information that correlates to a certain property from a very big amount of data. In qualitative and quantitative NIR analysis, the relevant part from the multivariate NIR spectral data is extracted without losing important information and to get rid of unwanted information (e.g., interferences or noise). NIR models describe how the measured multivariate spectral features (e.g., absorption values of samples measured at many different wavelengths) are related to properties of the analytes (e.g., the concentrations of the analytes; serving as an example throughout this chapter 4).

These so-called variables have to be arranged into matrices to perform the mathematical and statistical procedures. There are two types of variables: dependent and independent variables. Independent variables are those to be measured and dependent variables are those to be predicted. The independent variables are the measured absorbances. For each sample, the corresponding spectrum (absorbances at all wavelengths) is placed into the absorbance matrix (X matrix) as a column vector. The concentration matrix (Y matrix) holds the dependent variables, namely the known concentrations of the samples. It is the concentration matrix that is predicted for unknown samples as will be shown further below in the text.

4.2 Identification, calibration, and validation of NIR data

A comprehensive NIR data analysis starts with the classification and identification of a measured sample followed by qualification and quantitation (see also chapter 6 of this monograph for further details). Quantitative and qualitative analysis methods require prior calibration and validation. Representative samples as well as accurate results from reference analytical methods (e.g., concentrations of analyzed samples) are needed to calibrate and validate NIR absorption data. When it comes to determine the water content using NIR spectroscopy, for example, Karl Fischer titration can be used as a reference analytical method. For identification, qualification, and quantification, the sufficient number of samples needed to cover normal variation between sample to sample depends on the complexity of the sample matrix and/or the interference caused by the matrix on the signal of the analyte of interest.

4.3 Data pretreatment

Nonhomogeneous distribution of particles in a sample, particle size differences, sample density variations, sample morphology differences (shape and roughness of sample surface), ... lead to sample-to-sample variations in the overall path-length the photons have to travel before they reach the detector. These small physical differences from sample to sample lead to light scattering effects that influence the measured NIR spectra and result in baseline shifts and scaling variations (intensity variations). If the physical appearance of a sample is not of interest (e.g., as the determination of the composition of a sample), altered spectra can be detrimental to subsequent quantitative analysis as inaccurate results are obtained.

The pretreatment of NIR raw data is the first step of model development and optimization. Data pretreatment helps to correct unwanted systematic sample-to-sample variation in measured NIR spectral data. Additionally, pretreatment also helps to correct spectral variations due to intermolecular hydrogen bonds or spectrometer hardware. A suitable data pretreatment depends on the data to be analyzed (many measurements of a particular sample). Out of a number of tested pretreatments, the optimal pretreatment is characterized by the lowest statistical errors obtained in the data analysis step (linear regression). Optimal pretreatments help to eliminate interferences and facilitate information extraction from NIR spectral data. For example, pretreatment using multiplicative scatter correction (MSC) and standard normal variate (SNV) transformation are used. In literature, some studies applied orthogonal signal correction (OSC). Other classical methods of pretreatment include mean centering, detrending, first and second order derivatives, smoothing, auto-scaling, Pretreatment based on second order derivatives is very popular and can help to resolve nearby peaks and improve spectral resolution.

4.4 Reducing the quantity of NIR spectral data

Identification, calibration, and validation are based on spectra from dozens to hundreds of similar samples. Each NIR sample spectrum consists of data collected at hundreds to thousands of different wavelengths. This sums up in a vast amount of NIR spectral information that has to be processed. Moreover, it is typical for a NIR spectrum that the variables (absorptions at many wavelengths) are highly correlated which means that the measured absorbances at two or more wavelengths are not independent of each other. This is referred to as multicollinearity and can pose problems (poor prediction) with linear regression models used for calibration. Calibrations are developed through regressions of the measured NIR spectral data against the values of analyte concentrations determined by reference analytical methods (e.g., regression of measured absorptions against reference analyte concentration data).

To facilitate the classification and calibration steps, it is necessary to reduce the amount of data present in the pretreated spectra, to handle multicollinearity and to get a better overview of the data. Principal component analysis (PCA) is a very popular method for data compression (dimensionality reduction) and is used as an unsupervised pattern recognition technique in qualitative NIR analysis. However, when it comes to quantitative NIR analysis, it is not always necessary to apply PCA itself. The partial least squares regression model (PLS), which is the most popular linear calibration method used in quantitative NIR data analysis, simultaneously reduces the amount of data (dimension reduction of the NIR spectra matrix) and does the regression. Besides PCA, publications also mention genetic algorithms (GA) to extract relevant information from NIR spectra.

PCA is performed on data that was preprocessed at least by scaling (e.g., unit variance scaling) and mean centering. PCA itself is not dependent on results from reference analytical methods. The idea behind PCA is to mathematically decompose the originally correlated variables (absorptions at many wavelengths) into a much smaller set of uncorrelated latent variables (principal components, factors). From a mathematical point of view, principal components are vectors. The appropriate set of principal components represents the most common variations in the data set analyzed. The first principal component of this set is calculated such that it accounts for most of the variance in the data. Each succeeding principal component needed has to be orthogonal to the preceding principal

component and is calculated to have the highest remaining variance possible; this is done until only noise remains. «Overfitting» (choosing too many principal components so that one or more of them only represent noise or unwanted interferences) has to be carefully avoided as it is detrimental for the robustness and prediction performance of the involved model. It is well possible that an overfitted model may show excellent results for samples that are part of the calibration set but at the same time fails to predict properties of samples that are part of an external validation data set. However, «underfitting» should be avoided too, as important information could be lost by using too few principal components. Each of the newly calculated principal components holds information of the whole measured spectrum as it is a linear combination of the original variables. For every original variable (absorption at a wavelength) a score is obtained that describes the orientation of this variable to the principle component. The orientation of a principal component vector is characterized by so called loadings. Score plots help to interpret the NIR data subjected to PCA. As an example, outliers can easily be identified. More information on PCA can be obtained from Nørgaard (FOSS white paper) or Eriksson (2006).

4.5 Multivariate calibration methods and validation

Calibration is performed by using appropriate models that describe the relationship between dependent and independent variables. Let's assume first that the data is of univariate nature. This is the case, for example, when the data stems from single wavelength measurements. There, the independent variable is the measured absorbance at one single wavelength (X) that is measured for a number of calibration standards whose concentrations (dependent variable, Y) have been determined by a reference analytical method (e.g., by weighing). Provided that the Lambert-Beer law is followed, the relationship $Y = f(X)$ can be expressed by univariate calibration using linear regression. The calculated model parameters «slope» and «intercept» of the linear regression model precisely characterize this relationship and are used for the prediction of new samples. Multivariate data require multivariate calibration, which follows the same basic principle as just described for univariate data. The appropriate multivariate calibration model describes the relationship between the dependent and independent parameters. The independent variables are the absorbances at a number of wavelengths and more than one dependent variable (concentration values) can be accounted for. There are a number of tried and tested regression models that can be used for multivariate calibration like multiple linear regression (MLR) and partial least squares (PLS) regression. Literature also mentions principal component regression (PCR) as another linear calibration technique as well as artificial neural network (ANN) as a nonlinear calibration technique.

NIR data analysis using multiple linear regression (MLR) is normally affected by multicollinearity, which leads to a poor prediction performance. To overcome this problem and to obtain a strong and robust predictive model, the number of wavelengths used for the analysis has to be cut down considerably (e.g., only two to ten wavelengths are considered). To analyze only a carefully selected representative part of the recorded spectra, that is not at all or at least less influenced by interferences, can sometimes also lead to a better prediction performance with other regression models.

Compared to MLR, the two techniques PCR and PLS allow including more information in the calibration model and are not limited by multicollinearity.

PCR is a two-step multivariate calibration method obtained by consecutively applying PCA and MLR. The principal components obtained with PCA are the independent variables in the MLR step. These principle components were determined solely to account for most of the variance in the measured spectra (X matrix) but did not considering the Y variables at all. Thus, unfortunately, it is possible that these principle components of the PCR model are not always able to predict the component concentrations (Y variables) well.

PLS does not show this drawback. The PLS regression method performs the dimensionality reduction of the X matrix and the regression simultaneously. In analogy to the principle components of PCR, PLS uses so-called «PLS components» that are determined to maximize the covariance between the spectral data (independent variables) and the concentrations (dependent variables).

Literature describes the use of the ANN technique for NIR calibration when the relationship between the measured spectral data and the quantitative sample property to be analyzed is nonlinear. The linear calibration methods PCR and ordinary PLS would not work well under this condition.

It goes beyond the scope of this monograph to cover every mathematical detail of the different calibration models. Moreover, it is not necessarily a must to be a math wizard who has to know every detail about a used algorithm. It is the more important to know under which circumstances a model can successfully be used to solve a given analytical problem.

A calibration model has to be validated with a validation set of samples to get an impression of its predictive ability. It is important that these samples have not have been used in the calibration of the model. The model is characterized by a number of statistical parameters such as the root mean square errors of calibration and prediction (RMSEC and RMSEP). These are a measure for the deviation of the predicted values from the actual values obtained by reference analytical methods. Also specificity, linearity, range, accuracy, precision, robustness as well as limits of detection and quantification should be determined.

Further insight into chemometrics can be gained by consulting the many books and articles listed in chapter 12 of this monograph. Helpful information can also be found on the web under chemometrics.se or www.chemometry.com: If you are interested in PCR, then you can find a good tutorial on the webpages of the University of Brussels: <http://www.vub.ac.be/fabi/multi/pcr/>

5 NIR spectral region

The NIR region of the electromagnetic spectrum extends from the end of the visible spectral region (700 nm or 14285 cm^{-1}) to the beginning of the fundamental infrared (IR) spectral region (2500 nm or 4000 cm^{-1}). The most prominent absorption bands occurring in the NIR region are related to the overtone and combination bands of the fundamental molecular vibrations of C–H, N–H, O–H, and S–H functional groups observed in the mid-IR spectral region (Figure 1, Table 2). Thus, most chemical and biochemical species exhibit unique absorption bands in the NIR spectral region that can be used for both qualitative and quantitative purposes.

