

Technical University of Denmark



## Multispectral Image Analysis for Robust Prediction of Astaxanthin Coating

**Ljungqvist, Martin Georg; Frosch, Stina; Nielsen, Michael Engelbrecht; Ersbøll, Bjarne Kjær**

*Published in:*  
Applied Spectroscopy

*Link to article, DOI:*  
[10.1366/12-06823](https://doi.org/10.1366/12-06823)

*Publication date:*  
2013

*Document Version*  
Publisher's PDF, also known as Version of record

[Link back to DTU Orbit](#)

*Citation (APA):*  
Ljungqvist, M. G., Frosch, S., Nielsen, M. E., & Ersbøll, B. K. (2013). Multispectral Image Analysis for Robust Prediction of Astaxanthin Coating. *Applied Spectroscopy*, 67(7), 738-746. DOI: 10.1366/12-06823

## DTU Library

Technical Information Center of Denmark

---

### General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

# Multispectral Image Analysis for Robust Prediction of Astaxanthin Coating

Martin Georg Ljungqvist,<sup>a,b,\*</sup> Stina Frosch,<sup>b</sup> Michael Engelbrecht Nielsen,<sup>b</sup> Bjarne Kjær Ersbøll<sup>a</sup>

<sup>a</sup> Technical University of Denmark (DTU), Department of Informatics and Mathematical Modelling, Richard Petersens Plads, 2800 Kongens Lyngby, Denmark

<sup>b</sup> Technical University of Denmark (DTU), National Food Institute, Division of Industrial Food Research, Søtofts Plads, 2800 Kongens Lyngby, Denmark

The aim of this study was to investigate the possibility of predicting the type and concentration level of astaxanthin coating of aquaculture feed pellets using multispectral image analysis. We used both natural and synthetic astaxanthin, and we used several different concentration levels of synthetic astaxanthin in combination with four different recipes of feed pellets. We used a VideometerLab with 20 spectral bands in the range of 385–1050 nm. We used linear discriminant analysis and sparse linear discriminant analysis for classification and variable selection. We used partial least squares regression (PLSR) for prediction of the concentration level. The results show that it is possible to predict the level of synthetic astaxanthin coating using PLSR on either the same recipe, or when calibrating on all recipes. The concentration prediction is adequate for screening for all recipes. Moreover, it shows that it is possible to predict the type of astaxanthin used in the coating using only ten spectral bands. Finally, the most selected spectral bands for astaxanthin prediction are in the visible range of the spectrum.

Index Headings: Multispectral; Image analysis; Spectral imaging; NIR; Astaxanthin; Fish feed; Coating.

## INTRODUCTION

Astaxanthin is a naturally occurring carotenoid with a high antioxidant activity essential for growth and survival, and it is important for the development of color in salmonid fishes.<sup>1</sup> The primary use of astaxanthin in aquaculture is as a feed additive to ensure that farmed salmon and trout have a similar appearance to their wild counterparts.<sup>2</sup> Astaxanthin is very expensive,<sup>3</sup> and therefore optimization and quality control of its use in fish feed production is important.

Synthetic astaxanthin is more easily available and costs slightly less than natural astaxanthin, and it is therefore used more often in industry. However, there is a demand for natural astaxanthin for the organic salmonid fish market, where natural astaxanthin is mandatory.

An automatic vision system for at-line pigment quality control of concentration level and type would be of great benefit to the industry both in relation to process control and process optimization. Furthermore, the astaxanthin coating type can be verified by the customers in the organic salmonid fish market.

Multispectral imaging of astaxanthin coating has previously been done in an earlier study by Ljungqvist et al.,<sup>4</sup> which showed that it is possible to predict the presence of astaxanthin in fish feed pellet coating. However, this research did not use

different types of astaxanthin (synthetic vs. natural), nor were different concentration levels of astaxanthin coating used.

Besides this, to the authors' knowledge, no further work has previously been done on analyzing the coating of fish feed using image analysis. A study by Zhu et al.<sup>5</sup> used a spectroradiometer (325–1075 nm) for classification of four different fish feed pellet recipes; however, they did not use image analysis.

Multispectral image analysis has shown good results in previous biological applications where it is of interest to detect subtle differences in color and surface chemistry.<sup>6–17</sup>

The aim of this study is to investigate the possibility of predicting the type and concentration level of astaxanthin coating of aquaculture feed pellets using multispectral image analysis. Furthermore, the aim is to make this coating prediction robust for different underlying pellet recipes in order to make a model that can predict astaxanthin coating irrespective of the composition of the underlying pellet recipe.

Hypotheses: Using multispectral image analyses it is possible to do the following:

1. Discriminate between synthetic astaxanthin coating and natural astaxanthin coating
2. Predict the concentration level of synthetic astaxanthin coating using regression analysis
3. Predict the concentration level of synthetic astaxanthin coating using regression analysis irrespective of the composition of the pellet recipe

## MATERIALS AND METHODS

**Pellets.** Four different feed recipes with the same amount of protein (but of different origin) were investigated, and all but one contained the same amount of oil (see Table I). The four recipes were all based on normal commercial fish feed for salmonid fish. The recipes were named after their dominant ingredient; the standard recipe was called control (CON), then there was maize gluten (MG), soya protein concentrate (60%) (SPC), and a control recipe with a different amount of oil (CON31). All pellets had an approximate production diameter of 4.5 mm.

The feed material was extruded through a die plate with holes of a certain diameter, which determined the diameter of the pellets. On the other side of the disk a set of rotating knives cut the material into shorter cylinder-shaped pellets. After extrusion the pellets were vacuum coated. All pellets were stored dark and dry in sealed plastic bags for about 12 months between production and image acquisition.

Received 15 August 2012; accepted 27 February 2013.

\* Author to whom correspondence should be sent. E-mail: malj@imm.dtu.dk.

DOI: 10.1366/12-06823

**TABLE I.** Recipes of the pellets, with ingredient amounts shown as percentages. The recipe names are Control (CON), Maize gluten (MG), Soya protein concentrate (SPC), and Control 31 (CON31). All recipes have the same total amount of protein. CON, MG, and SPC have the same total amount of oil.

Ingredients	CON	MG	SPC	CON31
Fish oil	17	18	21	24
Wheat flour	23	13	11	14
Fish meal	60	29	28	62
Maize gluten	0	40	0	0
Soya (60%)	0	0	40	0
Total	100	100	100	100
Total content				
Protein	45	45	45	45
Oil	24	24	24	31

**Coating.** The pellets included in this study consist of several classes: pellets coated with fish oil using different amounts of added synthetic astaxanthin (20, 40, 60, 80 parts per million (ppm)), as well as pellets coated with fish oil with 20 ppm of added natural astaxanthin. Moreover, a coating of only fish oil without additional astaxanthin was used. Fish oil in itself typically contains a small amount of natural astaxanthin; however, this was assumed to be less than 1 ppm and here was referred to as a coating of 0 ppm concentration. Astaxanthin is commonly measured in parts per million, and it is measured in mass, so here parts per million corresponds to milligrams per kilogram.

