Evaluation of Vegetables Shelf Life Using Multispectral Scattering Method

Faisal Abdullah

Physics Department, Faculty of Mathematics and Natural Sciences, Syiah Kuala University
Darussalam-Banda Aceh 23111, Indonesia.
faisal@fmipa.unsyiah.ac.id

Abstract. This study was aimed to develop an algorithm based on the optical properties of some selected vegetables which can be used to evaluate shelf- life of some selected vegetables using multispectral scattering method. Multispectral algorithm will be developed to correlate light backscattering radiation of a vegetable with the level of shelf life conditions. The partial least square (PLS) regression models using three wavelengths were used to estimate the shelf life changes of the vegetable samples. The results showed that the developed multispectral scattering algorithms can be used to evaluate the shelf life changes of Chinese cabbage, carrot, chili, and cucumber using wavelengths at 880, 890, and 950 nm as the light sources respectively. The calibration and validation processes of the algorithm produced good accuracy measurements as represented by high R^2 values and low Root Mean Square Error (RMSE). For the Chinese cabbage and the carrot samples, the algorithm was not effective to predict the changes of the shelf life because the responses was kept increasing even after the samples were dried and deteriorated. Also, the algorithms cannot be used to evaluate the shelf life changes of the onion because the calibration and validation processes produced low accuracy of measurements.

Keywords: *shelf life, multispectral method, vegetables, partial least square, scattering, algorithms.*

Introduction

Vegetables shelf life can be defined as a period of time during which a vegetable may be stored and remain suitable for consumption or sale. Shelf life is often closely associated with freshness, primarily for fruit and vegetable products. Shelf life is an important indicator of quality parameters of vegetables and fruits for sellers and consumers. Freshness, appearance, flavour, texture, nutritive value, and defects are the primary quality factors of fruits and vegetables [1]. Freshness and appearance are also commonly used by inspectors and consumers to judge the quality of vegetables. Traditionally, the inspectors and consumers used their eyes to judge the freshness and appearance of fruits and vegetables but the appearance can trick our eyes in certain situations.

Light being reflected from a vegetable product carries information and is used by inspectors and consumers to judge several aspects of quality but human vision is limited only to small region of the electromagnetic spectrum. Some quality parameters, such as internal qualities, respond well to wavelengths outside the visible spectrum [10]. In order to cover these features the inspectors and consumers need to use a standard instrument such as spectrometer or spectrophotometer to the parameters. However, these instruments are very expensive and can only be used in a laboratory or a research institution.

Multispectral method is widely used technology in remote sensing, military surveillance, chemical identification and imaging. The method relies on the distinctive spectral signatures of objects under investigation. In order to increase the effectiveness of the method, an algorithm that suitable for specific purpose is needed to be established.

A study using a new wavelength selection algorithm to improve complex near infra red (NIR) calibration have found that the selection algorithms lead to accurate prediction with only a few selected wavelengths [2]. These findings suggest that the multispectral technique can be used to achieve optimum results when assessing the changes of vegetable freshness.

The aims of this study were to develop an algorithm based on the optical properties of some selected vegetables using multispectral

_____

scattering method and then used the algorithms to evaluate the shelf life of the vegetables. A multispectral algorithm will be developed to correlate light scattering radiation of a vegetable with the level of its shelf life conditions. Multiple regression techniques such as PLS will be used to improve the accuracy of the algorithms.

The principle of the method used in this study is to determine the backscattering signature of a sample in visible and near infrared region. When a light beam is incident upon a piece of sample, the majority of the lights penetrate into the sample tissue. Upon entering the tissue, photons scatter in different directions some of them are absorbed, some pass-through to the whole sample and emerge from the opposite side, and some scatter back and reemerge from the region adjacent to the incident center. While the absorption is related to certain chemical constituent of the sample, scattering is influenced by the density, compositions, cells and intercellular structures of samples and therefore can be useful for measuring samples freshness [3].