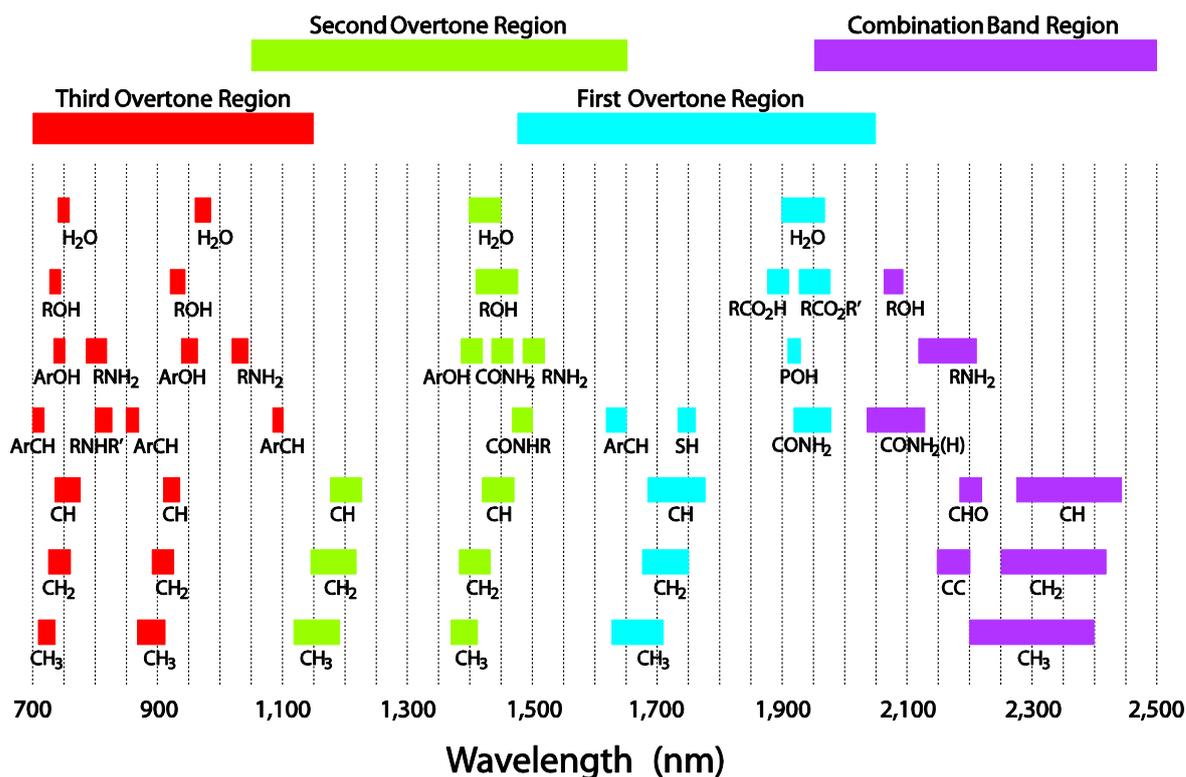


Figure 1: Major analytical bands and relative peak positions for prominent near-infrared absorptions. Most chemical and biological products exhibit unique absorptions that can be used for qualitative and quantitative analysis.

NIR spectral region

Table 2: Wavelengths in nm and wave numbers in cm^{-1} of some near-infrared bands of organic compounds and sulfur compounds

Wavelength [nm]	Wavenumber [cm^{-1}]	Assignment
2500	4000	combination S–H stretching
2200–2460	4545–4065	combination C–H stretching
2000–2200	5000–4545	combination N–H stretching; combination O–H stretching
1620–1800	6173–5556	first overtone C–H stretching
1400–1600	7143–6250	first overtone N–H stretching; first overtone O–H stretching
1300–1420	7692–7042	combination C–H stretching
1100–1225	9091–8163	second overtone C–H stretching
1020–1060	9804–9434	combination S=O stretching
950–1100	10526–9091	second overtone N–H stretching; second overtone O–H stretching
850–950	11765–10526	third overtone C–H stretching
775–850	12903–11765	third overtone N–H stretching
600–700	16667–14286	combination C–S stretching
450–550	22222–18182	combination S–S stretching

taken from B.H. Stuart, *Infrared Spectroscopy: Fundamentals and Applications* (2004)

NIR absorption bands are typically 10–100 times weaker than the corresponding fundamental mid-IR absorption band. This weak nature of NIR absorptions is actually an analytical advantage since it allows direct analysis of strongly absorbing, highly light-scattering matrices such as slurries, suspensions, pastes, and powders. NIR absorption bands are also very broad (bandwidths of 50–100 nm) and highly overlapped. NIR absorption bands occurring at longer wavelengths are stronger, sharper, and better resolved than their corresponding higher overtone bands, which occur at shorter wavelengths (Figure 2).

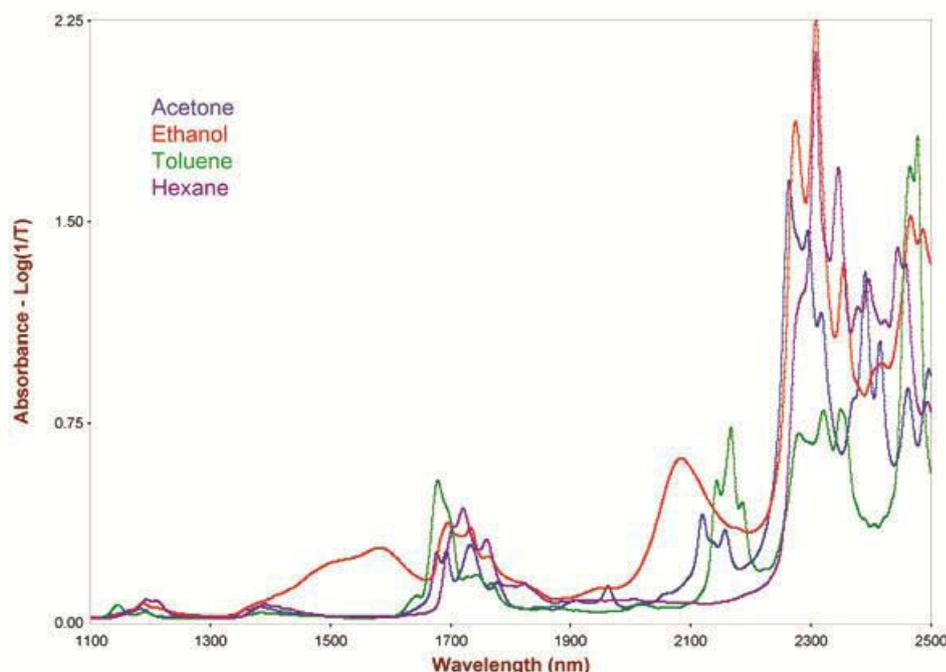


Figure 2: NIR (1100–2500 nm) diffuse transmittance ($\log(1/T)$) spectra of typical solvents

In addition to containing information about the chemical composition of a matrix, NIR spectra also contain physical information related to the sample. For example, particle size differences will produce a baseline offset; the coarser the particles, the greater the penetration of light and the greater the absorbance (Figure 3).

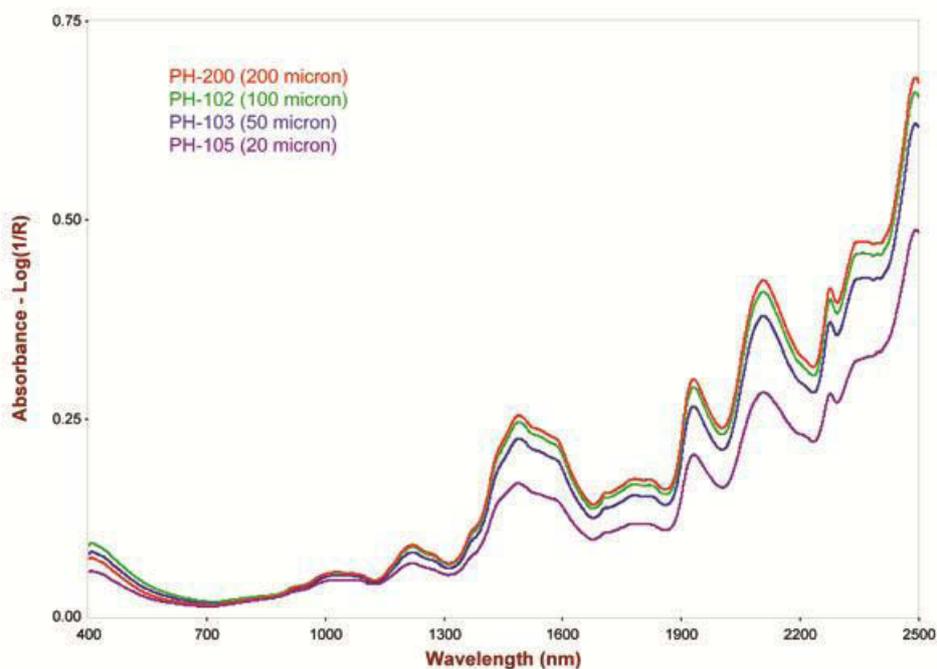


Figure 3: Visible (400–700 nm) and NIR (700–2500 nm) diffuse reflectance ($\log(1/R)$) spectra of differing grades of microcrystalline cellulose (Avicel™) which vary in particle size. Baseline increases as particle size increases due to changes in penetration depth and absorption.

