FOODINTEGRITY

Ensuring the Integrity of the European food chain

613688: Collaborative Project

Seventh Framework Programme
KBBE.2013.2.4-01: Assuring quality and authenticity in the food chain

Deliverable: 21.4

Demo results
WP19

Author(s): Dolores Pérez-Marín[1], Tom Fearn[2], Ana Garrido-Varo[1], Cecilia Riccioli[1]

Date of preparation: August 2018

Period covered: Months 27-55

Status: version 1

Dissemination level		
PU	Public	x
PP	Restricted to other participants	
RE	Restricted to a group specified by the consortium	
CO	Confidential, only members of the consortium	

The project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement No. 613688.

[1] Department of Animal Production, Non-destructive Sensor Unit, Faculty of Agricultural and Forestry Engineering, University of Córdoba, Campus Rabanales, N-IV, km 396, Córdoba 14014, Spain.

[2] Department of Statistical Science, University College London, Gower Street, London WC1E 6BT, UK.

TABLE OF CONTENTS

1. Introduction 3
2. Models to be validated 3
3. Results of the external validation 4
   3.1. Samples 4
   3.2. Quantitative calibration predicting 2018 samples 6
   3.3. Qualitative calibration predicting 2018 samples 10
4. Conclusions 12
5. Bibliography 13

Deliverable 21.4 Report of the demonstration study results WP19

1. Introduction

The advantages of Near Infrared Spectroscopy (NIRS) have resulted in more than 50% of the spectroscopic techniques currently used for the control of processes in the industry being based on the near infrared spectral region (SMA 2007). Technological advances in NIRS instrumentation have allowed the evaluation of on-site quality control applications of Iberian pork products as demonstrated in previous studies (Zamora-Rojas et al. 2012).

Work Package 19, framed within the Topic "Rapid, on-site, cost-effective methods for feed/food fraud detection", aims at evaluating a portable miniaturized NIRS spectrometer (MicroNIR Onsite Lite, Viavi Solutions Inc.) for on-site quantitative (fatty acid profile) and qualitative ("Premium" and "Non-premium" categories) of individual Iberian pork carcasses at the slaughterhouse. A system of "voluntary labelling" based on Near Infrared Spectroscopy (NIRS) in combination with information and communication technologies (ICTs) to be used in the Iberian Pig industry sector is also envisaged.

Work Package 21, on the other side, aims at both demonstrating the accuracy of the software in correctly identifying the fish families (WP15) and at deploying a portable miniaturized NIRS spectrometer for on–site application to Iberian pork carcasses at the slaughterhouse (WP19).

This deliverable reports on the activities carried out as part of task 21.2 that include the demonstration of assays in the plant with the final solution and the evaluation on-site of the advantages and limitations. The main goal of this deliverable is to describe the prediction accuracy of the best models developed in WP19 through a process of external validation, using new recorded samples in plant, not used for models development.

2. Models to be validated

Quantitative models were developed using PLS, Bayesian approach and LOCAL regression methods and were evaluated through the standard errors of cross validation (SECV) or standard errors of prediction (SEP) ranging from 0.83 to 0.84 for palmitic acid (C16:0), 0.94 to 0.99 for stearic acid (C18:0), 1.47 to 1.56 for oleic acid (C18:1) and 0.53 to 0.58 for linoleic acid (C18:2).

For classification purposes, we used the spectral data to make a direct classification as "Premium" or "Non-premium" classes, without going via a quantitative prediction of the fatty acids. Accepting from the outset that there will be samples for which the classification is uncertain, it seems important to be able to quantify that uncertainty. For this reason, the initial focus is on approaches whose output has the form of probabilities of class membership. There are many different algorithms that can be used to build such classification rules, too many to try them all on any one application. One major consideration in choosing an algorithm is that it is desirable that it should give probabilities for each category, not just a yes/no assignment. The appropriate action in the case of a report on a supposed

premium carcass that "This is Non-premium category with probability 0.55" is likely to be different from that following a report that "This sample is non-premium with probability 0.99". The focus has thus been on methods that return probabilities, and on the examination of those probabilities to see how well-calibrated they appear to be. Linear discriminant analysis, Quadratic discriminant analysis, a nonparametric approach and other classifiers (Logistic regression, Support Vector Machine, Ensemble subspace discriminant and Ensemble bagged trees) were evaluated.

