Talanta 116 (2013) 266–276

Contents lists available at ScienceDirect

Talanta

journal homepage: www.elsevier.com/locate/talanta

# Application of visible and near infrared hyperspectral imaging for non-invasively measuring distribution of water-holding capacity in salmon flesh

Di Wu, Da-Wen Sun *

*Food Refrigeration and Computerised Food Technology (FRCFT), School of Biosystems Engineering, University College Dublin, National University of Ireland, Agriculture & Food Science Centre, Belfield, Dublin 4, Ireland*

ARTICLE INFO

*Article history:*
Received 7 December 2012
Received in revised form
8 May 2013
Accepted 14 May 2013
Available online 20 May 2013

*Keywords:*
Hyperspectral imaging
Imaging spectroscopy
Fish
Salmon
Water-holding capacity
Chemometrics

ABSTRACT

Water-holding capacity (WHC) is a primary quality determinant of salmon flesh. One of the limiting factors for not having a direct measurement of WHC for salmon quality grading is that current WHC measurements are destructive, time-consuming, and inefficient. In this study, two hyperspectral image systems operated in the visible and short-wave near infrared range (400–1000 nm) and the long-wave near infrared range (897–1753 nm) were applied for non-invasive determination of four WHC indices, namely percentage liquid loss (PLL), percentage water loss (PWL), percentage fat loss (PFL), and percentage water remained (PWR) of salmon flesh. Two calibration methods of partial least square regression (PLSR) and least-squares support vector machines (LS-SVM) were applied, respectively, to establish calibration models of WHC indices based on the spectral signatures of salmon flesh, and the performances of these two methods were compared to determine the optimal spectral calibration strategy. The performances were also compared between two hyperspectral image systems, when full range spectra were considered. Out of 121 wavelength variables, only thirteen (PLL), twelve (PWL), nine (PFL), and twelve variables (PWR) were selected as important variables by using competitive adaptive reweighted sampling (CARS) algorithm to reduce redundancy and collinearity of hyperspectral images. The CARS–PLSR combination was identified as the optimal method to calibrate the prediction models for WHC determination, resulting in good correlation coefficient of prediction ($r_P$) of 0.941, 0.937, 0.815, and 0.970 for PLL, PWL, PFL, and PWR analysis, respectively. CARS–PLSR equations were obtained according to the regression coefficients of the CARS–PLSR models and were transferred to each pixel in the image for visualizing WHC indices in all portions of the salmon fillet. The overall results show that the laborious, time-consuming, and destructive traditional techniques could be replaced by hyperspectral imaging to provide a rapid and non-invasive measurement of WHC distribution in salmon flesh.



## 1. Introduction

Fresh Atlantic salmon is a frequently consumed seafood product worldwide. Eating salmon can measurably improve people's health. Nowadays, consumer's expectation of salmon products with high quality and safety has increased. Superior salmon products should be manufactured to conform to consumers' expectations, leading to the success in the highly competitive salmon industry. Salmon received by the process industries shows extreme diversity, because they are harvested from different farmers who follow different farm systems and deal with multiple breeds. It is difficult for the process industries to produce products with consistent quality that is critical for grading based on expected salmon palatability. Inconsistencies in quality would lead to the decline in customer satisfaction and subsequently in market shares. Consumers are willing to pay more for salmon products if high quality is guaranteed.

The measurement of the water-holding capacity (WHC) of muscle is a useful and effective way for describing quality changes in food and food products [1]. WHC is defined as the ability of muscle to retain water or resist water loss [2] and is of great significance for commercial value and consumer acceptance. The WHC of muscle to retain water is affected by several factors such as proteins oxidation, proteolytic activity of tenderising enzymes, and cross-linking of myofibrillar proteins [3]. The change of WHC to a large extent affects the texture of fish. Fish flesh usually becomes tougher when accompanied by a reduction of WHC [4]. Poor WHC will result in high drip and purge loss and is now of

* Corresponding author. Tel.: +353 1 7167342; fax: +353 1 7167493.
*E-mail address:* dawen.sun@ucd.ie (D.-W. Sun).
*URLS:* http://www.ucd.ie/refrig, http://www.ucd.ie/sun (D.-W. Sun).

0039-9140/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.talanta.2013.05.030

significant industry concern [5]. Therefore, it is critical to determine the WHC of salmon products throughout the entire process and management system to keep the products with superior quality and to understand how the different processing regimes influence the quality of the end product. Currently, the determination of WHC is commonly conducted by laborious and destructive techniques, such as gravimetric method [6], centrifugation method [2], filter paper wetness method [7], and cooking loss method [8]. However, these conventional methods to determine WHC are extremely time-consuming, which is a limitation for industrial online inspection. Therefore, there is a need for a rapid and non-destructive WHC evaluation method under industrial environment and the incorporation of such a method with the existing salmon inspection systems will bring in great benefits to the industry for efficient salmon product quality assessment.

Optical techniques are one of the most accessible ways instead of the destructive and time-consuming sensory and instrumental methods for predicting quality of food rapidly and non-invasively. Of the various non-invasive optical methods, near-infrared (NIR) spectroscopy has shown its ability to measure WHC of fish flesh [9] and is a potential method for on-line implementation. Indeed, visible and near-infrared (Vis–NIR) spectroscopy has also been intensively applied in quantitative evaluation of other quality attributes of salmon products, such as fat, moisture, and sodium chloride [10,11]. However, values of WHC vary along heterogeneous materials, such as salmon products. It is a difficult task to obtain the overall distribution of WHC for the salmon flesh at different positions using spectroscopy technique because of its small sample area (limited spatial information). In addition, traditional destructive techniques also cannot provide the WHC details for salmon flesh, but only its bulk value. The lack of knowledge on the concentration gradients and spatial distributions of WHC has seriously limited the ability of the salmon industry to provide consumers with products of consistent quality and safety. On the other hand, imaging techniques (generally refer to computer vision [71–73]) in the forms of monochromatic or colour images provide abundant information of a sample at a pixel-wise level and has been used for inspecting food quality [12,13]. However, they have limited capacity of inspecting constituent properties such as moisture, fat and protein, as they only work in the visible spectral range.

As an extension of both spectroscopy and imaging techniques, hyperspectral imaging technique has been widely accepted as a smart and promising analytical tool in quality evaluation and assurance of various agricultural and food products [14,15]. It can provide both spectral and spatial information of an object simultaneously, making hyperspectral imaging very powerful and advantageous in food quality control and assessment than conventional spectroscopic and imaging techniques [16]. Hyperspectral image data are made up of a series of congruent three dimensional "*hypercube*" ($x$, $y$, $\lambda$), which contains two-dimensional spatial information ($x$, $y$) as well as one-dimensional spectral information ($\lambda$). These data provide a large amount of information and can be analysed to characterize a specific object more objectively and reliably. Moreover, the obvious advantage of hyperspectral imaging is the ability of visualization of quality attributes distribution, which is not obtainable by spectroscopy or conventional destructive methods. Recently, hyperspectral imaging technique has received considerable applications for food quality and safety assessment [17,18], such as beef [19], pork [20], lamb [21], prawn [22], egg [74], citrus [23], and cucumber [24]. In particular, endeavours of using hyperspectral imaging have been reported for salmon quality evaluation in recent years, such as fat [25], colour [26], NaCl content [27], ice fraction [28], moisture [29], and microbial spoilage [30]. However, no reports of using hyperspectral imaging to determine WHC of salmon flesh have been found.

Given the limited information on the usefulness of hyperspectral imaging systems to determine WHC in fish or fish products, the main aim of this study was to investigate the feasibility of hyperspectral imaging system in the visible and near-infrared spectral region of 400–1753 nm for predicting WHC distribution of salmon fillets. The specific objectives of the current work were to (1) acquire hyperspectral images of salmon fillets in the spectral range of 400–1753 nm; (2) extract spectral information and identify the optimal wavelengths that are most correlated to WHC prediction; (3) build a quantitative relationship between the WHC values and spectral information; (4) compare the prediction abilities of spectral ranges of 400–1000 nm and 897–1753 nm, respectively; and (5) display the WHC distribution in the salmon fillets by developing an image processing algorithm for visualizing WHC in all pixels of the image.

