APPLICATION NOTE

No. AN52536

Acid and alcohol nir analysis as examples of food process control

Keywords
Acetic acid, alcohol, biological process, near-infrared, vinegar, real-time QC

Introduction

Vinegar is a widely used material that has been produced since antiquity. Babylonians are thought to have used it as a preservative 7000 years ago; while evidence exists of vinegar in 5000 year old Egyptian urns. While it historically has been used as a food preservative, condiment, and cleaning agent, it has more recently been used in the medical or veterinary fields as well as an organic herbicide. Currently U.S. supermarket sales of bottled vinegar top $220 million per year, which doesn't include vinegar found in salad dressings, other food stuffs, or chemical stocks. In addition to common distilled white vinegar, there is a great deal of growth in the introduction of specialty and flavored vinegars.

Thermo Scientific™ Antaris™ MX FT-NIR analyzer

Vinegar production is a biologically controlled process where specific bacterial cultures convert alcohol, commonly from wine feedstocks, to the end product; acetic acid. The oxidation of the alcohol into acetic acid requires the presence of oxygen which is distinct from anaerobic fermentation. Bacteria used for this process are typically Acetobacters, which first convert the alcohol to acetaldehyde then further oxidize it to acetic acid. The chemical reaction for this process is given in Figure 1. Vinegar manufacturers, both large scale and small scale specialty producers, use a material called "mother of vinegar." This is a naturally forming gel consisting of bacteria and cellulose and is used to inoculate each batch of feed stocks. This material is replenished from previous batches and can potentially be used indefinitely.

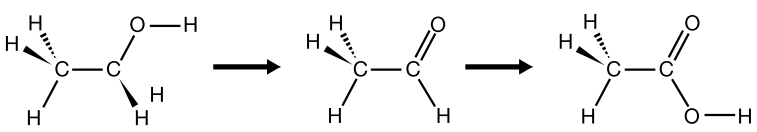


Figure 1: Oxidation of alcohol to acetic acid. This reaction is accomplished by the Acetobacter bacteria used in the production process.

Since it is a biological process, vinegar production is sometimes difficult to control. Often there is a great deal of variability from one bacterial culture to the next, which results in batch-to-batch non-uniformity. Additionally, the concentrations of acetic acid and alcohol are not always accurately predicted at any given time point in the process. Therefore samples from each lot are continually withdrawn and tested to monitor the progress of the oxidation process. These samples are tested for both acid content and alcohol residue. Alcohol content is desired to be very low in the final product and acid content is regulated by law to be above 5%. Unfortunately, laboratory tests for both acetic acid and alcohol require skilled technicians, hazardous chemicals, and specialized equipment. Acid content is typically performed using titration techniques where a sample is neutralized to an endpoint with a solution of sodium hydroxide.

ThermoFisher SCIENTIFIC

In addition to being extremely caustic, sodium hydroxide solutions are unstable and are required to be freshly prepared and standardized prior to use. Additionally, since the endpoint is determined colorimetrically, there is potentially a great deal of variability from technician to technician. Alcohol content, on the other hand, is determined using gas chromatography, which requires upkeep and a highly trained technician. Because of the shortcomings in current testing methods, a novel method of vinegar analysis using near-infrared spectroscopy (NIR) was explored.

Near-infrared spectroscopy uses the part of the electromagnetic spectrum between 10,000 and 4,000 cm^{-1} to measure absorbances caused by molecular vibrations. Molecules vibrate characteristically when they absorb light in the near- and mid-infrared regions; and that pattern of absorption can be used to identify the presence and concentration of specific chemicals. Fourier transform near-infrared (FT-NIR) instruments are particularly well suited for process analysis because they provide highly accurate spectra very rapidly with excellent resolution. Spectra can be obtained and interpreted with common computers and software in less than ten seconds. The Thermo Scientific Antaris MX system is a process-based FT-NIR analyzer that can easily be integrated into production streams or process tanks using probes and fiber optic cables. This system allows real-time monitoring of a process without the need to withdraw samples or analyze them in a laboratory.

Experimental

To validate the performance of the Antaris system for analyzing alcohol and acid content in vinegar, NIR spectra from multiple samples were collected. The acetic acid and alcohol content from these samples was verified using the primary techniques previously described above. The spectra were then used as a training set to build a chemometric model using the Thermo Scientific™ TQ Analyst™ Software package. The alcohol and acid content predicted by the TQ Analyst method were then compared to the results from the primary analysis methods to validate the NIR technique.