The three Bayesian methods evaluated provided acceptable results in terms of classification success (up to 98% of sample correctly classified, with 230 calibration samples with probabilities of 0.9 or over), though the probabilities from QDA are clearly too extreme. Logistic regression gave similar results.

The work reported in Deliverables 19.4 and 19.5 clearly demonstrates the feasibility of using the MicroNIR device for on-site classification of carcasses. However, as NIR calibrations are dynamic and the variabililty of agrofood products, in this case Iberian pig, is very high and changeable each year, it would be desirable to continue expanding the models with more samples from more producers and following seasons.

Given the above, in 2018 a total amount of new 199 pig carcasses belonging to 12 different producers were analysed. These were currently being used to validate the models and subsequently was added to the current dataset.

3. Results of the external validation

3.1. Samples

Studying the distribution of the four main fatty acids (palmitic, stearic, oleic and linoleic) among the calibration and the validation datasets (Figure 1), it can be observed that, in general, samples belonging to the validation set (thus collected in 2018) tend to have higher values of unsaturated fats (oleic and linoleic acids) and lower values of saturated values (palmitic and stearic acids).

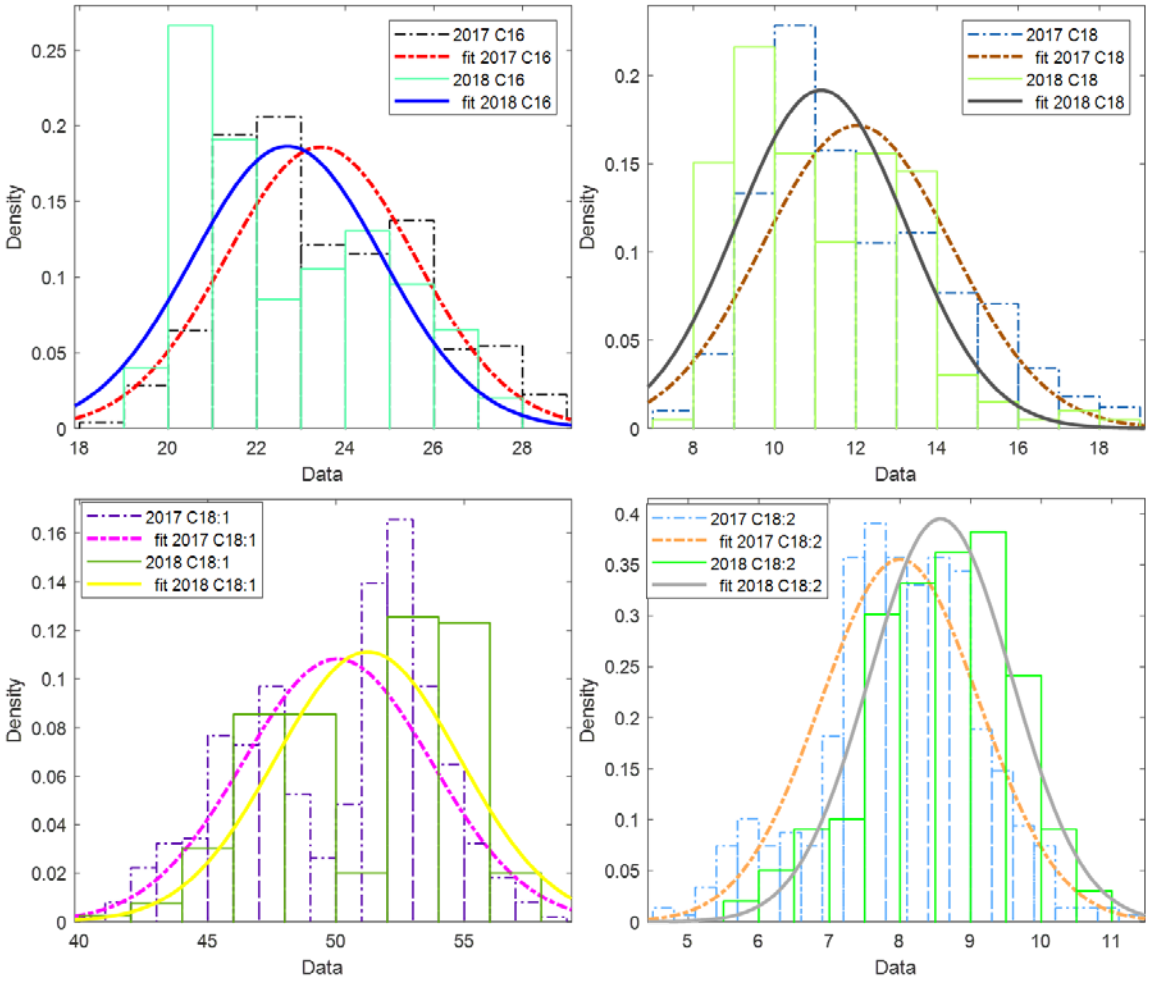