The total reflected intensity is the sum of specular and diffuse reflectance. But only the diffuse reflectance contains information about the composition of the materials. Near infrared (NIR) spectroscopy techniques were used to analyze several agricultural materials by some researchers [4]. The study found that the techniques was possible to measure representative sample portions with simple and often requires no sample preparation at all. The

techniques do not need a special sample preparation that makes this method preferred for analysis of agricultural and food materials. Several researchers used semi-empirical approach which relates spectral indices such as spectral vegetation indices to some characteristics of the materials. They established a simple relationship between the biochemical components of interest C and the material optical properties as a function of the wavelengths, (λ_n) [5].

$$C = f(...(\}_1), ...(\}_2), ..., ...(\}_n)) \quad (1)$$

The researchers used this relationship to determine their components of interest such as chlorophyll and nitrogen concentration and water content of a leaf material [6, 7, 8, 9]

Methods

In this study, the light sources and detector are positioned at the same side of the sample as shown in Figure 1. The fiber optic probe in this setup was function as a guide to the light travelling backward from the sample to the detector. When a sample illuminated by a light source, the scattered reflected radiation from the sample is guided by the fiber optic to the detector. The intensity of body scattering radiation of the sample will provide information about the quality level of the sample tissue. The body reflectance means the diffused light that come from sample internal features [1]. Based on the approximation made by Equation 1, the scattering method is tried to
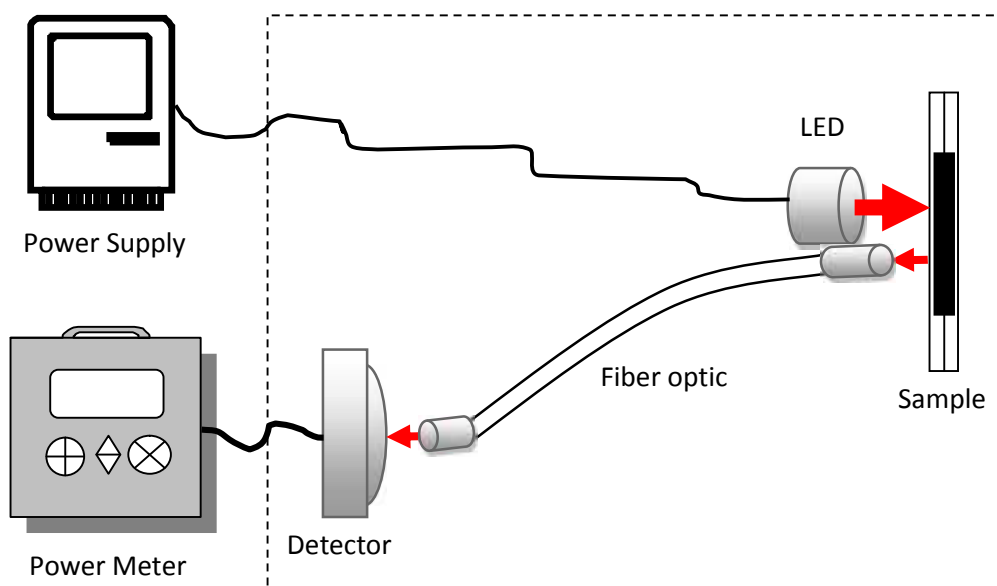

Figure 1 Setup of scattering measuring system.

_____

measure the reflectance intensities and then compare them with its reference intensity. The reference intensity for this study is the scattering radiation from a thick white paper which considered as a standard white illuminant.

Samples of selected vegetables were cut into a rectangular form and then put vertically onto a sample holder in front of a light source. Then a light beam perpendicularly illuminates the sample at its center as shown in Figure 1. To detect the transmitted photons from the samples, a photo detector type was positioned at the other side from the sample. The data measurements were repeated within the intervals of one hour. The acquired response data of the transmitted photons of the samples then plotted to get one-dimensional absorption profiles of the sample. Then these individual response were feed to PLS regression to produce a multispectral algorithm.

This study uses several type of fresh vegetables which are commonly consumed in daily food such as Chinese cabbage (*Brassica rapa chinensis*), carrot (*Daucus carota subsp. Sativus*), chilli (*Capsicum annuum*), cucumber (*Cucumis sativus*), and onion (*Allium sepa*). The Chinese cabbage was chosen to represent a green leaf vegetable which is mostly containing high chlorophyll. Carrot was chosen because of high carotenoids content while chili because of its red pigments. The cucumber was considered contain high water molecule and the onion because of its simple cell structure.