## 2. Materials and methods

### 2.1. Samples preparation

Salmon samples were prepared from local supermarkets in Dublin, Ireland. There were a total of eighteen farmed Atlantic salmon fillets (*Salmo salar*) originated from Norway (nine fillets) and Scotland (nine fillets). The salmon fillets were fresh and of superior quality. In order to establish a robust calibration model, a reasonable variation of attribute values should be guaranteed for the examined salmon samples. On the other hand, the WHC could vary within the same fillet and different WHC distributions could be found for the fillets with similar mean WHC values. Therefore, subsampling was conducted to obtain a wide span of WHC for the model calibration. In the laboratories of Food Refrigeration & Computerized Food Technology (FRCFT), University College Dublin (UCD), Ireland, each salmon fillet was first scanned by two hyperspectral imaging systems and then it was cut into several subsamples with a cuboid shape of 1.5 cm × 1.5 cm × 1.0 cm (length × width × thickness) from different locations within the fillet. A total of 98 cubed samples were obtained from the eighteen fillets and were divided into a calibration set and a prediction set. In the separating process, all samples were first sorted according to their respective reference WHC values. One sample of every five samples was then selected into the prediction set, and remaining four samples were selected into the calibration set. As a result, 79 samples were used for the calibration and the remaining 19 samples were used for prediction. The reference WHC values of these cubed samples were then determined immediately after the acquisition of hyperspectral images using the centrifugation method, in which each sample was first pre-weighed and then placed in a plastic tube with a pre-weighed filter paper, and was then centrifuged (1500 g, 10 °C, 10 min) and reweighed. The time between imaging and WHC measurement for each cuboid subsample was less than three minutes. Four weight indices, namely the weight of sample (S), the weight of pre-weighed filter paper (V1), the weight of the wet paper (V2), and the weight of the wet paper after dried to constant weight (V3), were obtained from the centrifugation method. Four indices of WHC were determined in this study, which were the percentage liquid loss (PLL), percentage water loss (PWL), percentage fat loss (PFL), and percentage water remained (PWR). The first three indices were calculated based on S, V1, V2, and V3 according to the equations in Ref. [31], and the forth index was the percentage of water retained in the salmon flesh after centrifugation. Table 1 summarizes the variations in four WHC indices of the salmon samples. There were wide variations obtained in WHC indices, which was important to establish accurate and robust calibration models [32]. Key steps for the whole procedure are shown in Fig. 1.

**Table 1**
Reference results of four water-holding capacity (WHC) indices, namely percentage liquid loss (PLL), percentage water loss (PWL), percentage fat loss (PFL), and percentage water remained (PWR) of salmon flesh samples measured by the centrifugation method.

| Statistics | Calibration set (100 samples) | | | | Prediction set (60 samples) | | | |
|---|---|---|---|---|---|---|---|---|
| | PLL (%) | PWL (%) | PFL (%) | PWR (%) | PLL (%) | PWL (%) | PFL (%) | PWR (%) |
| Maximum | 12.529 | 11.693 | 2.324 | 72.827 | 11.630 | 10.144 | 1.946 | 72.381 |
| Minimum | 1.143 | 0.863 | 0.004 | 56.446 | 1.735 | 1.605 | 0.057 | 56.867 |
| Mean | 6.405 | 5.683 | 0.724 | 64.838 | 6.330 | 5.630 | 0.689 | 64.707 |
| Standard deviation | 3.331 | 2.943 | 0.590 | 4.651 | 3.299 | 2.843 | 0.538 | 4.544 |
| Range | 11.387 | 10.830 | 2.320 | 16.380 | 9.895 | 8.540 | 1.889 | 15.514 |

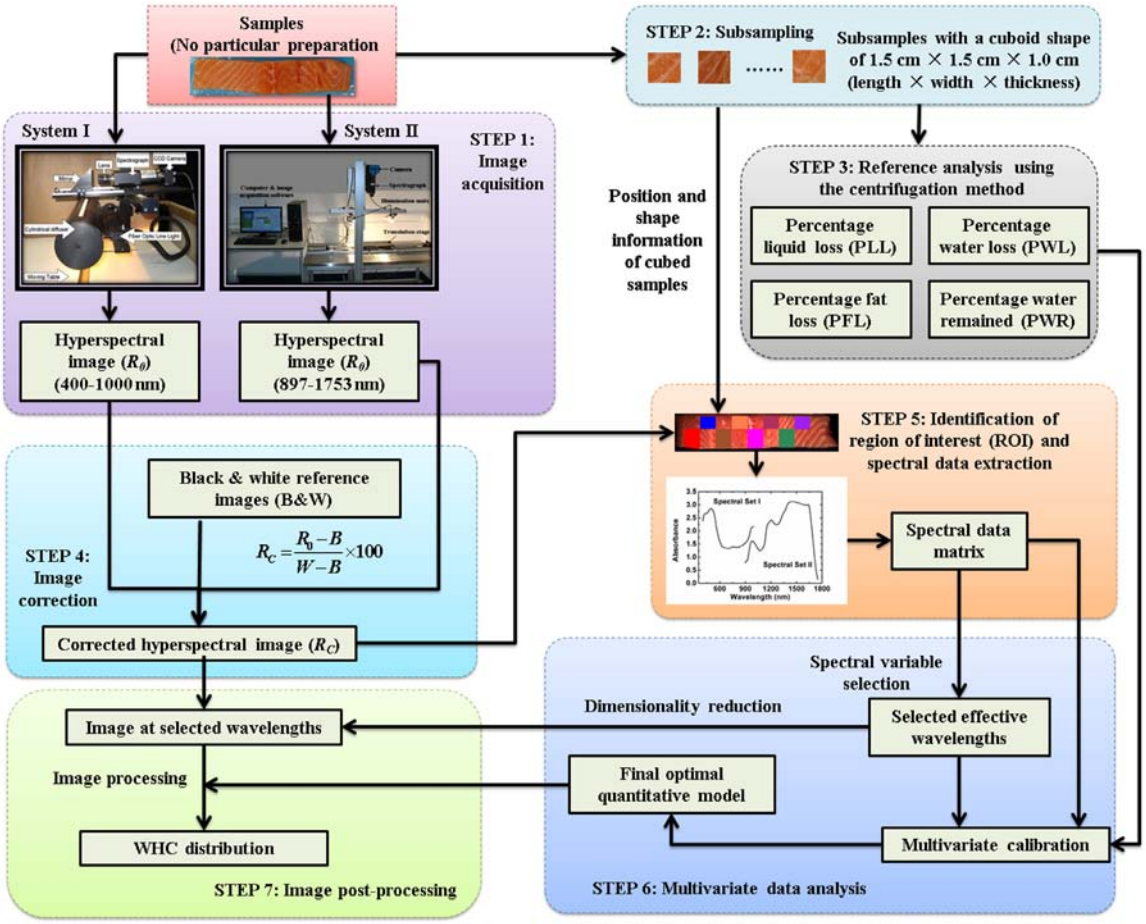


**Fig. 1.** Key steps for the whole procedure of hyperspectral image analysis. The picture of System I is from the literature [69] and the picture of System Ii is from the literature [70].

### 2.2. Hyperspectral imaging system

Hyperspectral images of salmon fillets were acquired by two hyperspectral imaging systems in reflectance mode at 400 −1000 nm (measured by System I) and 897–1753 nm (measured by System II), respectively. Both systems were in line-scanning configuration, which are also known as the pushbroom method. Such method records a whole line of an image as well as spectral information simultaneously corresponding to each spatial pixel in the line. The specific descriptions of Systems I and II could be found in the literatures [30] and [26], respectively. The spectral resolutions were 5 nm and 5.25 nm for System I and System II, respectively. The binning sizes for both System I (VNIR data) and System II were $1 \times 1$ binning.

### 2.3. Images acquisition and correction

After the surface moisture was wiped by paper towels, each salmon fillet was placed on the moving table to be scanned line by line to acquire the raw hyperspectral images ($I_O$) in the Band Interleaved by Line (BIL) format. The time of image acquisition for each fillet was about 2–3 s. The two 500 W halogen lamps in System II were placed about 40 cm away from the salmon sample being imaged. Moreover, the lamps in System II were blocked using two opaque wood planks during the intervals of imaging to prevent the moving table overheated. Therefore, there was few potential cooking affects for the fillets due to the lamps in System II. For each line-scan, a whole line of an image for the sample was recorded as well as spectral information simultaneously

corresponding to each spatial pixel in the line. As a result, a 2-D image ($y$, $\lambda$) was acquired for each line-scan, containing the whole spectral dimension ($\lambda$) with one spatial dimension ($y$). The sample was moving on the $x$ direction at a controlled speed and the line-scan was repeated along the direction of $x$ dimension of the sample. At last, a three dimensional hyperspectral image was created, recorded and stored with a two dimensional image ($x$, $y$) and one dimensional spectra ($\lambda$). The spectral data were used for the WHC prediction and the spatial information/image was used for generating the WHC distribution. The hyperspectral images measured using System I had 121 bands (400–1000 nm) in $\lambda$ direction, and Systems II acquired the hyperspectral images with 256 bands (897–1753 nm). There were 580 pixels and 320 pixels in the spatial dimension ($y$) of the hyperspectral images measured by System I and System II, respectively. The numbers of pixels in the spatial dimension ($x$, moving direction) of the hyperspectral images depended on the moving distances of the sample in the $x$ direction. As signal to noise ratios were good for the whole wavelength regions, no bands were eliminated for the full range spectral analysis. At last, two sets of spectral data were obtained, which had the wavelength ranges of 400–1000 nm (Spectral Set I, 121 wavelength variables) and 897–1753 nm (Spectral Set II, 256 wavelength variables), respectively. To calibrate the raw hyperspectral images ($I_O$), extra black and white reference images were required for correcting the detector signal intensity into reflectance spectra, resulting in minimizing the effects of illumination source and detector sensitivity. The black reference image ($B$) was acquired by recording a hyperspectral image after turning off the light source and covering the camera lens with its black cap completely. The white reference image ($W$) was obtained by collecting a spectral image of a uniform white calibration tile made of Teflon (Gilden Photonics Ltd., Glasgow, UK). The calibrated hyperspectral image ($I_C$) was calculated using the following equation:

$$I_C = \frac{I_O(\lambda) - B(\lambda)}{W(\lambda) - B(\lambda)} \times 100 \quad (1)$$

The corrected hyperspectral images were used for subsequent analysis of spectral extraction, variable selection, multivariate calibration, and distribution visualization purposes. The correction for the hyperspectral images were performed using the software ENVI v4.6 (Research Systems Inc., Boulder, CO, USA).