Eighty vinegar samples from various stages of production were analyzed and used in the development of the chemometric model. The acid content in the samples ranged from 4.65% to 10.70%, while the alcohol content ranged from 0.28% to 3.47%. The samples were scanned using transmission with a known constant pathlength.

The sample spectra were used without smoothing or taking derivatives. Two regions of the spectra were used in the method; 5100 to 4800 cm^{-1} and 7700 to 5400 cm^{-1}. These regions excluded the all absorbing water band found around 5000 cm^{-1}. A two point baseline correction was used to remove baseline shifts and other affects. The chemometric model used 65 samples for calibration, while the remaining 15 samples were used to validate the method. Representative spectra are shown in Figure 2.

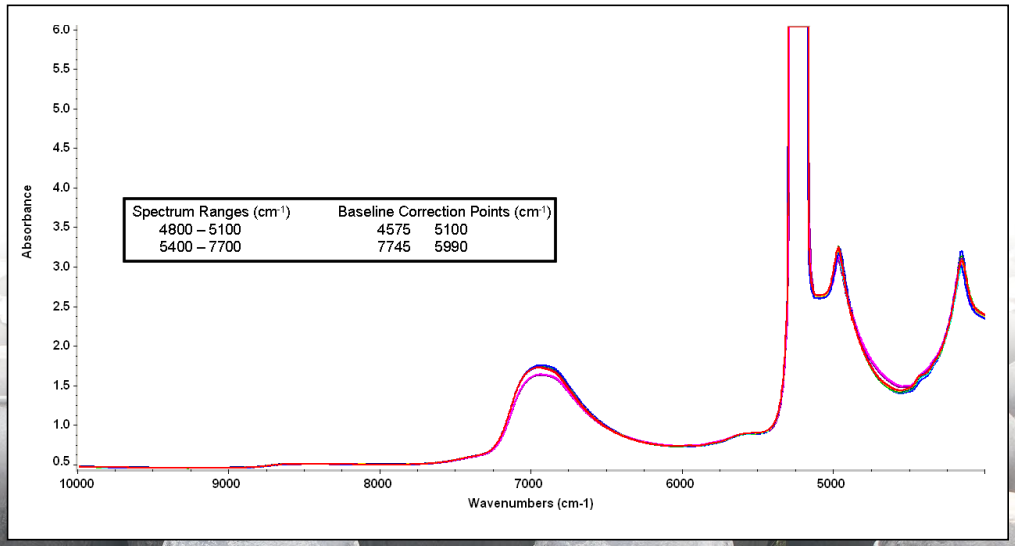


Figure 2: Representative spectra of vinegar samples. Spectra were analyzed in the regions indicated using two point baseline corrections without smoothing. The regions used avoid the all absorbing water peak at approximately 5000 cm^{-1}.

When developing a chemometric model on material with more than one component it is vital that the concentrations of those components are randomly distributed. If the concentrations of the components increase or decrease together, any spectroscopic variation between the standards cannot be attributed specifically to a particular component. TQ Analyst software uses a Pairwise Concentration diagnostic tool that plots the concentrations of the components against each other. Figure 3 shows the Pairwise Concentration results of the acetic acid and alcohol concentrations. Note that there is no strong trend indicating that the alcohol concentration does not track with the acetic acid concentration.

The chemometric model developed indicated there is excellent correlation between the calculated concentrations and the actual concentrations. Figure 4 shows the plots of the actual concentrations from the samples and the concentrations calculated from the chemometric model. Note that, for both acetic acid and alcohol, there is very close agreement between the calculated and the actual concentrations. Figure 4 also shows the correlation coefficients for the plots as well as the Root Mean Square Error of Calibration (RMSEC); Root Mean Square Error of Prediction (RMSEP) and Root Mean Square Error of Cross-Validation (RMSECV). The RMSEC is the error developed from the 65 samples used in the calibration of the model while the RMSEP was obtained from the 15 samples used in validating the model. The RMSECV is obtained by removing individual samples from the method, recalibrating without these samples, then calculating the concentration and subsequent error of the removed samples. This cross-validation technique removes and recalculates each sample sequentially to develop the final RMSECV. It is important to note that these three error values should be similar in magnitude to each other in order for the chemometric method to reliably predict unknown samples. The root mean error values for the acetic acid indicate they are accurate to less than 0.2%. Similar analysis shows the accuracy of the alcohol values to be better than 0.1%.