NIR spectral region

In the NIR spectral region, light-scattering efficiency and absorptivity increase with wavelength whereas band overlap and penetration depth decrease with wavelength (Figure 4). Hence, the optimal spectral region to be used for a particular analysis is determined by matching the spectral properties of the NIR region with the required analytical performance, the sampling requirements (laboratory, atline, or process), and the physical properties of the sample. For example, measurement of a component whose bands are not well resolved from bands of other components in the matrix, or is present at low concentrations, is most easily accomplished in the longer wavelength region, which requires the use of a shorter optical path length or decreased penetration depth.

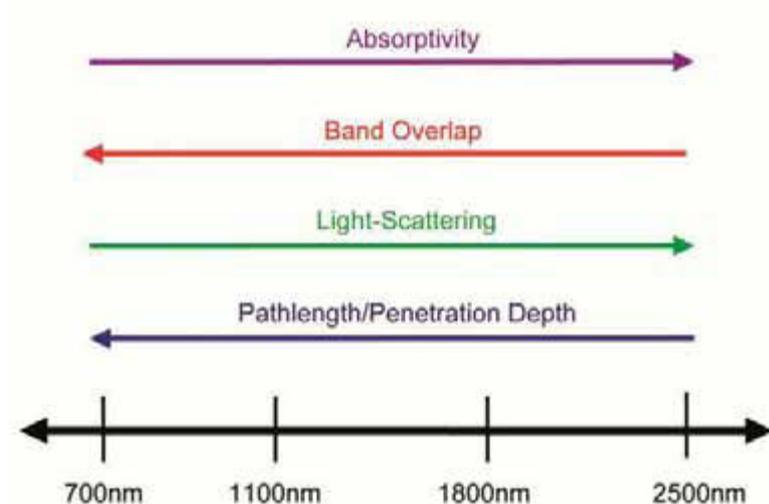


Figure 4: Chemical and physical characteristics of the NIR spectral region that vary with wavelength. These properties are considered when determining the ideal measurement mode and spectral region for performing an NIR analysis.

6 NIR analysis methods

NIR measurements are performed without dilution or the requirement of short optical path lengths or dispersion in nonabsorbing matrices that are necessary with traditional spectroscopic analysis methods such as UV/VIS or mid-IR spectroscopies. NIR spectra can be collected in either transmittance or reflectance mode. In general, transmittance ($\log(1/T)$) measurements are performed on translucent samples. Diffuse reflectance ($\log(1/R)$) measurements are performed on opaque or light scattering matrices such as slurries, suspensions, pastes or solids.

For NIR spectra collected in either measurement mode, it is often difficult to identify unique spectral features related to individual chemical components within a given matrix. To enhance spectral features and compensate for baseline offsets, it is common practice in NIR spectroscopy to apply a mathematical pretreatment to the spectral data. Frequently, the second-derivative of the absorbance data, with respect to wavelength, is calculated. In the second-derivative data, absorbance maxima are converted to minima with positive side-lobes. The apparent spectral bandwidth is also sharply reduced, allowing for resolution of overlapping peaks. Baseline differences between spectra are largely eliminated (Figure 5).

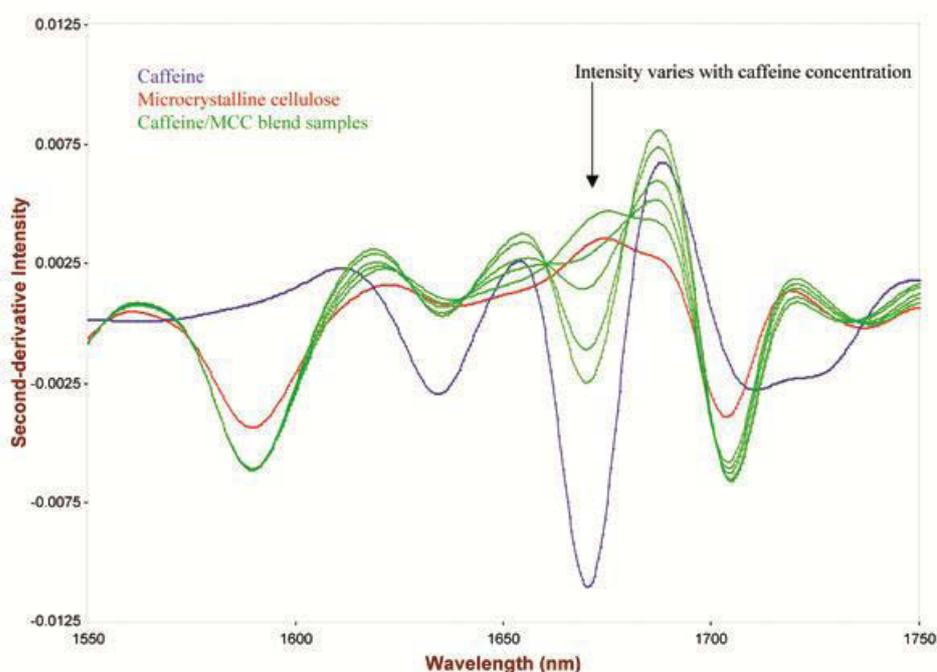


Figure 5: Expanded NIR first overtone second-derivative spectral data for caffeine, microcrystalline cellulose and two-component blends. Apparent bandwidth is reduced and scattering differences are eliminated. Region identified where the analyte of interest absorbs and other components have minimal or no absorption.

6.1 Qualitative analysis

Qualitative NIR analysis is performed using spectral library matching or pattern recognition algorithms to answer the question «is this material similar to an acceptable material that was previously analyzed». Using statistical matching algorithms, a qualitative NIR assessment will confirm that the NIR spectrum of an «unknown» material is (or is not) similar to the NIR spectrum of a material present in a library and whether or not it «fits» within the spectral variation derived from past lots of this material. Identity confirmation is performed by comparing the NIR spectrum of an «unknown» material to the composite spectra for all materials contained within a library. Qualitative analysis is performed by comparing the NIR spectrum of an «unknown» material with the spectral variance determined for each of the acceptable materials (of the same lot or type) contained within a library (Figure 6).

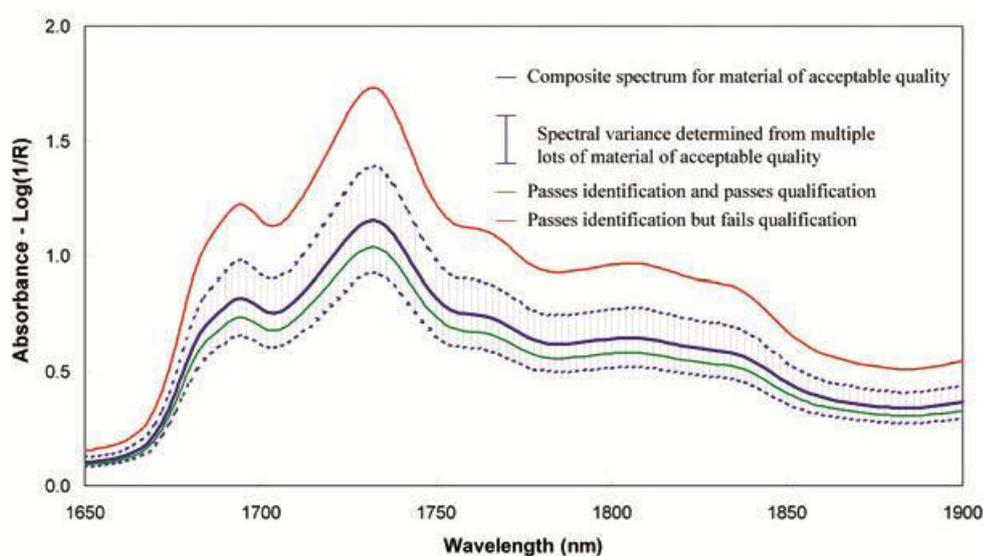


Figure 6: Spectral representation of identification and qualification approaches used for raw material inspection. Identification by correlation compares only the spectral «shape». Qualitative analysis routines use statistical methods to compare both the spectral mean and variance with an «unknown» material.

6.1.1 Library development

A library of known and qualified raw materials is initially created by collecting the NIR spectra of several different lots of a material of acceptable quality. Generally, an acceptable library can be created by collecting duplicate NIR spectra from duplicate samples for a minimum of three different lots of each material. Therefore, each raw material is represented in the library by at least twelve spectra (two scans/sample, two samples/lot, three lots/material).

To identify a material, methods such as Wavelength Correlation are employed to compare the NIR spectrum of an «unknown» material with the NIR spectra collected for known materials. From these methods, a correlation value (match index) is derived. The closer the correlation value is to 1, the greater the similarity of the NIR spectrum of the «unknown» material to the NIR spectrum of a material already present in the library.

A more rigorous qualitative assessment is performed by comparing the spectrum of the «unknown» material to the composite spectrum and spectral variance that is established for each of the materials contained within the library. In this case, statistical methods such as Mahalanobis Distance or residual variance are employed. Large portions of the NIR spectral region are often used to ensure that the method is sensitive to all possible contaminants or impurities.

6.1.2 Validation

Positive and negative control samples are used to validate the derived spectral library. For positive control, the NIR spectra of materials of acceptable quality are compared to the composite spectra and spectral variance for each product/material to ensure positive identification and qualification. For negative control, the NIR spectra of materials of unacceptable quality are compared to the composite spectra and spectral variance for each product/material contained within the library to ensure that these samples are «misidentified» or are flagged as being outside of the acceptable range of variability.

6.2 Quantitative analysis

Although unique NIR absorption bands can normally be identified in the spectra of samples that are consistent with pure component spectra, interfering absorption bands due to other components and matrix variations can affect the spectroscopic measurements. Therefore, traditional calibration approaches that utilize primary standards in simple solutions are not applicable. As a result, in quantitative NIR spectroscopy, empirical relationships are derived between the NIR spectra of a calibration set of samples and their corresponding reference analytical method for the constituents of interest.

