Fluor detection on race ready skis

For a quick and reliable field adoptable method to find PFAS related substances on prepared skis the use of Fourier Transform InfraRed (FTIR) Spectroscopy has been utilized. For the calibration model, multivariate methods have been developed according to the guidance[1] from FDA.

Abstract

The developed system is based on two parts: 1. The FTIR – spectrometer with a multivariate QUANT model and slider system and 2. The controlling procedure.
The detection for banned fluorinated ski-waxes is based on well proven technologies using infrared- reflected light which results in a unique finger-print like detection of the substances presented to the instrument. By using a mathematical model (QUANT) built up by a training set with known concentration of substances, as well of other related compounds used in newly developed ski-waxes, a reliable evaluation model with no false positives or negatives can be established. The system indicates for the presence of banned substances on a level of no competitive advantage for the athletes.

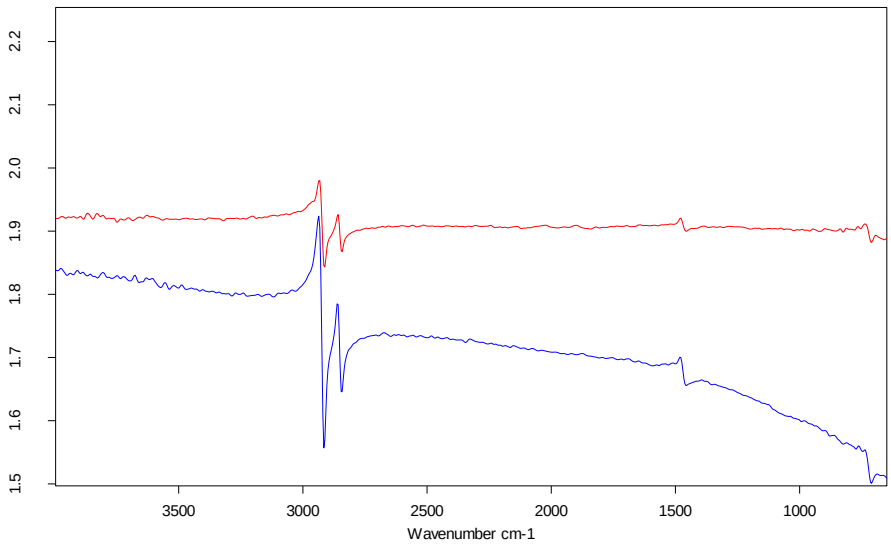
Experimental Design

To detect organic fluor substances several analytical methods can be used, some more sensitive than other, some more time consuming or expensive. However, in this case where with the following requirements apply; out-door use, easy handling for non-experience users, non-destructive measurements of the prepared ski-base , results within 1 minute and cost effective, the choices are very limited. The preferred choice ended up being FTIR reflection measurements.
During the trial phase of this project, we have also utilized: NMR, Mass-spectroscopy, XRF and other type of FTIR techniques as Attenuated Total reflection (ATR). This was done to check if the techniques were applicable for this application.
NMR and mass-spectroscopy fulfils the requirement that it can detect the substances, even at a very low level but apart from the that no other criteria were fulfilled. XRF indicate the presence of fluor as an element “F”. F is the 13th most common element on earth and can be found natural in many minerals/salts as well in such common things as toothpaste. Therefore, an indication of presence of the element F is not enough in this application.
The FTIR-ATR technique, which is more sensitive and more accurate than FTIR reflection, require optical contact between the matter of interest and the optical crystal. To get such good contact, a relatively high pressure must be applied to ensure a good spectrum. That is possible with ski-wax blocks or for testing a cut-out ski-base sample, but to measure a race prepared ski where no interference with the prepared ski base, ski wax or structure is allowed, turned out not to be possible. Even allowing disruption of the surfaces, the problems arise with applying a good pressure on the ski to get the required contact. The variation in contact created non reproduceable results, even using one single well-trained operator.