An additional metric for determining the validity of a chemometric model is through the use of Predicted Residual Sum of Squares (PRESS) plots. These plots graphically demonstrate the behavior of the cross-validation error as additional factors are used in the calibration. Each factor used in the calibration accounts for spectral variation found within the sample set. Ideally, the first factor will account for most of the variation with subsequent factors accounting for the remaining variation. An ideal PRESS plot will have errors that drop off dramatically to some minimum then stabilize at that minimum or increase slightly. Using more factors than is indicated by the PRESS plot often results in overfitting the model. Overfitting results in high correlations and low errors of calibration, but yields high validation errors, which limit the predictive ability of the model. Figure 5 shows the PRESS plots of the acetic acid and alcohol for the chemometric model. These plots show the initial reduction in error to a minimum with few factors as expected for a robust method.

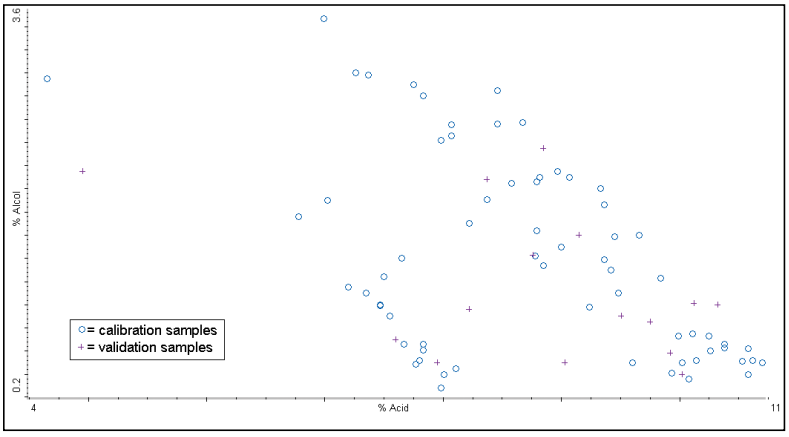


Figure 3: Results of Pairwise Concentration diagnostic indicating there is no great trend in plotting concentrations of acetic acid with concentrations of alcohol. Blue circles (o) represent spectra used in the calibration. Red crosses (+) represent validation spectra.

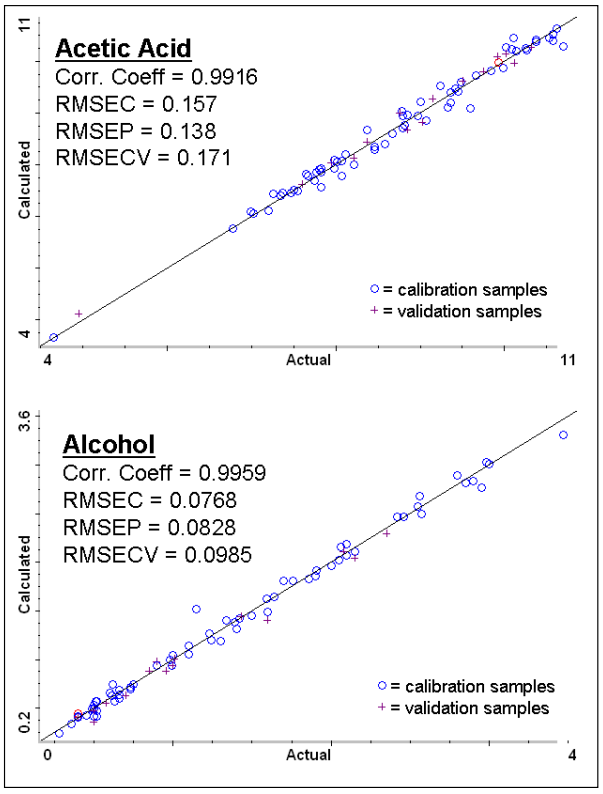


Figure 4: Correlation plots indicating how well the model predicts the actual acetic acid and alcohol concentrations.

