Quantitative analysis of resorcinol in aqueous solution by near infrared spectroscopy—from the laboratory to the production floor

S. Agyare Yeboah
Analytical Technology Division, Eastman Kodak Company, Rochester, NY 14650, USA.

Regis Cinier and Jean Guilment
Kodak European Research Division, Laboratoire d'Analyses, Chalon/Saone, 71102, France.

Introduction

In the photographic industry a large number of chemicals are generated and incorporated into a variety of product formulations. They comprise a powder, a mixture or a single analyTe dissolved in an appropriate solvent. Each of these chemicals must meet a set of pre-defined specifications with respect to some quantitatively measurable property such as concentration, potency, optical absorptivity at a known wavelength or a specific photographic activity. Most of the chemicals are stored as dry powders for long periods of time prior to their incorporation into products. For this reason, their long term stability, hygroscopicity and morphology are of interest to the formulation chemist who must ensure that they exist within tight limits of variability.

The potency of a chemical must be known to determine the appropriate amount to be used in any formulation. Potency is defined as the amount of the active ingredient present in a known weight of the dry chemical. It is important to contrast this measure with concentration which is merely the weight of the raw chemical in a known amount of solution or diluent. A number of analytical methods have been developed for the quantitative determination of the potency of the chemicals in solution (called doctor solutions). Typically, these methods utilize absorption in the ultraviolet (UV) or visible regions of the optical spectrum. Where the analyte of interest does not contain an organic chromophore, titration procedures are used. There are yet doctor solutions that cannot be easily quantitatively analyzed at-line with any of the traditional, easily implementable assay techniques. For the UV/visible spectroscopic methods, the solutions require large volumetric dilution factors, which have proved to be a major source of variability. This has led to the investigation of other techniques for determining potency.

One technique that has received much attention in our laboratories for this purpose is near infrared (NIR) spectroscopy.[1–4] The interest in, and the viability of, near infrared spectroscopy for the analysis of chemicals stem from a number of factors. Absorptions in the near infrared region arise from vibrational transitions to the second or higher energy states. Because of the very low probability of such transitions, absorption intensities are several orders of magnitude below those of the corresponding fundamental vibrations. Consequently, larger sample concentrations and thicker sample formats can be measured without the need for stringent sample preparation

From Near Infrared Spectroscopy: The Future Waves

procedures. Additionally, NIR has the advantage that aqueous solutions can be readily analyzed without much interference from water absorption. The simplicity of near infrared instrumentation readily permits analysis of various forms of samples. Fiber optic coupling further expands the sample handling capabilities by allowing samples to be measured *in situ*, on-line or at-line, eliminating the need to extract samples out of the process in which they are generated. The instruments combine high source intensity and sensitive detectors to provide high signal-to-noise measurement. They are very rugged, fast and lend themselves very well to on-site measurements.

This paper describes the development of near infrared methods for the analysis of doctor solutions at improved measurement precisions, and discusses the precautions necessary for the successful transfer of the methods from the analytical laboratory to the formulation facility. Results obtained for a solution of resorcinol at 300 g L^{-1} in water will be used to illustrate the approaches taken. Parameters such as temperature, optical pathlength and mathematical pre-processing of the spectral data are investigated in order to determine their effect on the predictive ability of the method and to find ways of developing a method insensitive to the lab environment.

Experimental

Samples were prepared gravimetrically within a weight range of ±10 % of the aim value (300 g L^{-1} in the case of resorcinol). Two stock solutions were prepared at the high and mid range concentrations of 110 and 100%, respectively. Aliquots of each solution were then gravimetrically diluted to obtain 21 samples, having a gaussian distribution around the aim concentration so as to minimize leverage effect of the extreme values[5] (see Table 1). The samples were split into two sets of 15 and 6 for the calibration and validation, respectively, of the quantitative methods. Other methods such as linear regression and K-Matrix were tested but do not give results as good as PLS for the samples; these methods will not be discussed in this paper. Finally, a batch sample from the manufacturing area was analyzed to check the repeatability of the method and to verify if there would be a bias between the NIR and the UV/visible method (linear regression) used in the plant.

NIR spectra were recorded at 8 cm^{-1} resolution from 4000 to 10,000 cm^{-1} (2500–1000 nm) with a Bruker IFS28/N spectrometer equipped with a tungsten filament source, a quartz beamsplitter and a Peltier cooled MCT detector. Measurements were made in the transflectance mode through a fiber optic bundle probe having an adjustable pathlength from 1 to 10 mm (equal to twice the gap between the sapphire window at the end of the fibers and the stainless steel reflecting mirror). The temperature of the samples was adjusted using a thermostatted double envelope cell. Five spectra at 50 scans per spectrum were collected in quick sequence to allow for short term instrument variation. Reference spectra collection was alternated between each set of sample spectra. The calibration methods were developed using PLS regression algorithms available in the Bruker OPUS software.