All of the selected fresh vegetables were then cut into a square shape sample with size of 4 cm^2. The leaf of the Chinese cabbage was carefully selected to produce two similar samples. The carrot, cucumber and chili were cut transversely and their outer shell was used to produce two similar samples. The onion samples were prepared by firstly peel off its outer dry shell and its fresh shell then carefully selected to produce two similar samples. The samples then divided into two groups. Both groups consist of similar type of the vegetable samples with approximately same level of freshness. The first group was used for collecting a calibration data set while the second group was used to produce a prediction data set.

The calibration of the algorithms to obtain a model for each sample were done using partial least square (PLS) regression method..The validation processes of the calibrated models

were done using cross validation method. Then all prediction residuals are combined to compute the validation residual variance and root mean square of prediction. The model then used to predict the self-life of each associated sample using the prediction data sets taken previously. The accuracy of the prediction can be evaluated using its variations values. The smaller variation values the higher the accuracy of the models.

Results and Discussion

Figure 2 shows the scatter plot of the response values of the Chinese cabbage sample for each individual wavelength. All of the responses keep increased over the course of the experiment even after the sample was badly deteriorated. The patterns are similar for all wavelengths. Table 1 shows the first three wavelengths which produced the higher R^2 values were 880 nm, 890 nm, and 950 nm. The correlation of these three wavelengths was particularly high and all of them located in the near infrared region.
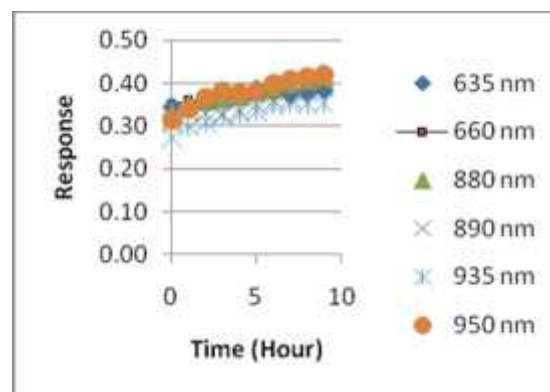

Figure 2 Scatter plot of the response values of the Chinese cabbage sample.

Table 1 The correlation values for Chinese cabbage samples

Wavelength (nm)	635	660	880	890	935	950
RSQ	0.742	0.802	0.987	0.949	0.896	0.909

The calibration processes of these three wavelengths produced a model of equation as shown in the Equation (2). The processes produced high R^2 value of 0.8421 and low RMSEC of 0.0064 as shown in Figure 3 (blue

_____

in color). The validation process of the model also produced high R^2 value of 0.7671 and low RMSEP of 0.0085. These results indicate that the accuracy of this model to detect the changes of the shelf life of Chinese cabbage was quite good. However the prediction results of the model are keep increasing even after the sample was badly deteriorated as shown in Figure 4. The variation values (vertical bars) produced from this process was fairly good but the model cannot be used to predict the shelf life of the Chinese cabbage.

Response = 0.1481 + 0.2974* (880) + 0.5299* (890) - 0.1850* (950)   (2)
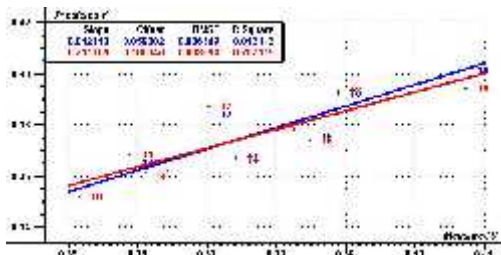
Figure 3 The results of calibration (blue) and validation (red) processes for the Chinese cabbage sample.
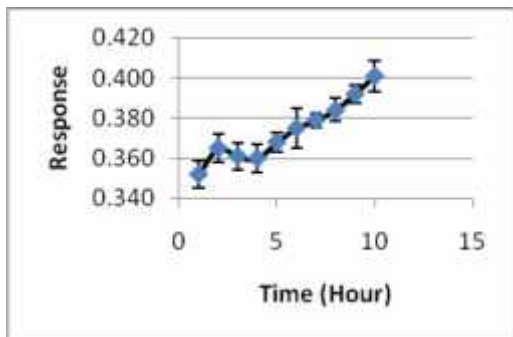

Figure 4 The prediction results of a new data set of the Chinese cabbage sample.

