A comparative study for improving prediction of total viable count in beef based on hyperspectral scattering characteristics

Feifei Tao, China Agricultural University
Carmen L. Gomes, Texas A&M University
Kuanglin Chao, United States Department of Agriculture
Jianwei Qin, United States Department of Agriculture

Available at: https://works.bepress.com/carmen-gomes/
A comparative study for improving prediction of total viable count in beef based on hyperspectral scattering characteristics

Feifei Tao a, Yankun Peng a,⁎, Carmen L. Gomes b, Kuanglin Chao c, Jianwei Qin c

a College of Engineering, China Agricultural University, 17 Qinghua East Road, Haidian, Beijing 100083, China
b Department of Biological & Agricultural Engineering, Texas A&M University, College Station, TX 77843-2117, USA
c Environmental Microbial and Food Safety Laboratory, Agricultural Research Service, United States Department of Agriculture, 10300 Baltimore Avenue, Beltsville, MD 20705, USA

ARTICLE INFO

Article history:
Received 14 December 2014
Received in revised form 5 April 2015
Accepted 7 April 2015
Available online 15 April 2015

Keywords:
Beef
Total viable count
Hyperspectral scattering imaging
Nondestructive detection
Lorentzian function
Parameter combination

ABSTRACT

The objectives of this study were to investigate the feasibility of hyperspectral scattering imaging to predict the bacterial contamination in meat nondestructively, and propose an optimal approach for detecting low levels of total viable count (TVC) contamination in beef. Fresh beef samples were obtained from a commercial slaughtering plant, and stored at 4°C for 0–12 days. The visible/near-infrared (VIS/NIR) hyperspectral images in the backscattering mode were acquired from 3–5 beef samples on each day of the experiment, in parallel with microbiological analysis to enumerate the TVC population. Lorentzian function was used to resolve the light scattering information within the hyperspectral image and consequently Lorentzian parameters, which represented different hyperspectral scattering characteristics were extracted. In this study, not only the individual Lorentzian parameters but also the parameter combinations were used to establish the multivariate statistical models for predicting beef TVC, based on the modeling methods of principal component regression (PCR), partial least squares regression (PLSR), and back propagation neural network (BPNN), respectively. The models established using individual Lorentzian parameters did not perform well in predicting low levels of TVC contamination in beef, and the best prediction result could only achieved with the correlation coefficient of prediction set (Rc) and root mean square error of prediction set (RMSEP) of 0.81 and 1.27 log CFU/g, respectively. Based on the parameter combinations, the best modeling results were achieved with Rc and RMSEP of 0.86 and 0.93 log CFU/g, 0.87 and 0.79 log CFU/g, 0.90 and 0.88 log CFU/g by PCR, PLSR, and BPNN methods, respectively, which confirmed the superiority of the parameter combination method. The results of this study demonstrated for the first time that hyperspectral scattering imaging combined with Lorentzian function and the proposed parameter combination method could be used to detect low levels of bacterial contamination in beef nondestructively.

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

Food quality and safety are directly related to the consumers' health and social progress, and thus are important issues throughout the world (Cen and He, 2007). There is an expectation throughout the society that the food supplied for human consumption is safe and nutritious to eat (Frewer et al., 2005; de Jonge et al., 2004). Beef is a commercially important and widely consumed muscle food, due to that it is a good source for proteins and other essential nutrients. However, it is also an ideal substrate for the growth of both spoilage and pathogenic microorganisms (Schirmer and Langsrud, 2010). As reported, microbial hazard is still considered to be one of the major challenges to meat safety in the 21st century (Sofos, 2008). The microbiological quality of meat mainly depends on the physiological status of the animal at slaughter, the spread of contamination during slaughter and processing, the temperature and other conditions of storage and distribution (Nychas et al., 2008). In general, the metabolic activity of microorganisms which prevail in a meat ecosystem leads to the manifestation of changes or spoilage of meat. Nychas et al. (2007) clarified that it was the microbial growth per se, rather than the activity of microbial enzymes and as a consequence, the accumulation of metabolic by-products of microorganisms that characterized food spoilage. Currently, the practice to assure microbial safety of meat still relies on regulatory inspection and sampling regimes. This approach, however, seems inadequate because it cannot sufficiently guarantee consumer protection, since 100% inspection and sampling is...
technically, financially and logistically impossible (Kodogiannis et al., 2014). In addition, most of the related microbiological detection/measurement methods are destructive, time-consuming, and giving retrospective information (Ellis and Goodacre, 2001), thus are not suitable to fast-paced processing environment in meat plants. In a modern food-processing environment, monitoring procedures need to give results in real-time so that corrective actions can be taken as soon as possible when implementing Hazard Analysis and Critical Control Point (HACCP) plan (Ellis et al., 2002).

