

FT-NIR analysis of Czech Republic beer: A qualitative and quantitative approach

Authors

Gabriela Budinova, Ph.D., Quintiles GesmbH, Czech Republic, Europe
Ivor Dominak, Ph.D., NICODOM Ltd, Czech Republic, Europe
Todd Strother, Ph.D., Thermo Fisher Scientific, Madison, WI, USA

Keywords

Antaris, beer, density, fermentation, NIR



Figure 2. Thermo Scientific Antaris II MDS system. This system includes a transmission cell used for analyzing liquids such as the beer samples used in the present application.

Background

Beer is one of the oldest known beverages commonly found throughout many cultures of the world with evidence of its production dating back over 7000 years. One of the chief advantages of beer in antiquity was its relative health benefits over water resources. This is because the alcohol content was high enough to preclude the proliferation of harmful microorganisms. Currently, beer production is a multi-billion dollar industry involving large multinational corporations worldwide. As shown in Figure 1, the largest per capita consumption of beer lies with the Czech Republic where almost 160 liters is consumed annually. This compares to the per capita consumption of Ireland (131 liters), Germany (116 liters) or Australia (110 liters). The Czechs are clearly leaders in the production and consumption of beer.

Czech beer is often classified as 10°, 11°, or 12°. This notation refers to the amount of sugars present in the liquid extract (wort) prior to fermentation that may later be transformed into alcohol by yeast. Thus a 12° beer will have more sugars and subsequently more alcohol than a 10° beer but will not contain 12% alcohol. Analysis of beer is required for proper reproducible production and labeling. Conventionally, a separate instrument is required for the analysis of each component in beer. For example, analysis of alcohol content may be performed using gas chromatography, while various specific gravity measurements require a hydrometer. Additionally, certain parameters require sample preparation such as extended boiling, which

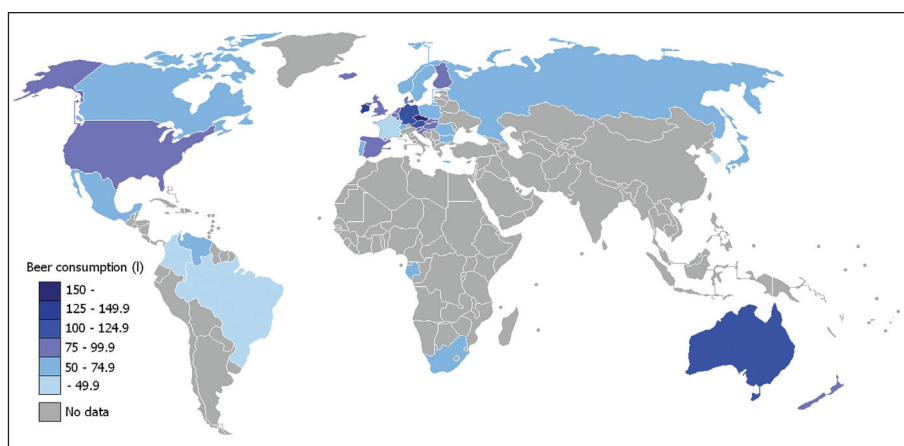


Figure 1: Graphic showing worldwide annual per capita consumption of beer

can take a substantial amount of effort and time. In contrast to these other methods, Fourier transform near-infrared (FT-NIR) spectroscopy is a much more rapid technique and was performed on beer samples from Czech Republic breweries to demonstrate its capacity to rapidly and accurately measure and predict multiple components simultaneously.

The Thermo Scientific™ Antaris™ II FT-NIR analyzer is well suited and is commonly used for the analysis of complex solutions. The Antaris II Method Development Sampling (MDS) system, shown in Figure 2, has a transmission cell often used with relatively clear liquids such as beer as well as an integrating sphere and fiber optic probe useful for obtaining diffuse reflectance spectra from solids, powders, and opaque liquids. The Antaris line of instruments collects and analyzes near infrared spectra between 12,000 and 4000 cm^{-1} . Light in this region interacts with the sample where absorption related to molecular vibration takes place. The remaining light that is not absorbed by the vibrations in the sample is collected at the detector and displayed for analysis. Near-infrared instruments can be coupled to fiber optic cables, can analyze multiple components in a mixture with a single spectrum, and can analyze materials directly without diluting or modifying them. These key features prove near-infrared analysis is a superior technique over other methods.

