Journal of the Science of Food and Agriculture

POSTHARVEST MONITORING OF ORGANIC POTATO (CV. Anuschka) DURING HOT-AIR DRYING USING Vis/NIR HYPERSPECTRAL IMAGING

Journal:	*Journal of the Science of Food and Agriculture*
Manuscript ID	Draft
Wiley - Manuscript type:	Research Article
Date Submitted by the Author:	n/a
Complete List of Authors:	Moscetti, Roberto; Tuscia University, Department of Innovation of Biological Systems, Food and Forestry Sturm, Barbara; University of Kassel, Postharvest Technologies and Processing Group, Department of Agricultural and Biosystems Engineering Crichton, Stuart; University of Kassel, Postharvest Technologies and Processing Group, Department of Agricultural and Biosystems Engineering Amjad, W.; University of Kassel, Postharvest Technologies and Processing Group, Department of Agricultural and Biosystems Engineering Massantini, Riccardo; Tuscia University, Department for Innovation in Biological, Agro-food and Forest system
Key Words:	Solanum tuberosum L., potato slice, convective air drying, smart drying, chemometrics

POSTHARVEST MONITORING OF ORGANIC POTATO (CV. Anuschka) DURING HOT-AIR DRYING USING Vis/NIR HYPERSPECTRAL IMAGING

Moscetti R[a], Sturm B.[b,c], Crichton S.O.J.[b], Amjad W.[d], Massantini R[a*]

[a]Department for Innovation in Biological, Agro-food and Forest system, Tuscia University, Via S. Camillo de Lellis snc, 01100, Viterbo, Italy

[b]Postharvest Technologies and Processing Group, Department of Agricultural and Biosystems Engineering, University of Kassel, 37213 Witzenhausen

[c]School of Agriculture, Food and Rural Development, Newcastle University, Newcastle upon Tyne, NE1 7RU

[d]Department of Energy Systems Engineering, University of Agriculture Faisalabad, Pakistan

* Corresponding author: Tuscia University, Department for Innovation in Biological, Agro-food and Forest system, S. Camillo De Lellis snc, 01100 Viterbo, Italy. Tel.: +39 0761 357496; fax: +39 0761 357498. E-mail address: massanti@unitus.it (Massantini, R.).

INTRODUCTION

With a production capacity of 368,096 Mio.t (2013) potatoes are amongst the most important staple foods worldwide[1] and their use in processed form has increased significantly over the last three decades, thus increasing the need for monitoring of quality throughout the involved processes.[2]

One of the most common means of preserving potatoes is drying (convective and conductive). It is well documented that extended heat treatment has a detrimental impact on the quality of the resulting product. These changes can be principally divided into chemical, microbial, nutritional and physical values. A comprehensive overview of these changes is given by Sturm and Hensel (2017).

In most industrial applications, changes within the product, including the current water content are not measured. The stopping criteria are usually time based, which regularly leads to over-drying of the product due to the concerns regarding shelf life. Whilst over-drying does not necessarily have a significant negative impact on product quality, the extended duration of the process leads to unnecessarily high energy demands as well as processing time, which negatively affects the productivity of the system.[4] As shown by a number of recent publications the approach of black-box evaluation and optimisation of drying processes, in particular, leaves a significant optimisation potential for the processes untapped.[5–8] This has severe consequences for the achievable product quality as well as energy and process efficiency.[9]

Thus, there is a great need for the development of appropriate methodologies and technologies for continuous process observation and consequently process control. Smart drying is one of the newest and most promising techniques amongst emerging drying technologies. Smart drying is a multi- and inter-disciplinary sector and its recent developments embrace the following R&D areas: artificial intelligence, biomimetic, computer vision, microwave/dielectric spectroscopy, visible (Vis) and near-infrared (NIR) spectroscopy, hyper-/multispectral imaging, magnetic resonance imaging, ultrasound imaging, electrostatic sensing and control system for the drying environment. Hyperspectral imaging (HSI), which allows for the non-invasive simultaneous spatial and spectral detection of process and product characteristics over a whole range of wavelengths, has proven to be a versatile technology.[10] Huang et al. (2014), Chen et al. (2016) and Ravikanth et al. (2017) give comprehensive overviews showing the wealth of recent work conducted in this field. Burger & Gowen (2011) give an exhaustive overview of the most common chemometrics methods that are useful in dealing with issues related to the data handling of HSI 3-D matrices (or hypercubes), which is usually affected by a “curse of dimensionality”. Thus, large data volumes result in the need for further development of data reduction approaches and development of fast algorithms if HSI is intended to be used for real-time monitoring of a process.[6] Once the minimal number of wavelengths is known, the system could potentially be simplified and either LB or CAV in combination with selective LEDs or filters could be used. Amjad et al. (2017) presented an approach for water content and chromaticity determination in potatoes during drying using only spectral data from HSI and various multivariate calibration approaches. Therefore, the main objective of the present experimentation was to further investigate the feasibility of visible / near-infrared (Vis/NIR) hyperspectral imaging as computer vision technology, which can be potentially used as smart-drying technology, to monitor chemical and physicochemical changes in organic potato slices during hot-air drying through the development of prediction models of low complexity, based on the combination of raw or at least minimally pre-processed spectral domain with spatial information from HSI. In this context, models were developed with the aim of monitoring both dry bases moisture content and browning development regardless of potato slice thickness.

MATERIALS AND METHODS

Sample preparation

Potatoes (*Solanum tuberosum* L. var. Anuschka) were purchased from the Hessische Staatsdomäne Frankenhausen (Grebenstein, Germany) on the evening before each trial, where they had been stored at 8±1 °C. After transport, they were stored at room temperature overnight (14 h) and processed on the following morning. Sampling was performed by selecting sound potatoes, with uniform size and shape. Potato slices, without peel, were prepared by washing, peeling and cutting the tuber into discs (thickness of 5, 7 and 9 mm) using an electric slicing machine mod. MAS62 (Bosch, München, Germany). A circular cutting mould was used to provide slices with an exact diameter of 45 mm. Prior to drying tests, potato slices were blanched in boiling citric acid solution (0.1% w/v) using a temperature controlled water bath mod. Wnb. 22 (Memmert, Schwabach, Germany). Samples were blanched for 3 min and then immediately cooled for 3 min in cold water. Finally, free water on the surface of slices was removed using a clean cloth.

Drying experiments

Drying experiments were performed at 50°C. Batch sampling was performed every 30 min until the average slice moisture content reached 12%. Each batch was subjected to weight measurement, hyperspectral scan and CIELab color analysis. Four replications were performed for each assessment.

Hyperspectral imaging system

Image acquisition was performed using a HSI system consisting of a Visible/Near Infrared (Vis/NIR) camera, an illumination source, a linear translation stage and a control system. An ImSpector V10E Vis/NIR camera (500÷1010 nm sensitivity, ~1.5 nm resolution) (Specim Spectral Imaging Ltd., Finland) was used, equipped with a 35-mm C-mount zoom lens mod. Xenoplan 1.9/35 (Schneider Optische Werke GmbH, Germany). The distance between camera lens and sample was set at 27 cm. The illumination source consisted of three 60-W halogen spots. Light spots were set at 45°. A desktop computer with the Spectral Imaging software v3.63.201 28R (Specim, Oulu, Finland) was used to control the camera. The linear translation stage speed was set to 8 mm s^{-1}, and images were consequently captured by the camera at intervals of 1.5 mm.

Hypercube acquisition

Variations in responsiveness of the camera, also known as 'pattern noise', were corrected by performing a reflectance calibration to account for the background spectral response and the 'dark' current of the camera. A white reference tile of 200×24 mm (H × W), which corresponds to a spatial resolution of 1700×1392 pixels (H×W), was used to collect the background spectral response by recording the spectral and pixel variations of the system's response. Moreover, the internal camera noise caused by the 'dark' current was acquired by covering the camera lens with a non-reflective opaque black cap. Reference and 'dark' images were acquired for each scan.

A binary mask was used to remove the background and the edges in each HSI image. The resulting Region Of Interest (ROI) was used to measure the mean reflectance spectrum of each potato slice.

Spectra pre-processing

During experimentation, spectra were pre-processed following a variety of spectral pre-treatments including the Standard Normal Variate (SNV), Multiplicative Scatter Correction (MSC), Savitzky-Golay first, second and third derivatives (D1*f*, D2*f* and D3*f*, respectively) with a second order polynomial fitted over a window of five, seven or nine variables and Mean Centering (MC).[15] Every possible combination of pre-processes was also tested (Supplementary Fig. 1).