An other part of the project was to build a library of different ski-waxes as well of cleaning liquids to ensure that the products used by the different ski-wax teams are indeed free from organic fluor. We have several cases where products are claimed to be either fluor free or containing fluor, but after further analyses we could show that it was not correct.

QUANT

The calibration model is based on both reference ski-waxes with added known amount of fluor compounds and clear fluor free (FF) products. All commercial FF products used have been separately tested for fluor to ensure that the spectra used to build up the model are correct.
During this work, we have notice that the fluor compounds sometimes create a very inhomogeneous blend in the ski-waxes block, even in one single block we clearly could find concentration variations, this in addition to the process of preparing the ski which also further give an un-even distribution of the fluor compounds along the ski. To mitigate this variation in the samples, we have prepared several skis for each reference wax and measured them all on several spots, to ensure that we build in this variation in the model. This variation can be attributed as the single major contributor to the error of prediction. Furthermore, ski-wax thickness, ski base structures and the addition of non-reflecting material in ski-waxes and ski bases will also influence the final signal with variated baseline slop, lower intensity etc, i.e reflecting more physical than chemical properties. However, with mathematical modelling including normalization of the spectra, these contributions to the final prediction of a value, were supressed to a minimum.



Effect of automatic pre-treatment including vector normalization of spectra before final evaluation.

The given prediction value is based on the given concentration of our reference's samples. When the predicting model is set the absolute error of the model can be calculated. All the involved errors described above contribute to this error. The absolute error + a tolerance/safety margin adds to the level where the threshold is set. This threshold has to be at a level where no competitive advantage are present in ordered to full fill the defined requirements. When the measured value is higher than the threshold a warning from the instrument for a suspected banned substance will occu

The development of the QUANT model is under constant development as new ski preparations emerge. Thus the release of new models will occur as new technologies have been added to the model. Special test sessions will be conducted where the wax manufacturers can test their range of FF products and in addition train the model.

FTIR Spectroscopy

Infrared spectroscopy (**IR spectroscopy** or **vibrational spectroscopy**) is the measurement of the interaction of infrared radiation with matter by absorption, emission, or reflection. It is used to study and identify chemical substances or functional groups in solid, liquid, or gaseous forms. It can be used to characterize new materials or identify and verify known and unknown samples. The method or technique of infrared spectroscopy is conducted with an instrument called an infrared spectrometer (or spectrophotometer) which produces an infrared spectrum. An IR spectrum can be visualized in a graph of infrared light absorbance (or transmittance) on the vertical axis vs. frequency, wavenumber or wavelength on the horizontal axis. Typical units of wavenumber used in IR spectra are reciprocal centimeters, with the symbol cm^{-1}. Units of IR wavelength are commonly given in micrometers (formerly called "microns"), symbol μm, which are related to the wavenumber in a reciprocal way. A common laboratory instrument that uses this technique is a Fourier transform infrared (FTIR) spectrometer. IR spectroscopy is often used to identify structures because functional groups give rise to characteristic bands both in terms of intensity and position (frequency).

Infrared spectroscopy is a simple and reliable technique widely used in both organic and inorganic chemistry, in research and industry. It is used in quality control, dynamic measurement, and monitoring applications such as the long-term unattended measurement of CO_2 concentrations in greenhouses and growth chambers by infrared gas analyzers.

It is also used in forensic analysis in both criminal and civil cases, for example in identifying polymer degradation. It can be used in determining the blood alcohol content of a suspected drunk driver.

IR-spectroscopy has been successfully used in analysis and identification of pigments in paintings and other art objects such as illuminated manuscripts.

With increasing technology in computer filtering and manipulation of the results, samples in solution can now be measured accurately (water produces a broad

absorbance across the range of interest, and thus renders the spectra unreadable without this computer treatment).

Some instruments also automatically identify the substance being measured from a store of thousands of reference spectra held in storage.