A promising way to overcome these limitations is to apply non-destructive optical technology (Sowoidnich et al., 2012). Among current emerging technologies, the optic-based methods were reported to have the greatest potential for online application (Shackelford et al., 1999; Vote et al., 2003). The “point” detection spectroscopic methods, such as Fourier transform infrared spectroscopy (FTIR), Near-infrared (NIR) spectroscopy and Raman spectroscopy, which can only detect small portions of the samples, commonly require the tested meat samples to be minced instead of meat chops or need many detected positions to ensure the representativeness of the obtained spectra, and such nature has significantly limited their widespread applications. Hyperspectral imaging is a new rapidly growing technique that integrates spectroscopic and imaging techniques together to provide both spectral and spatial information simultaneously, thus it could provide more adequate and comprehensive information of the object. Recently, studies have been reported on assessing the microbial contamination of meat, in terms of TVC, psychrotrophic plate count (PPC) and Enterobacteriaceae by using a NIR hyperspectral system (900–1700 nm) in the reflectance mode or a visible/NIR (VIS/NIR) hyperspectral scattering system (400–1100 nm) in the backscattering mode (Barbin et al., 2013; Feng and Sun, 2013; Peng et al., 2013; Peng et al., 2011; Tao et al., 2012; Tao and Peng, 2014, 2015). Different from the common hyperspectral imaging system, hyperspectral scattering system is based on the light-backscattering imaging (LBI) in which point light is applied as the illuminant source, and thus could capture the backscattering image of the object. During the process of meat spoilage, not only the bacterial load and chemical ingredients undergo changes, but also the microstructure of meat can be different, and these changes could be reflected by the derived light absorption and scattering features from hyperspectral scattering image. Light scattering is commonly reported to be due to physical characteristics (e.g., particle size, cellular structure, and density) of the tissue, whereas light absorption is related to the chemical constituents. And these two optical properties can be characterized by the reduced scattering coefficient, $\mu_s'$, and the absorption coefficient, $\mu_a$ (Tuchin, 2007).

Despite of our previous encouraging results (Peng et al., 2011; Tao et al., 2012; Tao and Peng, 2014, 2015; Wu et al., 2012), further studies are still necessary to fully explore the potentials and limitations of hyperspectral scattering imaging. Therefore, the objectives of the present study were to: (1) exploit the capabilities of the improved hyperspectral scattering imaging system, which was enhanced by introducing a laser displacement detector for accurate measurement of the object distance in this study for detecting low levels of TVC contamination in beef stored at 4 °C, to best simulate the real cold chain condition; (2) propose a data processing method for comprehensively utilizing the information within the three individual Lorentzian parameters, and explore its superiority on predicting low levels of TVC contamination in beef to individual parameters; (3) apply different chemometric methods for quantitative prediction of beef TVC and compare their modeling performances. Methods used in this study consisted of the linear regression methods of principal component regression (PCR) and partial least squares regression (PLSR), and the non-linear regression method of back propagation neural network (BPNN).

### 2. Materials and methods

#### 2.1. Preparation of beef samples

Two batches of beef (*Longissimus dorsi* muscle) were purchased from a local commercial slaughtering plant (Yuxiangyuan Husbandary Co., Ltd., Beijing, China) on two different days and immediately transported to the lab under refrigeration. All beef samples were from *Simmental* cattle breed. Totally, 47 beef samples were prepared aseptically by trimming into an uniform size of 9 cm × 5 cm × 2.5 cm (length × width × thickness), and packed separately in commercial food grade polyethylene bags. The samples were placed orderly in the refrigerator and stored at 4 °C for 0–12 days. On each day of the experiment, 3–5 samples were withdrawn randomly for hyperspectral imaging and reference microbiological analysis. Day 0 samples were used immediately before storage.

#### 2.2. Hyperspectral scattering imaging system

A laboratory VIS/NIR hyperspectral scattering imaging system in the spectral range of 400–1100 nm was used in this study. Compared to the previous studies on evaluating beef quality and safety attributes in our lab (Peng et al., 2011; Wu et al., 2012), the system used in this study was improved by introducing a laser displacement detector for accurate measurement of the object distance. Briefly, the enhanced hyperspectral imaging system consisted of a high-performance back-illuminated 12-bit charge coupled device (CCD) camera (Sensicam QE, PCO AG, Kelheim, Germany), an imaging spectrograph (ImSpector V10E, Spectral Imaging Ltd., Oulu, Finland), an illumination unit (Oriel Instruments, Stratford, USA) equipped with optical fiber, a laser displacement sensor (GV-H45, Keyence Corp., Shanghai, China) and a computer supported with a data acquisition and control software. The optical fiber was used to form point light in the scattering imaging system, and the diameter of the light beam formed was 5 mm. In addition, in order to minimize the effect of ambient light, the imaging system was enclosed in a shield box.

The system worked in a line scanning mode, and all scans were obtained at a position of 3 mm (from the incident light center to the scanned position) off the incident light center in order to avoid the signal saturation on CCD detector. The resolution of the imaging system was spectrally 2.8 nm with a 0.74 nm interval, and spatially less than 9 μm. The original image generated by this system was of 1376 × 1040 (spatial × spectral) pixels, with the CCD camera binning variable in the horizontal and vertical directions of 1, 2, 4, 8 and 1, 2, 4, 8, 16, respectively.