Methods

Near-infrared transmission spectra were collected for 86 degassed beer samples from Czech Republic breweries using a 2 mm transmission cell. Thirty eight samples were 10°, ten samples were 11°, and thirty six samples were 12°. This information was used to initially determine if the samples could be qualitatively classified as 10°, 11°, or 12°. The spectra were then analyzed quantitatively for alcohol, original gravity, real extract and apparent extract. Original gravity refers to the specific gravity of the wort prior to fermentation. Real extract is measured by gently boiling the beer sample to drive off the alcohol, adding distilled water back to make up the original volume, and then taking the specific gravity of the reconstituted

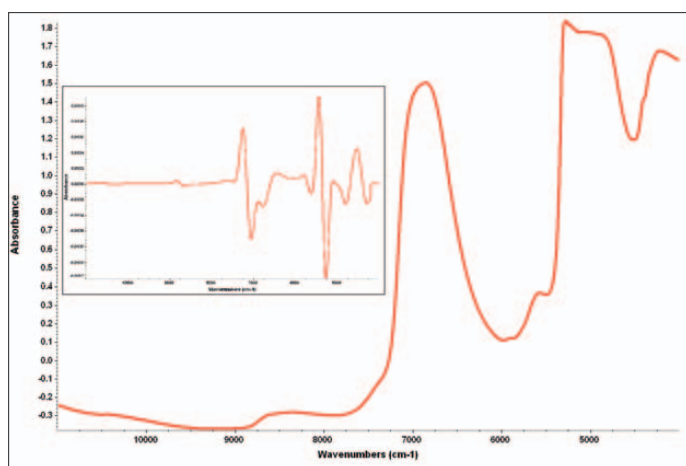


Figure 3. Representative spectrum from beer sample and second derivative of that spectrum (inset).

| | Qualitative classification | Quantitative analysis |
|-------------------------------------|--------------------------------------|--|
| Chemometric method | Discriminant analysis | Partial least squares |
| Spectral regions of analysis | 10,000–5400 cm^{-1} | 10,000–5400 cm^{-1} 4700–4100 cm^{-1} |
| Spectral treatment | Second derivative | None |
| Smoothing | Norris derivative (segment 9, gap 9) | None |
| Baseline correction | None | Average between 9000–9400 cm^{-1} |

Table 1. Summary of the parameters used for the analysis of Czech beer samples.

sample. This contrasts with apparent extract, which is the specific gravity of the sample without first removing the alcohol. The alcohol content ranged from 2.9% to 4.84% as measured by gas chromatography; the original gravity was converted to degrees Plato and ranged from 9.77 to 12.28; the real extract and apparent extract were also converted to degrees Plato and ranged from 2.77 to 5.42 and from 0.97 to 3.91 respectively. These four quantitative parameters and the qualitative classification were predicted using a single NIR spectrum for each of the samples. Data collection time for the qualitative and quantitative predictions of a single sample was 20 seconds.

Spectra were collected between 12,000 and 4000 cm^{-1} for both the qualitative and quantitative analysis. The spectra were signal averaged from 32 scans with a resolution of 16 cm^{-1} . For the qualitative determination, the spectra were treated to Discriminant Analysis using the second derivative in the range between 10,000 and 5400 cm^{-1} .

For the quantitative measurements, the four components were analyzed using a Partial Least Squares (PLS) method. Each component of this quantitative method was analyzed using two regions of the spectra; from 10,000 to 5400 cm^{-1} and from 4700 to 4100 cm^{-1} . These regions avoided the all absorbing water peak centered around 5000 cm^{-1} . The four components were analyzed using the raw spectra with a one point baseline correction. Figure 3 shows a representative spectrum used for the quantitative analysis and the second derivative spectrum (inset) used for the qualitative classification. Table 1 summarizes the parameters used for both the qualitative and quantitative analysis.

