Potato Crisp moisture determination using NIR data and a Back Propagation Neural Network

[1]Nigel Yee, [2]Paul Potgieter, [3]Steve Liggett
*[*1]Unitec Institute of Technology, nyee@unitec.ac.nz*
[2]Unitec Institute of Technology, ppotgieter@unitec.ac.nz
[3]Unitec Institute of Technology, sliggett@unitec.ac.nz

Abstract

Near infrared analysis is a tool used for non-destructive determination of material properties and the potato crisp production sector has been using the technique for determination of moisture content however near infrared spectral models suffer from problems associated with light scatter. Light scatter results from geometric irregularities in the samples geometry and this reduces the accuracy of near infrared calibration models without preprocessing for scatter removal. Quantitative calibration models have benefited from the development of artificial intelligence methods and the neural network is now a popular tool for quantitative calibration model formation. In this paper we compare the performance of a back propagation neural network calibration model using 3 forms of preprocessed data, orthogonal signal correction, standard normal variate and data with no scatter preprocessing prior. The correlation coefficient was used to determine the neural networks methods performance and it was found that a neural network using data with no scatter preprocessing yielded the best results.

Keywords: *Neural network; standard normal variate; orthogonal signal correction*

1. Introduction

Model calibration is critical for accurate prediction of material chemical properties when near infrared (NIR) spectral instrumentation is used. The efficient use of NIR spectroscopy is dependent on instrument calibration and the conditions under which the samples are obtained and presented to the instrument. NIR spectra can be obtained from a spectrograph or from an imaging spectrometer and results presented in this paper pertain to use of NIR potato crisp image spectra. A problem associated with NIR spectral analysis is light scatter due to geometric irregularities from the sample in its natural state interacting with light.

Different methods have been devised for removal of scatter in objects [1][2] and the most popular is the standard normal variate (SNV) method and this pre-processing method has been shown by a number of author to improve calibration model prediction abilities ([3],[4]). Orthogonal signal correction (OSC) is a newer technique which has been shown to improve regressive model predictive capabilities, and has been shown to compare well to multiplicative scatter correction in data sets containing scatter [5].

Neural networks are an alternative to regressive modeling and there are many different forms of neural networks that have been applied to NIR calibration problems ([6],[7]); a popular method is the back propagation algorithm[8]. Neural networks have the advantages in that the neural network can be tuned to remove noise and scatter and it has been shown that pre-processing can improve the neural networks predictive abilities however this has been shown to be dataset dependent [9].

This paper compares neural network models with different methods of preprocessing, OSC, SNV and no scatter preprocessing for the purpose of removing the effects of scatter from the data. Previous work has demonstrated the usefulness of OSC in removal of scatter with partial least squares regression [10] and the current work investigates the effectiveness of OSC with the neural network.

2. Signal treatments

2.1 Orthogonal signal correction

The OSC algorithm described in this paper, is the OSC Fearn algorithm [11]. OSC generates factors/components that contain the variation in X not correlated to y. OSC then uses a subtraction from the X matrix, of scores t_o and loadings p_o, that contain the largest amounts of variability in X, while still being correlated to y. Using vector loading weights w_o the problem is posed as $t=Xr$.

We maximize

$$Max(t^T t), \tag{1}$$

subject to constraints $t^T y=0,$ and $r^T r =1,$

Step 1 we use

$$M=I-X^T y(y^T XX^T y)^{-1} y^T X, \tag{2}$$

Step 2 we form a weighted matrix Z

$$Z=XM. \tag{3}$$

We now find all the eigenvalues of $Z^T Z$ and then use these eigenvalues, λ, and eigenvectors, c, to produce score vectors t_{oi} and loadings p_{oi} by iteration of steps 2 to 3 until the desired number of factors have been found.

Step 3 Score vectors are now calculated by

$$t_{oi}=\lambda^{-1/2} MX^T c_1, \tag{4}$$

Step 4 Loadings are then calculated

$$p_{oi}=X^T t_{oi}/(t_{oi}^T t_{oi}), \tag{5}$$

p_{oi} and t_{oi} are save for future predictions of data.

The OSC algorithm calculates the loading and score vectors specific to the X matrix by iteration of steps 3-4 solely.

2.2 Orthogonal signal correction prediction

The OSC, functions by subtracting only the orthogonal factors desired. In order to achieve this all the loading and score vectors are calculated for the X matrix, (4) and (5), the matrix containing all the loading vectors are then pre-multiplied by a binarized weight matrix (which behaves as a kernel) in order to find P_{new}

$$P_{new}=P\cdot diag(W_b), \tag{6}$$

where $W_b=[\ w_{b1},w_{b2},w_{b3},\ldots w_{bn}]$ and $diag(W_b)$ is a diagonal weight matrix of binarized vectors w_b, a w_b value of 1 denotes a selected orthogonal factor and 0 denotes a non-selected factor, n is the size of the calibration data set, $P=[p_{o1}\ , p_{o2}\ ,\ldots p_{on}\ ]$ is a matrix containing all the p load vectors.

