Universidad Pública de Navarra
Nafarroako Unibertsitate Publikoa

Department of Projects and
Rural Engineering

PhD THESIS

Near-infrared spectroscopy and hyperspectral imaging for non-destructive quality inspection of potatoes

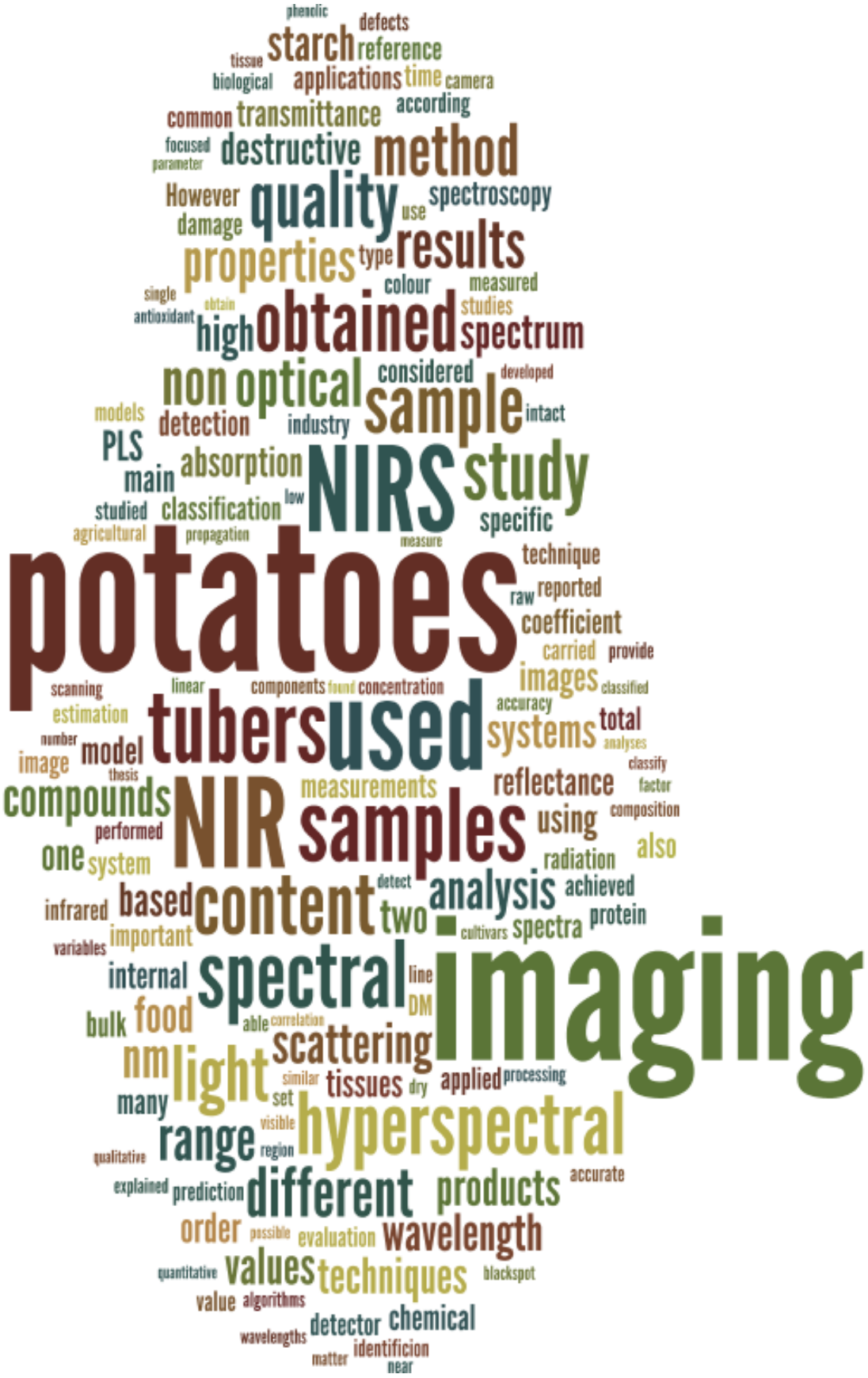
Supervisors

Carmen Jarén Ceballos

Silvia Arazuri Garín

Ainara López Maestresalas

Pamplona – Iruña, 2016


phenolic
defects
starch
reference
tissue
biological
applications
time
camera
according
common
transmittance
focused
parameter
destructive
method
However
quality
spectroscopy
damage
use
properties
type
results
colour
measured
single
obtain
antioxidant
high
obtained
studies
spectrum
non
optical
considered
developed
sample
models
detection
industry
intact
PLS
absorption
NIRS
study
main
studied
classification
low
specific
agricultural
propagation
measure
potatoes
technique
raw
reported
coefficient
carried
provide
images
classified
scanning
linear
components
found
concentration
accuracy
estimation
tubers
used
systems
total
number
model
analyses
image
classify
thesis
reflectance
factor
compounds
measurements
composition
NIR
samples
using
performed
one
system
radiation
also
detect
achieved
infrared
based
content
two
analysis
cultivars
spectra
protein
important
variables
internal
spectral
imaging
line
DM
bulk
food
able
correlation
scattering
nm
light
similar
tissues
dry
applied
processing
set
many
visible
hyperspectral
range
region
qualitative
explained
prediction
different
products
accurate
order
possible
evaluation
wavelength
quantitative
values
techniques
blackspot
value
algorithms
detector
chemical
wavelengths
identificion
matter
near

Dra. CARMEN JARÉN CEBALLOS, Profesora Titular del Departamento de Proyectos e Ingeniería Rural de la Universidad Pública de Navarra, y Dra. SILVIA ARAZURI GARÍN, Profesora Titular del Departamento de Proyectos e Ingeniería Rural de la Universidad Pública de Navarra.

INFORMAN: que la presente tesis doctoral titulada “Near-infrared spectroscopy and hyperspectral imaging for non-destructive quality inspection of potatoes”, y que constituye la memoria presentada por Ainara López Maestresalas ha sido realizada bajo nuestra dirección y que reúne las condiciones académicas exigidas por la legislación vigente para aspirar al grado de Doctor.

Y para que quede constancia, firmamos el presente informe en Pamplona a 29 de abril de 2016,

Fdo.: Profª. Dra. Carmen Jarén Ceballos

Fdo.: Profª. Dra. Silvia Arazuri Garín

Research is a high-hat word that scares a lot of people. It needn't. Research is nothing but a state of mind, a friendly welcoming attitude toward change.

- Charles F. Kettering –

Agradecimientos

El camino recorrido desde el inicio de esta tesis ha sido sumamente gratificante, tanto a nivel personal como profesional. Si bien es verdad que en todo recorrido aparecen baches que dificultan el trayecto, he tenido la enorme suerte de contar con mucha gente a mi alrededor para poder superarlos y continuar hacia adelante. No me gustaría olvidarme de ninguna de las personas que han contribuido a que este proyecto sea hoy una realidad. Muchas gracias a todas y todos por vuestra colaboración y ayuda.

Gracias a la Universidad Pública de Navarra por la financiación recibida durante estos años a través de una ayuda de investigación predoctoral sin la cual esta tesis no hubiera sido posible.

Me gustaría darles las gracias especialmente a mis directoras de tesis Carmen Jarén y Silvia Arazuri. Gracias por haberme animado a embarcarme en este viaje, por guiarme y motivarme en todo momento, por vuestro apoyo durante el desarrollo de esta tesis y por haber respetado siempre mis decisiones. Gracias Silvia y Carmen por vuestro esfuerzo para que este proyecto salga a la luz, que espero sea el primero de una larga lista.

Quiero agradecer a todos los miembros del grupo de investigación Mecatrónica Agraria y del departamento por haberme acogido desde el principio y por toda la ayuda prestada. En especial, a Txuma, por las clases magistrales, Alvarito, por esas imágenes tan bien hechas, Marce y Jaime por algún que otro café y muchas risas. A José por salvarme de los virus que han atacado en repetidas ocasiones mi ordenador. A Pedro por ser un libro abierto y por su paciencia. Y a Claudia por su gran ayuda y compañía sobre todo en esta etapa final de la tesis.

También a Carlos, muchas gracias por tu ayuda desinteresada (o eso espero), me ha encantado el sello personal que le has dado a esta tesis.

A Leire, Maite y Giorgio, por las charlas terapéuticas entre café y café o más bien entre caña y caña, eskerrik asko. Con vosotros los días duros se hacen mucho más llevaderos.

Gracias también al Instituto Vasco de Investigación y Desarrollo Agrario NEIKER y a sus miembros Nacho, Patrick, Roberto y Rakel, ya que es un placer trabajar con vosotros.

Quiero agradecer muy especialmente a José su acogida, y toda la ayuda prestada durante mi estancia en Copenhague. A pesar de haber sido corta, esta experiencia fue

muy intensa y puedo decir que de ella me llevo dos amigos para toda la vida, José y Ane, parranda bat dugu pendiente eta badakizue!.

I would like to thank Professor Wouter Saeys for his guidance and help during my stay in Leuven and to all the members of the MeBioS division of the Biosystem Department of KU Leuven. Many thanks to Professor Josse De Baerdemaeker, Ben Aernouts, Robbe Van Beers and Nghia Nguyen for being my patient teachers in that period. Especially, I would like to thank Mohammad Goodarzi and Janos C. Keresztes not only for their great support during the development of my thesis but also, and more important, for their friendship. Also, thanks to all the wonderful people I met in Leuven, Nadia, João (Pedrito), Nora and especially to Ioanna, I miss you my friend!.

Muchas gracias a mi cuadrilla, a todas y cada una de mis amigas, por ser las mejores que una podría tener, por esas charlas frente a un café arreglando el mundo o esas noches interminables en la ventana del mesón. Muy especialmente a Anabel que ha estado a mi lado desde que tengo uso de razón, siempre dispuesta a echar una mano, o las dos. Mucha suerte en esta nueva etapa que vas a comenzar, estoy segura de que lo harás genial, como todo.

También a mis txikas de Arden Street, la Peke, Meri, Lydi e Ido, porque sois un regalo para mí y porque sin vosotras la vida sería mucho más aburrida, estoy deseando veros!

No puedo olvidarme en este punto, de la parte más importante de mi vida, de la que más orgullosa estoy, mi familia. Muchas gracias ama, aita y tata, Mari Cruz, Esteban y Aintzane, por el amor incondicional que me demostráis a diario. También por haberme apoyado y respetado en todas las decisiones que he tomado en mi vida y por supuesto por haber intentado entender, a veces sin éxito, “qué es eso que hago con las patatas”. Y sobre todo, a la txikita de la casa, nuestro bien más preciado, gure Anetxu, gracias por hacer que se me pasen todos los males con sólo mirarte, asko maite zaitugu!. También a mis primas y primos, tías y tíos por ser parte de esta familia tan bien avenida que formamos. Por muchos años más compartiendo buenos momentos!.

Eta bukatzeko, zuri Josu, eskerrik asko bihotz bihotzetik, ez dut hitz nahikoa eskertzeko. Hasieratik hor egon zara, ni laguntzen, momentu zailenetan irribarre bat ateratzen saiatuz eta hori ezinbestekoa izan da niretzat. Zuri, nire bidelagunari, MILESKER!.

Abstract

Potato (*Solanum tuberosum* L.) is one of the most important crops in the world, occupying the fifth position in terms of production. Spain is one of the European countries with higher production and consumption of these tubers. Despite of being a highly appreciated product, potato industry faces the continuously growing demand of quality products, healthy, and free of damage, from consumers and regulatory bodies. The acceptance of these products in the market depends on several factors as the general aspect, texture, colour, absence of physical defects, etc. Therefore, it is essential to supply high quality products that result attractive to the final customer. With this aim, the present doctoral thesis is focused on offering tools for a quality control of potatoes in a non-destructive, effective and sustainable way. Different studies are presented where non-destructive spectroscopic techniques have been designed, developed and evaluated for the quality control of potato tubers. An optical characterization of potato flesh in the visible and near-infrared spectral range (500-1900 nm) was carried out using double integrating spheres, in order to understand the effects of light absorption and scattering in biological tissues to design more efficient sensors for the non-destructive evaluation of quality. A classification and determination of chemical compounds of potatoes by NIR spectroscopy and chemometric techniques was also accomplished. A system for the identification of internal damage in potato, invisible to the naked eye, was developed and evaluated using hyperspectral imaging systems in the visible, from 400 to 1000 nm, and near infrared (1000-2500 nm) spectral range. The results obtained in these studies show that absorption spectra of potato flesh are highly influenced by the water content of samples and also, that these tissues are highly scattering with anisotropy values close to 1 for the whole spectral range analysed. Moreover, an accurate classification of 18 different cultivars based on their phenolic and anthocyanins content, and hydrophilic antioxidant capacity was achieved with a precision of 86.1% by using an AOTF-NIR equipment. This would allow the incorporation of these methodologies to the selection of lines with high concentration of the desired compounds in breeding programs. Similarly, crude protein, nitrogen and phenolic content of a wide set of samples was successfully predicted by using the same non-destructive technique combined with a partial least square regression statistic method obtaining R^2 values of 0.86, 0.86 and 0.83 and SEP values of 0.68, 0.11 and 1.33 respectively. Finally, the results obtained with the hyperspectral imaging analysis allow a discrimination of healthy tubers from those with internal damage with an accuracy above 98% by applying a partial least squares discriminant analysis. In the

same way, a precise identification of internal damage in tubers 5 hours after bruising was achieved with 97.12% of bruised potatoes correctly classified. This allows the identification of damage at early stages of development, permitting these tubers to be driven to other industrial uses. The results obtained in this doctoral thesis could be used for a future implementation of these non-destructive spectroscopic techniques in real potato handling and packaging lines.

Resumen

La patata (*Solanum tuberosum* L.) es uno de los cultivos más importantes en el mundo, ocupando el quinto puesto en términos de producción. España es uno de los países europeos con mayor producción y consumo de estos tubérculos. Sin embargo, a pesar de ser un producto altamente valorado, la industria de la patata se enfrenta a la demanda cada vez más creciente de productos de calidad, sanos y libres de daños por parte de consumidores y organismo reguladores. La aceptación en el mercado de estos alimentos depende de varios factores, como son, el aspecto general, textura, color, ausencia de defectos físicos, etc. Resulta imprescindible por tanto suministrar productos de gran calidad que gusten al consumidor final. Con este objetivo, la presente tesis doctoral se centra en ofrecer herramientas para un mayor control de la calidad de manera no destructiva, eficaz y sostenible. Para ello, se presentan diferentes estudios en los cuales se han diseñado, desarrollado y evaluado técnicas espectroscópicas no destructivas para el análisis de calidad en tubérculos de patata. Se ha llevado a cabo en primer lugar una caracterización óptica de los tejidos de patata en el rango visible e infrarrojo (500-1900 nm) mediante el uso de Dobles Esferas Integradoras, para entender los efectos de absorción y dispersión que se producen en los tejidos biológicos y así poder diseñar sensores más eficientes para la evaluación de la calidad de forma no destructiva. Además, se ha realizado una clasificación y determinación de compuestos químicos en patata mediante espectroscopía NIR y el uso de técnicas quimiométricas de análisis multivariante. Por último, se ha diseñado y evaluado un sistema de detección de daños internos en patata, no apreciables a simple vista, mediante análisis de imágenes hiperespectrales obtenidas en el rango visible, de 400 a 1000 nm, e infrarrojo (1000-2500 nm) del espectro. Los resultados obtenidos en estos estudios muestran, por un lado, cómo los espectros de absorción están fuertemente influenciados por la alta concentración de agua de las muestras y, por otro, cómo los tejidos de patata son altamente dispersivos, con unos valores del factor de anisotropía muy cercanos a 1 a lo largo de todo el rango espectral estudiado. Además, se ha conseguido clasificar una colección de 18 cultivares de patata de acuerdo a su contenido en polifenoles, antocianinas y capacidad antioxidante, mediante el empleo de un equipo NIRS con tecnología AOTF, con una precisión de 86.1%. Esto permite la incorporación de estas metodologías en la selección de líneas con alta concentración de los compuestos químicos deseados en programas de mejora genética. De forma similar, se ha logrado estimar la concentración de proteína bruta (R^2: 0.86; SEP: 0.68), nitrógeno (R^2: 0.86; SEP: 0.11)

y polifenoles (R^2: 0.83; SEP: 1.33), en un grupo de muestras con alta variabilidad, de manera no destructiva empleando el mismo equipo combinado con análisis de regresión por mínimos cuadrados parciales. Finalmente, los resultados obtenidos en el análisis de imágenes hiperespectrales han permitido discriminar patatas sanas de aquéllas con daño interno con una fiabilidad superior al 98% mediante un análisis discriminante por mínimos cuadrados parciales. Al mismo tiempo, se ha logrado identificar la presencia de daño interno en tubérculos 5 horas después de haberlo inducido, con una precisión de 97.12%. Esto permite localizar tubérculos afectados en las primeras fases de desarrollo del daño y por tanto evitar que lleguen al consumidor final. Los resultados obtenidos en esta tesis doctoral podrían emplearse para una futura implementación de estas técnicas espectroscópicas no destructivas en líneas reales de manipulación y envasado de patata.

Laburpena

Ekoizpenaren arabera patata bosgarren postuan kokatuta dagoenez gero, mundu mailako labore nagusienetariko bat da. Europa mailan Espainiako ekoizpena eta kontsumoa handia bada ere, industria behartuta dago produktu osasuntsu eta kalitatezkoak lortzera, kontsumitzaile eta erakunde erregulatzaileen eskaeren eraginez. Produktu hauek merkatuan onartu ahal izateko hainbat ezaugarri bete behar dituzte: itxura orokorra, kolorea, testura eta kalte fisikorik eza, beste hainbaten artean. Laburbilduz, derrigorrezkoa da kalitate altuko produktuak ekoiztea.

Helburu honi jarraiki, doktorego-tesi honen xedea patata laborearen kalitate-kontrolerako tresnak eskaintzea da. Horretarako, espektroskopian oinarrituta hainbat ikerketa diseinatu, garatu eta aztertu dira, teknika hau erabiltzen duten tresnak ez baitira suntsitzaileak, eraginkorrak eta jasangarriak baizik.

Lehen ikerketa patataren ehuneko propietate optikoetan ardazten da, neurketak 500-1900 nm tartean egiten dira, hurbileko infragorri-ikusgai espektroan. Esfera bikoitzak erabiltzen dira ehun biologikoetan sortzen diren absortzio eta dispertsio efektuak ulertzeko, ondoren sentsore ez-suntsitzaile eta eraginkorragoak diseinatzeko. Bestalde, patatak sailkatzen dira eta hauen konposatu kimikoak aztertzen dira, hurbileko infragorri-espektroskopia (NIRS) eta kimiometria teknikako aldagai anitzeko analisiarekin konbinatuz. Azkenik, ikusgai ez dauden barruko kalteak detektatzeko sistema bat diseinatu eta aztertu da, 400-1000 nm (VIS) eta 1000-2500 nm (NIR) tarteetan hartutako irudi hiperespektralen bidez.

Lortutako emaitzek agerian uzten dute laginen ur-kontzentrazioak eragin handia duela absortzio espektroan eta horrez gain, patataren ehunak oso dispertsiboak direla, anisotropia faktoreko balioak 1etik gertu baitaude.

Azkenik, 18 patata ekoizpen desberdin sailkatu dira NIRS teknologiaren bidez, polifenol, antoziania eta antioxidatzaile gaitasunaren arabera (%86ko zehaztasunarekin). Hobekuntza genetikoen programetan kontzentrazio handiko ekoizpenak aukeratzen laguntzen digu emaitza honek, eta teknologia bera erabiliz, proteina, nitrogeno eta polifenol kontzentrazioa iragartzeko ereduak garatu daitezke.

Era berean, proteina, nitrogenoa eta polifenolak kalkulatzea lortu da, aldakortasun handiko laginetan modu ez-suntsitzailean, ekipamendu berarekin karratu txiki partzialen (PLS) erregresioarekin konbinatuz.

Azkenean, irudi hiperespektralen emaitzei esker barruko kalteak dituzten patatak aldendu ditzakegu (%98ko zehaztasunarekin), PLS-DA (karratu txiki partzialen analisi

diskriminatzailea) erabiliz. Horretaz gain, posible izan da kaltea sortu eta 5 ordu pasa ondoren hauek identifikatzea. Honek kaltetutako patatak aldentzen laguntzen digu kontsumitzailearengana ailegatu baino lehen.

Etorkizunean doktorego-tesiaren emaitzak erabil daitezke, espektroskopian oinarrituta teknika hauek benetako manipulazio eta ontziratze prozesuetan inplementatzeko.

Table of contents

List of figures

List of tables

List of abbreviations and symbols

λ	Wavelength
μ_a	Bulk Absorption Coefficient
μ_s	Bulk Scattering Coefficient
μ_s'	Reduced Scattering Coefficient
μ_t	Bulk Extinction Coefficient
α, ε, *κ*	Beer-Lambert Absorption Coefficient
ε	Molar Extinction Coefficient
ρ	Mass Concentration
c	Amount Concentration of Absorber
l	Path length of the Absorber
1D	One Dimension
2D	Two Dimensions
A	Absorbance
a.s.l.	Above Sea Level
AD	Adding-Doubling
ANN	Artificial Neural Networks
AOTF	Acousto-Optical Tunable Filter
BIL	Band Interleaved by Line
BIP	Band Interleaved by Pixel
BOP	Bulk Optical Properties
BSQ	Band Sequential
CC	Correctly Classified
CCBAT	Centro de Conservación de la Biodiversidad Agrícola de Tenerife
CCD	Charge-Coupled Device
CGE	cyanidin 3-O-glucoside equivalents
CGP	Coagulating Protein
CMOS	Complementary Metal Oxide Semiconductor
CP	Crude Protein
CV	Cross Validation
cv.	Cultivar
D	Dark References
DF	Dilution Factor
DIS	Double Integrating Spheres
DM	Dry Matter

DT	Detrend
DTGS	Deuterated Triglycine Sulfate
DW	Dry Weight
FAO	Food and Agriculture Organization of the United Nations
FIR	Far Infrared
FN	False Negatives
FP	False Positives
FTIR	Fourier Transform Infrared Spectroscopy
FW	Fresh Weight
g	Anisotropy Factor
GAE	Gallic Acid Equivalents
GC	Gas Chromatography
GLC	Gas-Liquid Chromatography
GM	Genetically Modified
HAC	Hydrophilic Antioxidant Capacity
HPLC	High-Performance Liquid Chromatography
HPTLC	High-Performance Thin-Layer Chromatography
HSI	Hyperspectral Imaging
IAD	Inverse Adding-Doubling
IMC	Inverse Monte Carlo
InGaAs	Indium Gallium Arsenide
INIA-FEDER	Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria-Fondo Europeo de Desarrollo Regional
InSb	Indium Antimonide
IPLSR	Interval Partial Least Squares Regression
IR	Infrared
IS	Integrating Spheres
IU	International Units
IWT	Agency for Innovation by Science and Technology
KM	Kubelka-Munk
KNN	K-Nearest Neighbour
KPLS	Kernel Partial Least Squares
KS	Kennard and Stone algorithm
KU Leuven	Katholieke Universiteit Leuven
LCTF	Liquid Crystal Tunable Filter
LDA	Linear Discriminant Analysis
LE	Lutein Equivalents

LEDs	Light Emitting Diodes
LR	Linear Regression
LRT	Logarithm Reference Transformation
LS-SVM	Least Squares-Support Vector Machines
LVs	Latent Variables
MC	Mean Centre
MCT	Mercury Cadmium Telluride
MeBioS	Mechatronics, Biostatistics and Sensors
MIR	Mid Infrared
MLR	Multiple Linear Regression
MPLS	Modified Partial Least Squares
MSC	Multiplicative Scatter Correction
MW	Molecular Weight
n	Number of Samples
N	Nitrogen
n/a	No Available Data
NDF	Neutral Density Filter
NEIKER	Instituto Vasco de Investigación y Desarrollo Agrario
NIR	Near Infrared
NIRS	Near Infrared Spectroscopy
Norm	Normalization
NPK	Nitrogen Phosphorus Potassium
OSC	Orthogonal Signal Correction
PbSe	Lead Selenide
PC	Principal Component
PCA	Principal Component Analysis
PCR	Principal Component Regression
PLS	Partial Least Squares
PLS-DA	Partial Least Squares Discriminant Analysis
PLSR	Partial Least Squares Regression
PRT	Power Reference Transformation
R	Reflectance
R	Correlation Coefficient
R^2	Determination Coefficient
RDA	Recommended Daily Intake
RF	Radio Frequencies
RGB	Red-Green-Blue

RI	Refractive Index
RMSECV	Root Mean Square Error of Cross Validation
RMSEP	Root Mean Square Standard Error of Prediction
ROI	Region of Interest
RP	Recoverable Protein
	Ratio of Prediction to Deviation
RS	Reducing Sugars
R_t	Total Reflectance
RTA	Proyectos de Investigación Fundamental Orientada
RTE	Radiative Transfer Equation
RTT	Radiative Transfer Theory
SD	Standard Deviation
SECV	Standard Error of Cross Validation
SEP	Standard Error of Prediction
SG	Savitzky-Golay
Si	Silica
SIMCA	Soft Independent Modelling of Class Analogies
SNR	Signal to Noise Ratio
SNV	Standard Normal Variate
SVM	Support Vector Machines
SWIR	Short Wave Infrared
T	Transmittance
T^2	Hotelling Values
T_c	Unscattered/collimated Transmittance
TC	Total Carotenoids
TE	Trolox Equivalents
TLC	Thin-Layer Chromatography
TMA	Total Monomeric Anthocyanin
TN	True Negatives
TP	True Positives
TRS	Total Reducing Sugars
TS	Total Sugar
TSP	Total Soluble Phenolics
T_t	Total Transmittance
UK	United Kingdom
UPNA	Universidad Pública de Navarra
USA	United States of America

USDA	United States Department of Agriculture
UV	Ultraviolet
UWW	Under Water Weight
V	Volume
VC	Vitamin C
VIS	Visible
W	Weight
WB	Weight Baseline
W_r	White References

Chapter 1.

Introduction and objectives

.

1.1. Importance of the potato industry

Potato (*Solanum tuberosum* L.) is one of the most important crops in the world being considered a staple food in many developing countries (Alva *et al.*, 2011).

In 2014, this crop occupied the fifth position worldwide in terms of production after sugar cane, maize, rice and wheat according to the Food and Agriculture Organization of the United Nations (FAO), with a production above 385 million tonnes (Mt) (FAO, 2014). The gross production value for this crop accounted for more than 88 billion dollars (USA terminology) with an annual consumption around 32 kg per person (FAO, 2013).

Spain occupied the 26th position in the world ranking of producer countries, with a production of around 2.3 Mt in 2014, below countries such as China (96.1 Mt), India (46.4 Mt) or Russia (31.5 Mt) (FAO, 2014). Within Europe, Spain was the 11th country with higher potato production, after Russian Federation, Ukraine and Germany among others. Nevertheless, Spain represented the 5th European importer country in 2013 with more than 600,000 tons of potatoes imported from other countries, principally France (FAO, 2013).

The importance of this crop also derives from the fact that potatoes can be used in many ways such as staple food, cash crop, animal feed, and as a source of starch for many industrial uses (López *et al.*, 2013).

Potato is a highly productive crop, generating more food per unit area and per unit time than maize, rice and wheat (Changchui, 2008). Its production almost reached 19 tons hectare^{-1} (t ha^{-1}) in 2013 followed by maize (5.52 t ha^{-1}), rice (4.52 t ha^{-1}) and wheat (3.26 t ha^{-1}) (FAO, 2013).

These tubers are rich in protein, calcium, potassium, and vitamin C, and have an especially good amino acid balance. Moreover, they supply high levels of energy due to their starch content (Woolfe & Poats, 1987). Raw tubers contain about 80% of water and 20% of dry matter. About 60 to 80% of the dry matter is starch, representing the major component in potato carbohydrates (Prokop & Albert, 2008).

The starch in raw potato cannot be digested by humans, for this reason, its consumption without any preparation is not a common practice, and normally potatoes are consumed after boiling (with or without the skin), baking or frying. The preparation methods employed affect potato composition in a different way, but in general terms they all reduce fibre and protein content. This is due to leaching into cooking water and oil, destruction by heat treatment or chemical changes such as oxidation (Prokop & Albert, 2008).

Boiling is the most common method of potato preparation worldwide and this practice causes a significant loss of vitamin C, especially in peeled potatoes. When potatoes are fried, either for French fries or chips preparation, they absorb high content of fat and reduce their mineral and ascorbic acid content. Losses of vitamin C are generally higher when baking rather than boiling due to the higher oven temperatures, but losses of other vitamins and minerals are lower (Prokop & Albert, 2008). Table 1.1 shows the main components of potato in its most common consumption preparations (USDA, 2016).

Table 1.1. Composition[a] of Raw and Cooked Potatoes (USDA, 2016).

Composition of potatoes per 100g					
	Raw	Baked without salt	Boiled without salt	Fried without salt	Chips plain
Water (g)	79.25	75.42	76.98	61.51	2.54
Energy (kcal)	77	93	87	172	559
Protein (g)	2.05	1.96	1.87	2.66	4.45
Fat (g)	0.09	0.10	0.10	5.22	38.41
Carbohydrates (g)	17.49	21.55	20.13	28.71	52.02
Fibre (g)	2.1	1.5	1.8	2.6	3.1
Potassium (mg)	425	391	379	451	751
Sodium (mg)	6	5	4	32	388
Phosphorus (mg)	57	50	44	97	125
Magnesium (mg)	23	25	22	26	43
Calcium (mg)	12	5	5	12	27
Vitamin C (mg)	19.7	12.8	13	13.3	8.2
Vitamin A (IU)	2	0	3	0	0
Vitamin B 6 (mg)	0.298	0.301	0.299	0.184	0.407
Niacin (mg)	1.061	1.395	1.439	2.218	3.240

[a]*g, grams; kcal, kilocalories; mg, milligrams; IU, international units.*

1.2. Quality of potatoes

Quality of potato and potato products is determined by several attributes that determine their final acceptance in the market. Both potato consumers and the retail sector prefer potatoes of high quality; therefore, potato industry faces the continuously growing demand of quality products (Peña *et al.*, 2001). Potato quality is determined by its chemical constituents and physical properties (microstructure).

Texture is considered an essential factor in the consumers' perception of the quality of potatoes (García-Segovia *et al.*, 2008). Texture of potatoes is determined by several mutually dependent factors including the chemical composition (dry matter content, starch, sugars, proteins, etc) and both the agronomic and the storage conditions (Kaur

et al., 2002; Liu *et al.*, 2009). Thus, according to Subedi and Walsh (2009) dry matter (DM) is a suitable index for determining potato quality.

Moreover, besides the nutritional value of potatoes, external appearance, flavour and the presence of defects caused by mechanical damage are important. According to Kays (1999), the appearance of foods directly affects the final purchase decisions of customers. Colour is a crucial attribute since it is the first characteristic that consumers take into consideration to evaluate the quality of many products.

Nowadays, in addition to be of high quality, potato products are required to be nutritious and healthy. In this respect, phenolic compounds have gained considerable interest in the last years due to their antioxidant nature that promote health benefits (Manach *et al.*, 2004). High content of these phytochemical compounds are found in potatoes (Al Saikhan *et al.*, 1995), especially in the red- and purple-fleshed varieties (Ezekiel *et al.*, 2013). For this reason, it is interesting to identify and characterize tubers presenting considerably phytochemical content that may result attractive to consumers.

Another important aspect that influences potato quality is the damage to tubers during harvesting. This is one of the most important causes of lower potato quality and value, and increases the incidence of losses and diseases during storage (Bentini *et al.*, 2006). The economic losses in the fruit and vegetable industry related to bruising are considerable (Van Zeebroeck *et al.*, 2003). According to Peters (1996), in the American potato industry, bruising represents substantial economic losses every year and 70% of total damage is caused by harvesting. Mathew and Hyde (1997) reported an estimated $20 to $60 million losses due to potato tuber bruising in the Washington State, the second major potato producer in the United States of America (USA), in a particularly bad year. An important factor contributing to the financial loss is the fact that many times, the affected tubers do not show external damage and are therefore processed as healthy ones, resulting in a waste and loss of confidence among consumers (Evans & Muir, 1999). However, the detection of those affected tubers, would allow these potatoes to be assigned to other industrial uses avoiding the use of damaged areas. Traditionally, sorting of damaged tubers is accomplished by human labour, but this type of handling has some disadvantages such as high cost, low efficiency and inability to sort internal defects (Rady & Guyer, 2015b).

On the other hand, the determination of internal composition of potatoes involves traditional “wet chemistry” methods, for instance gas chromatography (GC) and high-performance liquid chromatography (HPLC). These techniques are time-consuming, expensive, destructive, (Mehrubeoglu & Cote, 1997) and produce chemical waste (Manley, 2014).

For these reasons, in today's competitive market, where the importance of quality monitoring is increasing, it is becoming necessary to develop faster, more economical, safer and more versatile non-destructive techniques to efficiently measure quality properties of foods (López-Maestresalas *et al.*, 2015).

1.3. Non-destructive spectroscopic techniques

Development in many technology areas such as computing and optical devices has made possible actual production improvements (Jarén *et al.*, 2016). However, to assure a long-term food security is essential to progress in the state of the art in the non-destructive quality control of potatoes.

In this respect, near-infrared spectroscopy (NIRS) is considered one of the most advanced technologies for non-destructive quality assessments of fruits and vegetables (Cubeddu *et al.*, 2001; Cubero *et al.*, 2011; Nicolaï *et al.*, 2007; Wang & Li, 2013). Compared to reference methods, NIRS is a rapid, environmentally friendly and non-invasive technique. The main disadvantage is that relies on reference methods, but even though, NIRS measurements and predictions are considered more reproducible than those analytical methods (Manley, 2014).

Together with NIRS, hyperspectral imaging systems have recently started to be used in food quality assessment (Lorente *et al.*, 2012). They are capable of providing both spatial and spectral information simultaneously. Therefore, distribution of compounds within a sample can be provided. For this reason, hyperspectral imaging is currently being extensively investigated for quality and safety control of food products (ElMasry *et al.*, 2012b).

Moreover, in the last years, studies dealing with optical characterization of biological tissues are being carried out within the agro-food area in order to investigate how light propagates through tissues. The study of the behaviour of light is very useful in order to design more efficient NIR sensors to obtain information non-destructively from agricultural and food products (Zamora-Rojas *et al.*, 2014).

For all these reasons, the present doctoral thesis is focused on providing real and efficient solutions to the actual needs of the potato industry in terms of quality assessment. The possibilities of both non-destructive spectroscopic techniques are explored with the aim of providing food industry with adequate and advanced spectroscopic tools for assessment and quality control along the agro-food chain. For this, quality evaluation of potatoes is tackled from three different points of view. First by a study of the behaviour of light through potato tissues. Second by quantitative and qualitative analyses of potato compounds by near-infrared spectroscopy. And, finally by

the identification of internal damages on intact tubers using hyperspectral imaging systems.

1.4. Objectives

The main goal of this PhD thesis was to design, develop and evaluate non-destructive techniques in the visible and near-infrared spectral ranges for the quality control of potatoes. The main objective was accomplished with the following specific objectives:

1. To determine the optical properties of potato flesh in the visible and near-infrared spectral range in order to study light propagation in a biological tissue.
2. To develop and evaluate qualitative models for the classification of purple- and red-fleshed tubers according to their phytochemical content by near-infrared spectroscopy.
3. To develop and evaluate multivariate calibration models for the prediction of chemical compounds in potatoes by the use of near-infrared spectroscopy.
4. To identify intact tubers presenting internal damage (blackspot) at early stages by hyperspectral imaging techniques in the visible and near-infrared spectral range.
5. To map areas affected by blackspot in intact potatoes using hyperspectral imaging techniques in the visible and near-infrared spectral range.
6. To compare the performance of a Vis-NIR and a SWIR hyperspectral setups in the detection of blackspot in raw tubers.

1.5. Thesis structure

The doctoral thesis manuscript presented here is organized in 7 chapters as follows:

- **Chapter 1:** this introductory chapter establishes the context of the thesis by providing the background necessary to understand the objectives established and the progress of the work. In particular, chapter 1 highlights the importance of the potato crop in terms of production and consumption. Moreover, it sets out the current issues faced by the industry in terms of quality assessment. It also introduces different non-destructive techniques as an alternative to conventional methods for quality control of potatoes. Finally, the main and specific objectives are described in this chapter along with the structure of the thesis and the research framework of it.

- **Chapter 2:** in this chapter, the different techniques and methods used in this thesis are described in detail. This chapter points out the main characteristics, components and modes of measurements of the non-destructive spectroscopic techniques along with a description of the commonly used methods to measure optical properties of foods. Additionally, it reviews the state of the art on NIRS and imaging applications for the analysis of potatoes. It should be mentioned that section 2.4 is based on a review paper (López *et al.*, 2013), published after an extensive bibliographic revision carried out during the first year of the doctoral program. Nevertheless, this part was updated to include the latest contributions to the subject.

- **Chapter 3:** this chapter accomplishes the objective 1 explained in section 1.4. It includes a research paper recently published dealing with the study of light propagation through potato tissues. For this, the bulk optical properties of potato flesh were measured in the 500-1900 nm wavelength range by means of double integrating spheres and unscattered transmittance measurements.

- **Chapter 4:** this chapter addresses objectives 2 and 3. Qualitative analyses are carried out in order to classify a collection of purple- and red-fleshed potato varieties according to their phytochemical content. In the second part of the chapter quantitative analyses are developed to estimate chemical compounds of potatoes by NIRS. All the chemical analysis accomplished in this chapter are described in the Appendix, at the end of the document. This chapter includes a research paper already published.

- **Chapter 5:** this chapter complies with objectives 4, 5, and 6. In this chapter, firstly, an identification of potatoes affected by an internal damage (blackspot) using hyperspectral imaging system in the visible and NIR spectral range is completed. Moreover, a mapping of the affected areas in each tuber is accurately achieved. Finally, this chapter compares two hyperspectral setups for blackspot identification on raw tubers.

- **Chapter 6:** in this chapter a general discussion is presented. It summarizes main findings of this doctoral thesis and compares them with other authors' results highlighting the novelty provided by this research work. Additionally, possible future lines of research are presented.

- **Chapter 7:** this final chapter presents the overall and specific conclusions obtained in this PhD thesis.

Finally, a list of scientific publications resulting from this research work is included.

1.6. Research framework of the thesis

This doctoral thesis is the result of the research work accomplished by the author during the period 2011-2016, as a member of the Agrarian Mechatronics research group of the Department of Projects and Rural Engineering at the Universidad Pública de Navarra (UPNA).

This thesis is framed within two national research projects financed by the Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA) with the aim of applying NIR spectroscopy to a potato plant breeding program (INIA-FEDER RTA2011-00018-C03-03 and INIA-FEDER RTA2013-00006-C03-03). In both projects, the Agrarian Mechatronics research group works in conjunction with the Basque Institute for Agricultural Research and Development (NEIKER-Tecnalia) and with the Centro de Conservación de la Biodiversidad Agrícola de Tenerife (CCBAT).

Additionally, specific research dealing with the bulk optical properties of potatoes and the use of hyperspectral imaging techniques presented in this thesis is the result of a collaborative work with the Mechatronics, Biostatistics and Sensors (MeBioS) division of the Biosystems Department at the University of Leuven (KU Leuven). This work was completed during a research stay of 4 months at that department. The MeBioS division is focused on measurements of the electromagnetic properties of biological materials to develop novel sensor technologies and to combine these properties with advanced statistical techniques for quality assessment and optimization of the agro-food chain.

Chapter 2.

Literature review

Part of this chapter has been published as:

- López, A., Arazuri, S., García, I., Mangado, J., & Jarén, C. (2013). A review of the application of near-infrared spectroscopy for the analysis of potatoes. *Journal of Agricultural and Food Chemistry*, *61*(23), 5413-5424. Impact factor: 3.107.
- Jarén, C., López, A., & Arazuri, S. (2016). Advanced analytical techniques for quality evaluation of potato and its products. In J. Singh & L. Kaur (Eds.), Advances in Potato Chemistry and Technology (2nd ed., pp. 563-602): Elsevier.

Literature review

Abstract

In this chapter, a review of the non-destructive techniques used in this doctoral thesis is presented. First, an introduction of NIRS technology is explained along with the different instrumentation available and the modes of measurements commonly used. Then, NIR light interaction with biological tissues is described, the most frequent measurements setups are presented and the different algorithms generally used to extract optical properties of agricultural and food products are described. Hyperspectral imaging systems are also commented in this chapter, their main components and configuration setups. After that, chemometrics is explained and described and the most common methods used in quantitative and qualitative analyses are exposed. The last part of the chapter is focused on the applications of both NIR and imaging technologies for the analysis of potatoes. In this part the state of the art regarding non-destructive applications in quality control of potatoes is reviewed. For this, the different studies carried out since the first application of NIRS and hyperspectral imaging up to now are presented and commented. The literature review ends up with some conclusions made after the extended review process.

Keywords: NIRS; hyperspectral; spectroscopy; chemometrics; double integrating spheres; potatoes.

2.1. NIR technology

The use of near-infrared (NIR) spectroscopy (NIRS) for the analysis of food products began in the 1960s by the research group of Karl Norris in the U.S. Department of Agriculture (USDA) when studying the moisture of grain and seeds (Hart *et al.*, 1962; Massie & Norris, 1965; Norris & Hart, 1965).

This technique started to be introduced in the 1970s in various industries as an alternative to chemical and biological traditional methods (Cozzolino *et al.*, 2001; Murray, 1986; Norris *et al.*, 1976; Williams, 1975).

Nowadays, NIRS is considered to be an advanced spectroscopic technique for quality assessment of food in a non-destructive way (Magwaza *et al.*, 2012). Since its first application, NIRS technique has been successfully used for the quantitative and qualitative analysis of many agricultural and food products (Davies & Grant, 1987; Irudayaraj & Gunasekaran, 2001; Lister *et al.*, 2000; Nicolaï *et al.*, 2007).

2.1.1. Theory

Infrared radiation was first discovered by Herschel in 1800 while studying the colour responsible for the sunlight heat within the visible spectrum. With this aim, he carried out an experiment with a thermometer and a prism to separate the colours of the sunlight and he could observed that there was a rise in temperature when he placed the thermometer below the red end of the spectrum. He named this (invisible) radiation infrared radiation (IR) (Burns & Ciurczak, 1992).