Regression model development

Regression models were computed using the partial least squares (PLS) regression through the SIMPLS algorithm. In addition, the Interval PLS (iPLS) algorithm was also used to select a subset of wavelengths which could describe superior predictions compared to PLS models based on an all features dataset. The iPLS algorithm was configured in stepwise forward mode for the selection of a maximum of 10 wavelengths. In addition, features for use in PLS regression were extracted from the spectra as differences of raw reflectance values for each possible pair of wavelengths ($R[\lambda_1]$-$R[\lambda_2]$), and ratios between raw reflectance values for each possible pair of wavelengths ($R[\lambda_1]:R[\lambda_2]$). After this, difference and ratio values were mean centered. Furthermore, the PLS regression models were also computed by testing the combination of spatial and spectral information.

Model were individually computed for each sample thickness and for the global dataset of all samples (i.e. 5-, 7- and 9-mm thicknesses), in order to obtain models robust to the variance in slice thickness. In addition, with the aim of finding the optimal trade-off between under-fitting and over-fitting problems, it was essential to test each model by splitting the dataset as follows: 75% and 25% of the samples were assigned to the calibration set (C) and the prediction set (P), respectively. Each model was optimized by computing a venetian blinds cross-validation (CV) with 10 data splits.

Root Mean Square Error (RMSE) for calibration, cross-validation and prediction calculations were employed to evaluate each regression model with the purpose of circumventing unrealistic results.[16] Model performances were also evaluated using the adjusted coefficient of determination (adj-R^2).

To evaluate model robustness, Bias Control Limit (BCL) and Unexpected error Control Limit (UCL) were computed according to Eqs. 1 and 2, respectively:

(1) $$BCL = \frac{RMSEP}{RMSEC}$$

(2) $$UCL = \frac{BIASP}{RMSEC}$$

Model was assumed as not to be robust when BCL > 0.60 or UCL > 1.30.[17]

To evaluate statistical differences between models, the resulting variances (squared RMSEP) were compared using the Fisher's *F*-test.[18] Thus, the *F* value was computed according to Eq. 3:

(3) $$F\ value = \frac{RMSEP_2^2}{RMSEP_1^2}$$, where $RMSEP_2^2 > RMSEP_1^2$

Then, the *F*-value was compared with the *F* critical ($1 - \alpha$, df_1, df_2), which was obtained through the quantile function for the *F* distribution, where α corresponds to the test significance level (i.e. $\alpha = 0.05$) and both df_1 and df_2 the degrees of freedom of the compared models. The null hypothesis (H_0: $\mu_1 = \mu_2$) was rejected (i.e. models were assumed as different) when *F* value > *F* critical.

Reference measurements

Colour of potato slices was measured with a chroma meter (CR-410, Konica Minolta, Osaka, Japan). Four replications were carried out for each sample by performing four colour measurements on the top of each potato slice. The results were expressed according to the CIELab colour space and thus in terms of luminance (L^*), redness (a^*), yellowness (b^*), hue angle (h), chroma (C^*)[19] and luminance/yellowness ratio (L^*b^{*-1} ratio).

Moisture content was determined in 30 min intervals throughout the drying process by weighing the sample and after drying it was assessed on dry basis by the oven-drying method at 105°C for 24 h.

Potato surface area was extracted from each unprocessed HSI image as number of pixels. The relative area shrinkage was calculated according to Eq. 4:

$$S_b = \frac{S_t}{S_0} \quad (4)$$

where S_b corresponds to the 'relative area shrinkage', S_t represents the 'surface area' in pixels at the drying time t, and S_0 corresponds to the 'surface area' in pixels of the fresh sample.

Data handling, chemometrics and statistical analysis

A two-factor analysis of variance (ANOVA) was performed to evaluate the main effect of the thickness of potato slice and drying time, and the interaction between these factors. The Tukey's pairwise comparison method was performed, and the Honestly Significant Difference (HSD) was calculated for an appropriate level of interaction ($P \leq 0.05$).

Relationships between dry basis moisture content, luminance (L^*), redness (a^*), yellowness (b^*), hue angle (h), chroma (C^*) and L^*b^{*-1} ratio measurements were quantified by the coefficient of determination of linear, quadratic and exponential functions.

Matlab software R2015b coupled with 'Image Processing' toolbox was used to acquire the relative area shrinkage of each slice, while the PLS_Toolbox software v8.1 (Eigenvector Research Inc., WA, USA) was used for the spectral pre-processing, PLS model building and features selection. Data handling and other statistical analyses were performed using R v3.3.3 software in combination with 'dplyr' v0.5.0 and 'agricolae' v1.2-4 R-packages.

RESULTS AND DISCUSSION

Chemical and physicochemical changes during drying

Fig. 1 shows changes in the colour parameters (Fig. 1a-1f), relative area shrinkage (Fig. 1g) and dry basis moisture content (Fig. 1h) of potato slices dried at 50°C up to a final relative moisture content of 12%. Results from the two-way ANOVA showed that changes in chemical and physicochemical properties over drying time depend on slice thickness. Moreover, the analysis of data confirmed that drying time is proportional to the squared thickness of potato slice (adj-R^2 =

0.995) and consequently that thinner samples dried faster, as expected and already described by the 'proportional of thickness law' [20]. Thus, it is plain from Fig. 1h that the drying rate of the 5-mm potato slice was much faster than both 7- and 9-mm samples, as well as area shrinkage and colour development of product. In fact, it is evident from Fig. 1a, 1c and 1e that luminance (L^*), yellowness (b^*) and chroma (C^*) of 5-mm samples, respectively, had ascending trends in both pre-heating and first falling-rate periods and then broad descending trends toward the end of the drying process (i.e. second falling-rate period). The data also demonstrated that the rate of changes in redness (a^*) (Fig. 1b), hue angle (h) (Fig. 1d) and luminance/yellowness ratio (L^*b^{*-1} ratio) (Fig. 1f) as well as area shrinkage (S_b) (Fig. 1g) were significantly faster in the 5-mm slice than both 7- and 9-mm samples. This rapid colour degradation observed for the 5-mm sample should be mainly affected by non-enzymatic browning reactions (i.e. Maillard reaction, ascorbic acid oxidation and/or heat damage), as a consequence of surface overheating when drying entered the heating-up (or second falling-rate) period.[21] In fact, in the third drying period, moisture reduction slowed and then product temperature significantly increased.[22] Moreover, because of the existing positive relationship between sample thickness and drying time, a higher degree of non-enzymatic browning was tendentially observed in potato slices which required a longer second falling-rate period to reach the equilibrium moisture content. This means that thick potato slices (i.e. 9-mm sample) tended to be browner than thin potato slices (i.e. 5- and 7-mm samples).

The relative area shrinkage (S_b) on potato slices (Fig. 1g) changed in accordance to what was already observed by various Authors.[22–25] They demonstrated that, as drying proceeded and cellular rigidity increased due to moisture loss, potato slices tend to shrink faster, but also to bend upwards and thus to attain an irregular shape. The severity of the phenomenon is reduced at higher drying temperatures, which may induce an intense moisture gradient and, thus, to a rubber-glass transition of the slice surface, also known as "case hardening effect". In detail, it has largely been demonstrated that the "case hardening effect" fixes the shape of the potato slice, which consequently shrinks rather uniformly until the end of drying. Results from our experiment showed

201 the development of an irregular shape for the lowest thickness. In fact, the 5-mm potato slices
202 largely shrank because of reduced internal tissue stresses due to a comparably regular distribution of
203 moisture between the centre and the surface of the slice. Consequently, the shape of the 5-mm
204 samples was fixed towards the final stage of drying after bending upwards. This was particularly
205 evident by evaluating the changes in standard deviation of the relative area shrinkage (data not
206 shown), which increased by approximately a 5-fold factor starting from the third-last sampling time
207 (i.e. 240 min). At the highest thicknesses (i.e. 7 and 9 mm), the shape of samples was probably
208 fixed at an early stage of drying due to both the higher mechanical integrity of slice and the "case
209 hardening effect", which were responsible for smaller degree of shrinkage and no bending upwards.
210 In fact, unlike the behaviour of the 5-mm potato slices, the changes in standard deviation of the
211 average relative area shrinkage of both 7- and 9-mm samples did not show downward or upward
212 trends when the end of drying was approaching.

213 In addition, data were subjected to regression analysis to attempt to model possible
214 relationships between the dry basis moisture content against other variables by selecting the best-
215 fitting linear/quadratic/exponential equations. Selected models showed from good (≥ 0.85) to
216 excellent (≥ 0.95) adjusted coefficients of determination (adj-R^2) in describing the relationships
217 between the dry basis moisture content against the relative area shrinkage, hue angle and L^*b^{*-1}
218 ratio.