UV/visible absorption spectra were recorded at 1 nm resolution from 200 to 800 nm using 10 mm quartz cells on a Lambda 9 spectrometer from Perkin-Elmer with a 1 nm resolution and a scanning speed of 120 nm min^{-1}. All the samples were prepared or diluted using class A^{+} volumetric glassware.

Results and discussion

The application of multivariate mathematical techniques to the analysis of NIR spectra is necessitated by the fact that the spectra are generally complex due to broad and overlapping absorption bands. In a two component system, as in the aqueous solution of resorcinol, bands can, however, be found that are unique enough to warrant the use of simple linear regression techniques for quantitation. This is the case for resorcinol in water shown in Figure 1.

Table 1. Calibration and validation samples.

Samples	Aim	Actual	
1	110.0	110.0	Calibration
2	105.0	105.07	Calibration
3	**103.0**	**102.96**	**Validation**
4	102.0	102.02	Calibration
5	101.7	101.62	Calibration
6	**101.4**	**101.42**	**Validation**
7	101.0	101.03	Calibration
8	100.8	100.72	Calibration
9	**100.3**	**100.31**	**Validation**
10	100.0	100.03	Calibration
11	99.9	99.90	Calibration
12	99.6	99.61	Calibration
13	**99.5**	**99.48**	**Validation**
14	99.4	99.40	Calibration
15	99.0	99.00	Calibration
16	**98.7**	**98.72**	**Validation**
17	98.5	98.54	Calibration
18	98.0	97.98	Calibration
19	**96.0**	**96.03**	**Validation**
20	95.0	95.06	Calibration
21	90.0	89.98	Calibration

100% corresponds to 300 g L^{-1} of resorcinol in water.

The bands at 4700 and 6000 cm^{-1} are isolated sufficiently to permit the use of simple calibration procedures. But experience shows that in a production environment the effect of temperature and pathlength variations can lead to fast degradation of measurement precision and accuracy. For aqueous solutions, shifts in the water bands resulting from interaction of temperature with hydrogen bonds can adversely impact precision unless such shifts are accounted for in the calibration. The use of fiber optic probes not only limits the useful region of the spectrum to that

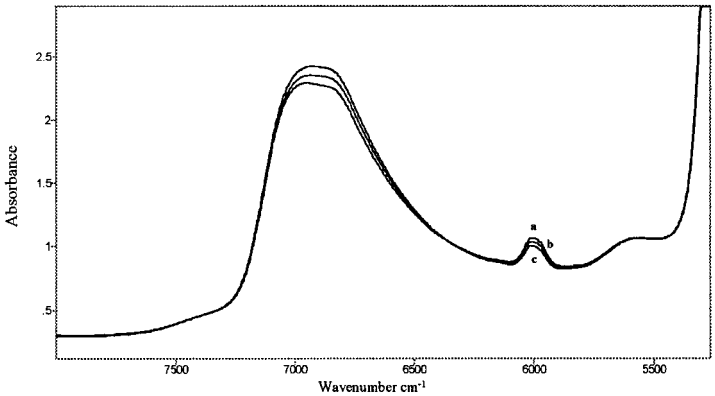

Figure 1. NIR spectra of aqueous solution of resorcinol at (a) 110%, (b)100% and (c) 90% potency.

above 5300 cm^{-1}, but can lead to small changes in pathlength due to wear and tear. For these reasons, we found partial least squares (PLS) approach to calibration to provide the best precision. Four methods based on PLS in combination with different spectral pre-processing tools were compared. These included vector normalization, derivative spectroscopy and internal standard, which were compared to evaluate their effectiveness in minimizing baseline shifts, pathlength changes and source fluctuation and thereby increase the measurement precision. The PLS prediction residual sum of squares (PRESS) and the standard error of prediction (*SEP*) formed the basis of the comparison.

Methods development

1. Vector normalization pre-processing of spectra is useful when small changes in pathlength in the analysis of liquid samples or films is unavoidable. The object is to reduce the intensity values (treated as vectors) of each spectrum to a constant value, typically unity.[6] This, however, has the disadvantage that artifacts, such as spectral nonlinearities resulting from absorption saturation, severely bias the results. In the approach used in the current analysis, normalization was achieved by calculating the average value of all of the intensity values in each spectrum. This average value was then subtracted from the spectrum so that the middle of the resulting spectrum attained a value equal to zero, i.e. $y = 0$. The spectrum was then divided by the square root of the sum of the squares of all the y values. A final spectrum was then obtained whose vector norm was unity. With this treatment, a *SEP* of 0.105% was able to be calculated from the validation data set using a rank of six.