## 2.4. Data analysis

Hyperspectral imaging has the capability of providing spectral information for each pixel in the image, which is its main advantage against common spectroscopy technique. Such detailed spectral information of each pixel provides an opportunity to show the spatial variation of the inspected attribute of food products by the way of generating distribution maps of the attribute. In order to display such variation, a qualitative or quantitative relationship should be established according to the correlation between spectral response (optical properties) and the target quality feature, usually a chemical component or physical attribute. Such relationship should be built to relate the spectrum according to its reference attribute value. Otherwise, such relationship would be meaningless. In practice, a certain size of salmon flesh was required for the WHC measurement using the centrifugation method mentioned above. This poses a problem for building a spectral model that it is practically impossible to provide the reference WHC value for each pixel spectrum using the centrifugation method. To overcome this problem, in this study, calibration models were established using the mean spectra over the region of salmon flesh (a cuboid shape of 1.5 cm × 1.5 cm × 1.0 cm) on which the reference WHC value was available to be measured.

### 2.4.1. Identification of region of interest (ROI) and spectral data extraction

After image acquisition and calibration, the pixels of the cubed samples should be selected from the acquired hyperspectral images of the fillets, so that the spectra of the cubed samples could be extracted, and the quantitative calibration model could be then established based on the spectra of these cubed samples and their corresponding WHC values. When the cubed samples were cut from the salmon fillets, the positions and shapes of these cubed samples were recorded. The pixels of the cubed samples were selected from the acquired hyperspectral images according to the recorded position and shape data using the Region of Interests (ROI) Function of ENVI v4.6 software (Research Systems Inc., Boulder, CO, USA). The pixels of a cubed sample were defined as the ROI of this sample. After the pixels of different ROIs were selected, the spectrum of each pixel within the identified ROI of each cubed sample was extracted. A mean spectrum of the ROI was then calculated based on the spectra of all pixels within this ROI, and was used as the spectrum of the cubed sample that the ROI belonged to. The extracted mean spectral data of each sample were then arranged in spectral matrix ($X$) where the columns of this matrix represent the wavelengths (121 variables for Spectral Set I, 256 variables for Spectral Set II) and the rows of this matrix represent samples (98 samples). In addition, one column vector ($Y$) containing the values of one WHC index of the samples was concatenated to this matrix. Prediction models were then established using multivariate data analysis between the values of different WHC indices and the spectral data extracted from all samples in two Spectral Sets, respectively. In future applications, the WHC indices at the pixel scale can be predicted directly from the measured spectrum of each pixel of salmon flesh based on the established prediction models.

### 2.4.2. Model calibration

On the basis of the extracted spectral information, the WHC indices were determined quantitatively using multivariate data analysis. The purpose of multivariate data analysis is to extract meaningful information from the massive quantity of hyperspectral cube data for determining the properties of objects in an efficient way. In this study, such determination was implemented by building calibration models to predict four WHC indices of salmon samples using their corresponding spectral information. Partial least squares regression (PLSR) was carried out to perform linear calibration between spectral matrix ($X$) and the values of one of the WHC indices ($Y$) obtained from the centrifugation method. The same spectral dataset was also used in least-squares support vector machines (LS-SVM) analysis, which is a nonlinear calibration algorithm. The performances of these two calibration methods were compared using the data of Spectral Sets I and II, respectively, to determine which method is the best for the WHC determination of salmon flesh. The optimal model was utilized for further WHC visualization.

PLSR is a widely used chemometric method for the calibration of a multivariate spectral model with explanatory or predictive purposes [33], and has been commonly used for spectral analysis [34–38]. By decomposing both $X$ and $Y$ features, PLSR extracts a set of orthogonal factors called latent factors (LVs), and searches the optimal functions by minimizing the error sum squares [39]. Different from principal component regression (PCR), such decomposition to estimate the LVs in PLSR is being carried out with consideration of $Y$-variable, so that the first LVs are those that are most relevant to predicting the $Y$-variable. Normally, most of the useful spectral information can be explained by the first few LVs while the remaining LVs mainly include random noise or linear dependencies between $X$-variables and $Y$-variables [40]. Support vector machine (SVM) is a

widely used supervised learning method proposed by Cortes and Vapnik [41]. It uses a kernel function to map the data input space to a high-dimensional feature space, where the non-linear separability problem becomes the linear separability problem [42]. LS-SVM is a least squares version of support vector machines (SVM) proposed by Suykens and Vandewalle [43], which applies least squares error in the training error function [44]. Different from classical SVM solving a convex quadratic programming (QP) problem, LS-SVM finds the solution by solving a set of linear equations. Radial basis function (RBF) kernel was used as the kernel function of LS-SVM, as it is a nonlinear function and a more compact supported kernel. Comparing with other kernel functions, RBF could handle the linear and nonlinear relationships between the spectral matrix ($X$) and target attributes ($Y$) [45], and therefore has been used in the establishment of LS-SVM models in many works [46,47]. There are also many other works show the SVM model with RBF kernel had more capability in prediction than other kernels [48–50] or had same prediction capability [51,52]. In order to determine the optimal parameter values of LS-SVM model, namely the regularization parameter $\gamma$ and the RBF kernel function parameter $\sigma^2$, a grid-search technique with leave one out cross-validation was used in the LS-SVM calibration based on the samples in the calibration set. All computations of model calibration and prediction were operated with Unscrambler 9.7 (CAMO PROCESS AS, Oslo, Norway) and Matlab 2011a software (The Mathworks, Inc., Natick, MA, USA).

In the model calibration, PLSR and LS-SVM were applied, respectively, to the establishment of calibration models according to the spectral information of samples in the calibration set with their reference WHC values. After the model was established, the prediction set was then analyzed in order to estimate the actual predictive capability of the established models, to minimize the concrete risk of overfitting and to avoid chance correlations. Therefore, the prediction set was independent to the established calibration model. The prediction set was used only after model establishment. The predictive ability of the model was evaluated according to some statistics, such as correlation coefficient of calibration ($r_C$), root mean square error of calibration (RMSEC) and coefficient of determination of calibration ($R_C^2$) for the calibration process, and correlation coefficient of prediction ($r_P$), root mean square error of prediction (RMSEP), residual predictive deviation (RPD), and coefficient of determination of prediction ($R_P^2$) for the prediction process. The standard for evaluating the performance of a model is that a good model should have high correlation coefficients ($r_C$ and $r_P$), high coefficient of determination ($R_C^2$ and $R_P^2$), and the low root mean square errors (RMSEC and RMSEP) as well as a small difference between RMSEC and RMSEP.

### 2.4.3. Selection of important variables

Hyperspectral images are characterized by its excessive degree of dimensionality with collinearity and redundancy among contiguous variables (wavelengths). Some congruent wavelengths are related to the similar constituents, and consequently contain much of the same information [53]. On the other hand, the wavelength variables correlated with the neighboring variables were considered as redundancy. Besides collinearity and redundancy among spectral variables, some bands may contain irrelevant information or noise for calibration model than pertinent information to quality attributes of samples [54]. Selecting some important spectral variables instead of the full-spectrum has showed better predictive results in some cases [55–57]. Another advantage is that the selected important variables could be used to build multi-spectral imaging systems suitable for on-line applications to meet the speed requirement of real-time production lines [58]. Therefore, it is important to select the important variables from the whole spectral range of samples that contain useful information for the spectral analysis and the establishment of the model.

In this study, competitive adaptive reweighted sampling (CARS) algorithm [59] was used to select the most important variables that had less redundancy and contributed most in WHC determination of the salmon flesh. Previous works show better predictions could be obtained by CARS when compared to Monte Carlo uninformative variable elimination (MC-UVE), moving window partial least squares regression (MWPLSR), successive projections algorithm (SPA), genetic algorithms [60–62]. Similar to the regression coefficient method that selects variables (wavelength) corresponding to a large regression coefficient value (regardless of the sign) of PLSR model, CARS also uses the absolute values of regression coefficients of PLSR model as an index for evaluating the importance of each wavelength. In general, the variable with larger absolute coefficients is considered as a good candidate for effective calibration. CARS has $N$ sequent Monte Carlo (MC) sampling runs and in each of them a PLSR model is established using not all the samples but a certain percentage of samples randomly selected. The purpose of such strategy is to select the variables that are of high adaptability regardless of the variation of training samples [59]. In each run, after each Monte Carlo model sampling, exponentially decreasing function (EDF) is used to eliminate the variables with relatively small absolute regression coefficients. There are two stages in EDF process, where a 'fast selection' is usually in the first stage and a 'refined selection' in the second stage. Variables are reduced in a rapid manner in the first stage and decreased gently in the second stage. Following EDF-based enforced variable reduction, adaptive reweighted sampling (ARS) is carried out to further select variables utilizing the principle of 'survival of the fittest' that is the basis of Darwin's Evolution Theory [63]. At last, root mean square error of cross validation (RMSECV) is calculated to determine the optimal variable set. As an example, the variation trends of some key parameters in CARS along with the increment of sampling runs based on the analysis of the spectra of samples in the calibration set in