Infrared spectroscopy is also useful in measuring the degree of polymerization in polymer manufacture. Changes in the character or quantity of a particular bond are assessed by measuring at a specific frequency over time.

Another important application of Infrared Spectroscopy is in the food industry to measure the concentration of various compounds in different food products.

Infrared spectroscopy is an important analysis method in the recycling process of household waste plastics, and a convenient stand-off method to sort plastic of different polymers (PET, HDPE, ...).

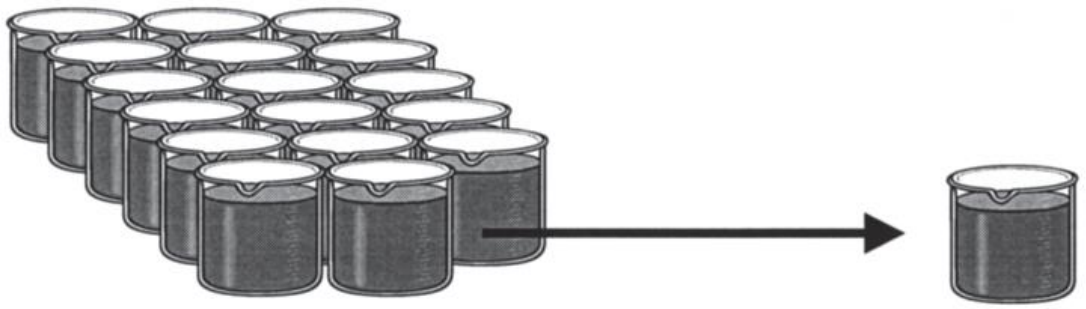
The instruments are now small, and can be transported, even for use in field trials.

Chemometric Models and their Validation

The purpose of QUANT is the quantitative analysis of an unknown multicomponent sample.
However, in order to perform an analysis, QUANT first has to “learn” about your system.
This means you have to develop a chemometric model, using a number of calibration samples of known composition that are representative for your system. The IR spectra of these samples will be used by QUANT to calculate a calibration function, which essentially is the model used for the analysis of unknown samples later. However,
the model has to be evaluated to test its reliability of prediction (validation).
There are two validation types: “Cross Validation” and “Test Set Validation”. While in the
latter case two different sets of samples are used, the Cross Validation uses the same set of samples for calibration and validation.

Cross Validation
Only one set of samples representative for your multicomponent system is used to calibrate
and validate your system. **Before** starting the calibration, one sample is excluded from the entity of samples. This sample is used for the validation. The remaining samples
are used to calibrate the system. **The sample used for validating the system must not be part of the calibration set.** Here is an example to illustrate this point: let’s

say you choose 100 samples of a known composition. From these samples you take sample number 67 and set it aside. The remaining 99 samples now make up your calibration
set and you will use them to create a chemometric model. After doing this you will test this model against sample 67. Then you repeat this cycle, this time separating a
different sample (e. g. #17) and so on, until all samples have been used for validation once. QUANT reiterates this cycle, starting with the first sample, until all samples have been used for validation.
The advantage of cross validation is the smaller number of samples required. Especially,
if the number of samples available is limited this method should be preferred upon the test set validation.



Calibration Set: Test Sample:

Figure 1: Cross Validation

Choosing Calibration Samples

The first step of building a chemometric model is to pick a sufficiently large number of samples to represent your system. These samples have to be quantitatively analyzed by
a reliable method to determine their components. Then the IR spectra of all samples are
taken and, depending on the type of validation method used, a calibration set and a test
set is formed of these spectra.
The following rules should be observed when forming a calibration set:
• No general recommendation can be given concerning the number of samples in a calibration set. As a rule of thumb, for a one component system a minimum of 20 samples should be measured. Multicomponent systems require a larger number of calibration samples.

Note: For setting up a calibration model using OPUS/QUANT you can use up to 60000

spectra maximum.