The pellets of recipe CON and CON31 had five coating concentration levels in total (0, 20, 40, 60, 80 ppm) of synthetic astaxanthin. The pellets of recipe MG and SPC had three concentration levels in total (0, 20, 60 ppm) of synthetic astaxanthin (see Table II).

The synthetic astaxanthin used was a cold water dispersible (BASF SE, Germany). It was dissolved in water.

The natural astaxanthin used consists of (a 50/50 mix of) monoesters and diesters of fatty acids. It was extracted from shrimp and was dissolved in oil.<sup>18</sup>

The distribution of the surface coating was unknown, and some amount of variation was likely to have occurred.

The common pellet compound for each recipe gives a spectral response that will be present for all concentration levels and astaxanthin types. Each pixel is thus a combination of the reflectance of a set of constituents. This mix was assumed to be of equal amount for each pellet recipe except for the difference of the astaxanthin coating concentration and astaxanthin type that we want to isolate in our model.

**Equipment.** The equipment used for image acquisition was a camera and lighting system called VideometerLab (Videometer A/S, Hørsholm, Denmark), which supports a multispec-

**TABLE II.** All 18 available pellet types. Coating concentration of astaxanthin dissolved in fish oil for the different pellet recipes. Amount of astaxanthin added to the fish oil coating in ppm. Fish oil coating is represented by 0 ppm.

Recipe	Astaxanthin	Concentration (ppm)				
CON	Synthetic	0	20	40	60	80
CON	Natural	—	20	—	—	—
MG	Synthetic	0	20	—	60	—
SPC	Synthetic	0	20	—	60	—
CON31	Synthetic	0	20	40	60	80

**TABLE III.** The wavelength of light sources in the VideometerLab device and their spectral representation.

Band	Wavelength (nm)	Color
1	385	Ultraviolet A
2	430	Violet
3	450	Violet, blue
4	470	Blue
5	505	Green
6	565	Green
7	590	Yellow, orange
8	630	Red
9	645	Red
10	660	Red
11	700	Red
12	850	NIR
13	870	NIR
14	890	NIR
15	910	NIR
16	920	NIR
17	940	NIR
18	950	NIR
19	970	NIR
20	1050	NIR

tral resolution of up to 20 spectral bands.<sup>19</sup> These are distributed over the ultraviolet A, visible, and first near infrared (NIR) region. The spectral range is from 385 to 1050 nm (see Table III). In comparison with spectroscopy, multispectral imaging includes both spectral information and spatial information simultaneously.

This system uses a Point Gray Scorpion SCOR-20SOM grayscale camera. The objects of interest are placed inside an integrating sphere (Ulbricht sphere) with uniform diffuse lighting from light sources placed around the rim of the sphere. All light sources are light-emitting diodes except for 1050 nm, which is a diffused laser diode. The curvature of the sphere and its matte-white coating ensure a uniform diffuse light so that specular effects are avoided and the amount of shadow is minimized. The device is calibrated radiometrically with a following light and exposure calibration. The system is also geometrically calibrated to ensure pixel correspondence for all spectral bands.<sup>20</sup>

The image resolution is 1280 × 960 pixels. Each file contains 20 images, one for each spectral band. In this situation one pixel represents approximately 0.072 × 0.072 mm. The Scorpion camera has a 12 bit analog-to-digital converter, and the system uses 8 bit data output from the camera. The correction for calibration gives reflectance intensity output of 32 bit precision.

**Spectroscopy.** In order to explore the spectral properties of astaxanthin further, a spectrometer was used. Absorption spectra of synthetic astaxanthin in a solution of fish oil along with plain fish oil were recorded in the visible and NIR range using a NIRSystems 6500 absorption spectrometer (Foss NIRSystems Inc., USA). The absorption spectra were transformed to reflection values using the standard relation  $A = -\log(R)$ , where  $A$  is the absorption values and  $R$  is the reflection values. Every second nanometer was recorded in the visible and NIR range.

**Image Acquisition.** A Petri dish filled with pellets arranged randomly resembles the at-line inspection that industry would desire for this application. The situation represents the disordered way in which the pellets lie on the conveyor belt and in the final bin at the end of the production line. In

comparison with spectroscopy that measures a point, multispectral imaging also includes spatial information, which covers many pellets simultaneously and arguably gives a more robust result.

For each of the total 18 pellet types (recipe, concentration level, and astaxanthin type) a total of four images of filled plastic Petri dishes (diameter 9 cm) of pellets were taken. Furthermore, a total of 60 images of CON pellets with 20 ppm synthetic astaxanthin (30 images) and 20 ppm natural astaxanthin (30 images) coating were captured. The pellets were at normal room temperature during image acquisition.

The pellets were segmented from the light gray background using a grayscale threshold in combination with morphological methods.<sup>21</sup> Optimally, the pellets should be segmented individually. However, since that was not feasible for this type of image, the segmented pellet cluster was divided into subregions of maximally  $100 \times 100$  pixels, and the mean of each region was used as a sample. One full subregion represented approximately the area of three pellets. A subregion rather than the individual pellet was therefore considered a sample. This was done in order to increase the number of samples to avoid an ill-posed problem so we could complete the statistical analysis described in the section below. On average one image consisted of about 120 subregions (samples).

Adjacent subregions tend to be positively correlated (spatial autocorrelation), since one pellet can be divided and can fall into several subregions. Yet since we had a large number of subregions, the proportion of autocorrelated subregions was reduced, and this was assumed not to affect the results.

Standard red-green-blue (sRGB) color image representations of the multispectral images for this paper were done using multispectral color mapping using penalized least-square regression described in Dissing et al.<sup>22</sup>

**Data Analysis.** The analyses of the images of the pellets were done in order to investigate hypotheses 1–3 presented above.

The reference values (ground truth) were that the pellets were coated with a specific amount of synthetic or natural astaxanthin: the nominal values. However, since the surface distribution was unknown, the coating variation from pixel to pixel was unknown. A way to solve this was to represent each image subregion using the pixel mean values as samples. In this manner the variance of all pixels in each pellet were evened out, and a distinct sample was achieved; each pellet became an observation.

The number of samples is denoted  $n$ , and the number of variables (the 20 spectral bands) is denoted  $p$ . The stored data of  $n$  samples and  $p$  variables is denoted as matrix  $\mathbf{X}$ . The ground truth reference values (the known concentration levels of astaxanthin) are stored in vector  $\mathbf{y}$  with length  $n$ . The predicted (estimated) value of  $\mathbf{y}$  is denoted  $\hat{\mathbf{y}}$ .

All image analyses and statistics were carried out using Matlab 7.9 (Mathworks Inc., Natick, MA, USA).

