

## Estimation of Chemical Composition of Quinoa (*Chenopodium quinoa* Willd.) by Near-Infrared Transmission Spectroscopy

Christian Encina-Zelada<sup>1,2,3</sup>, Vasco Cadavez<sup>1</sup>, Jorge Pereda<sup>2</sup>, Luz Gómez-Pando<sup>4</sup>, José A. Teixeira<sup>3</sup>, Ursula Gonzales-Barron<sup>1</sup>

<sup>1</sup>CIMO Mountain Research Centre, School of Agriculture, Polytechnic Institute of Braganza, Portugal. E-mail: cencina@ipb.pt

<sup>2</sup>Faculty of Food Industries, National Agricultural University La Molina, Lima, Peru

<sup>3</sup>Department of Biological Engineering, School of Engineering, University of Minho, Portugal

<sup>4</sup>Faculty of Agronomy, National Agricultural University La Molina, Lima, Peru

### Abstract

Quinoa (*Chenopodium quinoa* Willd.) is an Andean pseudo-cereal which has lately gained worldwide attention due to its high nutritional value. With views to speed up proximate analyses of quinoa grains, the aim of this study was to assess the feasibility of using Near-Infrared Transmittance (NIT) spectroscopy for rapid and accurate determinations. Peruvian-origin quinoa grains of 73 different accessions were subject to NIT spectrum analysis scanning the region 850-1050 nm. Moisture, protein, fat and ash contents were determined by approved methods. In order to assess the predictability of the NIT method, a cross-validation was performed using 10 random segments. Employing spectra pre-processed by standard normal variate, detrending or Savitzky-Golay second derivative and multiplicative scatter correction, higher proportions of explained variance were encountered for fat (0.786), protein (0.862) and ash (0.731) contents, than for moisture (0.369), when 14-16 partial least squares (PLS) components were retained. The cross-validation estimates of the root mean square error of prediction (RMSEP) were 0.707 for moisture, 0.537 for protein, 0.338 for fat and 0.246 for ash contents. Thus, while the technique exhibited an acceptable performance in predicting fat, protein and ashes, other signal pre-processing should still be tested in order to improve the accuracy of the rapid determination of moisture.

**Keywords:** NIR; grain quality; chemometrics; calibration

## 1. Introduction

Despite being a less well-known plant, there has been increasing interest in quinoa (*Chenopodium quinoa* Willd.) for the past 15 years owing to its perceived superior nutritional quality in comparison with other grains (Ferreira et al., 2015). Near infrared transmission (NIT) spectroscopy can presently provide rapid and low-cost whole kernel analysis of starch, moisture, protein, and oil percentages in cereals (Paulsen and Singh, 2004). Thus, the objective of this study was to assess the feasibility of accurately quantifying moisture, protein, fat and ash contents by NIT spectroscopy, applying partial least squares modelling.

## 2. Methodology

### 2.1 Proximal and NIT analyses

Quinoa crops (*Chenopodium quinoa* Willd.) harvested in Peru (from the National Agricultural University La Molina and Regional Development Centre – Highland), in different seasons between 2010 and 2012, were utilised. Quinoa grains samples amounted to 73 accessions which were of orange, beige and yellow colour. Moisture, protein, crude fat and ash contents were determined using the reference methods 925.10, 920.87 (conversion factor of 6.25), 923.05 (FOSS Soxhtec) and 923.03, respectively, described in AOAC (2000). Determinations, done in triplicate, were averaged and converted to dry basis.