• Choose your calibration samples in a way they cover a wider concentration range than you intend to analyze later. This helps to create a more stable model for analysis.
This becomes increasingly important if you expect outliers, with concentrations that largely deviate from your desired values, as this may be the case in quality control.
• The calibration samples should be spaced homogeneously across the concentration range. Do not include samples with concentrations well apart from the concentration field the majority of your samples span. In case you need to extend the concentration range, include a larger number of samples, so that the resulting range still retains the sample density.

• Do not try to correct external fluctuations, as this will be mirrored as concentration fluctuations in your samples. These fluctuations will be recognized as such by QUANT and accounted for in the calibration function. This will yield a more robust model. Keep in mind that an extensive sample preconditioning of the calibration samples will have to be repeated later for every sample to be analyzed.
Never try to account for deviations in the calibration set you can not correct for the samples you want to analyze. Rather increase the number of samples included in your calibration set.
• If your process conditions change later, there is no need to repeat the calibration, because the perturbations will be “filtered” by the PLS 1 algorithm. If your concentration
range expands in the future, simply add a sufficient number of samples to the calibration set, covering the new wider range.
• In case you prepare the samples for your calibration set in the lab, make sure that these samples show no collinearity, which means that they do not show a linear de- or increase in concentration of the components. Especially dilution series are not suited as calibration samples.
.

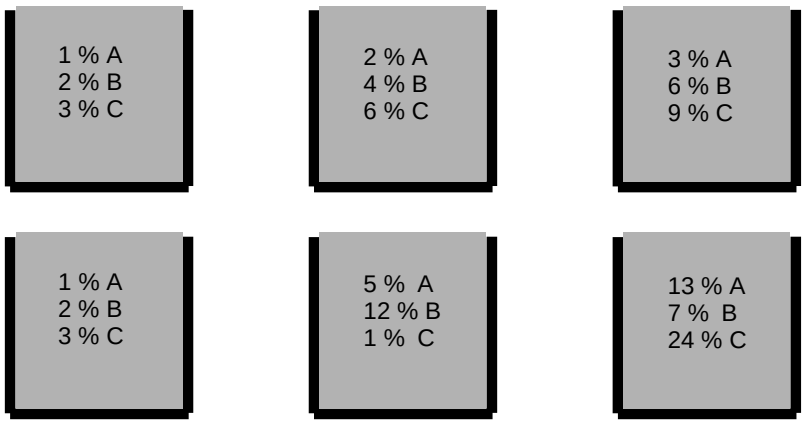


Figure 2: Example of collinear Samples and Samples showing no Collinearity

• When acquiring spectra from the calibration set, never measure the samples in increasing or decreasing order of their concentration. Otherwise, linear fluctuations in temperature (heating up or cooling of the samples) or concentration (evaporation of solvent) will not be recognized by the PLS 1 algorithm. If possible, repeat the measurements at a later point in time.
• Ensure that the reference method you use for the determination of the components concentration yields reliable results. Repeat these measurements to obtain statistical significance. Be sure to know the statistical error of your reference method

Acquiring Spectra and Data Preprocessing

After you have chosen a set of samples you need to acquire their IR spectra. Check the reproducibility of the measurements, for short and long time intervals, using a few test samples first. **Make sure to use the same parameter set during the measurements of the calibration set that you later want to use for the analysis.**
Now that you have all spectra at your hand, you should decide on whether you want to use the whole frequency region of the data and whether you want to perform some data preprocessing before starting the QUANT software.

Frequency Region

The PLS regression method is a “full spectrum method”; the chemometric model should improve with an increasing number of data points. However, in some cases spectral noise or additional components in the samples may cause the PLS algorithm to interpret these features, which can degrade the model. In these cases, as in our case, it is advisable to limit the frequency region used for the PLS regression. Usually this step is taken to improve a regression that did not yield a satisfactory model. When narrowing down a spectrum to a few absorption bands it is found, that in general bands between 0.7 and 1.0 absorbance units (AU) generate the best results. Values greater than 2.5 should not be used. Also, it is not necessary to identify substance specific peaks, but rather to include the complete frequency region of the functional groups (e. g. alcohols) from a spectrum. Nevertheless, in case of a minor component, it can be helpful to know the absorptions in the spectrum to find relevant frequency regions.