Figure 5 shows the scatter plot of the carrot sample for each individual wavelength. All of the responses increased slowly up to five hours of the experiment and then diverse in different ways. The response of 880 nm and 950 nm was constantly increased until the end of the experiment. The first three wavelengths which produced the higher R^2 values were 880 nm, 890 nm, and 950 nm as shown in the Table 3. These three wavelengths were then used in the multispectral setup to collect a validation data set.
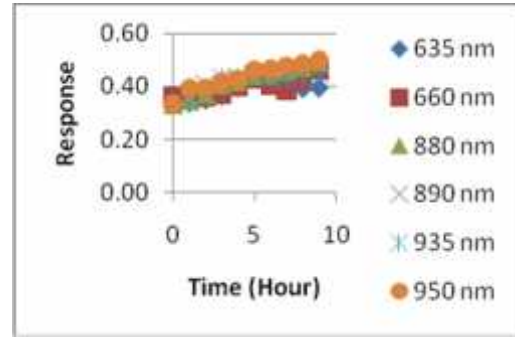

Figure 5 Scatter plot of the response values of the carrot sample.

Table 2 The correlation values of each wavelength for carrot sample.

Wavelengtht (nm)	635	660	880	890	935	950
RSQ	0.535	0.738	0.970	0.899	0.630	0.936

The calibration processes of the algorithm produced a model of equation as shown in Equation (3). The R^2 value produced by the process was quite high of 0.784 and low RMSEC of 0.0162 as shown in Figure 6 (blue color). The validation processes produced a fairly good R^2 value of 0.646 and low RMSEP as shown in Figure 6 (red color). Based on these values, the model still feasible in evaluating the shelf life changes of the carrot sample but the prediction processes provide the contrary as indicated by the continuously increasing response until the end of the experiment as shown in Figure 7. This indicated that this model cannot be used to evaluate shelf life changes of the carrot sample.

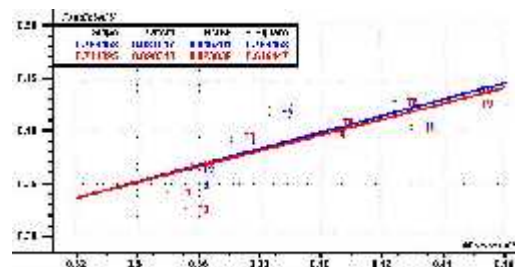Response = 0.1692 + 0.9865* (880) - 0.0830* (890) - 0.3326* (950)   (3)

Figure 6 The results of calibration (blue) and validation (red) processes for the carrot sample.
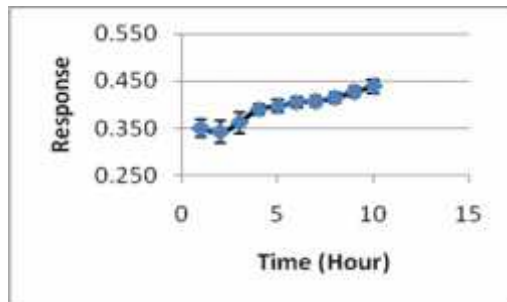
_____

Figure 7. The prediction results of a new data set of the carrot sample.

Figure 8 shows the scatter plot of the chilli sample for each individual wavelength. The responses of the wavelengths apparently increase linearly up to five hours of the experiment but then went to a diverse way. Wavelengths 880 nm, 890 nm and 950 nm keep increasing while the others become decreasing and increase slightly or still constant at the end of the experiment such as for both of the visible, 635 nm and 660 nm. In the visible range, the reflectance is mostly affected by pigments such as chlorophyll and carotenoids. The pigments usually deteriorated after the sample getting dry. These explained the decreasing of the wavelengths in the visible region. Table 3 shows that the first three wavelengths which produced the higher R^2 value were 880 nm, 890 nm, and 950 nm. These wavelengths were then used in the multispectral setup to collect a validation data set.
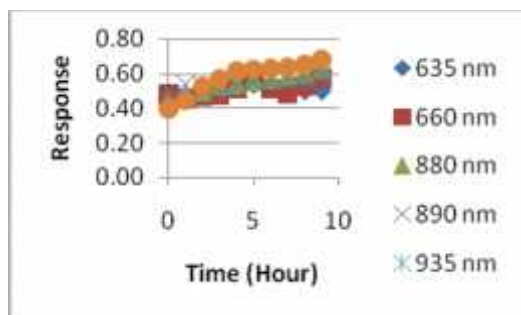

Figure 8. Scatter plot of the response values of the chili sample.