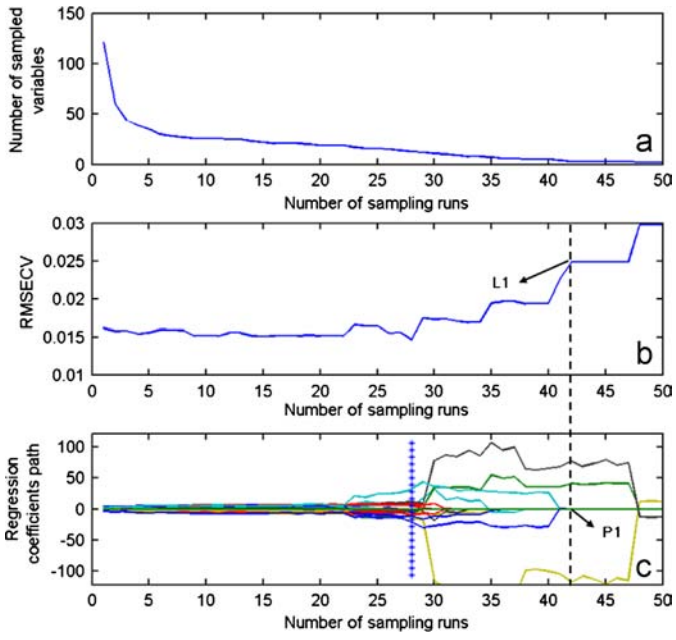


**Fig. 2.** Variation trends of some key parameters in CARS along with the increment of sampling runs based on the analysis of the spectra of samples in the calibration set in Spectral Set I for percentage liquid loss (PLL) analysis. (a) the number of sampled variables, (b) 5-fold RMSECV values, and (c) regression coefficients of each variables. The line (marked by asterisk) denotes the optimal point where 5-fold RMSECV values achieve the lowest.

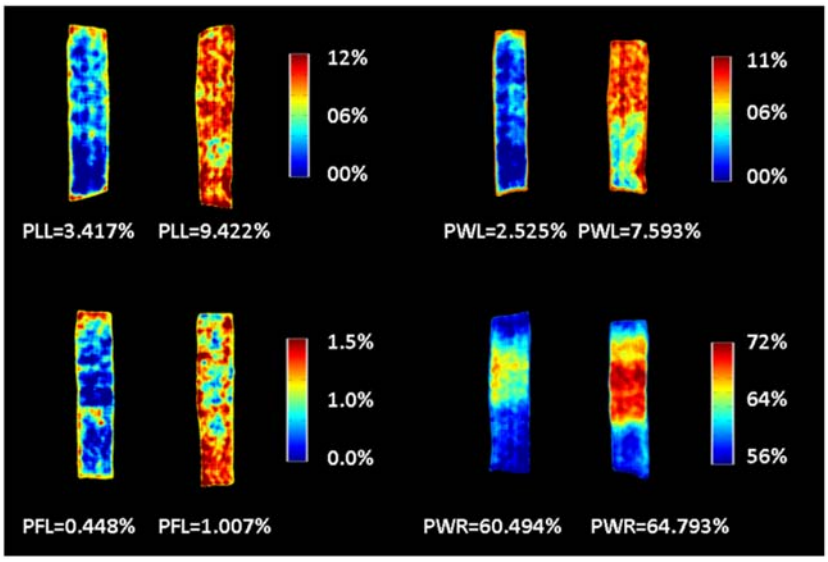


**Fig. 3.** Distribution maps of four water-holding capacity (WHC) indices, namely percentage liquid loss (PLL), percentage water loss (PWL), percentage fat loss (PFL), and percentage water remained (PWR) within some salmon fillets. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Spectral Set I for PLL analysis are shown in Fig. 2. Fig. 2a shows the variation trend of the number of sampled variables, where the important variables were kept while other insignificant variables were removed in a stepwise and efficient way. Fig. 2b shows the tendency of 5-fold RMSECV values along with the increase in the number of sampling runs. The RMSECV values did not change much from sampling runs 1–25, when the uninformative variables were being eliminated. Then the RMSECV reached the smallest value that was denoted by asterisk line to indicate the optimal subset. After that, the RMSECV increased abruptly due to the removing of some informative variables. For instance, the RMSECV rises up promptly indicated by dot line L1, because the wavelength (indicated by P1) is eliminated. These variables are important to the calibration model, which is proved by removing some of these variables, resulting in the increment of RMSECV value of the model established based on the remained variables. By analyzing the regression coefficient path of each variable shown in Fig. 2c, the process of CARS calculation could be understood more in details. Each line in Fig. 2c has its own regression coefficients at different sampling runs. The coefficients drop to zero when their corresponding variables were removed by CARS, which is somewhat like the incompetence species are exterminated. On the other hand, the coefficients of these important variables get larger and larger so that they will get more probability to survive, just like the 'survival of the fittest' in Darwin's Evolution Theory [59]. As shown in Fig. 2c, all the variables have small absolute values of regression coefficients at the beginning of the sampling run; and with the increase in the number of runs, the coefficients of some variables become larger, while others become smaller. In this study, the processes of CARS selection were performed with the aid of Matlab 2011a software (The Mathworks, Inc., Natick, MA, USA). The model establishment using the full range spectra (121 wavelengths for Spectral Set I and 256 wavelengths for Spectral Set II) was called Method I; while using only the important wavelengths selected by CARS for WHC determination was called Method II throughout this paper.

#### 2.4.4. *Visualization of WHC distribution within salmon fillets*

The visualization of WHC's spatial distribution is important to observe the variation in WHC from sample to sample and even from position to position within the same sample. By using hyperspectral imaging technique, a distribution map of different WHC indices within the whole fillet was generated based on the spatial position of every pixel and the corresponding colour values. Only the spectral data at the selected important variables/wavelengths were used to calibrate the prediction models for visualization, so that the redundancy of hyperspectral images could be reduced. After comparing calibration models established using different spectral sets and calibration methods, the optimal multivariate model was chosen and applied in a pixel-wise manner for the hyperspectral images of salmon fillets to produce distribution maps. There were two steps of visual representation. First, the value of one of four WHC indices of each pixel was calculated by inputting the spectrum of this pixel into the established prediction model. Such process was repeated for all pixels in the hyperspectral image of a fillet to generate a prediction matrix that had values of this WHC index in every pixel of the fillet. Then, a linear colour scale (colour bar in Fig. 3) was used to map the predicted values of every pixel into different colours (the pixels with high values were shown in red and those with low values were shown in blue.), resulting in generating a pseudo colour matrix (called distribution map). Resulted visualization maps were displayed in different colours to enhance the visualization display and to make them easily interpretable for discovering of the difference of WHC within one fillet. All visualization routines were programmed in Matlab 2011a software (The Mathworks, Inc., Natick, MA, USA).

## 3. Results and discussion

### 3.1. *Spectral features of salmon fillets*

The spectral profiles extracted from the hyperspectral images of the salmon samples in the Spectral Sets I (400–1000 nm) and II (897–1753 nm) are shown in Fig. 4, where the main components of salmon flesh were found having their own spectral absorption. This is important to predict quality attributes of salmon flesh in either qualitative or quantitative manner. The spectral pattern in the visible spectral region (400–700 nm) shows high absorbance in blue and green regions (about 400–550 nm) and low absorbance in red region (about 600–700 nm), which explains why salmon flesh gets its red/pink colour. On the other hand, there are

some absorption peaks over the entire near infrared spectral region (700–1753 nm), which were assigned to the main components of salmon flesh. Both Spectral Sets I and II show an absorbance peak around 980 nm due to presence of water (O–H stretching second overtone). In Spectral Set I, a weak absorbance was found around 760 nm that was assigned to O–H stretching third overtone of water. The absorbance peak of O–H stretching first overtone for water was presented around 1460 nm in Spectral Set II. Another obvious absorbance peak at around 1210 nm in Spectral Set II was due to fat presence (C–H stretching second overtone). Spectral Set I also has an absorbance peak for fat at around 930 nm assigned to the third overtone C–H stretch in the methylene group of fat, however it is immersed in water information. The information of protein is contained in Spectral Set II around 1500 nm due to the N–H stretching first overtone, leading to higher absorbance at the right side of the water absorbance peak (around 1460 nm) than the left side. Fig. 4 also shows the information about the range of the absorbance values. When many samples were considered, their spectral profiles had a similar pattern without prominently extraordinary peak. The only difference of absorbance was the magnitude owing to the variation of composition contents and uneven physical structure of the salmon fillets as well as unfixed scatter of the surface. In general, the spectral data contained the information of the main components of salmon flesh, namely water, fat, and protein. The contents of these components affect the WHC of salmon flesh. Therefore the spectral data containing the information of these components could be used for WHC determination. However, no typical feature peaks of WHC were found in the spectral profiles of salmon samples, causing the difficulty to directly observe spectra for the WHC determination. Chemometrics were therefore applied for mining spectral data, resulting in successfully establishing quantitative prediction models for WHC determination in salmon flesh.