#### 2.3. Image acquisitions

In order to eliminate the dark current effect of the imaging system, the dark image was first obtained by covering the camera lens before imaging for each beef sample. Sample surface with no visible fat or connective tissue was selected for imaging. Before each imaging, the object distance was first measured by the laser displacement sensor, and then kept to the set distance by adjusting the vertical translation stage. For each sample, 5 positions were selected for imaging, and for each imaging, 4 images were averaged to 1 automatically by setting the camera working parameter, which meant that actually 20 hyperspectral images were acquired for each sample.

In order to improve the signal-to-noise ratio (SNR), 2 × 2 binning was performed on the original images in this study, thus the resulting images were of 688 × 520 (spatial × spectral) pixels.
The images after pixel binning were saved in TIFF format for further analysis.

### 2.4. Microbiological analyses

After acquisitions of hyperspectral images, the microbiological tests were performed to determine the reference beef TVCs according to GB/T 4789.2-2010 (China National Standard, 2010). Briefly, 25 g samples were transferred aseptically into 225 mL sterilized 0.85% saline solution, and homogenized (10,000 rpm) in a homogenizer (JT-C, Luohu Jintian Test Equipment Institute, Luohu, China) for 60 s. Then, for microbial enumeration, 1 mL samples of 10-fold serial dilutions of beef homogenates were mixed with plate count agar (Beijing Aoboxing Bio-tech Co., Ltd., Beijing, China) in Petri dishes and incubated at 37 °C for 48 ± 2 h. Two replicates of at least two appropriate dilutions depending on the sampling day were enumerated. The number of beef TVC for each sample was then calculated according to GB/T 4789.2-2010. Finally, the TVC data was log-transformed, and expressed in log CFU/g for further analysis.

### 2.5. Data analysis method

With hyperspectral scattering imaging, a small continuous-wave light beam is incident on the sample, generating diffusely reflected images at its surface around the incident point as a result of light propagation and backscattering inside the sample (Mendoza et al., 2011). When applying this technique, it is crucial to exploit appropriate mathematical equations to describe the spectral scattering characteristics as well as to comprehensively utilize the information within the profiles. Our previous study showed that both Lorentzian function and Gompertz function could model the scattering profiles of meat effectively, and no obvious differences were observed from the fitting results by them (Tao and Peng, 2015); therefore, the emphasis of the present study was to exploit an effective method to comprehensively utilize the information within the scattering profiles.

Fig. 1 presents the methodology employed in this work, and the detailed data analysis method for each step is described in the following sections, respectively. All the data analysis was performed in the Matlab 7.0 software (Mathworks, Natick, USA).

#### 2.5.1. Extraction of hyperspectral scattering characteristics

A detailed description of the methodology employed in this section was presented elsewhere (Tao and Peng, 2015; Wu et al., 2012). In this study, the 3-parameter Lorentzian function, which was shown in Eq. (1) was used to fit the scattering profiles of beef.

\[
R_{wi} = a_{wi} + \frac{b_{wi}}{1 + (z/c_{wi})^2}
\]  

(1)

where \( R \) is the light intensity, in CCD Count; \( z \) is the distance from the detected position to the light incident center, in pixel; \( a \) is the asymptotic value of light intensity; \( b \) is the peak value of the scattering profile at the light incident center; \( c \) is the full scattering width of the scattering profile at one half of the peak value (FWHM), in pixel; \( w_i \) represents the designated wavelength in the spectral range of 400–1100 nm in accordance with \( i = 1, 2, 3, \ldots N \); and \( N \) is the number of wavelengths. By the procedure of curve fitting, the Lorentzian parameters \( a, b, \) and \( c \), which represent different spectral scattering characteristics of beef, can be obtained.

#### 2.5.2. Development of prediction models

The huge amount of information provided by hyperspectral data requires advanced data analysis approaches. This could be achieved through the integration of modern analytical platforms with computational and chemometric techniques (Miller and Miller, 2005). In this study, the multivariate statistical analysis methods including the linear regression methods of PCR, PLSR and the non-linear method of BPNN were used to develop the prediction models, as considering that the growth of microorganisms in meat is a complex process.

In this study, except for developing the prediction models using individual Lorentzian parameters, the parameter combinations were also exploited to comprehensively utilize the information within the three individual parameters. In essence, the parameter combination was the union set of two or more individual parameters, and referred to the union set of two and three individual Lorentzian parameters in this study. For instance, the combination of \( (a, b) \) represented the union set of parameters \( a \) and \( b \), and so forth. As three individual parameters were extracted from the above-mentioned curve fitting, there would be \( C_3^2 \) forms of two-parameter combination and \( C_3^3 \) form of three-parameter combination. The two-parameter combinations consisted of the combination forms of \( (a, b) \), \( (a, c) \) and \( (b, c) \), and the three-parameter combination referred to the combination form of \( (a, b, c) \).