219 The straight-line regression model was observed as best and simplest equation to describe
220 changes in the moisture content as a function of the relative area shrinkage for all thicknesses (Fig.
221 2a, 2d and 2g). However, a relative area shrinkage lower than ~0.65 was always paired with a
222 decrease in model linearity because the slice started to bend upwards and thus to acquire an
223 irregular shape. This trend was particularly evident only for potato slices of 5-mm thickness,
224 making the prediction of moisture content over a drying time of 240 min impossible.

225 Results from the regression analysis between the dry basis moisture content and the hue
226 angle, which describe the browning development, are shown in Figs. 2b, 2e and 2h. The type of

relationship was affected by slice thickness. In fact, the functional form, which was selected on the basis of the adj-R^2, changed from linear to exponential as thickness increases.

The dry basis moisture content exhibited a non-linear variation with respect to the L^*b^{*-1} ratio, as manifested in the form of a quadratic function (Fig. 2c, 2f and 2i). In general, the L^*b^{*-1} ratio tended to decrease during drying, and as a result the dried potato slice was darker than before drying.[26] By observing the steepness of each quadratic function, it is clear that the tendency for darkening was lower at higher thicknesses. However, it should also be noted that higher thicknesses corresponded to a lower proportion of variance in the dependent variable (i.e. dry basis moisture content) that is predictable from the independent variable (i.e. L^*b^{*-1} ratio), and thus to lower adj-R^2. On the basis of our results, any assumption on browning development may be affected by the fact that the quadratic function described the relationship between predictor and response variable less well as the thickness increased.

Finally, based on the Authors' best knowledge, potato drying has been widely addressed in literature; nevertheless, little insight is available on the effect of potato slice thickness on drying behaviour as well as on drying energy and exergy efficiencies, which seem to be affected by slice thickness.[27] Thus, in our opinion, the impact of slice thickness of potato drying deserves to be further investigated through mathematical modeling of thin-layer drying, which however was not the focus of the present work.

Regression models based on features extracted as full spectrum and iPLS-selected wavelengths

Table 1 summarizes the complete calibration, cross-validation and prediction performance metrics of both PLS and iPLS regression models, which were developed using spectral information as independent variables, categorized as "A models". In general, models showed promising results in terms of RMSEs and adj-R^2s for the prediction of the dry basis moisture content and brown/dark colour development, i.e. changes in hue angle and L^*b^{*-1} ratio. In addition, the models may fall within the definition of 'robust models' since both Bias Control Limit (BCL) and Unexpected error Control Limit (UCL) never exceeded the 0.60 and 1.30 threshold values, respectively. Thus, these control limits indicate that prediction accuracy of models may be reasonably considered moderately insensitive to unknown changes of external factors.[28]

Prediction models for moisture content (Table 1, models from #A01 to #A08) showed excellent metrics, with RMSEP and adj-R^2P values ranging from 0.11÷0.26 and 0.946÷0.990, respectively. Better results were achieved for those models individually developed for each sample thickness. Using all slice thicknesses in the same model calibration (i.e. #A07 and #A08) performance metrics deteriorated slightly, however, it still remained very good (RMSEP = 0.26; adj-R^2P ~= 0.95). Hence, both #A07 and #A08 models were able to capture variations in moisture content with both high precision and accuracy, regardless of sample thickness. Conversely, though characterized by very good or excellent results, the prediction models calibrated on colour changes did not show the same efficiency. The hue angle models (Table 1, models from #A09 to #A16) showed RMSEP and adj-R^2P values ranging between 0.88÷1.36 and 0.930÷0.961, respectively. The lowest adj-R^2Ps, which are related to the highest prediction errors, were observed in predicting browning development on the 9-mm thickness (i.e. Table 1, #A13 and #A14 models). The L^*b^{*-1} ratio models (Table 1, models from #A17 to #A23) performed worse than the hue angle models, despite showing a similar behaviour to that observed for the hue angle. The RMSEP appears to not be affected by thickness, while adj-R^2P showed a lower value when prediction was performed on datasets that include the 9-mm data (i.e. Table 1, models from #A21 to #A24). Results from both hue angle and L^*b^{*-1} ratio prediction models allow to speculate on the possibility that the lower model performances could be related to the weaker relationship observed among the moisture content and the colour indexes for the 9-mm samples. However, it can said with certainty that the inclusion of the 9-mm data into the global dataset of all samples had a negative impact on the efficiency of the models (i.e. Table 1, models #A15, #A16, #A23 and #A24).

Regarding spectral pre-treatments, since several mathematical transformations of spectra resulted in models with similar performance metrics, those that had the simplest combination of treatments at the lowest number of latent variables included in the model, were chosen. In this context, Savitzky-Golay smoothing filter (7-points window size) in combination with mean centering gave the best overall results. This suggests that regression models performed better when issues due to noise were reduced or removed, and spectral resolution enhancement was applied. The application of the other spectral pre-treatments such as derivatives, MSC and SNV always led to worse comparable results (data not shown), even though the spectra were probably affected by changes in forward scattering and backscattering by flesh tissue as a consequence of changes in its mechanical and textural properties[29] due to water loss and heat exposure during the dehydration process. Thus, further fundamental conclusions can be drawn since scatter correction was not effective in improving model performances. In fact, this may mean that variation in light scattering positively contributed in enhancing spectral differences related to changes in chemical and physicochemical parameters of potato slices during drying.

In general, PLS and iPLS algorithms produced similar models, evidenced by very closed RMSE, adj-R^2, BCL and UCL values as well as the number of latent variables. Thus, the experiment reported here demonstrates the feasibility of using feature subset selection for PLS regression models for monitoring the drying process of potato slices using a Vis/NIR hyperspectral setup. However, the effective number of wavelengths required by a iPLS model would be conditional on the spectral pre-treatment used. In fact, obtaining Savitzky-Golay spectra would require measurement of the neighbouring wavelengths to the selected features, depending on the number of points of the window size. In addition, correct application of the Savitzky-Golay filter while only using the features selected would require determination of the signal-to-noise ratio between the original and the Savitzky-Golay filtered spectra, which is related to the sensitivity of the chosen detector (i.e. Si and InGaAs or TE-InGaAs).[30] Therefore, the method presented here tried to balance the need for simplicity in terms of number of both features and pre-treatments selected with the increased accuracy of applying spectral pre-treatments. However, although very good results were obtained, further investigation will be required for successful development and transfer of each calibration model from the reference analytical instrument to a target hyperspectral setup embedded in a dryer.

Regression models based on features extracted as raw reflectance difference and ratio between each possible pair of wavelengths

Models based on raw reflectance difference and ratio between two wavelengths were categorized as 'B models' and the results are listed in Table 2. Each model was selected among a total number of $(n^2 - n) \times 2^{-1}$ = 49,141 possible models (i.e. pairs of features), which corresponded to the half-size triangular matrix with no diagonal entries (rows = columns) and where n is the total number of wavelengths (i.e. 314). Beyond model selection based on both cross-validation and prediction accuracy, each final best-fitted model was selected by visually evaluating the surface plot of the overall adj-R^2P values (Fig. 3a). Specifically, models not belonging to any surface area cluster (Fig. 3b) were likely discarded for circumventing over-optimistic results due to chance correlation. Overall, the proposed models were robust to various specifications (i.e. RMSE values, BCL and UCL), thus, indicating that outcomes could be accurately and precisely quantified with the proposed approach. In detail, the most accurate predictions were generally achieved using datasets of raw reflectance differences. Moreover, it is interesting to note that although "B models" were computed using only two wavelengths with no spectral pre-treatments, they showed similarities with the "A models". These similarities refer to the evident relationship between model performances and dataset composition (e.g. sample thickness, type of reference, etc.). Nevertheless, it should be highlighted that each "B model" was always out-performed by the corresponding "A model" in terms of both RMSE and adj-R^2 values. Thus, applying a "B model" resulted in reduced model complexity but also in a decreased prediction ability of approx. 56% on average, which nonetheless still remained acceptable.