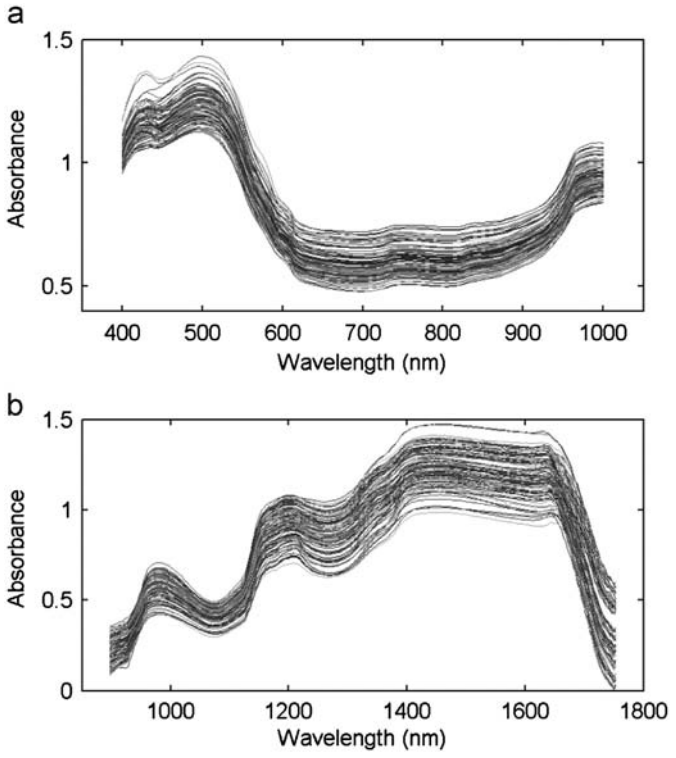


**Fig. 4.** Spectral profiles of the salmon chops in the Spectral Sets I (400–1000 nm) (a) and II (897–1753 nm) (b). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

### 3.2. Calibration of WHC indices using full range spectra

As no feature peaks were found for the WHC assessment in both Spectral Sets I and II as shown in Fig. 4, the model calibration was executed based on the full range spectra. Table 2 shows the prediction results of four WHC indexes (PLL, PWL, PFL, and PWR) of salmon samples by analyzing the full range spectra in both Spectral Sets I and II using two calibration algorithms of PLSR and LS-SVM, respectively. By analyzing Spectral Set I, the models for predicting three WHC indices (PLL, PWL, and PWR) had $r_C$ and $r_P$ higher than 0.9 for both PLSR and LS-SVM methods. The determination of WHC indices using Spectral Set II was not as good as that using Spectral Set I. In general, the average RMSEP of the PLSR and LS-SVM models increased by 63.4%, 68.9%, 30.1%, and 74.1% for PLL, PWL, PFL, and PWR, respectively, indicating that Spectral Set I was more suitable for the WHC determination of salmon flesh than Spectral Set II. This might be because water as the dominant component of salmon flesh had stronger absorptions in Spectral Set II than Spectral Set I. The spectral information of other components affecting WHC was immersed in the strong

**Table 2**
Prediction results of four water-holding capacity (WHC) indices, namely percentage liquid loss (PLL), percentage water loss (PWL), percentage fat loss (PFL), and percentage water remained (PWR) of salmon samples by analyzing the full range spectra.

| WHC index | Spectral Set | Calibration | Variable number | LVs[a] | Calibration | | | Prediction | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | $r_C$ | RMSEC (%) | $R_C^2$ | $r_P$ | RMSEP (%) | RPD | $R_P^2$ |
| PLL | I | PLSR | 121 | 11 | 0.939 | 1.143 | 0.881 | 0.922 | 1.256 | 2.567 | 0.845 |
| | I | LS-SVM | 121 | / | 0.946 | 1.076 | 0.894 | 0.925 | 1.222 | 2.628 | 0.855 |
| | II | PLSR | 256 | 8 | 0.895 | 1.480 | 0.800 | 0.729 | 2.336 | 1.377 | 0.471 |
| | II | LS-SVM | 256 | / | 0.969 | 0.833 | 0.937 | 0.843 | 1.728 | 1.859 | 0.711 |
| PWL | I | PLSR | 121 | 11 | 0.945 | 0.960 | 0.892 | 0.915 | 1.134 | 2.449 | 0.832 |
| | I | LS-SVM | 121 | / | 0.964 | 0.783 | 0.928 | 0.916 | 1.118 | 2.378 | 0.837 |
| | II | PLSR | 256 | 7 | 0.878 | 1.401 | 0.771 | 0.748 | 2.195 | 1.463 | 0.532 |
| | II | LS-SVM | 256 | / | 0.941 | 1.009 | 0.881 | 0.819 | 1.607 | 1.724 | 0.663 |
| PFL | I | PLSR | 121 | 7 | 0.799 | 0.353 | 0.638 | 0.789 | 0.342 | 1.536 | 0.576 |
| | I | LS-SVM | 121 | / | 0.856 | 0.304 | 0.731 | 0.809 | 0.317 | 1.672 | 0.633 |
| | II | PLSR | 256 | 6 | 0.724 | 0.405 | 0.524 | 0.654 | 0.432 | 1.311 | 0.322 |
| | II | LS-SVM | 256 | / | 0.839 | 0.328 | 0.687 | 0.724 | 0.425 | 1.448 | 0.340 |
| PWR | I | PLSR | 121 | 11 | 0.971 | 1.114 | 0.942 | 0.965 | 1.251 | 3.615 | 0.920 |
| | I | LS-SVM | 121 | / | 0.974 | 1.049 | 0.949 | 0.966 | 1.277 | 3.621 | 0.919 |
| | II | PLSR | 256 | 11 | 0.969 | 1.139 | 0.939 | 0.851 | 2.425 | 1.825 | 0.699 |
| | II | LS-SVM | 256 | / | 0.983 | 0.863 | 0.965 | 0.908 | 1.947 | 2.334 | 0.806 |

[a] Number of latent variables.

**Table 3**
Prediction results of four water-holding capacity (WHC) indices, namely percentage liquid loss (PLL), percentage water loss (PWL), percentage fat loss (PFL), and percentage water remained (PWR) of salmon samples by analyzing only the important wavelengths selected by competitive adaptive reweighted sampling (CARS).

| WHC index | Calibration model | Variable number[a] | LVs | Calibration | | | Prediction | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | $r_C$ | RMSEC (%) | $R_C^2$ | $r_P$ | RMSEP (%) | RPD | $R_P^2$ |
| PLL | PLSR | 13 | 13 | 0.948 | 1.059 | 0.898 | 0.941 | 1.147 | 2.888 | 0.872 |
| | LS-SVM | 13 | / | 0.953 | 1.000 | 0.909 | 0.947 | 1.060 | 3.101 | 0.891 |
| PWL | PLSR | 12 | 10 | 0.949 | 0.922 | 0.901 | 0.937 | 0.964 | 2.867 | 0.878 |
| | LS-SVM | 12 | / | 0.958 | 0.838 | 0.918 | 0.927 | 1.046 | 2.654 | 0.857 |
| PFL | PLSR | 9 | 9 | 0.826 | 0.331 | 0.681 | 0.815 | 0.312 | 1.682 | 0.645 |
| | LS-SVM | 9 | / | 0.824 | 0.333 | 0.678 | 0.814 | 0.321 | 1.643 | 0.624 |
| PWR | PLSR | 12 | 8 | 0.969 | 1.139 | 0.939 | 0.970 | 1.132 | 4.034 | 0.934 |
| | LS-SVM | 12 | / | 0.987 | 0.740 | 0.974 | 0.971 | 1.222 | 4.160 | 0.924 |

[a]Number of latent variables.

absorptions of water in Spectral Set II [64]. On the other hand, the effect of intense water bands in the long-wave near infrared spectral region (Spectral Set II) can be diminished in short-wave near infrared spectral region (a part of Spectral Set I) [65]. Therefore, Spectral Set I had less influence from water than Spectral Set II.

A previous work on determining WHC of broiler breast meat also shows that important wavelengths with high loadings were mostly in Spectral Set I rather than Spectral Set II [64].

Although Spectral Set II had worse results for WHC prediction than Spectral Set I, its prediction result for PWR ($r_P$ of 0.851) was better than that of the established PLSR models ($r_P$ of 0.70–0.72) in another study using near infrared spectroscopy (1000–2222 nm) to determine PWR in frozen cod [9]. This might be because the sampling positions of PWR and near infrared spectral measurements were not the same in that work. The WHC distribution maps within salmon fillets (Fig. 3) show that there were WHC variations along the fillet, which is due to the varying of flesh components. Therefore, sampling is an important factor for the point detection methods like the centrifugation method and spectroscopy technique in the WHC evaluation of salmon fillet because of its heterogeneity. On the other hand, based on the spatial information contained in the hyperspectral images, the spectral data could be extracted for every sampling point (cubed samples), whose reference WHC values could also be measured using the centrifugation method. In other words, the spectral data and the reference WHC values could have the same sampling position by using hyperspectral imaging, so that a better prediction was obtained in this work.