In addition, in order to overcome the magnitude differences among different Lorentzian parameters when performing the parameter combination method, the standardization approach was applied to the parameter combinations according to Eq. (2).

\[
Y_{ij} = \frac{X_{ij} - \overline{X}_j}{S_j}
\]

(2)

where \( X_{ij} \) is the data matrix of each parameter combination at the \( i \)-th variable of the \( j \)-th sample; \( \overline{X}_j \) is the mean value of all different variables for the \( j \)-th sample; \( S_j \) is the standard deviation of all different variables for the \( j \)-th sample; and \( Y_{ij} \) is the data matrix after standardization.

![Fig. 1. A schematic diagram of the data-processing routines.](image-url)
2.5.2.1. Principal component regression. Principal component analysis (PCA) is a well-known data compression method based on the correlation among variables (Massart et al., 1998). It relies on the linear transformation of the original set of measurements into a substantially smaller set of uncorrelated variables called the principal components (PCs), while retaining as much of information present in the original data set as possible (Breereton, 2003). In PCA, the independent data matrix (X, function parameters in this case) is decomposed into two matrices, T and P, such that X = TP. Where T is the “score” matrix, which represents the positions of the sample in the new coordinate system (PCs coordinate system); P is the “loading” matrix, whose columns describe how the new axes, i.e. the PCs, are built from the old axes. Hence, each spectrum will have its own unique set of scores, and a spectrum can be represented by its PCA scores in the factor space instead of intensities in the wavelength space (Park et al., 2003).

2.5.2.2. Partial least squares regression. Partial least squares (PLS) is a quantitative spectral decomposition technique that is closely related to PCA. However, in PLS the decomposition is performed in a slightly different fashion. Instead of first decomposing the spectral matrix into a set of scores and loadings, and regress for them against the dependent variable (Y, reference beef TVC in this case) as a separate step, PLS actually uses the information in Y during the decomposition process. The main idea of PLS is to get as much information in Y as possible into the first few loading vectors. The first latent variable (LV) conveys the largest amount of information, followed by the second LV, and so forth. The optimum number of LVs was determined by leave-one-out cross validation (Panagou et al., 2011), and therefore could be referenced elsewhere (Panagou et al., 2011), and therefore not shown here.

2.5.2.3. Back propagation neural network. The construction of BPNN was according to Hassoun (1995). Briefly, the system consisted of input, output and hidden layers. The output signal is obtained using the algebraic sum of the weighted inputs:

\[ Y_k = f \left( \sum_{i} W_{ij} X_i - \theta_j \right) \]  

(3)

where \( Y_k \) is the output signal at node \( k \); \( f(\cdot) \) denotes the transfer function; \( W_{ij} \) is the weight between the node \( i \) and node \( j \); \( X_i \) is the input signal at node \( i \); \( \theta_j \) is the bias at node \( j \). It works on the principle that after the information has gone through the network in a forward direction and an output has been produced, the error associated with this output is redistributed backwards through the model and weights are adjusted accordingly. Minimization of the error occurs through several iterations (training cycles) (Ham and Kostanic, 2001).

In this study, the PCs resulting from the above-mentioned PCA analysis were subjected to the BPNN model as the input layer, and the output layer contained one node for the prediction of beef TVC (in log CFU/g). The number of nodes in the hidden layer was optimized based on the empirical Eq. (4).

\[ m = \sqrt{n + 1 + a} \]  

(4)

where \( m \) is the number of nodes in the hidden layer, \( n \) is the number of nodes in the input layer, \( l \) is the number of nodes in the output layer, and \( a \) is a constant from 1 to 10 (Xu et al., 2013). For the development of BPNN models, the input X was normalized into \([-1, 1]\) or \([0, 1]\), depending on the original X being positive or negative, and the reference data Y was normalized into \([0, 1]\) uniformly. The transfer function was “tansig” for the hidden layer nodes and “purelin” for the output layer nodes in constructing the BPNN models in this study.

2.5.3. Evaluation of model performance
Statistical indices, including correlation coefficient of calibration set (\( R_c \)), root mean squared error of calibration set (RMSEC), correlation coefficient of prediction set (\( R_p \)), and root mean squared error of prediction set (RMSEP) were calculated to evaluate the performance of the established models. These parameters are defined as follows:

\[ R_c = \sqrt{\frac{\sum_{i=1}^{n_c} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{n_c} (y_i - \bar{y})^2}} \]  

(5)

\[ RMSEC = \sqrt{\frac{1}{n_c} \sum_{i=1}^{n_c} (y_i - \hat{y}_i)^2} \]  

(6)

\[ R_p = \sqrt{\frac{\sum_{i=1}^{n_p} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{n_p} (y_i - \bar{y})^2}} \]  

(7)

\[ RMSEP = \sqrt{\frac{1}{n_p} \sum_{i=1}^{n_p} (y_i - \hat{y}_i)^2} \]  

(8)

where \( y_i \) is the predicted value of the \( i \)-th observation; \( y_i \) is the measured value of the \( i \)-th observation; \( \bar{y} \) is the mean value of the calibration or prediction set; \( n_c \) is the number of observations in the calibration set; and \( n_p \) is the number of observations in the prediction set. Generally, a model with higher \( R_p \) and smaller RMSEP is more satisfactory.