**Discriminant Analysis.** Statistical discriminant analysis of the images was made for testing hypothesis 1. This analysis performed discrimination tests with only two groups of samples. Hotelling's  $T^2$  test was used in order to see if the two group means were significantly different.<sup>23,24</sup>

The classification methods used were linear discriminant analysis (LDA)<sup>23</sup> and sparse linear discriminant analysis (SLDA).<sup>25</sup>

The LDA method is based on the Mahalanobis distance and assumes that the observations in each group are normally distributed. LDA is based on a distance to the group mean weighted by the variance. Discrimination using LDA was trained and tested using multiple splits of the data.

In order to identify which wavelengths were of most interest for classification of coating types, the SLDA method was used. The SLDA method is designed for ill-posed problems where the number of variables is larger than the number of observations: a so-called large  $p$  small  $n$  problem ( $p > n$ ). That was not the case in this study, but since SLDA performs both variable selection and classification, it suits our purposes well on those aspects. SLDA uses the elastic net (EN) for variable selection.<sup>26</sup> The EN tends to select variables that are correlated with each other.

Two model parameters are needed for EN:  $\lambda_1$  for the  $L_1$  norm for determining the number of nonzero coefficients and  $\lambda_2$  for the Euclidean  $L_2$  norm for regularization. The regularization was less important here, since we did not have ill-posed problems. The model consists of the variable weights (coefficients) in  $\beta^{\text{en}}$ , see Eq. 1.

$$\beta_j^{\text{en}} = \arg \min_{\beta_j} (\|\mathbf{y} - \mathbf{X}\beta_j\|_2^2 + \lambda_2 \|\beta_j\|_2^2 + \lambda_1 \|\beta_j\|_1) \quad (1)$$

The estimated coefficients are then multiplied by  $(1 + \lambda_2)$  to get the final EN solution.

The two parameters were selected using cross-validation (CV) on the calibration set. The  $\lambda_1$  parameter steers the selection of variables and was calculated so that the number of selected variables was varied from 1 to 10. The  $\lambda_2$  parameter was varied with 12 logarithmic steps from  $10^{-7}$  to 10. The data  $\mathbf{X}$  were normalized before each calculation of the SLDA so that each variable got unit length.

For two groups, the SLDA algorithm calculates one sparse discriminant component that gives the best classification of the groups. The calibration of SLDA was done using CV and then validated on the validation set, and this was wrapped in multiple splits of the data for calibration and validation. This means that a discriminant model was calibrated using CV on 100 different randomized calibration sets of the data using 70% of the samples and likewise validated on the corresponding 100 validation sets.

**Regression.** The concentration level was analyzed using the partial least squares regression (PLSR) method.<sup>27,28</sup> The number of samples in the calibration set was  $n_c$ , and the number of components (factors)  $n_f$  tested in the calibration step was varied from 1 to 20 ( $p$ ). The data were mean centered, and the number of components to be used in the PLSR model was decided using CV on the calibration set, while minimizing the residual sum of squares (RSS). The “one standard error rule”<sup>27</sup> for selection was used: The least number of components with RSS value inside the range of one standard error of the lowest value of RSS was selected. This method tends to select a more parsimonious model with fewer components than just using the lowest value of RSS.

Then this model, using  $n_f$  components, was validated on the validation samples.

The coefficient of determination ( $R^2$ ) is a measure of how much variation is explained using the model and was calculated for the prediction of the validation set.  $R^2$  is basically the ratio of the RSS and the total sum of squares.

Furthermore, the ratio of the standard error of prediction (RPD) was calculated as a measure of how well the model predicts. The RPD is the ratio between the standard error of prediction (SEP) and the standard deviation (SD) of the reference values  $y$ :  $RPD = SD/SEP$ .

The SEP is equal to the root mean square error of prediction (RMSEP). An RPD value of 1.0 means that the model cannot predict accurately, since this means that the mean error is equal to the standard deviation of the reference values. An RPD value higher than 2.5 is considered satisfactory for screening, and values of 5–10 are adequate for quality control, according to Williams and Sobering.<sup>29</sup>

**Model Selection and Validation.** For parameter calibration of PLSR and SLDA methods a ten-fold cross-validation (CV) was used, where all samples were randomly partitioned into ten approximately equally sized parts. Then nine parts were used for training and one for testing, and this was repeated ten times.<sup>27</sup>

In order to calibrate the PLS regression model parameters, a ten-fold CV was done on a calibration set of 70% of the samples. Then the chosen model was validated on the remaining 30% of the validation samples. The calibration and validation set was chosen randomly, with all concentration levels present in both.

For discriminant analysis using SLDA and LDA, multiple splits of the data were used with 100 different randomized training (calibration) sets with 70% of the samples, and test (validation) sets using 30% of the samples. The multiple split method is a variant of CV and is also known as repeated random subsampling validation. The mean classification results of these multiple splits are reported.

**Recipes.** How the underlying pellet composition affects the image-based prediction of synthetic astaxanthin was also tested in order to make a robust model that can predict astaxanthin coating irrespective of the composition of the underlying pellet recipe.

Combinations of three recipes were used to calibrate the prediction model, and the remaining recipe was used for validation (prediction). Finally, all four recipes were used both

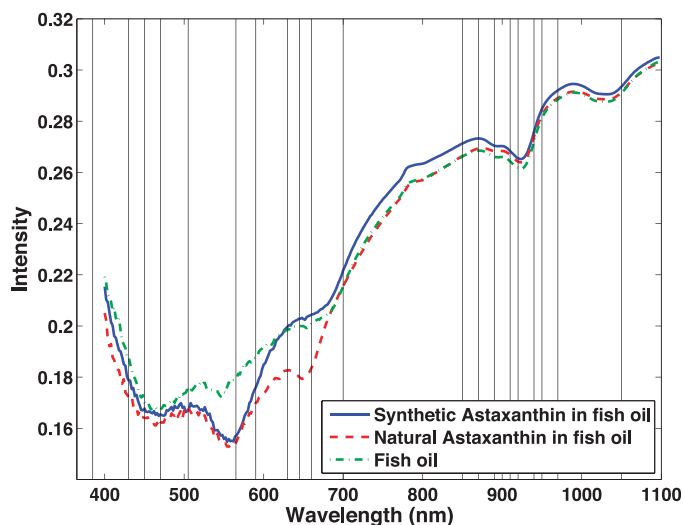


FIG. 1. Spectrometer spectra of synthetic astaxanthin in fish oil, natural astaxanthin in fish oil, and plain fish oil. The values are converted to reflectance. The wavelengths of the VideometerLab are marked using vertical lines.