2. First derivative spectroscopy, combined with 13 point smoothing, using the Savitzky–Golay algorithm, was employed as the second pre-processing tool. The primary advantage of derivative spectroscopy is to improve spectral resolution and correct for small local baseline shifts. This treatment resulted in a *SEP* of 0.088% at a PLS loading factor of eight. Figure 2 shows a plot of the actual versus predicted concentration values of the calibration and validation data sets.

3. A third pre-processing technique involving the use of the water band centered around 5500 cm^{-1} as an internal standard was evaluated. This band is known to be relatively insensitive to temperature. Lin and Brown found this band to coincide with the isosbestic point when the spectra of water measured at different temperatures were subtracted from that of a given temperature.[7] This band was sufficiently stable during the method development of the current study and was

From Near Infrared Spectroscopy: The Future Waves

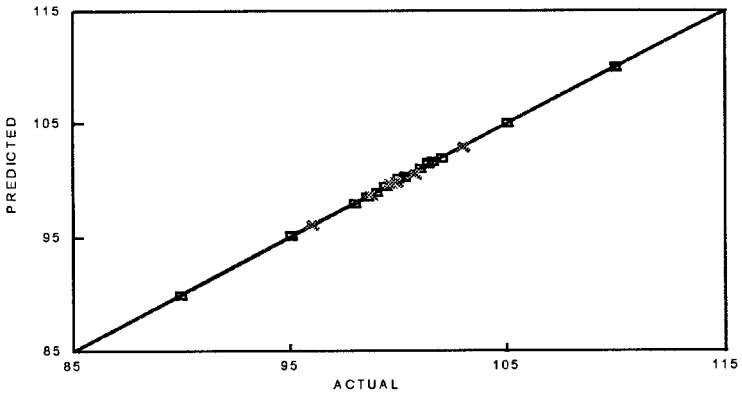

Figure 2. Predicted versus actual potency values obtained by Method 2.

therefore considered a useful internal standard. Its intensity is also in the same order of magnitude as the resorcinol band at 6000 cm^{-1} and therefore provides a good basis for normalization. This method gave a *SEP* value of 0.179% at a loading of four.

These results suggest that the first derivative transformation of the data yields the least error of the three methods. Up to this point, only results showing the laboratory evaluation of the methods under constant room temperature have been reported. However, criteria such as repeatability, operator variability and temperature insensitivity are good measures of the robustness of methods intended for use in production facilities. To this end, additional experiments were conducted using actual production samples of resorcinol solutions. One experiment tested the repeatability of the method; the second studies of the effect and incorporation of temperature variations on the method.

Repeatability

A batch sample of resorcinol obtained from the normal production process was tested for potency by a visible spectroscopic method. Fifteen measurements of the sample were performed by two operators over a period of three days, yielding a mean potency value of 100.77 with a standard deviation of 0.377. Because the sample was the output of a production process, it was impossible to determine an aim point. Near infrared spectra of the same sample were obtained over a period of five days with four sets of five measurements performed each day (two in the morning and two in the afternoon). A new reference was obtained prior to each set of sample measurements. Both the probe and sample temperatures were maintained at 20 ± 1°C. The spectra were then analyzed with first derivative pre-processing to obtain potency values. The results of the NIR measurements are graphically illustrated in Figure 3. The predicted potency averages about 97.5%, with very high precision within each set of five measurements. However, the precision drops if the first 90 measurements are considered together. This suggests that cleaning the probe between reference and sample measurements caused the sample measurement gap at the probe tip to change, leading to variations in pathlength and, consequently, predicted potency. In fact, when the probe gap is intentionally made loose and slightly larger than 2 mm, measurements 86 to 90 demonstrate that the predicted potency of the sample increases by about 20%. Quite clearly, slight variations in pathlength can be detrimental to the precision of the method and should be appropriately accounted for. That six loading factors were required for this method is a consequence of the non-systematic manner in which pathlength variations occurred.