Different from other WHC indices, the predictions of PFL were not satisfactory ($r_P$ were between 0.654 and 0.809) when either Spectral Set I or Spectral Set II was considered. This could be elucidated to the narrow range of PFL values of the samples (0.004–2.324%) involved in model calibration, meanwhile the ranges of other three WHC indices used for model calibration were 11.387% (PLL), 10.830% (PWL), and 16.380% PWR). Since salmon flesh is a very complex medium, various samples with a wide range in WHC should be used for the model calibration. Enlarging the range of target attribute by including more samples with different values of target values would play a great role for enhancing the calibration model [66].

The performances of two calibration methods were also compared according to the evaluation standard. There was no difference in using PLSR and LS-SVM for the determination of WHC indices when the full range spectra in Spectral Set I was considered. Different from the analysis of Spectral Set I, LS-SVM had better results than PLSR when Spectral Set II was considered, where the RMSEP values of PLSR models increased by 35.0% (PLL), 17.8% (PWL), 10.5% (PFL), and 28.0% (PWR), respectively. This might because that the dominant absorption of water immersed the spectral information of other components affecting WHC in Spectral Set II, so that the spectral data and WHC values were not linearly related and LS-SVM as a non-linear method could obtained a better prediction for WHC prediction than PLSR that is a linear method. On the other hand, the relationship was linear between the spectra in Spectral Set I and WHC values. Therefore, PLSR and LS-SVM had similar prediction of WHC in salmon flesh when the data of Spectral Set I was considered.

### 3.3. Selection of important variables

Establishing a simplified spectral model involves a selection of which variables should be used. The above analysis of using the full range spectra did not take into account that some spectral wavelengths might contain useless information with regard to the WHC of the samples. In this section, the important variables/wavelengths reflecting the characteristics of spectra for predicting WHC indices were selected using the CARS method by using the simple but effective principle 'survival of the fittest' on which Darwin's Evolution Theory is based. Only the Spectral Set I (400–1000 nm) was analyzed in the selection of important variables, because it had better prediction ability than Spectral Set II. After the CARS calculation, uninformative variables were eliminated and an optimal combination of some competent wavelengths was retained. As a result of the CARS calculation, there were thirteen, twelve, nine, and twelve variables selected as the important variables for the prediction of PLL, PWL, PFL, and PWR, respectively. The specific wavelength values of these identified variables are shown in the following Eq. (2) to Eq. (5).

As a consequence of the variable selection, new reduced spectral matrix was generated by selecting the hyperspectral images only at the important variables that contained the most relevant spectral information of WHC for salmon flesh. The new matrix was then used to replace the full range spectra for building new PLSR and LS-SVM models to determine four WHC indices. Table 3 shows the prediction results of four WHC indices (PLL, PWL, PFL, and PWR) of salmon samples by analyzing only the important variables selected by CARS. The new PLSR and LS-SVM models (Method II) had better results than the original models (Method I). In brief, the new reduced models resulted in the average RPD of the PLSR and LS-SVM models increased by 15.3%, 14.4%, 3.6%, and 13.2% for four WHC indices, respectively. Such improvements were achieved by using only 10.7% (13 vs. 121), 9.9% (12 vs. 121), 7.4% (9 vs. 121), and 9.9% (12 vs. 121) of variables in the full wavelength range, showing that CARS was efficient for selecting important variables in the WHC determination of salmon flesh. This is because the use of full wavelength range could imply the noise and unnecessary information, resulting in less accurate

models. A model with selected wavelengths that carry most valuable information might be equally or more efficient than a model of full wavelength range [67].

A high RPD value means the value of standard error of prediction of the target attribute is much smaller than the corresponding standard deviation of the reference values of this attribute. Therefore, a model with a high RPD value is potentially in agreement with the reference method in predicting the target attribute. As shown in Tables 2 and 3, it can be seen that hyperspectral imaging is an efficient alternative for determining WHC of salmon flesh rapidly and noninvasively. Among the analysis of four WHC indices, the best performances of hyperspectral imaging based on the important variables were achieved for the determination of PWR, which had RPD values of both PLSR and LS-SVM models higher than four (Table 3); followed by the determination of PLL and PWL that had RPD values around three; and the RPD values of PFL determination were just over 1.6. Fig. 5 shows the scatter plots of measured versus predicted value of the CARS–PLSR models for four WHC indices, namely PLL, PWL, PFL, and PWR. It was found that, the samples are distributed closely to the regression line in the PWR analysis (Fig. 5d), which shows an excellent performance of PWR prediction in salmon flesh using hyperspectral imaging. the PLL and PWL values predicted by the corresponding CARS–PLSR models also maintained appreciably lower deviations relative to their reference values (Fig. 5a and b), while the predicted PFL values showed consistently high deviations practically across the entire range of their reference values (Fig. 5d).

In addition, the performances of two calibration algorithms, PLSR and LS-SVM, were compared based on the consideration of only the important variables. It was found that the results of PLSR models were similar to those of LS-SVM models. Only for the PLL analysis, the PLSR model had a larger RMSEP with 8.1% increment compared with the corresponding LS-SVM model (Method II). Other new PLSR models had lower RMSEP decreased by 7.9%, 2.7%, and 7.3% for PWL, PFL, and PWR, respectively. Therefore, the CARS–PLSR combination was determined as the optimal strategy for multivariate calibration (Spectral Set I) in the assessment of WHC in salmon flesh using hyperspectral imaging technique. The quantitative equations of the CARS–PLSR models for the WHC determination of salmon flesh were obtained according to the regression coefficients of the CARS–PLSR models, and are shown as follows:

$$Y_{PLL} = 7.4 - 709.4X_{430nm} + 1608.0X_{445nm} - 656.4X_{450nm} + 2975.1X_{510nm} - 11888.1X_{605nm} + 8468.9X_{620nm} - 452.0X_{760nm} + 635.5X_{765nm} - 905.2X_{830nm} + 669.6X_{955nm} - 918.1X_{965nm} + 2244.0X_{975nm} - 1238.3X_{995nm} \quad (2)$$

$$Y_{PLL} = 8.2 - 1078.5X_{440nm} + 1599.5X_{450nm} - 759.5X_{520nm} + 707.1X_{595nm} - 1037.0X_{600nm} + 515.8X_{615nm} - 971.0X_{765nm} + 1615.6X_{830nm} - 1615.2X_{885nm} + 727.4X_{925nm} + 1113.8X_{975nm} - 936.8X_{995nm} \quad (3)$$

$$Y_{PFL} = 0.9 - 64.8X_{420nm} - 31.5X_{560nm} + 35.9X_{580nm} - 134.1X_{620nm} + 249.3X_{625nm} - 1263.4X_{695nm} + 1085.3X_{700nm} + 644.2X_{835nm} - 563.3X_{840nm} \quad (4)$$

$$Y_{PWR} = -59.8 - 246.5X_{600nm} + 657.1X_{605nm} - 518.3X_{615nm} + 951.2X_{755nm} - 1547.8X_{850nm} - 884.4X_{860nm}$$