Figure 1. Histograms of the fatty acids profiles of the calibration set (2017) and the validation set (2018). C16=palmitic acid. C18=stearic acid. C18:1=oleic acid. C18:2=linoleic acid.

Figure 2 shows a similar information and highlights the difference in terms of oleic acid content between “Premium” samples collected in 2016/2017 and used to build the model (thus belonging to the calibration set) and “Premium” samples collected in 2018 and used to validate the model (thus belonging to the validation set).

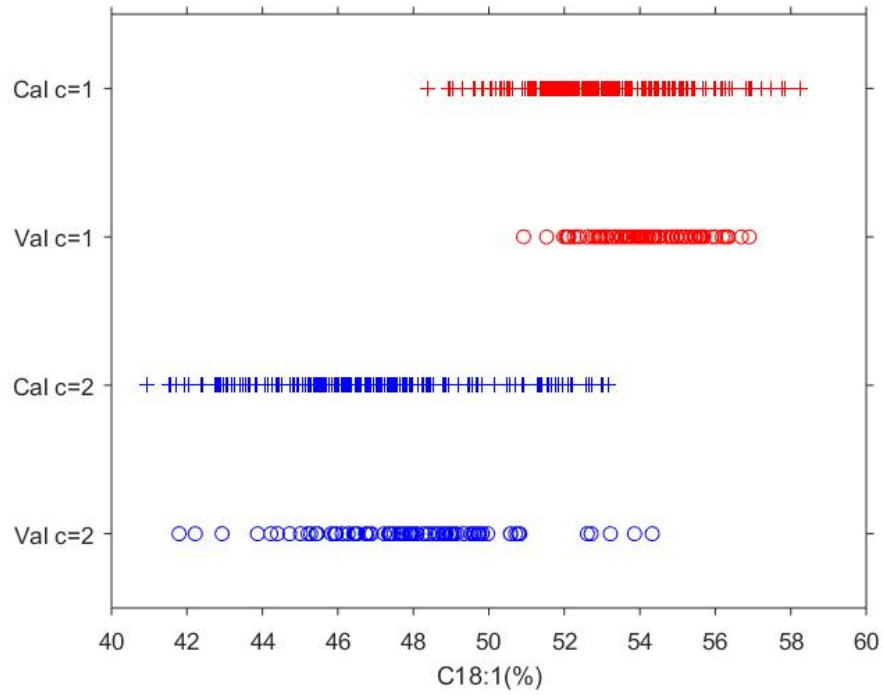

Figure 2. Oleic acid percentage in samples belonging to calibration (crosses) and validation (circles) sets. In red samples belonging to the category "Premium" and in blue sample belonging to the category "Non-Premium".

As previously stated, diets of Iberian pigs fed extensively prior to slaughter consists mostly of acorns, which is characterized by a very high oleic acid concentration (Ruiz et al. 1998). The 2018 campaign was characterized by a higher amount of acorn available for Iberian porks categorized as "Premium" and this is why the fatty acids profile changed from one year to the other. The high variability that can be expected from one year to another is key and makes the qualitative classification less reliable than the quantitative determination, unless a higher variability in terms of sampling campaigns is included in the model.