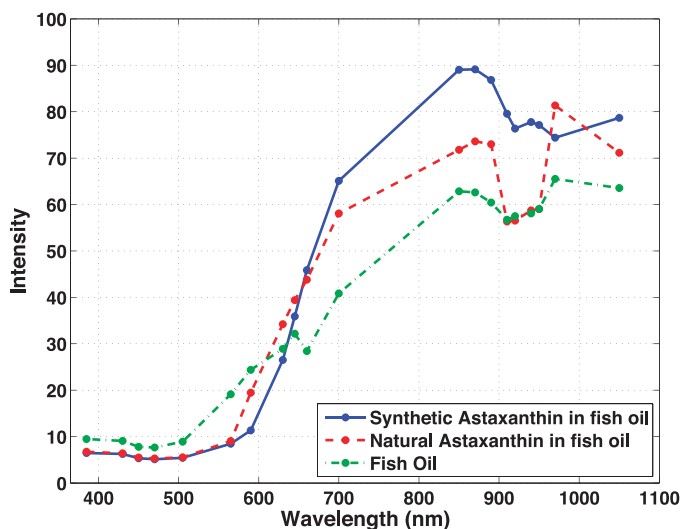


FIG. 2. VideometerLab reflectance spectra of synthetic astaxanthin in fish oil, natural astaxanthin in fish oil, and plain fish oil.

for calibration and for validation in the same manner as described above.

## RESULTS

The spectrometer results show a large spectral difference between fish oil with added (synthetic/natural) astaxanthin and plain fish oil in the range of 450–600 nm, as well as a spectral difference between fish oil with natural astaxanthin and fish oil with synthetic astaxanthin in the approximate range of 600–700 nm (see Fig. 1). The spectrometer results show similar characteristics and separation as the spectra from the VideometerLab images of astaxanthin in fish oil (see Fig. 2).

An image of a Petri dish filled with CON pellets is shown in Fig. 3, with subregions shown in Fig. 4. Pellets with the natural astaxanthin coating can be seen in Fig. 5. All concentration levels of the CON pellets are shown in Fig. 6, and the four different pellet recipes with plain fish oil coating can be seen in Fig. 7.

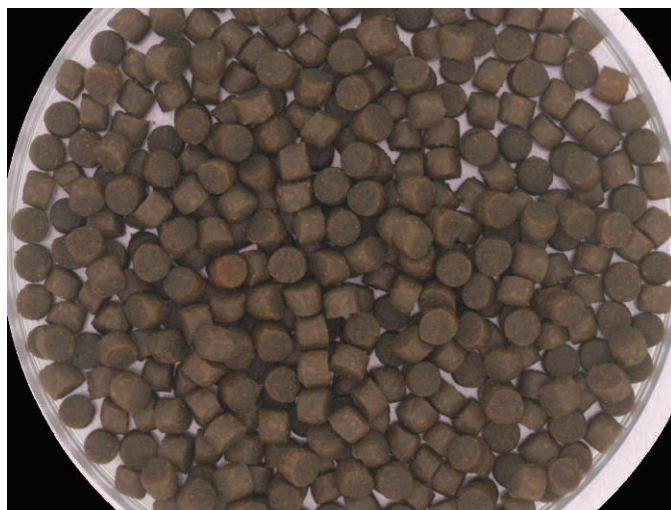


FIG. 3. Standard RGB version of an image of a Petri dish filled with CON pellets with coating using synthetic astaxanthin level of 20 ppm.

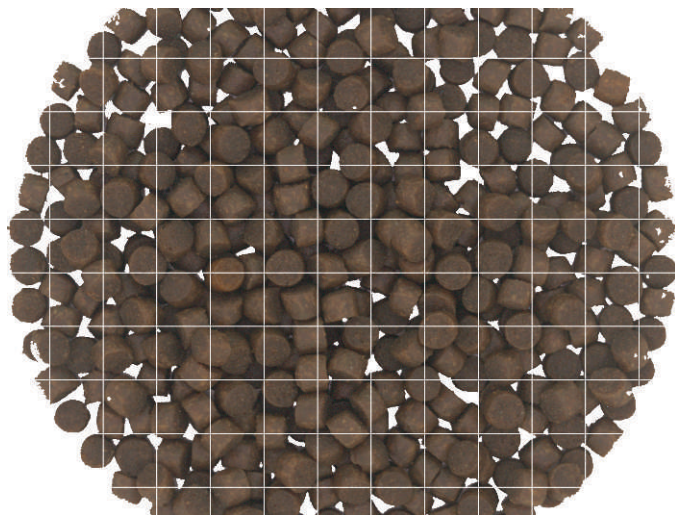


FIG. 4. The image in Fig. 3 (CON pellets with coating using synthetic astaxanthin level of 20 ppm), with the background segmented and subregions visualized.

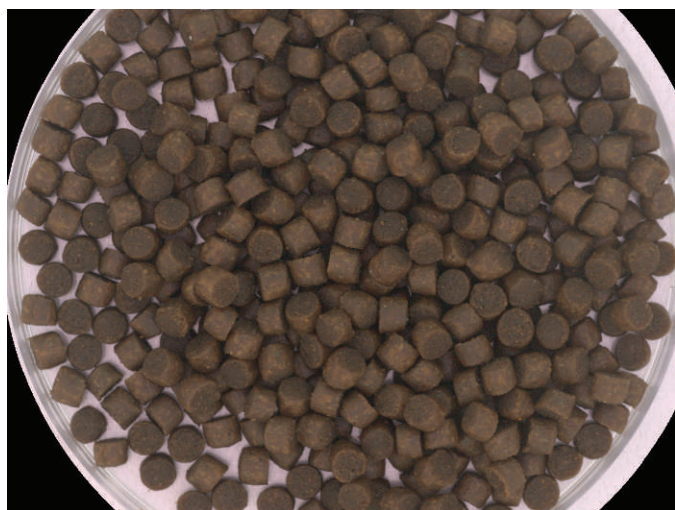


FIG. 5. Standard RGB version of an image of CON pellets with coating using added natural astaxanthin level of 20 ppm.

**Astaxanthin Type.** The investigation of hypothesis 1 is presented here: Is it possible to discriminate between synthetic astaxanthin coating and natural astaxanthin coating?

The spectral data of the two different astaxanthin coating types have group means that are significantly different at a

0.1% significance level using Hotelling's  $T^2$  test. This indicates that the spectral reflectance of the two groups is separable.

The discriminant analysis shows that LDA performs a 98% classification of synthetic astaxanthin coating and natural astaxanthin coating (see Table IV). The mean spectra of the two groups with one standard deviation marked can be seen in Fig. 8, where it shows that the two groups' mean spectra have a large overlap, while the standard deviations differs somewhat.

Using SLDA for classification of synthetic astaxanthin coating and natural astaxanthin coating gives a 98% classification using ten bands (see Table IV). The two groups projected on the sparse discriminant component can be seen in Fig. 9. The bands chosen cover the full range of the total 20 spectral bands used, with the majority of the most often selected bands in the visible range, and some bands in the NIR range. The bands that were selected in all validation sets were 385, 450, 505, 590, 660, 850 nm (see Fig. 10).