The new loading matrix P_{new} is used to calculate an orthogonally corrected data set X_{new}, given as

$$X_{new}=X-TP_{new}^{\ T}, \tag{7}$$

where $T=[t_1\ , t_2\ ,\ldots t_n\ ]$ is a matrix containing all the t vectors.

2.3 Standard normal variate

The SNV transformation removes the slope variation from spectra caused by scatter. The transformation is applied to each spectrum individually by subtracting the spectrum mean and scaling with the spectrum standard deviation.

$$x_{ij,SNV}=(x_{ij}-x_i)/((\Sigma_{i=1}{}^{p}(x_{ij}-x_i)^2/(p-1))^{0.5} \tag{8}$$

2.4 Back propagation neural network

The back propagation neural network is a supervised error driven learning method. The neural network used consists of three layers of units: an input layer, an intermediate layer hidden layer and an output layer and is configured in the feed forward direction (see Figure 1).

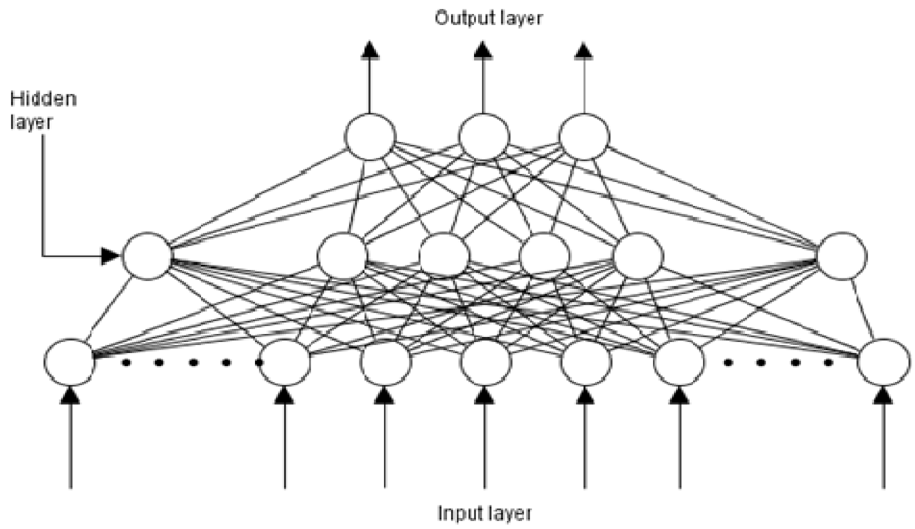

Figure 1. Feedforward back propagation neural network with input layer, hidden layer and output layer.

2.5 Measurement of model quality

The dot product correlation coefficient was chosen as a measure of the calibration model quality

$$R=\Sigma x_i y_i/(\ \Sigma x_i^2 \Sigma y_i^2)^{0.5}, \tag{9}$$

where x_i and y_i are the values of two spectra at wavelength i [12].

3.0 Experimental

An imaging spectrometer was built and mounted above a conveyor with an illumination system comprising of a single halogen lamp that flood lit the samples (shown in Figure 2). This instrument produced a spectral range from 700nm to 1100nm with 3nm separation between spectral data points on the spectral axis and 256 data points on the spatial axis. A single image was captured with no light entering into the spectrometer and this image was subtracted from each image in the data set to calibrate for background light. All the captured spectra were transformed using the Beer-Lambert transform (10) [13],

$$A=log(1/R). \tag{10}$$

A total of 120 potato crisp samples of different moisture contents were presented to the spectrometer and spectral images captured to form the data set. The chips then had the moisture content calculated using drying [14].

For the purpose of data analysis the 120 image data set was split into a calibration data set of 100 images and a prediction set of 20 images. A single spectra was then extracted from spatial position 72 in each image and used in the neural network model.