Table 3. The correlation values of each wavelength for chili sample.

Wavelength (nm)	635	660	880	890	935	950
RSQ	0.381	0.542	0.987	0.831	0.773	0.869

The calibration processes of the responses produced a model of equation as shown in Equation (4). The R^2 produced by the process was very high and the RMSEC value was quite

low as can be seen in Figure 8 (blue color). The validation processes of the model also produced a high R^2 value of 0.99073 and low RMSEP of 0.0157 (red color). These values indicated that the model is feasible to be used to detect the freshness condition of the chilli and it should produce a good accuracy of measurements. This is confirmed by the results of the prediction processes as shown in Figure 9. The variation values produced by the model were very low as depicted by the narrow vertical bars in the graph. The spectral responses of the prediction initially increase considerably up to 5 hours before it's become constant or almost constant until 8 hours. Therefore the sample was already dry at this time.

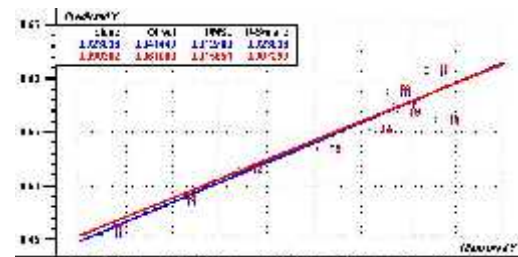$$Response = 0.3017 + 0.0431*(880) - 0.1746*(890) + 0.5443*(950) \quad (4)$$

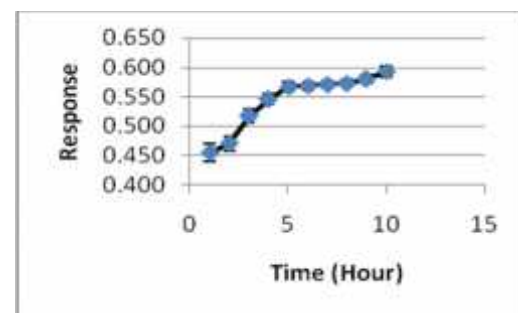Figure 9. The results of calibration (blue) and validation (red) processes for the chili sample.

Figure 10. The prediction results of a new data set of the chili sample.

Figure 11 shows the scatter plot from the responses of each individual wavelength for the cucumber sample. Mostly, the spectra responses were decreasing during the experiment, especially the responses of 880 nm and 950 nm. However the wavelengths of 635 nm and 890 nm decrease very slightly and tend to be constant. The first three wavelengths

which have higher correlation with the changes of shelf life condition are 950 nm, 890 nm, and 880 nm as shown in Table 4.

These three wavelengths which located in NIR region were then used in the multispectral setup to generate the validation data set.
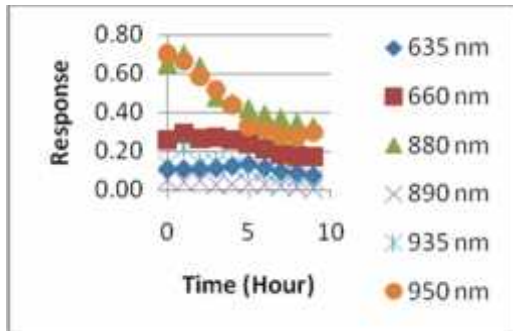

Figure 11 Scatter plot of the response values of the cucumber sample.

The calibration processes of the algorithm using the wavelengths produced a model of equation as shown in Equation (5). The R^2 value is 0.7349 which is reasonably high and the RMSEC is low (0.0217) as shown in Figure 12 represented by blue color graph. The validation processes also produced a fairly good R^2 values of 0.6669 and low RMSEP of 0.0271 as represented by red color graph.