Taking this into account, quantitative and qualitative predictions were performed, and results are shown in sections 3.2 and 3.3.

3.2. Quantitative calibration predicting 2018 samples

As previously indicated, an amount of 199 samples were predicted with models. From a preliminary observation of the spectral features, eight spectra in which the last spectral point was corrupted in some way were identified and initially predicted extremely badly. For the results reported below the last spectral point (125) was replaced by a linear extrapolation from the previous two points (123 and 124).

Calibrations using the whole set (495 samples in the training set) were calculated fixing the pre-treatment and numbers of factors as for the CV results reported in Deliverable 19.3. Then we predicted the 199 new ones with these calibrations. The results show that all the four fatty acids can be predicted with an acceptable accuracy. The RMSEPs (Root Mean Square Errors of Prediction) are

showed in Table 1 and correlation between predcted and measured values are showed in Figures 3-6.

Fatty acid	RMSEP
C16:0	1.17
C18:0	0.86
C18:1	1.83
C18:2	0.58

Table 1. RMSEP for the validation set (N=199)

The plots for the prediction of eqch fatty acid are presented in Figue 3 to Figure 6.

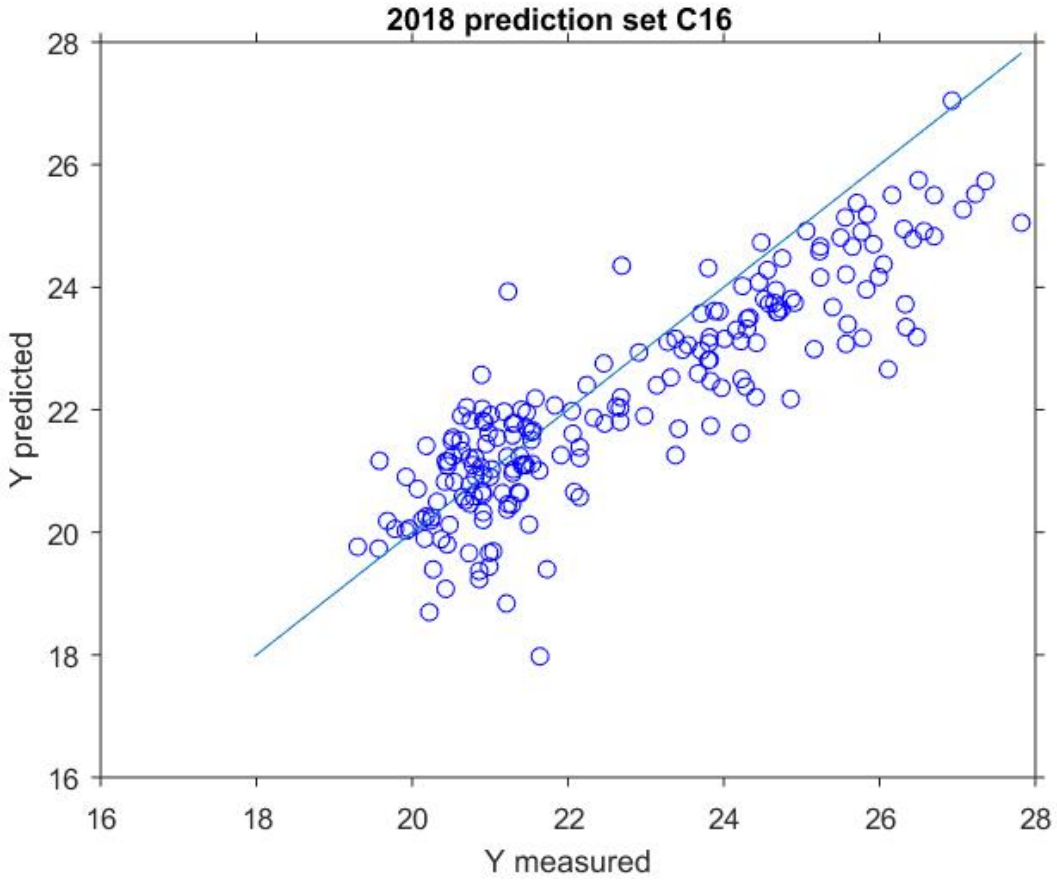

Figure 3. Palmitic acid predicted VS measured values.