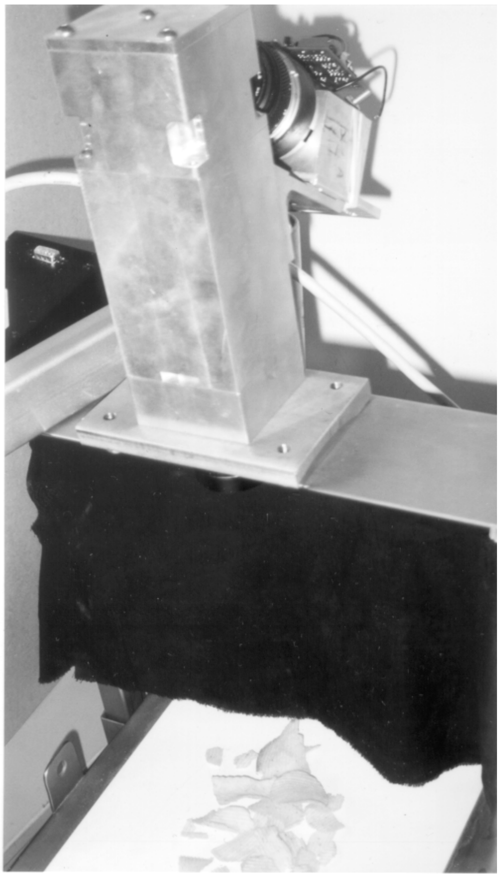

Figure 2. Experimental setup of spectrometer mounted above conveyor with sample product.

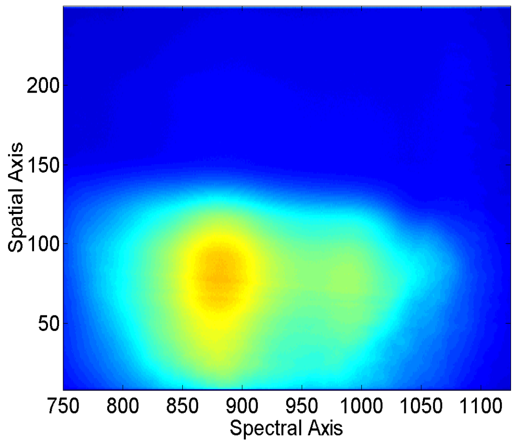

Figure 3. A spectral image captured for a potato crisp using an imaging spectrometer.

3.1 Software

The data analysis was performed on Matlab version 5.3 running on a Laptop with an Intel processor that had a clock speed of 2.6GHz.

4.0 Results and discussion

Presented in Figure 3, is multi-spectral image of a sample potato crisp. Within this image the potato crisp is located between positions 0 and 125 in the spatial dimension. Also within this image it can be seen that the lighting configuration introduces spectral non-linearity between the spatial axis positions 150 and 220 where no potato crisp is present.

Figure 4 shows a plot of the measured moisture content vs. the estimated moisture content for the prediction data set with no scatter correction preprocessing. Figure 5 shows a plot of the measured moisture content vs. the estimated moisture content for the prediction data set with SNV scatter correction preprocessing and Figure 6 shows a plot of the measured moisture content vs. the estimated moisture content for the prediction data set with OSC preprocessing.

The best correlation coefficient for the prediction set was obtained from the data that had no form of scatter correction preprocessing, followed by SNV preprocessing while OSC pre-processing produced the weakest correlation coefficient result. This result was surprising as we initially expected scatter correction preprocessing of the data prior to presenting the data to the neural network would improve the prediction performance. The neural network used a seeding algorithm to select the initial back propagation weight vectors. We noticed that between different runs the results were different within each form of preprocessing and this indicates that the strength of the neural network was dependent on the initial seeding of the weight vectors for the network and also indicates the existence of different optimal converged solutions for the neural network on the training set. The application of the neural network to this data set shows that pre-processing of data does not always produce a stronger network with strong prediction capabilities and that the application of scatter pre-processing to improve network prediction abilities is data set dependent.

It should be noted that this work has used the same number of hidden nodes in the hidden layer and no comparison was made with networks of a simpler or more complicated topology thus network complexity could be another factor that could influence the results

In future work we intend to investigate other neural networks with this data set and we also intend to capture a new dataset. The spectrometer used for this work was one that was custom made however off the shelf instrumentation offers the advantage of service guarantees and quality assurances in the manufacturing process and thus our near future work intends to move back to using off the shelf instrumentation to standardize data set capture.

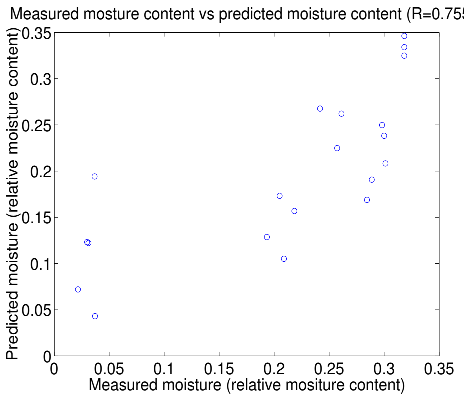

Figure 4. Predicted moisture content vs measured moisture content for data with no scatter preprocessing.