In order to investigate whether it is reasonable to use the image subregions as subsamples, the 60 images of CON pellets were analyzed in order to discriminate between synthetic and natural astaxanthin coating. The mean of the whole segmented pellet cluster in each image was used as samples ( $n = 60$ ), and this was compared with subregions on four images of each group ( $n = 1034$ ). The results of this show that, for the mean values of the full pellet clusters, LDA with multiple splits give 100% classification. For the subregions, the classification is about 98% for LDA (see Table IV). This shows that it is possible to get approximately the same result with subregions as without subregions for this classification problem, so therefore it is sound to use subregions in this study.

Also, with these 60 images, we tested whether one, two, three, or four subregions per image might be enough to discriminate between these two groups. These subregions were in the center of the segmented pellet cluster, and the four subregions were not adjacent to each other. Using LDA with multiple splits, we can see that using one subregion per image is enough to get a 100% correct classification of the type of astaxanthin coating used on the pellets. The same result was also achieved with two, three, and four subregions per image.

**Astaxanthin Concentration Level.** The investigation of hypothesis 2, concerning whether it is possible to predict the concentration level of synthetic astaxanthin coating, is presented here.

When calibrating the PLSR model, the RSS value drops significantly, until ten components approximately, and then flattens out for all tests. The PLSR results show good results for all recipes with  $R^2$  values above 0.89 and RPD values above 3, which is considered good enough for screening<sup>29</sup> (see Table V and Fig. 11).

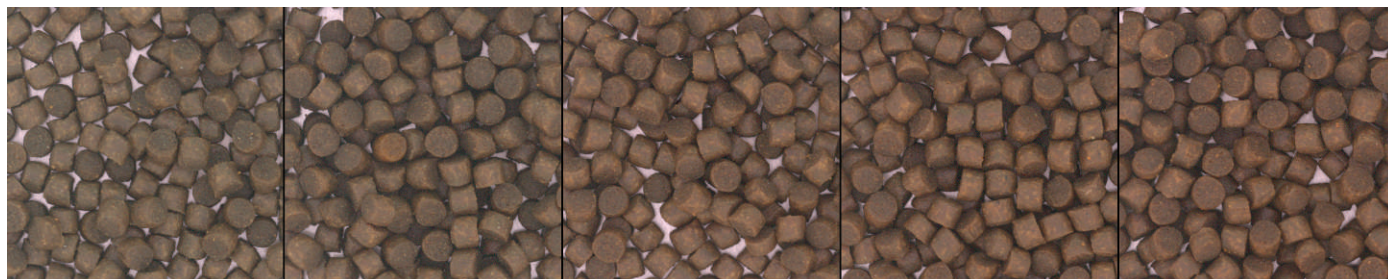


FIG. 6. Standard RGB version of images of CON pellets. The concentration level of synthetic astaxanthin is from left to right 0, 20, 40, 60, 80 ppm.

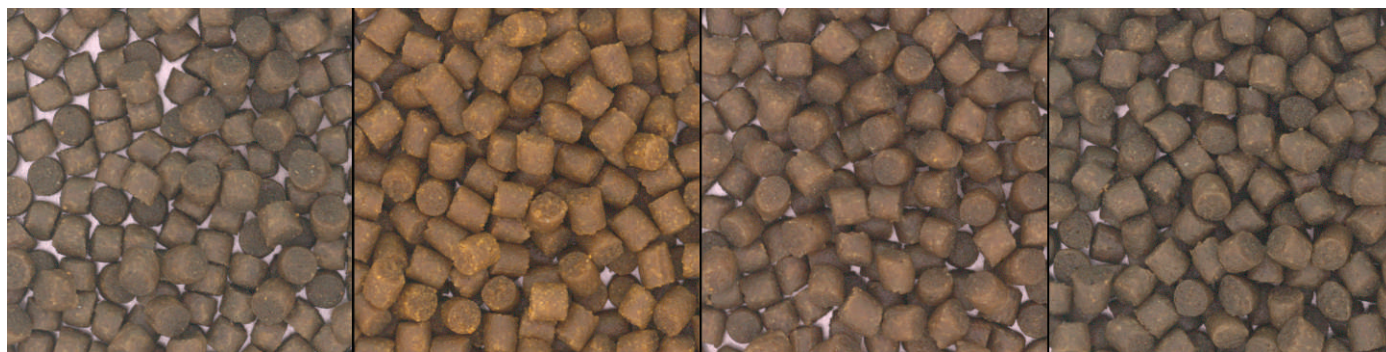


FIG. 7. Standard RGB version of images of pellets with plain fish oil as coating (0 ppm). The pellet recipe from left to right is CON, MG, SPC, CON31.

TABLE IV. Discriminant analysis of pellets with synthetic astaxanthin and natural astaxanthin coating. Classification test accuracy of pellet coating type using LDA and SLDA with multiple splits for training and test set. Also the standard deviation of the accuracy of the multiple splits is shown. The true group rates of classification are also presented (sensitivity and specificity). The total number of samples was 1034. All pellets were of CON recipe, and the amount of astaxanthin added to the fish oil coating was 20 ppm.

Method	Bands ( $p$ )	Accuracy	Accuracy (SD)	Natural rate	Synthetic rate
LDA	20	0.9874	0.0049	0.9825	0.9923
SLDA	10	0.9814	0.0077	0.9718	0.9911

The PLSR components' weights show the contribution of different spectral regions to the model, and thus their significance. Looking at the seven first components from the PLSR model of the CON pellets, we see that there are large weights between 400 and 600 nm, as well as at 700 nm (see Fig. 12).

**Recipe Tests.** The investigation of hypothesis 3 is presented here, namely, to see if it is possible to find a common component for astaxanthin prediction, independent of the underlying compounds in the recipes. It is of interest to see whether the spectral characteristic of synthetic astaxanthin is clearly identifiable, independent of the base spectra from the pellet recipe in order to make a robust model.

The mean spectral reflection of each recipe, with a coating of 20 ppm synthetic astaxanthin in fish oil, can be seen in Fig. 13.

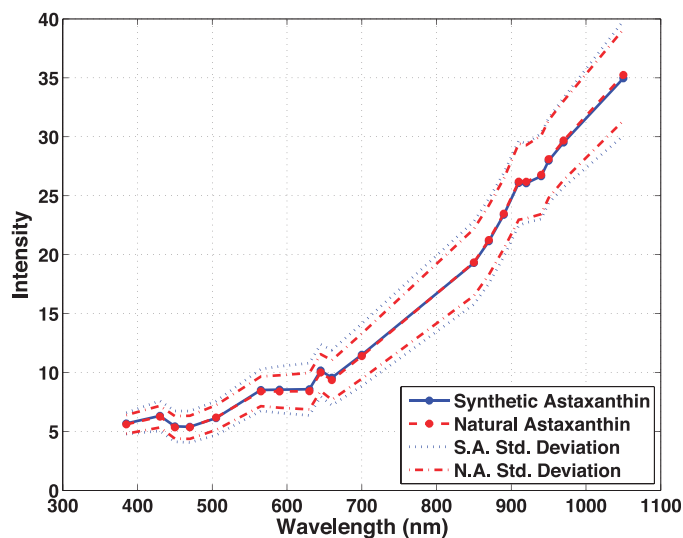


FIG. 8. VideometerLab reflectance spectra of CON image subregions. Images of pellets with added synthetic astaxanthin and added natural astaxanthin, with one standard deviation visualized.