Description of the selected spectral bands

"A models" computed using iPLS algorithms showed the most informative wavelengths at ~510, ~760- 790÷810-, ~880- and ~970-nm spectral bands (Supplementary Fig. 2a). Specifically, features around 510, 760 and 970 nm likely represent fifth (6ν), forth (5ν) and second (3ν) overtones of O–H stretching vibrations, respectively.[31] Methyl groups ($C–H_3$) exhibit the third stretching overtones (4ν) and bending (δ) combinations within both 790÷810- and ~880-nm bands.[31] Furthermore, features within neither the reddish nor greenish spectral region (i.e. 540÷660 nm) were exclusively paired to models developed for predicting changes in colour such as browning (i.e. decreases in hue angle) and yellow discoloration (i.e. decreases in L^*b^{*-1} ratio), which may be associated to losses of the total carotenoids content.[32,33]

Regarding "B models", based on our knowledge, there is no research on the monitoring of changes in both colour and moisture content in potato slices during drying using raw reflectance differences/ratios between pairs of wavelengths as correlation features. However, knowledge of the potato chemical composition allows some insight into the features found to give good regression models. Wavelengths from the reflectance range around 820÷920 nm, which likely represent the third (4ν) overtone and combinations of both C-H and $C\text{-}H_2$ stretches and deformations, were mainly selected from all datasets and frequently paired, either as a difference or a ratio (Supplementary Fig. 2b and 2c). Features computed in that region could be correlated with carbohydrate/starch content,[34] which has strong correlations with dry matter content,[35] discoloration[36] and change in internal structure of tissue due to starch gelatinization during drying.[37] Consequently, those close correlations could explain why most of "B models" mainly used the carbohydrate/starch signals. However, in the case of dataset of reflectance differences, features attributed to water and/or hydroxyl groups (i.e. ~510- and 970÷1000-nm) were also selected.

Regression models based on both spectral and spatial information

Hyperspectral imaging not only permits the measurement of chemical constituents and internal quality attributes of food, but also provides spatial distribution data (i.e. size and shape information) of product. Thus, in addition to the development of regression models based on the spectral profiles of samples, we also explored the possibilities offered by hyperspectral imaging for quantifying surface characteristics of each potato slice, measuring the relative area shrinkage (S_b)

during drying. In fact, it is fundamental to consider that changes in quality attributes of horticultural products during drying are successfully measurable on the basis of their variations in size and shape.[23–25,36,38–40]

As expected, models based on relative area shrinkage (S_b) failed in prediction of colour changes (data not shown). Conversely, the dry basis moisture content was successfully predicted through the spatial information with RMSEP = 0.31, adj-R^2P = 0.920, BCL = 0.07 and UCL = 1.13. However, an improved moisture prediction model was obtained by merging and then autoscaling the spatial data (i.e. S_b) and spectral data from the B07 model (i.e. R[511 nm] – R[994 nm]) (Supplementary Fig. 3). The resulting model ("B07+S_b") had lower RMSEP of 0.24 and higher adj-R^2P of 0.949, as well as improved BLC and UCL of 0.03 and 1.07, respectively. Consequently, although the analysis was restricted to the relative area shrinkage through pixel quantification, both spatial and spectral domains contributed toward obtaining a moisture prediction model insensitive to sample thickness. Nevertheless, the model must be further improved because of its inability to predict moisture content lower than ~0.2.

Statistical comparison of models

The upper triangular matrix represented in Fig. 4 summarizes the results from the F-test performed on models computed using the dataset of all sample thicknesses. The statistical analysis showed that "A models" had better prediction performances than "B models". Among "A models", no statistical differences were observed between model calibrated using full spectra (i.e. PLS model) and the corresponding model obtained towards features selection (i.e. iPLS model). Thus, results demonstrate that feature selection through the iPLS algorithm was highly beneficial in reducing model complexity though still maintaining the predictive ability of the model. In addition, in our experimentation it has been found that the use of raw reflectance differences/ratios as features did not produce statistically different models, except for hue angle, whereas the model based on raw reflectance difference (i.e. #B15) overwhelmed the model obtained using raw reflectance ratio (i.e. #B16).

Finally, particular attention should be given to the statistical comparison of the moisture prediction models. In fact, although performance metrics of both PLS and iPLS models (i.e. #A07 and #A08 models, respectively) were significantly superior to those of models based on the relative area shrinkage (i.e. S_b model), no statistical differences were noted when compared to the "B07+S_b" model. This noteworthy result demonstrates the feasibility of Vis/NIR hyperspectral imaging in predicting moisture content in potato slices at different dehydration phases by combining spatial and raw spectral domains in a simple way.

CONCLUSIONS

In this study the feasibility of Vis/NIR hyperspectral imaging (500÷1010 nm) as computer vision technology, which can be potentially used as smart-drying technology, to proactively and non-destructively detect and monitor chemical and physicochemical changes (i.e. moisture content, hue angle and L^*b^{*-1} ratio) in organic potato slices (*Solanum tuberosum L.* var. Anuschka) of various thicknesses (i.e. 5, 7 and 9 mm) during hot-air drying at 50°C was investigated.

The analysis of spectral features used in the best-performing models delivered valuable information for identifying the relevant parts of the spectra in monitoring the drying process of potato slices. Features for regression models comprising wavelengths that resulted in the best prediction results were generally in the ~510-, 760÷820-, 880÷920- and 970÷1000-nm spectral bands. Results suggest these are the predominant bands for detection of dry basis moisture content, which may be related to water and starch content, and colour changes due to non-enzymatic reactions (i.e. hue angle and L^*b^{*-1} ratio). Since the discoloration also exhibits features at 540÷660 nm, it is hypothesised that losses in total carotenoids content could be the underlying chemical basis for regression. Further research would be necessary for verification.

The best prediction results were obtained using a Savitzky–Golay filter with 7 smoothing points paired to mean centering, and datasets of features selected by using the forward-selection iPLS algorithm. However, both datasets of raw reflectance differences ($R[\lambda_1]$-$R[\lambda_2]$) and raw reflectance ratios ($R[\lambda_1]$:$R[\lambda_2]$) showed potential for the development of models with low complexity and thus more easily transferable to a low-cost dryer. Finally, yet importantly, the combination of spectral data (i.e. R[511 nm]-R[994 nm]) with spatial data (i.e. relative area shrinkage of potato slice) has proven to be a viable and preferable alternative to both the best PLS and iPLS models in predicting the dry basis moisture content.

Thus, the practical implication of this study is that modelling the data acquired during drying through hyperspectral imaging can provide useful information concerning the chemical and physicochemical changes of product. With all this information, the proposed approach lays the foundations for a more efficient smart dryer that can be designed and its process optimized for drying of potato slices. It can be further concluded that novel smart-dryer must be developed bearing in mind that thickness of slice crucially affects drying kinetics of product and thus the development of accurate, precise and robust prediction models. However, although the results obtained are promising, a larger validation sample must be used to address the additional possible variations expected from growing potatoes in different regions, crop years, agro-pedo-climatic conditions and degree of ripeness, in addition to other cultivars and drying conditions.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge CORE Organic Plus consortium (Coordination of European Transnational Research in Organic Food and Farming System, ERA-NET action), BLE BLE (Bundesanstalt für Landwirtschaft und Ernährung, Germany) and Mipaaf (Ministero delle politiche agricole alimentari e forestali, Italy) for financial support through the SusOrganic project titled: ‘Development of quality standards and optimized processing methods for organic produce’ (Nr: 2814OE006).

REFERENCES

1 FAO, FAO Statistical Pocketbook 2015. *Food Agric. Organ. United Nations*, (2015).

2 Rady A, Guyer D and Lu R, Evaluation of Sugar Content of Potatoes using Hyperspectral Imaging. *Food Bioprocess Technol.*, **8**:995–1010 (2015).

3 Sturm B and Hensel O, in *Handb. Dry. Veg. Veg. Prod.*, eds. Zhang M, and Mujumdar AS,

CRC Press, London, UK (submitted), pp. (2017).

4 Pask F, Lake P, Yang A, Tokos H and Sadhukhan J, Industrial oven improvement for energy reduction and enhanced process performance. *Clean Technol. Environ. Policy*, **19**:215–224 (2017).

5 Sturm B, Nunez Vega A-M and Hofacker WC, Influence of process control strategies on drying kinetics, colour and shrinkage of air dried apples. *Appl. Therm. Eng.*, **62**:455–460 (2014).

6 Kamruzzaman M, Makino Y and Oshita S, Hyperspectral imaging for real-time monitoring of water holding capacity in red meat. *LWT - Food Sci. Technol.*, **66**:685–691 (2016).

7 Nuñez Vega A-M, Sturm B and Hofacker W, Simulation of the convective drying process with automatic control of surface temperature. *J. Food Eng.*, **170**:16–23 (2016).

8 Defraeye T, Advanced computational modelling for drying processes – A review. *Appl. Energy*, **131**:323–344 (2014).

9 Sturm B, in *Intell. Control Dry.*, eds. Martynenko A, and Mujumdar AS, CRC Press, London, UK (submitted), pp. (2017).

10 Huang H, Liu L and Ngadi M, Recent Developments in Hyperspectral Imaging for Assessment of Food Quality and Safety. *Sensors*, **14**:7248–7276 (2014).