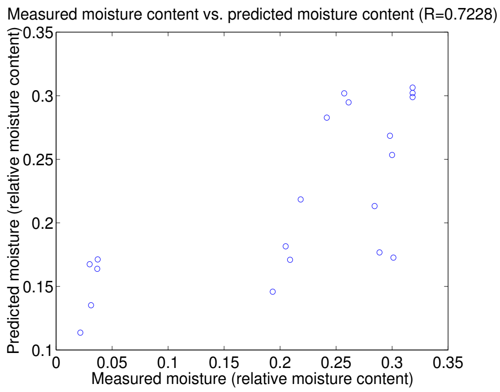

Figure 5. Predicted moisture content vs measured moisture content for SNV normalized data.

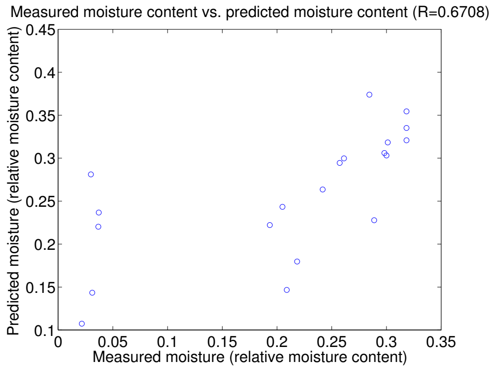

Figure 6. Predicted moisture content vs measured moisture content for OSC processed data.

5.0 References

[1] Martens, H., Nielsen, J., and Engelsen, S. B. "Light scattering and light absorbance separated by extended multiplicative signal correction. Application to near-infrared transmission analysis of powder mixtures", Analytical Chemistry, vol. 75, no. 3, pp. 394–404, 2003.

[2] Geladi, P., MacDougall, D., and Martens, H. "Linearization and scatter-correction for Near-Infrared Reflectance Spectra of Meat", Applied Spectroscopy, vol. 39, no. 3, pp. 491-500, 1985.

[3] Rinnan, A., van den Berg, F., and Engelsen, S. B. "Review of the most common pro-processing techniques for near-infrared spectra", TrAC Trends in Analytical Chemistry, vol. 28, no. 10, pp. 1201-1222, 2009.

[4] Barnes, R., Dhanoa, M., and Lister, S. J. "Standard normal variate transformation and de-trending of near infrared reflectance spectra", Applied Spectroscopy, vol. 43, no. 5, pp. 772–777, 1989.

[5] Pizarro, C., Esteban-Diez, I., Nistal, A., and J. Gonzalez-Saiz, J. "Influence of data pre-processing on the quantitative determination of the ash content and lipids in roasted coffee by near infrared spectroscopy", Analytica Chimica Acta, vol. 509, no. 2, pp. 217-227, 2004.

[6] Kim, J., Mowat, A., Polle, P., and Kasabov, N. "Linear and non-linear pattern recognition models for classification of fruit from visible-near infrared spectra", Chemometrics and Intelligent Laboratory Systems, 51, 2, pp. 201-216, 2000.

[7] Pavlidis, I., Morellas, V., and Papanikolopoulos, N. "A vehicle occupant counting system based on near-infrared phenomenology and fuzzy neural classification", IEEE Transactions on Intelligent Transportation Systems, vol. 1, no. 2, pp. 72-85, 2000.

[8] Schulz, H., Derrick, M., and Stulik, D. "Simple encoding of infrared spectra for pattern recognition Part 2. Neural network approach using back-propagation and associative Hopfield memory". Analytica Chimica Acta, vol. 316, no. 2, pp. 145-159, 1995.

[9] Dou, Y., Sun, Y., Ren, Y., and Ren, Y. "Artificial neural network for simultaneous determination of two components of compound paracetamol and diphenhydramine hydrochloride power on NIR Spectroscopy", Analytica Chimica Acta, vol. 528, no.1, pp. 55-61, 2005.

[10] Ni, W., Brown S., and Man, R. "The relationship between net analyte signal/preporcessing and orthogonal signal correction algorithms". Chemometrics and Intelligent Laboratory Systems, vol. 96, no. 5, pp. 97-107, 2009.

[11] Fearn, T. "On orthogonal signal correction", Chemometrics and Intelligent Laboratory Systems, vol. 50, no. 1, pp. 47-52, 2000.

[12] Draper, N. R., and Smith, H. Applied regression analysis, Wiley, NJ, USA, 1998.

[13] Swinehart, D. F. "The beer-lambert law", Journal of Chemical Education, vol. 39,no. 7, pp. 333-339, 1962.

[14] Nollet, L. 1996. Handbook of food analysis, Marcel Dekker, New York. 1996, 72-73.