6.2.1 Calibration

To ensure representative quantitative results are obtained, it is necessary to assure that the calibration set of samples is representative of the variation in the sample properties to be encountered in typical production lots. In general, the calibration sample distribution will contain a relatively even distribution of samples over the expected analyte range and a consistent number of representative samples at each analyte concentration level. This «boxcar» distribution will include samples that have variation in other chemical components within the matrix, as well as physical properties that may vary with the manufacture of this material (Figure 7). Generally, for industrial products, the calibration population will contain between 20–50 samples.

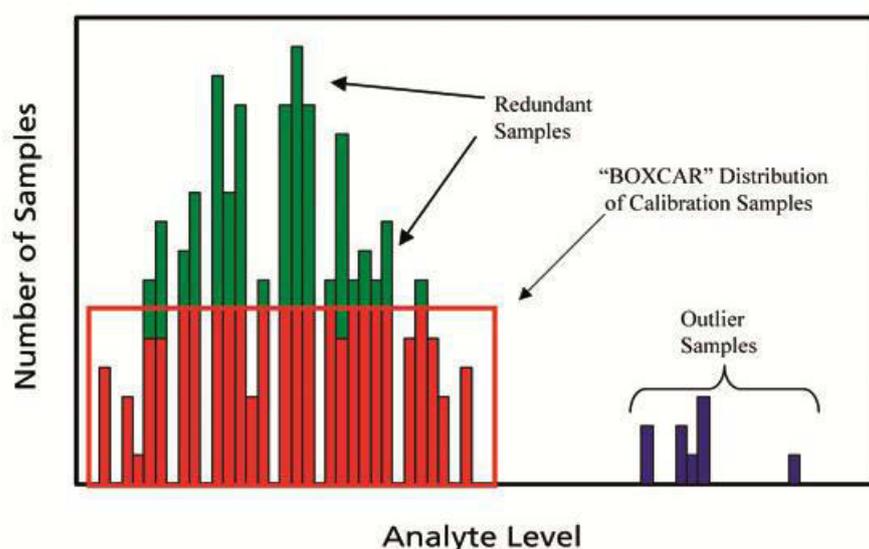


Figure 7: «Boxcar» calibration sample population used for developing robust NIR methods. Population should possess an even distribution of samples in terms of analyte range and other physical and chemical parameters that could vary with manufacture of this material.

A variety of chemometric algorithms, such as MLR, PLS, and PCR (chapter 4) can be used to derive suitable NIR spectroscopic models. In MLR, linear relationships are derived between the absorption of NIR energy at discrete wavelengths and the reference analytical chemistry for specific constituents. The primary wavelength is related to the analyte of interest. Subsequent wavelengths can be included to compensate for overlapping, interfering absorption bands due to other components, or changes in the light scattering intensity. Multivariate methods such as PLS extract spectral factors over a broad range of wavelengths (50–250 nm or more) that simultaneously relate to analyte concentration and compensate for interfering absorption bands, scattering differences and shifts in band position. Generally, MLR methods are reserved for simple chemical systems – where unique analyte bands are readily observable. PLS and PCR methods are utilized for more complex chemical systems – where analyte bands are highly overlapped by other matrix components. While PLS and PCR methods are better able to extract useful quantitative information, they are susceptible to unmodeled spectral variations. In contrast, since MLR methods are derived using only a few wavelengths, they tend to be less sensitive to subtle matrix variations. Although generally dependent upon the analyte range and distribution of the calibration sample population, a suitable NIR calibration model will have a standard error calibration (SEC) that is of the same order of magnitude as the error of the reference method and a correlation coefficient (R) that approaches 1 (Figure 8).

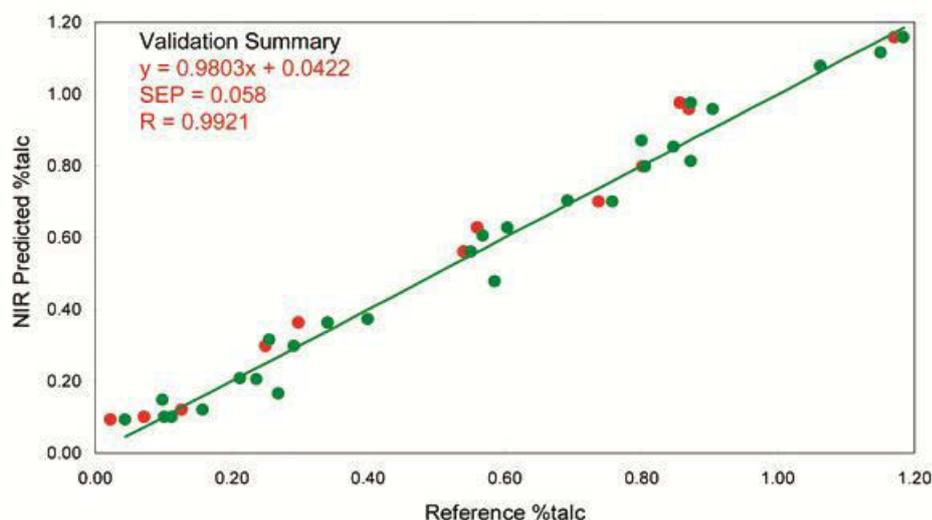


Figure 8: NIR calibration and validation scatter plot for analyzing talc in a three-component talc/methylcellulose/lactose matrix. Calibration model: MLR using second-derivative intensity at 1390 nm; R = 0.9921, SEC = 0.04%.

6.2.2 Validation

The integrity and applicability of the derived calibration model is totally dependent upon the set of data used to create the model. To ensure that the derived NIR spectroscopic model is both representative of the analyte under investigation and compensates effectively for interfering absorptions and scattering differences to be encountered, a second independent set of samples collected from additional production runs (or lots) is used to validate the performance of the calibration model. The correlation coefficient (R), the standard error of prediction (SEP), and the slope and y-intercept of the regression line are used to judge the adequacy of the calibration method. In general, for a similar sample distribution, a robust NIR model will yield an R and a slope that approaches 1, a y-intercept that approaches zero and an SEP that is similar to the SEC of the derived NIR calibration model (Figure 8).

7 NIR instrumentation

A number of technologies exist that can be used to separate the polychromatic NIR spectral region into monochromatic frequencies for both qualitative and quantitative purposes. Broadband, discrete filter photometers, or light-emitting diode (LED) based instruments provide spectral coverage over a narrow spectral region (50–100 nm). Diffraction grating, interferometer, diode-array or acousto-optic tunable filter (AOTF) based instruments provide full-spectral coverage. Selection of the appropriate technology is usually based upon the required analyte sensitivity and selectivity, reliability, ease-of-use, and implementation needs. Furthermore, pharmaceutical and chemical manufacturers demand cost-effective and efficient method development and implementation.

NIRS XDS analyzers use a combination of internal performance standards to maintain instrument stability and response and NIST-traceable external standards placed directly at the sample location to precisely match the band pass, photometric and wavelength response for all analyzers at the sample location. By precisely matching the performance for all instruments, a quantitative calibration model or a qualitative library developed on one XDS NIR analyzer can be used to predict quantitative or qualitative results on subsequent analyzers (of similar configuration) or the same analyzer after service (lamp or component change) without requiring a bias or slope adjustment or any other data manipulation.

7.1 Laboratory analyzers

The most significant advantage for laboratory NIR analysis is the elimination of sample handling and manipulation. Thus, analyses can be performed rapidly and reproducibly by a number of analysts. Laboratory analyzers can be operated in QC, R&D, and plant laboratories. Since NIR analysis is performed on unmodified samples, presenting the sample to the instrument is the most important aspect of NIR analysis. The modular design of laboratory analyzers ensures that analyses are optimized for specific sample types. The NIRS XDS MasterLab Analyzers perform true transmission measurements on whole solid dosage forms. The NIRS XDS RapidContent Analyzer performs reflectance analysis on materials in bags, vials, or sample cups and transmittance analysis of liquids in beakers. The NIRS XDS RapidContent Analyzer with Solids Module provides large area sampling for inhomogeneous powders, fibrous materials, flakes, and pellets. The NIRS XDS RapidLiquid Analyzer performs temperature controlled transmission analysis of liquids in cuvettes or vials. The NIRS XDS SmartProbe Analyzer enables the analyses of solids and liquids to be performed directly in shipping containers.

7.2 Process analyzers

Process NIR analyzers provide near real-time chemical process information while operating in harsh manufacturing conditions. The process sample interface is dictated by the sample type and process conditions. Contact transmission and reflectance probes are used for analyzing clear to opaque liquids and solids. Noncontact reflectance measurements are performed on materials transported in hoppers, transport and conveyor lines. Generally, NIR light from the instrument is transferred to the process sample interface using fiber optics. To maintain analytical performance, as the light scattering properties of the process sample increases, the number of fibers used in the fiber optic bundle is increased (Table 3). Single fiber process NIR analyzers are typically employed to analyze clear liquids. Micro-bundle process NIR analyzers monitor slightly scattering liquid media, suspensions, and drying processes. Full-bundle process NIR analyzers are used for the most challenging of applications such as monitoring the drying of hydrated media or analysis of low-level constituents. The length of the fiber optic interface can be from 1 m (large fiber bundle) to 150 m (single fiber optic). The use of longer fiber optic lengths can enable a process analyzer to be located outside of electrically classified or safety classified areas, or harsh operating conditions such as those with large temperature variations. Up to nine separate process streams or sampling points can be monitored using a multiplexed

process NIR analyzer from Metrohm NIRSystems. Multiplexing decreases both the cost/measurement point and the overall implementation cost for a process NIR analyzer. However, a risk-assessment should be completed to ensure the economic benefits clearly support the increased liability per measurement point.

Table 3: Comparison of fiber optic interface, fiber bundle size and measurement mode. To maintain analytical performance, fiber count is increased as the turbidity of process sample increases.