This IR region is defined as the segment of the electromagnetic spectrum ranging from 780 to 1,000,000 nanometres (nm). Moreover, different sections can be distinguished inside this IR, depending on the wavelength range considered (Burns & Ciurczak, 1992). Thus, NIR is the referred type of radiation that extends between 780 and 2500 nm. This range corresponds to wavenumbers between 12,820 and 4,000 cm^{-1} (Williams & Norris, 2001) lying between visible light with shorter wavelengths and the mid infrared (MIR) with longer wavelengths (figure 2.1).

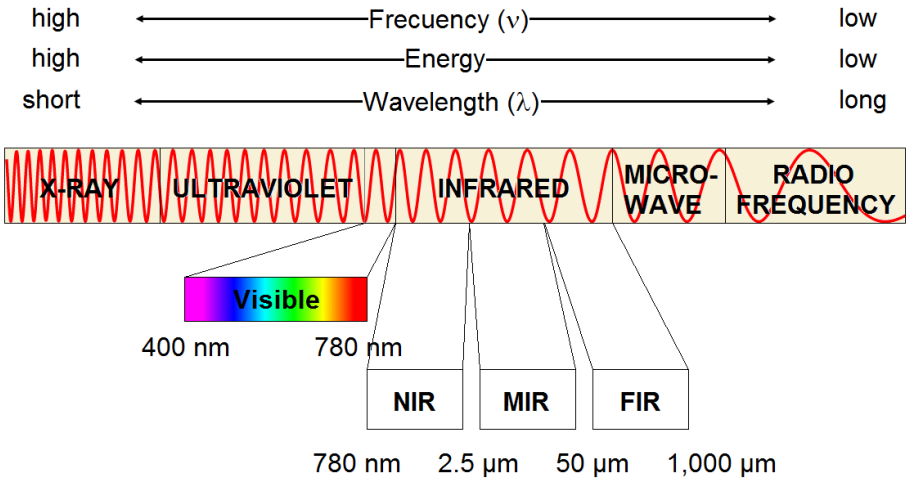


Figure 2.1. Electromagnetic spectrum (Adapted from Pavia *et al.* (2008)).

Regardless, this IR region remained of little interest until a century later when W. W. Clobentz built an IR spectrometer with a rock salt prism. He needed a full day to acquire an spectrum while a full spectrum in the NIR region can be obtain in less than a minute for more than 20 years now (Burns & Ciurczak, 1992).

Near-infrared spectroscopy studies the interaction between electromagnetic radiation and matter. It consists in the radiation of a sample with one or more wavelength bands between 780-2500 nm. This radiation penetrates into the sample and interacts with the bonds of the molecules present in it, especially with those of the form -CH, -NH, -OH and -CO that will absorb light according to their specific vibration frequencies (Burns & Ciurczak, 1992). The main modes of molecular vibrations are two, stretching and bending. Stretching involves movement along the axes of bonds, making the distance between atoms to change rhythmically and it can be symmetric or asymmetric, while bending result in a change in bond angle and can occur either in the plane of the molecule or out of it. There is IR absorption by a molecule only when vibrations cause a change in the dipole moment of the molecule (Miller, 2001) and when this occurs, the amplitude of vibration of the chemical bonds is increased. When light is absorbed, it cannot be measured directly as light; therefore, the transmittance or diffuse reflectance are common measurements used to correlate with absorbance through the following equations (Dahm & Dahm, 2007):

$$A = \log(1/T) \quad (1)$$

$$A = \log(1/R) \quad (2)$$

where A is the absorbance, T is the transmittance, or the quotient of the intensity of the transmitted radiation and the incident radiation, and R is the reflectance, or the ratio of the intensity of light reflected from the sample and that reflected from a background or reference reflective surface.

The interaction between energy and matter follows the Beer-Lambert´s Law. According to it, absorbance at any wavelength is proportional to the number or concentration of absorbing molecules present in the path of the radiation (Burns & Ciurczak, 1992; Osborne & Fearn, 1986; Osborne *et al.*, 1993; Shenk & Westerhaus, 1993; Williams & Norris, 2001). Therefore, Beer-Lambert equations can be written as follows (Dahm & Dahm, 2007):

$$A = \varepsilon\, c\, l \tag{3}$$

where ε is the Beer-Lambert absorption coefficient, sometimes referred to as α or κ; c is the amount concentration of absorber, that it can also be expressed as mass concentration (ρ); and l is the path length of the absorber (Bertie, 2002).

A NIR spectrum is the graphical representation of the intensity of light transmitted or reflected by the sample at the different wavelengths. This region is characterized by overtone and combination bands of the fundamental molecular vibrations (Siesler *et al.*, 2008). Therefore, a NIR spectrum is shown as a combination of overlapping bands difficult to interpret. This makes necessary the use of multivariate methods in order to extract information from NIR spectra (Naes *et al.*, 2002). Those multivariate methods will be explained in section 2.3.

2.1.2. Instrumentation and measurement modes

There are four different measurement setups commonly used for NIR analysis: transmittance, reflectance, transflectance and interactance (Shenk & Westerhaus, 1993), shown in figure 2.2. In transmittance mode, used for gaseous, liquid, semi-liquid and solid samples, the detector is located at the opposite site of the incident light and it measures the light transmitted through the sample. In reflectance mode, detector measures both the specular reflectance from the sample, with no information of its internal composition, and the diffuse reflectance, that is, the reflected light that has interacted with the internal particles of the sample. This measurement mode is used for solid and semi-solid samples. Transflectance mode is a combination of the previous two, in which a reflectance surface is located behind the sample so when light crosses the sample, it reflects on the surface and is forced to cross the sample again to reach the detector. In interactance mode, the detector is separated from the light source but both are in direct contact with the surface of the sample (Osborne, 2000). The main difference of the latter with the reflectance mode is that the interactance setup avoids measuring specular reflectance.

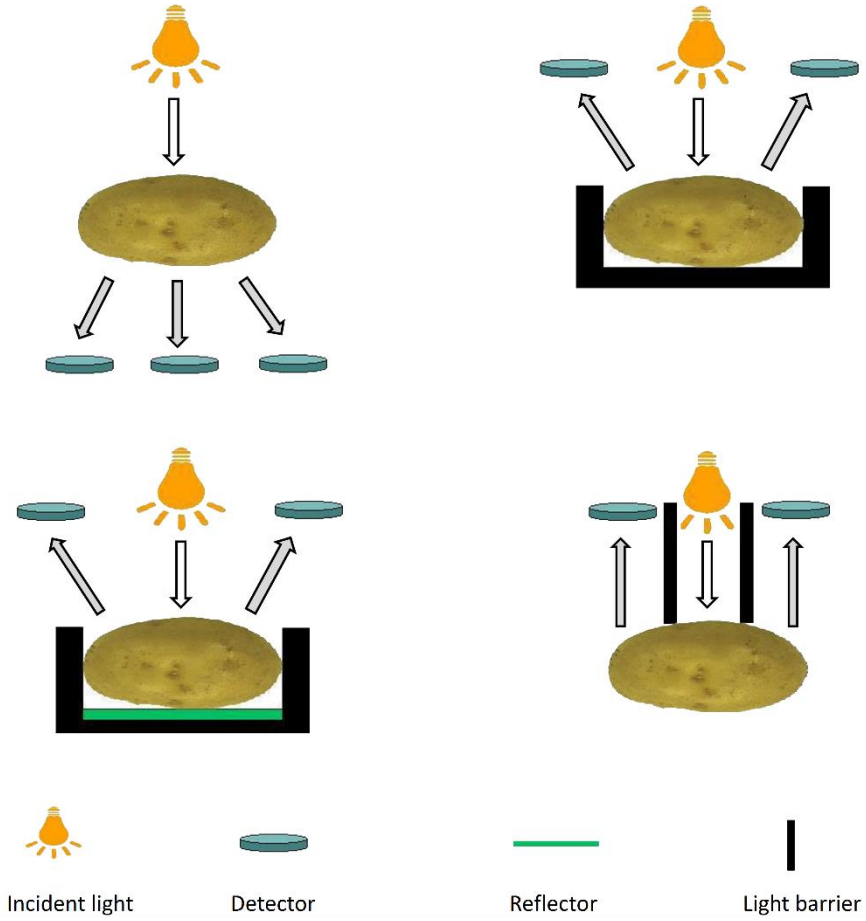


Figure 2.2. Measurement setups: (a) transmittance, (b) reflectance, (c) transflectance, and (d) interactance.

A typical NIRS instrument consists of mainly 5 components: a radiation source, a wavelength selection device such as monochromator, a sample holder, a photoelectric detector to measure the intensity of detected light to convert it to electrical signals, and a system to acquire and processes spectral data (Lin *et al.*, 2009).

- Radiation source: there are two main types of radiation sources used in spectroscopy, continuum sources or line sources. Continuum sources emit light with a continuous intensity over a wide wavelength range (Robinson *et al.*, 2005). Tungstone halogen lamps are the most common continuum sources used in NIR spectrometers because they can provide high intensity at low cost. However, they generate large amount of heat and have a relatively short life time (Cen & He, 2007).

  Line sources are generally used for applications where the full spectrum is not required. This type of source includes lasers, laser diodes and Light Emitting Diodes (LEDs) (McClure *et al.*, 2002).

- Wavelength selection device: according to the type of wavelength selection device NIR instruments are generally classified in two groups: filter-based instruments and monochromator-based instruments (Davies, 1999; Wetzel, 2001).

  Monochromators are used to disperse light with wide range of wavelengths into monochromatic light at different wavelengths. The elements that compose a monochromator are: a dispersion element for instance prisms or gratings, an entrance and an exit slit, and lenses or mirrors. In these instruments, NIR light enters through the entrance slit, penetrates into the dispersion element where is dispersed into the different wavelengths and then, the exit slit allows a very narrow wavelength to go through and reach the detector.

  In filter-based instruments, two types of filters are used to select wavelengths of the light, optical interference filters and electronically tunable filters. Interference filters are simpler and cheaper and are composed of thin layers with different refractive properties (Wetzel, 2001). In the electronically tunable filters, the spectral transmission can be electronically adjusted. There are two types of filters within this configuration, liquid crystal tunable filter (LCTF) and acousto-optical tunable filter (AOTF). The LCTF consists of several birefringent liquid crystal plates sandwiched between two linear polarisers. An AOTF system, consists of a crystal in which acoustic waves at radio frequencies (RF) are used to separate a single wavelength of light from a broadband source (Gat, 2000). In this type of filter, by adjusting the frequency of the RF applied to the crystal, the wavelength of the filtered light can be selected (Lin *et al.*, 2009).
- Sample holder: is the place where the sample is positioned and where the interaction between NIR and the object of study occurs. Different configurations have been developed by the manufacturers for the successful adjustment of the different type of materials. A correct design is essential in order to measure good NIR spectra.
- Detector: is the sensor that measures the intensity of NIR radiation and converts it into electrical signals. According to their operation, two main types of detectors are distinguished, thermal detectors and photon-sensitive detectors or photodiodes. Thermal detectors include those devices based on deuterated triglycine sulfate (DTGS) along with thermocouples, bolometers and thermistors among others, while photon-sensitive detectors comprise silicon photodiode, indium gallium arsenide (InGaAs), lead selenide (PbSe), mercury cadmium telluride (MCT) and indium antimonide (InSb) detectors (Robinson *et al.*, 2005). The detectors most commonly used in NIR systems are the PbSe, InGaAs and silicon photodetectors (Lin *et al.*, 2009).

- System to acquire and process data: this is normally a microprocessor or computer capable of processing and analysing spectral data using specific software. There are currently many types of user-friendly software specifically designed for spectral data acquisition and subsequent multivariate statistical analysis (Lin *et al.*, 2009).

More designs and configurations of NIR spectrometers can be found in several contributions to books as those by Workman and Burns (1992), McClure (2001) and Wetzel (2001).

The different measurement modes explained above and the instruments currently available for analysis provide NIR with many advantages:

- It is a non-destructive and non-invasive technique.
- Very little sample preparation is needed.
- Analysis can be performed very fast.
- It is environmentally friendly since it does not generate any residues.
- It is very versatile: the same technology can be used for different products.
- It allows cost-effective analysis.
- A single NIR spectrum contains both physical and chemical information from a sample.

However, this method presents some drawbacks that can affect its implementation, such as:

- Its dependence on reference methods based on chemical analysis.
- The relatively high price of the equipment.
- The difficulty in transferring models from one instrument to another.
- Its limitation since the results obtained are only applicable to samples with similar characteristics that those used to develop them.
- The requirement of qualified personnel for multivariate analyses.

2.1.3. NIR light interaction with biological tissues

Absorption in the NIR spectral range is related with the physical and chemical characteristics of samples; however, light scattering is due to the microstructure and texture of the tissues analysed (Chen & Thennadil, 2012; Saeys *et al.*, 2008).

In potatoes, as in many other agricultural foods, both skin and flesh are highly structured tissues at the microstructural level (Aguilera, 2000), causing multiple scattering of the light that crosses through them (Saeys *et al.*, 2008). NIR spectroscopy is generally applied on potatoes to study the properties of their flesh; however,

radiation has to travel through the skin first and therefore, results can be influenced by its properties. In the case of thin-peeled food products like potatoes, some authors pointed out that the presence of the peel does not change the correlation between the spectral and chemical data (Dull *et al.*, 1989; Krivoshiev *et al.*, 2000). Even though, NIR spectra are often dominated by scattering effects and so, it is essential to study how light propagates through biological tissues in order to obtain good quality spectra when analysing intact products in a non-destructive way (Aernouts *et al.*, 2013; Saeys *et al.*, 2008).

Light propagation through biological tissues is determined by the bulk optical properties (BOP): bulk absorption coefficient (μ_a), bulk scattering coefficient (μ_s), reduced scattering coefficient (μ_s'), with physical units of cm^{-1} or mm^{-1}, and anisotropy factor (g). Bulk absorption and scattering coefficients are defined as the probability of respectively absorption and scattering per unit infinitesimal path length (Aernouts *et al.*, 2015). The optical properties of turbid media have to be estimated as cannot be measured directly (Aernouts *et al.*, 2013). The g and μ_s can be combined to obtain μ_s' (Aernouts *et al.*, 2013; Arimoto *et al.*, 2005; Lino *et al.*, 2003; Zamora-Rojas *et al.*, 2013a). BOP are requisite parameters in the Radiative Transfer Equation (RTE) that can be used to describe light propagation in homogeneous turbid media like potatoes. RTE is a complex integro-differential equation that accurately describes photon migration through biological tissues (Martelli, 2012). However, it does not exist a general analytical solution of it (Van Gemert *et al.*, 1995), so, a diffusion approximation of the RTE is often used because closed-form analytical solutions can be obtained (Prahl, 1995).

Biological tissues are highly diffuse and low absorbing media ($\mu_s >> \mu_a$). Scattering results from the many different diffracting interfaces present in the cells of the tissue. Scattering due to cells is mainly anisotropic (Schulz & Semmler, 2008). An anisotropy factor of 0 implies isotropic scattering, while a value of 1 denotes exclusive forward scattering. If BOP of potato tissues are known, mathematical methods can be applied to numerically model NIR light interaction with potatoes in order to design appropriate optical systems to monitor potato quality (Wang & Li, 2013).

Some authors have described different methods for measuring optical properties of biological tissues in the biomedical area (Cheong *et al.*, 1990; Tuchin, 2007). They can be classified into two categories: *ex vivo* and *in vivo,* differing on the way the sample is prepared (Kim & Wilson, 2010). The *ex vivo* method is based on the total reflectance and transmittance measurements of samples required to be of certain shapes and sizes, that are then combined with indirect iterative techniques to obtain the BOP

(Wang & Li, 2013). The *in vivo* method is normally based on measurements of diffuse reflectance or backscattering directly made on the sample in the time domain, frequency domain or continuous wave domain and used to approximate the RTE (Kim & Wilson, 2010).

Indirect techniques use a theoretical model of light propagation in a medium to solve the inverse scattering problem. They are divided into non-iterative and iterative models. Non-iterative techniques are based on the two-flux Kubelka-Munk (KM) model and multi-flux models (Bashkatov *et al.*, 2011).

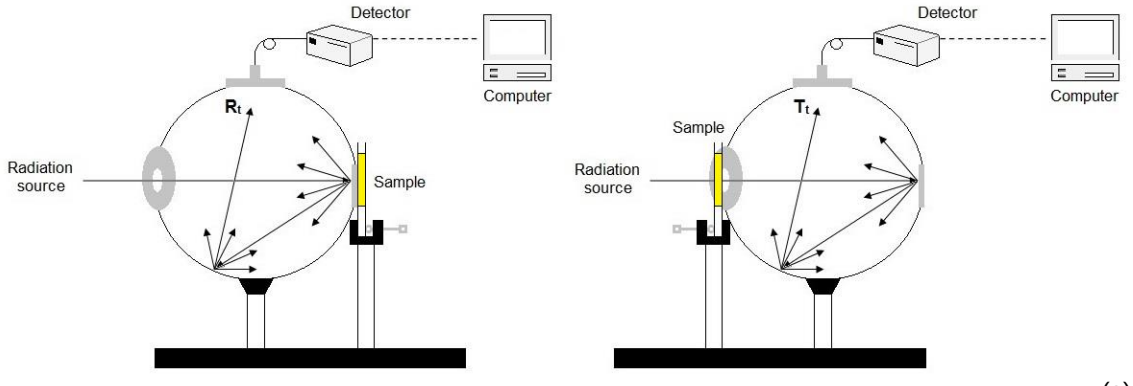
Iterative techniques used to retrieve BOP include the Inverse Adding-Doubling (IAD) and Inverse Monte Carlo (IMC) algorithms. Those are indirect iterative methods that require some experimental measures to calculate the optical parameters of a tissue:

- Total transmittance (T_t) for collimated or diffuse radiation;
- Total reflectance (R_t) for collimated or diffuse radiation;
- Absorption by a sample placed in an integrating sphere;
- Unscattered/collimated transmittance (T_c);
- Angular distribution of radiation scattered by the sample.

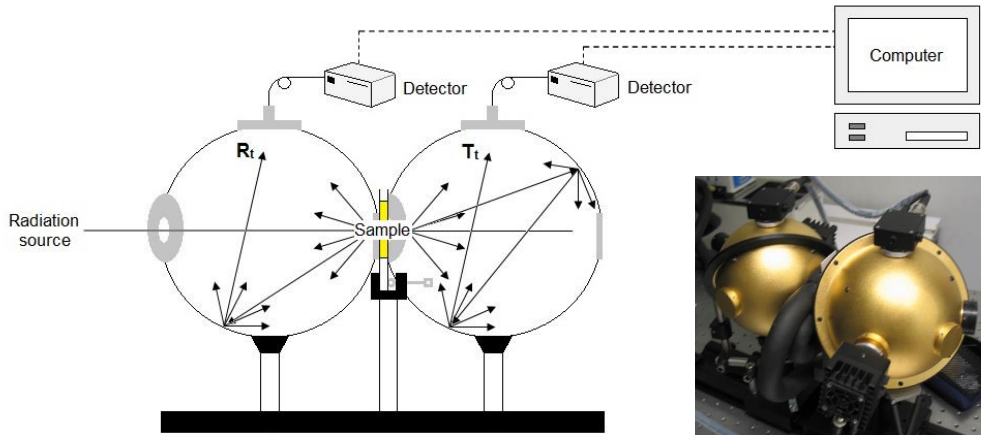
Usually, by measuring T_t and R_t, and assuming a value for the anisotropy factor, absorption and reduced scattering coefficients (μ_a and μ_s') can be determined (Saeys *et al.*, 2008). Moreover, if measuring any three of the five parameters, all optical properties can be calculated (Cheong *et al.*, 1990) with no need for anisotropy factor assumption (Tuchin, 2007) by using iterative methods like IAD or IMC. In addition, by an accurate measurement of the unscattered transmittance, the bulk extinction coefficient (μ_t) is obtained and all bulk optical properties (μ_a, μ_s and g) can be derived (Prahl, 2011). The unscattered transmittance should be measured in a configuration especially optimized to minimize the collection of scattered photons (Aernouts *et al.*, 2013; De Vries *et al.*, 1999; Prahl, 2011). Otherwise, the unscattered transmittance could be overestimated, leading to an underestimation of μ_t (Aernouts *et al.*, 2013; Prahl, 2011). Since optical properties depend on the wavelength range considered, total reflection, total transmittance and unscattered transmittance measurements should be made for all the wavelength range of interest.

Measurements of T_t and R_t are generally accomplished by the use of integrating spheres (IS). These spheres are characterized by an spherical internal design with a surface covered by a highly reflectance material (Hanssen & Snail, 2002). Therefore, when incident light strikes any part of the sphere, it is uniformly distributed to the entire surface. If a detector is located on that surface, it will receive a flow directly proportional to the incident light of the sphere, independently of its location. A single IS setup

consists of one sphere that can only measure one parameter (T_t or R_t) at the same time; while a double integrating spheres (DIS) system comprises two spheres allowing simultaneously measurements of both T_t and R_t (Pickering *et al.*, 1993). Such systems provide more robust and accurate estimates of the bulk optical properties compared to single integrating sphere measurements. DIS systems are considered the "golden standard" for *ex vivo* optical characterization of biological tissues (Bashkatov *et al.*, 2005a; Pickering *et al.*, 1993; Prahl, 2011). Figure 2.3 shows the different configurations of IS: single (a) and double (b).



(a)



(b)

Figure 2.3. The schematic of the system configuration for measuring total reflectance (R_t) and total transmittance (T_t) in: (a) single integrating spheres and (b) double integrating spheres. (Adapted from Wang and Li (2013)).

After measuring total reflectance, total transmittance, or any of the other experimental measures previously mentioned, using one of the iterative methods in their inverse form, optical properties can be obtained.

The adding-doubling (AD) method allows calculating T_t and R_t very accurately, while maintaining a high degree of flexibility and time efficiency. This method uses the RTE to obtain the reflectance and transmittance for a single 'infinitesimally' thin sample layer. Accordingly, this layer is 'doubled' and T_t and R_t of this doubled layer are

calculated. Finally, this process is repeated until the desired thickness of the homogeneous sample is reached (Saeys *et al.*, 2008; Zamora-Rojas *et al.*, 2013a). As the AD method calculates T_t and R_t for a tissue layer with known bulk optical properties, it has to be inverted to allow extraction of the bulk optical properties (μ_a and μ_s') from DIS measurements. The inverse AD (IAD) algorithm, developed by Prahl *et al.* (1993), estimates the optical properties by iteratively changing their values in the AD simulations until the simulated reflectance and transmittance values match with the measured ones (Prahl, 2011; Saeys *et al.*, 2008; Zamora-Rojas *et al.*, 2013b).

The Inverse Monte Carlo (IMC) algorithm is more flexible than the IAD and allows calculations for more complex geometries and closer to the real structure of the tissue (Bashkatov *et al.*, 2011). The procedure is based on the Monte Carlo algorithm, developed by Wang and Jacques (1992), applied inversely. Monte Carlo approach is a statistical method that simulates photon migration in random media by sending photon "packets" to a sample and tracing their trajectories (Yamada, 2000). It is based on the principle of the Radiative Transfer Theory (RTT); in which the absorption and scattering coefficients are defined as a probability of a photon to be absorbed per unit length, and the probability of a photon to be scattered per unit length respectively (Bashkatov *et al.*, 2011).

In the IMC, from the experimental measures of R_t and T_t, optical properties are estimated by iterative calculations that follow the same structure that simulation of light propagation by Monte Carlo (Yamada, 2000).

In general terms, all these approaches are focused on providing solutions to the RTE that describes light propagation in a turbid media, in order to design more efficient sensors for non-destructive quality evaluation of food products.

Optical properties of many agricultural and food products have been an object of research in last years as it is explained in chapter 3; however, only a few authors have focused on studying the optical properties of potatoes. Karagiannes *et al.* (1989) carried out a study for the determination of absorption and reduced scattering coefficients of potatoes, assuming a fixed value for *g*. The results of that study are further described in chapter 3.

2.2. Hyperspectral imaging

The potential of NIRS has extended further with the use of multispectral and hyperspectral imaging systems that have the advantage of including spatial information along with spectral data (Lu, 2003; Martinsen & Schaare, 1998).

While conventional Red-Green-Blue (RGB) colour cameras acquire three monochromatic, broad-band images to obtain a colour image; both multispectral and hyperspectral are systems able to acquire a set of monochromatic images at a few wavelengths. However, while in the former the number of spectral bands is usually fewer than 10 and do not necessarily have to be continuous nor narrow, hyperspectral systems involve a large number of narrow bands (100 to 250) over a continuous spectral range (ElMasry & Sun, 2010). Therefore, hyperspectral images are composed of hundreds of contiguous wavebands for each spatial location of a sample. As a result, each pixel in a hyperspectral image comprises the spectrum of the corresponding position (Gowen *et al.*, 2007). With multispectral systems on the other hand, the spectra of the pixels cannot be obtained (ElMasry & Sun, 2010).

Hyperspectral imaging technique was originally developed for remote sensing (Goetz *et al.*, 1985), but has since found applications in other research fields, such as precision agriculture (Erives & Fitzgerald, 2005; Muhammed, 2005) or food quality control (Park *et al.*, 1998; Sun, 2010).

Hyperspectral images offer more information and a more reliably characterization of a product than that by a single monochromatic image or a conventional RGB image (ElMasry *et al.*, 2012b). The number of images acquired depends on the spectral resolution of the system used. Those images are combined forming a three-dimensional data cube (hypercube) containing two-dimensional spatial information as well as spectral information. Figure 2.4 illustrates one example of the hypercube extracted from a hyperspectral image of a potato. The raw hyperspectral image consists of a series of contiguous sub-images; each one representing the intensity and spatial distribution of the sample at a certain waveband.

Hypercubes can provide physical and chemical information of the product examined, such as size, colour, texture, along with water, fat, and other constituents (Lawrence *et al.*, 2003). However, these systems have a disadvantage as they produce very large data volumes that suffer from collinearity problems. Such data volumes complicate the storage, management, and further image processing and analyses (Sun, 2010).

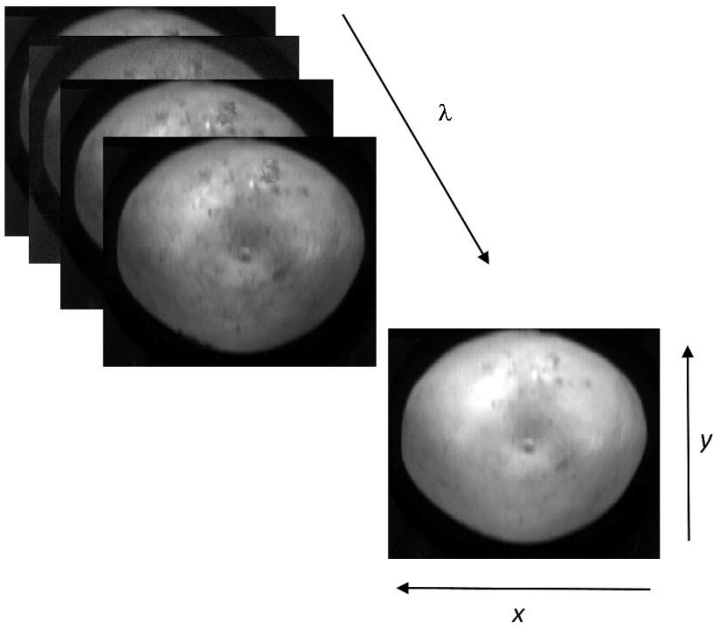


Figure 2.4. Hyperspectral image cube of a potato. The two spatial dimensions are *x* and *y*, and the spectral dimension is λ.

A hyperspectral imaging system consists of three main elements: light sources, a detector and a wavelength selection device.

- Light sources: halogen lamps are the most common type of illumination sources used in visible (VIS) and NIR spectral regions (Qin, 2010). They must be placed in an arrangement in which diffuse illumination is provided achieving a uniformly illumination of the object (Gómez-Sanchis *et al.*, 2008b).
- Detector: a two-dimensional detector, typically a monochromatic charge-coupled device (CCD) or complementary metaloxide semiconductor (CMOS) camera to simultaneously collect the spectral and spatial information are generally used (ElMasry & Sun, 2010).
- Wavelength selection device: such as an imaging spectrograph (ElMasry & Sun, 2010), liquid crystal tunable filters (LCTF) (Evans *et al.*, 1998; Gómez-Sanchis *et al.*, 2008a; Wang *et al.*, 2012); or acousto-optic tunable filters (AOTF) (Bei *et al.*, 2004; Vila-Francés *et al.*, 2010).

Depending on the type of wavelength selector employed, the selection of the wavelengths can be performed by dispersing the light beam into individual wavelengths (e.g. spectrograph) or by filtering the radiation allowing only the desired wavelength to reach the image detector (e.g. LCTF and AOTF).

Depending on the detector type, hyperspectral images can be acquired in the ultraviolet (UV), NIR or IR regions of the electromagnetic spectrum (ElMasry *et al.*, 2012b). The most widely system used in food analysis applications is in the VIS and VIS-NIR range, 380-800 nm or 400-1000 nm. Nowadays, there are hyperspectral

systems in the range 900-1700 nm that provide the accuracy required in most challenging applications in food analysis. Moreover, numerous manufacturers are currently producing hyperspectral imaging systems that cover the shortwave-infrared (SWIR) region (900-2500 nm) in order to be used in many applications in food and agricultural analyses, chemical imaging, and process analytical technologies (ElMasry & Sun, 2010). Since hyperspectral systems can be manufactured with a sensitivity up to about 2500 nm, both the wavelength selector and the image sensor must be sensitive to the same spectral range (Cubero *et al.*, 2011).

The most common mode of hyperspectral imaging acquisition is reflectance, although measurements can be carried out in transmission or fluorescence modes as well (Gowen *et al.*, 2007).

There are three conventional configurations of hyperspectral imaging systems: area scanning, point scanning, and line scanning. As these systems capture only a one- or two-dimensional subset of the hypercube, they require scanning to obtain the remaining components (ElMasry & Sun, 2010)

- Area-scanning design: it is also known as focal plane scanning imaging. In this configuration the image field of view is kept fixed, and the images are obtained one wavelength after another. This mode requires a tunable filter to acquire an image at different wavelengths and the resulting hypercube data is stored in Band Sequential (BSQ) format.
- Point-scanning method: it is also known as whiskbroom. This setup measures the spectrum of a single point; then the sample is moved to take another spectrum and so on to produce a complete hyperspectral image. Hypercube data obtained with this configuration are stored as Band Interleaved by Pixel (BIP) format.
- Line scanning: it is also called pushbroom. Within this configuration a whole line of an image rather than a single pixel is recorded by an array detector; and the resultant hypercube is stored in the Band Interleaved by Line (BIL) format. This is the most appropriate method for conveyor belt systems used in many food industry applications.

Whiskbroom and pushbroom scanning methods are considered spatial-scanning methods as the object is scanned in the spatial domain by moving the object point-by-point or line-by-line respectively. Conversely, area scanning are spectral-scanning methods (ElMasry & Sun, 2010).

Once configuration, mode of acquisition and illumination have been chosen and a hyperspectral image for a tested sample has been acquired, there are a number of steps to follow to extract information from that image (ElMasry & Sun, 2010). Firstly,

this image should be calibrated with the help of white and dark hyperspectral images. Then, the spectral data are extracted from different regions of interest (ROIs) in the calibrated image. These spectral data should be pre-processed in order to reduce noise, improve the resolution of overlapping data, and to minimize contributions from imaging instrument that no represent features of the sample being analysed. Finally, once spectral data have been pre-processed, multivariate analysis can be employed to extract information from the hyperspectral data (section 2.3).

2.3. Chemometrics

The term *Chemometrics* was first introduced more than 40 years ago by Wold (1972). It was later defined by Massart *et al.* (1988) as "the chemical discipline that uses mathematical, statistical, and other methods employing formal logic, (1) to design or select optimal measurement procedures and experiments, and (2) to provide maximum relevant chemical information by analysing chemical data" (Massart *et al.*, 1988, p. 5). In a paper published by Wold and Sjöström (1998), authors highlighted that after almost 30 years, chemometrics had become an essential application not only in chemistry but also in other fields.

As explained in section 2.1.1, it becomes challenging to extract information from spectral data, especially in the NIR region, due to the difficult interpretation of overlapping bands. A NIR spectrum is a complex data matrix resulting from the combination of different compounds and properties of the sample studied. There is a need for some chemometric tools capable of relating the electromagnetic information (spectrum) with the information of the physical and chemical composition (reference method), or with the different categories or classes of interest using mathematical algorithms through the application of different statistical models. This process is known as development of quantitative or qualitative models (Martens & Martens, 2001). Multivariate analysis method is referred to as one of the most commonly applied group of chemometric tools for the analysis of spectral data (Wold & Sjöström, 1998).

Prior to model building, spectral pre-processing techniques are commonly applied to the spectral data to avoid the influence of unwanted effects such as light scattering that can negatively affect the reliability of the multivariate model (Barbin *et al.*, 2012). Several pre-processing methods have been developed for this purpose (Martens & Naes, 1989; Naes *et al.*, 2002). In this context, the techniques which have shown to be effective in classical spectroscopy are also frequently applied to hyperspectral imaging data (Vidal & Amigo, 2012). The principal spectral pre-processing techniques are smoothing, derivatives and scatter correction methods.

The smoothing technique allows removing some of the instrumental noise. Among the different algorithms available, Savitzky-Golay (SG) is the most popular for this purpose (Amigo, 2010).

Both Standard Normal Variate (SNV) and Multiplicative Scatter Correction (MSC) are techniques capable of reducing the effects of light scattering on the acquired spectra (Rinnan *et al.*, 2009), which typically provide similar results. SNV is a method of spectral normalization, which establishes a common scale for all spectra by centering each spectrum on its mean value and scaling it by its standard deviation. In this way, it corrects for additive and multiplicative variation between spectra. MSC estimates the multiplicative and additive effects within a set of data by regressing each spectrum onto a reference spectrum, which is typically the mean spectrum. The spectrum is then pre-processed by subtracting the estimated intercept value and dividing by the estimated slope value (Dhanoa *et al.*, 1995). Detrend (DT) is an algorithm normally applied in conjunction with SNV. Detrend fits a least square quadratic polynomial to each spectrum previously corrected by SNV and it then generates a new spectrum as the difference between the SNV-corrected spectrum and the polynomial (Delwiche & Reeves, 2004; Guthrie *et al.*, 2005).

Other scatter correction techniques include normalization which is similar to SNV and MSC, and, orthogonal signal correction (OSC) that removes the variability in the spectral signal orthogonal to the information of interest (Wold *et al.*, 1998).

As for derivative transformations, the most habitually applied is the Savitzky-Golay algorithm that simultaneously performs a spectral smoothing and calculates the derivative of the data (Savitzky & Golay, 1964). Both first and second derivatives remove baseline offsets in the data, while the latter is also useful for separating overlapping peaks (Burger & Geladi, 2007).

Another commonly applied pre-processing method is Mean Centre (MC), a basic technique that reduces the systematic noise (Barbin *et al.*, 2012). By MC, the average value of each variable is calculated and then subtracted from the data.

Spectral data, either being pre-processed by any of the above methods or raw, can be used in both quantitative and qualitative analyses. However, before model building, samples can be analysed by means of exploratory methods such as Principal Component Analysis (PCA).

PCA is a method able to extract the main sources of variability in the data (Amigo *et al.*, 2013). It transforms the variables into Principal Components (PCs) which are linear combinations of the spectral data describing most of the variation in the original variables (Kamruzzaman *et al.*, 2012). The first PC is defined as the linear combination

of the original variables which captures the largest part of the variation in the data, the second captures as much as possible variation orthogonal to the first PC, and so on (Otto, 1999). The number of principal components is usually less while it could be equal to the number of original variables.

PCA helps reducing redundant information in the spectral data and it can also be used to identify outliers, these are observations that are extreme or that do not fit the PCA model. It is important to observe them since outliers can negatively influence the predictive capacity of models and the interpretation of results (Bro, 2003). Therefore, detection, interpretation, and possible elimination of those outliers is an essential stage in multivariate analysis of spectral data (Naes *et al.*, 2002). There are different procedures for their detection based on the model residuals (Q) and Hotelling's (T^2) obtained in a PCA analysis; or those based on the measure of the distance between samples or to the centre of a distribution in a multivariable space. Mahalanobis distance, or a scaled version (*H*), measures the distance of a data point from the centre of a distribution (Krzanowski, 2000). A *H* distance of 3.0 from the mean indicates an outlier sample and is used as a boundary to exclude that sample from the population (Shenk & Westerhaus, 1991). However, it should be mentioned that Mahalanobis distance measurements work well when data have a normal distribution but they may not work when data is non-normally distributed, since Mahalanobis distance could fail to reflect the real position of the points (Fearn, 2011).

The model Q residuals are the X-variation that was not captured by the PCA model, and the Hotelling's T^2 measures the Mahalanobis distance of the corresponding vector from the sample mean vector (Williams *et al.*, 2006).

2.3.1. Quantitative analysis

Quantitative analysis models correlate spectral information with concentration or percentage of one or more compounds in a sample or with a physical parameter of a product. These algorithms can be classified as linear and non-linear. A brief description of the most common algorithms used in multivariate analysis of spectral data is given below.

2.3.1.1. Linear algorithms

The main statistical techniques for this process are the Multiple Linear Regression (MLR), Principal Component Regression (PCR) and Partial Least Squares (PLS) Regression. These chemometric techniques establish a linear mathematical relationship between spectral data and the parameter of the sample studied. This

relationship can then be used to predict the parameter value in unknown samples. This procedure assumes that the correlation between analyte or property to be measured and its absorbance follows Beer's law (Agelet & Hurburgh, 2010).

- MLR: this algorithm calculates a regression model taking into account all spectral variables available and the reference values (Sternberg *et al.*, 1960). Thus, it does not account for wavelength multicollinearity. If variables are correlated, like in spectral data, the resulting calibration is unstable and provides several solutions. Therefore, this method is only appropriate if non-correlated wavelengths are selected (Agelet & Hurburgh, 2010).
- PCR: in this method, the spectral data matrix is decomposed by a PCA, that generates new non-correlated orthogonal variables (PCs) and a regression process by least squares with the reference data matrix is accomplished. The main drawback of this algorithm is that it only uses spectral data to calculate the new variables being the ones that best represent X data matrix but they may not be appropriate to estimate Y data matrix (Vandeginste *et al.*, 1998).
- PLS: in PLS regression, orthogonal linear combinations of the original variables are defined, which maximally capture the covariance between the X and Y variables. These linear combinations are referred to as Latent Variables (LVs) or PLS components. The main difference with the PCR is that the new variables are calculated considering both data matrixes X and Y (Faber & Bro, 2002; Martens & Naes, 2001; Wold *et al.*, 2001). Two PLS algorithms have been used in chemometrics: PLS1 and PLS2 (Manne, 1987). PLS1 develops a PLS model for each of the parameters independently while PLS2 deals with several chemical variables at the same time.

2.3.1.2. Non-linear algorithms

Sometimes, the relationship between spectral and reference data is not linear. This may be due to multiplicative effects of the instrumental signal or due to the nature of the parameter object to study (Bertran *et al.*, 1999). Those non-linearities can occasionally be corrected with some data pre-treatment. However, when it is not possible to handle them, there is a need for specific non-linear algorithms. Some of these algorithms include Artificial Neural Networks (ANN) and Support Vector Machines (SVM). These algorithms are beyond the scope of this work and are therefore not explained here, but more information about them can be found in a review published by Pérez-Marín *et al.* (2007) and papers such as those by Devos *et al.* (2009) and Marini (2009).

2.3.2. Qualitative analysis

Qualitative analysis consists of comparing spectra from unknown samples with spectra from samples with known characteristics (Mark, 1992). This type of analysis is generally used to classify samples according to categorical variables like for example their variety, geographical origin or presence of defects. All classification methods or pattern recognition techniques are group in two classes: supervised and non-supervised learning algorithms and they differ in the knowledge about the objects to be grouped. Thus, non-supervised algorithms do not require any previous knowledge about the samples to be grouped (Heise & Winzen, 2002).

2.3.2.1. Non-supervised methods

The main objective of non-supervised algorithms is to identify sources of grouping in the available samples considering only the information from their spectral data (Downey, 1996; Naes *et al.*, 2002). Such methods are normally applied in the early steps of data exploration. Cluster analysis is an example of a non-supervised method (Massart *et al.*, 1997).

- Cluster: this method can be either performed visually or by one of the hierarchical methods. These methods identify samples that are close to each other using distances between the objects, generally by the Euclidean or Mahalanobis distances. As a result, a dendrogram is created which is a graphical representation of the different groups identified (Naes *et al.*, 2002).

2.3.2.2. Supervised methods

Supervised methods are commonly referred to as discriminant analysis. In these methods, apart from the spectral information of samples, there is information about the groups or categories existing in the sample set. These methods correlate spectral information with information of the categorical variable, building classification rules for the number of predefined classes. These rules are then used to assign new and unknown samples to the most probable class (Naes *et al.*, 2002). K-nearest Neighbour (KNN), Linear Discriminant Analysis (LDA), Soft Independent Modelling of Class Analogies (SIMCA) and Partial Least Squares Discriminant Analysis (PLS-DA) are examples of supervised techniques.

- KNN: this is one of the simplest classification methods. It is commonly based on the Euclidean distance between samples. This approach uses the training data at

the test time to make predictions looking at the K most similar training samples (Naes *et al.*, 2002).

- LDA: this method looks for linear combination of features that best explained the data and can separate two or more classes. LDA is based on the assumption of equal covariance matrix for each group (Naes *et al.*, 2002).
- SIMCA: is a supervised classification technique that has been successfully applied to solve many pattern recognition problems (Massart *et al.*, 1988). Being a supervised method, SIMCA requires knowledge on the class membership of the samples in the training set. Therefore, a classification model is built by using a training data set of samples with known class affiliation and is then evaluated using external samples (Martens & Naes, 1989). In SIMCA, a separate PCA model is built for each class. Samples are projected onto the different PCA models and a metric is used which combines the distance from the model (Q-residual) with the distance from the centre of the model within the model (Hotelling's, T^2) in order to calculate class membership. As a consequence, it is theoretically possible that samples are classified in multiple classes or in none.
- PLS-DA: is a pattern recognition technique where the class memberships are predicted from the sample spectra by means of PLS regression (Höskuldsson, 1988; Wold, 1966). In order to be able to use PLS regression for discrimination purposes, the class variable must be transformed into a binary-coded dummy matrix with the same number of rows as X and the same number of columns as there are classes. Thus, the first column of Y will be a vector with all values equal to zero except for the samples belonging to the first category that will be equal to 1. Then, in the same way as for the regression method, the model will give a calculated Y that will not have either 1 or 0 values perfectly. So, a threshold has to be defined to decide if an object is assigned to the category or not (Ballabio & Todeschini, 2009).