3. Results and discussion

3.1. Results of microbiological analyses

Fig. 2 depicts the reference TVCs of beef samples tested on each day of the experiment, while it needs to be pointed out that the average TVC on each day were not plotted to describe the growth kinetics of total bacteria here, as two batches of meat samples were included in this study and consequently their initial bacteria counts might be different which would lead to different TVCs on the following tested days.

The total bacteria count on day 0 represented the level of initial contamination on beef surfaces, with the mean data of 2.81 ± 0.34 log CFU/g which indicated an acceptable hygienic level of fresh meat samples (Fig. 2). During the first 4 storage days, the reference beef TVC showed an increase from 2.81 log CFU/g to 5.70 log CFU/g, especially between the storage time of day 2 and day 4, which depicted a pronounced “log phase” together with a short time of...
“lag phase” of bacteria growth (Fig. 2). After day 6, the population of total bacteria started to exceed 7.0 log CFU/g for some samples which suggested that these samples had undergone microbiological spoilage; however, the average TVC of the tested samples was less than 7.0 log CFU/g (Fig. 2). The indicator level of 7.0 log CFU/g was used for bacterial spoilage of beef in this study, as the levels at which bacterial spoilage takes place in meats were related to the muscle type and pH (Schreurs, 2000).

The statistical data of reference TVCs for all 47 beef samples were calculated and shown in Table 1. From Table 1, it can be seen that the minimum TVC was 2.30 log CFU/g and the maximum TVC was 8.20 log CFU/g, with the mean value and standard deviation (SD) of 5.81 and 1.72 log CFU/g, respectively, which shows that a wide range of meat TVCs was covered in this study, and it could represent the complete process of beef microbial spoilage. In addition, as the samples were divided into calibration and validation sets for model development and evaluation in the following sections (see Sections 3.3 and 3.4), the statistical data were also calculated for each set, and the results are shown in Table 1.

3.2. Resolving of hyperspectral image

3.2.1. Curve fitting to the scattering profiles

Fig. 3 shows the hyperspectral image of beef in 3-dimension (3-D) format, and the axis represents the information of light intensity, wavelength, and scattering distance, respectively. The spectral range was between 400 and 1100 nm, and the spatial distance was from −30 to 30 mm with the incident light center at 0 mm (Fig. 3).

In order to apply the chemometric techniques for data analysis, the 3-D hyperspectral image needs to be converted to a 2-D matrix, and therefore, the method of spatially-resolving hyperspectral image (Tao and Peng, 2015) was employed in this study to extract the light scattering characteristics of beef. Fig. 4 depicts the extracted scattering profiles of beef on day 0 at the wavelengths of 500, 600 and 800 nm, and it could be seen from this figure that the scattering profiles of beef at different wavelengths are conspicuously similar in shape just with different light intensities, which is in agreement with those reported previously (Peng et al., 2011; Tao et al., 2012; Tao and Peng, 2014, 2015; Wu et al., 2012). Therefore, it is reasonable to analyze such light scattering profiles using appropriate mathematical equations.

![Fig. 3. Hyperspectral image of beef in 3-D format.](image)

From Fig. 4, it could also be observed that the scattering profiles of beef are symmetrical to the incident light center; therefore, only halves of the scattering profiles were used for curve fittings in this study. The fitting results by Lorentzian function between 470 and 1000 nm are shown in Fig. 5, and it was observed that the fitting coefficients were all higher than 0.96, which indicated an effective interpretation to the scattering information of beef by Lorentzian function; and consequently, Lorentzian parameters extracted in the above-mentioned spectral range were used for further analysis.

3.2.2. Lorentzian parameters

Fig. 6 shows the extracted Lorentzian parameters of all beef samples. It could be seen from this figure that Lorentzian parameters a, b and c of beef show respectively, similar spectral patterns and only differing in values among different samples. Overall, the spectral patterns of Lorentzian parameters a, b and c of beef were similar to those extracted from pork meat (Tao and Peng, 2015).

Fig. 7 depicts the average Lorentzian parameters of fresh and spoiled beef samples on day 0 and day 10, which corresponded to the average TVCs of 2.81 and 7.99 log CFU/g, respectively. From Fig. 7, significant spectral differences could be observed between fresh and spoiled beef, specifically in the spectral range of 470–585 nm. In the average spectrum of parameter a of fresh samples (day 0), clear peaks or valleys could be noticed around the wavelengths of 513, 530, 555 and 566 nm, while they were not observed for spoiled samples (day 10) (Fig. 7a). The spectral differences of parameter b between fresh and spoiled samples were
similar to those of parameter a (Fig. 7a and b). While, for parameter c, the spectral peaks and valleys appeared around 502, 525, 544 and 563 nm for fresh samples, instead of the above-mentioned four wavelengths in the spectra of parameters a and b, and similarly disappeared in spoiled samples (Fig. 7c). Bowen (1949) reported that the phenomenon of myoglobin oxidation would decrease the absorbance value at 555 nm, while increase around the wavebands of 542 and 578 nm. The characteristic wavelengths of 555 and 542 nm could be identified from the above-mentioned spectral differences between fresh and spoiled beef, which indicated that myoglobin oxidation occurred during beef storage in this study. Additionally, according to Millar et al. (1996), the characteristic wavelength of 502 nm in the spectrum of parameter c of fresh beef (Fig. 7c) could be assigned to the absorbance of metmyoglobin (MetMb) in meat.