11 Chen Y-N, Sun D-W, Cheng J-H and Gao W-H, Recent Advances for Rapid Identification of Chemical Information of Muscle Foods by Hyperspectral Imaging Analysis. *Food Eng. Rev.*, **8**:336–350 (2016).

12 Ravikanth L, Jayas DS, White NDG, Fields PG and Sun D-W, Extraction of Spectral Information from Hyperspectral Data and Application of Hyperspectral Imaging for Food and Agricultural Products. *Food Bioprocess Technol.*, **10**:1–33 (2017).

13 Burger J and Gowen A, Data handling in hyperspectral image analysis. *Chemom. Intell. Lab. Syst.*, **108**:13–22 (2011).

14 Amjad W, Crichton S, Munir A, Hensel O and Sturm B, Hyperspectral imaging for the

determination of potato slice moisture content and chromaticity during the convective hot air drying process. *Biosyst. Eng.*, **(submitted**:(2017).

15 Moscetti R, Haff RP, Stella E, Contini M, Monarca D, Cecchini M and Massantini R, Feasibility of NIR spectroscopy to detect olive fruit infested by Bactrocera oleae. *Postharvest Biol. Technol.*, **99**:58–62 (2015).

16 Moscetti R, Saeys W, Keresztes JC, Goodarzi M, Cecchini M, Danilo M and Massantini R, Hazelnut Quality Sorting Using High Dynamic Range Short-Wave Infrared Hyperspectral Imaging. *Food Bioprocess Technol.*, **8**:1593–1604 (2015).

17 Shenk JS, Workman JJJ and Wsterhaus MO, in *Handb. Near-Infrared Anal.*, eds. Burns DA, and Ciurczak EW, CRC Press, London, UK, pp. 380–382 (2008).

18 Liu C, Liu W, Lu X, Chen W, Yang J and Zheng L, Potential of multispectral imaging for real-time determination of colour change and moisture distribution in carrot slices during hot air dehydration. **195**:110–116 (2016).

19 Moscetti R, Carletti L, Monarca D, Cecchini M, Stella E and Massantini R, Effect of alternative postharvest control treatments on the storability of ‘Golden Delicious’ apples. *J. Sci. Food Agric.*, **93**:2691–2697 (2013).

20 Karel M and Lund DB, Physical Principles of Food Preservation. *J. Food Process. Preserv.*, Marcel Dekk, New York, NY, USA, pp. (2003).

21 Leeratanarak N, Devahastin S and Chiewchan N, Drying kinetics and quality of potato chips undergoing different drying techniques. *J. Food Eng.*, **77**:635–643 (2006).

22 Lu L, Tang J and Ran X, Temperature and Moisture Changes During Microwave Drying of Sliced Food. *Dry. Technol.*, **17**:414–431 (1999).

23 Yadollahinia A and Jahangiri M, Shrinkage of potato slice during drying. *J. Food Eng.*, **94**:52–58 (2009).

24 Senadeera W, Bhandari BR, Young G and Wijesinghe B, Modeling Dimensional Shrinkage of Shaped Foods in Fluidized Bed Drying. *J. Food Process. Preserv.*, **29**:109–119 (2005).

25 Khraisheh M a. ., Cooper TJR and Magee TR a., Shrinkage Characteristics of Potatos Dehydrated Under Combined Microwave and Convective Air Conditions. *Dry. Technol.*, **15**:1003–1022 (1997).

26 Álvarez MD, Fernández C and Canet W, Effect of freezing/thawing conditions and long-term frozen storage on the quality of mashed potatoes. *J. Sci. Food Agric.*, **85**:2327–2340 (2005).

27 Amjad W, Hensel O, Munir A, Esper A and Sturm B, Thermodynamic analysis of drying process in a diagonal-batch dryer developed for batch uniformity using potato slices. *J. Food Eng.*, **169**:238–249 (2016).

28 Nicolaï BM, Beullens K, Bobelyn E, Peirs A, Saeys W, Theron KI and Lammertyn J, Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: A review. *Postharvest Biol. Technol.*, **46**:99–118 (2007).

29 Romano G, Nagle M, Argyropoulos D and Müller J, Laser light backscattering to monitor moisture content, soluble solid content and hardness of apple tissue during drying. *J. Food Eng.*, **104**:657–662 (2011).

30 Moscetti R, Haff RP, Aernouts B, Saeys W, Monarca D, Cecchini M and Massantini R, Feasibility of Vis / NIR spectroscopy for detection of flaws in hazelnut kernels. *J. Food Eng.*, 1–7 (2013).

31 Workman J and Weyer L, Practical Guide to Interpretive Near-Infrared Spectroscopy. CRC Press, London, UK, pp. (2008).

32 Suvarnakuta P, Devahastin S and Mujumdar AS, Drying Kinetics and ß-Carotene Degradation in Carrot Undergoing Different Drying Processes. *J. Food Sci.*, **70**:520–526 (2005).

33 Sablani SS, Drying of Fruits and Vegetables - Retention of Nutritional/Functional Quality. *Dry. Technol.*, **24**:123–135 (2006).

34 Chen J, Miao Y, Zhanga H and Matsunaga R, Non-destructive determination of carbohydrate content in potatoes using near infrared spectroscopy. *J. Near Infrared Spectrosc.*, **12**:311

(2004).

35 Haase NU, Prediction of potato processing quality by near infrared reflectance spectroscopy of ground raw tubers. *J. Near Infrared Spectrosc.*, **19**:37–45 (2011).

36 Marquez G and Anon MC, Influence of Reducing Sugars and Amino Acids in the Color Development of Fried Potatoes. *J. Food Sci.*, **51**:157–160 (1986).

37 Devahastin S and Niamnuy C, Modelling quality changes of fruits and vegetables during drying: A review. *Int. J. Food Sci. Technol.*, **45**:1755–1767 (2010).

38 Sankat CK and Castaigne F, Foaming and drying behaviour of ripe bananas. *LWT - Food Sci. Technol.*, **37**:517–525 (2004).

39 Diamante LM and Munro PA, Mathematical modeling of hot air drying of sweet potato slices. *Int. J. Food Sci. Technol.*, **26**:99 (1991).

40 Mayor L and Sereno AM, Modelling shrinkage during convective drying of food materials: A review. *J. Food Eng.*, **61**:373–386 (2004).

TABLE 1

Parameter	Model #	Algorithm	Dataset		LVs[c]	RMSE[d]			Adjusted-R^2			Control limits	
			Slice thickness	Number of features		C[e]	CV[f]	P[g]	C	CV	P	BCL[h]	UCL[i]
Moisture (d.b.)	A01	PLS[a]	5 mm	314 (full spectrum)	5	0.20	0.21	0.19	0.971	0.968	0.975	0.26	0.95
	A02	iPLS[b]		6	6	0.18	0.19	0.18	0.977	0.975	0.979	0.34	0.99
	A03	PLS	7 mm	314 (full spectrum)	6	0.13	0.13	0.15	0.985	0.984	0.981	0.01	1.16
	A04	iPLS		5	5	0.11	0.12	0.13	0.988	0.988	0.987	0.05	1.10
	A05	PLS	9 mm	314 (full spectrum)	6	0.15	0.15	0.15	0.982	0.981	0.982	0.05	1.02
	A06	iPLS		6	5	0.13	0.13	0.11	0.986	0.985	0.990	0.11	0.88
	A07	PLS	all	314 (full spectrum)	6	0.24	0.25	0.26	0.953	0.952	0.946	0.03	1.07
	A08	iPLS		6	5	0.24	0.24	0.26	0.954	0.953	0.948	0.04	1.07
Hue angle (h)	A09	PLS	5 mm	314 (full spectrum)	4	0.98	1.02	0.97	0.949	0.945	0.954	0.28	0.99
	A10	iPLS		5	5	0.78	0.82	0.88	0.967	0.964	0.961	0.28	1.12
	A11	PLS	7 mm	314 (full spectrum)	3	1.04	1.07	0.97	0.946	0.944	0.953	0.06	0.93
	A12	iPLS		6	4	0.91	0.93	0.91	0.959	0.957	0.958	0.07	1.00
	A13	PLS	9 mm	314 (full spectrum)	5	1.21	1.25	1.36	0.944	0.940	0.923	0.10	1.12
	A14	iPLS		5	5	1.18	1.20	1.26	0.946	0.944	0.934	0.10	1.07
	A15	PLS	all	314 (full spectrum)	3	1.19	1.20	1.24	0.937	0.936	0.930	0.07	1.05
	A16	iPLS		4	4	1.20	1.21	1.22	0.935	0.935	0.932	0.06	1.01
$L^* b^{*-1}$ ratio	A17	PLS	5 mm	314 (full spectrum)	6	0.09	0.10	0.09	0.933	0.923	0.939	0.08	0.93
	A18	iPLS		7	6	0.08	0.08	0.07	0.952	0.948	0.957	0.28	0.96
	A19	PLS	7mm	314 (full spectrum)	6	0.08	0.08	0.08	0.942	0.937	0.942	0.11	1.01
	A20	iPLS		9	5	0.08	0.08	0.07	0.939	0.937	0.947	0.14	0.94
	A21	PLS	9 mm	314 (full spectrum)	6	0.09	0.09	0.10	0.916	0.911	0.896	0.06	1.15
	A22	iPLS		9	8	0.08	0.08	0.08	0.923	0.919	0.920	0.02	0.95
	A23	PLS	all	314 (full spectrum)	7	0.10	0.10	0.10	0.909	0.906	0.897	0.02	1.00
	A24	iPLS		8	7	0.10	0.10	0.10	0.911	0.907	0.895	0.00	1.04