2.4. Application of NIRS for the analysis of potatoes

Although literature concerning NIR applications in potato is not as extended as in other type of vegetables (Fernández-Ahumada *et al.*, 2006), NIR applications in this industry were initiated in the 1980s. For this reason, it is important to compile the information that has been published relating NIRS applications to the quality determination of potato and potato products since its beginning.

2.4.1. Internal composition determination

Water, dry matter and starch: One of the first applications of NIR in the potato industry was to measure the moisture content of chips. McDermott (1988) reported good results with a correlation coefficient (R) of 0.95 and a standard error of estimation of 0.15.

Similar studies have been conducted since then on different commercial potato chips samples obtaining results in agreement to the previous report, with standard errors of prediction (SEP) between 0.20 and 0.26 and correlation coefficients between 0.95 and 0.98 when PLS regression (PLSR) statistic method was applied (Ni *et al.*, 2011; Shiroma & Rodriguez-Saona, 2009). Nevertheless, Ni *et al.* (2011) obtained better results for moisture prediction using Least Squares-Support Vector Machines (LS-SVM) and Kernel Partial Least Squares (KPLS) methods. The coefficients of correlations were 0.99 and 0.98 with a root mean square standard error of prediction (RMSEP) of 0.07 and 0.10 for LS-SVM and KPLS respectively. Based on the results obtained, the authors concluded that it was possible to determine moisture content of chips in a fast and accurate way.

Some authors have used NIRS as the unique method for the estimation of moisture content of potatoes in their on-line routine analyses. Broothaerts *et al.* (2007) determined the water content of freeze-dried samples of potatoes by NIRS. The investigation was focused in the development of a certified reference material for Genetically Modified (GM) potatoes with altered starch composition and the water content of both non-GM and GM potato samples was determined by using an acousto-optical tunable near-infrared spectrometer (AOTF-NIR) integrated on-line into their instrumentation. NIR data were evaluated by a PLS1 regression model based on meat calibrations previously evaluated by Kestens *et al.* (2008) and able to predict water content of the samples. Other authors have focused on the determination of DM and starch in potatoes. DM content provides information on both water and starch concentration (Subedi & Walsh, 2009). A great percentage of DM in potato is in form of starch and it is considered a very important constituent of this food since final quality of

potato products is directly related to this component. Moreover, the European Potato Starch Industry bases its payments to the farmers on the starch concentrations of tubers (Haase, 2003). In the decade of the 1980s the underwater weight (UWW) method was the best practical and non-destructive method for estimating DM content of potatoes. By this method, the specific gravity was calculated first, and then, using its correlation with DM, the latter was determined. Although those two compounds were known to be highly correlated (Wilson & Lindsay, 1969), there was a need for a much more rapid and accurate non-destructive technique. Therefore, from this decade through the 1990s and until today some authors have studied the correlation between spectral and both DM and starch content of tubers. The most common method used to establish the correlation was to combine NIR data with PLSR statistic method.

Table 2.1 shows the determination coefficients (R^2) as well as the RMSEP obtained for DM and starch content in potatoes reported by several authors. A wide range of potato varieties have been studied in diverse sample presentations: intact, mashed, freeze-dried, etc. Also, different wavelength regions have been used to develop the calibration models ranging from 734 to 931, 750 to 950, 800 to 1000, 1100 to 2500, 770 to 2500, 850 to 2500, 1000 to 2500 nm, or including the visible range 400 to 2500 and 460 to 1040 nm. It is known that starch have bands at 1200, 1700, 1720 and 1780 nm (Osborne *et al.*, 1993). Because of these facts, a wide range of RMSEP values have been obtained and consequently, it is difficult to establish either the type of sample presentation or a specific wavelength to predict the content of DM and starch. However, it seems that the lower RMSEP values for both components occurred when the sample is mashed and homogenized and the range between 1100 and 2500 nm is used (Brunt & Drost, 2003, 2010; Fernández-Ahumada *et al.*, 2006; Haase, 2011; Hartmann & Büning-Pfaue, 1998). DM content was also highly predicted in potato chips (Root mean square error of cross validation, RMSECV: 0.84) where the percentage of this component is significantly higher than in intact and mashed potatoes (Pedreschi *et al.*, 2010). Other authors have focused on the study of the optimum region in potato tubers to be scanned by NIRS in order to predict DM content of the whole tuber. As a result, Peiris *et al.* (1999) stated that DM content of potatoes was greater toward the surface of the tuber than at the centre. This result matched with that obtained by Scanlon *et al.* (1997) where the most closely correlated values were those of the centre outside section of the tuber. Helgerud *et al.* (2012) also obtained better results for the measurements taken at the centre of the longest axis in a study about the ability of NIR to rapidly estimate DM content of intact potatoes. They compared the performance of two different NIR instruments to the performance of the traditional specific gravity measurement method. First, they used a 1D NIR interactance equipment for stationary

analysis and then, a commercial 2D NIR interactance system to provide on-line estimation. The specific gravity was calculated based on equation (4).

$$\text{Specific gravity} = \frac{\text{Weight in air}}{\text{Weight in air - Weight in water}} \quad (4)$$

Then, following equation (5) provided by Lunden (1956), they calculated DM content.

$$\text{DM} = 215.73\ (\text{Specific gravity} - 0.9825) \quad (5)$$

A standard normal variate (SNV) pre-treatment was applied only to the data obtained by the 1D NIR interactance equipment.

Table 2.1 shows that authors achieved the lowest RMSECV and highest R^2 when the specific gravity method was employed. However, this could only be used with small sample volumes and took much more time than the other two systems capable of recording multiple spectra per second. Moreover, the underwater weight method was the one with the largest deviation among the three methods used. Nevertheless, a direct relationship exists between specific gravity and cooking quality of these tubers. Some authors have reported that specific gravity could be used as a direct measure of quality characteristics of potatoes (Bewell, 1937; Clark *et al.*, 1940; Komiyama *et al.*, 2002; Smith & Nash, 1940). Chen *et al.* (2005) examined the correlation between NIR spectroscopy and specific gravity of intact potatoes. 250 samples of potatoes from three different varieties were used for this study. Samples were scanned by NIR in interactance mode in the 700 to 1100 nm wavelength range, and the specific gravity of each sample was measured by using specific gravity measurement equipment based on the principle of liquid displacement. For the development of the calibration equation samples were divided into two groups: 150 for calibration model and 100 to validate it. PLSR method was applied in order to obtain the prediction models. They obtained correlation coefficient (R) values ranging from 0.93 to 0.94 for the prediction set of samples with SEP values between 0.0044 and 0.0047 g/cm^3. The highest R values and lower SEPs were those obtained for raw and normalized spectra. Based on the results obtained, authors concluded that NIR spectroscopy was able to accurately measure the specific gravity of intact potatoes.

Table 2.1. Overview of applications of NIR spectroscopy to measure dry matter and starch content of potatoes[a].

	Type of sample	Number of samples	Cultivar	Wavelength range (nm)	Pre-process	Validation	Analysis	Mode	Range %	R^2	RMSEP/SEP	References
DM & Starch	Mashed & homogenized	116 116	Granola & Nicola	1100-2500	1st der.	Cross & external	MPLS & PCA	Reflectance	15.6-21.0 10.0-14.8	0.97 0.93	0.19 0.28	Hartmann & Büning-Pfaue (1998)
	Mashed & homogenized	81	Nicola Irene Bintje	1100-2500	None	Full cross-validation	PLSR & PCA	Reflectance	16.9-30.2 10.8-21.8	0.86 0.90	1.66 0.96	Van Dijk *et al.* (2002)
	Freeze-dried	628	n/a	850-2500	None 1st der. 2nd der. None 1st der. 2nd der.	External	PCR	Reflectance	16.7-33.4 12.1-27.3 	0.98 0.98 0.97 0.96 0.96 0.95	0.518 0.514 0.568 0.638 0.622 0.678	Haase (2003)
	Mashed & homogenized	504	n/a	1100-2500	Smoothing & MSC	External	PLSR	n/a	19.5-29.2 14.2-23.4	0.93 0.84	0.22 0.39	Brunt & Drost (2003)
	Mashed	275	n/a	1000-2500	1st der. & MSC	Cross & external	PLSR	Interactance-reflectance	19.5-29.8 13.4-21.9	0.92 0.81	0.45 0.50	Fernández-Ahumada *et al.* (2006)
	Mashed & homogenized	219	18 different	1100-2500	Smoothing	External	PLSR	Reflectance	19.8-34.0	0.93 0.84	0.47 0.63	Brunt & Drost (2010)
	Mashed, homogenized & freeze-dried	2517	10 different	850-2500	SNV & DT	Cross & external	MPLS	Reflectance	14.1-35.4 12.8-27.2	0.99 0.96	0.39 0.47	Haase (2011)
DM	Intact Sliced	910 907	Russet Burbank	800-1000	2nd der.	External	LR	Transmittance	14.1-32.5	0.84 0.90	1.52 1.69	Dull *et al.* (1989)
	Intact	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	0.93	n/a	Scanlon *et al.* (1997)
	Flesh	200	Shepody & Russet Burbank	770-2500	2nd der.	External	MLR	Reflectance	n/a	0.77-0.58	1.3-1.5	Scanlon *et al.* (1999)
	Intact	49	n/a	734-931	None 2nd der.	Cross validation	PLSR & MLR	Reflectance	$\bar{x}$ = 13.4	0.79 0.76	1.04 1.09	Walsh *et al.* (2004)
	Intact Clean Peeled Sliced	100 100 100 100	Nicola, Spunta, Golden Delight, Sebago, Russet Burbank	750-950	2nd der.	Cross & external	PLSR	n/a	17.6-25.5	0.85 0.81 0.90 0.86	1.52 1.64 1.13 1.08	Subedi & Walsh (2009)
	Chips	60	Saturna	460-1040	SNV	Full cross-validation	PLSR	Interactance	82.9-98.6	0.94	0.84	Pedreschi *et al.* (2010)
	Intact	114	Asterix, Bruse, Celine, Folva, Saturna	760-1040	SNV (1D) None (2D) None (SG)	Cross validation	PLSR	Interactance	14.4-30.5	0.95 0.83 0.97	0.91 1.68 0.65	Helgerud *et al.* (2012)
Starch	Mashed	268 269 272	n/a	850-2500	SNV & DT+None 1st der. 2nd der.	External	MPLS	Reflectance	9-30	0.90 0.89 0.88	0.74 0.75 0.79	Haase (2006)
	Mashed & homogenized	126	Aveka, Festien, Karakter, Kamiko, Mercator, Seresta, Valiant	400-2500	MSC	Cross validation	PLSR	Reflectance	18.0-23.4	n/a	0.4	Brunt *et al.* (2010)

[a]DM, dry matter; PLSR, partial least squares regression; n/a, no available data; MPLS, modified partial least squares; PCA, principal component analysis; PCR, principal components regression; MSC, multiplicative scatter correction; SNV, standard normal variate; DT, detrend; LR, Linear regression; MLR, multiple linear regression.

The estimation of the components of potato starch by NIRS has also been a matter of research. Thygesen *et al.* (2001) carried out a study for the determination of phosphate content and viscosity behaviour of potato starch by NIRS combined with PLSR method. The viscosity behaviour of the starch is one of the most important quality parameter for potato starch. Since this parameter is normally measured by viscograms that are time and sample consuming, the aim of that project was to evaluate the feasibility of NIR to predict viscosity behaviour in a faster way. 97 samples of potato were used prepared from a set of 100 potato samples. They were measured by NIR and by a viscogram for viscosity determination. Phosphate content of samples was determined by wet oxidation with sulphuric acid and colorimetric determination of the formed inorganic phosphate according to Stuffins (1967). Phosphate content of the samples was between 0.029-0.11%. The results obtained showed that NIR combined with PLS for the prediction of phosphate content was possible with a RMSECV value of 0.006% on a basis of standardized sample preparation. Moreover, prediction of viscosity was also possible as this parameter was highly related to phosphate content in the data set.

Protein: Protein content of potatoes is considerably smaller than dry matter and starch content (0.5-2.5%), therefore, it seems difficult the prediction of this component by NIRS. Table 2.2 shows several studies carried out in order to predict protein content of potatoes. Best results were obtained for the estimation of coagulating protein with an RMSEP value of 0.06 (Brunt & Drost, 2003, 2010) whereas crude and recoverable protein were harder to predict achieving higher SEP values and lower R^2 (Brunt *et al.*, 2010; Fernández-Ahumada *et al.*, 2006; Haase, 2006).

Some authors attributed these lower values to the reduced range of these constituents and the high error values of the reference methods (Fernández-Ahumada *et al.*, 2006).

Furthermore, the ability of NIRS to qualitative classify samples according to their protein content has been assessed. In a research developed by Fernández-Ahumada *et al.* (2006) a discriminant analysis was performed in order to classify samples in two categories regarding to their protein content. Samples were split into two groups: one including the samples with low recoverable protein content (<14mg g^{-1}) and the second with the samples that presented high values (≥14mg g^{-1}) for that parameter. A total of 184 samples were used for the study and an overall of 161 of them (87.5%) were correctly classified (CC). The results obtained demonstrated that, in spite of the low protein content, it was possible to classify potato samples by NIRS according to that parameter (table 2.2).

Table 2.2. Overview of applications of NIR spectroscopy to measure protein content of potatoes[a].

Type of sample	Parameter	Number of samples	Cultivar	Wavelength range (nm)	Pre-process	Validation	Analysis	Mode	Range %	R^2	RMSEP/SEP	References
Mashed & homogenized	CGP	504	n/a	1100-2500	Smoothing & MSC	External	PLSR	n/a	0.65-1.58	0.84	0.0036	Brunt & Drost (2003)
Mashed	P CP	176 174 173 187 190 188	n/a	850-2500	SNV & DT + None 1st der. 2nd der. None 1st der. 2nd der.	External	MPLS	Reflectance	0.85-2.91 0.85-2.91	0.61 0.59 0.58 0.25 0.22 0.13	0.20 0.20 0.21 0.09 0.08 0.10	Haase (2006)
Mashed	CP RP	275	n/a	1000-2500	1st der. & MSC	Cross & external	PLSR	Interactance -reflectance	1.23-3.21 0.58-2.05	0.62 0.46	2.4 1.7	Fernández-Ahumada et al. (2006)
Mashed & homogenized	CGP	219	18 different	1100-2500	Smoothing	External	PLSR	Reflectance	0.87-1.53	0.84	0.0036	Brunt & Drost. (2010)
Mashed & homogenized	CP	126	n/a	400-2500	MSC	Cross validation	PLSR	Reflectance	0.69-1.95	n/a	0.025	Brunt et al. (2010)

[a]CP, crude protein; RP, recoverable protein; CGP, coagulating protein.

Sugars: Sugars are carbohydrate compounds presented in narrow concentrations in potatoes; therefore, their estimation based on NIRS might not be as accurately as in other compounds such as DM and starch (Hartmann & Büning-Pfaue, 1998). The main sugars in potato are glucose, fructose and sucrose. Glucose and fructose are reducing sugars, each representing about 0.15-1.5% of fresh weight (FW) of tubers while sucrose accounts for about 0.4-6.6% FW (Rady & Guyer, 2015b; Storey, 2011).

Monitoring sugar levels in potatoes before and during storage is very important since higher levels of reducing sugars favoured the Maillard reaction between them and the amino acid asparagine resulting in a dark brown colour (Storey & Davies, 1978). Moreover, an increase of sucrose content due to storing tubers at temperatures below 4ºC, causes a sweetening flavour in potato chips and French fries (Storey, 2011).

To this respect, several authors have reported estimation of sugars by NIR (table 2.3). Mehrubeoglu and Cote (1997) while investigating the on-line application of NIR to estimate total reducing sugars (TRS) of potatoes achieved a RMSECV and RMSEP values that complied with the specifications of less than 0.15% of TRS. Thus, authors concluded that NIR spectroscopy met the requirements to be used for the on-line real-time measurements of these compounds. On the other hand, Scanlon *et al.* (1999) reported poor ability of NIRS to predict fructose content, results (not shown) that differed from those reported by the former authors. Other authors have studied NIRS prediction of individual sugars components such as glucose, fructose and sucrose along with the estimation of TRS content. They concluded that the prediction of these individual components needed to be improved (Hartmann & Büning-Pfaue, 1998). Nevertheless, Rady and Guyer (2015c) have recently obtained good results in the prediction of glucose and fructose in intact potatoes. Other authors also conducting experiments in intact tubers resulted in higher RMSEP values (Chen *et al.*, 2010; Rady *et al.*, 2014). In any case, it seems that the estimation of TRS gave lower RMSEP values than the prediction of glucose and fructose separately. This fact might be useful since TRS content seems to be technologically more important than the content of the single compounds (Kolbe, 1997).

Table 2.3. Overview of applications of NIR spectroscopy to measure carbohydrate content of potatoes[a].

Type of sample	Parameter	Number of samples	Cultivar	Wavelength range (nm)	Pre-process	Validation	Analysis	Mode	Range %	R^2	RMSEP/SEP	References
Sliced	TRS	39	Russet Chipping	n/a	n/a	Cross & external	PLSR	Interactance	0.04-0.2 0.002-0.12	0.57 0.62	0.0009 0.0001	Mehrubeoglu & Cote (1997)
Mashed & homogenized	Glucose Fructose Sucrose Σ red. sugars	116 100 109 134	Granola & Nicola	1100-2500	1st der.	Cross & external	MPLS & PCA	Reflectance	0.148-0.520 0.101-0.439 0.136-0.399 0.249-0.790	0.70 0.89 0.62 0.82	0.04 0.02 0.03 0.06	Hartmann & Büning-Pfaue (1998)
Intact	Carbohydrates	250	Irish-Cobbler, May-Queen & Kita-akari	700-1100	None 2nd der.	Cross & external	PLSR	Interactance	11.1-22.6	0.86 0.86	0.98 0.98	Chen *et al.* (2004)
Intact	Glucose Fructose	100	May-Queen	600-1100	2nd der.	Cross validation	PLSR	Interactance	0.5-2 0.1-1.1	0.65 0.71	0.46 0.26	Chen *et al.* (2010)
Mashed, homogenized & freeze-dried	Red. sugars Sucrose TS	2517	10 different	850-2500	SNV & DT	Cross & external	MPLS	Reflectance	$9*10^{-5}$-$9*10^{-3}$ $1*10^{-3}$-$22*10^{-2}$ $1.2*10^{-3}$-$27*10^{-2}$	0.43 0.71 0.66	$38.9*10^{-6}$ $96.9*10^{-6}$ $135*10^{-6}$	Haase (2011)
Intact	Glucose Sucrose Glucose Sucrose	200 200 	Russet Norkotah Frito Lay 1879 	446-1125	SNV+MC+PRT WB+MC WB+MC+PRT MC	Cross & external	PLSR	Interactance	0.05-0.83 0.05-0.23 0.01-0.03 0.02-0.06	0.79 0.26 0.88 0.81	0.152 0.205 0.062 0.043	Rady *et al.* (2014)
Sliced	Glucose Sucrose Glucose Sucrose	200 200 	Russet Norkotah Frito Lay 1879 		Norm+MC+OSC Norm+LRT 2^{nd} der+MC+PRT 1^{st} der+MC		PLSR			0.95 0.5 0.90 0.81	0.078 1.023 0.051 0.043	
Sliced	Glucose Sucrose Glucose Sucrose	200 200 	Russet Norkotah Frito Lay 1879 		MC 1^{st} der+MC+LRT MC+LRT 1^{st} der+MC		PLSR	Transmittance		0.87 0.63 0.66 0.57	0.192 0.896 0.051 0.058	
Intact	Glucose Sucrose Glucose Sucrose	167 74 503 193	Russet Norkotah Frito Lay 1879 		MC MC MC OSC	Cross & external	IPLSR	Interactance	0.02-0.05 0.1-0.49 0.03-0.07 0.03-0.12	0.97 0.94 0.81 0.80	0.019 0.108 0.015 0.038	Rady & Guyer (2015b)

[a]TRS, total reducing sugars; TS, total sugar; Norm, normalization; MC, mean centre; PRT, power reference transformation; WB; weight baseline; OSC, Orthogonal signal correction; LRT, log reference transformation; IPLSR, interval partial least square regression.

Fat and acrylamide: Determination of acrylamide contents in potato chips is currently necessary due to its potentially toxic attributes and the fact that very high concentrations can be produced in amylaceous fried foodstuffs (Rosén & Hellenäs, 2002). Moreover, consumer awareness of the fat content in potato products is increasing worldwide as do so the seeking for low fat products. Some studies have been developed for the determination of both fat and acrylamide content in potato processed products.

Segtnan *et al.* (2006) investigated the determination of acrylamide contents in potato chips using process variable settings and NIRS. Acrylamide is normally present at elevated concentrations in different types of heat treated foods and is considered a carcinogen constituent. For this study potato samples were sliced, fried and ground before NIR analysis. Then PLSR method was applied to build the spectral prediction models. A correlation coefficient between predicted acrylamide values and reference values was 0.952 with a RMSECV of 246.8 µg/kg. The high correlation coefficient along with the low RMSECV suggested that NIR spectroscopy could be accurate enough for determining the acrylamide contents in processed potato chips.

Another related study was accomplished by Pedreschi *et al.* (2010) for the on-line monitoring of different constituents in potato chips using near infrared interactance and visual reflectance imaging. The objective of the study was to determine DM, fat and acrylamide contents in potato chips by NIR in routine analysis. Raw potatoes were hydrogenated with palm oil, cut into slices and fried at different durations resulting in 60 samples analysed by NIR, visible spectroscopy (VIS), combination of both and reference methods. The corresponding correlation between predicted values by NIR and reference values was 0.99 for fat with a SEP value of 0.99% (w/w). Therefore, on-line NIR interactance technology was found to predict fat content of potato chips with high accuracy. For acrylamide content the best model resulted from the use of NIR and VIS (both spectral regions) with a correlation coefficient of 0.83 and a SEP value of 266 µg/kg. Pedreschi *et al.* (2010) concluded that the acrylamide estimation error was a little high and thus, they suggested that the system should be used for classification of samples with high and low acrylamide contents rather than for prediction.

Shiroma and Rodriguez-Saona (2009) investigated the potential of NIR combined with chemometrics to determine fat and moisture content in potato chips and its capacity to classify samples based on their composition. A total of 15 commercial potato chips fried from different sources according to their label were used. PLSR method was used for the prediction models and SIMCA was used for qualitative analysis. The correlation coefficient of CV obtained was 0.97 for fat with a SECV value of 1.54. The classification

model based on SIMCA was able to differentiate potato chips by source of frying oil. Based on these results authors concluded that it was possible to determine fat content in potato chips as well as classify them according to their composition by a fast, simple and accurate method.

Ni *et al.* (2011) investigated NIR application in potato chips for the prediction of the following quality parameters: fat, moisture, acid and peroxide values. The aim of the investigation was to compare the performance of calibration models developed using NIR spectra and PLSR method with non-linear KPLS and LS-SVM models for the determination of the above parameters. For this purpose samples of four commercial brands were analysed by chemical methods and by NIR. The results showed that both KPLS and LS-SVM methods performed well for the four parameters with correlation coefficients for CV ranging from 0.930 to 0.996 with RMSEP values between 0.076 and 0.518. The highest R value for independent validation was obtained for fat content by LS-SVM with a RMSEP of 0.211. On the other hand, PLS calibrations performed well for three parameters but the results for peroxide value were poor with the lowest correlation coefficient (0.762) and highest RMSEP (0.772). Authors summarized that NIR spectroscopy combined with the use of chemometric was able to accurately predict quality parameters in potato chips.

According to these studies, it may be assumed that NIR spectroscopy performs well for the parameter fat whereas for acrylamide content more robust models need to be built. In the meantime, NIR spectroscopy is a useful tool for classifying samples according to this last constituent.

Carotenoids: Benefits of carotenoids have been reported by several authors. Carotenoids are well known for their health promoting functions to the immune system and reduction of the risk of degenerative diseases (Fraser & Bramley, 2004; Rodriguez-Amaya, 2001; Van den Berg *et al.*, 2000). Due to these advantages consumer concern for products with high carotenoids concentration is growing in the same way as the industry interest for the screening and development of food crops with increased concentrations of those components (Seddon *et al.*, 1994; Stahl & Sies, 2005).

Bonierbale *et al.* (2009) examined the potential of NIR to estimate total and individual carotenoid concentrations in cultivated potatoes. 189 potato tubers were used for the development of NIRS calibrations and external validation. Samples were freeze-dried and milled prior to NIRS analysis. The individual carotenoids analysed were anteraxanthin, violaxanthin, lutein, zeaxanthin and β-Carotene. The concentration of total carotenoids ranged between 440 and 8560 $\mu 100g^{-1}$ dry weight (DW) and of

individuals from 0 to 2240 µ100g^{-1} DW. Coefficients of determination obtained ranged from 0.60 to 0.92. Best results were obtained for total carotenoids estimation (R^2: 0.91) and zeaxanthin (R^2: 0.92) with SEP values of 610 and 410 µ100g^{-1} DW (dry weight) respectively. Results demonstrated that NIR had the potential to accurately predict total carotenoids and zeaxanthin and the rest of the individual carotenoids with relatively good accuracy.

2.4.2. Other characteristics

Crop yield evaluation: NIRS major applications in potatoes are generally focused on determining internal components (DM, starch, soluble solids, carotenoids, etc); but also, sensory texture of cooked potatoes has been evaluated (Thybo *et al.*, 2000). Despite this, in general terms, the scope of NIRS covers a wide range of applications nowadays such as the determination of physiological indices of crops (Schmilovitch *et al.*, 2000) or the optimal date for fruit picking (Peirs *et al.*, 2001).

Jeong *et al.* (2008) studied the correlation between sprouting capacity in potato tubers and NIRS. They used 380 potato tubers divided into four groups, two groups of the same variety (cv. Superior) harvested at two consecutive years, a group of another variety (cv. Atlantic) and the last group containing the total number of samples. The sprouting capacity of the four calibration sets ranged between 0.24 and 7.70 with a standard deviation between 1.03 and 1.95. NIR spectra were measured in reflectance mode in the 400-2500 nm wavelength range. First derivative and SNV-DT pre-treatments were applied to the data. Modified partial least squares (MPLS) was used as a regression method to correlate spectral data and sprouting capacity. The R^2 obtained for cross and external validation ranged from 0.69 to 0.93 with SECV and SEP values between 0.40 and 0.68. Based on the results obtained Jeong *et al.* (2008) concluded that it was possible to predict the sprouting capacity of potato tubers by NIRS with a reliable accuracy. That fact was an important discovery and could have significant implications in the potato industry.

Other authors found that the relationship between the absorption of nitrogen and the total fresh weight in potato crops could be used to calculate proportions of supplemental nitrogen fertilizer (Young *et al.*, 1997). Consequently, the correlation between NIR spectroscopy and nitrogen absorption of potato plants was investigated in order to be used for this purpose. As an example, Young *et al.* (1997) developed a study for the NIR determination of nitrogen in potato tissues. They used samples of two different potato varieties grown under six nitrogen treatments. Samples of leaves, stems and tubers with nitrogen concentrations between 0.60 to 3.65%, 0.68 to 5.88%

and 2.25 to 8.00% (on a DW basis) respectively were scanned by NIR and by reference method (Dumas combustion). They obtained low SEP values of 0.11, 0.03 and 0.09% for leaf, stem, and, tuber respectively. Authors concluded that the estimation of nitrogen concentration in potato crops by NIR techniques was cheaper and safer than comparable chemical methods. These results were in accord with those reported by Váradi *et al.* (1987); however, they studied NIR reflectance for the determination of total nitrogen in ground grape leaf samples rather than in potato tubers.

Likewise, MacKerron *et al.* (1995) published a report about the influence of particle size, milling speed and leaf senescence to the assessment of total nitrogen in potato tissues by comparing near infrared reflectrometry and Dumas combustion methods. Samples of leaf, stem, and tubers were used for this research with nitrogen concentration levels between 0.55-1.35%. The results obtained reported particle size as a source of error in NIR analysis whereas milling speed within the range examined did not appear to be an important variable. The R^2 obtained at two different milling speeds were 0.79 and 0.89 with RMSEP values of 0.05 and 0.16 for tuber and leaf material respectively. Later in the same year, the second part of that experiment was carried out. At that time authors compared the influence of operator, moisture and maturity class in the assessment of total nitrogen comparing the two methods explained before. Once again, samples of leaf, stem, and tubers were analysed by Dumas combustion and by NIRS, and in the following two years, by a number of operators who made estimates of nitrogen concentrations. Authors achieved a good correlation between NIRS and Dumas combustion methods in the determination of nitrogen for the different operators who tested the samples with SEP values between 0.02 and 0.11 (Young *et al.*, 1995).

Texture: Texture of potatoes at the time of consumption is an important factor related to products quality. Consumers associate the quality of potatoes according to the texture they perceive when consuming. This sensory perceived quality is normally measured using a panel of trained judges. These procedures require a considerable amount of time as well as an important investment. Therefore, much research has been developed to determine the texture of potatoes by instrumental technologies rather than methods based on human's perception (Hesen & Kroesbergen, 1960).

Some authors have studied the correlation between NIRS and texture profiling of potatoes. Researches have focused on cooked potatoes as the most consumed potato based food product. Hesen and Kroesbergen (1960) carried out a project determining the correlation between NIRS and texture profiling of steam cooked potatoes. The

texture of steam-cooked potatoes samples was sensory evaluated at 1, 3, and 6 months after storage and NIR spectra were measured. Authors used 87 samples in the range between 1100 and 2500 nm. A quantitative model based on PLS was developed for the parameters firmness, moistness, waxiness and mealiness with R^2 ranging between 0.79 and 0.91 and SEP values 8.64-14.64. Authors stated that NIR was able to evaluate the texture of cooked potatoes with good accuracy.

Another study with similar characteristics was carried out by Thybo *et al.* (2000), but at that time measurements were made in raw and water boiled potatoes. 24 samples of six different potato varieties were used. NIR measurements were made in reflectance mode in the 1100-2500 nm wavelength range and then, PLSR method was applied. The correlation coefficients for the sensory texture attributes: firmness, mealiness and moistness were lower than those obtained by Hesen and Kroesbergen (1960) for the same attributes (R: 0.67-0.73). Moreover, the range of values of the sensory attributes was much smaller than in the previous work and therefore, SEP values were greater in comparison (Thybo *et al.*, 2000).

Van Dijk *et al.* (2002) studied the relationship between DM content, sensory-perceived texture and NIRS in steam cooked potatoes. 81 potato tubers representing different types of cooking behaviour were used for this assessment. Sensory texture analysis was accomplished by a panel of 16 trained judges. The results obtained were very similar to those reported by Thybo *et al.* (2000) for the same parameters (R^2: 0.82-0.92).

Damages evaluation: Damage to potato tubers either by mechanical harvesting or transporting causes a great loss of quality of the final product. More than 60 years ago it was established that almost two-thirds of the potatoes sold in the American market showed external or internal damages (Larsen, 1962); and, in spite of the fact that there have been several investigations focused on reducing the degree of damage, it seems that there is still a need to continue working in this field. Economic losses due to tuber's damages are thus significant (Salar, 2009).

Evans and Muir (1999) investigated the feasibility of NIR spectroscopy as a method for determining the discoloration of potatoes associated with bruising in a non-destructive way. Bruising is considered one of the biggest problems in the potato industry since it causes very important economic losses (Peters, 1996). Therefore, investigation in this field is always welcome. For that research, samples of cv. Record susceptible to bruising were used. Tubers were given a consistent impact and then were stored for 16 h. NIR spectra were measured in both unpeeled and peeled tubers as in bruised and unbruised sites. The results showed that reflectance spectra from unpeeled

bruised tubers had higher reflectance in the NIR than unbruised tubers. Moreover, in peeled tubers, the differences were higher in those regions. Based on these results, authors suggested that bruise detection by NIRS may be possible in unpeeled tubers and almost certainly in peeled tubers. Nevertheless, they stated that the method required an improvement in order to be a reliable technique.

A different type of research was developed by Kemsley *et al.* (2008) when they studied the feasibility of NIR diffuse optical tomography to monitor quality of fresh fruit and vegetables. For that study, a NIR tomograph built from relatively low cost components was used along with potato samples as model specimens or phantoms to develop the image reconstruction approach. Authors found that NIR tomography had the potential to monitor internal defects in agricultural products. That conclusion entailed an important discovery given that the determination of internal damages such as bruising in potatoes before reaching the market could save a lot of money since as stated before, internal bruising is one of the main concerns in the potato industry and causes many annual losses.

From those studies until now, investigations of potato defects and diseases have been focused on using non-destructive imaging-based techniques as described in the next section 2.5. These systems have the advantage that they could provide whole images of the tubers, resulting in a more effective damage determination than using point-based technologies.

2.5. Application of imaging techniques for the analysis of potatoes

Machine vision techniques have gained many interest within the food industry in recent years, resulting in an increase of their applications (Sun, 2011) in many agricultural and fruit products (Cubero *et al.*, 2011). These systems are able to provide spatial information; useful for quality evaluation and sorting of agricultural products (Zheng *et al.*, 2006) based on shape, size, texture, and external defects. However, since such vision systems are generally based on RGB colour cameras, they are usually unable to provide information about internal composition of products (Lorente *et al.*, 2012). To overcome this lack of information, hyperspectral imaging (HSI) is becoming an emerging technology in the field of food quality assessment (Sun, 2010).

In any case, imaging systems either based on RGB cameras, multispectral or hyperspectral have been successfully applied in many studies regarding potatoes in the last years. Rady and Guyer (2015b) have recently reviewed the state of the art in

non-destructive quality evaluation of potatoes from the first application in the sixties up to now. However, no reports were found on the non-destructive detection of blackspot damage in intact potatoes.

2.5.1. Defects and diseases detection

Detection of defects such as greening in potatoes was carried out by Tao *et al.* (1995). For this, they used a machine vision system based on a Sony XC-711 CCD colour camera to acquire the images. The method proved highly effective for colour evaluation and image processing. The vision system achieved over 90% accuracy in inspection of greening potatoes.

Jin *et al.* (2009) studied the detection of external defects in a sample set composed of 104 potato tubers belonging to three different cultivars by a vision system including a CCD Olympus camera (C5060WZ). They were able to identified healthy tubers from those with defects with more than 90% accuracy. In a similar study, blemishes in raw potatoes were detected with a 89.6% accuracy by the use of a Sony DSRL-A350K colour camera with a resolution of 1536x1024 pixels (Barnes *et al.*, 2010).

Noordam *et al.* (2005) carried out a study to compared multispectral imaging and RGB colour imaging for the identification of defects and diseases on raw French fries. The system consisted of a Sony 3-CCD colour camera for acquiring colour images and a monochrome PMI-1400EC camera with a mounted ImSpector V9 spectrograph for multispectral images. Authors also investigated the effect of different pre-processing on the spectral data. Best classification performance was given by SVM classifier for multispectral images and KNN for RGB images. Detection of latent greening was only successfully achieved by multispectral images since this type of defect is not visible in the colour images.

Dacal-Nieto *et al.* (2011b) investigated the capability of a hyperspectral imaging technology to detect the presence of an internal disorder (hollow heart) in potatoes. A Xenics Xeva 1.7-320 camera coupled with an SWIR-NIR spectrograph (Specim ImSpector N17E) was used to obtain the images from raw tubers in the spectral range from 900 to 1700 nm. A sample set of 234 potato tubers belonging to Agria variety was scanned with the hyperspectral imaging systems and, after that, tubers were cut to check the presence of hollow heart. 4 different classification algorithms were tried giving similar results with a percentage of CC samples above 86%. Best classification performance was achieved by using SVM (89.1%). In a similar study accomplished by the same authors, the detection of common scab by hyperspectral imaging was investigated (Dacal-Nieto *et al.*, 2011a). Same hyperspectral setup as in the previous

study was used to scan 234 raw tubers with different degrees of common scab. SVM and Random Forest classifiers were used both achieving very high accuracy with more than 96% of CC samples. However, in that study, once again, SVM performed slightly better with a 97.1% of accuracy. Other authors also studied detection of scab diseases in potato by image processing (Samanta *et al.*, 2012). In this case, authors used potato images from an RGB camera and by applying a novel histogram approach, they were able to detect potatoes affected by scab disease with more than 87% accuracy.

Razmjooy *et al.* (2012) published a paper dealing with real-time detection of defects in potatoes using machine vision. A set of 50 bags of potatoes were randomly selected including healthy, unhealthy, standard, and malformed potatoes. A Sony cyber-shot DSC-W350 was used to acquire a total of 500 images. Using SVM as the classification algorithm, authors achieved a 95% correct classification rate of external defects.

Gao *et al.* (2013b) carried out a comparative study of transmission and reflection hyperspectral imaging technology for potato damage detection. Potato images from three directions (damage part facing front, back, and side to the camera) were acquired by transmission and reflection with a hyperspectral system. Best classification performance was achieved by transmittance information with the damage part facing front and back (100%) and very high accuracy was obtained when tubers faced side to the camera (99.53%).

2.5.2. Sorting methods

Elimination of tubers with external or internal defects is essential during processing of potatoes for fresh market. As commented before, sorting is traditionally carried out by human labour, with the disadvantages this practice has associated. For these reasons, some authors have focused on studying potato sorting by imaging systems.

Thus, Al-Mallahi *et al.* (2008) developed a machine vision system to discriminate raw tubers from soil clods. For this, authors used a RGB colour camera (Canon, Powershot A80) to take colour images and a hyperspectral camera (Specim, ImSpector V10). A sample set consisting of 230 yellow potatoes from three different varieties was used. Measurements were performed under dry and wet conditions. Authors achieved 92.8% of CC samples under wet condition and 73.4% under dry using RGB colour images. Better results of classification were obtained when using hyperspectral images with 99.9% and 97.4% of CC samples under wet and dry conditions respectively.

Later on, same authors developed a machine vision system based on ultraviolet imaging to detect potato tubers on the harvester. A system consisting of an UV 1-CCD camera (Sony, XC-EU50) with a wavelength range between 300 and 420 nm was used

to capture the images of a set of samples comprising 60 tubers and 60 clods. Results obtained showed that above 98% of both tubers and clods were successfully detected (Al-Mallahi *et al.*, 2010).

Other authors have focused on the in-line implementations of such vision system for the detection of irregular potatoes. Thus, ElMasry *et al.* (2012a) developed an image processing algorithm to detect misshapen potatoes in a conveying system. They designed and constructed a machine vision system consisted of a CCD camera (Sony XC003P), a frame grabber (Matrox Meteor), a lighting chamber and a roller conveying system. A sample set of 410 tubers was measured. A stepwise linear discriminant analysis was carried out resulting in an overall accuracy of 96% in the detection of malformed tubers.

Another type of study was that performed by Nguyen Do Trong *et al.* (2011) to predict optimal cooking time for cooked potatoes by hyperspectral imaging. The hyperspectral imaging setup mainly consisted of an ImSpector V10 Spectrograph (Spectral Imaging Ltd., Oulu, Finland) coupled with a monochrome CCD camera KP-F120 (Hitachi Denshi Ltd., UK) covering the wavelength range 400-1000 nm. Samples were boiled at 0 (raw), 3, 6, 9, 12, 15, 18, 21, 24, 27, and 30 min and one half of each potato was scanned. PLSDA was performed to discriminate between cooked regions and raw regions of each potato. Image processing techniques were applied to detect the cooking front and then, by modelling the evolution of the cooking front over time, authors were able to estimate the optimal cooking time for boiled potatoes.

2.6. Conclusions

The different studies described in this chapter demonstrate the potential of non-destructive spectroscopic techniques in the quality control of potatoes. In section 2.4 a review of the many applications of NIRS in potatoes is presented. As a conclusion, it can be said that most of the studies have been focused on the estimation of internal compounds, mainly dry matter and starch but also others such as sugars, protein, etc. Moreover, it was explained in section 2.1.3 that biological tissues are a highly scattering media and therefore, NIR spectra are dominated by these effects. For this reason, it is important to determine the optical properties of such scattering products in order to know how light propagates through them and to identify in which way this affect NIR spectra. Despite the importance of this fact, only few authors have focused on studying the propagation of light through potato tissues.

In addition, in section 2.5, the application of hyperspectral imaging systems in the quality evaluation of potatoes is described. Hyperspectral imaging is a more recent

technology compared to NIR spectroscopy and therefore, its number of applications in potato quality control is considerably smaller.

Although NIR and hyperspectral imaging applications in the potato industry are generally implemented at the laboratory level, some scientific studies have demonstrated their viability for on-line and in-line applications.

In the case of NIRS, it should be mentioned that even this technique has been applied for the estimation of internal composition of potatoes since the 1980s, there are still some compounds which prediction becomes difficult due to their small presence in potatoes.

Regarding hyperspectral imaging, it seems that there is a need for further investigation on the detection of internal damage; more so, taking into account the advantages this technology presents over conventional NIRS for such detection, and the serious issues this type of damage entails.

These conclusions are made within the current objective of satisfying consumers' and industry quality expectations and specifications, an aspect included in the present doctoral thesis.

Acknowledgments

The funding of this work has been covered by the Universidad Pública de Navarra through the concession of a predoctoral research grant and by the National Institute for Agricultural and Food Research and Technology (INIA) project: "Genetic improvement of Potato. Characterization of the material by NIRS technology" RTA 2011-00018-C03-03.

Chapter 3.

Bulk optical properties of potato flesh in the 500-1900 nm range

This chapter has been published as:

- López-Maestresalas, A., Aernouts, B., Beers, R., Arazuri, S., Jarén, C., Baerdemaeker, J., & Saeys, W. (2015). Bulk Optical Properties of Potato Flesh in the 500-1900 nm range. *Food and Bioprocess Technology,* 9(3), 463-470. Impact factor (2014): 2.691.

Bulk optical properties of potato flesh in the 500-1900 nm range

A. López-Mestresalas[1], B. Aernouts[2], R. Van Beers[2], S. Arazuri[1], C. Jarén[1], J. De Baerdemaeker[2], W. Saeys[2]

[1] Department of Projects and Rural Engineering, Universidad Pública de Navarra, Campus de Arrosadia, 31006 Navarra, Spain

[2] KU Leuven Department of Biosystems, MeBioS, Kasteelpark Arenberg 30, 3001 Leuven, Belgium

Abstract

In this study the optical properties of potato flesh tissue were estimated using double integrating sphere (DIS) measurements combined with an inverse adding-doubling (IAD) light propagation model. Total reflectance, total transmittance and unscattered transmittance were measured for the wavelength range 500-1900 nm with 5 nm resolution. From these measurements, the bulk optical properties (absorption coefficient, scattering coefficient and anisotropy factor) of 53 potato tubers of the Hermes cultivar were estimated. The estimated absorption coefficient spectra were dominated by water and starch absorption bands, the main chemical components of potato tissue. Comparison of these values to those reported in literature for similar products showed comparable absorption profiles. The obtained scattering coefficient spectra showed a smooth decrease from 166 to 160 cm^{-1} in the NIR spectral range with increasing wavelength, which is common for biological tissues. The anisotropy factor spectra obtained for the full wavelength range studied ranged between 0.949 and 0.959 with a maximum variability of 0.009 among the set of samples used. The information obtained is essential to understand the effects of absorption and scattering on the propagation of light through the potato tubers in order to design more efficient sensors for non-destructive quality evaluation.