3.2.3. Correlation analysis between Lorentzian parameters and beef TVC

In order to study the linear relationship between Lorentzian parameters and beef TVC, the Pearson correlation coefficient (R) was calculated at each wavelength in the spectral range of 470–1000 nm and the results are shown in Fig. 8. From Fig. 8, it was interesting to notice that the correlation pattern between parameter a and beef TVC was similar to that between parameter b and beef TVC, and these two patterns seemed to be line-symmetric in the selected spectral range. Additionally, it was also observed from Fig. 8 that the correlation results between parameters a, b and beef TVC were superior to those between parameter c and beef TVC, with the correlation coefficients varying in the ranges of −0.28 to 0.50, −0.50 to 0.10 and 0.12–0.39 for parameters a, b, and c, respectively.

By comparing the correlation results between Lorentzian parameters and beef TVC with those of pork meat reported previously (Tao and Peng, 2015), it was easy to notice a similarity in the correlation patterns between beef and pork meat, even though the correlation coefficients between Lorentzian parameters and beef TVC were not as good as those of pork meat (Fig. 8). The inferiority of correlations might be due to a smaller ratio of light scattering
and RMSEC of 0.93 for predicting pork meat TVC based on individual parameters. First, the individual Lorentzian parameters were determined based on the lowest RMSEP of 0.78 and RMSEC of 0.57 and 1.40 log CFU/g. Additionally, by comparing the modeling results by PCR method, parameter b performed best among the three individual parameters, with \( R_p \) and RMSEP of 0.72 and 1.47 log CFU/g (Table 2). Additionally, it could also be observed from Table 2 that PCR modeling could perform better than PCR method for all individual parameters. Peng et al. (2011) have investigated the capabilities of individual Lorentzian parameters for predicting beef TVC by PLSR method, and obtained the \( R_p \) and standard error of prediction set (SEP) values of 0.80, 1.51 log CFU/g, and 0.74, 0.99 log CFU/g using parameters a, b, c, respectively. Our previous study has also reported a \( R_p \) of 0.93 for predicting pork meat TVC based on individual Lorentzian parameter b (Tao and Peng, 2015). It was in accordance with the previous studies that parameter b performed best among the three individual parameters for predicting meat TVC; however, the \( R_p \) could only reach 0.72 for detecting low levels of TVC contamination in beef (Table 2). The Pearson correlation coefficients between Lorentzian parameters and beef TVC (Fig. 8) may explain the inferiority of the modeling results using individual parameters. As the linear regression methods of PCR and PLSR did not give satisfactory prediction results using individual parameters, the BPNN method which is one of the most commonly used non-linear regression methods was applied in this study. The BPNN model was established using the PCs from PCA analysis, and the node number of hidden layer was examined in the range of 3–13 for each Lorentzian parameter, respectively. For parameter a, the optimum node number of hidden layer was determined to be 9 based on the modeling results of different node numbers, and the corresponding modeling result was achieved with \( R_p \) and RMSEP of 0.78 and 1.18 log CFU/g (Table 3). By following similar procedures to parameter a, the BPNN models using parameters b and c were established and optimized, and their optimum node numbers of hidden layer, model training and prediction results are shown in Table 3. From Table 3, it could be seen that among the three individual parameters, parameter c with 3 nodes of hidden layer performed best for predicting low levels of beef TVC, with \( R_p \) and RMSEP of 0.81 and 1.27 log CFU/g. Additionally, by comparing the modeling results in Tables 2 and 3, it could be observed that BPNN models performed better than PCR and PLSR methods using parameters a and c, while inferior to PLSR method when using parameter b.

### 3.3. Modeling results using individual parameters

As mentioned above, three multivariate methods including PCR, PLSR and BPNN were used to establish the quantitative prediction models using individual parameters and parameter combinations, respectively. First, the individual Lorentzian parameters were tested for their capabilities in detecting low levels of TVC contamination in beef.

For PCR method, the optimum PCs were determined by explaining most (>99%) of the total variation in their spectral variables. The first five PCs accounted for 99.47% of the total variation of parameter a (76.85%, 18.68%, 2.93%, 0.54%, and 0.47% for PC1, PC2, PC3, PC4, and PC5, respectively), so they were used to establish the regression model for predicting beef TVC. The PCR modeling result using parameter a was achieved with \( R_p \) and RMSEC of 0.79 and 1.08 log CFU/g. However, the calibration result was not good enough to predict low levels of beef TVC contamination in meat, as the sizes of collagen fiber, myofibrils and sarcomeres in meat are all close to or larger than the used optical wavelengths (visible and short-wave near infrared light).