Near-infrared transmission (NIT) spectra were acquired by placing the grains directly in an Infratec 1241 grain analyzer (Module Foss Tecator), using 60-mm quartz cuvettes, and scanning the region 850-1050 nm. The spectra were recorded at scanning step intervals of 2 nm to give 100 data points per sample. A total of 10 frequency scans were performed per sample, and carefully assessed for consistency. Frequency scans were averaged, and all raw spectral data were then mean-centered and linked to the chemical analyses data. To minimise the effect of changes in the baseline, the raw spectra were firstly pre-processed using the following filters separately: multiplicative scatter correction (MSC), first and second derivatives using the Savitzky-Golay method (SG1, SG2), standard normal variate (SNV) and detrending (DT). In addition, combinations of filters with MSC were also tested: SG1+MSC, SG2+MSC, SNV+MSC and DT+MSC.

### 2.2 Statistical analysis

The extraction of information from quinoa grain's pre-processed spectra to estimate moisture, protein, fat and ash contents was performed by Partial Least Squares (PLS) analysis. Separate PLS analyses were carried out using moisture, protein, fat and ash contents as dependent variables, and statistical inferences were computed using the jack-knife cross-validation method. For a specified number of PLS components, the cross-validation was set to randomly remove 10 samples at once (prediction set), and estimate the root mean square error of prediction (RMSEP) and the coefficient of determination ( $R^2$ ) for the plot between the values predicted from the NIT model and the chemical analyses observations. For each of the nine pre-processing filters, RMSEP and  $R^2$  statistics were obtained for a number of PLS components ranging from 10 to 20. Thus, the optimal numbers of Partial Least Squares (PLS) factors were deduced for every pre-processing filter. The entire NIT spectra analysis was conducted using the "pls" and the "prospectr" packages in R version 3.2.2 (R Development Core Team).

## 3. Results

Tables 1 and 2 compile the PLS prediction results of quinoa fat and protein, respectively (results for moisture and ash contents are not shown). For fat,  $R^2$  ranges from 0.588 to 0.815, while RMSEP ranges from 0.423 to 0.607%. For protein,  $R^2$  ranges from 0.681 to 0.897, while RMSEP ranges from 0.523 to 0.627%. In all cases, it was observed that as more PLS components were retained, higher  $R^2$  were achieved. However, as the prediction ability of the model cannot be based solely on  $R^2$  values, the RMSEP values were primarily evaluated. RMSEP was found to exhibit a different behaviour (not

shown): in most cases RMSEP steadily decreased until attaining a minimum value (at an optimal number of PLS components), at which point they increased at a faster pace as more PLS components were retained.

Filter applied	Number of components			
	12	14	16	18
SG2	0.437 (0.588)	0.443 (0.641)	0.452 (0.703)	<b>0.432</b> <b>(0.728)</b>
SG2+MSC	0.429 (0.601)	0.444 (0.669)	0.472 (0.71)	0.496 (0.732)
SNV	0.449 (0.535)	0.437 (0.666)	<b>0.438</b> <b>(0.754)</b>	0.460 (0.804)
DT	0.423 (0.635)	0.428 (0.722)	<b>0.451</b> <b>(0.786)</b>	0.502 (0.828)
DT+MSC	0.505 (0.653)	0.542 (0.721)	0.592 (0.775)	0.607 (0.815)

Table 1: RMSEP adj (%) Statistic and R<sup>2</sup> (in Brackets) Obtained from Cross-Validation for Fat Quantification Using Individual and a Combination of Pre-Processing Filters: SG2 (Savitzky-Golay Second Derivative), SG2+MSC (Multiplicative Scattering Correction), SNV (Standard Normal Variate), DT (Detrend), DT+MSC.

Subsequently, to find an optimal number of PLS components, RMSEP was kept to a minimum while R<sup>2</sup> was targeted to a reasonable value (R<sup>2</sup>>0.70). Thus, in terms of model predictability for fat content, in general SG2, SNV and DT filters led to more accurate results than the other filters (Table 1). Specifically, extracting 18 PLS components from the SG2 pre-processed spectra (RMSEP=0.432; R<sup>2</sup>=0.728) produced comparable results to extracting 16 PLS components from the SNV-processed spectra (RMSEP=0.438; R<sup>2</sup>=0.754) and the DT-processed spectra (RMSEP=0.451; R<sup>2</sup>=0.786). The DT+MSC filter, however, yielded the poorest prediction capacity among the filters shown in Table 1. Furthermore, MSC applied in combination with the other filters did not consistently enhance the prediction capacity of the models for fat determination.