Data Preprocessing

Data preprocessing is an important stage in performing a calibration. To ensure the reproducibility of the calibration samples, several spectra of each sample have to be acquired. If the spectra of the same sample are not identical, a data preprocessing procedurmust be chosen to bring them into line with each other. Data preprocessing can
eliminate variations in offset or different linear baselines.
In quantitative analysis, it is assumed that the layer thickness (i. e. the effective pathlength
of the infrared light in the sample) is identical in all measurements. A lack of reproducibility in sample preparation can easily cause variations in sample thickness. If
the thicknesses are different or unknown, this effect can be eliminated by a normalization
of the spectra. The purpose of data preprocessing is to ensure a good correlation between the spectral data and the concentration values. The following methods can be applied:

• *Linear Offset Subtraction:* shifts the spectra in order to set the y-minimum to zero.
• *Straight Line Subtraction:* fits a straight line to the spectrum and subtracts it. This accounts for a tilt in the recorded spectrum.
• *Vector Normalization:* normalizes a spectrum by first calculating the average intensity value and subsequent subtraction of this value from the spectrum. Then the sum of the squared intensities is calculated and the spectrum is divided by the square root of this sum. This method is used to account for different samples thickness, for example.
Min-max Normalization: first subtracts a linear offset and then sets the y-maximum to a value of 2 by multiplication with a constant. Used similar to the vector normalization.
• *Multiplicative Scatter Correction*: performs a linear transformation of each spectrum for it to best match the mean spectrum of the whole set. This method is often used for spectra measured in diffuse reflection.
• *First Derivative*: calculates the first derivative of the spectrum. This method emphasizes steep edges of a peak. It is used to emphasize pronounced, but small features over a broad background. Spectral noise is also enhanced.
• *Second Derivative*: similar to the first derivative, but with a more drastic result.
No general recommendation can be given whether a given data set should be preprocessed
or which method is suited best for it. Therefore, the optimal data preprocessing method can only be found empirically by applying several methods to your spectral data
and comparing the results.

Validating the Model

At this point the model needs to be validated. If a sufficient number of samples have been measured, it is possible to divide the samples into two sets of about equal number,
a calibration set and a test set. The calibration set is used to build up a model which is

then tested with the test set. This procedure is called test set validation. The distribution
of the concentration values should be similar for both sets. A test set validation requires
less computational time than a cross validation.
If only a limited number of samples is available, use a cross validation (see above). To perform a good cross validation the number of spectra per sample should be equal for all
calibration standards.

➢ **Important:** Repetitive spectra of one sample must be assigned as “one sample”!

A matrix is formed from the spectral data of the calibration set. The matrix will be transformedby the PLS 1 algorithm into a result matrix consisting of eigenvectors (factors)
only, as mentioned above. These factors are sorted in decreasing order according to their contribution to the spectral features. Factors which present a large contribution to the spectrum are found in the top rows of the matrix, while factors listed towards the bottom
rows mainly reflect spectral noise and fluctuations. Thus not all factors are needed to explain the spectral features of the components (the contributions representing noise
can be omitted). The quality of the chemometric model now depends on the choice of the correct number of factors needed; this is also called the rank of the model. Choosing
a too small rank results in underfitting so that not all features can be explained by the model. On the other hand, including too many factors (rank too high) leads to overfitting
and only adds noise, in fact degrades the model.