When predicting astaxanthin coating using PLSR with three recipes for calibration and the remaining recipe as validation, the results are somewhat robust, with  $R^2$  between 0.76 and 0.87. However, when using all four recipes for calibration and validation, we get even better results, with an  $R^2$  value of 0.89 and an RPD above 3 (see Table VI).

## DISCUSSION

The spectrometer results and the reflection spectra from the VideometerLab images partly corresponds to previous studies of astaxanthin,<sup>30–32</sup> where they found absorbance peaks of astaxanthin around 450–505 nm and also secondary peaks around 500–600 nm for various solvents, as well as at 870 nm.

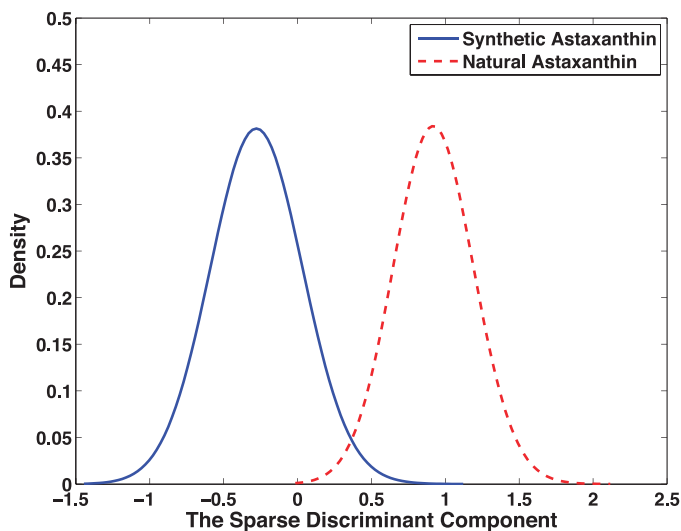


FIG. 9. Estimated density of CON pellets with synthetic astaxanthin and natural astaxanthin, projected on the sparse discriminant component (SD) (consisting of ten selected spectral bands) calculated using SLDA.

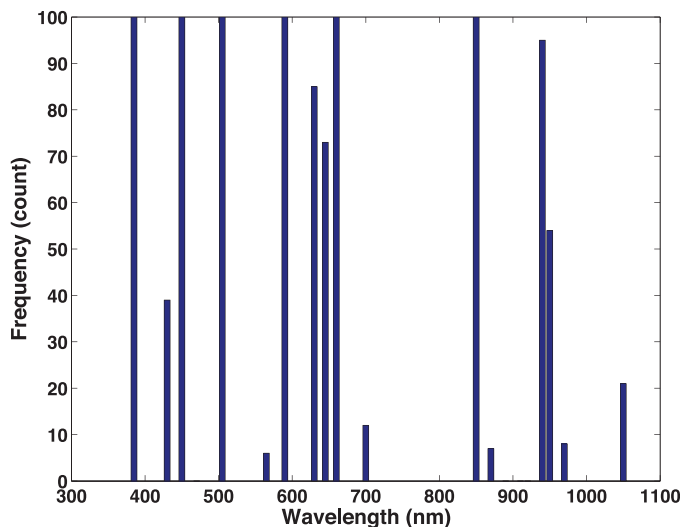


FIG. 10. Frequency of spectral bands selected using SLDA with 100 multiple splits for discrimination between CON pellets with synthetic astaxanthin and natural astaxanthin.

Hotelling's  $T^2$  calculates the statistical significance of the difference between group means. However, since we have a large number of correlated samples, this affects the significance value in an optimistic way, and the difference might not necessarily be significant in practice.

The classification between synthetic astaxanthin and natural astaxanthin coating works well using ten bands, while the regression problem for concentration level of synthetic astaxanthin needs ten or more components. Creating a model for the concentration level seems more complex, probably because the coating layer is relatively thin and the reflection of the astaxanthin coating is mixed with the background (the pellet compound). Also, a PLSR component is a linear combination of all 20 spectral bands, so it is not exactly comparable with selected bands of the SLDA classification method.

Looking at the PLSR components, we see that components 1 and 2 include a small amount of all spectral bands. PLSR component 4 onward show considerable characteristics for synthetic astaxanthin coating concentration prediction using the CON pellets, as can be seen in Fig. 12. We also see that there are large weights in the visible region of the spectrum, which well corresponds to the spectrometer results and with the band selection results using SLDA shown in Table IV and Fig. 10.

The storage time of the pellets between production and imaging in this study will affect the astaxanthin concentration in the coating, since astaxanthin oxidizes with time. This might affect the spectral response of astaxanthin coating; however, it could not be found in the literature how much astaxanthin

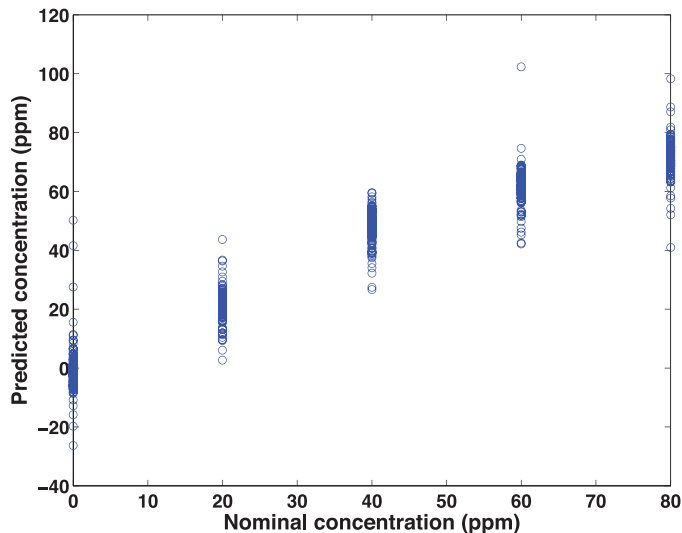


FIG. 11. PLSR prediction of synthetic astaxanthin concentration level (ppm) on the validation set, using images of CON pellets and 13 PLSR components.