### Table 2

<table>
<thead>
<tr>
<th>Parameter</th>
<th>PCR method</th>
<th>PLSR method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Optimum no. of PCs</td>
<td>Optimum no. of LVs</td>
<td>RMSEC</td>
</tr>
<tr>
<td>( a )</td>
<td>5</td>
<td>0.79</td>
</tr>
<tr>
<td>( b )</td>
<td>4</td>
<td>0.84</td>
</tr>
<tr>
<td>( c )</td>
<td>4</td>
<td>0.41</td>
</tr>
</tbody>
</table>

### Table 3

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Optimum node number of hidden layer</th>
<th>( R_p )</th>
<th>RMSEC</th>
<th>( R_p )</th>
<th>RMSEP</th>
</tr>
</thead>
<tbody>
<tr>
<td>( a )</td>
<td>9</td>
<td>0.82</td>
<td>1.02</td>
<td>0.78</td>
<td>1.18</td>
</tr>
<tr>
<td>( b )</td>
<td>5</td>
<td>0.87</td>
<td>0.88</td>
<td>0.57</td>
<td>1.47</td>
</tr>
<tr>
<td>( c )</td>
<td>3</td>
<td>0.68</td>
<td>1.33</td>
<td>0.81</td>
<td>1.27</td>
</tr>
</tbody>
</table>

### Table 4

<table>
<thead>
<tr>
<th>Parameter combination</th>
<th>Original parameter combination</th>
<th>Standardized parameter combination</th>
</tr>
</thead>
<tbody>
<tr>
<td>Optimum no. of PCs</td>
<td>RMSEC</td>
<td>( R_p )</td>
</tr>
<tr>
<td>((a, b))</td>
<td>5</td>
<td>0.88</td>
</tr>
<tr>
<td>((a, c))</td>
<td>7</td>
<td>0.86</td>
</tr>
<tr>
<td>((b, c))</td>
<td>6</td>
<td>0.90</td>
</tr>
<tr>
<td>((a, b, c))</td>
<td>7</td>
<td>0.89</td>
</tr>
</tbody>
</table>
and RMSEP of 0.81 and 0.94 log CFU/g, and RMSEP of 0.86 and 0.98 log CFU/g, respectively. Using the determined PCs, the PCR modeling results were calculated and shown in Table 4. By comparing the PCR results by original and standardized parameter combinations. Among all the original parameter combinations, the parameter combination (a, b, c) performed best based on PCR method, with $R^2$ and RMSEP of 0.87 and 0.79 log CFU/g, respectively (Table 5). Furthermore, by comparing the results in Tables 4 and 5, it could be observed that for all original parameter combinations, the PLSR modeling method performed better than PCR method, which was in accordance with those using individual parameters.

Table 5 also shows the PLSR modeling results using the standardized parameter combinations. Compared to the results using original parameter combinations by PLSR method, it was observed that the model performances of parameter combinations (a, b) and (a, b, c) were enhanced by the data standardization, with $R^2$ improving from 0.78 and 0.78 to 0.83 and 0.85, respectively. Although the modeling result using the standardized (b, c) was not improved, it still achieved the best prediction result among all the standardized parameter combinations by PLSR method, with $R^2$ and RMSEP of 0.86 and 0.84 log CFU/g, respectively (Table 5).

Table 6 shows the BPNN modeling results using the original and standardized parameter combinations. By comparing the results in Tables 3 and 6, we found that for BPNN method, the models established using the original parameter combinations were not as good as those developed using individual Lorentzian parameters, while after the pre-processing of standardization, the model performances were significantly improved. Among all the original and standardized parameter combinations by BPNN method, the best modeling result was achieved by the standardized (a, b, c) with $R^2$ and RMSEP of 0.90 and 0.88 log CFU/g, respectively (Table 6). Overall, the standardization approach on the parameter combinations played a more important role in the BPNN method than in PCR and PLSR methods.

3.4. Modeling results using parameter combinations

As mentioned above, in order to comprehensively utilize the information within the three individual Lorentzian parameters, the parameter combinations were exploited for detecting low levels of bacterial contamination in beef in this study. In this section, the parameter combinations (a, b), (a, c), (b, c), and (a, b, c) were applied to develop the quantitative models for predicting beef TVC, by PCR, PLSR, and BPNN methods, respectively.

By implementing similar procedures to individual parameters, the optimum number of PCs for each original parameter combination was determined. The optimum numbers of PCs for the original parameter combinations (a, b), (a, c), (b, c), and (a, b, c) were 5, 7, 6, and 7, respectively. Using the determined PCs, the PCR modeling results were calculated and shown in Table 4. By comparing the modeling results in Tables 2 and 4, it was observed that all the original parameter combinations performed better than individual parameters for predicting low levels of TVC contamination in beef, which indicated the superiority of the parameter combinations based on PCR modeling method. Among all the original parameter combinations, the parameter combination (b, c) performed best based on PCR method, with $R^2$ and RMSEP of 0.81 and 0.94 log CFU/g, respectively (Table 4).

In order to overcome the magnitude differences among different Lorentzian parameters when performing the parameter combination method, the pre-processing method of standardization was applied in this study. The PCR modeling results using the standardized parameter combinations were shown in Table 4. By comparing the PCR results by original and standardized parameter combinations in Table 4, it was noticeable that the approach of standardization could significantly improve the model performance of (a, b), while no significant improvement on the model performances of other parameter combinations. The best PCR modeling result using the standardized parameter combinations was achieved by the combination (a, b), with $R^2$ and RMSEP of 0.86 and 0.93 log CFU/g, respectively (Table 4).