Keywords: Optical properties; Double integrating spheres; Scattering; Absorption; Inverse adding-doubling; Potatoes.

3.1. Introduction

As commented section 2.1, NIRS technology was introduced in the industry in the 1970s, and has been used since then for moisture, protein, and fat content determination among other compounds in many agricultural and food products (Davies & Grant, 1987; Irudayaraj & Gunasekaran, 2001; Nicolaï *et al.*, 2014). The detection of these chemical components relies on the wavelength-specific NIR absorbance which is linearly proportional to the concentration of the absorbing constituents. Nevertheless, next to the sample's chemical composition, the absorbance of the photons also depends on the travelled path length. In agricultural and food products, light is typically scattered by the numerous local non-uniformities (physical microstructure): cells, cell organelles, air pores, fibrous structures, fat globules, etc. (Wang & Li, 2013). Because these light deflections increase the photon's path length, the NIR absorbance increases to an unknown extend. In practice, the measured NIR reflectance or transmittance spectra are often dominated by these scattering effects. Accordingly, a decrease in NIR reflectance or transmittance caused by a change in scattering might be misinterpreted as an increase of the sample's actual absorption coefficient, while the latter should be normalized for the photon's path length to follow a linear relation with the sample composition.

Simulation of the light propagation through the tissues of interest could provide insight in the specific relations between scattering, absorption, and measured spectroscopic signals, promoting a more efficient NIR sensor design, data processing, and data interpretation (Steponavičius & Thennadil, 2009, 2011, 2013; Zamora-Rojas *et al.*, 2014). Light propagation in homogeneous turbid media such as potatoes can be described with the Radiative Transfer Equation (RTE) which as explained in section 2.1.3 uses the bulk optical properties of the tissue (μ_a, μ_s and g) as input parameters (Aernouts *et al.*, 2013; Arimoto *et al.*, 2005; Bashkatov *et al.*, 2005a; Lino *et al.*, 2003; Zamora-Rojas *et al.*, 2013a). The bulk absorption coefficient relates linearly to the sample's composition, while the bulk scattering coefficient and scattering anisotropy factor define the level and the angular distribution of the light deflections (Xia *et al.*, 2007). Knowledge of the potential range of the sample's bulk optical properties is essential to perform detailed light propagation studies (Arimoto *et al.*, 2005; Lino *et al.*, 2003; Zamora-Rojas *et al.*, 2013a).

As stated previously, double integrating spheres measurements, either alone or in combination with unscattered transmittance measurements, are the most convenient method to obtain accurate estimates of the optical properties for thin slices of biological tissues (Pickering *et al.*, 1993). A DIS setup allows measuring the total reflectance and

total transmittance simultaneously, providing more robust and accurate estimates of the bulk optical properties compared to single integrating sphere measurements (Bashkatov *et al.*, 2011; Pickering *et al.*, 1993; Prahl, 2011). The theory and practicability of integrating spheres for measuring the reflectance of a sample have been widely studied (Pickering *et al.*, 1993).

Since 2000, the bulk optical properties of many food products have been characterized in the visible (VIS) and near infrared (NIR) spectral ranges. Cubeddu *et al.* (2001) measured the internal optical properties of apples by time-resolved reflectance spectroscopy, while Fraser *et al.* (2003) studied the light distribution inside mandarins using a custom-made probe connected to a spectrometer. Wang & Li (2013) and Saeys et al. (2008) used integrating sphere techniques to determine the bulk optical properties of, respectively, onions and apples. Regarding potatoes, Karagiannes *et al.* (1989) studied the bulk absorption and reduced scattering coefficients of Idaho potatoes in the 340-1360 nm spectral range, assuming a wavelength-independent value for g. Nevertheless, both the bulk scattering coefficient and the anisotropy factor are essential inputs for simulation studies to develop and evaluate novel optical quality sensors for the non-destructive analysis of tubers. Additionally, the NIR range above 1400 nm includes important absorption bands related to potato constituents (e.g., water, starch …). Therefore, the aim of this study was to measure all three bulk optical properties (μ_a, μ_s, and g) of potato flesh in the 500-1900 nm wavelength range. The goal of this chapter complies with the specific objective 1 stated in section 1.4.

3.2. Materials and methods

3.2.1. Design of the experiment

Figure 3.1 shows the different steps followed to determine the bulk optical properties of potatoes. Samples from cv. Hermes were provided by The Basque Institute for Agricultural Research and Development (NEIKER-Tecnalia) in Spain, and were sent to KU Leuven Department of Biosystems, MeBioS, Leuven, Belgium, for the measurements. Samples were sliced to be measured with the DIS and unscattered transmittance system to obtain the total reflectance (R_t), total transmittance (T_t) and unscattered/collimated transmittance (T_c). This system was controlled with LabVIEW 8.5 software (National Instruments Corporation, Austin, TX). After that, from these values, the absorption coefficient (μ_a), reduced scattering coefficient (μ_s'), and anisotropy factor (g) were calculated using the IAD software implemented in MATLAB

(ver 7.10, The Mathworks Inc., Massachusetts, USA). This process is further explained in the following sections.

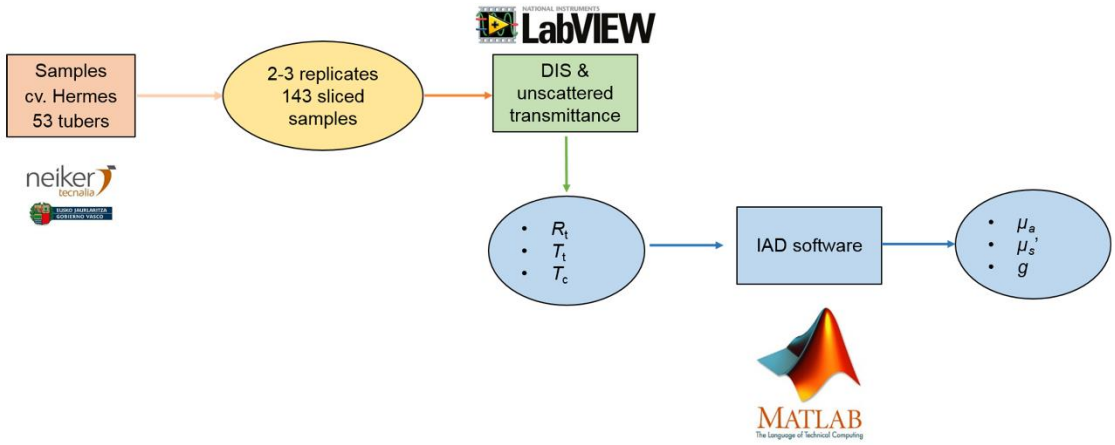


Figure 3.1. Flowchart for the measurement of the bulk optical properties of potatoes.

3.2.2. Sample set

A total of 53 different tubers belonging to cultivar Hermes were used in this study, and 2-3 replicates of each tuber were analysed accounting for a total of 143 samples of flesh. Potato tubers were kept in a refrigerator at 4°C before analysis. Then, tubers were washed, weighed and peeled. Potato flesh samples were sliced into 550-µm-thick portions of 30-mm diameter using a meat slicer (Junior sup 19 mod 30-595A, CAD, ITALY). The thin portions were sandwiched between two optical borosilicate glass plates after adding distilled water to remove air and consequently reduce the refractive index (RI) mismatch at the glass-sample boundary. The thickness of the sample slides was measured with a calliper with an accuracy of 5 µm. The thickness of the cuvette, together with the thickness of the sample slab and glass plates (1100 µm), was used to calculate the thickness of the present water layer.

3.2.3. DIS and unscattered transmittance measurements

To acquire total reflectance, total transmittance and unscattered transmittance, a wavelength-tuneable DIS and unscattered transmittance measuring system was used as shown in figure 3.2. Detailed information about the used system can be found in Aernouts *et al.* (2013). The measurements were taken in the wavelength range 500-1900 nm with an interval step of 5 nm.

In this setup, illumination of the samples is provided by a high-power (4 Watts) supercontinuum laser source with a 460-2400 nm range (SC450-4, Fianium Ltd., Southampton, UK) coupled into a high precision monochromator (Oriel Cornerstone™ 260×, Newport, Irvine, USA) covering the 450-2800 nm wavelength range in order to

obtain a high signal-to-noise ratio (SNR) for DIS measurements on biological tissues. A flip mirror, located behind the monochromator, reflected or passed the collimated light toward either the DIS or the unscattered transmittance measurement path.

The thin potato tissue samples were placed between two optical borosilicate glass slides (1.1-mm wall thickness), separated by a spacer of 550 µm (path-length). For DIS measurements, the sample (tissue + glass plates) was positioned between the two Infragold®-coated integrating spheres (RT-060-IG, Labsphere, Inc., North Sutton, USA; 700-20,000 nm wavelength range; 6 inch inner diameter). The sample port had a diameter of 1 inch, while all detector ports were 0.5 inch in diameter (Aernouts *et al.*, 2013; Zamora-Rojas *et al.*, 2013a). Two detectors were located on the wall of each sphere: a large-area Silica (Si) detector (PDA 100A, Thorlabs Inc., New Jersey, USA) for the 500-1050 nm range and a one-stage Peltier-cooled extended-InGaAs detector (PDA 10DT-EC, Thorlabs Inc., New Jersey, USA) for measurements in the 1050-2250 nm range. To avoid detection of reflected or transmitted light coming directly from the sample, baffles were located on the inner sphere wall between the sample and the detectors (Aernouts *et al.*, 2013).

For the unscattered transmittance measurements, three round slits were optically aligned: one immediately before and after the sample and a third one at a 1.5-m distance behind the sample. This design ensures a maximum collection angle of 5 mrad in order to minimize the number of scattered photons captured. The sample was placed in a sample holder, perpendicular to the incident collimated light beam. An automated flip mirror, located behind the third slit, reflected or passed the light respectively toward a Si or an extended-InGaAs detector with the specifications mentioned above.

Instrumental calibration and sample analysis were performed following the methodology described by Prahl (2011). To prevent detector saturation during calibration of the unscattered transmittance setup, a neutral density filter (NDF, Qioptiq Limited, Luxembourg) of optical density 3.0 was located in the sample holder (Zamora-Rojas *et al.*, 2013a). The system (laser, monochromator, flip mirror, and detectors) and data collection were automated and controlled with LabVIEW software (Aernouts *et al.*, 2013).

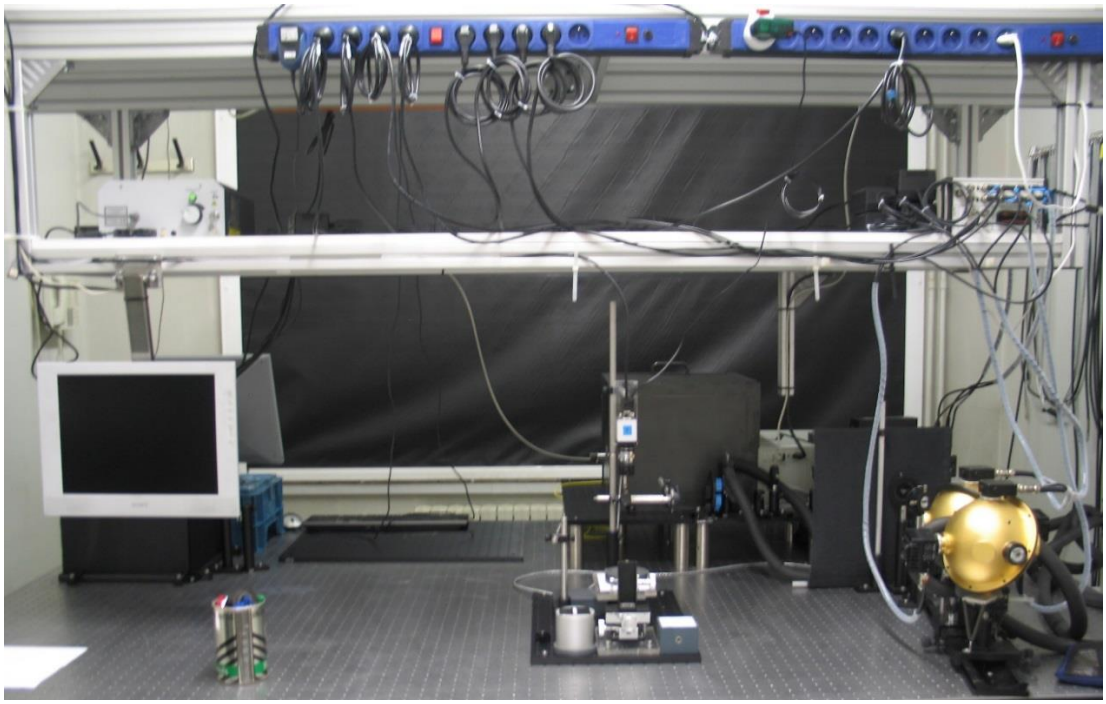


Figure 3.2. Double integrating spheres and unscattered transmittance measuring system.

3.2.4. Optical properties estimation

R_t, T_t, and T_c were calculated from the DIS and unscattered transmittance measurements in MATLAB, as described by Aernouts *et al.* (2013). From these values, the absorption coefficient (μ_a), reduced scattering coefficient (μ_s'), and anisotropy factor (g) were calculated using the IAD software provided by Prahl (2011) and implemented in MATLAB. To correct for reflections at the glass-sample boundaries, the refractive index (RI) of potato tissue samples had to be supplied to the algorithm in order to solve the inverse problem. The wavelength-dependent RI values were estimated by adding the difference between the RI of potato and water at 589.29 nm to the wavelength-dependent RI values for water reported by Hale and Querry (1973) and Segelstein (1981). According to Birth (1978), the RI value for potato at 589.29 nm is 1.4, while the RI of water is 1.33. The scattering coefficient (μ_s) was calculated from μ_s' and g using the similarity relation (Tuchin, 2007):

$$\mu_s = \frac{\mu_s'}{(1-g)} \tag{6}$$

Additionally, through the unscattered transmittance measurement of 1 mm water, the μ_a spectrum of pure water was derived. Taking into account this water spectrum, as well as the thickness of the water layer and the sample slab in the cuvette, the μ_a spectrum of the sample slab alone was obtained. Furthermore, the μ_s and μ_s' spectra were

rescaled with the thickness of the sample slab relative to the total thickness (sample slab + water layer).

3.3. Results and discussion

In figure 3.3, the mean and standard deviation (vertical lines) of the total reflectance (R_t), total transmittance (T_t), and unscattered transmittance spectra (T_c) acquired for 143 potato tissue samples in the 500-1900 nm wavelength range are illustrated.

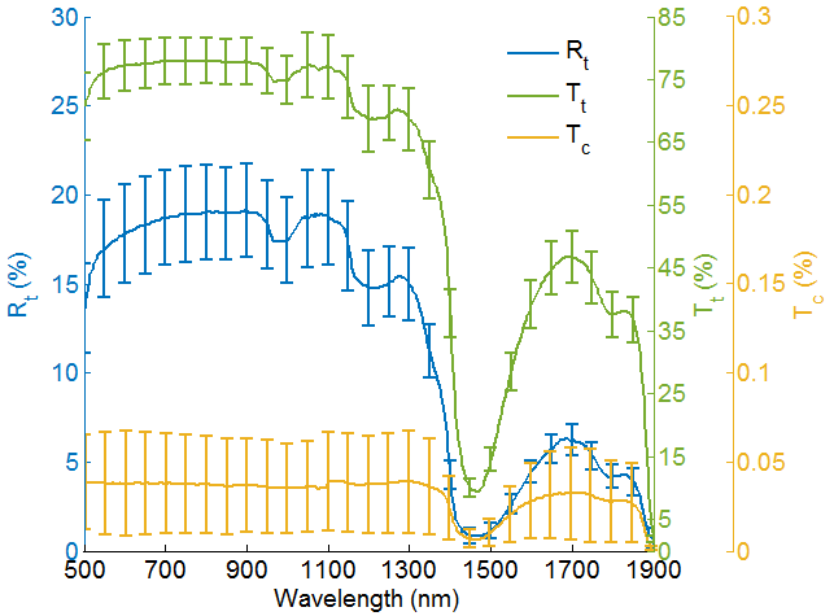


Figure 3.3. Measured R_t, T_t, and T_c values of potato flesh samples in a DIS and unscattered transmittance measuring.

All the measured spectra (R_t, T_t, T_c) show clear valleys around 970, 1210, 1440-1490, and at 1900 nm, all associated with water absorption bands related to the second and first overtones of OH stretching and OH combination bands. As raw tubers usually have a water content of approximately 80% (Büning-Pfaue, 2003), it is not surprising that the acquired spectra were dominated by the absorption signature of water. At 1765 nm, another valley can be observed which can be related to the first overtone of the CH stretching. This can be attributed to starch which represents approximately 60-80% of the dry matter of raw potato tubers and contains both CH and CH_2 groups. Due to the high water absorption, the measurement values were close to zero in the 1900-2250 nm wavelength range (data not shown). As a result, the system failed to estimate the bulk optical properties at these wavelengths. In the 500-1900 nm range, T_c values varied between 0 and 0.04 ± 0.01%. These low values indicate that only a very small fraction of the photons passes through the tissue without being scattered. This limits the SNR of the obtained T_c (Zamora-Rojas *et al.*, 2013a).

In figure 3.4, the mean and standard deviation of the bulk optical properties (μ_a, μ_s, and g) calculated with the IAD method are illustrated. As mentioned above, no reliable values could be obtained in the 1900-2250 nm wavelength range (not shown) as the signal-to-noise ratios of the DIS and unscattered transmittance measurements were too low. As can be seen in the plot, the absorption coefficient (μ_a) spectrum shows absorption peaks at 970, 1210, and 1440-1490 nm due to the high water content in raw potatoes. A small peak around 1765-1800 nm is observed which can be related to the starch content. Additionally, an increase in the absorption of visible light can be noticed towards 500 nm. This is more clearly presented in the detail plot in figure 3.5. The absorption at these wavelengths corresponds to the yellow colour of potato flesh and is probably related to the presence of beta-carotene (Du *et al.*, 1998; Penner, 2010). The bulk scattering coefficient (μ_s) spectrum changes only gently with increasing wavelength, decreasing from 166 to 160 ± 7.35 cm^{-1} for wavelengths below 1900 nm. This decreasing tendency with increasing NIR wavelength is typical for the scattering characteristics of biological tissues (Bashkatov *et al.*, 2005b). The estimated anisotropy factor values were quite stable ranging from 0.959 to 0.949. It should be noted that the accuracy of the estimated anisotropy factor values was negatively influenced by the strong water absorption at 1440-1490 nm, as can be seen from the larger standard deviations.

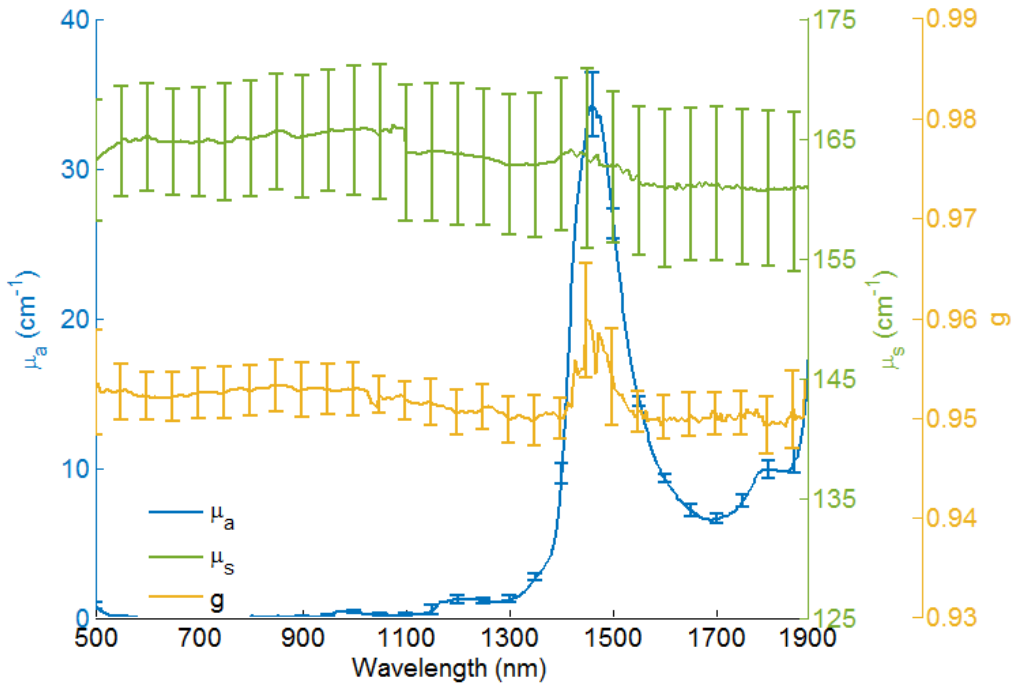


Figure 3.4. Bulk optical properties of potato flesh estimated by the IAD method. The vertical lines indicate the standard deviation values.

In figures 3.5 and 3.6, the bulk optical properties of potato flesh of this study are plotted together with the values reported by Karagiannes *et al.* (1989) for Idaho potato tubers. In the latter study, potato tuber samples of 3.5 mm were measured in air in the 340-1360 nm spectral range with a two-channel integrating sphere spectrophotometer. The optical properties were estimated from these data with two formulations of the 1D diffusion approximation differing in the utilized phase function, assuming a refractive index for potato equal to 1.36 (Karagiannes & Grossweiner, 1988) and a fixed anisotropy factor *g* of 0.

In figure 3.5, the average absorption coefficient spectrum obtained and the one reported by Karagiannes *et al.* (1989) are plotted together with the absorption coefficient spectrum of water from 500 to 1900 nm (Segelstein, 1981). Although the same absorption bands for water appear in both potato spectra, the values at the absorption peaks are on a different level. The estimated values at 970 and 1200 nm reported by Karagiannes *et al.* (1989) considerably exceed the absorption coefficient values of pure water. These values are highly questionable as water is the most important NIR-absorbing component in potatoes. Therefore, it is hypothesized here that these values were overestimated due to an inferior separation between scattering and absorption. This can additionally be noticed at 800 nm, where a baseline is clearly present in the μ_a reported by Karagiannes *et al.* (1989), while nearly no absorption is expected as potato constituents do not absorb significant amounts of light at this wavelength. In the 600-1400 nm range, the μ_a spectrum obtained is very close to the absorption coefficient spectrum of water (figure 3.5), suggesting that the sample consists mainly of water. As explained before, some water was added to the samples to reduce the RI mismatch at the boundaries, probably causing an increase in water content of the samples. In the 1500-1790 nm range, the average μ_a spectrum of the samples is higher compared to the water spectrum. This can be attributed to the first overtone absorption bands of the CH and CH_2 bands in the starch molecules. The μ_a values in the 1400-1500 nm range are probably slightly overestimated due to imperfect separation of the scattering and absorption coefficient values in this region, induced by the low signal levels (figure 3.3). Although the results in this study are not perfect, they are much closer to reality than those reported by Karagiannes *et al.* (1989).

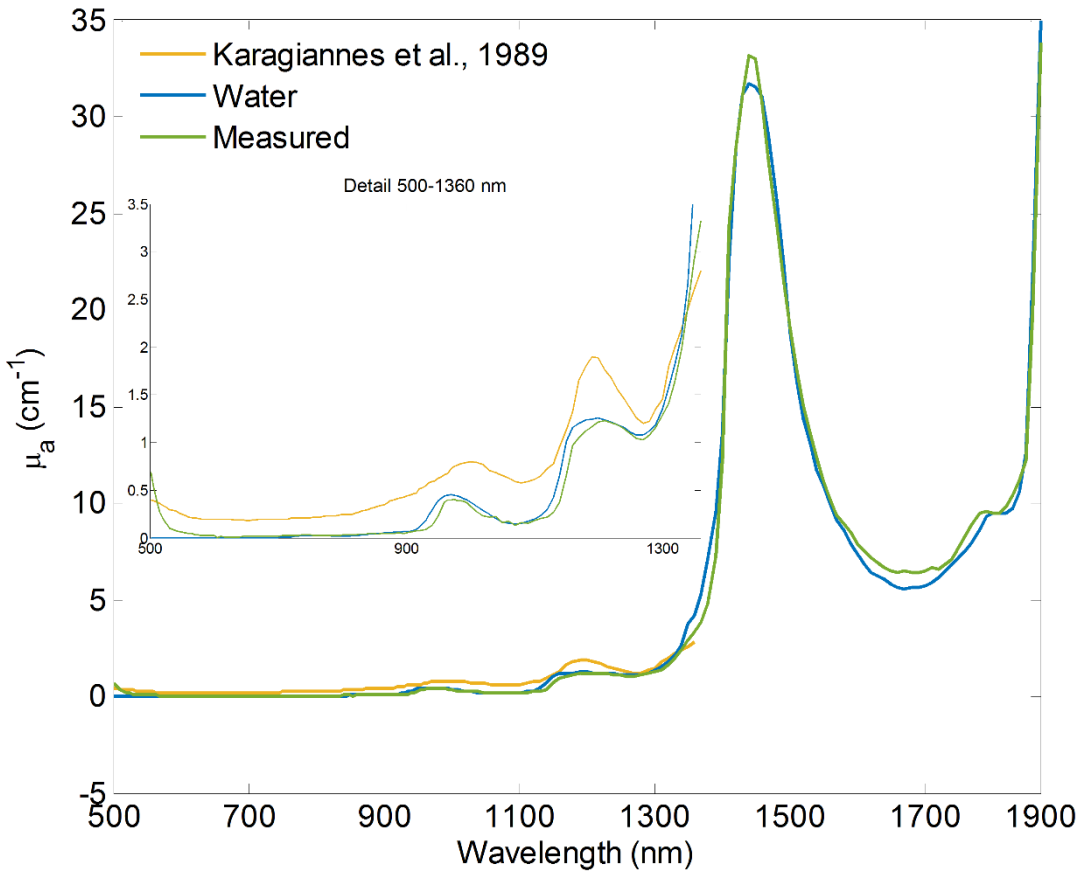


Figure 3.5. Comparison of potato flesh and water absorption coefficient in the 500-1900 nm spectral range. The detail plot (500-1360 nm) zooms the spectral region analysed by other authors.

In figure 3.6 the scattering coefficient μ_s' values in the 500-1900 nm spectral range estimated in both studies are shown. It can be observed that the μ_s' spectrum obtained is relatively steady along the entire spectral range with values between 7.5-8.3 cm^{-1} except at the 1440-1490 nm range where it is influenced by water and values decrease down to 6.6 cm^{-1}. The reduced scattering coefficient values obtained by Karagiannes *et al.* (1989) are smaller, which may be attributed to a difference in the microstructure properties of the cultivars and samples considered in both studies. However, as no microstructure information on the samples is available, this hypothesis cannot be verified. On the other hand, as their absorption coefficient values were most likely overestimated, it is likely that they underestimated the reduced scattering coefficient μ_s' values due to imperfect separation of the absorption and scattering properties (Zamora-Rojas *et al.*, 2013a).

Recently, Wang and Li (2013) measured the bulk absorption coefficients, reduced scattering coefficients, and scattering anisotropy of skin and flesh of four different onion cultivars at 633 nm. They obtained absorption coefficient values between 0.32 and 1.2 cm^{-1} for the flesh of the different cultivars and bulk scattering coefficients in the 5.8-6.6 cm^{-1} range for white and yellow onions. These values are similar to those obtained

in this study for potato flesh. The estimated anisotropy values at 633 nm for onion flesh tissues varied from 0.35 to 0.73, which is considerably smaller than those obtained for potato tissue in the present study.

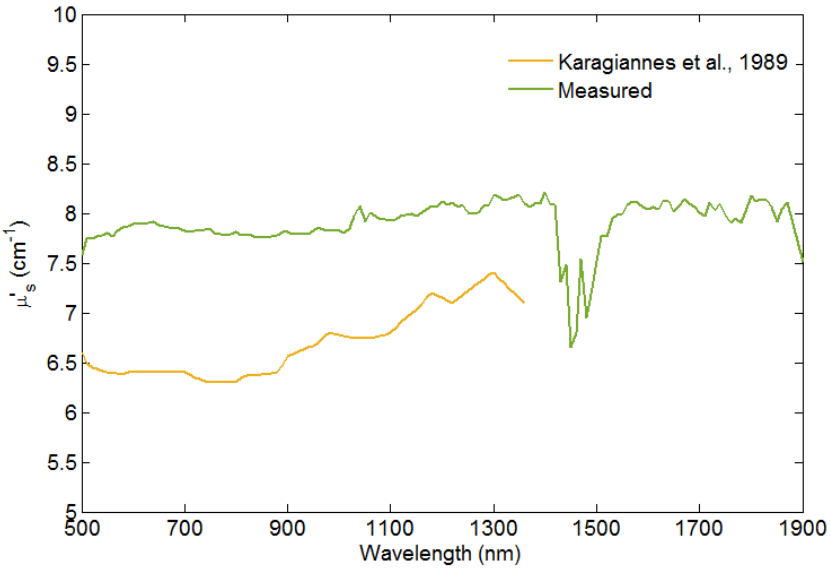


Figure 3.6. Comparison of potato flesh reduced scattering coefficient in the 500-1900 nm spectral range.

Saeys *et al.* (2008) studied the optical properties of both skin and flesh for three different apple cultivars at 1450 nm by using a single integrating sphere measurements in combination with IAD. The reduced scattering coefficient values for apple flesh obtained in their study ranged from 10 to 15 cm^{-1} in the 350-1900 nm range, while those for apple skin ranged from 30 to 60 cm^{-1} in the same wavelength range. This suggests that both apple flesh and skin tissue are more optically dense than potato flesh tissue.

The fact that the bulk optical properties estimated in this study are more reliable than those obtained by Karagiannes *et al.* (1989) can most likely be attributed to the use of a DIS system in combination with a high-power supercontinuum laser, rather than the use of a single integrating sphere in combination with a halogen light source. The DIS system allows for simultaneous and robust measurement of R_t and T_t without replacing the sample or the sphere. The use of a supercontinuum laser light source makes it possible to focus a high-power beam on a very small spot on the sample, reducing the lateral light losses. Another possible explanation for the difference between the bulk optical properties obtained in both studies may be related to the use of different light-propagation models, the instrument calibration, and the heterogeneity of the samples (Tuchin, 2007).

3.4. Conclusions

The bulk optical properties of potato flesh were measured in the 500-1900 nm wavelength range by means of double integrating spheres and unscattered transmittance measurements. This resulted for the first time in a wavelength-dependent estimate of the anisotropy factor for potato flesh over that wavelength range. The obtained results show that the bulk absorption coefficient spectra are mainly dominated by water, especially in the NIR region, while the starch absorption bands are less pronounced, but still clearly visible. The bulk scattering coefficient spectra slightly decreases with increasing wavelength. The estimated values for the anisotropy factor were high (>0.94) over the considered spectral range, which indicates that potato tissue is highly forward scattering. Compared to the results obtained by other authors, the absorption and reduced scattering coefficient spectra for potato tissue retrieved in this study are considered more reliable. Moreover, the optical properties were obtained in the visible and extended NIR range from 500 to 1900 nm, compared to 340-1360 nm. The added value of the extended range is significant as the relevant absorption peaks for water and starch are located in the NIR region of the spectrum. Additionally, the obtained optical properties were comparable to the values obtained for other products (apple and onion flesh and skin) reported in literature.

The optical characterization of potato flesh elaborated provides information about the absorption, scattering, and the angular scattering distribution of light propagating through the tissue. This information is essential to understand the effects of absorption and scattering on the reflectance and transmittance spectra measured with VIS-NIR spectroscopy.

Acknowledgments

The funding of this work has been covered by the Universidad Pública de Navarra through the concession of both a predoctoral research grant and a mobility grant, by the National Institute for Agricultural and Food Research and Technology (INIA) project: “Genetic improvement of Potato. Characterization of the material by NIRS technology” RTA 2011-00018-C03-03, and by the Agency for Innovation by Science and Technology in Flanders (IWT) through the Chameleon (SB-100021) project. The authors would also like to thank the Basque Institute for Agricultural Research and Development (NEIKER-Tecnalia) for supplying the potato samples.

Chapter 4.

Non-destructive classification and determination of potato compounds by NIRS

Part of this chapter has been published as:

- Tierno, R., López, A., Riga, P., Arazuri, S., Jarén, C., Benedicto, L. and Ruiz de Galarreta, J. I. (2016). Phytochemicals determination and classification in purple and red fleshed potato tubers by analytical methods and near infrared spectroscopy. *Journal of the Science of Food and Agriculture, 96,* 1888-1899. Impact factor (2014): 1.714.
- López, A., Arazuri, S., Jarén, C., Mangado, J., Arnal, P., Ruiz de Galarreta, J. I., Riga, P., & López, R. (2013). Crude Protein content determination of potatoes by NIRS technology. *Procedia Technology*, *8,* 488-492.

Non-destructive classification and determination of potato compounds by NIRS

Abstract

Over the last two decades, the attractive colours and shapes of pigmented tubers and the increasing concern about the relationship between nutrition and health have contributed to the expansion of their consumption and a specialty market. Thus, in this study, first the concentration of health promoting compounds such as soluble phenolics, monomeric anthocyanins, carotenoids, vitamin C, and hydrophilic antioxidant capacity, in a collection of 18 purple- and red-fleshed potato accessions was quantified and then, cultivars were classified into 3 groups (low, mid and high) regarding their phytochemical composition by using NIRS and chemometric tools.

Cultivars and breeding lines high in vitamin C, such as Blue Congo, Morada and Kasta, were identified. Deep purple cultivars Violet Queen, Purple Peruvian and Vitelotte showed high levels of soluble phenolics, monomeric anthocyanins, and hydrophilic antioxidant capacity, whereas relatively high carotenoid concentrations were found in partially yellow coloured tubers, such as Morada, Highland Burgundy Red, and Violet Queen. Moreover, very good classification rates were obtained regarding samples concentration of soluble phenolics, monomeric anthocyanins and hydrophilic antioxidant capacity.

In addition, a quantitative analysis was performed to estimate potato compounds by NIR spectroscopy in a wide range of samples. Estimation of dry matter (DM), starch, reducing sugars (RS), crude protein (CP), nitrogen (N), total soluble phenolics (TSP) and hydrophilic antioxidant capacity (HAC) in potato tubers grown in two consecutive years was carried out using NIRS and PLS regression. Good PLS models were obtained for the prediction of CP, N and TSP proving the capability of NIRS in estimating internal compounds in potato tubers. Also, the PLS model obtained in the estimation of HAC could be used for screening processes in order to identify samples with high or low HAC.

Keywords: pythochemicals; crude protein; nitrogen; NIRS; TSP; HAC; quantitative analysis; qualitatitve analysis.

4.1. Introduction

As explained in section 2.1, phenolic compounds have been reported as antioxidant compounds that could promote health benefits (Guisti & Wrolstad, 2001; Manach *et al.*, 2004). Potatoes are an excellent source of vitamins and phytochemicals, such as vitamin C, phenolic acids, flavonoids and carotenoids, which are elements beneficial for health (Navarre *et al.*, 2009). Thus, identification and characterization of tubers with high phytochemical content is interesting in order to attract consumers.

Phenolics are commonly classified into three important groups: phenolic acids, flavonoids and tannins (Rice-Evans *et al.*, 1996). The major phenolic acid in potato is chlorogenic acid (Brown, 2005). Anthocyanins are phenolic pigments which constitute the main sub-class among flavonoids. The most common anthocyanidins (the de-glycosylated forms of anthocyanins) found in potatoes are malvidin, petunidin, delphinidin and peonidin in purple tubers and pelargonidin in red ones (Kita *et al.*, 2013).

Carotenoids are a widespread family of lipophilic organic pigments (Rodriguez-Amaya & Kimura, 2004). Based on epidemiological studies a positive link is suggested between higher dietary intake and tissue concentrations of carotenoids and lower risk of chronic diseases (Rao & Rao, 2007). The orange flesh colour found in some potatoes is due to zeaxanthin whereas lutein concentration correlates with the intensity of yellow coloration. Cultivated diploid potatoes derived from *Solanum stenotonum* Juz. & Bukasov and *Solanum phureja* Juz. & Bukasov have been reported to be a great source of zeaxanthin and lutein (Burgos *et al.*, 2009; Griffiths *et al.*, 2007; Lu *et al.*, 2001).

Purple- and red-fleshed potato cultivars are attractive to consumers. Besides their exotic pigmentation, coloured genotypes show significantly higher contents of phenolic compounds (Hamouz *et al.*, 2006; Reyes *et al.*, 2005). Specialty potato food products, such as coloured chips, crisps, purees, canned potatoes, and ready meals are becoming more and more widespread. With varying degrees of success, some companies and research centres are trying to develop new coloured cultivars to support increasing demand of the specialty potato market, placing particular emphasis on their superior nutritional profile, attractive colours and different textures. To this end, the selection of appropriate parents to be used in artificial crosses is one of the main decisions faced by breeders. Thus the identification of promising genotypes is a key step in potato breeding.

The analytical methods commonly employed to determine the main compounds of potatoes in order to identify valuable genotypes require a lot of time and are destructive. Therefore, these methods seem to be not suitable for in-line applications in the food industry (Chen *et al.*, 2010). As commented in section 2.1 NIR spectroscopy presents some advantages over reference methods for quantitative analysis of agricultural and food products. Moreover, NIRS can be used for qualitative analysis, where the aim is to classify samples on the basis of its spectral features rather than estimating the components present in them (Lister *et al.*, 2000).

Despite the fact that quantitative analyses of potatoes are widely available, there is not a great deal of literature concerning qualitative analysis, to the best of our knowledge. Some authors have focused on the classification of potato tubers according to their chemical components. Thus, in a research study developed by Fernández-Ahumada *et al.* (2006), a discriminant analysis was performed in order to classify samples into two categories regarding to their protein content. The results obtained demonstrated the accuracy of NIR to classify potato samples in groups of low and high protein content. More recently, Rady and Guyer (2015a), carried out a study to classify potato samples based on sugar levels. Authors classified sliced and whole tubers of two different varieties regarding their content of glucose and sucrose.

In this study, the concentration of phenolic compounds, monomeric anthocyanins, carotenoids, vitamin C, and hydrophilic antioxidant capacity by analytical methods in a collection of 18 purple- and red-fleshed potato cultivars was analysed. In addition, NIRS combined with chemometric tools was used to classify samples into 3 groups (low, mid and high) according to their phytochemicals content.

On the other hand, and considering, as previously mentioned, that the analytical methods currently employed to determine internal composition of vegetables and fruits include gas-liquid chromatography (GLC), HPLC and UV-VIS spectrophotometry (Friedman, 1997), and that these techniques have a few drawbacks; in this study, quantitative analyses were carried out in order to determine main compounds of potatoes using NIR spectroscopy.

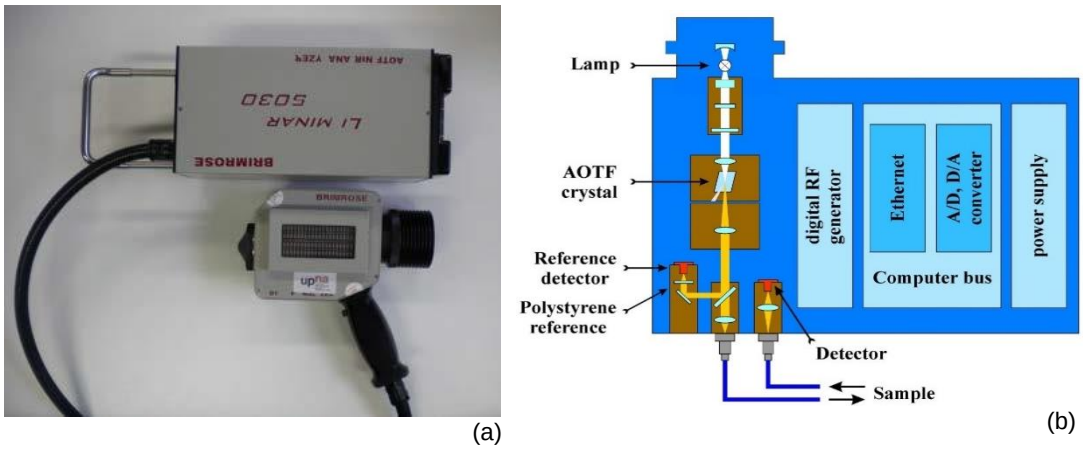
The aims of this study were, first to classify a collection of purple- and red-fleshed potato varieties according to their phytochemical concentration by NIR spectroscopy; and, second, to estimate potato compounds by NIR spectroscopy in a wide range of samples. These targets address the specific objectives 2 and 3 proposed in section 1.4.

4.2. Materials and methods

4.2.1. NIR spectrometer

A Luminar 5030 Miniature 'Hand held' AOTF-NIR (Acousto-Optic Tunable Filter-Near Infrared) Analyzer (Brimrose, Baltimore, MD, USA) was used for both qualitative and quantitative analyses (figure 4.1a). This is a rugged portable spectrometer capable of conducting non-destructive and contact/non-contact tests. This system comprises an InGaAs detector and pre-aligned lamp assembly. This spectrometer allows diffuse reflectance measurements and liquid transmission with probe attachment in the wavelength range from 1100 to 2300 nm. Moreover, it has a scanning speed of 16,000 wavelength/s and a wavelength increment adjustable from 1 to 10 nm. This system can acquire 26 spectra per second. Figure 4.1b shows the schematics of the AOTF-NIR spectrometer.

In this study, the signals were acquired with Brimrose Analytical Software SNAP32! with Brimrose MACRO language and transferred later on to either MATLAB R2014a (The MathWorks, MA, Natick, USA) or Unscrambler X (version 10.3, CAMO software AS, Oslo, Norway).



(a) (b)

Figure 4.1. NIR spectrometer: (a) Luminar 5030 AOTF-NIR analyzer and (b) schematic of the AOTF-based spectrometer. Source: (Biofotonica, 2013).