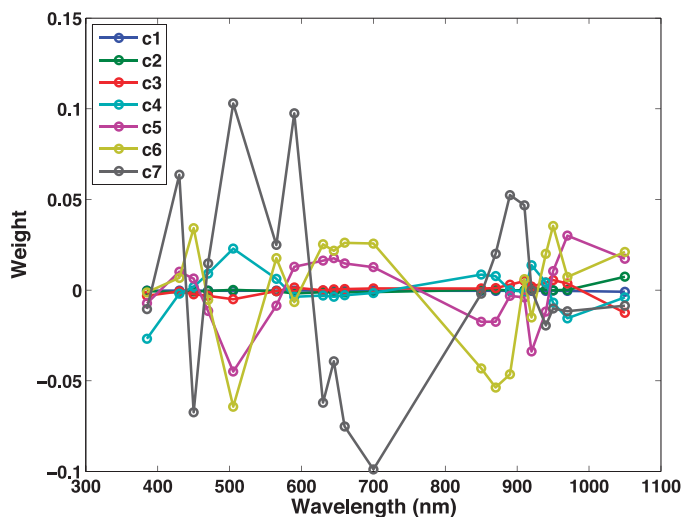


FIG. 12. PLSR components from the prediction of synthetic astaxanthin concentration level on the validation set, using images of CON pellets. Here showing the first seven of the 13 total PLSR components.

oxidizes with time. There might be a difference in the astaxanthin coating spectra characteristics in comparison with a real-time production situation. The results presented here indicate that there are still detectable amounts of astaxanthin after storage.

The results are satisfactory for screening the astaxanthin concentration level, but they are not good enough for quality

TABLE V. PLS regression for synthetic astaxanthin concentration level prediction. For CON and CON31 using synthetic astaxanthin with the concentration levels: 0, 20, 40, 60, 80 ppm, and for MG and SPC the levels are 0, 20, 60 ppm.

Recipe	Comp.	$R^2$ calibr.	$R^2$ pred.	SE calibr.	SE pred.	RPD
CON	13	0.9119	0.9151	8.4469	8.2519	3.4346
CON31	13	0.9148	0.9117	8.2980	8.4137	3.3675
MG	14	0.9330	0.9239	6.5283	6.9120	3.6283
SPC	10	0.9340	0.8959	6.7168	8.3997	3.1035



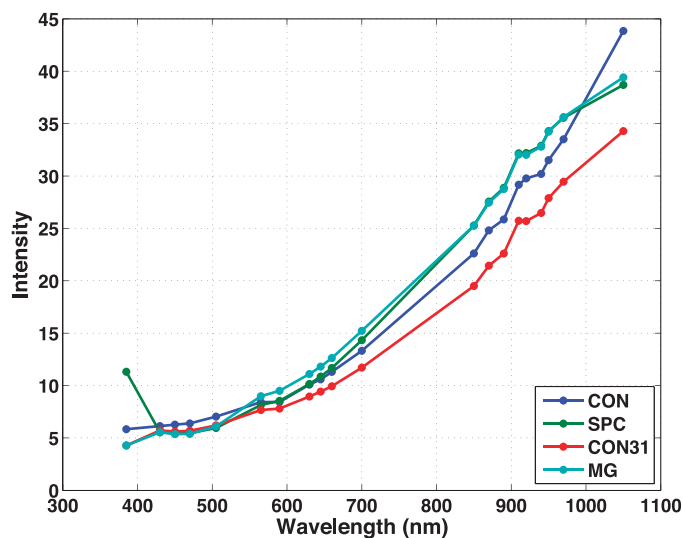


FIG. 13. VideometerLab mean spectra of pellets from each recipe, using images of pellets with coating using synthetic astaxanthin level of 20 ppm.

control. The high prediction error (SEP) in most tests is also indicated by the RPD values (see Table V). The concentration level predictions have an error in the approximate range of 7–9 ppm, which is considered high for the industry. It would be of interest to see whether more spectral bands in the critical regions produce better results. Since the visual inspection system is intended to be at-line, there would be time to capture more wavelengths if necessary. Also it would be of interest to investigate the astaxanthin coating in the production, without having been stored for a time, which likely would improve the inspection results.

It seems as though the best regression results for different recipes are achieved when calibrating and validating on the same pellet recipe or when calibrating on all recipes simultaneously (see Tables V and VI). Since this procedure could also be valuable when using the same recipe but with varying ingredient batches, the raw material ingredients of any one recipe may well vary depending on availability and price.

## CONCLUSIONS

The results show that it is possible to discriminate between synthetic astaxanthin coating and natural astaxanthin coating with 98% accuracy using spectral imaging. Similar results were achieved using SLDA with only ten spectral bands and using LDA with all 20 spectral bands.

The most frequent spectral bands selected using SLDA for classification are in the visual range of the spectrum, with some bands in the NIR regime. Corresponding spectral weights can

be seen in the PLS regression components for synthetic astaxanthin concentration prediction.

Moreover, we have shown that it is possible to predict the synthetic astaxanthin concentration in the coating using multispectral image analysis. The best regression results for different recipes are achieved when calibrating and validating on the same pellet recipe or when calibrating on all recipes simultaneously. So in order to make a prediction model that is robust for the underlying pellet compound, all relevant recipes should be included in the calibration step. In this way it is possible to predict the astaxanthin coating concentration level irrespective of the composition of the pellet recipe. The results are adequate for screening the synthetic astaxanthin coating concentration level.

## ACKNOWLEDGMENTS

The work presented has received funding from BioMar A/S and the EU under the Seventh Framework Programme FP7/2007-2013 under grant agreement number 214505.10.

1. A.P. Simopoulos. "Omega-3 Fatty Acids in Health and Disease and in Growth and Development". *Am. J. Clin. Nutr.* 1991. 54: 438-463.
2. O.J. Torrisen, R.W. Hardy, K.D. Shearer. "Pigmentation of Salmonids—Carotenoid Deposition and Metabolism.". *Rev. Aquat. Sci.* 1989. 1(2): 209-225.
3. R.T.M. Baker, A.-M. Pfeiffer, F.-J. Schöner, L. Smith-Lemmon. "Pigmenting Efficacy of Astaxanthin and Canthaxanthin in Fresh-Water Reared Atlantic Salmon, *Salmo salar*". *Anim. Feed Sci. Technol.* 2002. 99(1-4): 97-106.
4. M.G. Ljungqvist, B.K. Ersbøll, M.E. Nielsen, S. Frosch. "Multispectral Image Analysis for Astaxanthin Coating Classification". *J. Imaging Sci. Technol.* 2012. 56(2): 020403.
5. F. Zhu, S. Cheng, D. Wu, Y. He. "Rapid Discrimination of Fish Feeds Brands Based on Visible and Short-Wave Near-Infrared Spectroscopy". *Food Bioprocess Tech.* 2011. 4(4): 597-602.
6. B. Park, Y.R. Chen, M. Nguyen. "Multi-Spectral Image Analysis Using Neural Network Algorithm for Inspection of Poultry Carcasses". *J. Agric. Eng. Res.* 1998. 69(4): 351-363.
7. J.P. Wold, F. Westad, K. Heia. "Detection of Parasites in Cod Fillets by Using Simca Classification in Multispectral Images in the Visible and NIR Region". *Appl. Spectrosc.* 2001. 55(8): 1025-1034.
8. T. Brosnan, D.-W. Sun. "Improving Quality Inspection of Food Products by Computer Vision—A Review". *J. Food Eng.* 2004. 61(1): 3-16.
9. K.C. Lawrence, W.R. Windham, B. Park, D.P. Smith, G.H. Poole. "Comparison Between Visible/NIR Spectroscopy and Hyperspectral Imaging for Detecting Surface Contaminants on Poultry Carcasses". *Proc. SPIE* 2004. 5271(1): 35-42.
10. L.H. Clemmensen, B.K. Ersbøll. "Multispectral Recordings and Analysis of Psoriasis Lesions". MICCAI 06—Workshop on Biophotonics Imaging for Diagnostics and Treatment. Technical University of Denmark: Kgs. Lyngby, Denmark, October 6, 2006.
11. S.K. Stormo, A.H. Sivertsen, K. Heia, H. Nilsen, E. Elvevoll. "Effects of Single Wavelength Selection for Anisakid Roundworm Larvae Detection Through Multispectral Imaging". *J. Food Prot.* 2007. 70(8): 1890-1895.
12. D.D. Gomez, L.H. Clemmensen, B.K. Ersbøll, J.M. Carstensen. "Precise Acquisition and Unsupervised Segmentation of Multi-Spectral Images". *Comput. Vis. Image Underst.* 2007. 106(2-3): 183-193.
13. L.H. Clemmensen, M.E. Hansen, J.C. Frisvad, B.K. Ersbøll. "A Method