[a] Partial Least Squares; [b] Interval Partial Least Squares; [c] Latent variables; [d] Root Mean Square Error; [e] calibration; [f] cross-validation; [g] prediction; [h] bias control limit; [i] unexpected error control limit.

TABLE 2

Parameter	Model #	Dataset		RMSE[c]			Adjusted-R^2			Control limits	
		Slice thickness	Features	C[d]	CV[e]	P[f]	C	CV	P	BCL[g]	UCL[i]
Moisture (d.b.)	B01	5 mm	Diff[a]	0.28	0.28	0.26	0.947	0.945	0.959	0.21	0.94
	B02		Ratio[b]	0.29	0.29	0.26	0.943	0.945	0.959	0.31	0.90
	B03	7 mm	Diff	0.25	0.25	0.23	0.948	0.947	0.954	0.02	0.92
	B04		Ratio	0.27	0.27	0.24	0.938	0.938	0.949	0.06	0.90
	B05	9 mm	Diff	0.26	0.27	0.25	0.943	0.943	0.944	0.09	0.94
	B06		Ratio	0.33	0.33	0.29	0.910	0.909	0.925	0.15	0.87
	B07	all	Diff	0.38	0.38	0.37	0.886	0.885	0.888	0.03	0.99
	B08		Ratio	0.41	0.41	0.39	0.867	0.867	0.877	0.07	0.96
Hue angle (h)	B09	5 mm	Diff	1.13	1.15	1.04	0.930	0.928	0.949	0.33	0.92
	B10		Ratio	1.10	1.11	0.95	0.938	0.936	0.955	0.42	0.87
	B11	7 mm	Diff	0.97	0.98	1.06	0.954	0.953	0.950	0.07	1.09
	B12		Ratio	1.26	1.27	1.28	0.917	0.915	0.928	0.12	1.02
	B13	9 mm	Diff	1.36	1.38	1.34	0.929	0.927	0.929	0.05	0.98
	B14		Ratio	1.53	1.54	1.38	0.903	0.902	0.922	0.13	0.90
	B15	all	Diff	1.32	1.32	1.28	0.923	0.923	0.925	0.05	0.97
	B16		Ratio	1.50	1.50	1.46	0.900	0.900	0.904	0.03	0.97
L*/b* ratio	B17	5 mm	Diff	0.14	0.14	0.14	0.847	0.845	0.877	0.12	1.01
	B18		Ratio	0.14	0.14	0.13	0.852	0.849	0.880	0.31	0.99
	B19	7 mm	Diff	0.12	0.13	0.13	0.842	0.839	0.860	0.15	1.03
	B20		Ratio	0.14	0.14	0.13	0.808	0.804	0.857	0.21	0.92
	B21	9 mm	Diff	0.15	0.15	0.13	0.751	0.748	0.795	0.09	0.84
	B22		Ratio	0.16	0.16	0.13	0.731	0.729	0.783	0.10	0.84
	B23	all	Diff	0.16	0.16	0.15	0.763	0.762	0.769	0.09	0.93
	B24		Ratio	0.17	0.17	0.16	0.738	0.737	0.742	0.15	0.94

[a] model computed by using feature dataset comprising raw reflectance differences for all possible pairs of wavelengths; [b] model computed by using feature dataset comprising raw reflectance ratios for all possible pairs of wavelengths; [c] Root Mean Square Error; [d] calibration; [e] cross-validation; [f] prediction; [h] bias control limit; [i] unexpected error control limit.

TABLE CAPTIONS

Table 1. *Summary of the characteristics and performance metrics for the combinations of pre-processing and PLS and iPLS models complexity which gave the best results.*

Table 2. *Summary of the characteristics and performance metrics for the best-fitting models obtained using features extracted from both raw reflectance differences and ratios for each possible pair of wavelengths.*

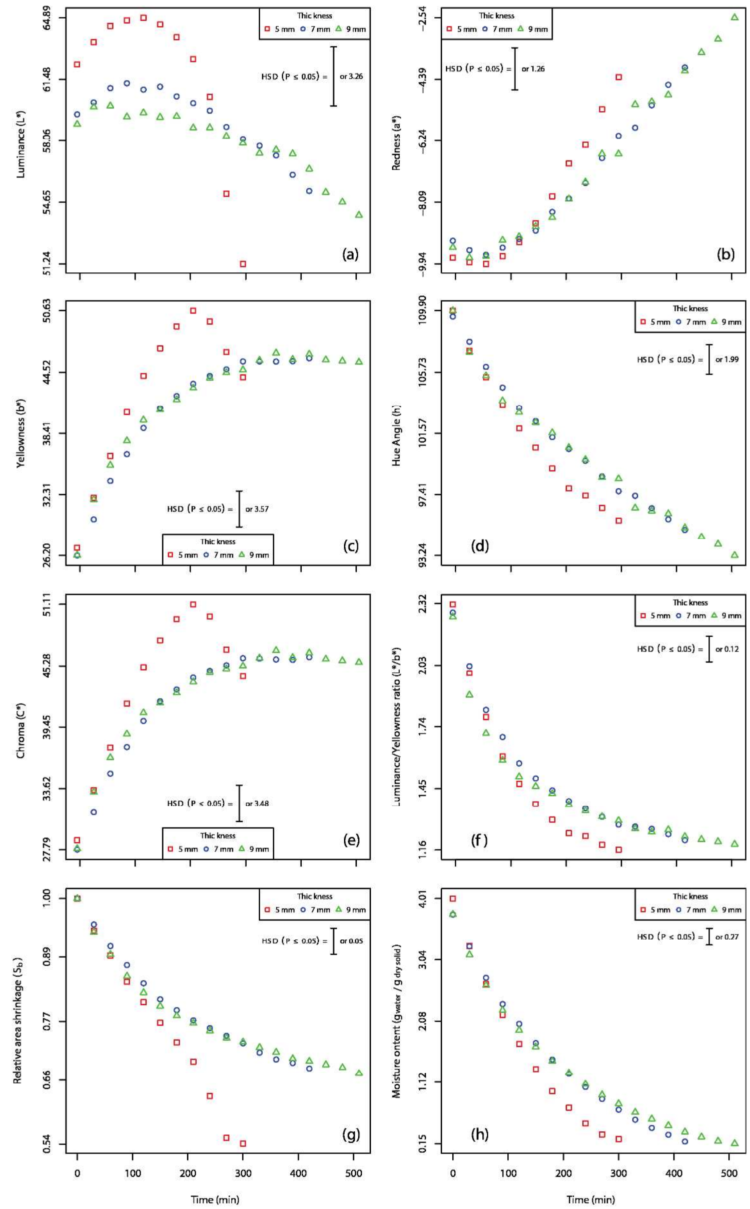


Figure 1. Changes in (a) luminance (L*), (b) redness (a*), (c) yellowness (b*), (d) hue angle (h), (e) chroma (C*), (f) luminance/yellowness ratio (L*b*-1 ratio), (g) relative area shrinkage and (h) moisture content (g water / g dry solid) of potato slices of 5-, 7- and 9-mm thickness during hot-air drying at 50°C. HSD, honestly significant difference (P ≤ 0.05).