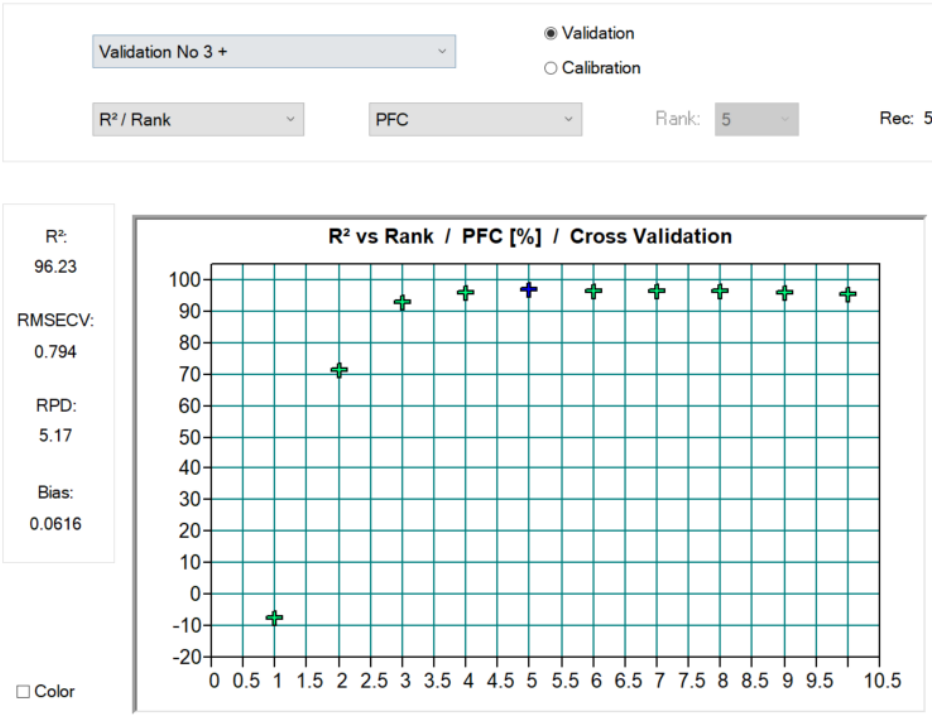


Figure 3:

As a consequence, there is an optimum number of factors for every system, i. e. an optimum
rank. A criteria for determining the optimum rank is to look at the **r**oot **m**ean **s**quare **e**rror of **p**rediction resulting from an analysis of the cross validation. If the RMSEP is depicted against the rank used in each model, a minimum can be observed in this graph, indicating the optimum rank.