4.2.2. Qualitative analysis

4.2.2.1. Design of the experiment

In figure 4.2, a flowchart of the steps followed for the qualitative analysis of samples by NIRS is shown. A sample set consisting in 18 potato cultivars were selected for the study.

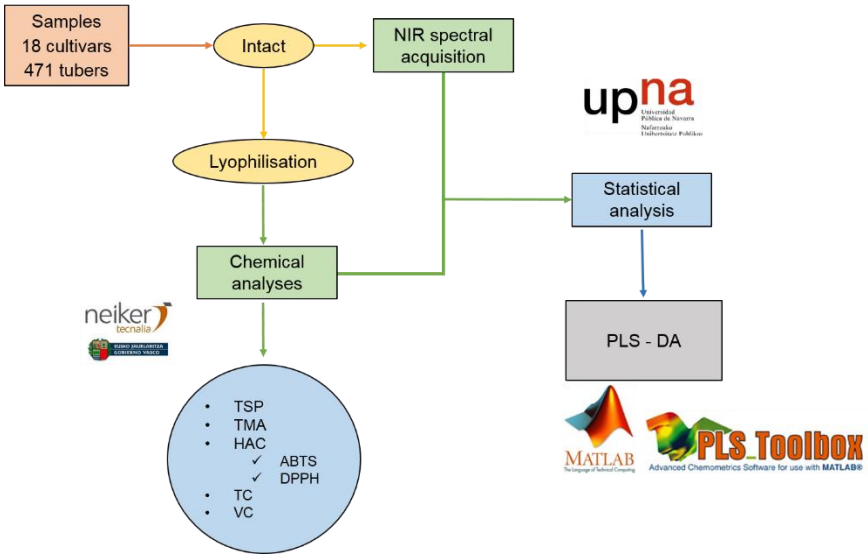


Figure 4.2. Flowchart for the qualitative analysis using NIRS

NIR spectral acquisition was accomplished on intact tubers first. After that, chemical analyses were carried out on lyophilized tubers. As a part of the collaborative work under the research project INIA-FEDER RTA2011-00018-C03 “Mejora genética de patata: Caracterización del material por tecnología NIRS”, the process of lyophilisation along with the chemical analysis of samples were accomplished at the Basque Institute for Agricultural Research and Development (NEIKER-Tecnalia). Total soluble phenolics (TSP), total monomeric anthocyanins (TMA), hydrophilic antioxidant capacity (HAC) quantified by two methods, ABTS and DPPH, (explained in section 1 of the Appendix), total carotenoids (TC) and vitamin C (VC) were extracted and quantified for further analysis. The final step was carried out at the Universidad Pública de Navarra, a Partial least Squares Discriminant Analysis (PLS-DA) model was built to classify the 18 cultivars used according to their phytochemical concentration in groups of low, mid and high content.

4.2.2.2. Plant material

A collection of 18 purple- and red-fleshed potato cultivars and breeding lines was selected on the basis of the contrasting flesh colour of tubers (table 4.1), accounting for a total of 471 samples (*n*=471). Tubers were selected from potato accessions (Potato

Germplasm Collection, NEIKER) grown during the year 2013 in a precise field trial in Arkaute (Álava) in the north-east of Spain (550 m a.s.l.) with humid climate and annual rainfall of about 800 l m^{-2}. The soil with a clay loam texture was previously subjected to conventional wheat cropping. Plants were grown from mid-May to mid-October 2013 after pre-sowing fertilization with 800 kg ha^{-1} (NPK 4-8-16). Watering was performed using an automatic spray irrigation system.

Table 4.1. Collection of 18 purple- and red-fleshed potato cultivars or breeding lines

Cultivar or breeding lines	*Origin*	*Status*	*Skin/Flesh type*[1]
Bleu de La Manche	France	Cultivar	P/P
Blue Congo	Sweden-UK	Cultivar	P/P
Blue Star	Netherlands	Cultivar	P/P
British Columbia Blue	Canada-UK	Cultivar	P/P
Fenton	Canada-UK	Cultivar	P/P
Highland Burgundy Red	France	Cultivar	R/R
Kasta	Spain	Cultivar	P/P
Morada	Spain	Cultivar	P/P
Morea	Spain	Cultivar	P/P
NK-08/349	Spain	Breeding line	P/P
NK-08/360	Spain	Breeding line	P/P
NK-08/362	Spain	Breeding line	P/P
Purple Peruvian	Peru	Cultivar	P/P
Rosa Roter	Peru	Cultivar	R/R
Rouge de Flandes	Belgium	Cultivar	R/R
Valfi	Sweden-UK	Cultivar	P/P
Violet Queen	Netherlands	Cultivar	P/P
Vitelotte	France	Cultivar	P/P

[1] *Key to skin and tuber flesh types: R=Red, P=Purple.*

4.2.2.3. NIR spectral acquisition

Prior to lyophilisation NIRS measurements were made in whole intact tubers (unpeeled and free of soil). No previous preparation of the samples was carried out for spectral acquisition. Since the main objective in the near future is to implement these techniques for real-time measurements in potato handling lines, it is considered essential to analyse the tubers as they are currently being manipulated in those lines. Samples were scanned at four different points along the equatorial area (figure 4.3) and the average spectrum was used for the analysis. This criterion was adopted from Dull *et al.* (1989) who declared that NIR measurements taken at the centre point of the long axis of potatoes were the most representative sampling of the whole tubers.

For each sample, reflectance measurements were acquired from 1100 to 2300 nm with a 2 nm wavelength increment. Each spectrum provided is the average of 50 scans.

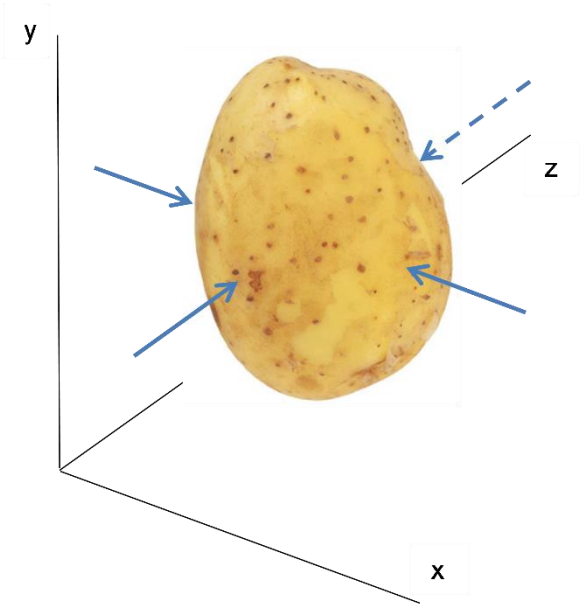


Figure 4.3. The 4 points measured by the NIR spectrometer on each potato.

4.2.2.4. Sample preparation for chemical analysis

Five raw tubers of each accession were washed and patted dry with paper towels, and subsequently peeled and diced. Diced tubers were divided into two equal parts. One part was frozen with liquid nitrogen, kept frozen at -80°C, and later freeze dried, milled by an automatic mortar grinder, and stored at -30°C until analysis. Lyophilisation was used for the determination of dry matter. Total soluble phenolics and total monomeric anthocyanins were analysed using freeze-dried tubers. Total carotenoids, vitamin C, and hydrophilic antioxidant capacity were determined immediately using the fresh portion. All assays were performed in triplicate. Extraction and quantification of the chemical compounds were carried out at NEIKER-Tecnalia and the details of the procedures are explained in section 1 of the Appendix.

4.2.2.5. Statistical analysis

In order to classify the varieties into three groups according to their phytochemical concentration level and called: low content (LC), mid content (MdC) and high content (HC), 3 PLS-DA models were performed. The first PLS-DA model corresponded to varieties grouped according to their content of total soluble phenolics (TSP), total monomeric anthocyanins (TMA) and hydrophilic antioxidant capacity (HAC), since these three appeared to be well correlated in this study. Correlation between these three phytochemicals was also found by other authors (Reyes *et al.*, 2005). The

second PLS-DA performed comprised groups categorized by their total carotenoids (TC) content. Finally, a third PLS-DA was accomplished covering varieties grouped by means of vitamin C (VC) level.

Consequently, a 3-column response Y matrix was introduced for each PLS-DA in which varieties defined as having low content of phytochemicals were described by the dependent vector [1 0 0], varieties belonging to MdC group, by the vector [0 1 0] and the ones belonging to the HC group by the vector [0 0 1]. Samples of the different varieties were randomly divided into calibration and prediction sets corresponding to 70% and 30% of samples, respectively. A classification model was built with the calibration data set and it was externally evaluated using the prediction data to prove its accuracy in classifying new samples. Data were pre-processed by Multiplicative Scatter Correction (MSC) and Mean Centre (MC). As explained in section 2.3, pre-processing methods are commonly used to reduce or avoid the influence of unwanted effects in the data that could negatively affect the consistency of the model (Amigo, 2010; Barbin *et al.*, 2012). MSC reduces scatter effects in the data (Rinnan *et al.*, 2009) while MC reduces systematic noise (Barbin *et al.*, 2012). The cross-validation (CV) method employed was Venetian Blinds with 10 data (splits) subsets since this method is considered simple and easy to implement. Both pre-processing of data and PLS-DA were performed in the PLS-Toolbox (Eigenvector Research Inc., Wenatchee, WA, USA) working under MATLAB R2014a (The Mathworks, MS, Natick, MA, USA).

Performance of the classification model was evaluated by the percentage of correctly classified (CC) samples in each group, the root mean square standard error of prediction (RMSEP) and the coefficient of determination (R^2) (Skibsted *et al.*, 2004).

4.2.3. Quantitative analysis

4.2.3.1. Design of the experiment

Potato tubers grown in two consecutive years (2012 and 2013) were used. Figure 4.4 shows the different steps followed for the quantitative analysis. NIR spectral acquisition was first acquired on intact tubers and then, samples were lyophilized. The process of lyophilisation was accomplished at the Basque Institute for Agricultural Research and Development (NEIKER-Tecnalia). In year 1 (2012), one sample of lyophilized tubers was prepared for each cultivar resulting in a total of 135 lyophilized samples while more than one sample per variety was prepared in year 2 resulting in 228 lyophilized samples. Lyophilized samples were also scanned by NIRS, two replicates per samples were acquired and the mean spectrum of each sample was used for the analysis.

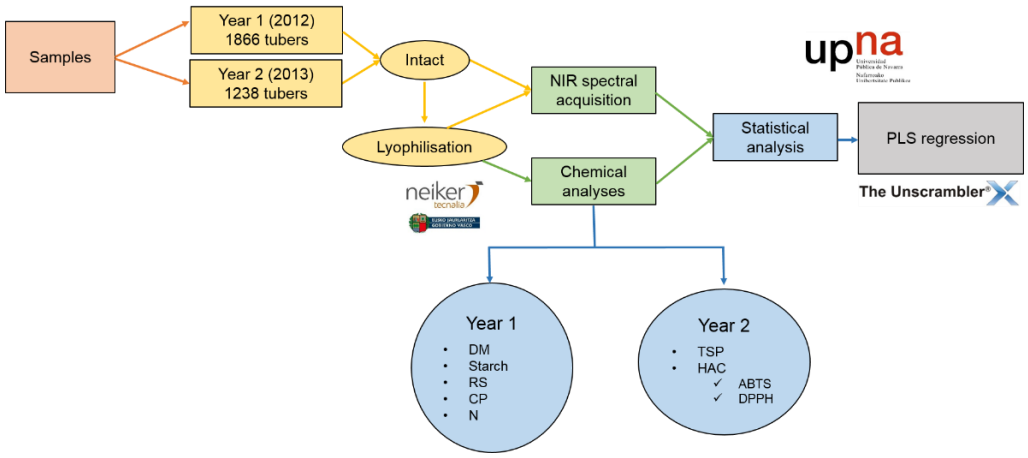


Figure 4.4. Flowchart for the quantitative analysis using NIRS.

Chemical analyses were carried out at the Basque Institute for Agricultural Research and Development (NEIKER-Tecnalia). In year 1 (2012) the following parameters were quantified: dry matter (DM), starch, reducing sugars (RS), crude protein (CP) and nitrogen (N), whereas in year 2, total soluble phenolics (TSP) and hydrophilic antioxidant capacity measured following two different procedures, ABTS and DPPH, explained in section 2 of the Appendix, were extracted and quantified. Then, statistical analyses were performed at the Universidad Pública de Navarra. For this, PLS regression models were developed in order to predict the above chemical compounds of tubers by NIRS.

4.2.3.2. Plant material

Similar to the previous study, samples were grown in the same precise field trial in Arkaute (Álava). Some of the varieties used in this study have been developed within the same institute. Moreover, some of the varieties included in this work are currently undergoing breeding programs and some of the others are clones derived from those programs. A total of 1866 tubers were scanned by NIRS in year 1 (2012) corresponding to 135 different varieties. Spectra for every variety were averaged and one spectrum per variety was used for the analyses. Therefore, a total of 135 samples from 135 different cultivars belonging to year 1 (2012) were used for analysis (n_{year1}=135). Similarly, NIR spectra from 1238 tubers were acquired in year 2 (2013) corresponding to 99 different varieties. More than one spectrum per cultivar was used. Thus, in year 2 (2013), a total of 228 samples corresponding to 99 different varieties were used for quantitative analyses (n_{year2}=228). A total of 80 cultivars were grown in both years, while 55 were exclusively grown in year 1 and 19 only in year 2. Table 4.2 gathers the information of the different cultivars included in this study. The range of samples studied covers white-, red-, yellow- and purple-fleshed varieties.

Table 4.2. List of the cultivars used in this study, harvest year, status and skin and flesh colour

Cultivar or breeding lines	*Year*	*Origin*	*Status*	*Skin/Flesh type*[1]
Adriana	2012	France	Cultivar	Y/LY
Agata	2012/2013	Holland	Cultivar	Y/Y
Agria	2012	Germany	Cultivar	Y/Y
Aladin	2012/2013	Holland	Cultivar	R/Y
Alava	2012	Germany	Cultivar	Y/Y
Albata	2012/2013	Spain	Cultivar	Y/Y
Alda	2012	Spain	Cultivar	Y/W
Alegria Oro	2012/2013	Germany	Cultivar	Y/W
Almera	2012	Holland	Cultivar	Y/LY
Amalia	2012/2013	Holland	Cultivar	Y/Y
Amanda	2012/2013	Sweden	Cultivar	Y/C
Ambition	2012	Holland	Cultivar	Y/LY
Amora	2012	Holland	Cultivar	Y/DY
Amorosa	2012/2013	Holland	Cultivar	R/Y
Antina	2012/2013	Germany	Cultivar	Y/R
Arizona	2012	Holland	Cultivar	Y/LY
Armada	2012	Germany	Cultivar	Y/Y
Arrow	2012	United Kingdom	Cultivar	Y/C
Arsenal	2012	Holland	Cultivar	Y/LY
Artemis	2012	Holland	Cultivar	Y/LY
Asun	2012/2013	Spain	Cultivar	Y/LY
Aurea	2012/2013	France	Cultivar	Y/Y
Ayala	2012/2013	Spain	Cultivar	Y/C
Baraka	2012/2013	Holland	Cultivar	Y/Y
Bleu de La Manche	2013	France	Cultivar	P/P
Blue Congo	2013	Sweden-UK	Cultivar	P/P
Blue Star	2013	Netherlands	Cultivar	P/P
Bonita Negra	2012	Spain	Cultivar	B/LY
Brda	2012/2013	Poland	Cultivar	Y/Y
British Columbia Blue	2013	Canada-UK	Cultivar	P/P
Buesa	2012/2013	Spain	Cultivar	Y/Y
Bufet de Mata Alta	2012	Spain	Cultivar	P/C
Camilla	2012	Germany	Cultivar	Y/-
Candela	2012/2013	Germany	Cultivar	Y/-
Carrera	2012/2013	Holland	Cultivar	Y/Y
Casablanca	2012	United Kingdom	Cultivar	Y/C
Cazona	2012/2013	Spain	Cultivar	Y/C
CIP-700234	2013	Peru	Cultivar	Y-R/Y-R
Clon 23-3-2	2012	Spain	Breeding line	R/-
Corine	2012	Holland	Cultivar	Y/Y
Corrida	2012/2013	Poland	Cultivar	R/LY
Daniela	2012	Germany	Cultivar	Y/Y
Desiree	2012/2013	Holland	Cultivar	R/Y
Duquesa	2012	Spain	Cultivar	Y/Y
Edurne	2012/2013	Spain	Cultivar	Y/Y
Elfe	2012/2013	Germany	Cultivar	Y/DY
Esta	2012/2013	Germany	Cultivar	Y/Y
Europrima	2012/2013	Germany	Cultivar	Y/W
Faluka	2012/2013	Holland	Cultivar	Y/C
Fenton	2013	Canada-UK	Cultivar	P/P
Fina de Carballo	2012	Spain	Cultivar	Y/W
Fontane	2012	Holland	Cultivar	Y/Y
Gorbea	2012	Spain	Cultivar	Y/W
Goya	2012	Spain	Cultivar	Y/Y
H-88-31/34	2012/2013	Spain	Breeding line	Y/Y
Heidrun	2012/2013	Germany	Cultivar	Y/Y
Hermes	2012	Austria	Cultivar	Y/Y
Highland Burgundy Red	2013	France	Cultivar	R/R
Idoia	2012/2013	Spain	Cultivar	Y/Y
Integra	2012/2013	Holland	Cultivar	Y/LY
Inova	2012	Holland	Cultivar	Y/Y
Irati	2012/2013	Spain	Cultivar	Y/C
Isla	2012/2013	Spain	Cultivar	Y/Y
Isle of Jure	2012	United Kingdom	Cultivar	Y/LY
Iturrieta	2012/2013	Spain	Cultivar	Y/Y
Jaerla	2012/2013	Holland	Cultivar	Y/W
Jesus	2013	Spain	Cultivar	P/P-Y
Jimena	2012/2013	Spain	Cultivar	Y/Y
Kasta	2012/2013	Spain	Cultivar	P/W-P
Kenita	2012	Holland	Cultivar	Y/LY
Kennebec	2012/2013	USA	Cultivar	Y/W
Kondor	2012/2013	Holland	Cultivar	R/LY
L-37	2012/2013	Spain	Breeding line	Y/Y
Labadia	2012	Holland	Cultivar	Y/Y
Lady Claire	2012/2013	Holland	Cultivar	Y/Y
Leire	2012/2013	Spain	Cultivar	Y/Y
Lora	2012	Spain	Cultivar	Y/Y

LT-8	*2012/2013*	*Peru*	*Breeding line*	*Y/W*
LT-9	2012/2013	Peru	Breeding line	Y/W
Lutetia	2012/2013	Holland	Cultivar	Y/Y
Madeleine	2012/2013	Holland	Cultivar	Y/LY
Manitou	2012	Holland	Cultivar	R/LY
Marfona	2012/2013	Holland	Cultivar	Y/LY
Mayka	2012/2013	Spain	Cultivar	Y/Y
Melibea	2012/2013	Spain	Cultivar	Y/Y
Melody	2012/2013	Holland	Cultivar	Y/LY
Merida	2012/2013	Germany	Cultivar	Y/LY
Miranda	2012/2013	Germany	Cultivar	Y/Y
Miss Bianka	2012	Germany	Cultivar	Y/LY
Monalisa	2012	Holland	Cultivar	Y/Y
Montico	2012/2013	Spain	Cultivar	Y/W
Morada	2012/2013	Spain	Cultivar	P/P
Morea	2013	Spain	Cultivar	P/P
Murato	2012/2013	Holland	Cultivar	R/LY
Musica	2012/2013	Holland	Cultivar	Y/Y
Mustang	2012	Austria	Cultivar	R/LY
N-180	2012	Germany	Breeding line	B/-
Naga	2012/2013	France	Cultivar	Y/Y
Nagore	2012/2013	Spain	Cultivar	R/Y
Nela	2012/2013	Spain	Cultivar	Y/LY
Nerea	2012/2013	Spain	Cultivar	Y/W
NK-08/349	2013	Spain	Breeding line	P/P
NK-08/360	2013	Spain	Breeding line	P/P
NK-08/362	2013	Spain	Breeding line	P/P
Omega	2012/2013	France	Cultivar	Y/Y
Onda	2012	Spain	Cultivar	Y/C
Opal	2012/2013	Germany	Cultivar	Y/LY
Orchestra	2012	Holland	Cultivar	Y/R
Orla	2012/2013	United Kingdom	Cultivar	Y/LY
Palogan	2012/2013	Germany	Cultivar	Y/Y
Panda	2012/2013	Germany	Cultivar	Y/Y
Pecaro	2012	Holland	Cultivar	R/Y
Pedro Muñoz	2012/2013	Spain	Cultivar	Y/W
Presto	2012	Germany	Cultivar	Y/Y
Priamos	2012	Germany	Cultivar	Y/C
Primavera	2012	Peru	Cultivar	Y/LY
Purple Peruvian	2013	Peru	Cultivar	P/P
Ramses	2012/2013	-	-	Y/Y
Red Baron	2012/2013	Holland	Cultivar	R/Y
Red Scarlett	2012	Holland	Cultivar	R/Y
Riviera	2012/2013	Holland	Cultivar	Y/LY
Roja Ojosa	2013	Peru	Cultivar	R/Y-R
Roja Riñon	2012/2013	Spain	Cultivar	R/Y
Romula	2012	Germany	Cultivar	Y/-
Rosa Gold	2012	Canada	Cultivar	R/Y
Rosa Roter	2013	Peru	Cultivar	R/R
Rouge de Flandes	2013	Belgium	Cultivar	R/R
Rudolph	2012/2013	United Kingdom	Cultivar	R/W
Saline	2012	Holland	Cultivar	Y/DY
San	2012	Poland	Cultivar	Y/Y
Savanna	2012/2013	Ireland	Cultivar	C/W
Saviola	2012/2013	Holland	Cultivar	Y/Y
Simply Red	2012/2013	Holland	Cultivar	R/C
SL-994005	2012	Spain	Breeding line	Y/-
Sofia	2012/2013	Holland	Cultivar	Y/LY
Soprano	2012/2013	Holland	Cultivar	Y/R
Spunta	2012	Holland	Cultivar	Y/Y
SR-260	2012	Spain	Breeding line	Y/-
Stemster	2012/2013	United Kingdom	Cultivar	R/LY
Taurus	2012	Holland	Cultivar	Y/LY
Tebina	2012	Belgium	Cultivar	Y/Y
Valetta	2012	Germany	Cultivar	Y/Y
Valfi	2013	Sweden-UK	Cultivar	P/P
Valnera	2012/2013	Spain	Cultivar	Y/Y
Victor	2012/2013	Spain	Cultivar	R/C
Violet Queen	2013	Netherlands	Cultivar	P/P
Vitelotte	2013	France	Cultivar	P/P
Vivaldi	2012	Holland	Cultivar	Y/LY
Voyager	2012/2013	Holland	Cultivar	B/Y
Zafira	2012	Holland	Cultivar	Y/LY
Zela	2012/2013	Spain	Cultivar	Y/W
Zepa	2012/2013	Spain	Cultivar	Y/C
Zorba	2012/2013	Spain	Cultivar	Y/W
Zuri	2012	Spain	Cultivar	B/W

[1] *Key to skin and tuber flesh types: Y=Yellow, LY=Light Yellow, R=Red, W=White, C=Cream, P=Purple, B=Blue, DY=Deep Yellow.*

4.2.3.3. NIR spectral acquisition

As explained in section 4.2.3.1 NIR spectra of samples were taken before and after lyophilisation. For intact tubers, NIR measurements were acquired along the equatorial area at four points and the average spectrum was used for the analyses. In the same way, once samples were lyophilized, two NIR measurements per sample were acquired and the mean spectrum was used for further studies. A total of 135 samples were used in year 1 (2012) and 228 in year 2 (2013). As commented before (section 4.2.3.1), those samples corresponded to 135 different cultivars in the first year and to 99 in the second. Data from every cultivar was used to build a representative calibration set containing all the variability of samples under study with the objective to improve the prediction capabilities.

The 1100-2300 nm spectral range with 601 points (2 nm steps) was used to obtain the spectra at room temperature. Each spectrum measured is the average of 50 spectra.

4.2.3.4. Sample preparation for chemical analysis

For the lyophilized of samples, tubers were cut lengthwise in order to obtain representative samples of the different tissues. Pieces from 5 to 8 different tubers were then lyophilized in a freeze-dryer Alpha d1-4 (CHRIST, Germany) until they reached 250 g of fresh weight. This process was carried out at -50ºC and 0 atmospheres until the samples lost their whole water content. After that, they were ground with liquid nitrogen up to fine dust and stored at -20ºC until their use. After lyophilisation, extraction and quantification of the chemical compounds described in section 4.2.3.1 were carried out in NEIKER-Tecnalia and the details of the procedures are explained in section 2 of the Appendix.

4.2.3.5. Statistical analysis

A PLS regression model was carried out for every chemical compound analysed. For this, samples of the different varieties were randomly divided into calibration and prediction sets. Only the calibration data set (70% of samples) was used to build the quantitative model while the prediction data set (30% of samples) was used to externally evaluate its capability to estimate the concentration of compounds in new samples. In this case different pre-processing of the data was accomplished including MSC, derivatives and MC.

Venetian Blinds cross-validation method was applied to define subsets of the calibration set of samples. Both pre-processing of data and PLS regression models were carried out in Unscrambler X (version 10.3, CAMO software AS, Oslo, Norway).

The performance of the models was evaluated based on the coefficients of determination (R^2) for predicted *versus* measured parameters in cross-validation and prediction, on the prediction error of the model defined as the root mean square error for cross validation (RMSECV), and on the root mean square error for prediction (RMSEP) (Naes *et al.*, 2002) defined in equation (7):

$$\text{RMSECV or RMSEP} = \sqrt{\frac{\sum_{i=1}^{n_p}(\hat{y}_i - y_i)^2}{n_p}} \quad (7)$$

where n_p is the number of validated objects, and $\hat{y}_i$ and y_i the predicted and measured value of the i^{th} observation in the test set, respectively.

The number of latent variables (LV) in the calibration model was determined as that which minimizes the RMSECV.

The coefficient of determination (R^2) indicates the percentage of the variance in the *Y* variable that is accounted for by the *X* variable. An R^2 value between 0.50 and 0.65 indicates that discrimination between high and low concentrations of the parameter measured can be made. A value for R^2 between 0.66 and 0.81 indicates that model can be used for screening and "approximate" calibration, while, an R^2 between 0.82 and 0.90 can be used for most applications. Calibration models with an R^2 value higher than 0.91 are considered to be excellent (Williams, 2003); and a value for R^2 above 0.98 can be used in any application (Williams, 2001).

The standard error of cross-validation (SECV) and the standard error of prediction (SEP), defined in equation (8), were also reported.

$$\text{SEP} = \sqrt{\frac{\sum_{i=1}^{n_p}(\hat{y}_i - y_i - b)^2}{n_p}} \quad (8)$$

where *b* is the model bias. The bias, also called the error of means is defined as the difference between the mean of the predicted *versus* the mean of the true values of the response variable (Bellon-Maurel *et al.*, 2010).

Finally, the ratio of prediction to deviation (RPD) was also analysed. RPD is the ratio of the standard deviation (SD) of the response variable (concentration of a parameter) to the RMSEP or RMSECV. In this study, the criteria for RPD levels was adopted from Saeys *et al.* (2005). Authors established 5 levels of RPD classification. Thus, an RPD value below 1.5 means very poor predictions and consequently, that model should not

be used; an RPD value between 1.5 and 2.0 indicates that the model could be used for screening purposes as it could differentiate high and low values; an RPD value between 2.0 and 2.5 indicates that approximate quantitative predictions are possible, while a value between 2.5 and 3.0 indicates a good model and above 3.0 the prediction performance is considered excellent.

The evaluation of outliers was made according to the X-residuals and leverage. The leverage is directly related to the Hotelling's T^2 statistic (Velleman & Welsch, 1981). Leverage in the calibration spectra may be compared to a threshold value (Nicolaï *et al.*, 2007). In this study, any observation with abnormal residual or leverage was studied and eliminated if necessary, and then, the calibration model was repeated.

4.3. Results and discussion

4.3.1. Qualitative analysis

4.3.1.1. Chemical analysis

The tested potato cultivars showed considerable differences among each other in terms of total soluble phenolics (TSP), total monomeric anthocyanins (TMA), total carotenoids (TC), hydrophilic antioxidant capacity (HAC) and vitamin C (VC), thus genotype greatly affected all measured phytochemicals (table 4.3). TSP concentrations, ranging from 0.14 ± 0.0208 to 2.78 ± 0.0512 g GAE kg^{-1} FW, were higher in the genotypes Violet Queen, Purple Peruvian and Highland Burgundy Red, while lowest TSP contents were found in NK-08/360, NK-08/349 and NK-08/362. TMA values ranged from 0.0001 ± 0.00001 to 1.33 ± 0.0111 g CGE kg^{-1} FW. Highest TMA concentrations were also measured in the cultivars Violet Queen, Purple Peruvian, Highland Burgundy Red and Vitelotte, while lowest TMA concentrations were found in NK-08/360, Kasta, Morada and Morea. Regarding hydrophilic antioxidant capacity (HAC), the highest value was measured in the cultivar Violet Queen (HAC-ABTS=0.009 ± 0.0002; HAC-DPPH=0.009 ± 0.0001 mol TE kg^{-1} FW) while lowest value corresponded to NK-08/360 (HAC-DPPH=0.002 ± 0.0001 mol TE kg^{-1} FW). TC values ranged from 0.009 ± 0.0012 to 0.0590 ± 0.0043 g LE kg^{-1} FW. TC concentrations were higher in the cultivars Morada and NK-08/362, whereas Bleu de La Manche and Blue Congo showed the lowest TC values. VC concentrations ranging from 0.037 ± 0.0149 to 0.107 ± 0.0204 g kg^{-1} FW showed a three-fold variation among the collection of purple- and red-fleshed tubers. VC values were higher in Blue Congo, Morada and Kasta, while Rosa Roter, NK-08/349 and Highland Burgundy Red showed the lowest VC concentrations.

In the present study, the VC concentrations were significantly lower than those reported by other authors, such as Love *et al.* (2004) or Han *et al.* (2004) who found concentration values ranging from 0.115 to 0.42 g kg^{-1} FW in white- or yellow-fleshed commercial cultivars and breeding lines from North America. Similar to the data of this research, Jiménez *et al.* (2009) reported lower VC concentrations ranging from 0.0754 to 0.286 g kg^{-1} FW in a collection of seven Andean potato cultivars. According to the CD2008/100/EC (2008) an average edible portion (100 g of peeled tubers) contains about 4.6 to 13.3% of the recommended daily intake (RDA) of VC.

Moreover, considering the high losses during thermal processing from 25% to 40% (Hägg *et al.*, 1998; Vanderslice *et al.*, 1990) these cultivars should be considered only as a relatively poor source of VC. However, numerous studies have shown that VC levels are highly dependent on many factors, such as culture conditions, wounding and storage, which can alter rapidly its concentration in tubers (Dale *et al.*, 2003; Tudela *et al.*, 2002). According to Oba *et al.* (1998) the VC contents in potato tubers stored at 4°C for 1 month may experience a decrease from 46 to 57%. Thus, the effect of cold storage on VC could explain why our values were significantly lower than those reported by other authors.

TSP values showed wide variability among the collection of purple- or red-fleshed cultivars or breeding lines showing the great genotype effect on these phytochemicals. The results obtained are in agreement with previous works reporting higher TSP values in coloured potato tubers; for instance, Madiwale (2012) found TSP levels about 0.25 and 2.87 g GAE kg^{-1} FW, while Stushnoff *et al.* (2008) reported values between 0.9 and 4.00 g GAE kg^{-1} FW in a collection of coloured accessions.

Total monomeric anthocyanins (TMA) values found in the present tubers (from 0.0001 to 1.33 g CGE kg^{-1} FW) were similar to most published total monomeric anthocyanin values in purple- or red-fleshed potato tubers. Reyes *et al.* (2005) found TMA values between 0.21 and 0.55 g CGE kg^{-1} FW in red-fleshed tubers, and between 0.11 and 1.75 g CGE kg^{-1} FW in purple-fleshed ones.

TMA values were well correlated with TSP and hydrophilic antioxidant capacity. The correlation between antioxidant capacity and phenolics has also been reported (Andre *et al.*, 2007; Hamouz *et al.*, 2006). Present values of hydrophilic antioxidant capacity were higher than that of white- or yellow-fleshed potatoes (Cevallos-Casals & Cisneros-Zevallos, 2003; Wu *et al.*, 2004), but the values of this study are similar to those obtained by Brown *et al.* (2005) in coloured potato breeding lines.

In relation to the total carotenoid concentrations, values ranging from 0.009 to 0.059 g LE kg^{-1} FW are comparable to those obtained by Lachman *et al.* (2003) using the same

colorimetric method (0.0002–0.025 g LE kg^{-1} FW). The lipophilic character of most carotenoids can explain the lack of correlation between TC and HAC. With the exception of Violet Queen, higher TC values were detected in tubers with a perceptible partially yellow coloration, such as Morada, Rosa Roter, Rouge de Flandes and Highland Burgundy Red. Relatively low TC values were found in most of the medium or deep purple cultivars. According to Kotíková *et al.* (2007) deep purple potato cultivars generally have lower ability to synthesise and accumulate carotenoids when compared to yellow-fleshed potato cultivars.

4.3.1.2. Partial Least Squares Discriminant Analysis

In order to perform a PLS-DA, once the phytochemical content of each variety was determined, samples were divided on the basis of their phytochemical concentration into three groups as “Low Content (LC)” (1-33 percentile), “Mid Content (MdC)” (34-66 percentile) or “High Content (HC)” (67-100 percentile). As previously mentioned, 3 different PLS-DA models were performed. Table 4.3 shows the total soluble phenolics (TSP), total monomeric anthocyanins (TMA), total carotenoids (TC) and vitamin C (VC) obtained for the different purple- and red-fleshed potato cultivars and breeding lines.

Table 4.3. Total soluble phenolics (TSP), total monomeric anthocyanins (TMA), total carotenoids (TC), hydrophilic antioxidant capacity measured by the two procedures, ABTS and DPPH, and vitamin C (VC) in peeled tubers of 18 purple- and red-fleshed potato cultivars and breeding lines.

Cultivar/ Breeding line	*TSP*	*TMA*	*HAC-ABTS*	*HAC-DPPH*	*TC*	*VC*
Bleu de La Manche	0.915±0.0159	0.230±0.0264	0.006±0.0003	0.005±0.0001	0.009±0.0012	0.083±0.0095
Blue Congo	1.180±0.0476	0.178±0.0235	0.005±0.0001	0.005±0.0001	0.010±0.0013	0.107±0.0204
Blue Star	1.080±0.0117	0.129±0.0143	0.005±0.0001	0.005±0.0001	0.019±0.0007	0.091±0.0036
British Columbia Blue	1.120±0.0833	0.228±0.0316	0.005±0.0001	0.005±0.0003	0.016±0.0020	0.058±0.0097
Fenton	1.000±0.0914	0.172±0.0126	0.005±0.0002	0.005±0.0001	0.009±0.0022	0.085±0.0135
Highland Burgundy Red	2.140±0.0489	0.350±0.0102	0.006±0.0001	0.004±0.0002	0.036±0.0036	0.057±0.0096
Kasta	1.270±0.0530	0.001±0.0003	0.006±0.00004	0.005±0.0001	0.025±0.0024	0.097±0.0046
Morada	0.514±0.00984	0.007±0.0018	0.004±0.0003	0.003±0.0001	0.059±0.0043	0.107±0.0061
Morea	0.400±0.0122	0.032±0.0096	0.005±0.0003	0.003±0.0001	0.029±0.0057	0.091±0.0219
NK-08/349	0.376±0.0275	0.055±0.0163	0.005±0.0002	0.004±0.0001	0.024±0.0028	0.050±0.0144
NK-08/360	0.140±0.0208	0.0001±0.00001	0.004±0.0002	0.002±0.0001	0.019±0.0015	0.070±0.0120
NK-08/362	0.428±0.0252	0.158±0.0046	0.004±0.0007	0.003±0.0001	0.037±0.0016	0.087±0.0182
Purple Peruvian	2.150±0.0102	0.408±0.0170	0.007±0.0004	0.007±0.0001	0.019±0.0011	0.093±0.0097
Rosa Roter	1.090±0.0110	0.036±0.0060	0.005±0.0003	0.005±0.0003	0.034±0.0052	0.037±0.0149
Rouge de Flandes	1.810±0.0239	0.270±0.0058	0.004±0.0001	0.003±0.0002	0.035±0.0016	0.068±0.0109
Valfi	0.521±0.0162	0.123±0.0034	0.004±0.0001	0.003±0.0001	0.020±0.0022	0.059±0.0179
Violet Queen	2.780±0.0512	1.330±0.0111	0.009±0.0002	0.009±0.0001	0.035±0.0012	0.067±0.0110
Vitelotte	1.640±0.0901	0.436±0.0074	0.007±0.0004	0.007±0.0001	0.028±0.0048	0.093±0.0182

Means and standard deviations of 1) TSP are expressed as g GAE kg^{-1} FW; 2) TMA are expressed as g CGE kg^{-1} FW; 3) & 4) HAC are expressed as mol TE kg^{-1} FW; 5) TC are expressed as g LE kg^{-1} FW; 6) VC are expressed as g VC kg^{-1} FW.

4.3.1.3. Total soluble phenolics (TSP), total monomeric anthocyanins (TMA) and hydrophilic antioxidant capacity (HAC) model

Figure 4.5 shows the varieties included in each of the three groups in the first PLS-DA carried out. It should be mentioned that only 16 varieties were included in this model, excluding Kasta and Rosa Roter since those two did not show correlation between TSP, TMA and HAC values. Samples (n=429) were randomly divided between calibration and validation sets (n_{cal}=302, n_{val}=127). As it is shown in figure 4.5, the LC group comprised 165 samples belonging to 6 categories defined as having less than 0.9 g GAE kg^{-1} FW of TSP, 0.125 g CGE kg^{-1} FW of TMA and 0.004 mol TE kg^{-1} FW of HAC. It should be noted that variety NK-08/362 was included in this group only attending to their TSP and HAC content, because if according to its TMA content, it must have been included in the second group (MdC). It should be mentioned that cv. Violet Queen was not taken into account for TMA percentile calculation since this cultivar had a TMA value much higher than the rest of the cultivars.

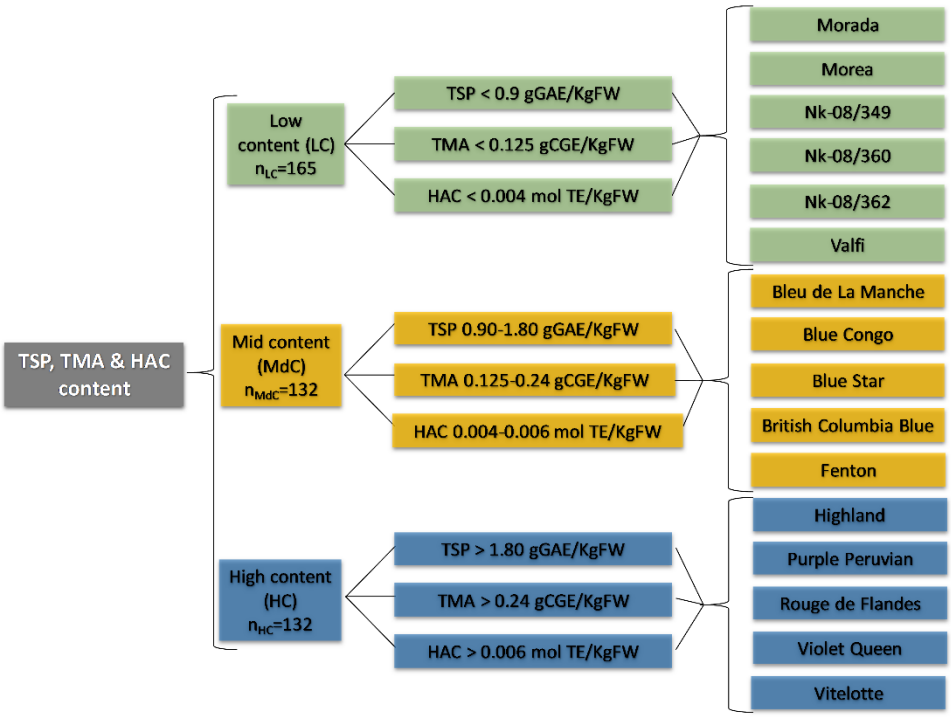


Figure 4.5. Flowchart of the distribution of varieties according to their content of TSP, TMA, and HAC.

The second group, MdC, covered 132 samples belonging to 5 varieties with levels of TSP, TMA and HAC of 0.9-1.8 g GAE kg^{-1} FW, 0.125-0.240 g CGE kg^{-1} FW and 0.004-0.006 mol TE kg^{-1} FW respectively.

Lastly, HC group also contained 132 samples from 5 different varieties with levels above 1.8 g GAE kg^{-1} FW, 0.24 g CGE kg^{-1} FW and 0.006 mol TE kg^{-1} FW of TSP, TMA and HAC respectively. In this case, cv. Vitelotte was included in this group

attending to its content of TMA and HAC only, because according to its TSP content this cultivar should belong to MdC group. Nevertheless, this cultivar had a considerable higher TSP value than the rest of the cultivars comprised in that MdC group.

As revealed before, PLS-DA models were evaluated in terms of correctly classified samples in each of the three groups (LC, MdC and HC), the RMSEP and the R^2. The best PLS-DA prediction model corresponding to the LC group had an R^2 value of 0.75 with RMSEP of 0.23. In table 4.4 the confusion matrix obtained with the PLS-DA can be seen. It shows the percentage of correctly classified samples in calibration (Cal), cross-validation (CV), and prediction (Pred). The diagonal of each confusion matrix represents the percentage of samples correctly classified into the group they belong to while the values outside it correspond to samples wrongly classified into a different group. A perfect classification corresponds to a matrix with a diagonal full of 100% surrounded by 0% rates.

Table 4.4. Confusion matrix of the three groups: low content (LC), mid content (MdC) and high content (HC), in the TSP, TMA and HAC PLS-DA model for Cal, CV and Pred.

		Actual Group (%)		
Set	**Predicted group (%)**	LC	MdC	HC
Calibration set (Cal)	LC	**91.2**	2.47	0.00
	MdC	4.90	**87.6**	9.09
	HC	3.92	9.88	**90.9**
Cross-validation set (CV)	LC	**90.2**	4.94	0.00
	MdC	6.86	**82.7**	10.4
	HC	2.94	12.3	**89.6**
Prediction set (Pred)	LC	**97.6**	0.00	3.85
	MdC	2.38	**83.8**	19.2
	HC	0.00	16.2	**76.9**

Values in bold type correspond to the same group (actual and predicted).