Qualitative classification results

The discriminant analysis method developed resulted in proper separation and classification of the samples based on degree type (e.g. 10°, 11°, or 12°). Figure 4 shows an example of the Mahalanobis distance plots for the calibration. These plots show the Mahalanobis distances of the individual spectra from the different classes. For the example plot shown, all of the 10° samples are well grouped together and each falls closer to the center of the 10° degree group than the center of the

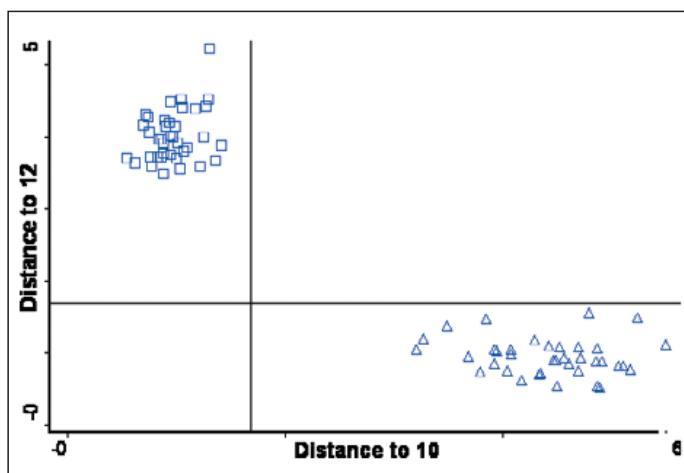


Figure 4. Example of a Mahalanobis distance plot used in the Discriminant Analysis. Note there is clear separation of the different sample groups. In this example plot the triangles (Δ) represent 12° samples; the rectangles (\square) represent 10° samples.

12° group. This is indicated by the rectangles (\square) appearing closer to the X-axis origin (center of the 10° group) than the Y-axis origin (center of the 12° group). Likewise the triangles (Δ), representing the 12° spectra, fall closer to the Y-axis origin than the X-axis origin. For all similar Mahalanobis distance plots in the analysis, the individual spectra yielded correct classification. Figure 5 lists a sampling of the 86 different spectra with their actual and calculated classes. The calculated classes were determined by assigning the spectra to the lowest Mahalanobis distance from the three choices. Note that all of the spectra are properly assigned to their actual class and no errors or misclassification occurred.

Quantitative analysis results

The samples were also analyzed using a PLS method to predict the alcohol content, original gravity, and real and apparent extracts. Figure 6 graphically demonstrates the degree of correlation between the chemometric method's calculated values and the actual values for the four components. In each case there is high correlation and low error. The Root Mean Square Error of Calibration (RMSEC) values were calculated from the samples used in the chemometric model. The Root Mean Square Error of Prediction (RMSEP) values were determined by using independent validation samples not used in building the model. Finally, the Root Mean Square Error of Cross-Validation (RMSECV) values were determined by removing four samples from the chemometric model, calculating the error of those samples from the model, and then repeating the calculation for new sets of four samples until all of the standards had been measured. Note that the RMSEC, RMSEP and RMSECV for each of the components are similar in magnitude. This characteristic indicates the chemometric model is robust, is not overfitting the data and can accurately predict unknown samples.

| File Name | Actual Class | Calculated Class | Distance | Next Distance |
|-----------|--------------|------------------|----------|---------------|
| 743.jdx | 10 | 10 | 1.2170 | 2.4048 |
| 744.jdx | 10 | 10 | 0.9898 | 3.1836 |
| 745.jdx | 10 | 10 | 1.1683 | 4.1900 |
| 746.jdx | 10 | 10 | 0.8338 | 3.2018 |
| 747.jdx | 10 | 10 | 0.8837 | 2.7484 |
| 748.jdx | 10 | 10 | 0.6236 | 2.8279 |
| 749.jdx | 10 | 10 | 0.8839 | 3.2659 |
| 750.jdx | 10 | 10 | 1.2880 | 3.8033 |
| 822.jdx | 11 | 11 | 0.9534 | 2.4676 |
| 823.jdx | 11 | 11 | 0.9531 | 2.1132 |
| 824.jdx | 11 | 11 | 0.9537 | 2.2653 |
| 825.jdx | 11 | 11 | 0.9541 | 1.9792 |
| 826.jdx | 11 | 11 | 0.9521 | 2.0769 |
| 827.jdx | 11 | 11 | 0.9515 | 2.1336 |
| 828.jdx | 11 | 11 | 0.9534 | 1.8494 |
| 829.jdx | 11 | 11 | 0.9535 | 1.9138 |
| 783.jdx | 12 | 12 | 1.2036 | 2.1646 |
| 784.jdx | 12 | 12 | 0.9830 | 3.6879 |
| 785.jdx | 12 | 12 | 0.7321 | 2.8764 |
| 786.jdx | 12 | 12 | 1.0283 | 3.2912 |
| 787.jdx | 12 | 12 | 1.0432 | 3.6340 |
| 788.jdx | 12 | 12 | 0.9441 | 3.6823 |
| 789.jdx | 12 | 12 | 1.0473 | 3.2080 |
| 790.jdx | 12 | 12 | 0.8484 | 3.9200 |

Figure 5. Representative summary of the classification of the beer samples. All of the samples in the entire data set were correctly classified according to degree of fermentable sugars. The Distance column indicates the sample's Mahalanobis distance to the center of the nearest group. The Next Distance column indicates how close the sample is to the center of the next nearest group.