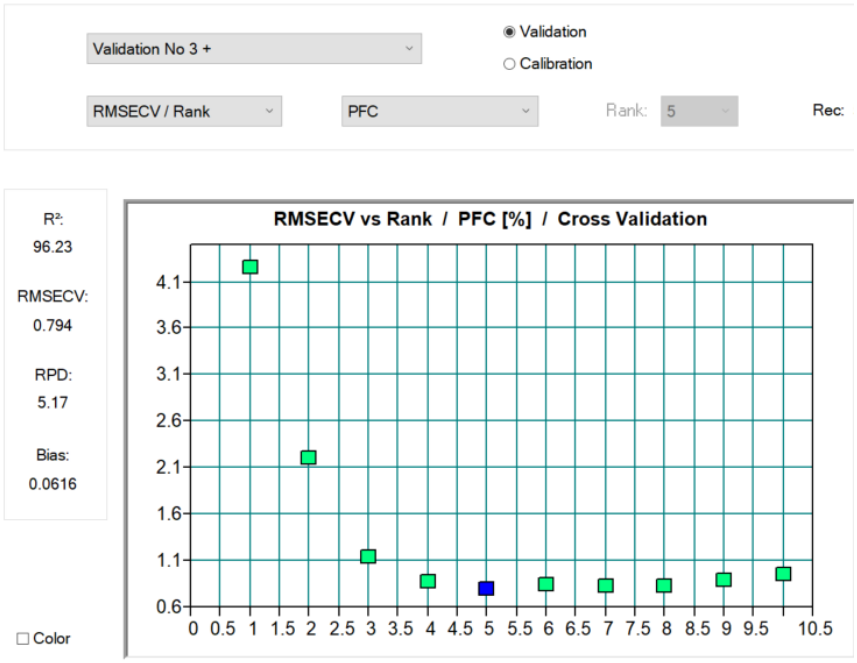


Figure 4: Display of RMSECV against the rank

As the quality of the model improves, it becomes increasingly difficult to distinguish the errors of prediction judging from these plots only. A better way of determining the optimum
rank is plotting the RMSECV values versus the rank. Switch to the *RMSECV/Rank* plot. Apparently, the model improves drastically up to the rank 5. However, ranks higher
than 6 barely improve the model and basically represent the addition of fluctuations (noise, emperature differences of the samples etc.) which, in fact, eventually leads to a degradation
of the result. It also becomes clear that a calculation up to rank 10 would have been sufficient to determine the optimum rank. Restricting the calculation to lower ranks saves processing time as
the calibration set contains more samples

Figure 5 shows the diagrammatic representation of the validation result. By default, the predicted concentration values versus the true concentration values are displayed. Outliers are marked in red. The recommended rank *Rec.* is in our case *5*. The results of the predicted concentration values are displayed for this rank, but the display can be changed by selecting a different rank in the *Rank* drop-down list. In addition, the name of the validation, the component for which the result is shown, as well as the values for RMSECV (**r**oot **m**ean **s**quare **e**rror of **c**ross **v**alidation**)** and R2 (coefficient of determination) are displayed.

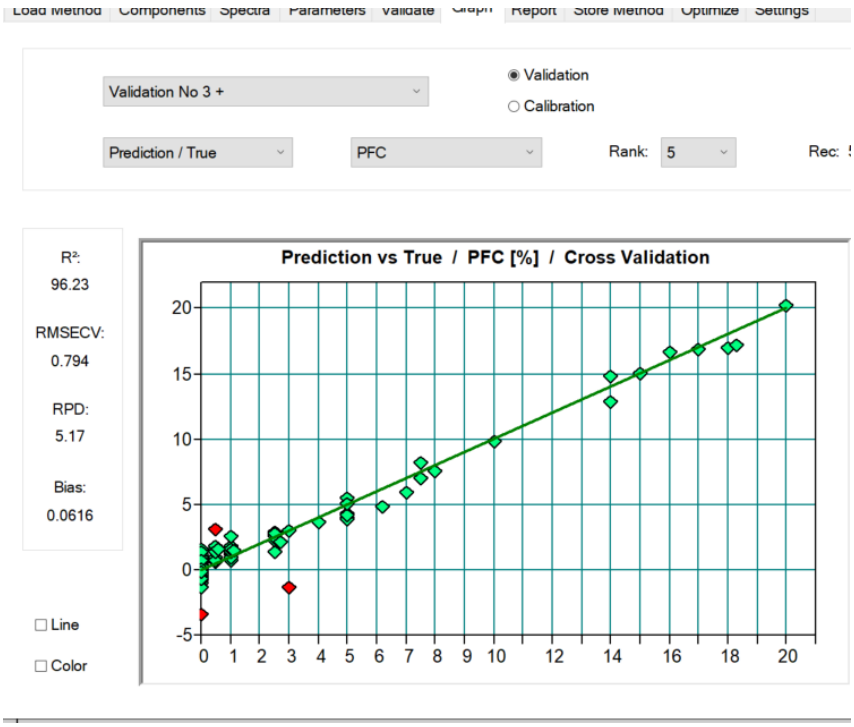


Figure 5: Display of predicted values against the true values

Anders Nilsson Ph.D.
Bruker Nordic AB

Board member of the Chemometric Division, Swedish Chemical Society
Board member of the Vibrational Spectroscopic Division, Swedish Chemical Society

Referenses

[1] *Development and Submission of Near Infrared Analytical Procedures, Guidance for Industry, U.S. Department of Health and Human Services. August 2021*