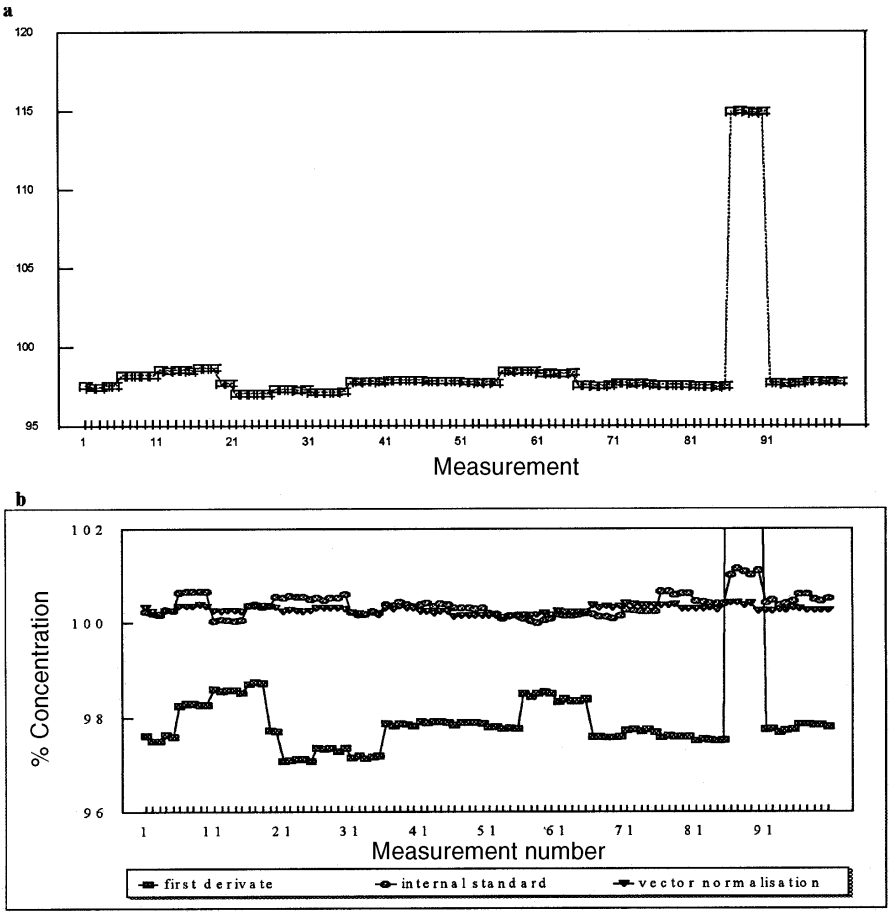

Figure 3. Measurement repeatability of (a) Method 2 and (b) Methods 2, 3 and 1.

The same spectra were analyzed with the two other methods, namely, pre-processing with vector normalization and internal standard. As Figure 3(b) shows, both methods not only provide a better correction for pathlength variations, but also give a better estimate of the potency of the solution. In addition, the vector-normalized spectra yield precisions that are three and four times better than the near infrared internal standard and visible spectroscopic methods, respectively. Table 2 lists the precision obtained by each method for the production sample. In comparison, near infrared spectroscopy, coupled with vector normalization, yields the best overall results. Near infrared spectroscopy also has the advantage of speed in that spectra can be generated in less than 1 minute, compared to the 20 minutes required for visible spectroscopy.

Temperature effect

Having established that pathlength variations, leading to changes in predicted potency, can be corrected for by use of vector normalization of the spectra, we proceed to a study of the effect of temperature. The effect of temperature on the near and mid-infrared spectra of water and aqueous solutions has been widely published.[8–14] Changes in the positions and relative intensities of the

Table 2. Effect of spectral pre-processing on the measurement repeatability of production resorcinol sample.

Mathematical treatment	Average value	Standard deviation (%)
First Derivative	98.98	3.756
Internal standard	100.38	0.2441
Vector-normalization (21°C)	100.32	0.078
First derivative (except measurements 86-90)	97.82	0.426
Vector-normalization (different temperatures)	99.90	0.141

water bands occur as a result of the weakening and breaking of hydrogen bonds. In particular, Lin and Brown[7] and McCabe *et al.*[8] have established that a variation in temperature causes a shift in the equilibria between the different species of water. Thus, it is important that temperature variations be taken into account when developing quantitative methods for aqueous systems, if sample or instrument temperature is expected to vary during the application of the methods. To evaluate the impact of temperature on the measurement of the resorcinol solutions, two experiments were conducted. First, a set of spectra of the same samples used for the repeatability study were recorded after exposing the sample to different temperatures between 9 and 34°C. Secondly, a selected set of the calibration and validation samples were recorded again at 9, 15, 24.5 and 34°C, so that a span of five temperatures, including 21°C, could be included in the development of a second calibration. In this case, temperature is considered an independent variable in the PLS modeling. As Figure 4 shows, the potency of the production sample, previously determined to be 100.77% by visible spectroscopy, shifts from 102.25% at 9°C to 100.25% at 34°C when the values are predicted using Method 1, that is, applying vector-normalization with all the samples at 20°C. Such an error is significant, especially since temperature swings from day to day in a production environment can be large. However, when the second calibration, with temperature as an