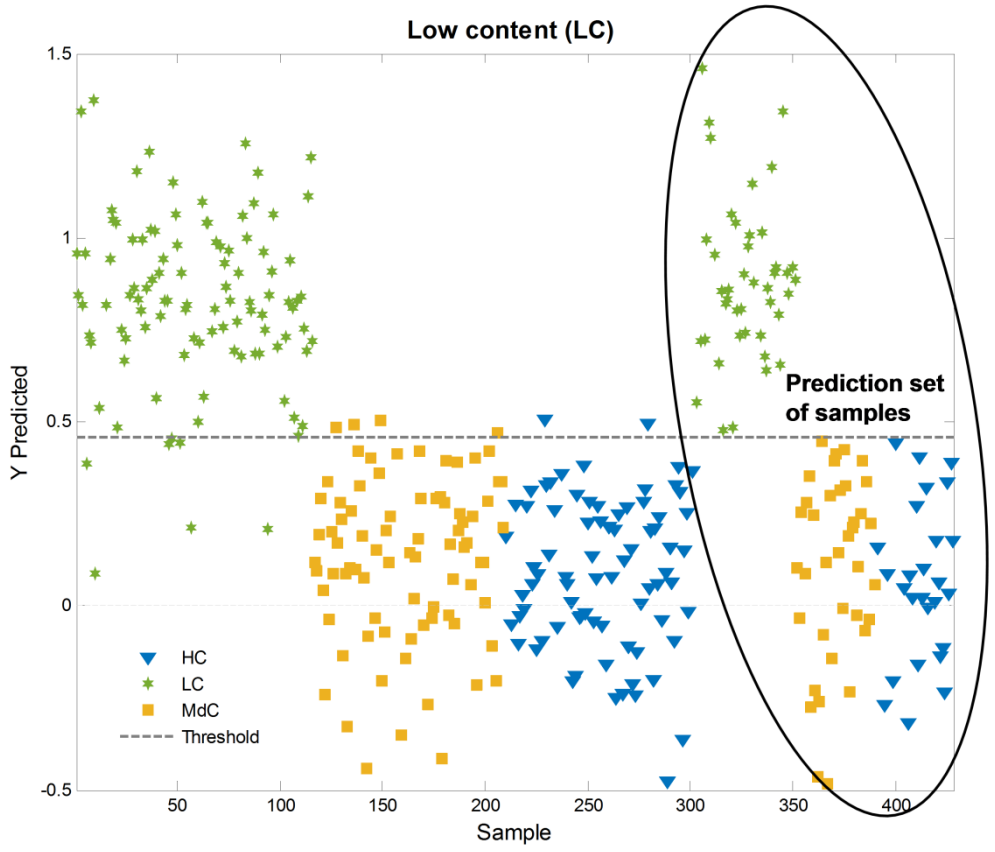
A total of 9 LVs were used explaining 99.9% of the total variance. Slightly better classification rates were obtained for the LC group than for the rest of the groups. This was expected since samples into LC group presented values far away from the other two groups that had more similar values between each other. In any case, good rates above 80.0% of correctly classified samples were achieved for Cal and CV sets. However, LC group was the best classified for the three Cal, CV and Pred sets with more than 90.0% of samples correctly classified. It should be mentioned that the percentages of samples from LC group that were badly classified into the MdC group corresponded to Valfi and NK-08/362 varieties, the two cultivars with the highest values of phytochemicals inside this group. Moreover, as mentioned before if only taking into

account TMA content of samples, NK-08/362 should be included in MdC group; therefore, some misclassification of this cultivar was expected. As a whole, 86.1% of samples were correctly classified in the prediction set.

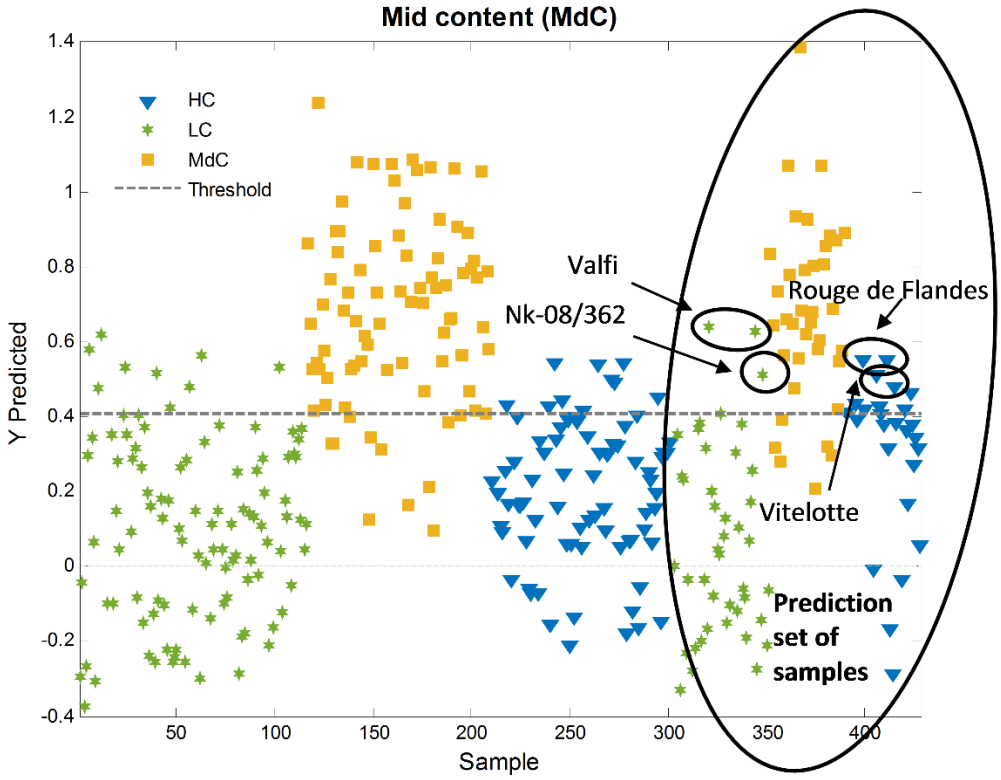
Some misclassification was also found in MdC group in which a few samples belonging to Blue Star and Fenton varieties, with the lowest values of phytochemicals, were classified into the LC group. On the other hand, a few samples belonging to British Columbia Blue and Blue Congo were identified as having HC. Finally, in HC group, a small number of samples from Vitelotte and Rouge de Flandes were incorrectly classified into MdC. In this group, again, those varieties were the ones with the lowest values of phytochemicals in the group. As commented before, cv. Vitelotte was first included in HC group but according to its TSP content should have been comprised in MdC so, it is not surprising that some of the samples belonging to this cultivar were misclassified.

The results obtained demonstrated that PLS-DA classification technique enabled identification of different potato cultivars with low, mid and high content of TSP, TMA and HAC.

Figure 4.6 is the graphical representation of the confusion matrix for each group included in the study. The horizontal black dotted line indicates the threshold above a sample is assigned to a particular class. Therefore, in figure 4.6a all samples located above the threshold are classified as belonging to the LC group. It can be seen that there are a few samples belonging to the MdC and HC groups that according to this plot were classified as LC. Similarly, all the samples above the threshold in figure 4.6b and c were classified as belonging to MdC and HC respectively. In these plots, the prediction set of samples appeared inside a circle for the sake of easier visualization.


Low content (LC)
Y Predicted
Sample
Prediction set of samples
HC
LC
MdC
Threshold

(a)


Mid content (MdC)
Y Predicted
Sample
HC
LC
MdC
Threshold
Valfi
Nk-08/362
Rouge de Flandes
Vitelotte
Prediction set of samples

(b)

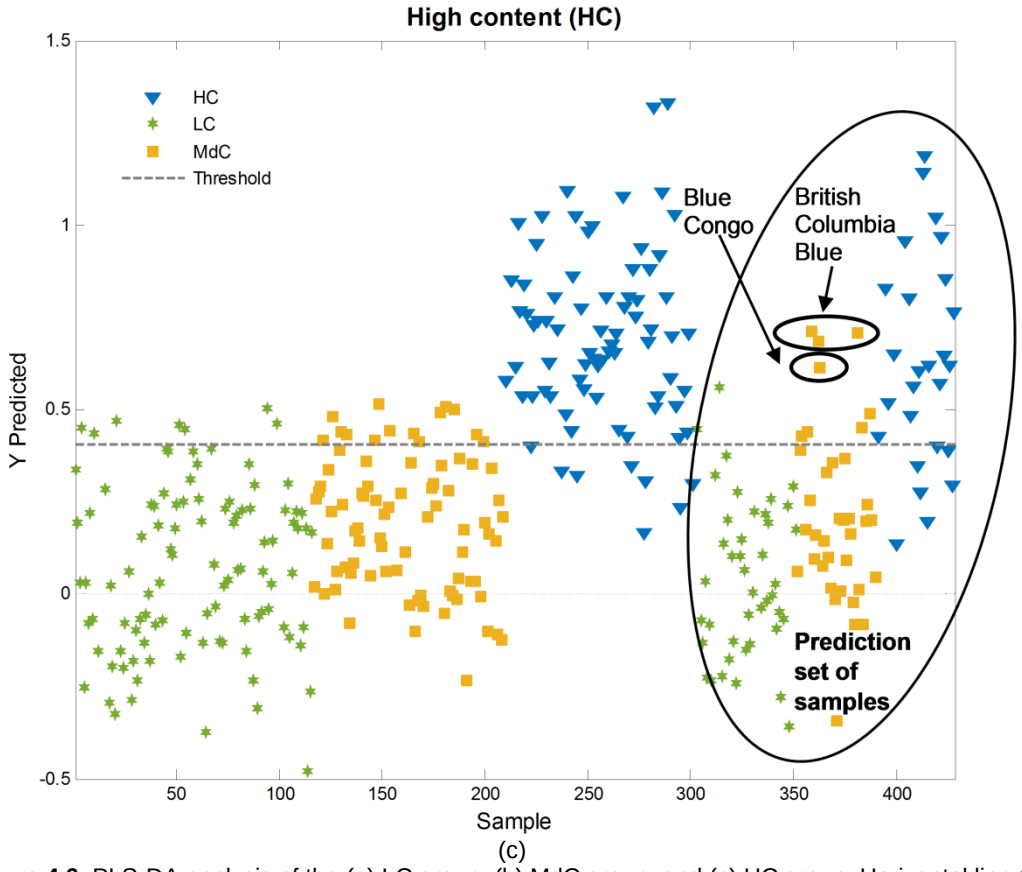


Figure 4.6. PLS-DA analysis of the (a) LC group, (b) MdC group, and (c) HC group. Horizontal line (---) indicates the threshold above which a sample is assigned to a particular group.

4.3.1.4. Total carotenoids model

Figure 4.7 shows the varieties included in each of the three groups in the second PLS-DA. All 18 varieties were included in this model. 70% of the total number of samples (n=471) were randomly selected for the calibration set (n_{cal}=315) and 30% for validation (n_{val}=156). As can be seen in figure 4.7, the LC group comprised 192 samples belonging to 7 cultivars with TC content below 0.02 g LE kg^{-1} FW, while the second group (MdC) had a total of 132 samples from 5 varieties with TC values between 0.02 and 0.03 g LE kg^{-1} FW. Lastly, the HC group was formed by 147 samples from 6 different varieties with TC levels above 0.03 g LE kg^{-1} FW.

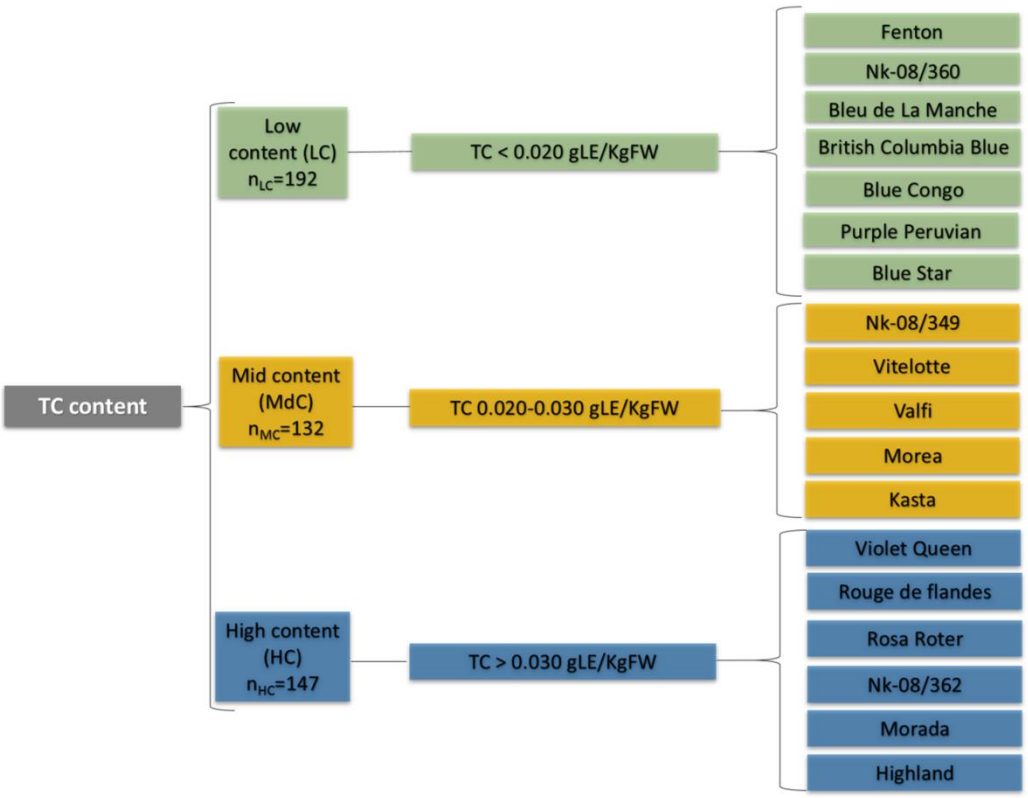


Figure 4.7. Flowchart of the distribution of varieties according to their content of total carotenoids (TC).

A total of 8 LVs were selected in this PLS-DA model explaining the 99.96% of variance. Table 4.5 shows the percentage of correctly classified samples into the three groups. It is observed that it was not possible to obtain an accurate classification of both LC and MdC groups since only a small number of samples were correctly classified into the group they belonged to and the fact that a considerable percentage of samples were classified in either of the other two groups. Thus, 17.2% and 31.2% of samples belonging to LC were classified as MdC and HC respectively, while 39.1% and 30.43% of samples of MdC group were classified as LC and HC respectively in the Pred set. On the other hand, a high percentage of correctly classified samples was achieved in the HC group of around 80%. These results, suggests that NIRS technology could be used for screening processes in order to identify samples with high content of TC.

Table 4.5. Confusion matrix of the three groups: low content (LC), mid content (MdC) and high content (HC), in the TC PLS-DA model for Cal, CV and Pred.

		Actual Group (%)		
Set	**Predicted group (%)**	LC	MdC	HC
Calibration set (Cal)	LC	**45.6**	26.8	4.25
	MdC	20.0	**43.9**	14.9
	HC	34.4	29.3	**80.8**
Cross-validation set (CV)	LC	**46.4**	25.6	4.25
	MdC	20.0	**42.7**	14.9
	HC	33.6	31.7	**80.8**
Prediction set (Pred)	LC	**51.6**	39.1	6.52
	MdC	17.2	**30.4**	13.0
	HC	31.2	30.4	**80.4**

4.3.1.5. Vitamin C model

In figure 4.8 the cultivars included in each of the three groups in the third PLS-DA are shown. Once again, samples (n=471) were randomly divided between calibration and validation sets (n_{cal}=315, n_{val}=156). In this analysis, the LC group comprised 174 samples belonging to 7 varieties having less than 0.07 g kg^{-1} FW of VC, the second group, MdC, covered 162 samples belonging to 6 cultivars with VC content between 0.07 and 0.092 g kg^{-1} FW and finally, HC group contained 135 samples from 5 different varieties with VC concentrations above 0.092 g kg^{-1} FW.

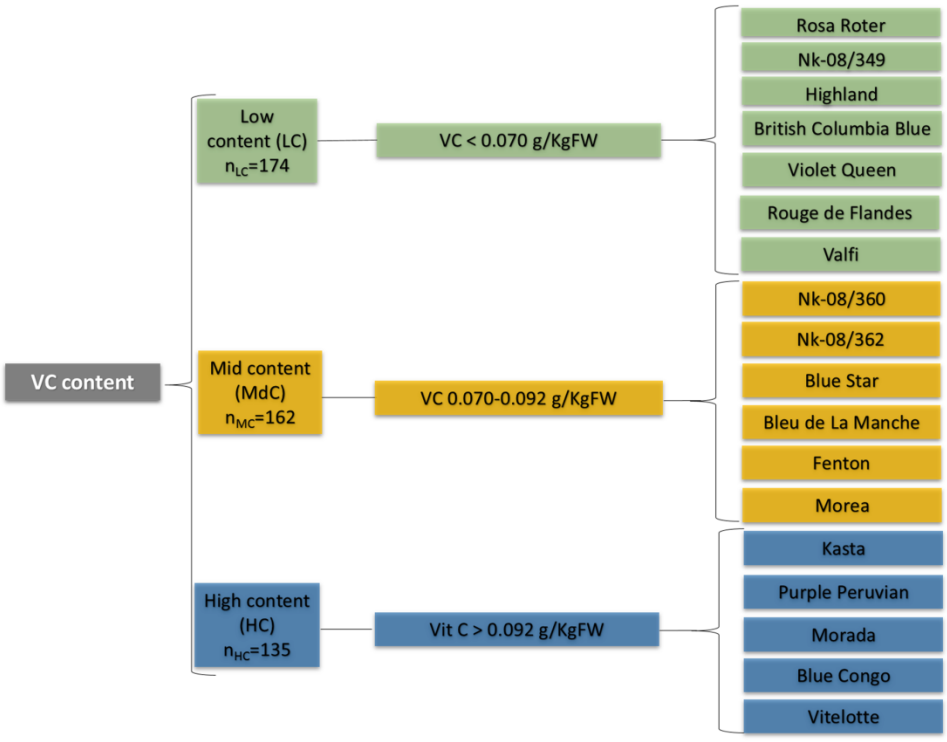


Figure 4.8. Flowchart of the distribution of varieties according to their content of vitamin C (VC).

A total of 5 LVs were selected explaining the 99.8% of the variance. Low classification rates, below 63.0%, were obtained for this PLS-DA (data not shown) for CV and Pred sets. Besides, high misclassification was found between LC and HC groups. Around 30.0% of samples belonging to HC group were wrongly classified as LC in CV and Pred sets, suggesting that the results were not reliable. It is worth mentioning that vitamin C content in potatoes highly depends on many factors and can vary considerably from one area to another. Thus, it becomes very challenging to perform robust classification methods for this compound. Moreover, significantly lower contents of VC were found compared to other authors as mentioned before, due to the effect of cold temperatures during storage. Therefore, these facts were considered to be responsible for the low classification rates obtained in this PLS-DA model.

As explained before, there is little literature concerning qualitative analysis of potatoes by NIR spectroscopy despite its potential. Even so, some authors successfully investigated the discrimination of two categories of potato samples regarding their recoverable protein content. An overall 87.5% of correctly classified samples was achieved in that study (Fernández-Ahumada *et al.*, 2006). Other authors were able to classify potato chips by source of frying oil by combining NIR spectroscopy and SIMCA (Shiroma & Rodriguez-Saona, 2009).

Recently Rady and Guyer (2015a) studied different discriminant algorithms to classify potato samples based on their sugar level. Potato samples from two different varieties, Frito Lay 1879 (FL) and Russet Norkotah (RN) and from two seasons (2009 and 2011), accounting for a total of 990 tubers were used in that study. Classification rates of 81% and 100% for FL and RN were obtained for whole tubers based on glucose levels using combined classifiers and PLS-DA respectively. For sliced samples, 83% and 81% of CC samples for FL and RN were achieved using LDA and PLS-DA. Lower classification rates were obtained based on sucrose levels with 71% and 79% of CC samples for whole tubers of FL and RN respectively by PLS-DA and LDA. For sliced samples, 75% and 82% of FL and RN were correctly classified with LDA and PLS-DA algorithms respectively.

A successfully classification of samples into three levels based on their content of TSP, TMA and HAC was achieved. Figure 4.6a-c confirms the results obtained, as it shows that the majority of the samples from each group were correctly classified into them. However, there was some misclassification, as previously described, but in overall, very good rates of discrimination were obtained.

Therefore, according to this study, NIR spectroscopy combined with PLS-DA was capable of accurately identifying samples containing different levels of TSP, TMA and HAC belonging to this collection of 18 purple- and red-fleshed potatoes.

Regarding TC content, it was found that NIRS technique was only capable of identifying samples with high content of this compound among the varieties analysed. Information that could be useful for screening processes.

Finally, the classification results achieved according to the content of vitamin C suggested that it was not possible to obtained a reliable classification of varieties regarding their VC content by NIRS technology.

Nevertheless, these findings are of great importance considering the continuously increasing demand for quality control of food products among consumers and authorities. The outcome of this research could be considered as a screening step for future potato breeding programmes. Further research is advisable including a larger set of samples comprising not only coloured varieties but yellow-skinned tubers as well.

4.3.2. Quantitative analysis

As mentioned in section 4.2.3.2, the dataset used comprised 135 spectra in year 1 (2012) and 228 spectra in year 2 (2013). An overview of the cultivar characteristics is given in table 4.2 and table 4.6 provides an overview of the concentration of the different parameters analysed in calibration and prediction sets.

Table 4.6. Overview of concentration of the different parameters and number of samples (*n*) in the calibration and prediction datasets.

		Calibration Set				***Prediction Set***			
	Parameter	*n*	Range	Mean	SD	*n*	Range	Mean	SD
2012	DM	90	17.02-25.05	19.85	1.86	45	17.04-24.00	19.90	1.86
	Starch	90	51.04-72.83	63.53	4.20	45	51.58-72.38	63.18	4.88
	RS	90	0.22-1.83	0.76	0.35	45	0.20-1.34	0.73	0.31
	CP	90	6.58-14.2	10.15	1.67	45	7.00-14.58	10.02	1.63
	N	90	1.05-2.26	1.62	0.27	45	1.12-2.33	1.60	0.26
2013	TSP	152	0.75-14.28	3.95	3.12	76	0.91-14.44	4.20	3.38
	HAC-ABTS	152	2.95-39.89	16.11	6.61	76	3.82-40.59	16.69	6.55
	HAC-DPPH	152	0.90-40.37	9.55	8.53	76	1.40-39.82	9.78	8.31

As it can be seen in table 4.6, the range of TSP content was wide as it was the range of potatoes used in this study. As stated before, purple- and red-fleshed potatoes are a much richer source of TSP than yellow-fleshed ones. Same behaviour was observed for HAC analysed by the two procedures previously described.

Figure 5.9a-d shows the spectra of samples after the different pre-treatments applied. Spectra of lyophilized samples belonging to year 2 (2013) from 1100 to 2300 nm are plotted. Figure 4.9a shows the MSC-treated NIR reflectance spectra of the 228 potato samples in which bands representative of specific functional groups were observed. These groups are mainly related to moisture and carbohydrate components of tubers (Shiroma & Rodriguez-Saona, 2009). Thus, a sharp absorbance peak was observed at around 1440 nm related to water and starch, the two main components of potatoes and at 1940 nm related to water. Furthermore, absorption bands were observed around 1200 nm related with CH second overtone of the CH_2 and CH_3 groups (Osborne *et al.*, 1993). Figure 4.9b shows the MSC and 1st derivative spectra of samples. First derivative by Savitzky-Golay (SG) method was calculated by second order polynomial and 9 window points. Figure 4.9c shows the 1st derivative spectra by SG calculated by second order polynomial and 7 window points. In figure 4.9d the SNV followed by detrend pre-treated spectra are shown.

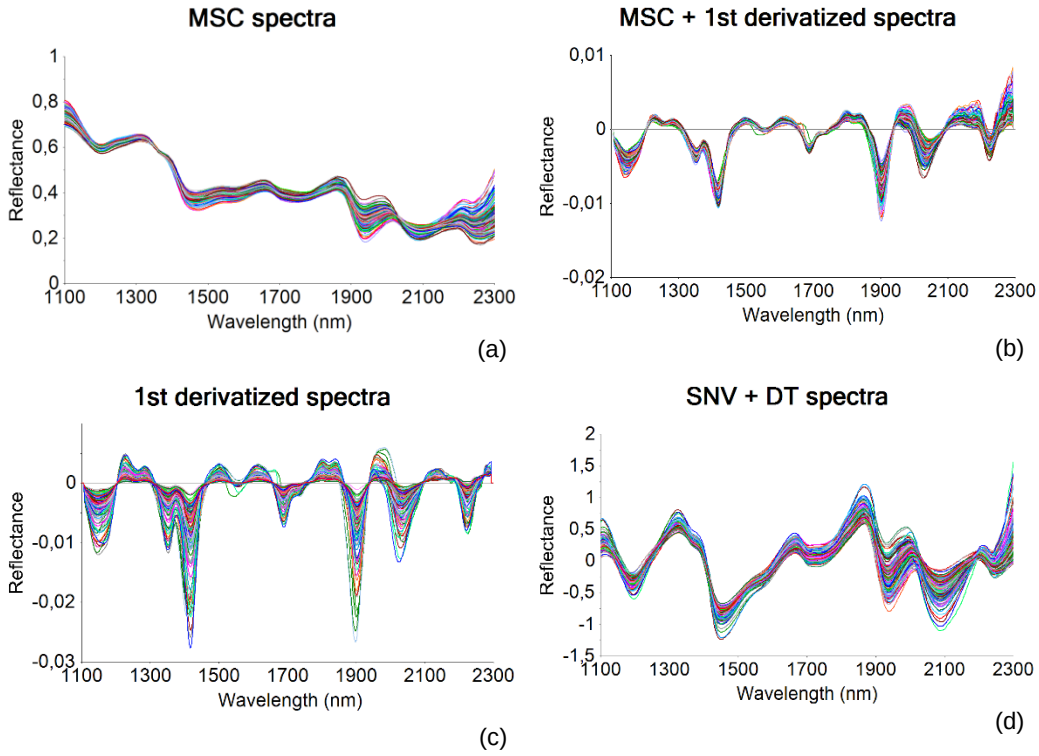


Figure 4.9. Pre-treated NIR spectra of the lyophilized potato tubers belonging to year 2 (2013)., (a) MSC-treated NIR spectra, (b) MSC + 1st derivatized spectra, (c) 1st derivative-treated spectra, and (d) SNV + Detrend-treated spectra.

Table 4.7 shows the results obtained in cross-validation and prediction datasets from the two years analysed. The latent variables (LV) required for models varied from 6 to 14. The number of samples (*n*) shown on the table is computed after identification and elimination of a few outliers.

Only results obtained with lyophilized samples are shown in table 4.7, as good PLS models could not be obtained when working with intact tubers. This could be due to the following reasons; first, NIR spectral acquisition was performed on intact unpeeled tubers, while lyophilisation was accomplished with peeled tubers, it was explained in section 2.1.3 that skin of potatoes is a highly structured tissue at the microstructural level (Aguilera, 2000) and thus, light crossing through it can be highly scattered (Saeys *et al.*, 2008). Therefore, the correlation between NIR spectral data acquired from intact tubers and the chemical composition quantified from lyophilized peeled tubers may not give the expected results. In addition, for the lyophilisation of samples, more than one tuber were selected for each cultivar, so, chemical concentration was quantified not for each sample separately but for a group of samples and these data was then correlated with the spectral information for each sample. One possible solution for this for next time could be to lyophilize each sample separately and quantified its chemical composition in order to reduce the laboratory errors and build more robust calibration models.

Nevertheless, it can be seen that in year 1 (2012) good PLS models were obtained for CP and nitrogen with an R^2 value of 0.86 for the both prediction datasets and a RPD value of 2.83 and 2.70 respectively. According to the RPD classification established by Saeys *et al.* (2005), those models are considered good and regarding their R^2 value, they could be used for most applications. Figure 4.10 shows the plot of weighted regression coefficients for the PLS regression model develop to estimate crude protein content of lyophilized potatoes. This plot indicates that multiple spectral wavelength regions along the NIR spectral range played an important role in the PLS model for prediction of CP. Figure 4.11 shows the PLS model develop for the prediction of CP content of samples.

Some authors have studied the determination of CP content in mashed and homogenized potatoes by NIRS obtaining relatively low correlation coefficients ranging between 0.36 and 0.78 (Fernández-Ahumada *et al.*, 2006; Haase, 2006). Authors attributed these low values to the reduced range of this constituent in their samples and the high values of the reference method errors.

Other authors have investigated the correlation between NIRS and nitrogen absorption of potato plants to calculate proportions of supplemental nitrogen fertilizer in potato crops obtaining low SEP values (0.09%) (Young *et al.*, 1997).

On the other hand, good models for DM, starch and RS could not be obtained, probably due to the way these compounds were quantified. DM and starch were calculated indirectly by the UWW measurement first, followed by equation (15) adopted

from Barredo (1993) as explained in section 2 of the Appendix. The majority of the authors that have studied the estimation of DM by NIRS as shown in section 2.4 calculated DM directly by oven dried. Therefore, this could be the reason for the low correlations obtained. In the case of sugars, as commented in section 2.4, due to the small presence of them in potato tubers, its estimation by NIRS is complicated. The results obtained in this study are similar to those obtained by Mehrubeoglu and Cote (1997) and Haase (2011).

Table 4.7. Cross-validation and validation statistics for PLS Models developed to predict different compounds in potatoes.

	Parameter	Pre-process	*n*	LV	R^2_{CV}	RMSECV	R^2_P	RMSEP	SEP	RPD
2012	DM	SNV+ DT	122	14	0.56	1.16	0.43	1.37	1.36	1.37
	Starch	MSC+1st der	122	7	0.61	2.11	0.40	3.18	3.29	1.28
	RS	1st der+SNV	113	7	0.49	0.25	0.42	0.24	0.25	1.24
	CP	MSC+1st der	128	11	0.90	0.49	**0.86**	0.59	0.68	**2.83**
	N	MSC+1st der	128	11	0.90	0.08	**0.86**	0.10	0.11	**2.70**
2013	TSP	1st der (9p)	219	8	0.84	1.20	**0.83**	1.41	1.33	**2.54**
	HAC-ABTS	1st der (7p)	219	9	0.51	4.29	0.58	4.15	4.10	1.59
	HAC-DPPH	1st der (9p)	228	6	0.67	4.84	0.70	4.52	4.51	1.84

It can also be seen in table 4.7 that PLS models developed with samples belonging to year 2 (2013) reported good results for the prediction of TSP with an R^2 value of 0.83 for the validation set and a RPD value of 2.54 (figure 4.12). Therefore, according to these two statistics, this is considered a good model. On the other hand, the estimation of HAC by the DPPH method reported better results than ABTS with an R^2 value of 0.70 and a RPD value of 1.84. These results suggests that this model could be used for screening and approximate calibrations while the PLS model developed to estimate HAC quantified by the ABTS method, with an R^2 value of 0.58 and a RPD of 1.59, could only be used to distinguish samples with high and low content of HAC.

High correlations for cross-validation and prediction were obtained with small RMSECV and RMSEP values regarding phenolic content. Similarly, in a research developed by Shiroma-Kian *et al.* (2008), estimation of the polyphenol compounds in lyophilized potatoes using Fourier transform infrared spectroscopy (FTIR) was performed. Authors achieved excellent performance statistics with correlations >0.99 for the cross-validated PLS models with a SECV value of 4.17.

In a related work, Bonierbale *et al.* (2009) estimated the total and individual carotenoid concentration in *Solanum phureja* potatoes by NIRS obtaining R^2 values between 0.63 and 0.92 with standard errors of predictions ranging from 20 to 610 µg/100g DW.

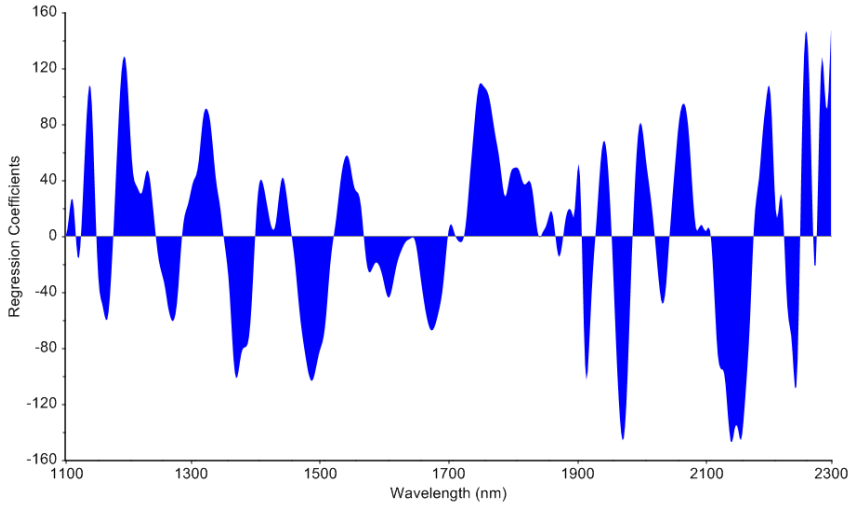


Figure 4.10. Weighted regressions plot showing important variables in modelling CP content in lyophilized potato samples from year 1 (2012).

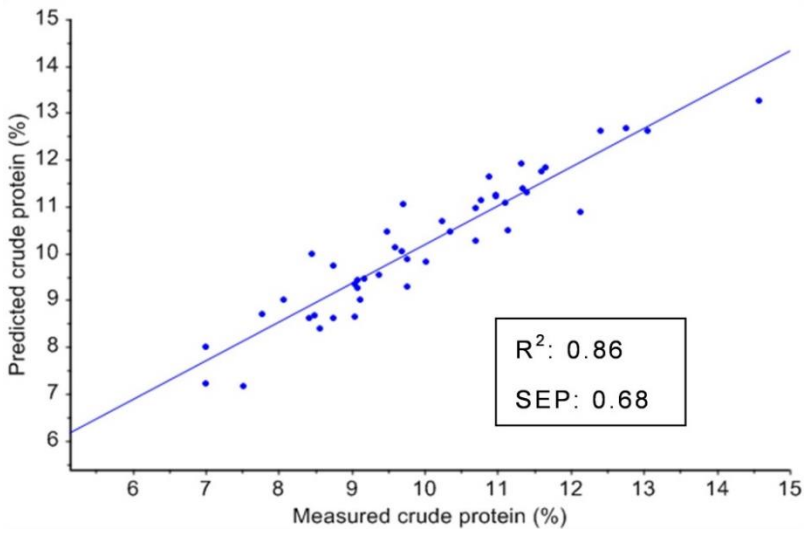


Figure 4.11. PLS regression plot of predicted *versus* measured CP content (%) in the 1100-2300 nm spectral range pre-processed with MSC + 1st derivative.

In figure 4.13 the plot of weighted regression coefficients for the PLS regression model develop to estimate TSP content of lyophilized potatoes is shown. Similar to figure 4.10, in this plot can be seen that multiple spectral wavelength regions were important in the PLS model for prediction of TSP and specially the spectral region between 1700 and 1850 nm associated with 1st stretching overtone of C=H and the 2000-2100 nm spectral region related to starch absorption bands (Osborne *et al.*, 1993).

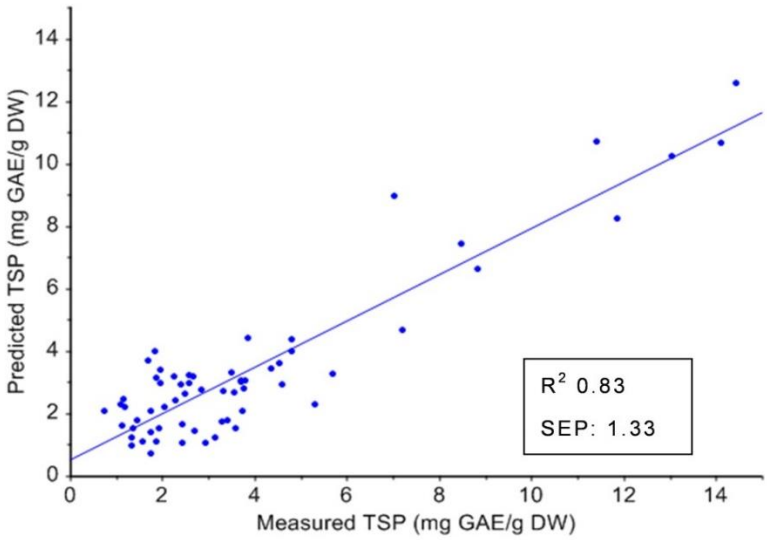


Figure 4.12. PLS regression plot of predicted *versus* measured TSP content (%) of lyophilized samples in the 1100-2300 nm spectral range pre-processed with 1st derivative.

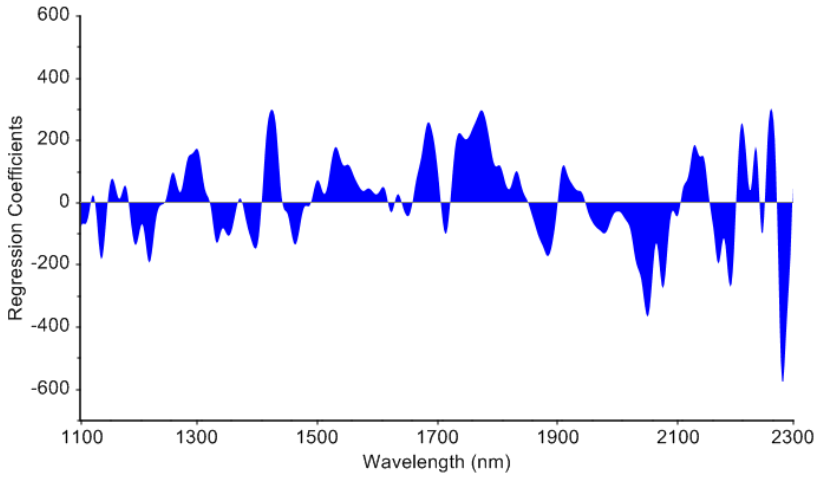


Figure 4.13. Weighted regressions plot showing important variables in modelling TSP content in lyophilized potato samples from year 2 (2013).

From figure 4.12 can be observed from that there are two groups of samples according to their TSP content. The larger one with TSP values up to 6 mg GAE/g DW that includes yellow- and red-fleshed cultivars, and a second group of samples with larger TSP values between 6-15 mg GAE/g DW corresponding to some red- and purple-fleshed tubers.

4.4. Conclusions

The identification of potato genotypes rich in phenolic compounds, carotenoids and antioxidant capacity is a key step for both identifying phytochemical-rich food products and for developing breeding lines with high concentrations of bioactive compounds. The automatic and non-destructive characterization of cultivars and breeding lines with

different levels of bioactive compounds through near-infrared spectroscopy has been shown useful for discriminating and classifying potato tubers in terms of phytochemical content.

On the other hand, quantitative analyses of different potato compounds have shown that good PLS regression models could be obtained for the prediction of CP, nitrogen and TSP in a large collection of potato varieties. Moreover, PLS models obtained for the estimation of HAC could be used for screening and approximate calibrations. This is a very important point, since most of the varieties used in this study are included in breeding programs where it becomes essential to identify samples with high and low hydrophilic antioxidant capacity.

Acknowledgments

This work was financed within the frame of INIA's project RTA2013-00006-C03-01-03, the Basque Government and the Universidad Pública de Navarra through the concession of a predoctoral research grant.

Chapter 5.

Non-destructive detection of blackspot in potatoes by VIS-NIR and SWIR hyperspectral imaging

This chapter is under revision as:

- López-Maestresalas, A., Keresztes, J. C., Goodarzi, M., Arazuri, S., Jarén, C., & Saeys, W. (2016). Non-destructive detection of blackspot in potatoes by VIS-NIR and SWIR hyperspectral imaging. *Food Control.*

Non-destructive detection of blackspot in potatoes by VIS-NIR and SWIR hyperspectral imaging

Ainara López-Maestresalas[1], Janos C. Keresztes[2], Mohammad Goodarzi[2], Silvia Arazuri[1], Carmen Jarén[1], Wouter Saeys[2]

[1] Department of Projects and Rural Engineering, Universidad Pública de Navarra, Campus de Arrosadia 31006, Navarra, Spain

[2] KU Leuven Department of Biosystems, MeBioS, Kasteelpark Arenberg 30, 3001 Leuven, Belgium

Abstract

Blackspot is a subsurface potato damage resulting from impacts during harvesting. This type of bruising represents substantial economic losses every year. As the tubers do not show external symptoms, bruise detection in potatoes is not straightforward. Therefore, a non-destructive and accurate method capable of identifying bruised tubers is needed. Hyperspectral imaging (HSI) has been shown able to detect other subsurface defects such as bruises in apples. This method is non-destructive, fast and can be fully automated. Thus, its potential for blackspot detection in potatoes was investigated in this study. Two HSI setups were used, one ranging from 400 to 1000 nm, named Visible-Near Infrared (VIS-NIR) and another covering the 1000-2500 nm range, called Short Wave Infrared (SWIR). 188 samples belonging to 3 different varieties were divided in two groups. Bruises were manually induced and samples were analysed 1, 5, 9 and 24 hours after bruising. PCA, SIMCA and PLS-DA were used to build classifiers. The PLS-DA model performed better than SIMCA, achieving an overall correct classification (CC) rate above 94% for both hyperspectral setups. Furthermore, more accurate results were obtained with the SWIR setup at the tuber level (98.56 *vs.* 95.46% CC), allowing the identification of early bruises within 5 hours after bruising. Moreover, the pixel based PLS-DA model achieved better results in the SWIR setup in terms of CC samples (93.71 *vs.* 90.82%) suggesting that it is possible to detect blackspot areas in each potato tuber with high accuracy.

Keywords: VIS-NIR; SWIR; hyperspectral imaging; potato; blackspot; damage.

5.1. Introduction

Blackspot bruise in potato is an internal damage mainly produced from impacts either between the tubers and hard surfaces or between each other during mechanical harvesting and subsequent handling (Fluck & Ahmed, 1973; Hesen & Kroesbergen, 1960; Mathew & Hyde, 1997). This type of bruise appears at the subsurface and most frequently at the stem end of the tubers due to the fact that the radius of curvature is smaller there (Sawyer & Collin, 1960). The resulting blue-black discoloration of the damaged tissue is a consequence of oxidation of tyrosine by polyphenol oxidase (Dean *et al.*, 1993; Mohsenin, 1986). The damaged tissue tends to absorb more oil during frying (Baritelle *et al.*, 2000), resulting in after-cooking darkening, one of the most undesirable effects reported by consumers (Wang-Pruski & Nowak, 2004).