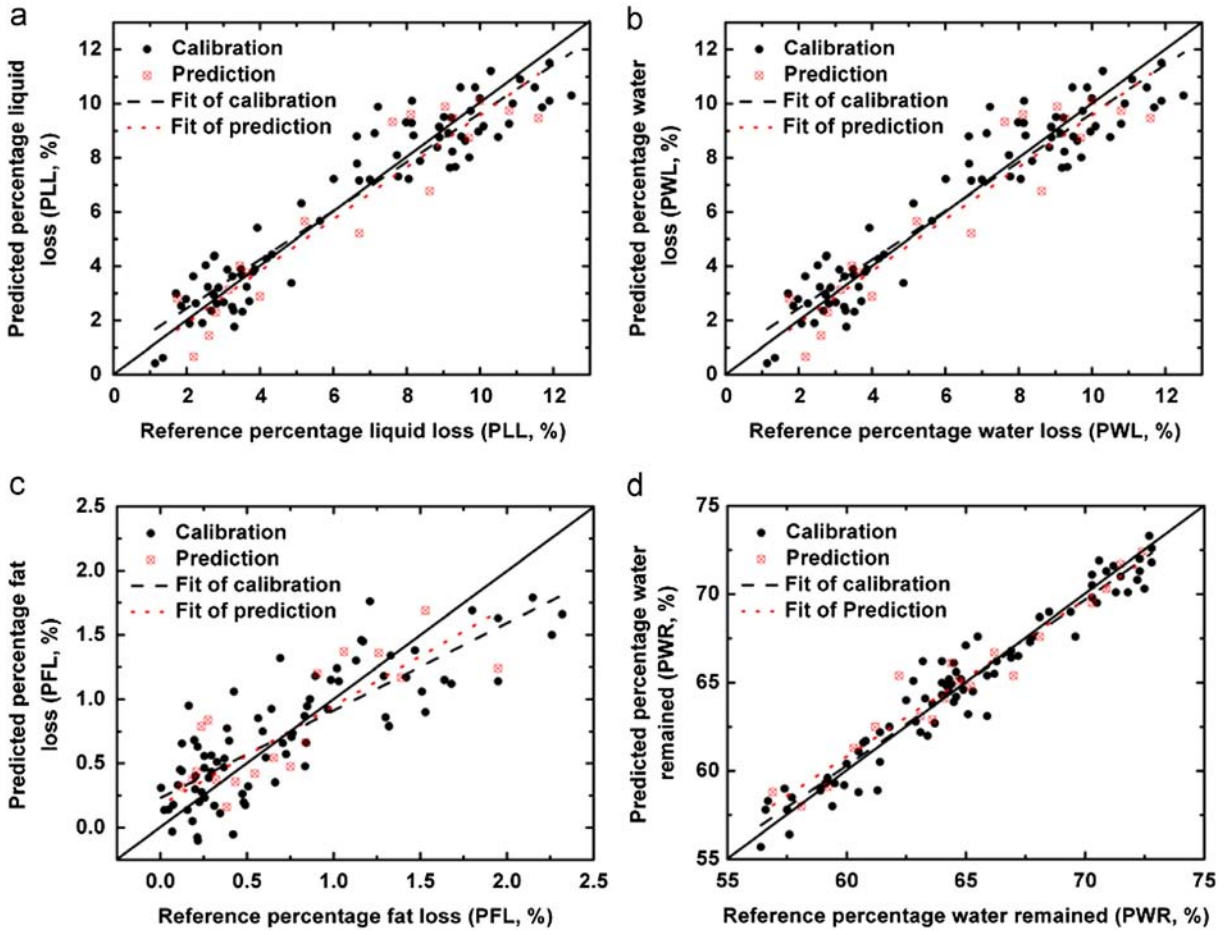


**Fig. 5.** Reference and predicted water-holding capacity (WHC) indices, namely percentage liquid loss (PLL), percentage water loss (PWL), percentage fat loss (PFL), and percentage water remained (PWR) by CARS–PLSR models.

$$-477.9X_{865nm} + 2843.3X_{890nm} + 275.1X_{935nm} - 903.6X_{940nm}$$

$$-891.5X_{960nm} + 775.1X_{995nm} \quad (5)$$

where $X_{inm}$ is the reflectance spectral value at the wavelength of $i$ nm and $Y_{PLL}$, $Y_{PWL}$, $Y_{PFL}$, and $Y_{PWR}$ are the predicted values of the WHC indices. In order to further evaluate the performances of the CARS–PLSR models, the mean WHC value of each ROI was computed over all pixels in this ROI based on the corresponding CARS–PLSR models. The results indicated that the quantitative relationships between the reference WHC values of ROIs and their mean WHC values calculated based on the above mentioned way were similar to those in calibrated models. The correlation coefficients and root mean square errors were 0.945 and 1.077% for PLL, 0.947 and 0.930 for PWL, 0.822 and 0.328% for PFL, and 0.969 and 1.138% for PWR. Therefore, the established CARS–PLSR models were very efficient in predicting WHC of salmon flesh.

### 3.4. Visual representation of WHC indices

Besides the quantitative prediction, the spatial distribution of WHC within salmon fillets was also realized by using hyperspectral imaging, which is the main advantage of hyperspectral imaging against the traditional spectroscopy. In a hyperspectral image, each pixel had its own spectrum that could be characterized as the fingerprint of the physicochemical features at this pixel of the sample. The hidden quality attributes of all pixels within the hyperspectral images of the samples could be predicted using the established calibration model for analyzing the spectra of these pixels, and further visualized with the aid of a developed image processing algorithm. In this study, the CARS–PLSR models (Spectral Set I) were used to obtain the distribution map of WHC indices. The spectrum of each pixel at the important variables (Method II) in the hyperspectral image of salmon fillet was multiplied by CARS–PLSR regression coefficients of a WHC index (shown in Eq. (2) to Eq. (5)), resulting in the value of the WHC index of this pixel. Fig. 3 shows the distribution maps of four WHC indices, namely PLL, PWL, PFL, and PWR within some salmon fillets. For each WHC index shown in Fig. 3, the right-side fillet has higher average WHC value, while the left-side fillet has lower WHC value, indicating that different fillets had various WHC values. On the other hand, there were also considerable ranges of WHC values within one fillet, which is hard to be observed using traditional spectroscopy technique. However, such differences in WHC indices from location to location within the same fillet could be easily discerned using hyperspectral imaging. This advantage is very important as all of these variations cannot be inspected by naked eyes. Indeed, WHC is an essential attribute of evaluating the ability of the muscle to hold water for describing the quality of fish. From economic point of view, low WHC means a potential weight decrease due to loss of water. Besides, low WHC will lead accumulated exudate that is unattractive to the consumers [68]. Therefore, accurate determination of WHC value and its distribution in salmon flesh is very important for correct pricing, authentication and categorization of salmon products, providing economic benefits for producers by increasing consumer confidence in the supplied salmon products. Competitiveness of the salmon industry would also be improved as products could be marketed as guaranteed quality. As the determination of WHC of salmon flesh by the traditional centrifugation method that was used to measure the reference WHC values in this paper appears to be a destructive and laborious method, the technique of hyperspectral imaging provides a rapid and non-destructive WHC assessment of salmon flesh in order to establish prime quality grades.

## 4. Conclusions

The WHC distribution within salmon fillets was visualized at a pixel-wise level using hyperspectral imaging technology in a rapid and non-invasive manner. The successful determination of four WHC indices, namely PLL, PWL, PFL, and PWR lies in the spectral data provided by hyperspectral images, which contain the molecule information of chemical compositions in salmon flesh. On the other hand, the spatial information contained by each pixel within the hyperspectral image provides the feasibility of generating the distribution maps of WHC, which is important for understanding the WHC variation of salmon flesh. The results and illustrations presented in this work suggest that both spectral and spatial information are essential for the quantitative generation of WHC distribution within salmon fillets. That is the reason why hyperspectral imaging is required for measuring distribution of WHC indices in salmon flesh, rather than traditional spectroscopy or computer vision techniques. The multivariate analysis indicated that the visible and short-wave near infrared range of 400–1000 nm (Spectral Set I) was more suitable than the long-wave near infrared range of 897–1753 nm (Spectral Set II) for the WHC prediction. CARS algorithm was conducted to select the important variables for reducing the excessive degree of dimensionality with redundancy in hyperspectral images. The combination of CARS and PLSR was considered as the optimal method for the spectral calibration, resulting in good RPD values of 2.888, 2.867, 1.682, and 4.034 for PLL, PWL, PFL, and PWR, respectively. On the basis of important variables, an optimized multispectral imaging system could be developed for industrial on-line purpose. As a comparison to a hyperspectral image system that has voluminous data and high cost, such optimized multispectral imaging systems have the potential to meet the speed requirement of production lines and would not be significantly different from RGB cameras in speed and cost. By implementing hyperspectral imaging as a key component of an automated optic-based system, it could enable the salmon industry to sort, label, and price salmon products truthfully according to their attributes, in opposite to conventional methods. It is believed that the results of this work will promote more endeavours to investigate the potential of hyperspectral imaging for assessing additional quality and safety attributes of salmon and salmon products.

## Acknowledgements

The authors would like to acknowledge the financial support provided by the Irish Research Council for Science, Engineering and Technology under the Government of Ireland Postdoctoral Fellowship scheme. We also thank Professor Colm O'Donnell for lending us the hyperspectral imaging equipment and Dr. Carlos Esquerre Feranandez for excellent technical assistance.

## References

[1] G.B. Olsson, R. Ofstad, J.B. Lødemel, R.L. Olsen, LWT Food Sci. Technol. 36 (2003) 771–778.
[2] G.B. Olsson, M.A. Seppola, R.L. Olsen, LWT Food Sci. Technol. 40 (2007) 793–799.
[3] M.N. Lund, M. Heinonen, C.P. Baron, M. Estevez, Mol. Nutr. Food Res. 55 (2011) 83–95.
[4] V.M. Ocano-Higuera, E. Marquez-Rios, M. Canizales-Davila, F.J. Castillo-Yanez, R. Pacheco-Aguilar, M.E. Lugo-Sanchez, K.D. Garcia-Orozco, A.Z. Graciano-Verdugo, Food Chem. 116 (2009) 933–938.
[5] K.L. Pearce, K. Rosenvold, H.J. Andersen, D.L. Hopkins, Meat Sci. 89 (2011) 111–124.
[6] K.O. Honikel, Meat Sci. 49 (1998) 447–457.
[7] J. Brøndum, L. Munck, P. Henckel, A. Karlsson, E. Tornberg, S.B. Engelsen, Meat Sci 55 (2000) 177–185.