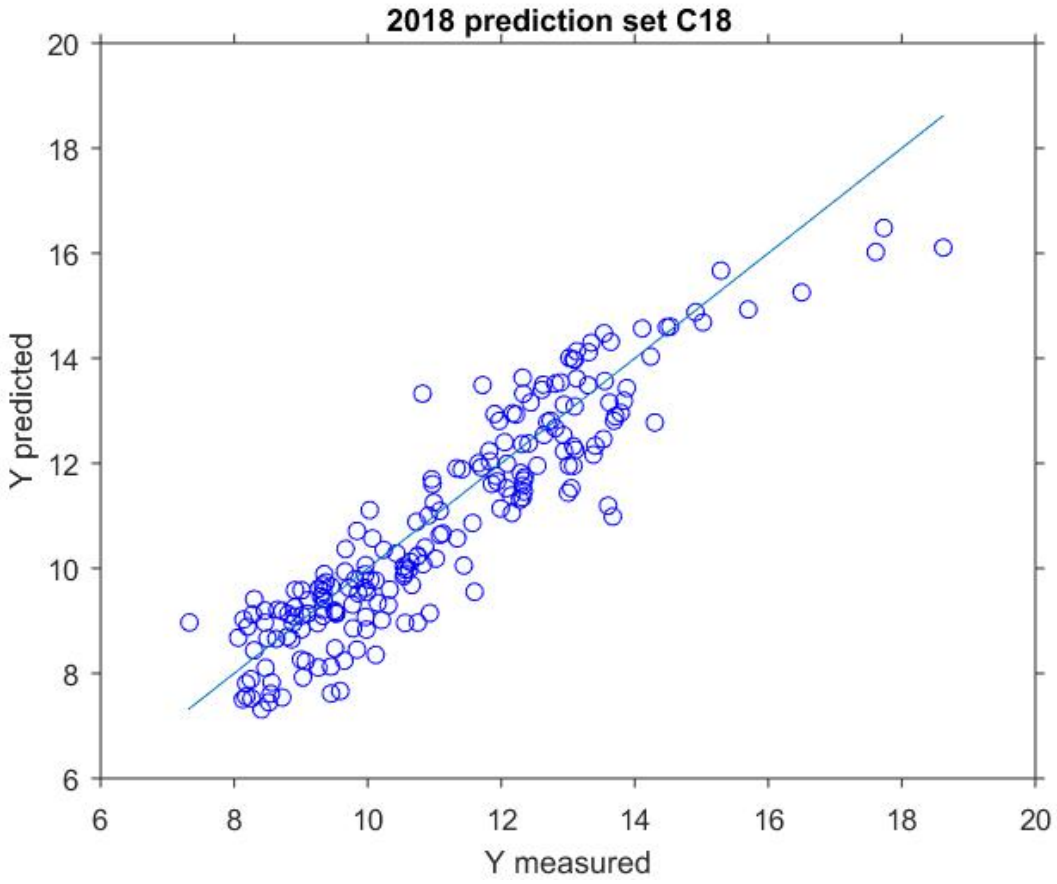

Figure 4. Stearic acid predicted VS measured values.