As explained in section 1.2, damage to tubers results in important economic losses. Therefore, to reduce the current losses and regain customers' trust, the potato industry demands a non-destructive technique able to detect this internal damage before tubers reach the market. Hyperspectral imaging, a technique combining the principles of spectroscopy and imaging has been applied to subsurface defect detection in fruit and vegetables, such as apples (ElMasry *et al.*, 2008; Lu, 2003; Xing & De Baerdemaeker, 2005; Xing *et al.*, 2007), pears (Zhao *et al.*, 2010) and mushrooms (Gowen *et al.*, 2008). As reported in section 2.5, in the case of potatoes the usefulness of hyperspectral imaging has been reported for the discrimination between potato tubers and clods (Al-Mallahi *et al.*, 2008; Al-Mallahi *et al.*, 2010), the detection of hollow heart (Dacal-Nieto *et al.*, 2011b) and the detection of common scab (Dacal-Nieto *et al.*, 2011a). Thybo *et al.* (2004) were able to identify internal bruises in potato slices of cultivar Saturna by applying magnetic resonance, another type of imaging technique.

Considering the promising results reported for hyperspectral imaging techniques in identifying subsurface defects in different fruit and vegetables and other potato defects, the objective of this study was to evaluate the potential of hyperspectral imaging to identify and map blackspot affected areas in intact potatoes in order to comply with the specific objectives 4 and 5 from section 1.4. For this, two wavelength regions of the electromagnetic spectrum, Visible-Near Infrared (VIS-NIR, 400-1000 nm) and Short Wave Infrared (SWIR, 1000-2500 nm) were considered and compared. The latter goal addresses the specific objective 6 explained in section 1.4.

5.2. Material and methods

5.2.1. Design of the experiment

Samples of 3 different cultivars were used accounting for a total of 188 tubers. Some of the samples were provided by The Basque Institute for Agricultural Research and Development (NEIKER-Tecnalia). The sample set was divided into two groups of 94 tubers each for further analyses. One group was kept intact while the other was induced a bruise. After that, all of the tubers were scanned with the hyperspectral imaging setup controlled with LabVIEW 8.5 software (National Instruments Corporation, Austin, TX), at different times as it is explained in the following sections. These steps were carried out at the University of Leuven, KU Leuven in Belgium. Once all the images were acquired and processed, statistical analyses were performed at Universidad Pública de Navarra to discriminate healthy tubers from bruised-induced ones using MATLAB and PLS_Toolbox software. Figure 5.1 shows the steps followed in this study to detect blackspot in potatoes.

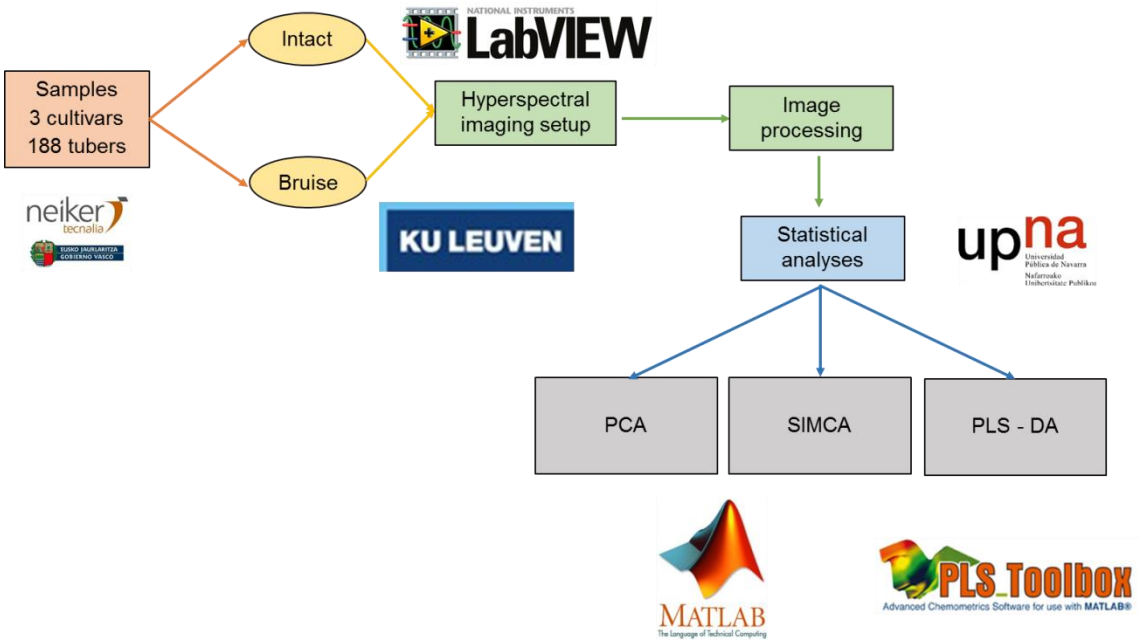


Figure 5.1. Flowchart for the blackspot detection in potatoes by hyperspectral imaging systems.

5.2.2. Sample preparation

A total of 188 potato tubers of three different cultivars (Hermes, Bintje and Magnum) harvested in 2013 were analysed. Samples from cv. Hermes (109 tubers) were provided by The Basque Institute for Agricultural Research and Development (NEIKER-Tecnalia), Spain and were sent to KU Leuven Department of Biosystems, MeBioS, Leuven, Belgium, for the measurements. Samples from cv. Bintje and cv.

Magnum, consisting of respectively 44 and 35 tubers, were supplied by a local farmer from Leuven, Belgium. The tubers were randomly divided in two groups of equal size. The samples of the first group (n_b=94) were subjected to impact in order to induce internal bruising, while the others (n_h=94) served as the control group. Prior to analysis samples were kept in a refrigerator at 4°C. Then, samples were washed and weighed.

In order to induce the bruises, the tubers were dropped inside a cylinder 300 mm above an impactor facing the stem end (figure 5.2). They were left to fall free and hit a hemispherical head of 25 mm in diameter attached to a circular flat plate. The calculated impact energy varied between 303 mJ and 994 mJ depending on the mass of the potatoes. After impact, tubers were kept in a hot climate control chamber (Weiss WKL 100, Weiss Umwelttechnik GmbH, Reiskirchen-Lindenstruth, Germany) at a temperature of 34°C with 95% relative humidity for 24 hours, because combination of both high temperature and high humidity has been reported to promote a faster development of the bruises (Baheri, 1997).

Samples were then scanned with the hyperspectral imaging system 1, 5, 9 and 24 hours after the impacts, as will be explained in the following section. All the samples were placed with the stem end facing the hyperspectral imaging system. After all measurements, samples were peeled and photographed with a standard RGB camera (Powershot 1100D, Canon Corporation, Japan) to check whether the healthy ones had already bruises and the bruise-induced ones had developed them or not.

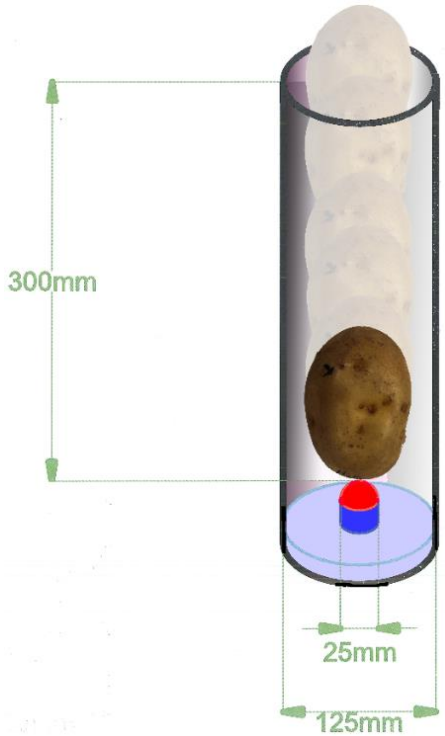


Figure 5.2. Schematic diagram of the system used to induce bruises.

Samples were divided into 2 classes for statistical analysis. Class 1 (n_{c1}=94) corresponding to the Healthy group and Class 2 (n_{c2}=376) including bruised samples measured at the different times after bruising: 1 hour group, 5 hours, 9 hours and 24 hours. Each group containing 94 samples.

5.2.3. Hyperspectral imaging

The hyperspectral imaging was performed at the KU Leuven Department of Biosystems, MeBioS, Leuven, Belgium. Potato samples were scanned on a setup for hyperspectral imaging that consisted of: a transportation plate, an illumination unit, two hyperspectral cameras (one for the VIS-NIR range from 400 to 1000 nm and one for the SWIR range from 1000 to 2500 nm) and a computer. The VIS-NIR hyperspectral camera used in this study consisted of a CCD camera (TXG14, Baumer, Germany) with a 1392 by 1040 pixel image resolution coupled to a prism-grating-prism-based imaging spectrograph (ImSpector V10, Spectral Imaging Ltd., Oulu, Finland), and a focusing lens (Canon TV Lens, VF 25 mm, f/0.95, Japan). The SWIR hyperspectral camera (HS SWIR XS-M320C4-60, Headwall Photonics Inc., Fitchburg, MA) consisted of an MCT camera (XEVA MCT-2140, Xenics, Leuven, Belgium) with a 320 by 256 pixel resolution with a reflective concentric grating (HS SWIR XS-M320C4, Headwall Photonics Inc., Fitchburg, MA), a slit of 60 µm, and focusing lens (Oles 22.5, Specim Ltd, Oulu, Finland) with a focal length of 22.5 mm. Six halogen lamps (DECOSTAR ALU 12 V-20 W-36°, OSRAM, Germany) were used to illuminate the samples. Four of them were arranged on an arc frame, while the other two were set one at the front of the sample and another at the back of it to achieve homogeneous illumination of the scanned area (figure 5.3). The transportation plate used for moving the samples under the cameras consisted of a computer controlled translation stage (TLA 15-400, Franke GmbH, Aalen, Germany). The entire setup was controlled by a computer equipped with LabVIEW software.

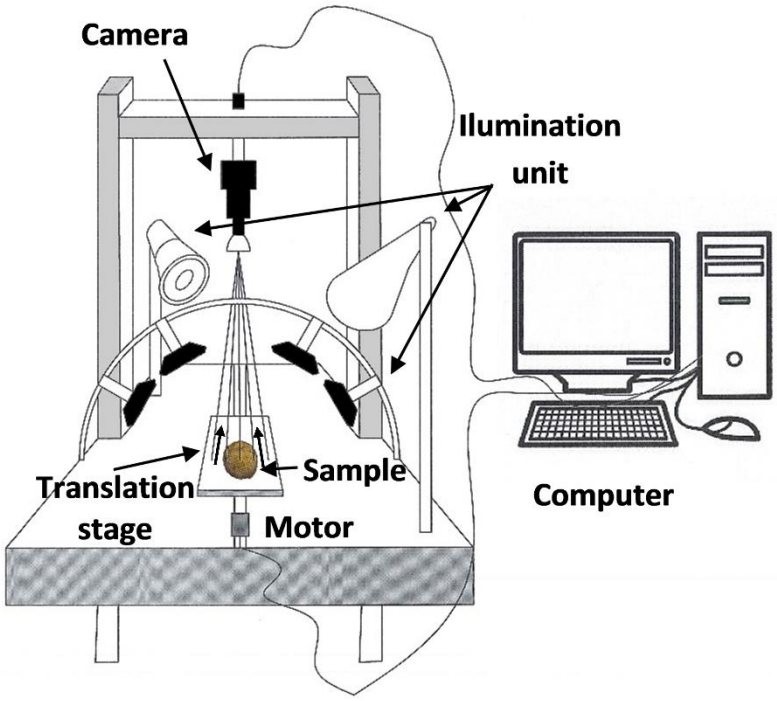


Figure 5.3. Schematic diagram of the hyperspectral imaging system used for potato scanning.

The exposure time was optimized at 35 ms and 2 ms for the VIS-NIR and SWIR cameras, respectively, in order to maximize the spectral signal-to-noise ratio while avoiding saturation of specular reflective regions. The translation stage speed was set to 100 mm/s and 200 mm/s and images were captured consequently in intervals of 0.3 mm and 0.1 mm for the VIS-NIR and SWIR cameras, respectively.

5.2.3.1. Reflectance calibration

Three images were acquired for the reflectance calibration with both the VIS-NIR and SWIR cameras. First, white reference measurements (W_r) were obtained using a white calibration tile for both the VIS-NIR and SWIR range (Spectralon® Reflectance Standards 75%, RSS-08-010, Labsphere, North Sutton, USA). Then, dark references (D) were acquired with the illumination switched off and the camera lens covered by a cap. Both white and dark references were acquired every 1, 5, 9 and 24 hours. Finally, all images of samples were first scanned with one hyperspectral setup at each time (1, 5, 9 and 24 hours) and then with the other hyperspectral setup. Equation (9) was used to convert the raw intensity values in the hyperspectral images into relative reflectance values:

$$R = \frac{I_s - I_D}{I_{Wr} - I_D} \quad (9)$$

where, R is the relative reflectance, I_S corresponds to the intensity value acquired on the sample, I_D is the intensity acquired for the dark reference and I_{Wr} is the intensity acquired on the white reference tile.

5.2.3.2. Segmentation

The first step in image segmentation consists in separating the region of interest, namely the potato, from the background and saturated areas. In this study, two masks were applied to remove the background and the saturated pixels in each image, as these pixels do not contain any information of the quality of the tubers. It should be mentioned that no specular reflections were observed on the bruised areas. To remove the background in the VIS-NIR hypercubes a threshold of 0.10 was applied to the image at 854 nm, where the intensity was the highest, while a threshold of 0.09 was applied to the image at 1106 nm for the SWIR hypercube. All pixels with values below those thresholds were labelled as background. This step produced a binary image of the whole potato including the saturated pixels (figure 5.4b). Then, a high mask was applied to select saturated pixels by thresholding at a value of 0.55 and 0.57 for VIS-NIR and SWIR respectively to produce a binary image of saturated pixels only (figure 5.4c). Finally, the interesting areas of potato were isolated by subtracting saturated pixels (figure 5.4c) from the first binary image (figure 5.4b) to produce a mask containing only the no-saturated areas of the potato in a black background (figure 5.4d). This isolated potato was then used as the region of interest (ROI). The sequential procedure for image segmentation is displayed in figure 5.4.

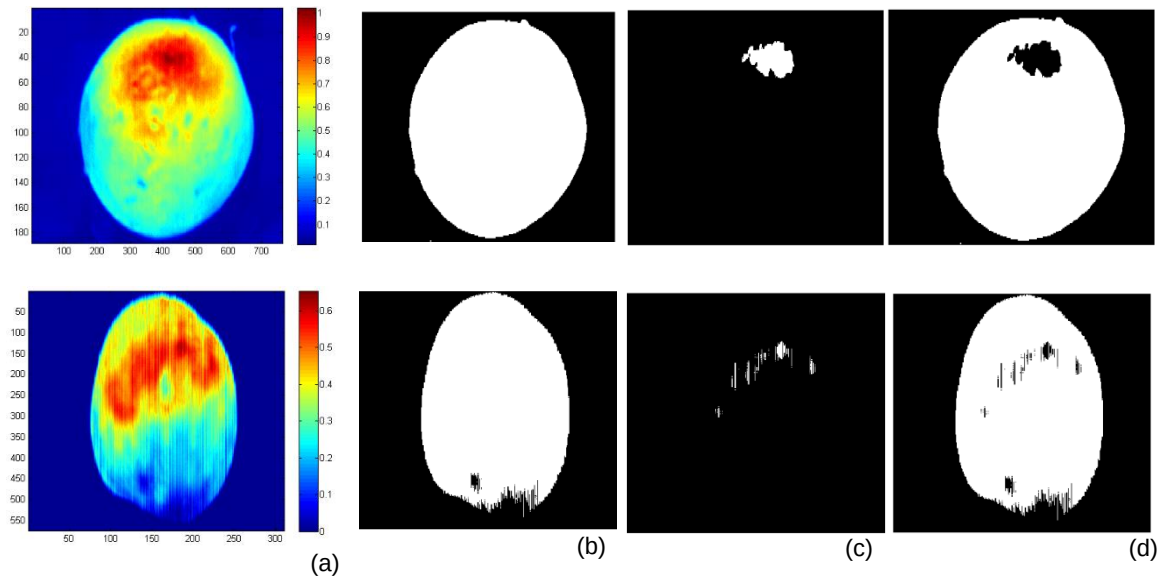


Figure 5.4. Low and high masks of a potato sample for VIS-NIR and SWIR images, (a) images in bands 854 and 1106 nm for VIS-NIR and SWIR, respectively, were used to select a threshold, (b) images after applying the low mask, (c) images after applying the high mask, (d) images after applying the low and high masks for VIS-NIR and SWIR setups.

5.2.4. Multivariate data analysis

Data pre-processing and classification modelling were performed in MATLAB R2014a (The MathWorks, Natick, MA) using the PLS_Toolbox (Eigenvector Research Inc., Wenatchee, WA).

5.2.4.1. Spectral pre-processing techniques

Before any other pre-treatment, each mean spectrum of every sample was smoothed by a 15 points Savitzky-Golay filtering operation. Then, different combinations of the methods described above were used for model building. The effect of no pre-treatment at all was also analysed. 1st and 2nd derivatives by Savitzky-Golay (SG) method were calculated by second order polynomial and 15 window points. Finally, all pre-treated data, as well as the non pre-treated data, were Mean Centred (MC).

5.2.4.2. Unsupervised analysis (PCA)

Principal Component Analysis (PCA) was used in the first place to understand the data by analysing the differences existing between the samples and identifying possible outliers as well as to visualize any possible segregation or clustering among different classes. As described in section 2.3, this method extracts the main sources of variability in the data (Amigo *et al.*, 2013) by transforming the variables into Principal Components (PCs) which are linear combinations of the spectral data describing most of the variation in the original variables (Kamruzzaman *et al.*, 2012). In this study, the ability of PCA to separate the different groups was examined visually by inspecting the scores plots.

5.2.4.3. Soft Independent Modelling of Class Analogy (SIMCA)

Soft Independent Modelling of Class Analogy (SIMCA) analysis was used for potato classification regarding their health condition. SIMCA was applied to the NIR spectra of the 188 tubers (table 5.1) to classify the samples as either healthy or bruised. Therefore, the two different potato classes described in section 5.2.2 were used: intact samples no subjected to any damage (Class 1) and samples measured after impacts (Class 2).

5.2.4.4. Partial Least Squares Discriminant Analysis (PLS-DA)

Moreover, a Partial Least Squares Discriminant Analysis (PLS-DA) was carried out to distinguish healthy and bruised tubers. For this, a 2 column response matrix Y was

introduced in which samples belonging to the first Class (Healthy) were described by the dependent vector [1 0] and likewise, samples belonging to the second Class (Bruised), by the vector [0 1].

5.2.4.5. Blackspot detection

The ability of VIS-NIR and SWIR hyperspectral systems to detect blackspot areas in each potato tuber was also investigated. This was performed with the objective to use a detector capable of mapping out the sound areas and the blackspot affected ones for each potato (each hypercube) at the final stage of the potato manufacturing process. In order to develop this mapping, each pixel of the hypercube was individually classified (was taken as one sample) in the PLS-DA model. Then, with the results of this classification model, a map of the affected areas in each tuber was created.

With that aim, 10 and 5 potatoes belonging to Bruised Class and analysed 24 hours after impacts were selected as the calibration and external validation set, respectively. Same tubers were selected for both hyperspectral ranges. That selection was made in accordance with the results obtained in the PLS-DA based on the mean spectrum where the clearest discrimination results were obtained for the 24 hours group of samples.

For area labelling, the roipoly (region inside a polygon) MATLAB function (MATLAB Version: 8.3.0.532 (R2014a); The MathWorks, Natick, MA) was applied to each tuber to manually select a polygonal ROI within the image corresponding to the bruised area. After selection of the desired ROI, this function creates a mask with the same size as the roipoly. The roipoly consists of a binary image with 1 and 0 inside or outside the polygon. In this study, that binary image was then selected as the bruised mask. Finally, the resulting product from the subtraction between the whole mask and the bruised mask was selected as the healthy mask.

Since that selection led to a large number of pixels, the Kennard and Stone (KS) algorithm was applied to select a representative number of pixels in each mask (Kennard & Stone, 1969). This algorithm has recently been successfully applied for pixel selection in NIR spectroscopy (Casale *et al.*, 2010; Zhu *et al.*, 2010) and hyperspectral imaging (Fernández Pierna *et al.*, 2012; Riccioli *et al.*, 2011). In this work, the KS algorithm was applied to select half of the pixels in both Classes (Healthy and Bruised) in the VIS-NIR and SWIR spectral ranges. The resulted data matrix after applying KS algorithm consisted of 463,995 rows and 220 columns in which the healthy area was represented by 452,622 rows and the bruised area by the remaining 11,373 rows in the VIS-NIR spectral range. This data matrix was comprised by 10 potatoes

and used as the training set. Accordingly, the test set consisted of 5 potatoes individually analysed representing an overall data matrix of 178,560 rows and 220 columns, where the Healthy Class accounted for 175,029 rows and the Bruised Class for the remaining 3,531 rows.

In the same terms, after applying KS algorithm to the SWIR spectral data, the resulting data matrix represented by 10 tubers and used as the training set consisted of 269,938 rows and 150 columns, where the Healthy Class covered 263,094 rows and the remaining 6,844 were covered by the Bruised Class. Consequently, the test set represented by 5 tubers accounted for a total data matrix of 114,410, in which the Healthy Class represented a total of 107,921 rows while the Bruised Class covered 1,893 rows.

5.2.4.6. Model validation and accuracy

For the supervised classification methods (SIMCA and PLS-DA), potato samples were randomly divided into training and test sets consisting of respectively 70% and 30% of the tubers. Classification models were built using the training data set, while the test data set was used to test the model capability of classifying new samples.

The cross-validation (CV) method chosen was Venetian Blinds with 10 data subsets (splits). In this type of CV, each test set is determined by selecting every s^{th} object in the data set, starting at objects numbered 1 through s.

The comparison between the two spectral ranges used and different pre-processing techniques was based on the overall accuracy of the classification model in the training and test sets. This accuracy was determined by the percentage of correctly classified (% CC) samples and the sensitivity and specificity of each Class for both SIMCA and PLS-DA classification techniques.

For any Class A or B, the sensitivity is defined as the proportion of samples belonging to that Class A that are correctly classified and is calculated as equation (10), where True Positives (TP) are the number of samples belonging to Cass A that are correctly classified into their Class, while False Negatives (FN) are the number of samples from Class A incorrectly classified as Class B. The specificity describes the ability of the model to discard samples of all other classes, and can be written as equation (11), where True Negatives (TN) are the number of samples belonging to Class B that are correctly classified as Class B, while False Positive (FP) are the number of samples from Class B incorrectly classified as Class A (Ballabio & Consonni, 2013):

$$Sensitivity = \frac{TP}{TP + FN} \quad (10)$$

$$Specificity = \frac{TN}{TN + FP} \quad (11)$$

Sensitivity and specificity take values between 0 and 1. The closer to 1 the sensitivity and specificity of a given class, the better the classification performance of the model.

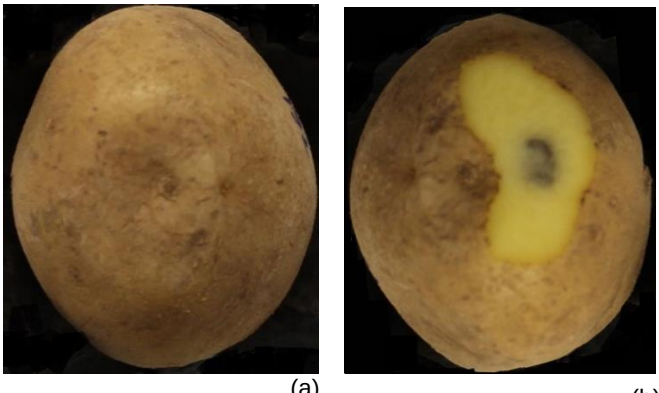
5.3. Results and discussion

In table 5.1 the physical characteristics of the sample set of tubers are summarized, such as the skin and flesh colour and the weight. Also the resistance against internal bruising according to the European cultivated potato database is reported (SASA, 2015).

Table 5.1. Characteristics of the potato samples.

Cultivars	*n*	*Skin colour*	*Flesh colour*	*Resistance to internal bruising*	*Weight (g)*
Hermes	109	White to yellow	Yellow	Very low	103-212
Bintje	44	White to yellow	Light yellow	High	148-338
Magnum	35	White to yellow	White	Not reported	112-315

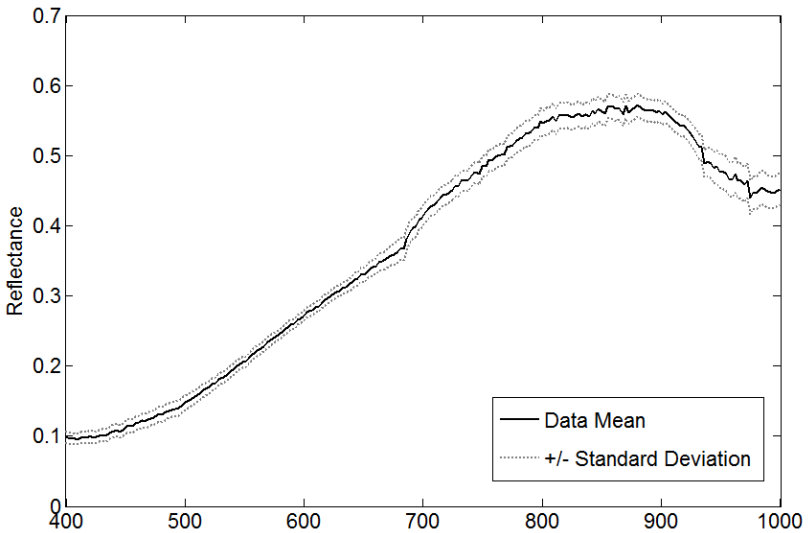
Once all the peeled tubers were photographed, it was observed that 15.9% of the tubers subjected to impacts had not developed any blackspot damage, while 12.7% of the healthy tubers presented some kind of damage in the scanned area including blackspot (70.5%) and internal fissures and crushing (29.5%), according to the classification system for impacts made by Baritelle *et al.* (2000). An RGB image of a potato sample 24 hours after impact, before (a) and after peeling (b) is shown in figure 5.5. In this figure, it can clearly be observed that before removing the skin (figure 5.5a) it is not possible (in some tubers) to visually detect any bruise, while the blackspot can be clearly seen after peeling (figure 5.5b).



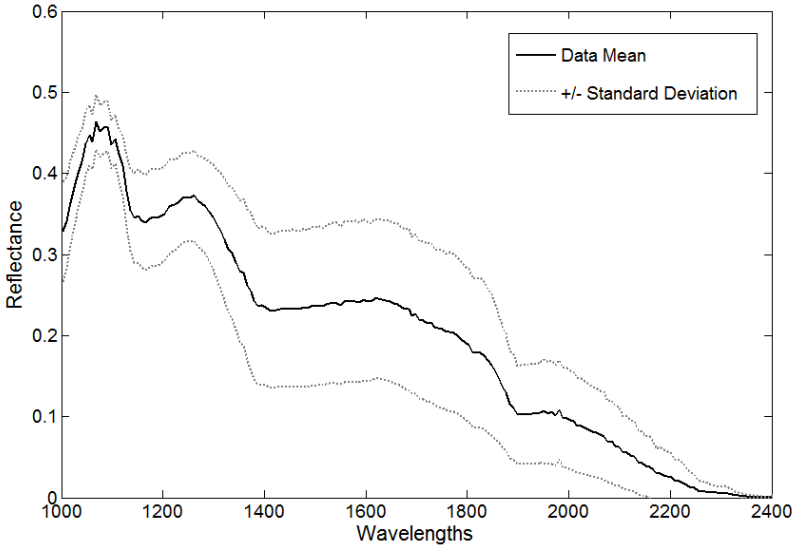
(a) (b)

Figure 5.5. One sample photographed 24 hours after bruising (a) before and (b) after peeling.

In figure 5.6a-b the mean ± standard deviation reflectance spectra obtained from the VIS-NIR and SWIR hyperspectral imaging systems are shown. It can be seen that the variation is higher in the SWIR hyperspectral range. In figure 5.7 the mean spectra of the different groups obtained with both setups are plotted together in order to investigate any possible differences between the measurement times. From figure 5.7a it can be seen that the mean spectra for the different times after bruising overlap in the VIS-NIR spectral range, while the mean spectrum of the healthy group can be distinguished from the rest. Compared to the other groups, the healthy tissue has the highest reflectance from 600 to 900 nm. This is in accordance with Porteous *et al.* (1981) who obtained a notably reduced reflectance in areas of potato with brown lesions compared to normal tissue along the same spectral range. Same behaviour was observed in apples by Xing *et al.* (2007) where the absorption by water was initially high (500-800 nm) in a bruised area as the water was set free from the cells, but after some time, the absorption decreased because that water was lost through evaporation. Moreover, figure 5.7a shows a higher reflectance of the healthy group at the water peak around 970 nm which could also be attributed to water loss from the bruised tissue.



(a)



(b)

Figure 5.6. Mean ± standard deviation reflectance spectra of potato samples for (a) VIS-NIR and (b) SWIR spectral range.

In figure 5.7a no clear temporal hierarchy in the reflectance spectra for the different groups after bruising can be noticed. The 24 hours group shows the lowest reflectance in the region between 600-900 nm.

The mean spectra of the different groups in the SWIR region (figure 5.7b) appeared overlapped too. Here, there is also no clear correlation between the spectra and the time after bruising. It can be observed from the figure that the mean spectra of the healthy and 1 hour groups are overlapped along the wavelength range and separated from the rest, especially between 1400 and 1700 nm where they show lower reflectance than the rest. As the 1400 nm region is characteristic for absorption by water, figure 5.7b suggests that the water content in potatoes from the healthy and 1 hour groups was higher than in bruised (blackspot affected) ones. This could be due to a loss of water as a consequence of the bruise. A positive correlation between blackspot and specific gravity has been reported by several researchers. Authors found that potatoes with a higher specific gravity were more susceptible to blackspot (Massey, 1952; Scudder, 1951). The higher the specific gravity of the samples, the lower the water content (Hegney, 2005). Similarly, Workman and Holm (1984) reported a positive correlation between blackspot susceptibility and dry matter content. However, they only observed this behaviour for recently harvested tubers, while such correlation was not observed for long stored tubers.

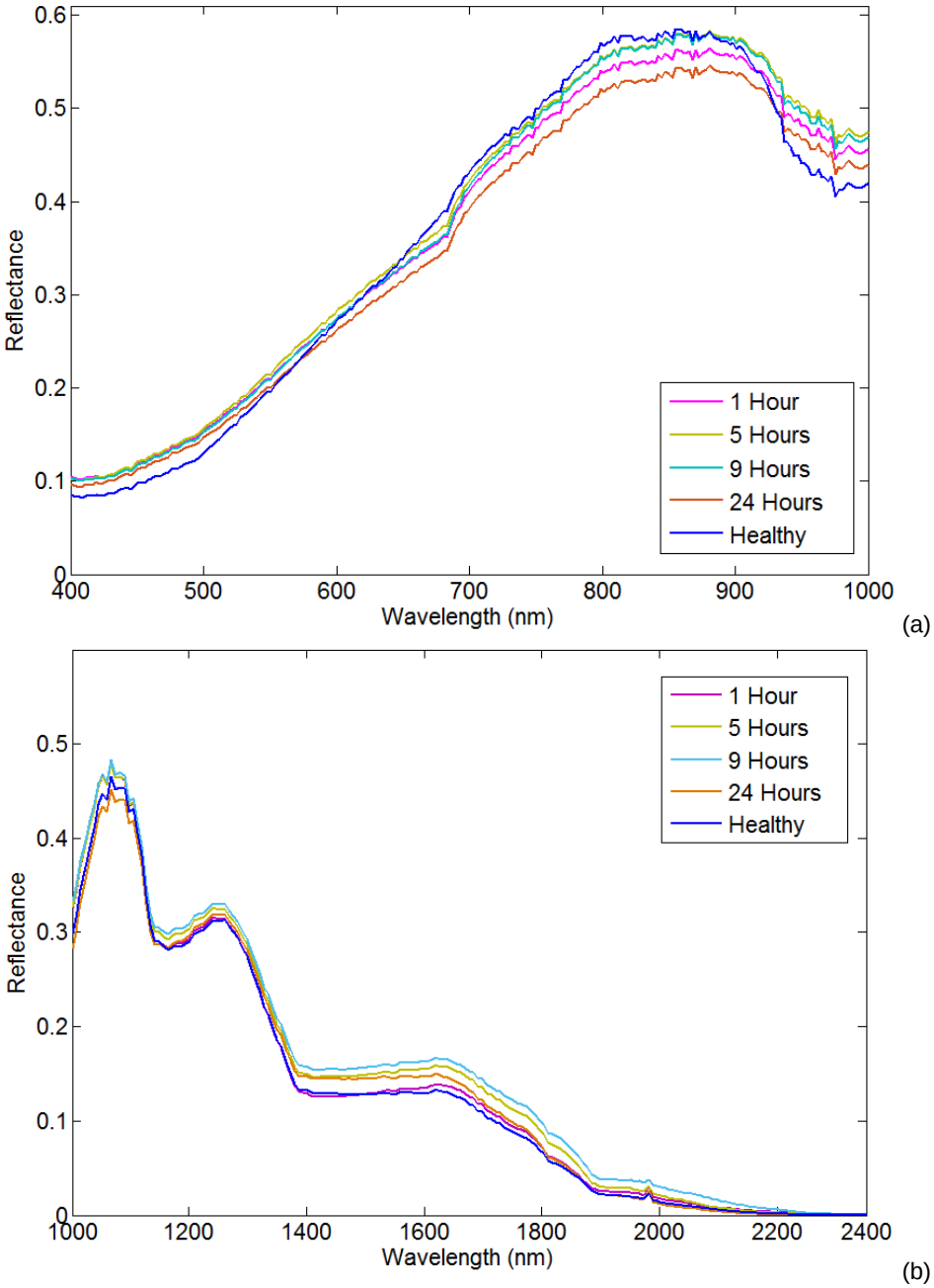


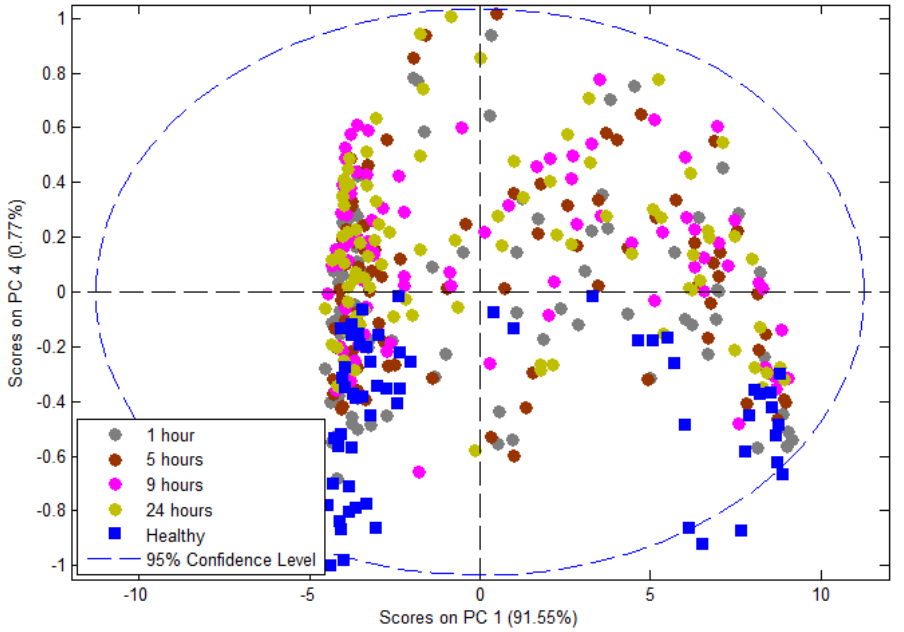
Figure 5.7. Mean spectra for each group of samples for (a) VIS-NIR and (b) SWIR spectral range.

5.3.1. PCA

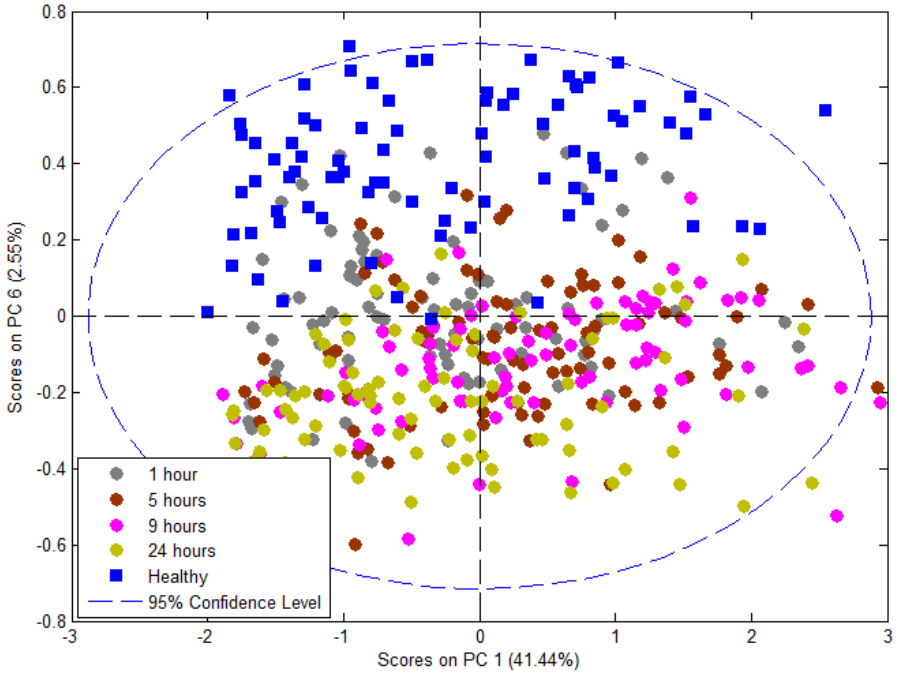
A PCA was carried out in order to explore the spectral differences between the two groups of samples. Moreover, all combinations of the pre-processing techniques formerly described were studied. Six PCs were selected for the VIS-NIR range explaining 98.54% of the variance, while seven PCs were chosen in the SWIR range, representing 96.52% of the variance. In figure 5.8a, PC 1 is plotted versus PC 4 in the VIS-NIR spectral range, showing that all healthy samples had negative values in this PC 4 and could be separated along it. Although PC 1 and PC 2 represented the main

part of the data variance (91.55 and 3.31%, respectively), no Class separation was observed by plotting those PCs together (not shown).

In figure 5.8b, the score plot of PC 1 against PC 6 is shown, from which it is observed that PC 6 plays an important role in separating the Healthy from the 24 hours group. This result is similar than the one in the previous plot, as even though PC 1 and PC 2 represented the main part of the data variance (41.44 and 29.17%, respectively), no group separation was observed by plotting these two (not shown).



(a)



(b)

Figure 5.8. Score plot of (a) PC 1 *versus* PC 4 for bruise detection in VIS-NIR and (b) PC 1 *versus* PC 6 in SWIR images for 2nd der + SNV pre-processing technique.

5.3.2. SIMCA

SIMCA models were assessed in terms of correctly classified (CC) samples and sensitivity and specificity of each class. Only best results corresponding to the combination of smoothing, second derivative, SNV and MC pre-processing technique are displayed.

In table 5.2 the % CC samples for the two classes are summarized for both hyperspectral setups. A total of 92.59% healthy samples and 75% bruised in the test set were correctly classified based on the VIS-NIR spectra. From these 25% of bruised samples wrongly classified as healthy, 56.86% corresponded to 1 hour group, 19.61% to the 5 hours group, 11.76% to the 9 hours group and the rest were samples from the 24 hours group (11.76%).

A better classification rate was achieved in the SWIR spectral range with 100% and 77.23% CC tubers in healthy and bruised classes, respectively. From the bruised samples that were incorrectly classified as healthy, 65.22% corresponded to the 1 hour group, 30.43% were samples from the 5 hours group and the rest (4.35%) corresponded to samples analysed 9 hours after bruising. No misclassifications were observed for the samples measured 24 hours after bruising.

Table 5.2. The % CC tubers by hyperspectral imaging in SIMCA and PLS-DA classification models.

	SIMCA				PLS-DA			
Classes	VIS-NIR		SWIR		VIS-NIR		SWIR	
	Train	Test	Train	Test	Train	Test	Train	Test
Healthy	98.48	**92.59**	100	**100**	96.97	**96.29**	100	**100**
Bruised	80.68	**75**	81.12	**77.23**	95.07	**94.64**	97.63	**97.12**

The detection of internal damage in fruit and vegetables by applying SIMCA was also investigated by other authors. Pholpho *et al.* (2011) studied the capability of visible spectroscopy to detect bruised longan fruits. They were able to correctly classify 86% of them when applying SIMCA. Liu *et al.* (2006) also obtained very good results by the use of a hyperspectral imaging in the 447 to 951 nm range, coupled with SIMCA for the detection of chilling injury in cucumbers with almost 92% CC samples.

Regarding potatoes, Gao *et al.* (2013a) conducted a study for the detection of black heart in raw tubers by the use of transmission hyperspectral imaging. They reported an accurate identification of black heart of 100% in the range between 400-1000 nm. More recently, Zhou *et al.* (2015) also investigated the identification of black heart in potatoes

using VIS-NIR transmittance spectroscopy combined with PLS-DA in the 513-850 nm region. They achieved overall classification rates above 96% in the validation set.

In table 5.3, the sensitivity and specificity values for both spectral ranges with SG + 2nd der + SNV + MC pre-processing techniques are summarized. A total of 9 PCs were selected for both classes (Healthy and Bruised) in the VIS-NIR explaining 99.45 and 99.35% of the variance, respectively. Besides, 5 and 6 PCs were chosen for Healthy and Bruised Classes explaining 96.17 and 96.44% of the variance respectively for the SWIR setup. As shown in table 5.3, the sensitivity value of Healthy Class in the VIS-NIR was higher than that of Bruised Class and close to 100% in both training and test sets. In comparison, higher sensitivity and specificity values were obtained in the SWIR spectral range for both Classes. These results suggest that SIMCA allows to discriminate healthy from bruised potato tubers based on the acquired hyperspectral images.