Filter applied	Number of components			
	10	12	14	16
SG1+MSC	0.538 (0.695)	0.542 (0.716)	0.627 (0.754)	0.626 (0.813)
SG2	0.596 (0.693)	0.614 (0.757)	0.610 (0.777)	0.603 (0.796)
SG2+MSC	0.523 (0.737)	0.556 (0.78)	0.606 (0.819)	<b>0.598</b> <b>(0.859)</b>
DT	0.546 (0.695)	0.587 (0.7905)	<b>0.545</b> <b>(0.833)</b>	0.619 (0.875)
DT+MSC	0.547 (0.681)	0.582 (0.816)	<b>0.597</b> <b>(0.862)</b>	0.617 (0.897)

Table 2: RMSEP adj (%) Statistic and R<sup>2</sup> (in brackets) Obtained from Cross-Validation for Protein Quantification Using Individual and a Combination of Pre-Processing Filters: SG1 (Savitzky-Golay First Derivative)+MSC (Multiplicative Scattering Correction), SG2 (Savitzky-Golay Second Derivative), SG2+MSC, DT (Detrend) and DT+MSC.

In the case of protein estimation, greater predictability of the models was generally achieved by the SG2 and DT filters (Table 2). Specifically, extracting 16 PLS components from the SG2+MSC pre-processed spectra produced statistics (RMSEP=0.598;  $R^2=0.859$ ) that were comparable to the extraction of 14 components from the DT-processed spectra (RMSEP=0.545;  $R^2=0.833$ ) and the DT+MSC-processed spectra (RMSEP=0.597;  $R^2=0.862$ ). Unlike the models for fat estimations, the application of MSC in conjunction with other filters (SG2, DT) consistently enhanced the prediction capacity of the PLS models. Applying a SG2+MSC, and retaining 14 PLS components, the predictability of ash content ( $R^2=0.731$ ) was optimised but not that of moisture ( $R^2=0.369$ ).

## 4. Conclusion

The NIT technique in conjunction with PLS modelling can be effectively used for predicting protein, fat and ash contents of quinoa whole grains, although slightly better accuracy was obtained for fat content. A comparison of pre-processing filters through PLS cross-validation indicated that, in general, SNV, SG2 and DT are comparably good filters for fat estimation while DT and SG2 are better for protein estimation. Protein estimation can be enhanced by additionally applying MSC to the filtered signals. NIT technique with no waste generation, low cost, reduced time and no sample preparation can replace the laborious methods of analysis that are presently used for proximate analysis.

## Acknowledgments

Eng. Encina-Zelada acknowledges the financial aid provided by the Peruvian National Programme of Scholarships and Student Loans (PRONABEC) in the mode of PhD grants (N°183308). Also, Dr. Gonzales-Barron wishes to acknowledge the financial support provided by the Portuguese Foundation for Science and Technology (FCT) through the award of a five-year Investigator Fellowship (IF) in the mode of Development Grants (IF/00570).

## References

- AOAC. (2000). Official methods of analysis of AOAC International. In W. Horwitz, & G. W. Latimer (Eds.), AOAC International (17th ed.) Gaithersbrug, Maryland.
- Ferreira, D. S., Pallone, J. A. L., & Poppi. R. J. (2015). Direct analysis of the main chemical constituents in *Chenopodium quinoa* grain using Fourier transform near-infrared spectroscopy. *Food Control*, (48), 91-95.
- Paulsen, M. R., & Singh, M. (2004). Calibration of a Near-infrared Transmission Grain Analyser for Extractable Starchin Maize. *Biosystems Engineering*, 89(1), 79–83.