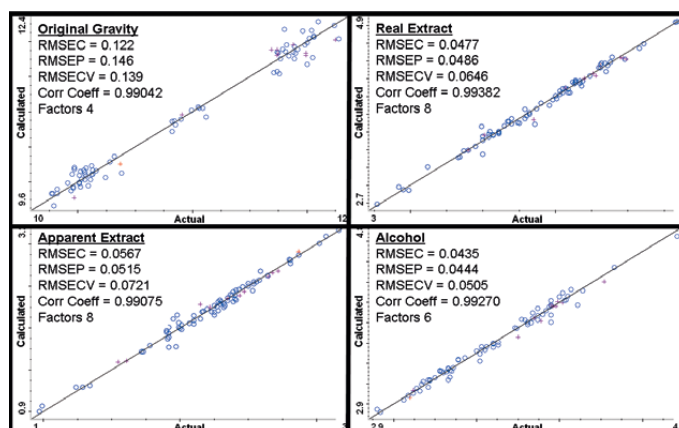


Figure 6. Correlation plots for the four components indicating close agreement between the chemometric model's calculated values and the actual sample values.

Predicted Residual Error Sum of Squares (PRESS) plots were also generated from the cross-validation metric. An indication of the validity of a chemometric model lies in the shape of the PRESS plot. PRESS plots that show a rapid decrease to a minimum and then stabilization are indicative of a robust chemometric model, suitable for accurate prediction. Figure 7 shows the PRESS plots for the four components measured in the beer samples which follow the expected pattern for an acceptable chemometric model.

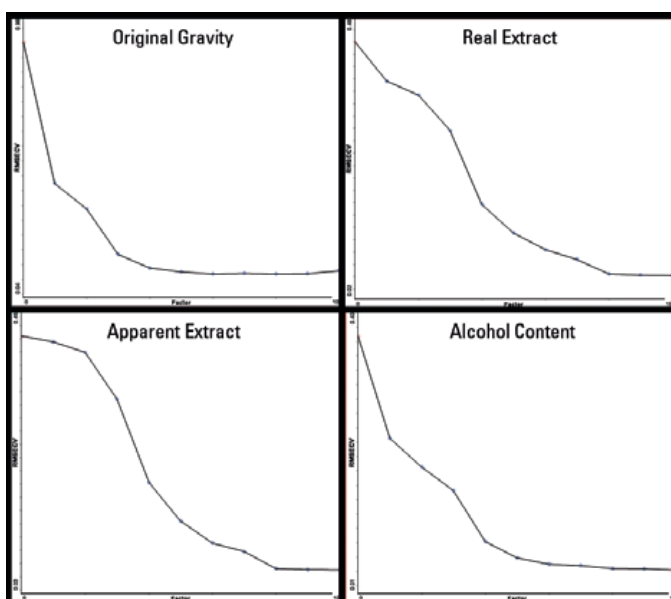


Figure 7. PRESS plots of the four components measured in the beer samples. The PRESS plots follow the expected behavior of a robust chemometric model.

Conclusions

NIR analysis of a variety of beers from the Czech Republic with the Antaris FT-NIR Analyzer successfully predicted both qualitative as well as quantitative parameters from a single data collection taking 20 seconds. The samples were subjected to Discriminant Analysis and correctly classified according to degree of fermentable sugars (i.e. 10°, 11° or 12°). Additionally, multicomponent quantitative analysis was performed on the samples using a PLS model. The four components: alcohol content, original gravity, real extract, and apparent extract were correctly predicted with correlation coefficients better than 0.990 and low root mean square errors indicating the model is appropriate and robust. Multiple component analysis was demonstrated with the Antaris FT-NIR much more rapidly than conventional techniques and without the need for sample preparation.

Learn more at thermofisher.com/nirfood

thermoscientific