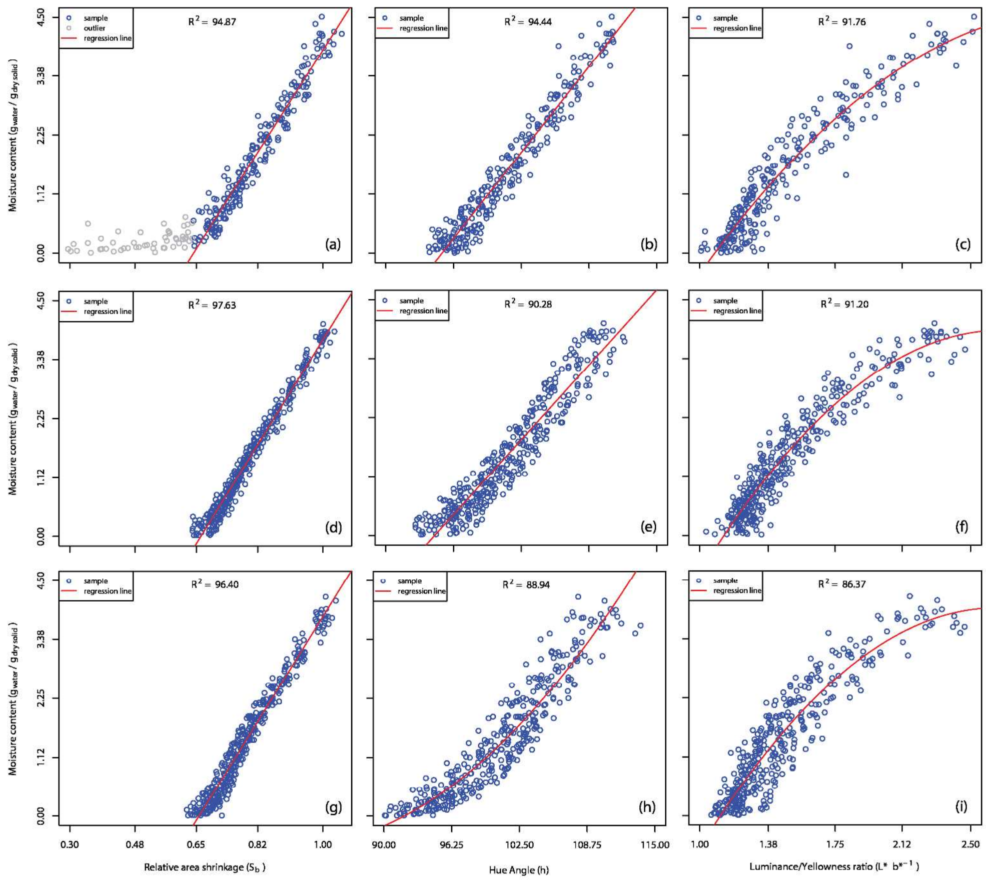


Figure 2. Linear (a, b, d, e and g), exponential (h) and quadratic (c, f and i) relationships having the relative area shrinkage, hue angle and L*b*-1 ratio, respectively, as predictors and moisture content (g water / g dry solid) as response variable. Plots refer to results obtained for 9-mm potato slices dried at 50°C up to an average moisture content of 12%.

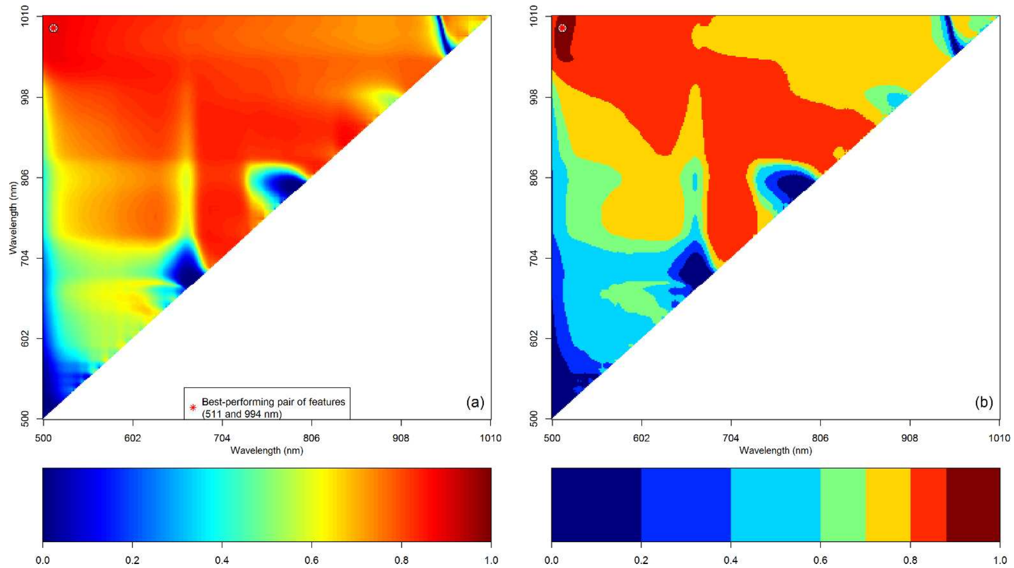


Figure 3. 2D-surface plots of the adjusted proportion of variance (adj-R2) explained by the all possible models computed from the raw reflectance difference (R[λ1]-R[λ2]) between two wavelengths. Results refer to the global dataset of all thicknesses. Figure (a) was plotted using the full color ramp and figure (b) shows a color ramp with breaks of 0.0, 0.2, 0.4, 0.6, 0.8, 0.85 and 1.0.