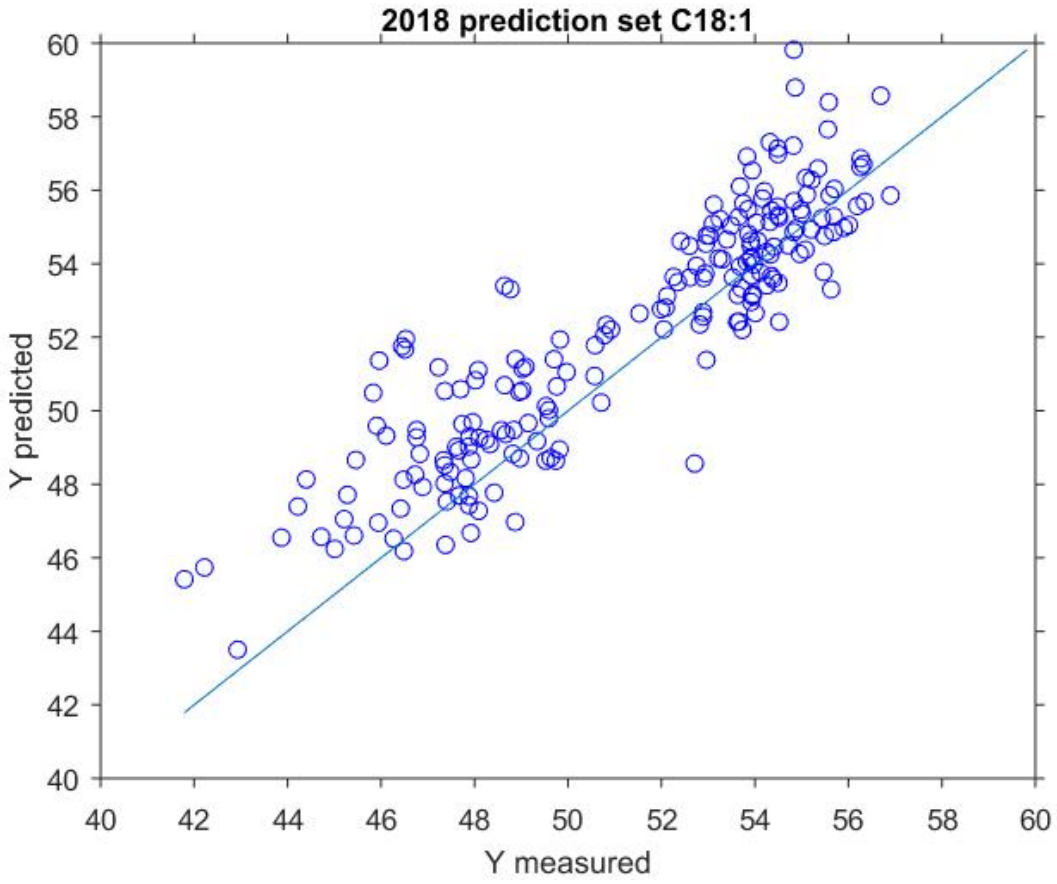

Figure 5. Oleic acid predicted VS measured values.