The modeling results using the original parameter combinations by PLSR method were shown in Table 5. Comparing these results to those using individual parameters in Table 2, it was clear that all original parameter combinations performed better than only using individual parameters, which was in accordance with those by the above-mentioned PCR method. Another similarity to PCR method was that, the parameter combination (b, c) performed the best for predicting low levels of TVC contamination in beef, with $R^2$ and RMSEP of 0.87 and 0.79 log CFU/g, respectively (Table 5).
models established in those studies could not be directly applied to meat chops. Additionally, the beef TVCs used for model development were high in the above-mentioned studies, with the minimum TVC values around 5.0 log CFU/g, which suggested a difficulty in predicting low levels of bacterial contamination in beef by FTIR technique.

Most recently, Argyri et al. (2010) applied the FTIR technique in tandem with multilayer perceptron neural network to detect the microbial spoilage of beef fillets, and based on the calculated value of bias factor ($B_f$), it could be inferred that the network under-estimated TVC in semi-fresh and spoiled samples ($B_f < 1$), whereas for fresh samples, over-estimation of microbial population was evident ($B_f > 1$). In addition, the calculated value of accuracy factor ($A_f$) showed that the predicted TVCs were 18.1%, 12.2%, and 8.4% different (either above or below) from the observed values for fresh, semi-fresh, and spoiled meat samples, respectively; which indicated the inferior performance of the network for microbial count prediction in fresh beef. Kodogiannis et al. (2014) applied the Extended Normalized Radial Basis Function (ENRBF) neural network to analyze the FTIR spectra of TVC growth in beef fillets, and obtained relatively satisfactory prediction results. However, it should be noted that even though the samples were prepared in beef fillets, the collection of FTIR spectra still required the assistance of ZnSe ATR (Attenuated Total Reflectance) crystal for detection, which indicated the difficulty in application to on-line detection. Compared to the above-mentioned studies on detecting beef TVC, this study shows the potential for nondestructively detecting low levels of bacterial contamination in beef by hyperspectral scattering imaging and the presented data analysis method, with the best modeling results in $R_p$ and RMSEP of 0.86 and 0.93 log CFU/g, 0.87 and 0.79 log CFU/g, 0.90 and 0.88 log CFU/g by PCR, PLSR, and BPNN methods, respectively. The lower prediction results may be due to the relative nonuniformity of bacterial growth in beef chops, compared to the sample preparation of minced meat in previously reported studies (Ammor et al., 2009; Argyri et al., 2013; Ellis et al., 2004).

4. Conclusions

Meat TVC is an important reference index for the sanitary quality and safety evaluation. In this study, hyperspectral scattering imaging combined with Lorentzian function, and PCR, PLSR, BPNN multivariate methods were used to detect low-level bacterial contamination in beef. The prediction models established using individual Lorentzian parameters did not perform well for detecting low levels of TVC contamination in beef, and the best prediction result could only be achieved with $R_p$ and RMSEP of 0.81 and 1.27 log CFU/g, respectively.

In this study, the approach of parameter combination was exploited in order to comprehensively utilize the information within the three individual Lorentzian parameters, and the results revealed that the method of parameter combination could significantly improve the model performances for predicting low-level TVC contamination in beef based on PCR and PLSR methods, while for BPNN method, the prediction results could only be enhanced after the standardization of parameter combinations. The results also showed that for PCR and PLSR linear regression methods, the improvement of model performance by the standardization of parameter combinations was not as significant as that for non-linear BPNN method. Based on the parameter combination method, the best modeling results for determining beef TVC were achieved with $R_p$ and RMSEP of 0.86 and 0.93 log CFU/g, 0.87 and 0.79 log CFU/g, 0.90 and 0.88 log CFU/g by PCR, PLSR, and BPNN methods, respectively. Imaging more positions of the sample will be considered to improve the model accuracy in the next step. Meanwhile, future research requires more accurate hardware system, and also advanced data processing algorithm could provide more precise prediction equation to promote the development of this technique for industrial application.

Overall, this study indicated that hyperspectral scattering imaging in tandem with the proposed parameter combination method could be used to predict the TVC contamination in beef, which could cover low levels of bacterial contamination, compared to other reported studies. Hyperspectral scattering imaging is promising to become a valid tool for nondestructively monitoring the bacterial contamination in meat during the distribution chain.

Acknowledgements

The authors gratefully acknowledge the Special Fund for Agro-scientific Research in the Public Interest Program (Project No. 201003008), the National Science and Technology Support Program (Project No. 2012BAH04B00), and the China Postdoctoral Science Foundation (Project No. 2014MS61096) for supporting this research.

References


Eng. 142, 118–131

Feng, Y.-Z., Sun, D.-W., 2013. Determination of total viable count (TVC) in chicken breast fillets by near-infrared hyperspectral imaging and spectroscopic transforms. Talanta 105, 244–249.