thermo scientific

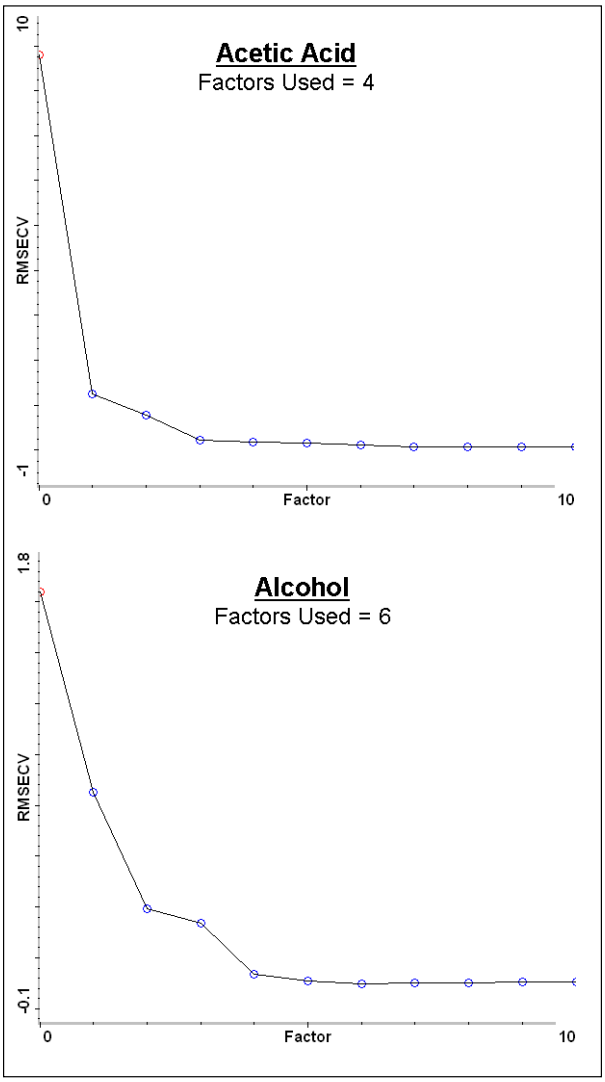
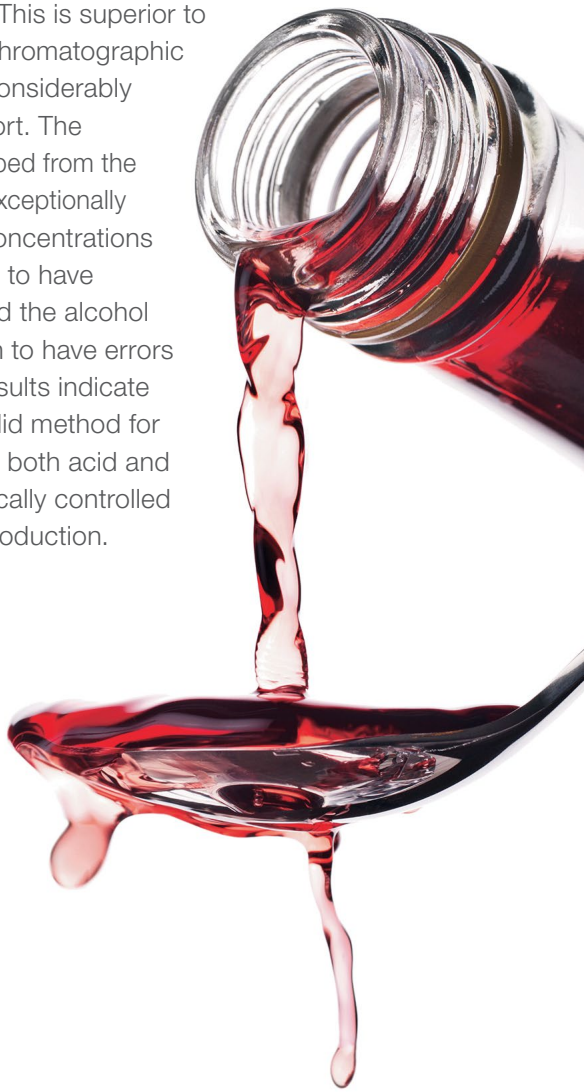


Figure 5: PRESS plots showing the decrease in cross-validation error as additional factors are used in the calibration. The low number of factors and the magnitude of the error indicate there is no overfitting of the data.

Conclusion

Analysis of vinegar stocks for acetic acid and alcohol content was successfully accomplished using the Antaris FT-NIR analyzer. Each spectrum was automatically collected within a few seconds allowing for rapid real-time determination of acid and alcohol concentrations. This is superior to the standard titration and chromatographic techniques which require considerably more time and technical effort. The chemometric model developed from the spectral samples provided exceptionally good results. Acetic acid concentrations in the samples were shown to have errors of less than 0.2% and the alcohol concentrations were shown to have errors of less than 0.1%. These results indicate that FT-NIR analysis is a valid method for simultaneously determining both acid and alcohol content in a biologically controlled process such as vinegar production.



Find out more at **thermofisher.com/nir**

ThermoFisher SCIENTIFIC