Fiber optic interface	Fiber size/count	Fiber length [m]	Sample type	Mode
Single fiber	600 μm , 1 illumination/1 collection	1–150	Clear liquids, thin films, gases	Transmission
Small fiber bundles	200 μm , 40 illumination/40 collection	1–75	Turbid liquids and suspensions	Transmission
Small fiber bundles	200 μm , 40 illumination/40 collection	1–75	Powders and films	Reflectance
Large fiber bundles	200 μm , 210 illumination/210 collection	1–15	Pastes, slurries, pellets, fibers	Reflectance

8 NIR analysis strategies

Today's NIR instrumentation can be installed in the laboratory, atline or directly into a process stream, dryer, extruder, or reactor. Overall, the most appropriate NIR measurement mode and location of the NIR analyzer is dictated by the optical properties of the samples, the required analyte selectivity and sensitivity, the duration of the process run, and the monitoring and control requirements (Figure 9).

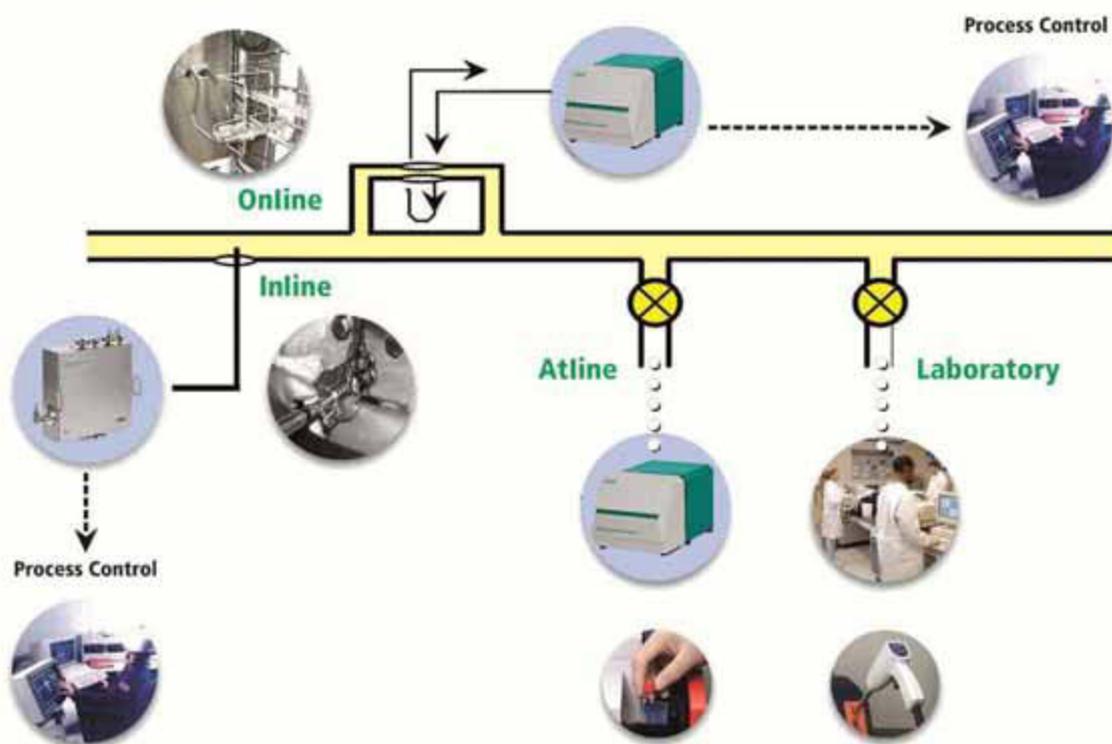


Figure 9: NIR analysis strategies. The process monitoring and control requirements and measurement performance determine ideal strategy.

8.1 Laboratory analysis

Situated in a centralized support laboratory, a single laboratory NIR instrument can support multiple manufacturing operations from raw material inspection to in-process analysis to finished product release. This configuration provides the greatest analysis flexibility as both reflectance and transmittance measurements can be performed on process samples, feedstocks, ingredients, raw materials, effluents, and finished products. Laboratory analysis strategies require manual sampling and therefore the turnaround time for results can take up to 1 hour, although the NIR analysis requires less than a minute. This analysis strategy is most useful for supporting pilot plant/scale up facilities, or measurement of long duration runs where process information is required infrequently or when lengthier analytical techniques can be eliminated.

8.2 Atline analysis

In this analysis strategy, an NIR analyzer is located close to a process stream. This analyzer configuration is dedicated to performing a particular analysis on a specific sample type, e.g., transmittance measurements of solid dosage forms. Atline analysis strategies require manual sampling and the analyzer must meet appropriate classifications such as IP55-NEMA12. The turnaround time for results can be less than 5 minutes. This strategy fits well in process monitoring and control strategies, and in manufacturing operations.

8.3 Online analysis

In this analysis strategy, a process NIR analyzer is interfaced to the process using a sample-loop and performs a particular analysis on a specific sample type, e.g., transmittance analysis on a chemical reaction. NIR spectral measurements are performed on a continuous flow of sample as it passes through a flow-cell. Side-streams provide a convenient means for sample conditioning such as heating, filtering, or debubbling, for performing maintenance and allow calibration and check samples to be analyzed, all while the process is operating. Online analysis strategies provide unattended near real-time (results in < 10 s) analysis on specific media and are optimized for that particular analysis. This strategy fits well in closed-loop monitoring and control strategies for scale-up and manufacturing operations.

8.4 Inline analysis

In this analysis strategy, a process NIR analyzer is interfaced directly to the process using fiber optics and a stainless steel (or other material) probe that is inserted directly into a port that is installed in the process stream or vessel. This analyzer configuration is dedicated to performing a particular analysis on a specific sample type, e.g., reflectance analysis in a dryer. A direct interface requires minimal supporting hardware but maintenance cannot be performed unless the process is shut down. In addition, for transmission analysis, the narrow optical path lengths combined with the fluid dynamics of process streams can sometimes make this analysis strategy difficult to implement. Inline analysis strategies provide unattended near real-time (results in < 10 s) analysis on specific media and is optimized for that particular analysis. This strategy fits well in closed-loop monitoring and control strategies for scale-up and manufacturing operations.

8.5 Data reporting

For laboratory and atline analyses, results are normally reported using manual entry or data-logging of results for process monitoring. Process NIR analyzers are normally interfaced directly to a distributed control system (DCS) or process logic control (PLC) system. Results can be conveyed using analog current loops, digital I/O modules, and digital communication protocols such as Modbus or OPC (Figure 10).

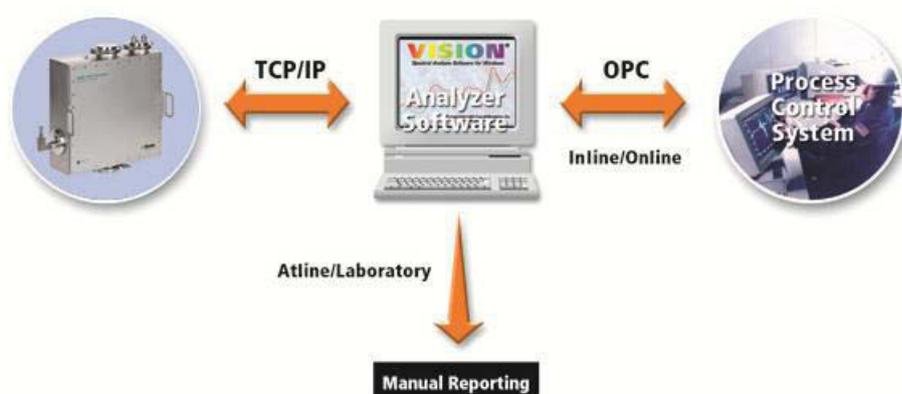


Figure 10: Reporting and communication of NIR results and data. Manual reporting for laboratory or atline analysis, integrated closed-loop communication for online/inline analysis.

8.6 Maintenance

Over time, changes in raw materials, process improvements, or other «unknown» factors can potentially cause the performance of an NIR method to be compromised. It is also possible that the instrument is no longer operating appropriately. Thus, routine control tests should be performed to monitor the process and the analyzer to create confidence in the NIR measurement and to verify that the analyzer is functioning properly (Figure 11).

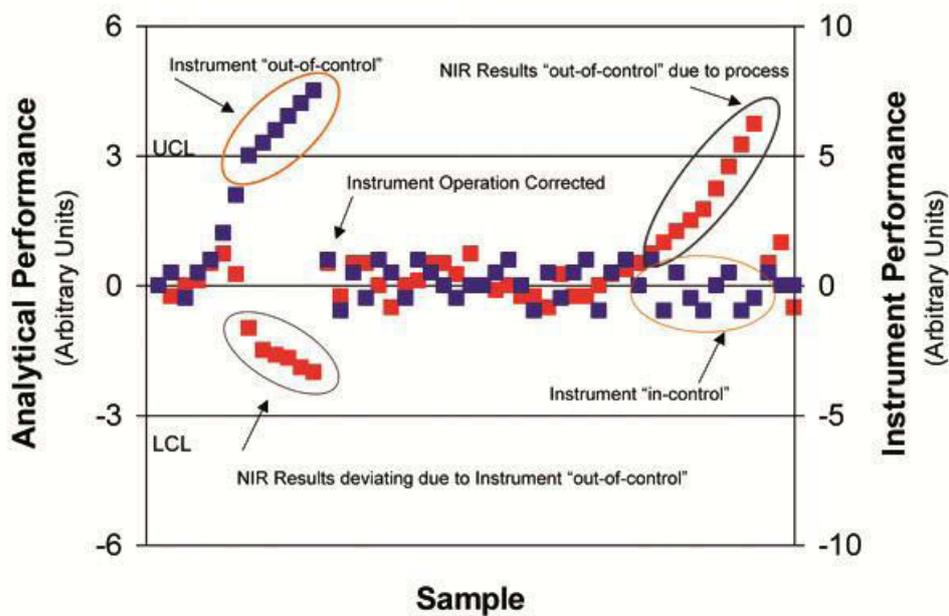


Figure 11: NIR Analyzer and method control charts for monitoring both the analyzer and method performance and to identifying «out-of-control» situations.