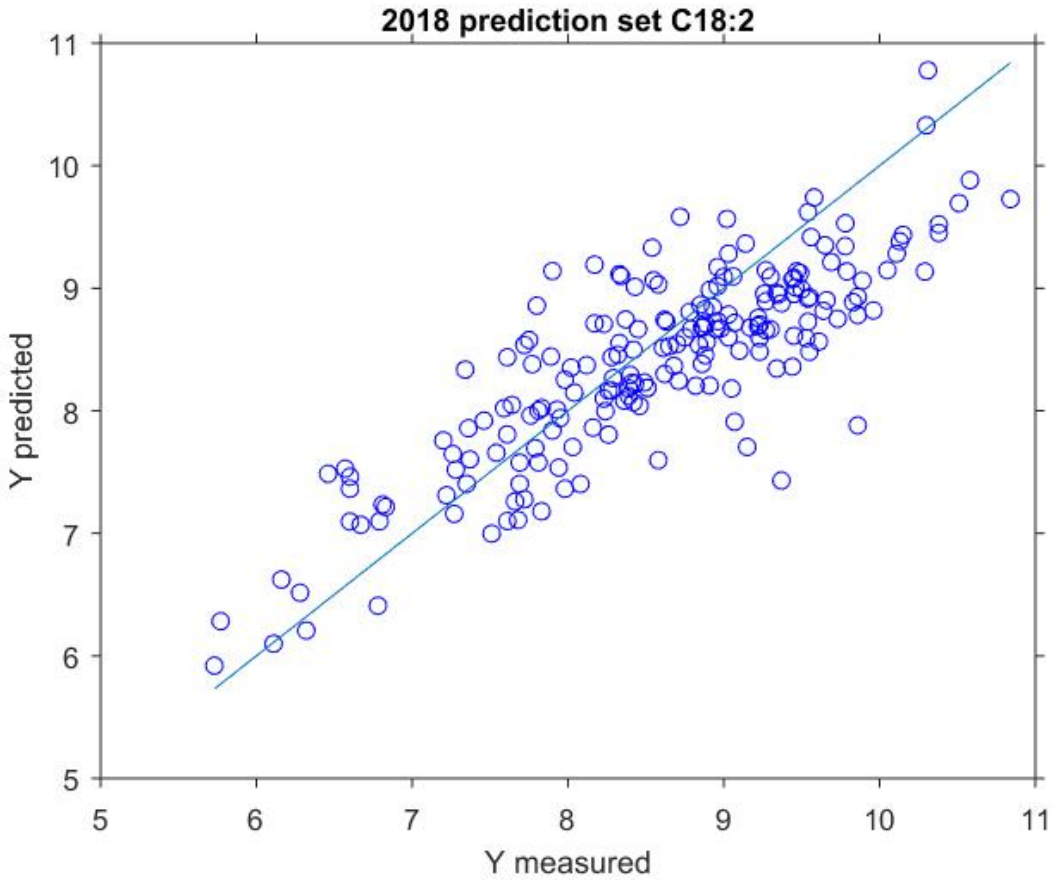

Figure 6. Linoleic acid predicted VS measured values.

In general terms, quantitative results obtained with the MicroNIR portable instrument are only a bit worse than the obtained previously using more expensive monochromator instruments at-line (Zamora-Rojas et al. 2013). Improvements can be expected by expanding calibrations with this new samples of 2018.

3.3. Qualitative calibration predicting 2018 samples

When the qualitative calibration by LDA for the classes Premium/Non-Premium developed for samples from 2016 and 2017 was used to predict 2018 samples, the results were very poor in terms of samples correctly classified. In particular, many Non-Premium samples were predicted as Premium. Given that the quantitative calibrations for fatty acids help up when used to predict 2018, it is unlikely that the problem lies with either the instrument or the spectra. The probable source of the problem is that, as previously stated, a number of samples with lower than normal levels of C18:1 were classified as “Premium” in the two earlier years, something that did not happen to the same extent in 2018.

In the light of this, the following experiment was carried out. The 2016/17 calibration set was modified by removing all the Premium samples with lab values of C18:1 less than 52.5%. This value was chosen taking into account the previous Spanish legislation (BOE 2007) on Iberian pig products

that fixed threshold values for the four fatty acids to classify iberian pig animals according to their feeding regime. This left 369 samples in the calibration set.

The calibration exercise was repeated, using leave-one-out cross validation to choose the number of principal components (PCs) for the LDA and compare six pre-treatments on these samples. The results are shown in Table 2. Plots of numbers of errors versus numbers of PCs tended to show a local minimum with around 10-15 PCs and a global one with nearer 40 PCs. Results for both are reported in the table. The first (D1) and second (D2) derivative pre-treatments both used a window of 5 spectral points. For the cross-validation the prior probabilities of class membership were taken to be the proportions in the calibration set, which is now unbalanced as a result of the deletions. For the prediction of the 2018 samples, the prior probabilities were taken as 0.5 for each class.

	Calibration set (2016/7)			Validation set (2018)	
Pre-treat	No of PCs	Errors/369	Errors (%)	Errors/199	Errors (%)
Raw	13	18	5	12	6
	38	11	3	21	11
SNV	16	21	6	25	13
	37	8	2	19	10
D1	11	22	6	19	10
	34	12	3	24	13
D1+SNV	10	20	5	18	9
	27	16	4	14	7
D2	15	20	5	11	6
	38	10	3	14	7
D2+SNV	13	19	5	12	6
	38	12	3	10	5

Table 2. Calibration and validation performance for prediction of Premium/Non-Premium class using LDA

Evaluating all of these calibrations on the 2018 samples does weaken the status of these as a validation set, but once the initial prediction failed that was already the case. The multiple evaluations do enable some general conclusions to be drawn.