[8] M.D. Aaslyng, C. Bejerholm, P. Ertbjerg, H.C. Bertram, H.J. Andersen, Food Qual. Preference 14 (2003) 277–288.
[9] I.E. Bechmann, B.M. Jørgensen, LWT Food Sci. Technol. 31 (1998) 648–652.
[10] M.R. Brown, P.D. Kube, R.S. Taylor, N.G. Elliott, Aquac. Res. (2012), Doi: 10.1111/are.12021, in press.
[11] Y. Huang, A.G. Cavinato, D.M. Mayes, L.J. Kangas, G.E. Bledsoe, B.A. Rasco, J. Food Sci. 68 (2003) 482–486.
[12] P. Jackman, D.-W. Sun, P. Allen, Trends Food Sci. Technol. 22 (2011) 185–197.
[13] D. Wu, D.-W. Sun, Trends Food Sci. Technol. 29 (2012) 5–20.
[14] D. Wu, D.-W. Sun, Innovative Food Sci. Emerg. Technol. (2013), Doi: 10.1016/j.ifset.2013.04.014, in press.
[15] D. Wu, D.-W. Sun, Innovative Food Sci. Emerg. Technol. (2013), Doi: 10.1016/j.ifset.2013.04.016, in press.
[16] D.-W. Sun, Hyperspectral Imaging for Food Quality Analysis and Control, Academic Press/Elsevier, San Diego, California, USA, 2010.
[17] D. Lorente, N. Aleixos, J. Gomez-Sanchis, S. Cubero, O.L. Garcia-Navarrete, J. Blasco, Food Bioprocess Technol. 5 (2012) 1121–1142.
[18] L.S. Magwaza, U.L. Opara, H. Nieuwoudt, P.J. Cronje, W. Saeys, B. Nicolaï, Food Bioprocess Technol. 5 (2012) 425–444.
[19] D. Wu, S. Wang, N. Wang, P. Nie, Y. He, D.-W. Sun, J. Yao, Food Bioprocess Technol. (2013), Doi: 10.1007/s11947-012-0928-0, in press.
[20] D.F. Barbin, G. ElMasry, D.-W. Sun, P. Allen, Anal. Chim. Acta 719 (2012) 30–42.
[21] M. Kamruzzaman, G. ElMasry, D.-W. Sun, P. Allen, Anal. Chim. Acta 714 (2012) 57–67.
[22] D. Wu, H. Shi, S. Wang, Y. He, Y. Bao, K. Liu, Anal. Chim. Acta 726 (2012) 57–66.
[23] D. Lorente, N. Aleixos, J. Gómez-Sanchis, S. Cubero, J. Blasco, Food Bioprocess Technol. 6 (2013) 530–541.
[24] D.P. Ariana, R.F. Lu, Comput. Electron. Agric. 74 (2010) 137–144.
[25] V.H. Segtnan, M. Hoy, F. Lundby, B. Narurn, J.P. Wold, J. Near Infrared Spectrosc. 17 (2009) 247–253.
[26] D. Wu, D.-W. Sun, Y. He, Innovative Food Sci. Emerg. Technol. 16 (2012) 361–372.
[27] V.H. Segtnan, M. Hoy, O. Sorheim, A. Kohler, F. Lundby, J.P. Wold, R. Ofstad, J. Agric. Food Chem. 57 (2009) 1705–1710.
[28] S. Ottestad, M. Hoy, A. Stevik, J.P. Wold, J. Near Infrared Spectrosc. 17 (2009) 77–87.
[29] H.-J. He, D. Wu, D.-W. Sun, Innovative Food Sci. Emerg. Technol. 18 (2013) 237–245.
[30] D. Wu, D.-W. Sun, Talanta 111 (2013) 39–46.
[31] A. Rørå, O. Einen, J. Food Sci. 68 (2003) 2123–2128.
[32] P. Berzaghi, R. Riovanto, Ital. J. Anim. Sci. 8 (2009) 39–62.
[33] R.W. Gerlach, B.R. Kowalski, H.O.A. Wold, Anal. Chim. Acta-Comput. Tech. Optim. 3 (1979) 417–421.
[34] X.J. Chen, X.X. Lei, J. Agric. Food Chem. 57 (2009) 334–340.
[35] X. Chen, D. Wu, Y. He, S. Liu, Anal. Chim. Acta 638 (2009) 16–22.
[36] X. Zhu, G. Huang, S. Luo, X. Guan, X. Chen, Anal. Lett. 46 (2013) 671–681.
[37] V. Sinija, H. Mishra, Food Bioprocess Technol. 4 (2011) 136–141.
[38] F. Antonucci, F. Pallottino, G. Paglia, A. Palma, S. D'Aquino, P. Menesatti, Food Bioprocess Technol. 4 (2011) 809–813.
[39] D. Wu, J. Chen, B. Lu, L. Xiong, Y. He, Y. Zhang, Food Chem. 135 (2012) 2147–2156.
[40] G. ElMasry, N. Wang, A. ElSayed, M. Ngadi, J. Food Eng. 81 (2007) 98–107.
[41] C. Cortes, V. Vapnik, MLear 20 (1995) 273–297.
[42] C.J.C. Burges, Data Min. Knowl. Discovery 2 (1998) 121–167.
[43] J.A.K. Suykens, J. Vandewalle, Neural Process. Lett. 9 (1999) 293–300.
[44] J.A.K. Suykens, T. Van Gestel, J. De Brabanter, B. De Moor, J. Vandewalle, Least Squares Support Vector Machines, World Scientific, Singapore, 2002.
[45] S. Ji-yong, Z. Xiao-bo, H. Xiao-wei, Z. Jie-wen, L. Yanxiao, H. Limin, Z. Jianchun, Food Chem. 138 (2013) 192–199.
[46] X.J. Chen, D. Wu, Y. He, S. Liu, Food Bioprocess Technol. 4 (2011) 753–761.
[47] Y.D. Liu, R.J. Gao, Y. Hao, X.D. Sun, A.G. Ouyang, Food Bioprocess Technol. 5 (2012) 1106–1112.
[48] E. Kakaei Lafdani, A. Moghaddam Nia, A. Ahmadi, J. Hydrol. 478 (2013) 50–62.
[49] D. Wu, H.Q. Yang, X.J. Chen, Y. He, X.L. Li, J. Food Eng. 88 (2008) 474–483.
[50] P.C. Nie, D. Wu, Y. Yang, Y. He, J. Food Eng. 109 (2012) 155–161.
[51] S.J. Wang, K.S. Liu, X.J. Yu, D. Wu, Y. He, J. Food Eng. 109 (2012) 531–537.
[52] D. Wu, X. Chen, X. Zhu, X. Guan, G. Wu, Anal. Methods 3 (2011) 1790–1796.
[53] U. Indahl, T. Naes, J. Chemom. 18 (2004) 53–61.
[54] D. Wu, P.C. Nie, Y. He, Y.D. Bao, Food Bioprocess Technol. 5 (2012) 1402–1410.
[55] D. Wu, X.J. Chen, P.Y. Shi, S.H. Wang, F.Q. Feng, Y. He, Anal. Chim. Acta 634 (2009) 166–171.
[56] G. ElMasry, A. Iqbal, D.-W. Sun, P. Allen, P. Ward, J. Food Eng. 103 (2011) 333–344.
[57] X.J. Chen, D. Wu, Y. He, Spectroscopy-Us 26 (2011) 42–47.
[58] G. Elmasry, D.F. Barbin, D.-W. Sun, P. Allen, Crit. Rev. Food Sci. Nutr. 52 (2012) 689–711.
[59] H. Li, Y. Liang, Q. Xu, D. Cao, Anal. Chim. Acta 648 (2009) 77–84.
[60] X. Wei, N. Xu, D. Wu, Y. He, Food Bioprocess Technol. (2013), Doi: 10.1007/s11947-013-1053-4, in press.
[61] W. Fan, Y. Shan, G.Y. Li, H.Y. Lv, H.D. Li, Y.Z. Liang, Food Anal. Methods 5 (2012) 585–590.
[62] T. Sun, W.L. Xu, J.L. Lin, M.H. Liu, X.W. He, Spectrosc. Spect. Anal. 32 (2012) 3229–3233.
[63] C. Darwin, G. Beer, The Origin of Species, Oxford University Press, 1998.
[64] D. Samuel, B. Park, M. Sohn, L. Wicker, Poult. Sci. 90 (2011) 914–921.
[65] J. Reeves III, J. Near Infrared Spectrosc. 2 (1994) 199–212.
[66] G. ElMasry, D.-W. Sun, P. Allen, J. Food Eng. 110 (2012) 127–140.
[67] G. ElMasry, D.-W. Sun, P. Allen, J. Food Eng. 117 (2013) 235–246.
[68] G.B. Olsson, R.L. Olsen, R. Ofstad, LWT Food Sci. Technol. 36 (2003) 125–133.
[69] M. Taghizadeh, A.A. Gowen, C.P.O Donnell, Sens. Instrum. Food Qual. 3 (2009) 219–226.
[70] G. Elmasry, D.-W. Sun, Principles of hyperspectral imaging technology, in: D.-W. Sun (Ed.), Hyperspectral Imaging for Food Quality Analysis and Control, Academic Press/Elsevier, San Diego, California, USA, 2010, pp. 3–44.
[71] D.-W. Sun, T. Brosnan, Pizza quality evaluation using computer vision - part 1 - Pizza base and sauce spread. JOURNAL OF FOOD ENGINEERING 57 1 81-89 Article Number: PII S0260-8774(02)00275-3 10.1016/S0260-8774(02)00275-3 Published: 2003.
[72] Du, CJ; Sun, D-W. Comparison of three methods for classification of pizza topping using different colour space transformations. JOURNAL OF FOOD ENGINEERING 68 3 277-287 10.1016/j.jfoodeng.2004.05.044 Published: JUN 2005.
[73] Jackman, Patrick; Sun, Da-Wen; Du, Cheng-Jin; Allen, Paul. Prediction of beef eating quality from colour, marbling and wavelet texture features. MEAT SCIENCE 80 4 1273-1281 10.1016/j.meatsci.2008.06.001 Published: DEC 2008.
[74] L. Liu, M. Ngadi, Food Bioprocess Technol., (2013) Doi: 10.1007/s11947-012-0933-3, in press.