9 Industry sectors and applications

Biotechnology

In bioreactors, product titer yield can be enhanced by predicting key analytes in real-time, allowing for timely control and optimization of nutrient levels as well as checking ammonia and lactate concentration. Both mammalian and microbial cell cultures have been monitored with our laboratory and process instruments with the accuracy and reliability of offline chemistry methods.

Common biotechnology analytes include: glucose, lactate, ammonia, glutamine, glutamate, product titer, acetate, and biomass.

Chemistry

NIR is extremely sensitive to the O–H absorbance; moisture determination is therefore a common application. Additionally, NIR is sensitive to N–H, C–H and S–H bond absorbance, for what reaction monitoring and end point determination are good applications for organic chemistry reactions and processes. Many applications that have historically been measured with physical property testing, like viscosity, may be measured with NIR if it is dependent on an intrinsic chemical characteristic, such as chain length or cross linking.

Common chemical applications include the determination of the acid value, hydroxyl number, adhesive content, antioxidant and UV inhibitor content, cure, melt index, HDPE/LDPE, melamine content, plastic identification, polymer analysis, alcohol detection/analysis and residual solvent detection.

Petrochemistry

In petrochemistry, monitoring of the blending process by NIR has proven huge economical savings, since crucial properties such as water content, density, viscosity, additive content, and hydroxyl number can be measured in less than one minute.

Gasoline parameters measured with NIR cover research octane (RON, ASTM D 2699), motor octane (MON, ASTM D 2700), and road octane number (RdON), and volume percentage or even mole percent of individual components (paraffins, isoparaffins, aromatics, naphthenes, and olefins; PIANO). Common diesel parameters measured with NIR include specific gravity, viscosity, flash point, cold filter plugging point (CFPP), pour point, and cloud point.

Pharma

The manufacture of pharmaceutical products from raw material identification to the measurement of content uniformity of dosage forms can be assisted by the implementation of NIR methods. The United States Food and Drug Administration (FDA) and European Union health guidelines have increased the workload and rigor associated with receiving inspection, blending, and content assay. With the advent of 100% container testing for receiving inspection of raw materials in Europe and Canada, NIR technology can reduce the time and skill level required to meet the increased challenge of compliance.

With the PAT (Process Analytical Technology) initiative, the FDA aims to bring about an increase in efficiency in pharmaceutical production, including a trend away from final checks towards real-time process analysis and control. The initiative requires rapid analytical techniques that allow a comprehensive online and inline monitoring of the manufacturing process. To this end, NIRS is the most powerful analytical tool that is currently dominating all PAT projects.

Common pharmaceutical applications using NIR include: receiving inspection of excipients and active pharmaceutical ingredients (API), blend uniformity, granulation, drying and coating, and particle size analysis. Additionally, NIR is an invaluable tool for the detecting of counterfeit drug products and the determination of water and residual solvent content. NIRS is described in the European (Ph.Eur.) and Japanese (JP) Pharmacopoeia as well as in the United States Pharmacopeia (USP).

Polymers

Polymer attributes have been measured with NIR since the 1950s. Due to NIR requiring no sample preparation and being nondestructive, many polymer and plastic attributes can be measured rapidly inline or atline for qualitative as well as quantitative parameters. Thermoplastics production, raw material purity, and moisture content can be analyzed with NIR. The disappearance of double bonds can be monitored real time in reactions using inline process NIR. Residual solvents and monomers can also be detected using NIR analysis. Additives like antioxidants, UV inhibitors and ones that prevent discoloration are measured using NIR. It is an excellent tool for determining the incoming raw materials and reaction endpoints, reducing over-processing of product and improving production consistency. As in other sectors, it facilitates the quick determination of the moisture content.

NIRS analysis also allows for a determination of physical properties such as molecular weight, degree of branching, tacticity, melting point, particle size, density, and viscosity. Moreover, hydroxyl number as well as acid and amine value can be determined.

Pulp and paper

The pulp and paper industry has utilized NIR analysis for many years providing qualitative and quantitative information about incoming timber materials and lignin content. Discriminant NIR analysis can be used to determine species, hardwoods from softwoods and sapwoods from heartwoods.

Common pulp and paper attributes measured with NIR include: kappa number, lignin content, kraft pulp yield, tall oil, moisture, resin, brightness, wood species, hardwood/softwood ratio, coatings and component analysis (clay, titanium dioxide, fillers, ash, etc.).

Education

Although NIR is primarily used in production-related applications, it has found great utility in research and development facilities. Pharmaceutical R&D often uses NIR in pilot plants operations to help evaluate daily processes. Work aims at understanding blend uniformity, granulation drying, dissolution profile prediction, and lyophilization processes. Chemical R&D facilities use NIR to optimize process performance and product quality.

Common R&D parameters measured include: moisture content, coating thickness, roller compaction hardness, blend uniformity, granulation, particle size and prediction of dissolution profile.

Textiles (Polymers)

NIR has been long used in the textile industry to differentiate fiber types for carpet recycling. Blend analysis of different polymers fibers can be analyzed with NIR. Real-time analysis of the application of polyvinyl alcohol (PVA or PVOH) sizing to warp yarn has been done with NIR online process analyzers. Common fiber identified with NIR include: cotton/linen, merchandized cotton, acrylic, modified acrylic, acetate, triacetate, Nomex[®], Kevlar[®] (K-29, K49, and K129), nylon-6, nylon-6,6, silk, polyester, cationic and disperse dyeable polyester, polypropylene, PVA and PVC.

Other NIR applications

Many applications for NIR do not fall under common headings. NIR can be applied nondestructively and with no sample preparation to many different fields. NIR applications work best on material that have strong OH, CH, NH or SH absorbers. Because NIR models are empirically developed, NIR is most useful in applications where the same product is measured on a routine basis. Wood products industries that make particle board panels and products use NIR for moisture analysis and adhesive content analysis. Further applications in the NIR measurements include: moisture in soaps, adhesive level in particle board chips.

Moreover, there is an increasing demand to combine NIR imaging techniques with well-established brain imaging techniques such as magnetic resonance imaging (MRI) and positron emission tomography (PET). Additionally, there are several applications in the environmental sector.

Table 4: Applications that highlight NIR spectroscopy (under <http://www.metrohm.com/nirs.html>)

Application	Title	Industry
AB-358	Analysis of residual moisture in a lyophilized pharmaceutical product by near-infrared spectroscopy	Pharma
AN-NIR-001	Determination of active ingredients in solid (pharmaceutical) dosage forms using solid-state standard additions	Pharma
AN-NIR-002	Nondestructive, single tablet analysis using the NIRS XDS RapidContent Analyzer	Pharma
AN-NIR-003	Analysis of copolymer levels in polymer pellets by near-infrared spectroscopy	Polymers
AN-NIR-004	Determination of additives in polymer pellets by near-infrared spectroscopy	Polymers
AN-NIR-005	Determination of finish on nylon fibers by near-infrared spectroscopy	Polymers/textile
AN-NIR-006	Near-infrared analysis of polyols	Polymers
AN-NIR-007	The near-infrared analysis of polyols: process monitoring in a hostile environment	Polymers
AN-NIR-008	Analysis of lignin in wood pulp	Paper & pulp
AN-NIR-009	Determination of hardwood/softwood content in wood products by near-infrared spectroscopy	Paper & pulp

10 Standards

The impetus to develop globally accepted standards is greater than ever before. Standards are technical documents of agreed-upon solutions that have already been time-tested and proven. Norms and standards help translate research results into products and processes with a short time-to-market.

Pharmacopeias ensure safety of drugs

Specifications and test methods for the commonly used active ingredients and excipients are monographed in detail in national pharmacopeias in more than 38 states according to the World Health Organization (WHO). NIRS has emerged as a powerful tool for analysis of pharmaceuticals including qualitative and quantitative evaluation of ingredients, additives, or water contents. Moreover, it can be used for the evaluation of physical properties and for controlling the pharmaceutical manufacturing process.

As a secondary test method, NIRS appears in the European Pharmacopoeia (2.2.40), the Japanese and the US Pharmacopeia (USP<1119>).

Dissemination of NIRS methods by citation in standards

Every year several tens of million NIRS analyses are performed in the pharmaceutical, (petro-)chemical, and polymer laboratories worldwide. The potential of this former «sleeping giant» is by far not yet exhausted. The success of NIRS is linked with the understanding of chemometrics and the applied predictive models. The more NIRS finds its way into international standards, the more the unique strength of this diagnostic tool emerges. Table 4 shows a selection of some international standards where NIRS is used.

Table 5: A selection of international standards that use NIRS

Standard	Title	Industry
ASTM D 6122	Standard practice for validation of the performance of multivariate online, atline, and laboratory infrared spectrophotometer based analyzer systems	Instrumentation
ASTM E 1655	Standard practices for infrared multivariate quantitative analysis	General
ISO 11151	Lasers and laser-related equipment – standard optical components – part 1: components for the UV, visible and near-infrared spectral ranges	Instrumentation
ISO/TS 10868	Nanotechnologies – characterization of single-wall carbon nanotubes using ultraviolet-visible-near infrared (UV-Vis-NIR) absorption spectroscopy	Nanotechnology
ISO 15063	Plastics – polyols for use in the production of polyurethanes – determination of hydroxyl number by nir spectroscopy	Polymers

11 Glossary

Accuracy: The measure of the degree of agreement between the estimated and the true value. The accuracy of an NIR method is usually comparable to the accuracy of reference method used to derive the NIR calibration.