These results indicate that the model can be used to detect the shelf life conditions of the cucumber. The prediction processes confirmed that the model produced an acceptable variation values as shown in Figure 12.

The responses of the prediction results are slowly decreasing and become constant at the end of the experiment when the sample was deteriorated.

Table 4. The correlation values of each wavelength for cucumber sample.

Wavelength (nm)	635	660	880	890	935	950
RSQ	0.302	0.858	0.888	0.889	0.736	0.907

Response = 0.1165 + 0.1123* (880) + 0.0112* (890) + 0.1379* (950)   (5)
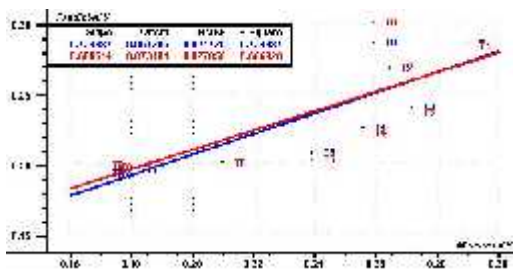

Figure 11. The results of calibration (blue) and validation (red) processes for the cucumber sample.
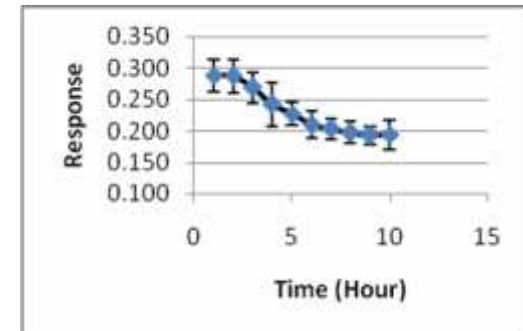

Figure 12 The prediction results of a new data set of the cucumber sample.

Figure 13 shows the scatter plot of the responses from each individual wavelength for the onion sample. The plot shows that few spectral responses were decreasing during the experiment, especially at the wavelength of 635 nm and 935 nm. The response values of these wavelengths were higher at the beginning and then decreased significantly up to five hours before becoming constant for the next three hours. Two wavelengths (660 nm and 950 nm) decrease slowly until to the end of the experiment while the other two (880 nm and 890 nm) almost constant. Table 5 shows that the first three wavelengths which have higher R^2 values are 635 nm, 660 nm, and 935 nm. These three wavelengths were then used in the multispectral setup to generate validation data set for the onion sample.
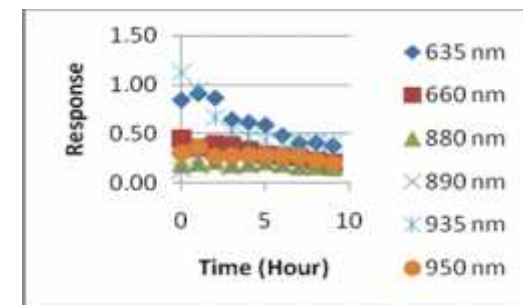

Figure 13 Scatter plot of the response values of the onion sample.

Table 5. The correlation values of each wavelength for onion sample.

Wavelength t (nm)	635	660	880	890	935	950
RSQ	0.921	0.937	0.500	0.201	0.711	0.646

_____

The calibration processes the onion data produced a model of equation as shown in Equation (6). This model can be used to predict the shelf life condition of the onion sample. The R^2 and RMSE values produced by the calibration processes (blue color) were quite low and the values of R^2 and RMSE produced by the validation processes (red color) were even lower as shown in Figure 14. These results indicate that the algorithm is not suitable to evaluate the shelf Life conditions of the onion sample. The prediction processes also produced high variation as represented by wide vertical bars in Figure 15. These results might be caused by simple structures and transparency feature of the onion tissues. One or two wavelength may be adequate to evaluate the shelf life changes of the onion sample.

$$\text{Response} = 0.1841 + 0.0643\ *(635) + 0.0210\ *(660) + 0.0729\ *(935) \qquad (6)$$
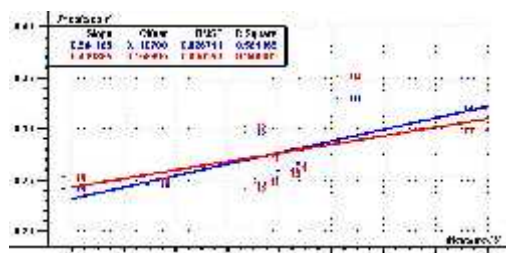