TABLE VI. PLS regression for synthetic astaxanthin concentration level prediction. For these tests the concentration levels of synthetic astaxanthin are 0, 20, 60 ppm. Calibrated SE and predicted SE can be interpreted as the error of the model, measured as the amount of synthetic astaxanthin added to the fish oil coating in ppm.

Recipe train	Recipe test	Comp.	$R^2$ calibr.	$R^2$ pred.	SE calibr.	SE pred.	RPD
CON, MG, SPC	CON31	15	0.9219	0.7685	7.0935	12.0394	2.0790
CON, MG, CON31	SPC	15	0.9042	0.7677	7.7556	12.5367	2.0754
CON, SPC, CON31	MG	12	0.9027	0.7656	7.9125	12.1267	2.0660
MG, SPC, CON31	CON	16	0.8925	0.8732	8.3271	8.8985	2.8088
All	All	14	0.9025	0.8911	7.9017	8.3344	3.0312

- for Comparison of Growth Media in Objective Identification of Penicillium Based on Multi-Spectral Imaging". *J. Microbiol. Methods* 2007. 69(2): 249-255.
14. L.H. Clemmensen, M.E. Hansen, B.K. Ersbøll. "A Comparison of Dimension Reduction Methods with Application to Multi-Spectral Images of Sand Used in Concrete". *Mach. Vis. Appl.* 2010. 21(6): 959-968.
  15. K. Kobayashi, Y. Matsui, Y. Maebuchi, T. Toyota, S. Nakauchi. "Near Infrared Spectroscopy and Hyperspectral Imaging for Prediction and Visualisation of Fat and Fatty Acid Content in Intact Raw Beef Cuts". *J. Near Infrared Spectrosc.* 2010. 18(5): 301-315.
  16. B.S. Dissing, M.E. Nielsen, B.K. Ersbøll, S. Frosch. "Multispectral Imaging for Determination of Astaxanthin Concentration in Salmonids". *PLoS One*. 2011. 6(5): e19032.
  17. D. Unay, B. Gosselin, O. Kleynen, V. Leemans, M.-F. Destain, O. Debeir. "Automatic Grading of Bi-Colored Apples by Multispectral Machine Vision". *Comput. Electron. Agric.* 2011. 75(1): 204-212.
  18. M.E. Nielsen, H. Mikkelsen, L.B. Nielsen, O. Joensen. "By-Product Based Production of Natural Astaxanthin (NAX)". 7th Joint meeting: 50th Annual Atlantic Fisheries Technology Conference and 29th Annual Seafood Science and Technology Society of the Americas. Norfolk, VA; November 6-9, 2005.
  19. J.-M. Carstensen, J. Folm-Hansen. "An Apparatus and a Method of Recording an Image of an Object". Patent EP1051660. Filed 1999. Issued 2003.
  20. J. Folm-Hansen. *On Chromatic and Geometrical Calibration*. [Ph.D. Thesis]. Kgs. Lyngby, Denmark: Technical University of Denmark, 1999.
  21. J. Serra. *Image Analysis and Mathematical Morphology*. London, UK: Academic Press, 1982. Pp. 50.
  22. B.S. Dissing, J.M. Carstensen, R. Larsen. "Multispectral Colormapping Using Penalized Least Square Regression". *J. Imaging Sci. Technol.* 2010. 54(3): 0304011-0304016.
  23. B.K. Ersbøll, K. Conradsen. *An Introduction to Statistics*, vol. 2. Kgs. Lyngby, Denmark: DTU Informatics, 2007. 7th ed. Pp. 263, 198-200, 241-252.
  24. H. Hotelling. "Relations Between Two Sets of Variates". *Biometrika*. 1936. 28(3/4): 321-377.
  25. L.H. Clemmensen, T. Hastie, D. Witten, B.K. Ersbøll. "Sparse Discriminant Analysis". *Technometrics*. 2011. 53(4): 406-413.
  26. H. Zou, T. Hastie. "Regularization and Variable Selection via the Elastic Net". *J. Roy. Stat. Soc. B*. 2005. 67: 301-320.
  27. T. Hastie, R. Tibshirani, J. Friedman. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. New York: Springer, 2009. 2nd ed. Pp. 80-82, 242, 547.
  28. M. Sjöström, S. Wold, W. Lindberg, J.-Å. Persson, H. Martens. "A Multivariate Calibration Problem in Analytical Chemistry Solved by Partial Least-Squares Models in Latent Variables". *Anal. Chim. Acta*. 1983. 150: 61-70.
  29. P.C. Williams, D.C. Sobering. "Comparison of Commercial Near Infrared Transmittance and Reflectance Instruments for Analysis of Whole Grains and Seeds". *J. Near Infrared Spectrosc.* 1993. 1(1): 25-32.
  30. M. Buchwald, W.P. Jencks. "Optical Properties of Astaxanthin Solutions and Aggregates". *Biochemistry*. 1968. 7(2): 834-843.
  31. J.-P. Yuan, F. Chen. "Identification of Astaxanthin Isomers in *Haemato-coccus lacustris* by HPLC-Photodiode Array Detection". *Biotechnol. Tech.* 1997. 11(7): 455-459.
  32. S. Amarie, U. Förster, N. Gildenhoff, A. Dreuw, J. Wachtveitl. "Excited State Dynamics of the Astaxanthin Radical Cation". *Chem. Phys.* 2010. 373(1-2): 8-14.