Anharmonicity: Refers to the vibration of atoms and describes the deviation from harmonic oscillation.

Bias: A systematic error between the predicted NIR results and the reference analytical chemistry. This could be related to changes in the instrument performance or due to changes in the sample matrix.

Calibration: Step used to derive statistical quantitative equations that relate the changes in the NIR spectra to variation in components contained within the actual samples.

Calibration, instrument: See Instrument Calibration.

Calibration data set: Set of actual samples that have been previously analyzed for component of interest. Samples are chosen to adequately reflect variation in properties to be encountered with respect to the range of analyte values, matrix variations and interfering constituents. Used to generate a calibration equation.

Correlation coefficient (R): A measure of the strength of the linear relationship that exists between the NIR spectral data and constituent values and should be as close to 1 as possible.

Harmonic oscillator: Model in classical mechanics that describes the harmonic vibration of a diatomic molecule. The model describes the vibrations of two balls (atoms) that are linked with a spring (bond).

Hooke's Law: For small displacements, a spring experiences a restoring force that is proportional to the displacement ($F = -kx$ with $F =$ force, $k =$ spring constant, $x =$ displacement). In this model, the spring stands for the chemical bond of a diatomic molecule.

Identification: Step used to confirm the identity of a material by comparing the NIR spectrum of an "unknown" material to the composite spectra for all materials contained within a library data set.

Instrument calibration: XDS instruments use a traceable wavelength reference to adjust the wavelength response of each instrument to known, repeatable values. In reflectance, the wavelength standard is traceable to NIST SRM-1920a. In transmission, the wavelength standard is traceable to NIST SRM-2035. This process enhances transferability between instruments, by virtually eliminating any difference in wavelength response between similar XDS instruments.

Instrument performance certification (IPC): A program offered by Metrohm NIRSystems, which provides complete testing, documentation, and certification of the instrument every six months. This program serves as the *Instrument Operational Qualification* (IOQ) as required in the USP Chapter on NIR Spectrophotometry.

Library data set: Set of actual samples of known and qualified raw materials. Generally, an acceptable library can be created by collecting duplicate NIR spectra from duplicate samples for a minimum of three different lots of each material. This ensures that representative variations (particle size, moisture, color) are included in the library.

Library development: Step used to derive pattern recognition algorithms to identify and qualify raw ingredients or products.

Mahalanobis distance: A statistical method used in qualitative analysis that determines the separation between two components or spectra. For identical materials, the distance approaches zero.

Maximum distance: A statistical method used in qualitative analysis that determines the location of a spectrum relative to all library product spectra. For similar materials, the distance approaches zero.

Multilinear regression (MLR): A statistical method used in quantitative analysis that uses information at more than one wavelength to relate analyte concentration (and compensate for interferences) to NIR spectra. Primary wavelength usually relates to spectral variation related to changes in analyte concentration. Subsequent wavelengths are included to correct for overlapping, interfering absorptions or light-scattering differences.

NIST: National Institute of Standards and Technology, based in the United States. This organization defines characteristics and response of standards used for instrument testing. NIST has two standards

suitable for NIR wavelength testing, SRM-1920a for reflectance, and SRM-2035 for transmission. NIST has also published methods and procedures that are used to qualify photometric standards for NIR use, although no NIST-issued standards exist for NIR photometric testing.

Outlier: A spectrum that is statistically different than the majority of spectra in a population. Outlier samples should not be used in the calibration population. In routine operation, outliers could indicate a change in the manufacturing process and should be investigated.

Partial least squares (PLS): A statistical method used in quantitative analysis that extracts spectral factors over broad segments of the NIR spectrum that relate to both spectral variations related to changes in analyte concentration and differences due to variations in other matrix components. Can yield greater analyte sensitivity than an MLR method but is more sensitive to unmodeled spectral variations.

Performance verification tests: USP has defined instrument tests and acceptance specifications for regular, periodic verification of instrument performance. Includes tests of instrument noise, wavelength standard response, and photometric standard repeatability. Metrohm NIRSystems has included these tests and acceptance specifications in the instrument operating software.

Photometric standards: A set of standards of known photometric values, used to test an instrument for correct photometric (absorbance) response. In reflectance, the material is usually a perhalopolyethylene mixture, doped with carbon black. In transmission, neutral density filters are used. For tablet standards, a high-absorbance perhalopolyethylene mixture is used.

Precision: The reproducibility of data under identical conditions. The precision of an NIR method is usually better than the precision of reference method used to derive the NIR algorithm.

Principal component analysis (PCA): A statistical method used in qualitative analysis that describes unique differences within a set of spectra using a linear combination of wavelengths.

Qualitative analysis: Qualitative NIR analysis is performed using spectral library matching or pattern recognition algorithms to answer the question «is this material similar to an acceptable material that was previously analyzed».

Quantitative analysis: Quantitative NIR analysis is performed using statistical regression techniques to provide chemical information about the composition of a particular matrix.

Reference, instrument: A photometric reference, used in reflectance, that serves as a «background» for the sample scan. This may be a ceramic tile or some other stable, spectroscopically reflective material.

Reference method: Analytical method used to perform primary analysis on the calibration and validation samples. All errors in the reference method will be included in the derived NIR method and therefore the reference method should be as accurate and precise as possible.

Reference standardization: The process of measuring the true spectroscopic value of an instrument reference, and thereby adjusting the true response of the instrument to a known scale of absolute absorbance. This process enhances transferability between instruments, by virtually eliminating the measured differences which may be attributed to individual instrument reference materials.

Standard error of calibration (SEC): A measure of how well the calibration equation can reproduce constituent values of spectra used to create the calibration. SEC is the standard deviation of the differences between the reference values and the NIR values derived in the calibration step. The SEC should be similar in value to the standard error of the reference method.

Standard error of prediction (SEP): A measure of how well the calibration equation can reproduce constituent values of spectra not contained in the calibration data set. The standard deviation of the differences between the reference values and the NIR values predicted for the samples contained in the validation data set using the algorithm derived in the calibration step. For identical distributions, the SEP will be similar to the SEC but provides a better estimation of the true accuracy of the NIR calibration algorithm.

USP (United States Pharmacopeia): The USP issues guidelines and recommendations for NIR instrument testing and use. The specific chapter relating to NIR is <1119>. Metrohm NIRSystems has adopted USP instrument test guidelines for users, and has included the test methods in the

instrument operating software, for «Performance Verification Tests.»

Validation: Step used to ensure that the derived qualitative and/or quantitative algorithm is representative of the matrix/analyte under investigation and not unique to the samples present only in the library/calibration data sets.

Validation data set: Set of actual samples that resemble the calibration data set (but not used for calibration or library development) in terms of analyte range and sample variability. Samples are chosen to adequately reflect variation in actual samples to be encountered with respect to the range of analyte values, matrix variations and interfering constituents.

Wavelength certification: The software test that uses the wavelength standard to verify correct wavelength response of the instrument.

Wavelength correlation: A measure of spectral similarity used for identification of raw materials. For spectrally similar materials, the correlation approaches 1.

Wavelength standard: A material, which, when measured on the NIR instrument, exhibits measurable absorbance peaks in known areas of the spectrum. The «as-measured» peak positions are compared to the known, or «nominal» peak positions. If the «as-measured» peak positions agree with the nominal peak positions within a defined wavelength tolerance, the instrument is known to be reading correctly on the wavelength scale. Reflectance standards use a mixture of rare-earth oxides, and are traceable to NIST SRM-1920a. Transmission standards may be composed of rare-earth glass, or a mixture of stable, narrow-band absorbers, and are traceable to NIST SRM-2035.

12 References

Near-infrared spectroscopy

- H.P.R. Aenugu, D.S. Kumar, Srisudharson, N. Parthiban, S.S. Ghosh, and D. Banji, *Near Infra Red Spectroscopy – An Overview, International Journal of ChemTech Research* **3** (2011) 825–836.
- V. Baeten and P. Dardenne, *Spectroscopy: Developments in Instrumentation and Analysis, Grasas y Aceites* **53** (2002) 45–63.
- M. Blanco and I. Villarroya, *NIR spectroscopy: a rapid-response analytical tool, Trends in Analytical Chemistry* **21(4)** (2002) 240–250.
- D.A. Burns and E.W. Ciurczak, *Handbook of Near-Infrared Analysis*, 2nd Edition, Practical Spectroscopy Series, Volume 27, Marcel Dekker, Inc., New York, USA (2001).
- S.L. Cantor, S.W. Hoag, C.D. Ellison, M.A. Khan, and R.C. Lyon, *NIR Spectroscopy Applications in the Development of a Compacted Multiparticulate System for Modified Release, AAPS PharmSciTech* **12(1)** (2011) 262–278.
- D. Cozzolino, W.U. Cynkar, R.G. Damberg, L. Janik, and M. Gishen, *Near infrared spectroscopy in the Australian grape and wine industry, CRCV Technology Application Note*, published online in April 2007: [http://www.crcv.com.au/viticare/vitinotes/Viti-Notes/near infrared spectroscopy technology for grape and wine quality measurement/Near infrared spectroscopy in the Australian grape and wine industry.pdf](http://www.crcv.com.au/viticare/vitinotes/Viti-Notes/near%20infrared%20spectroscopy%20technology%20for%20grape%20and%20wine%20quality%20measurement/Near%20infrared%20spectroscopy%20in%20the%20Australian%20grape%20and%20wine%20industry.pdf)
- European Pharmacopoeia 5.0 (2005), 2.2.40 Near-Infrared Spectrophotometry.
- B.G. Osborne, T. Fearn, and P.H. Hindle, *Practical NIR Spectroscopy with Applications in Food and Beverage Analysis*, 2nd Edition, Longman Technical and Scientific, London (1993).
- P. Pasquini, *Near Infrared Spectroscopy: Fundamentals, Practical Aspects and Analytical Applications, J. Braz. Chem. Soc.* **14(2)** (2003) 198–219.
- G. Reich, *Near-infrared spectroscopy and imaging: Basic principles and pharmaceutical applications, Advanced Drug Delivery Reviews* **57** (2005) 1109–1143.
- H.W. Siesler, Y. Ozaki, S. Kawata, and H.M. Heise, *Near-Infrared Spectroscopy – Principles, Instruments, Applications*, Wiley-VCH Verlag GmbH, Weinheim, Germany (2002).
- B.H. Stuart, *Infrared Spectroscopy: Fundamentals and Applications*, John Wiley & Sons, New York, USA (2004).
- United States Pharmacopoeia USP 35-NF 30 (2013), General chapter <1119> Near-Infrared Spectroscopy.
- P. Williams and K. Norris, *Near-Infrared Technology in the Agriculture and Food Industries*, 2nd Edition, American Association of Cereal Chemists, St. Paul, MN, USA (2001).