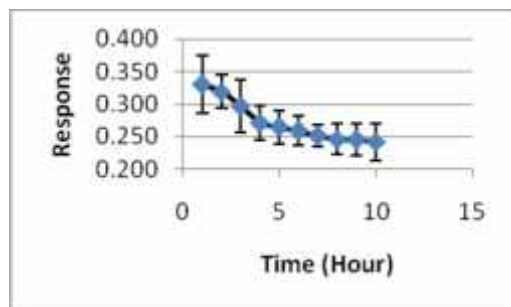Figure 14 The results of calibration (blue) and validation (red) processes for the onion sample.

Figure 15 The prediction results of a new data set of the onion sample.

Conclusion

The scattering setup can be used to detect the shelf life changes of some selected vegetables. Generally, the calibration and validation of the scattering algorithms produced high R^2 values and low RMSE except for the onion sample which only 0.564 and 0.0267 respectively. This exception maybe caused by simple structure of the onion cells which can be detected by using only one or two wavelengths.

The shelf life changes of Chinese cabbage, carrot, chilli, and cucumber can be evaluated using the wavelengths of 880, 890 and 950 nm as its light sources. The scattering algorithm developed in this study can be used to evaluate and to predict the shelf life changes of carrot, chili and cucumber with a good accuracy. However for the Chinese cabbage and carrot, the algorithm is not effective because the prediction responses produced by the algorithm kept increasing even after the samples were dried or deteriorated. Based on these results, the scattering algorithm is considered not suitable to evaluate shelf life changes of Chinese cabbage and carrot samples.

References

[1]. Choi, K.-H., K.-J. Lee, and G. Kim, 2007. *Nondestructive Quality Evaluation Technology for Fruits and Vegetables Using Near-Infrared Spectroscopy*, National Institute of Agricultural Engineering, Rural Development Administration, Republic of Korea: Suwon.

[2}. Michael J. McShane, Cameron, B, D., Cote, G. L., Motamedi, M. Spiegelman, C.H. 1999. *A novel peak-hopping stepwise feature selection method with application to Raman spectroscopy.* Analytica Chimica Acta,. **388**(3): 251-264.

[3]. Renfu Lu and Y. Peng, 2005. *Multispectral scattering measures fruit texture*, in *Laser Focus World*, PennWell: Nashua, NH. p. 99 - 103.

[4]. Eric C. Newgard, 2004. *Near-Infrared Spectroscopy for Analysis of Agricultural Material*. Final Report for Physics 598 OS (Optical Spectroscopy) Fall 2004 Available from: http://online.physics. uiuc.edu/courses/phys598OS/fall05/.

[5]. Stephane Jacquemoud and S.L. Ustin. 2001. *Leaf Optical Properties: A State Of The Art*. in *8th Int. Symp. Physical Measurements & Signatures in Remote Sensing.*. Aussois, France.

[6]. Barbara J. Yoder and R.E. Pettigrew-Crosby, 1995. *Predicting nitrogen and chlorophyll content and concentrations from reflectance spectra (400-2500 nm) at leaf and canopy scales.* Remote Sensing of Environment, **53**(3) 199-211.

[7]. Anatoly A. Gitelson and M. Merzlyak, 1996. *Signature analysis of leaf*

_____

reflectance   spectra: algorithm development   for   remote   sensing   of chlorophyll. Journal of Plant Physiology, **148**(3): 494-500.

[8]. Joseph Penuelas and I. Filella, 1998. *Visible   and   near-infrared   reflectance techniques   for   diagnosing   plant physiological status.* Trends in Plant Science, **3**(4): 151-156.

[9]. Pietro Ceccato, Flasse, S., Jacquemonud, S., Gregoire, M and Tarantola, S. 2001. *Detecting vegetation leaf water content using reflectance in the optical domain.* Remote Sensing of Environment, 2001. **77**(1): 22-33.

[10].   Judith   A.   Abbott.   1999.   Quality measurement of fruits and vegetables. *Postharvest   Biology   and   Technology,* 15(3) 207-225.