- Of the pairs of calibrations, the one with the smaller number of PCs predicts better for 3 of the 6 pre-treatments. For the three where it does not (the ones with SNV) the prediction accuracies are similar for the two choices of numbers of PCs. The belief that it is worth sacrificing a little performance on the calibration in return for the robustness likely to result from the use of a much smaller number of PCs was the reason for retaining both options, and this is supported by the performance on prediction.
- The calibrations using the smaller number of PCs all give 5 or 6% classification errors by cross-validation on the training set and there is no clear winner. The results of these 6 calibrations on the 2018 set range from 6 to 13% classification errors. Had we chosen the "best" cross-validation performance (18 errors with raw spectra) it would have given 12 errors on the 2018 samples, a 6% error rate compared with 5% in training. However, this best calibration performance is almost indistinguishable from the 21 cross-validation errors given by SNV, and this makes 25 errors (13%) in the 2018 samples.

The probabilities associated with the predictions of the 13 factor LDA calibration for the 2018 samples using raw spectra are promising. They are shown in Table 3.

Probability of premium	Number of samples	Number of Premium	Proportion of Premium
0-0.1	75	2	0.03
0.1-0.5	17	3	0.17
0.5-0.9	15	11	0.73
0.9-1	92	89	0.97

Table 3. Proportions of true premium samples according to probability assigned to premium by 13 factor LDA using raw spectra, 2018 samples

The same exercise was repeated for both QDA and the nonparametric Bayes algorithms previously calibrated on the 2016/7 data. The prediction results of the existing calibrations were similar to those for LDA with large numbers of Non-Premium samples classified as Premium. However, unlike LDA, the exercise of recalibrating on a reduced training set did not solve the problem, with the numbers of errors on the 2018 samples remaining high. The obvious conclusion is that these more flexible methods are over-fitting the training samples and are not robust to the changes from year to year. However, there may be other explanations, and this is still under investigation.

4. Conclusions

After validation testing of models in 2017, the system was industry tested with samples belonging to 12 different producers and collected in 2018. The parameters that were selected for industrial trials were the commercial category and main fatty acids profile. The latter showed the most promising levels of accuracy in validation testing. The subsequent in-line industrial tests of fatty acids quantification produced a level of accuracy that is suitable for industrial use, with mean square error of prediction (RMSEP) of 1.17 for palmitic acid, 0.86 for stearic acid, 1.83 for oleic acid and 0.58 for linoleic acid. Previous validation testing in the laboratory produced a RMSECV of 0.84 for palmitic acid, 0.94 for stearic acid, 1.47 for oleic acid and 0.58 for linoleic acid. This compares well with measurements made using more expensive and non-portable laboratory instruments at UCO (RMSEP of 0.63 for palmitic acid, 0.76 for stearic acid, 1.10 for oleic acid and 0.47 for linoleic acid).

It is worth to mention that NIR models built so far are based on spectral and reference data, and these data sources changes with time due to several reasons (i.e. fat composition seasonal, weather or genetic, reference methods, sensor/spectrometer drifts by temperature and aging of electronic components, etc.). Because of all these changes, NIRS requires extensive application calibration and validation on an ongoing basis. Thus, new NIR data should be collected continuously and should be used to expand the previous calibration, fill the matrix gaps and increase robustness

To conclude, results suggest that the system holds the potential for a higher level of accuracy given further model refinement. It is envisaged that the model will be continually refined and improved with additional annual samples.

5. Bibliography

BOE. 2007. *Orden APA/3653/2007, de 13 de diciembre, por la que se publican los valores de ácidos grasos aplicables a las designaciones de alimentación «Bellota» y «Recebo», para la campaña 2007-2008.*,

Ruiz J et al. 1998. Prediction of the feeding background of Iberian pigs using the fatty acid profile of subcutaneous, muscle and hepatic fat. *Meat Sci.* **49**, 155–163.

SMA. 2007. *Life Sciences Process Spectroscopy: at –line and on-line spectroscopy in the food and pharmaceutical industries: 2002-2007* Rasmusson, Whilley, (Ed.), Strategic Directions International Inc.

Zamora-Rojas E et al. 2012. In-situ Iberian pig carcass classification using a micro-electro-mechanical system (MEMS)-based near infrared (NIR) spectrometer. *Meat Sci.* **90**, 636–642.

Zamora-Rojas E et al. 2013. Prediction of fatty acids content in pig adipose tissue by near infrared spectroscopy: at-line versus in-situ analysis. *Meat Sci.* **95**, 503–11.