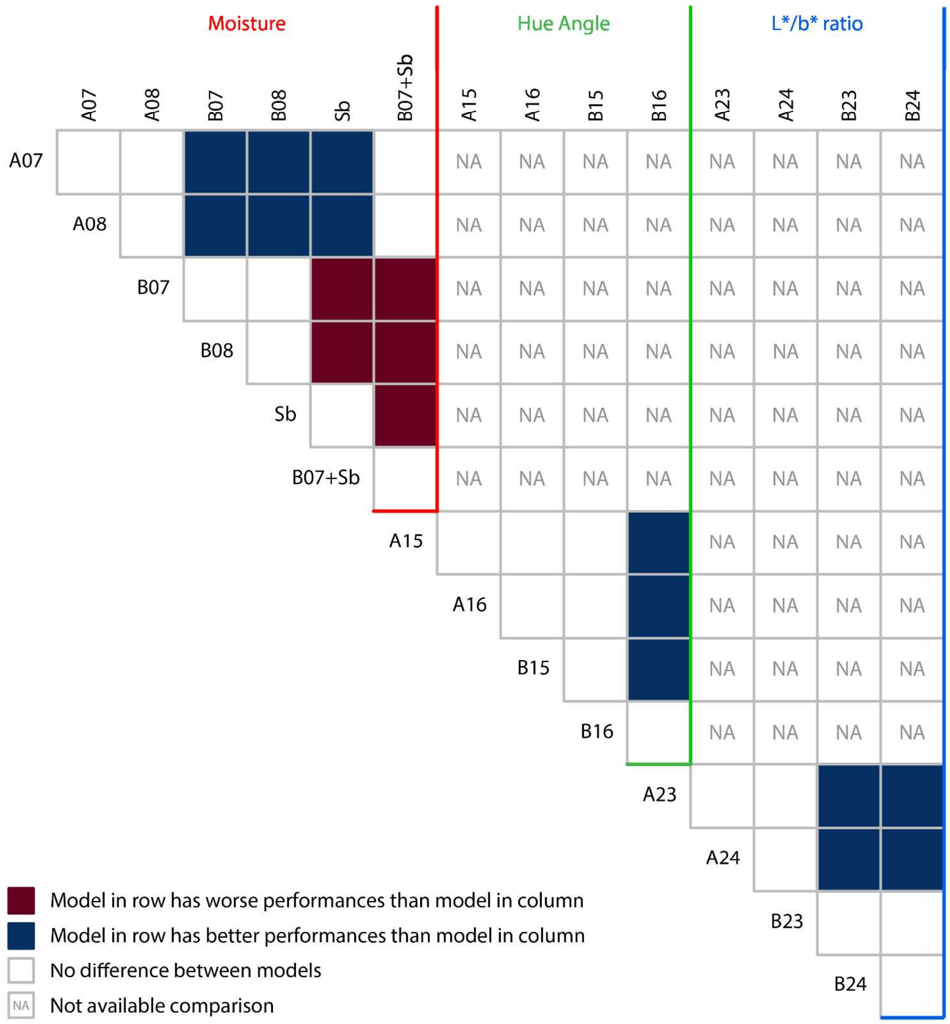
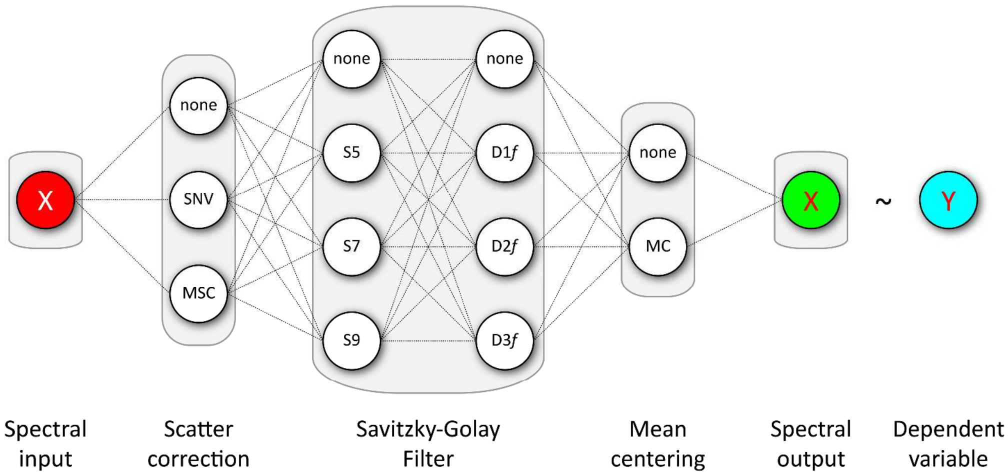
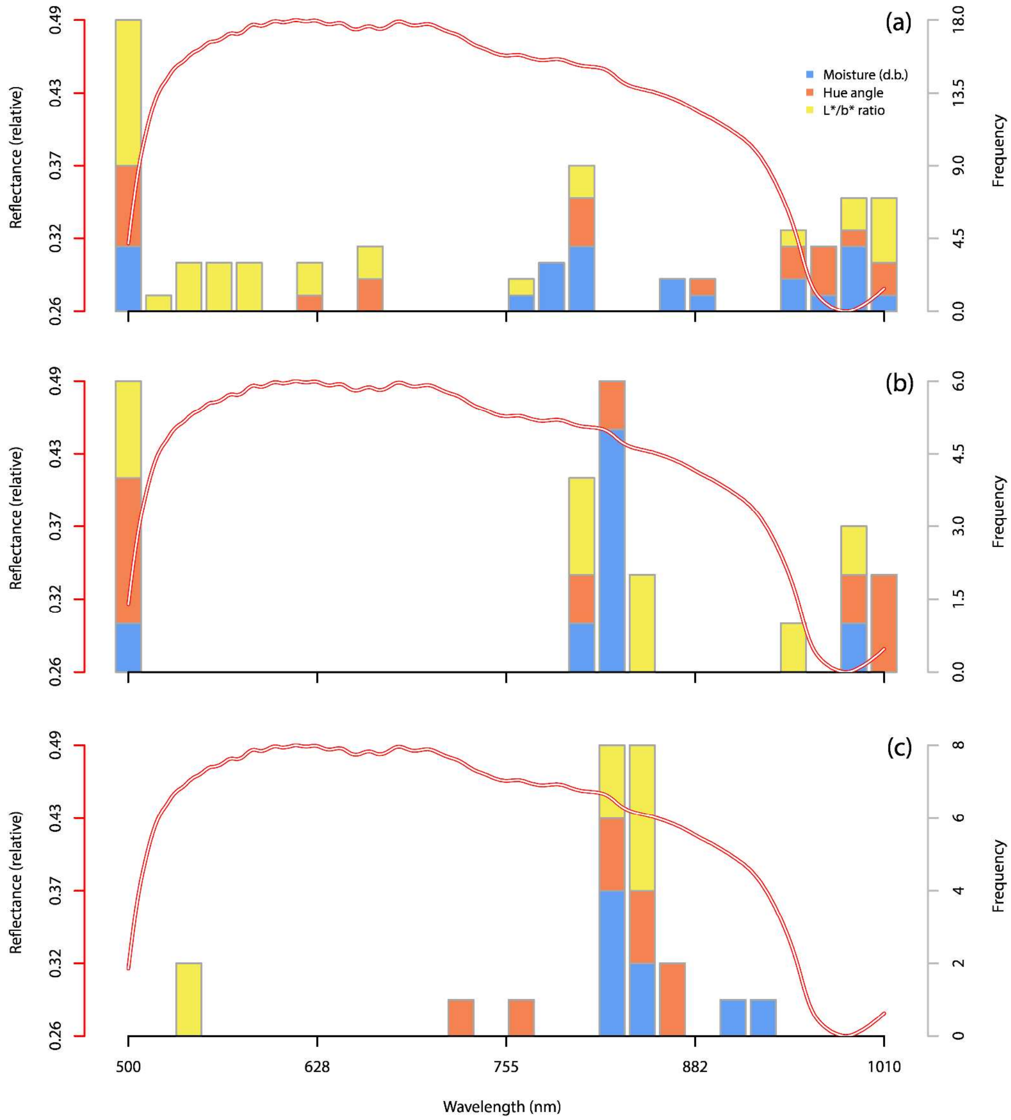


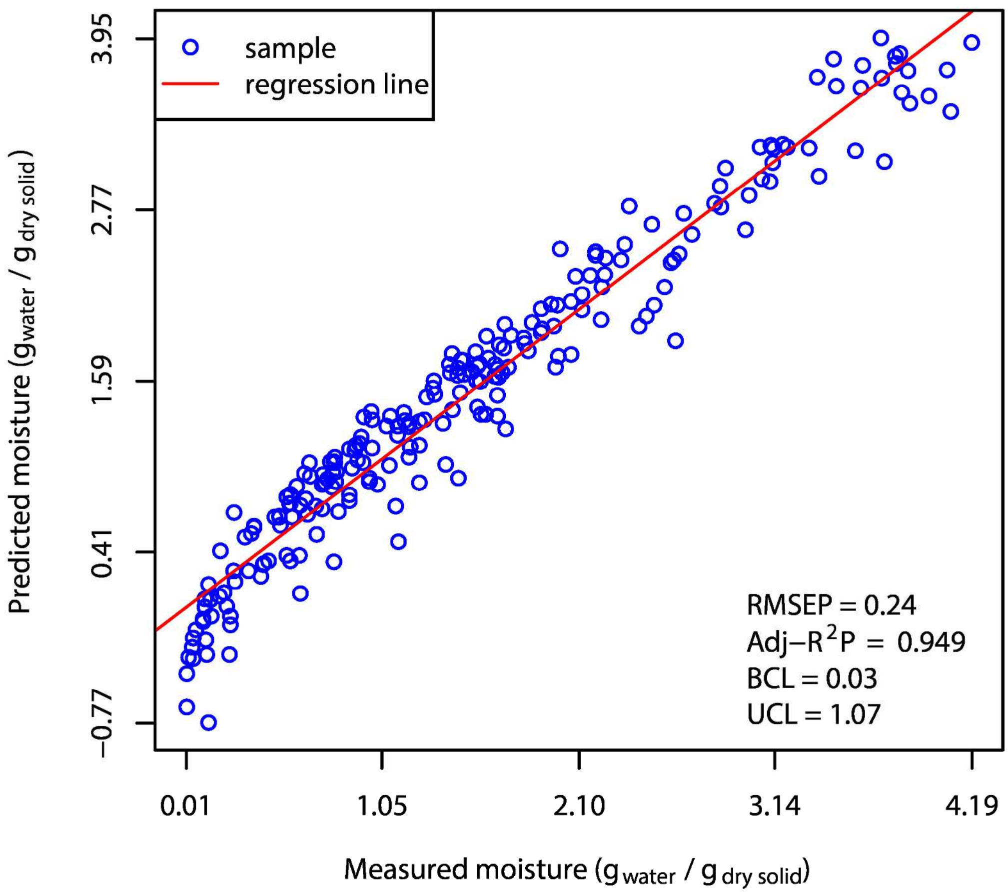
Figure 4. Upper triangular matrix used to compare models' performances through the Fisher's F-test.



Supplementary Figure 1.
Flowchart of the algorithm used to optimize the spectral pre-treatment selection procedure.



Supplementary Figure 2.
Frequency plot of features extracted in the 500÷1010-nm spectral range using (a) the iPLS algorithm, (b) the raw reflectance difference for each pair of wavelengths and (c) the raw reflectance ratio for each pair of wavelengths. Red line in each frequency plot represents the average raw reflectance spectrum.



Supplementary Figure 3.
Measured versus predicted values of moisture content for the external prediction set. Model was computed by using spatial data (Sb) in combination with raw reflectance difference (Table 2; model #B07; R[511 nm]–R[994 nm]) from the 5-, 7- and 9-mm potato dataset.