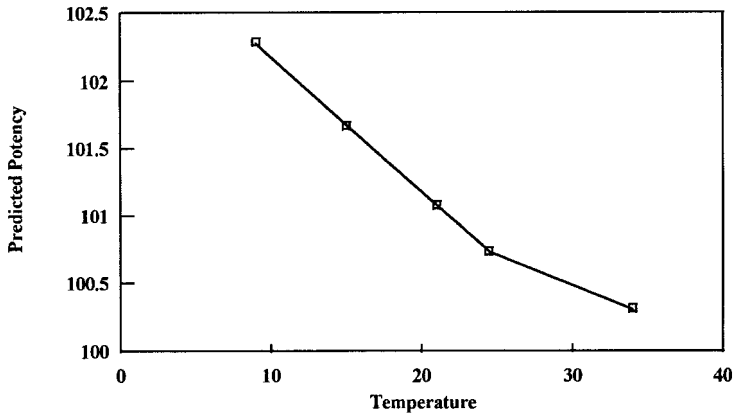

Figure 4. Potency predicted by Method 1. Samples were measured at different temperatures.

independent variable, is used to predict the potency of the production sample, a mean value of 100.1% is obtained, with a standard deviation of 0.158. This calibration yields a *SEP* of 0.141 (shown in Table 2), which, although worse than the previous calibration, is still three times better than that of the visible spectroscopic method.

Conclusion

The calibration procedure described in this report demonstrates that NIR is useful for the quantitative determination of resorcinol in aqueous solution with a precision of 0.1% (one sigma). Improvements in *SEP* can be gained by including both temperature and pathlength variations in the model development. The proper selection of spectral pre-processing tools, such as the application of derivative spectroscopy and vector-normalization, provides additional enhancements. The successful transfer of environment-insensitive methods from the laboratory to the production facility hinges, among other things, on how well process conditions and instrument variations are accounted for in the calibration model. Such procedures are of importance in the transfer of methods from the laboratory to the production facility.

A related subject is the transfer of calibration from one instrument to another.[14–16] Our initial attempts have proved feasible, but more investigation is needed to avoid degradation of precision and ensure transparency to the end-user. The use of FT-NIR instruments, with their frequency accuracy and repeatability, should help avoid the need for frequent wavelength linearization.

References

1. A.M.C. Davies, *European Spectroscopy News* **73,** 10 (1987).
2. K.H. Norris and J. Ben-Gera, *Isr. J. Agric. Res*. **18,** 125 (1968).
3. K.H. Norris, K.J. Kaffka, F. Kulscar and I. Draskovits, *Acta Aliment*. **11,** 271 (1982).
4. K.H. Norris, *Agricultural Handbook* **6,** 643 (1989).
5. H. Martens and T. Næs, *Multivariate Calibration*.John Wiley and Sons, Chichester (1989).
6. S.N. Deming, J.A. Palasota and J.M. Nocerino, *J. Chemometrics* **7,** 393 (1993).
7. J. Lin and C.W. Brown, *Vibrational Spectroscopy* **7,** 117 (1994).
8. W.C. McCabe, S. Subramanian and H.F. Fisher, *J. Phys. Chem*. **74(25),** 4360 (1970).
9. P. Geladi and B.R. Kowalski, *Anal.Chem. Acta* **19,** 185 (1986).
10. J.J. Berzas Nevado, C. Guiberteau, A.M. Cabanillas and C. Salcedo, *Analysis* **22,** 5 (1994).
11. J. Lin and C.W. Brown, *Trends in Analytical Chemistry* **8,** 13 (1994).
12. J. Lin and C.W. Brown, *Environmental Science & Technology*, **27,** 1611 (1993).
13. K.H. Hazen, M.A. Arnold and G.W. Small, *Appl. Spectrosc*. **48(4),** 477 (1994).
14. C.A.Young, K. Knutson and J.D. Miller, *Appl. Spectrosc*. **47(1),** 7 (1993).
15. J.J. Workman, *Handbook of Near Infrared Analysis*. Marcel Dekker, New York (1992).
15. M. Forina *et al., Chemometric Intelligent Lab. System* **27,** 189 (1993).
16. E. Bouveresse, D.L. Massart and P. Dardenne, *Anal. Chem*. **67,** 1381 (1995).