Chemometrics

- T. Azzouz, A. Puigdoménech, M. Aragay, and R. Tauler, *Comparison between different data pre-treatment methods in the analysis of forage samples using near-infrared diffuse reflectance spectroscopy and partial least-squares multivariate calibration method, Analytica Chimica Acta* **484** (2003) 121–134.
- R.G. Brereton, *Chemometrics: Data Analysis for the Laboratory and Chemical Plant*, Wiley, Chichester, UK (2003).
- R.G. Brereton. *Chemometrics for Pattern Recognition*, Wiley, Chichester, UK (2009).
- N. Broad, P. Graham, P. Hailey, A. Hardy, S. Holland, S. Hughes, D. Lee, K. Prebble, N. Salton, and P. Warren, *Guidelines for the Development and Validation of Near-Infrared Spectroscopic Methods in the Pharmaceutical Industry, Handbook of Vibrational Spectroscopy*, published online on 15. August 2006: <http://onlinelibrary.wiley.com/doi/10.1002/0470027320.s8303>
- D.A. Burns and E.W. Ciurczak, *Practical Spectroscopy Series Vol. 35: Handbook of Near-Infrared Analysis*, 3rd Edition, CRC Press, Boca Raton, FL, USA (2008).

- J. Chen and X.Z. Wang, *A New Approach to Near-Infrared Spectral Data Analysis Using Independent Component Analysis*, *J. Chem. Inf. Comput. Sci.* **41** (2001) 992–1001.
- H. Chung, H. Lee, and C.-H. Jun, *Determination of Research Octane Number using NIR Spectral Data and Ridge Regression*, *Bull. Korean Chem. Soc.* **22**(1) (2001) 37–42.
- C.B. Davis, *Application of Chemometric Analysis to UV-Visible and Diffuse Near-Infrared Reflectance Spectra*, Thesis, Baylor University, Waco, TX, USA (2007).
- N. Dahlman, *Near Infrared Spectroscopy – An Introductory Study On Measurement Techniques And Tools For Analysing Moisture Content In Pulp And Paper Media*, Thesis, Karlstad University, Karlstad, Sweden (2012).
- L. Eriksson, E. Johansson, N. Kettaneh-Wold, J. Trygg, C. Wikström, and S. Wold, *Multi- and Megavariate Data Analysis Part I: Basic Principles and Applications*, 2nd revised and enlarged edition, UMETRICS AB, Umeå, Sweden (2006).
- P. Fardim, M.M.C. Ferreira, and N. Durán, *Multivariate Calibration for Quantitative Analysis of Eucalypt Kraft Pulp By NIR Spectrometry*, *Journal of Wood Chemistry and Technology* **22**(1) (2002) 67–81.
- J. Fortier, J. Rodgers, J. Foulk, and D. Whitelock, *Textile Technology: Near-Infrared Classification of Cotton Lint, Botanical and Field Trash*, *The Journal of Cotton Science* **16** (2012) 72–79.
- J. Gayo, *Species Authenticity and Detection of Economic Adulteration of Atlantic Blue Crab Meat Using VIS/NIR Spectroscopy*, Thesis, North Carolina State University, Raleigh, NC, USA (2006).
- P. Geladi, *Chemometrics in spectroscopy. Part 1. Classical chemometrics*, *Spectrochimica Acta Part B* **58** (2003) 767–782.
- Guideline on the use of Near Infrared Spectroscopy (NIRS) by the pharmaceutical industry and the data requirements for new submissions and variations*, Draft from 20. January 2012, EMEA/CHMP/CVMP/QWP/17760/2009 Rev2, European Medicines Agency, London, UK (2012).
- J.A. Guthrie, *Robustness of NIR Calibrations for Assessing Fruit Quality*, Thesis, Central Queensland University, Rockhampton, Australia (2005).
- T. Hasegawa, *Chemometrics for spectroscopic analysis*, *Anal. Bioanal. Chem.* **375** (2003) 18–19.
- V.S. Hollis, *Non-Invasive Monitoring of Brain Tissue Temperature by Near-Infrared Spectroscopy*, Thesis, London, UK (2002).
- H.C. Howland, *Application of NIR Spectroscopy and Fluorescence Spectroscopy for Monitoring of Curing of Sustained Release Coatings*, Thesis, University of Maryland, Baltimore, USA (2011).
- J. Huang, S. Romero-Torres, and M. Moshgbar, *Practical Considerations in Data Pre-treatment for NIR and Raman Spectroscopy*, *American Pharm Review* (Sept./Oct. 2010).
- R. Kramer, *Chemometric Techniques for Quantitative Analysis*, Marcel Dekker, Inc., New York, USA
- E.R. Malinowski, *Factor Analysis in Chemistry*, 3rd Edition, Wiley-Interscience, New York, USA (2002).
- H. Mark, *Principles and Practice of Spectroscopic Calibration*, John Wiley & Sons, New York, USA (1991).
- H. Mark H and B. Campbell, *An introduction to near infrared spectroscopy and associated chemometrics*, The Near Infrared Research Corporation (2008).
- H. Martens and M. Martens, *Multivariate Analysis of Quality: An Introduction*, John Wiley & Sons, New York, USA (2001).
- H. Martens and T. Naes, *Multivariate Calibration*, John Wiley & Sons, New York, USA (1989).
- A. Jorgensen, *Clustering excipient near infrared spectra using different chemometric methods*. Technical Report, Dept. of Pharmacy, University of Helsinki, <http://www.pharmtech.helsinki.fi/seminaarit/vanhat/annajorgensen.pdf> (published on 18. September 2000).

- D.L. Massart, B.G.M. Vandeginste, L.M.C. Buydens, S. de Jong, P.J. Lewi, and J. Smeyers-Verbeke, *Handbook of Chemometrics and Qualimetric: Part A. Data Handling in Science and Technology: 20A*, 3rd impression, Elsevier, Amsterdam, The Netherlands (2003).
- E. Morgan, *Chemometrics: Experimental Design*, ACOL Series, John Wiley & Sons, New York, USA (1991).
- L. Meier, *Multivariate Statistics, Chapter 5: Multivariate Analysis of Spectra*, Seminar Script from ETHZ, Zurich, Switzerland, <http://stat.ethz.ch/~meier/teaching/cheming/cheming-e-part-III.pdf> (found online on 1 February 2013).
- L. Nørgaard, R. Bro, and S. Balling Engelsen, *Principal Component Analysis and Near Infrared Spectroscopy*, White paper from FOSS, Copenhagen, Denmark, <http://www.foss-analytical.co.in/industry-solution/feed-and-forage/papers> (found online on 1 February 2013).
- Andrew D. Richardson, James B. Reeves, and Timothy G. Gregoire, *Multivariate analyses of visible/near infrared (VIS/NIR) absorbance spectra reveal underlying spectral differences among dried, ground conifer needle samples from different growth environments*, *New Phytologist* **161** (2003) 291–301.
- R. Riovanto, *Near infrared spectroscopy in food analysis: qualitative and quantitative approaches*, Thesis, University of Padova, Padova, Italy (2011).
- Y. Roggo, P. Chaluz, L. Maurer, C. Lema-Martinez, A. Edmond, and N. Jent, *A review of near infrared spectroscopy and chemometrics in pharmaceutical technologies*, *Journal of Pharmaceutical and Biomedical Analysis* **44** (2007) 683–700.
- I. Sen, *Spectroscopic Determination of Major Nutrients (N, P, K) of Soil*, Thesis, Izmir Institute of Technology, Izmir, Turkey (2003).
- S. Turnbull, *The Use of Multivariate Analysis in Process Analytical Chemistry*, Thesis, University of Strathclyde, Glasgow, UK (2008).
- P. D. Wentzell and L. V. Montoto, *Comparison of principal components regression and partial least squares regression through generic simulations of complex mixtures*, *Chemometrics and Intelligent Laboratory Systems* **65** (2003) 257–279.

Process analytical chemistry

- K. A. Bakeev, *Near-Infrared Spectroscopy as a Process Analytical Tool, Part I: Laboratory Applications*, *Spectroscopy* **18(11)** (2003) 32–35.
- K. A. Bakeev, *Near-Infrared Spectroscopy as a Process Analytical Tool, Part II: At-line and On-line Applications and Implementation Strategies*, *Spectroscopy* **19(1)** (2004) 39–42.
- C.P. Beyers, *In-line Monitoring of Controlled Radical Copolymerisation Reactions with Near Infrared Spectroscopy*, Thesis, Eindhoven University of Technology, Eindhoven, The Netherlands (2003).
- J.M. Chalmers, *Spectroscopy in Process Analysis*, Sheffield Academic Press, Sheffield, UK (2000).
- F. McLennan and B.R. Kowalski, *Process Analytical Chemistry*, Blackie Academic & Professional, London, UK (1995).

www.metrohm-nirs.com



All rights reserved, including translation rights.
Printed in Switzerland by Metrohm Ltd., CH-9101 Herisau, Switzerland.
8.108.5026EN – 2013-02