Table 5.3. The obtained sensitivity and specificity results by SIMCA and PLS-DA models.

Models	Setup	VIS-NIR					SWIR				
	Classes	**PC/ LV**	**Sensitivity**		**Specificity**		**PC/ LV**	**Sensitivity**		**Specificity**	
			Train	**Test**	**Train**	**Test**		**Train**	**Test**	**Train**	**Test**
SIMCA	**Healthy**	9	0.984	**0.925**	0.806	**0.750**	5	1	**1**	0.811	**0.772**
	Bruised	9	0.806	**0.750**	0.984	**0.925**	6	0.811	**0.772**	1	**1**
PLS-DA	**Healthy**	6	0.969	**0.962**	0.950	**0.946**	4	1	**1**	0.976	**0.971**
	Bruised		0.950	**0.946**	0.969	**0.962**		0.976	**0.971**	1	**1**

5.3.3. PLS-DA

PLS-DA models were also evaluated in terms of % CC samples and sensitivity and specificity of each group. In table 5.2 the % CC samples for each Class are summarized for training and test sets for both spectral ranges. In the VIS-NIR spectral range the mean classification success in terms of prediction was above 94%. From the 5.36% of bruised samples which were wrongly classified as healthy, 66.67% of the samples corresponded to the 1 hour group, 16.67% to the 5 hours group and another 16.67% to the 9 hours group. These results suggest that by the use of VIS-NIR hyperspectral imaging and PLS-DA classification it is possible to accurately discriminate healthy tubers from bruised potatoes 24 hours after bruising, while there is some misclassification for shorter times after impact.

A better classification rate of healthy samples was obtained in the SWIR range where 100% of the tubers were correctly classified. On the other hand, 97.12% of the tubers from the bruised group were correctly identified as bruised. An overall classification for the prediction set of 98.56% was achieved. It should be noted that all the misclassified samples corresponded to the 1 hour group, which suggests that hyperspectral imaging in the SWIR range in combination with PLS-DA allows the detection of bruises at early stages of development, in this case, at 5 hours after bruising.

Similar classification results were reported by Gowen *et al.* (2008) using hyperspectral imaging and PCA for bruise damage detection in mushrooms. Additionally, as stated previously, bruise detection in apples has been widely investigated. Several authors have reported correct classification rates above 77% by using different spectral regions (Lu, 2003; Xing & De Baerdemaeker, 2005; Xing *et al.*, 2007). Also, ElMasry *et al.* (2008) were able to detect bruises in apple as early as 1 hour after bruising. Non-destructive detection of bruises in potatoes has been less extensively investigated than in apples. However, as described in section 2.5.1 Dacal-Nieto *et al.* (2011b) studied the application of hyperspectral imaging and chemometrics for determining the presence of hollow heart, an internal defect in potato tubers, achieving a correct classification rate of 89% for healthy and affected tubers. Moreover, Jin *et al.* (2009) were able to classify more than 91% of tubers showing external defects by using computer vision technology.

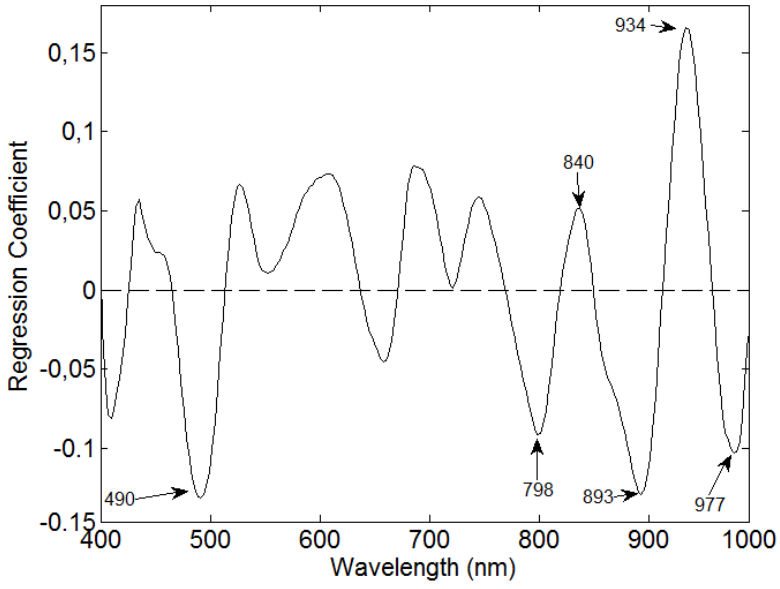
In table 5.3 the sensitivity and specificity values are summarized for the PLS-DA models. A total of 6 and 4 LVs were used to build the PLS-DA model explaining 97.54 and 83.88% of the spectral variance in VIS-NIR and SWIR spectral ranges, respectively. Sensitivity and specificity values obtained for Classes 1 and 2 were above 94% in both training and test sets for the VIS-NIR range. It can be seen from table 5.3 that sensitivity values for Class 1 (Healthy) were 100% for both training and test sets in the SWIR spectral range. Furthermore, sensitivity values above 97% were obtained for Class 2 (Bruised) in both training and test sets. The fact that the PLS-DA model for the VIS-NIR data performs worse in predicting the Y-variance even though it captures more spectral variance suggests some kind of overfitting.

These results suggest that the combination of hyperspectral imaging techniques and PLS-DA allows to accurately discriminate healthy from bruised potatoes. Moreover, blackspot affected tubers could be identified within 5 hours after bruising by means of SWIR hyperspectral imaging.

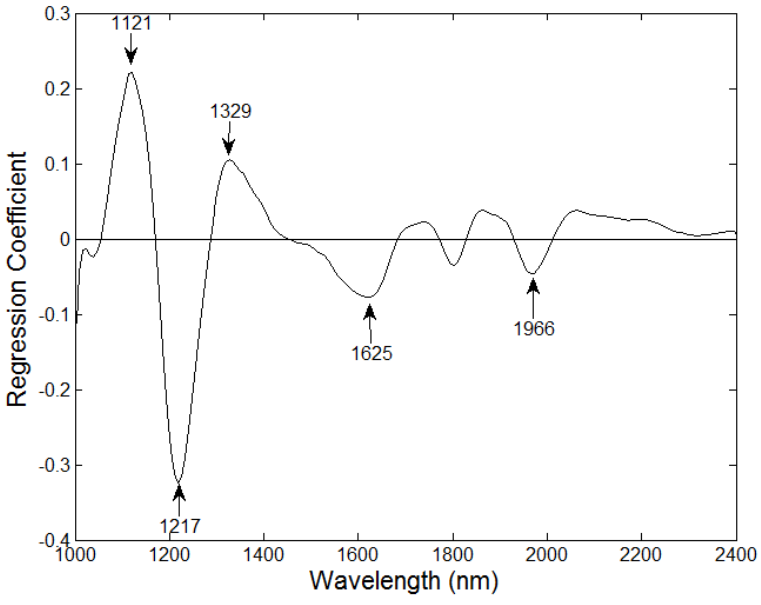
In order to compare both classification methods, it should be mentioned that PLS-DA achieved more accurate results for classification of both classes in the VIS-NIR and

SWIR spectral ranges. In figure 5.9 the regression coefficients are shown for the PLS-DA model built for the VIS-NIR (figure 5.9a) and SWIR (figure 5.9b) range after using SG + 2nd der + SNV + MC pre-processing. As shown in figure 5.9a, the important wavelengths are located along the entire 400-1000 nm spectral region. Based on figure 5.9a, the most informative wavelengths for the PLS-DA model are 490, 798, 840, 893, 934 and 977 nm. The 934 nm can be associated to the third overtone of CH stretching modes (Osborne *et al.*, 1993). The 977 nm can be related to the second overtone of OH stretching, normally associated with water content of the samples (Porteous *et al.,* 1981). As shown in figure 5.9b, the important wavelengths in the SWIR spectral region are mainly located at the beginning of the spectral region. According to figure 5.9b, the most informative wavelengths for the PLS-DA model are 1121, 1217, 1329, 1625 and 1966 nm. The 1121, 1217 and 1329 nm bands could be assigned to the influence of CH stretching modes (Osborne *et al.*, 1993). This could be a result of the formation of intermediates during the complete conversion from tyrosine to melanin that occurs in the course of blackspot formation as a result of the harvesting and managing of tubers. That blackspot formation occurs due to the oxidation of the polyphenols present in the tubers. The initial reaction by polyphenol oxidase catalyses the oxidation of *o*-diphenols to produce *o*-quinones, that are highly reactive and suffer a succession of non-enzymatic reactions to produce melanin pigments responsible for potato browning (Busch, 1999).

The 1966 nm is associated with water absorptions bands due to second and first overtones of OH stretching and OH combination bands. This may be associated with water loss from the tissue in and around the bruised zone (Porteous *et al.*, 1981).



(a)



(b)

Figure 5.9. Regression coefficient plot for (a) VIS-NIR and (b) SWIR setups using SG + 2nd der + SNV + MC pre-processing.

5.3.3.1 Mapping of the affected area

In table 5.4 the sensitivity and specificity values and the % CC samples for the pixel based PLS-DA model with SG + MC pre-processing technique are shown. A total of 5 LVs were used to build the model in both VIS-NIR and SWIR spectral ranges explaining respectively 99.81 and 98.89% of the variance. Although 5 tubers (5 hypercubes) were individually used as a test set, only averaged results are presented here. Slightly better results were obtained in the SWIR spectral range in terms of % CC samples and sensitivity and specificity. In any case, % CC pixels above 90% were obtained in both spectral ranges with high sensitivity and specificity values. It should be mentioned that the labelling of the images was quite hard, so part of the misclassifications may be due to incorrect labelling of the pixels at the edge of the bruised area.

Table 5.4. The sensitivity, specificity and % CC results by the pixel based PLS-DA model.

Setup		Classes	*n*	LV	Sensitivity	Specificity	% CC
VIS-NIR	**Train**	Healthy/Bruised	463,995	5	0.967	0.912	93.92
	Test	Healthy/Bruised	178,560		**0.912**	**0.898**	**90.82**
SWIR	**Train**	Healthy/Bruised	269,938	5	0.914	0.930	92.16
	Test	Healthy/Bruised	109,814		**0.947**	**0.927**	**93.71**

In figure 5.10 the sequence for blackspot detection from the acquired hypercubes is schematically illustrated. In figure 5.10a the first image on the left corresponds to the

original hypercube of a sample measured with the VIS-NIR spectral range 24 hours after impact in which the bruised mask is represented only by its edges. It can be seen that this blackspot area is not easily identified and could be easily confused with any other mark present on the tuber. However, the impact procedure was carried out in a controlled way that allowed us to know exactly where the bruises should be located. The second figure on the left corresponds to the segmented hyperspectral cube of the entire sample. The third figure on the left corresponds to the mapping of the bruised area, where some misclassified pixels can be seen. However, the blackspot area in the centre was correctly identified. The last figure shows the segmented hyperspectral cube with the mapping of the bruised area and the edges of the bruised mask. In figure 5.10b the operation of the blackspot detection system in the SWIR spectral range is illustrated as well. The first image on the left corresponds to the original hypercube and the edges of the bruised mask of a sample measured 24 hours after impact, there is a bruised area located in the centre. The second image on the left shows the segmented hyperspectral cube of the tuber. In the third figure, the mapping of the bruised area is shown. Similar to the VIS-NIR spectral range, there were some misclassified pixels, while the blackspot was well identified. Finally, the last figure shows the segmented hyperspectral cube with the mapping of the blackspot affected area and the edges of the bruised mask in order to compare the results.

As commented previously, there are no records of the application of hyperspectral imaging systems for the detection of blackspot in potatoes as far as we are concerned.

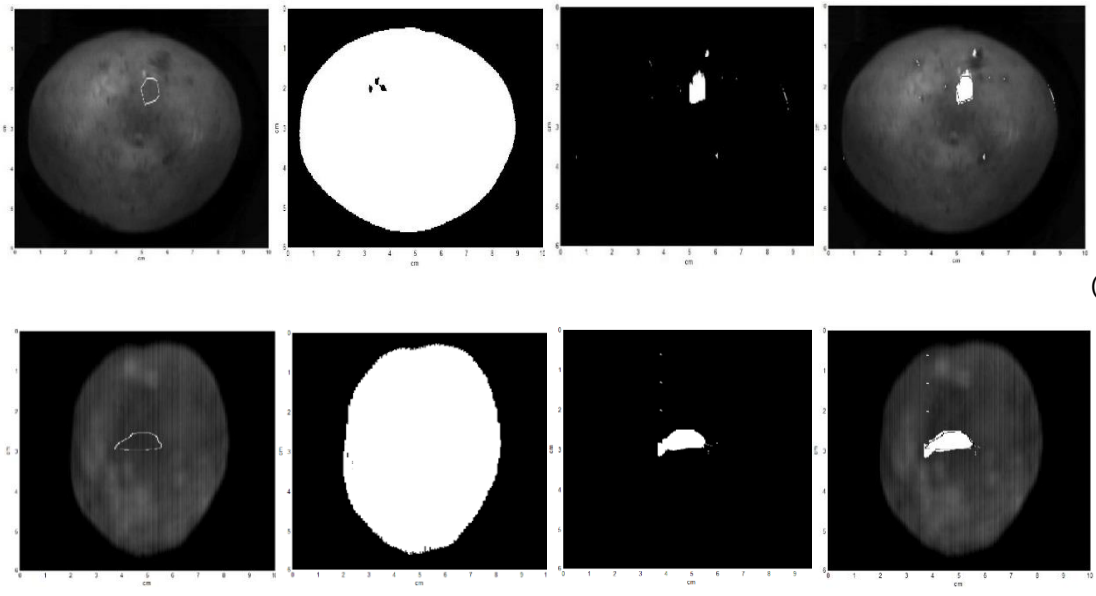


Figure 5.10. An example of the operation of the blackspot detection system in (a) VIS-NIR and (b) SWIR spectral range. From left to right: original hypercube with the edges of the bruised mask, segmented hyperspectral cube, mapping of the bruised area, segmented hyperspectral cube with the mapping of the bruised area and the edges of the bruised mask.

5.4. Conclusion

The results obtained suggest that it is possible to identify raw potatoes affected by blackspot by combining hyperspectral imaging and chemometric techniques. The correct classification rate was between 2 and 7% higher for the SWIR range (1000-2500 nm) than for the VIS-NIR range (400-1000 nm) for Classes 1 and 2 for both SIMCA and PLS-DA discrimination methods. Furthermore, it was observed that more accurate discrimination of healthy and bruised tubers was achieved by applying PLS-DA to SWIR hyperspectral imaging data where all samples belonging to 5, 9 and 24 hour groups and the healthy set were correctly classified into their corresponding group. Therefore, it can be concluded that SWIR coupled with PLS-DA is able to accurately identify early bruises in potatoes within 5 hours after bruising.

Moreover, in this study, a map of the areas affected by blackspot in a set of potato samples with a pixel classification was performed with an accuracy above 93% when using the SWIR spectral range. In this case, SWIR hyperspectral imaging was able to correctly classify a 0.61% more samples than VIS-NIR.

According to these results, the use of SWIR hyperspectral imaging coupled with PLS-DA to detect blackspot in raw potatoes has potential for on-line quality control of tubers in industry. The application of NIR hyperspectral imaging could enable the development of a fast, reliable and non-destructive method for internal damage detection of potatoes avoiding the commercialization of affected tubers. This would be beneficial to both potato industry and consumers. However, it should be mentioned that prior to the implementation of this system in the potato industry, those results must be validated on a larger sample set covering a wide range of potato varieties grown in different areas and under different conditions.

Acknowledgments

The funding of this work has been covered by the Universidad Pública de Navarra through the concession of both a predoctoral research grant and a mobility grant, by the National Institute for Agricultural and Food Research and Technology (INIA) project: “Mejora genética de la patata: caracterización reológica y por tecnología NIRS del material.” RTA2013-00006-C03-03, and by the Agency for Innovation by Science and Technology in Flanders (IWT) through the Chameleon (SB-100021) project. The authors would also like to thank the Basque Institute for Agricultural Research and Development (Neiker Tecnalia) for supplying some of the samples used in this study.

Janos Keresztes and Mohammad Goodarzi have been funded respectively as PhD student and postdoctoral researcher on the IWT Chameleon project (SB-100021).

Chapter 6.

General discussion

General discussion

The aim of this thesis was to apply non-destructive technologies in the visible and near-infrared region of the electromagnetic spectrum in the quality control of potatoes. This goal arises due to the current scenario in which agricultural and food products must complied with quality standards required by consumers and the retail sector.

Potato is one of the most consumed agricultural products worldwide and thus, potato industry needs techniques capable of efficiently measuring its quality properties in a safe and economically feasible way. For these reason, in this doctoral thesis, quality analyses of potatoes with non-destructive spectroscopic techniques were developed and evaluated. In the beginning, some areas of research were identified in which a more deep study was needed, and consequently, the present work was focused on solving some of the current issues.

In this thesis, quality was evaluated from three points of view directly related with each other. First, after a revision process, it was detected that optical characterization of potatoes was hardly studied in spite of its importance. Second, when a review of the NIRS applications for the analyses of potatoes was performed, it was observed that although dry matter and starch content were accurately predicted by many authors, other compounds not so much. This was the case for crude protein, for which authors could not obtained reliable results using NIRS, and for sugars, for which authors found high RMSEP values. Moreover, it was detected that some other internal compounds were barely studied. Finally, once hyperspectral imaging applications for potato analyses were inspected, it was seen that even though a lot of work was carried out to detect defects and diseases in potatoes, there was no record of an identification of blackspot by imaging systems being one of the main responsible for economic losses within potato industry. Therefore, in the present thesis, optical characterization of potato tissues was accomplished in the first place, along with a qualitative and quantitative analysis of tubers ending up with an application of hyperspectral imaging to detect internal damaged tubers.

6.1. Optical properties characterization

Potato tissues are highly scattering in the NIR region due to the microstructure and texture of their tissues. In general, when analysing potato products by NIRS, the Beer-Lambert Law is followed, which relates the absorbance of a medium to the concentration of the absorbing compounds present in that medium (Swinehart, 1962). However, according to this law, an increase in absorbance by a product would

correspond to an increase in the concentration of the parameter measured in that product, while this is not always the case, as reflectance measurements provide information from both absorption and scattering phenomena. Thus, a change in scattering properties of a product could be wrongly seen as a variation in the product composition according to Beer-Lambert Law (Aernouts *et al.*, 2011).

In the present thesis, an optical characterization of potato flesh was for the first time accomplished with a wavelength-dependent estimate of the anisotropy factor. Moreover, the bulk optical properties of potato flesh were measured in the 500-1900 nm wavelength range with a setup consisting of double integrating spheres and unscattered transmittance measurement system which are considered the reference criteria for optical characterization of biological tissues. The results obtained show that the bulk absorption coefficient spectra were dominated by water, more specifically in the NIR region of the spectrum, a result that was not surprising since raw tubers have an 80% of water content. Also an absorption peak at around 1800 nm probably related to starch content was observed. On the other hand, the bulk scattering coefficient (μ_s) spectrum changes only slightly from 166 to 160 ± 7.35 cm^{-1} with increasing wavelength, a tendency considered as typical for the scattering characteristics of biological tissues (Bashkatov *et al.*, 2005b). Finally, the estimated values for the anisotropy factor were high ranging from 0.959 to 0.949 along the spectral range studied, pointing out the highly forward scattering nature of potato tissues. Comparing the results obtained in this study by those of other authors, it should be mentioned that the absorption and reduced scattering coefficient spectra for potato tissue retrieved here are considered more reliable for several reasons. First, optical properties of potatoes have only been measured by Karagiannes and Grossweiner (1988) and Karagiannes *et al.* (1989) by the use of a single integrating sphere which in general provides less accurate results compared to DIS. Second, these authors estimated optical properties in the 400 to 1100 nm range and in the 340 to 1360 nm respectively while in this study, optical properties were obtained in the visible and extended NIR range from 500 to 1900 nm, which makes a difference as the relevant absorption peaks for water and starch are located in the NIR region of the spectrum. Moreover, for both studies, authors assumed a fixed value of g=0 which implies isotropic scattering whereas the results in the present thesis suggest an almost exclusive forward scattering due to the high values of g obtained. Finally, Karagiannes *et al.* (1989) reported absorption values at 970 and 1200 nm that exceeded the absorption coefficient values of pure water. Since water is the most important NIR-absorbing component in potatoes, these values give the impression that there were overestimated. In addition, same authors reported a

baseline in the absorption coefficient spectrum at 800 nm where no absorption is expected.

However, it should be noted that absorption values in the 1400-1500 nm range of the spectrum were probably slightly overestimated as a consequence of imperfect separation of the scattering and absorption coefficient values in this region. In any case, the results of this study are considered more reliable than those reported by Karagiannes *et al.* (1989). This fact is confirmed with the comparable values obtained for other similar products (onion flesh and skin, and apple) reported by Wang and Li (2013) and Saeys *et al.* (2008) in literature.

6.2. Application of NIRS for the analysis of potatoes

An identification of potato cultivars rich in phenolic compounds, anthocyanins and antioxidant capacity by non-destructive NIRS technique was carried out. Such identification is considered very important in order to develop breeding lines with high concentrations of bioactive compounds to satisfy the current demand of not only high quality products but also healthy ones. Therefore, a successful classification of 16 cultivars according to their concentration of total soluble phenolics, total monomeric anthocyanins and their hydrophilic antioxidant capacity was achieved by the use of NIRS on intact tubers. The samples used belonged to the Potato Germplasm Collection of the Basque Institute for Agricultural Research and Development (NEIKER-Tecnalia), and some of them were breeding lines developed within the institute.

NIRS technique was combined with a PLS-DA chemometric method to categorize the selected cultivars as having low, medium, or high content of phytochemicals. An overall classification rate of 86.1% for the validation group was obtained, being the cultivars belonging to the Low Content group the best classified ones. Some misclassification was observed predominantly in cultivars with the highest or lowest concentrations of phytochemicals into their group. These values are comparable to those obtained by other authors when classifying varieties according to their recoverable protein content or sugar level. Fernández-Ahumada *et al.* (2006) obtained similar percentage of correctly classified potato samples (87.5%) as either having high protein content (≥14 mg g^{-1}) or low (<14 mg g^{-1}). However, it should be mentioned that the former authors used mashed potatoes for the qualitative analysis while in the present study NIR spectral acquisition was performed on intact tubers. More recently, Rady and Guyer (2015a) classified potato samples based on glucose and fructose level achieving

classification rates between 71 and 100% for glucose level on whole tubers and between 75 and 83% for fructose levels on sliced samples

Moreover, even though it was not possible to classify samples into 3 levels based on their carotenoids content, an identification of tubers presenting high carotenoid concentrations was achieved with 80% of accuracy. Considering the significance of a successful characterization of tubers and breeding lines with high contents of these healthy, beneficial organic pigments, the findings of this study are considered of great importance for both potato breeders and researchers.

On the other hand, in this research a quantitative analysis was performed to estimate different chemical compounds of potatoes. An accurate estimation of dry matter and starch by NIRS was not achieved and the reason for this was probably the reference method used to measure those compounds. DM and starch were indirectly calculated by underwater weight measurements. Indirect methods are less accurate than direct methods (Goñi *et al.*, 1996); therefore, the error of the reference method increases and hinders the correlation with NIR spectral data, since much of the error found in any NIR calibration is due to the error of the reference method (Hruschka, 2001). It was commented in section 2.4.1 that best predictions of DM and starch by NIRS found in literature were obtained when using a mashed and homogenized sample in the 1100-2500 nm spectral range; however, it should be added, that the reference method has a great influence on final results, since in the majority of the studies reviewed, DM content was calculated directly by oven drying (Brunt & Drost, 2003, 2010; Dull *et al.*, 1989; Fernández-Ahumada *et al.*, 2006; Haase, 2003, 2011; Hartmann & Büning-Pfaue, 1998; Pedreschi *et al.*, 2010; Scanlon *et al.*, 1999; Subedi & Walsh, 2009; Walsh *et al.*, 2004). In only two studies among all reviewed in section 2.4.1 DM content of potatoes was calculated by both direct (oven dried) and indirect methods (UWW) obtaining similar results (Helgerud *et al.*, 2012; Van Dijk *et al.*, 2002); however, Helgerud *et al.* (2012) observed that the highest deviation occurred when the indirect method was used.

On the contrary, in the present thesis, very good results were obtained for the estimation of crude protein, nitrogen and total soluble phenolics. The good results obtained for the prediction of CP with an RPD value of 2.84 provide an important novelty since no other author had achieved an accurate estimation of CP by NIRS. Fernández-Ahumada *et al.* (2006) obtained an RPD of 1.45 indicating a very poor prediction according to levels established by Saeys *et al.* (2005), while Haase (2006) achieved a better result (RPD: 1.61) which indicate that the model could be used for screening purposes; but in both studies, RPD values were further away from the results

of the present research. Authors considered that if the variance in the reference data is low, the values of R^2 and RPD obtained are also low (Fernández-Ahumada *et al.*, 2006). The better results achieved compared to other authors could be due to the great variability of the data that allows developing of more robust calibration models.

Prediction of phenolic compounds in potatoes by NIRS has been less studied than in other compounds. Shiroma-Kian *et al.* (2008) studied the determination of total phenolics, anthocyanins, and antioxidant capacity in polyphenolic extracts from potato by Fourier transform infrared spectroscopy. Prediction models obtained provided reliable predictions as compared to reference values with deviations lower than 2% from the target value. Accurate models were obtained for the estimation of total soluble phenolics by NIRS. These results provide a significant tool for the selection of potato breeding lines with enhanced phenolic content. Moreover, models obtained for the prediction of hydrophilic antioxidant capacity can be used to improve the screening process in order to select and further develop potato varieties with a higher antioxidant capacity.

6.3. Application of hyperspectral imaging for the analysis of potatoes

In this thesis, an accurate identification of blackspot in potatoes by the use of hyperspectral imaging was for the first time achieved with classification rates above 98%.

Internal damage in potatoes is one of the main responsible for a decrease of quality and involves substantial economic losses yearly. Thus, a non-destructive tool capable of monitoring tubers quality safely and economically is needed.

Although NIRS technology is one of the most appropriate non-destructive techniques to measure quality of agricultural and food products, it becomes challenging to use it for internal damage identification due to its limitations being a point-based scanning instrument. On the other hand, imaging-based systems offer an advantage over NIRS providing spatial information along with spectral one. For this reason, an identification of blackspot by hyperspectral imaging systems was accomplished with very good results. An identification of raw potatoes affected by blackspot by combining hyperspectral imaging in two spectral ranges (400-1000 nm and 1000-2500 nm) and chemometric techniques was accomplished. Nevertheless, better classification results were obtained when using SWIR hyperspectral imaging setup that works between 1000 and 2500 nm. Among chemometric techniques, PLS-DA showed better potential

than SIMCA for damage identification. Main findings were that early bruise detection (< 5 hours) was achieved by combining SWIR hyperspectral imaging and PLS-DA statistical analysis with an accuracy above 97%. Other authors were able to detect bruises in apple as early as 1 hour after impacts (ElMasry *et al.*, 2008); however, such identification in potatoes had not been accomplished before. The results obtained indicate that this technique can be used to effectively detect bruises on potato surfaces in early stage of bruising. This fact provides an important innovation since those tubers identified as having internal damage, could be derived to other industrial uses rather than fresh use.

Moreover, a mapping of the surface areas affected by blackspot in each tuber of a small representative sample was achieved with an accuracy of more than 93% by the use of SWIR hyperspectral imaging setup. These results are similar to those obtained by Dacal-Nieto *et al.* (2011a) for common scab detection in potatoes. Authors were able to map surfaces areas affected by common scab in raw tubers with a 96% of accuracy. However, such mapping was performed with the use of Support Vector Machines, a complex classifier (Steinberger & Tesar, 2007).

The results obtained in this study are considered relevant since hyperspectral imaging systems used can be fully automated which make them suitable for on-line quality control of tubers at industry level. Since the efficiency of the method used here was demonstrated on three potato varieties with different characteristics and grown under different conditions, it is considered that this procedure has the potential for being extended to other cultivars.

6.4. Future perspectives

The continuously growing demand for quality control of food products in recent years, together with the consumer concerns about the methods of handling and processing of these products, had led to a very severe control of the nutritional contents of many foodstuffs. Moreover, since potato represents a pillar in human nutrition, the mechanization and optimization of tools for quality control at the delivery point are essential.

In this thesis, some tools for quality inspection of potato tubers are provided. However, further investigation is recommended for optimization and future establishment of the methods proposed.

In chapter 3, the bulk optical properties of potato flesh are estimated and a study of light propagation through them is accomplished. The results obtained are considered of high relevance since the bulk optical properties of potatoes have for the first time been

calculate without assumption of the anisotropy factor. However, these results depend on the structure of potato tissues that are determined by several factors like the chemical composition, grown conditions, length of storage, etc. Therefore, the results obtained should be validated in a larger set of samples including different cultivars with varying cellular structures.

In the qualitative analysis performed in chapter 4, very good classification of cultivars based on their phenolic concentration, anthocyanins content and antioxidant capacity was achieved. Nevertheless, this study could be further extended by including yellow- and white-fleshed cultivars in order to increase the variability of the data and the applicability of the classification model to breeding programs.

In the quantitative analysis carried out in the same chapter for the estimation of main chemical compounds of potatoes, the need for a standardization of procedures for the determination of the internal composition was highlighted. In this analysis, main chemical compounds of potatoes, dry matter, starch and sugars among others, were estimated, and so, their correlation with NIR spectral data did not provide reliable results. Thus, it would be interesting for future analyses to establish a protocol for the chemical determination and extraction of chemical compounds to achieve accurate estimations by NIRS.

In the study for the internal damage identification carried out in chapter 5, only a part of the whole potato surface was scanned by the hyperspectral imaging system to determine the presence of blackspot and then, a classification methodology was built to label potatoes as healthy or bruised with the information from that part of the surface. It would be interesting to design an element capable of examining the whole potato surface in order to ensure the complete absence of damage. In the same way, only the surface facing the hyperspectral setup was mapped with healthy and bruised areas and even though the stem end is the most sensitive part and most of the time blackspot appeared there, the whole potato surface should be map to provide a more accurate inspection of quality.

Finally, it should be mentioned that the challenge for coming years within this field is directed to the in-line implementation of those non-destructive spectroscopic techniques for real time quality control along the food chain. The progress of computer hardware and efficiency of software would lead to a reduction of the actual time needed to develop some of the processes presented in this thesis. This fact would promote the implementation of these technologies in potato manufacturing lines for the monitoring of quality of raw products during handling and packaging.

Chapter 7.

Conclusions

7.1. Conclusions

Considering the planned objectives and the results obtained in this thesis, the following conclusions can be drawn:

1) The bulk absorption coefficient spectra were mainly dominated by water, especially in the NIR region, while the starch absorption bands were less pronounced. The bulk scattering coefficient spectra behaved as other biological tissues decreasing over the wavelength range. The anisotropy factor values obtained were high (>0.94) indicating the highly forward scattering nature of potato tissues.
2) The optical characterization of potato tissues obtained in this study was considered more reliable than that reported by other authors.
3) The information obtained from this study may be used to design more efficient sensors for non-destructive quality evaluation of potato tubers.
4) An accurate classification (>86% CC) of cultivars based on their phenolics, anthocyanins content and hydrophilic antioxidant capacity was achieved by using NIRS combined with PLS-DA. In addition, NIRS technology was able to identify cultivars with high content of carotenoids.
5) The results obtained in this study may be used for the selection of potato breeding lines with enhanced polyphenol content.
6) A successful non-destructive prediction of crude protein in potatoes was for the first time achieved using NIRS and a PLS regression method (RPD: 2.83). Accurate estimations of nitrogen (RPD: 2.70) and phenolics (RPD: 2.54) concentration in potatoes were also accomplished by NIRS. Moreover, adequate models for screening were obtained for the hydrophilic antioxidant capacity (RPD: 1.84).
7) An identification of raw potatoes affected by blackspot at early stages (5 hours after bruising), was achieved with an accuracy above 98% by combining hyperspectral imaging and chemometric techniques.
8) Blackspot areas in a set of potato samples were mapped with an accuracy over 93% in the 1000 to 2500 nm spectral range by the use of hyperspectral imaging and a PLS-DA pixel-based model.
9) In all the statistical analyses carried out, SWIR hyperspectral imaging performed better than VIS-NIR in terms of specificity and sensitivity values and percentage of correctly classified samples, with between 0.61% and 7% more samples properly classified.

7.2. Conclusiones

Teniendo en cuenta los objetivos planificados y los resultados obtenidos en esta tesis, se pueden extraer las siguientes conclusiones:

1) Se observó que el coeficiente de absorción estaba altamente influenciado por la presencia de agua en las muestras, especialmente en la región NIR, mientras que las bandas de absorción del almidón eran menos pronunciadas. Los valores del coeficiente de dispersión disminuyeron a lo largo del espectro, tal y como ocurre en otros tejidos biológicos. Se obtuvieron valores altos para el factor de anisotropía (>0.94), indicando la naturaleza altamente dispersiva de los tejidos de patata.
2) La caracterización óptica de los tejidos de patata obtenida en esta tesis se consideró más fiable que la estimada por otros autores.
3) La información obtenida en esta tesis podría utilizarse para diseñar sensores más eficientes en la evaluación de la calidad en tubérculos de patata de manera no destructiva.
4) Se consiguió una clasificación precisa de cultivares (>86% de muestras bien clasificadas) basada en su contenido de fenoles, antocianinas y capacidad antioxidante mediante el empleo de la espectroscopia en el infrarrojo cercano (NIRS) combinada con el análisis discriminante por mínimos cuadrados parciales (PLS-DA). Además, se consiguió identificar mediante las mismas técnicas, cultivares con alto contenido en carotenos.
5) Los resultados obtenidos en esta tesis podrían ser usados en la selección de nuevas variedades con alto contenido en polifenoles dentro de programas de mejora genética de patata.
6) Se consiguió por primera vez, una estimación precisa del contenido en proteína bruta en patatas mediante el uso de tecnología NIRS y modelos de regresión por mínimos cuadrados parciales (PLS) (RPD: 2.83). Asimismo, se consiguieron predicciones fiables del contenido de nitrógeno (RPD: 2.70) y fenoles solubles (RPD: 2.54) con el empleo de estas técnicas. Además, se obtuvieron resultados adecuados para su utilización en procesos de evaluación (screening processes) de la capacidad antioxidante de las muestras (RPD: 1.84).
7) Se identificaron patatas con daño interno (blackspot) en las primeras fases de desarrollo del mismo (5 horas después del impacto) combinando análisis de imágenes hiperespectrales y técnicas quimiométricas, con una precisión superior al 98%.

8) Se mapearon zonas con daño interno en patatas con una precisión del 93% en el rango espectral 1000-2500 nm mediante el uso de análisis de imágenes hiperespectrales y un modelo PLS-DA basado en pixeles.
9) En todos los análisis estadísticos desarrollados, SIMCA y PLS-DA, el análisis hiperespectral SWIR obtuvo mejores resultados que el análisis VIS-NIR en términos de especificidad y sensibilidad y porcentaje de muestras bien clasificadas, con entre un 0.61 y 7% más de muestras clasificadas correctamente.

References

Aernouts, B., Nguyen Do Trong, N., Watté, R., Bruggeman, W., Tsuta, M., Verboven, P., Nicolaï, B., & Saeys, W. (2011). Food quality control by combining light propagation models with multiple vis/NIR reflectance measurements. *NIR News, 22*(2), 14-16.

Aernouts, B., Van Beers, R., Watté, R., Huybrechts, T., Jordens, J., Vermeulen, D., Van Gerven, T., Lammertyn, J., & Saeys, W. (2015). Effect of ultrasonic homogenization on the Vis/NIR bulk optical properties of milk. *Colloids and Surfaces B: Biointerfaces, 126*, 510-519.

Aernouts, B., Zamora-Rojas, E., Van Beers, R., Watté, R., Wang, L., Tsuta, M., Lammertyn, J., & Saeys, W. (2013). Supercontinuum laser based optical characterization of Intralipid® phantoms in the 500-2250 nm range. *Optics Express, 21*(26), 32450-32467.

Agelet, L. E., & Hurburgh, C. R. (2010). A tutorial on near infrared spectroscopy and its calibration. *Critical Reviews in Analytical Chemistry, 40*(4), 246-260.

Aguilera, J. (2000). Structure-property relationship in foods. In J. Lozano, C. Añón, E. Parada-Aries & G. Barbosa-Cánovas (Eds.), *Trends in Food Engineering* (pp. 1-14). USA: Lancaster.

Al-Mallahi, A., Kataoka, T., & Okamoto, H. (2008). Discrimination between potato tubers and clods by detecting the significant wavebands. *Biosystems Engineering, 100*(3), 329-337.

Al-Mallahi, A., Kataoka, T., Okamoto, H., & Shibata, Y. (2010). Detection of potato tubers using an ultraviolet imaging-based machine vision system. *Biosystems Engineering, 105*(2), 257-265.

Al Saikhan, M., Howard, L., & Miller, J. (1995). Antioxidant activity and total phenolics in different genotypes of potato (Solanum tuberosum, L.). *Journal of Food Science, 60*(2), 341-343.

Alva, A., Fan, M., Qing, C., Rosen, C., & Ren, H. (2011). Improving nutrient-use efficiency in Chinese potato production: experiences from the United States. *Journal of Crop Improvement, 25*(1), 46-85.

Amigo, J. M. (2010). Practical issues of hyperspectral imaging analysis of solid dosage forms. *Analytical and Bioanalytical Chemistry, 398*(1), 93-109.

Amigo, J. M., Martí, I., & Gowen, A. (2013). Hyperspectral Imaging and Chemometrics: A Perfect Combination for the Analysis of Food Structure, Composition and Quality. In F. Marini (Ed.), *Chemometrics in Food Chemistry* (Vol. 28, pp. 496). Amsterdam: Elsevier.

Andre, C. M., Ghislain, M., Bertin, P., Oufir, M., del Rosario Herrera, M., Hoffmann, L., Hausman, J.-F., Larondelle, Y., & Evers, D. (2007). Andean potato cultivars (Solanum tuberosum L.) as a source of antioxidant and mineral micronutrients. *Journal of Agricultural and Food Chemistry, 55*(2), 366-378.

Arimoto, H., Egawa, M., & Yamada, Y. (2005). Depth profile of diffuse reflectance near-infrared spectroscopy for measurement of water content in skin. *Skin Research and Technology, 11*(1), 27-35.

Baheri, M. (1997). Development of a method for prediction of potato mechanical damage in the chain of mechanized potato production. *Dissertationes de Agricultura (Belgium)*.

Ballabio, D., & Consonni, V. (2013). Classification tools in chemistry. Part 1: linear models. PLS-DA. *Analytical Methods, 5*(16), 3790-3798.

Ballabio, D., & Todeschini, R. (2009). Multivariate classification for qualitative analysis (pp. 83-104). Burlington, MA, USA: Elsevier.

Barbin, D. F., ElMasry, G., Sun, D. W., & Allen, P. (2012). Predicting quality and sensory attributes of pork using near-infrared hyperspectral imaging. *Analytica Chimica Acta, 719*, 30-42.

Baritelle, A., Hyde, G., Thornton, R. E., & Bajema, R. (2000). A classification system for impact-related defects in potato tubers. *American Journal of Potato Research, 77*(3), 143-148.

Barnes, M., Duckett, T., Cielniak, G., Stroud, G., & Harper, G. (2010). Visual detection of blemishes in potatoes using minimalist boosted classifiers. *Journal of Food Engineering, 98*(3), 339-346.

Barredo, A. (1993). *Desarrollo y análisis de métodos de selección de variedades de patata para la producción industrial de productos destinados a la alimentación.* Universidad del País Vasco, Alava (Spain).

Bashkatov, A. N., Genina, E. A., Kochubey, V. I., & Tuchin, V. V. (2005a). Optical properties of human skin, subcutaneous and mucous tissues in the wavelength range from 400 to 2000 nm. *Journal of Physics D: Applied Physics, 38*(15), 2543-2555.

Bashkatov, A. N., Genina, E. A., Kochubey, V. I., & Tuchin, V. V. (2005b). Optical properties of the subcutaneous adipose tissue in the spectral range 400–2500 nm. *Optics and Spectroscopy, 99*(5), 836-842.

Bashkatov, A. N., Genina, E. A., & Tuchin, V. V. (2011). Optical properties of skin, subcutaneous, and muscle tissues: a review. *Journal of Innovative Optical Health Sciences, 4*(1), 9-38.

Bei, L., Dennis, G. I., Miller, H. M., Spaine, T. W., & Carnahan, J. W. (2004). Acousto-optic tunable filters: fundamentals and applications as applied to chemical analysis techniques. *Progress in Quantum Electronics, 28*(2), 67-87.

Bellon-Maurel, V., Fernandez-Ahumada, E., Palagos, B., Roger, J.-M., & McBratney, A. (2010). Critical review of chemometric indicators commonly used for assessing the quality of the prediction of soil attributes by NIR spectroscopy. *TrAC Trends in Analytical Chemistry, 29*(9), 1073-1081.

Bentini, M., Caprara, C., & Martelli, R. (2006). Harvesting damage to potato tubers by analysis of impacts recorded with an instrumented sphere. *Biosystems Engineering, 94*(1), 75-85.

Bernfeld, P. (1955). Enzymes of carbohydrate metabolism. *Methods in Enzymology, 1*, 149-150.

Bertie, J. E. (2002). Glossary of terms used in vibrational spectroscopy. *Handbook of Vibrational Spectroscopy*.

Bertran, E., Blanco, M., Maspoch, S., Ortiz, M., Sánchez, M., & Sarabia, L. (1999). Handling intrinsic non-linearity in near-infrared reflectance spectroscopy. *Chemometrics and Intelligent Laboratory Systems, 49*(2), 215-224.

Bewell, E. (1937). The determination of the cooking quality of potatoes. *American Potato Journal, 14*(8), 235-242.

Biofotonica. (2013). Spettrometria Infrarossa per applicazioni Industriali Brimrose Corporation. Retrieved 01/03/2016, from http://www.biofotonica.it/industria/Brimrose

Birth, G. S. (1978). The light scattering properties of foods. *Journal of Food Science, 43*(3), 916-925.

Bonierbale, M., Grüneberg, W., Amoros, W., Burgos, G., Salas, E., Porras, E., & Zum Felde, T. (2009). Total and individual carotenoid profiles in *Solanum phureja* cultivated potatoes: II. Development and application of near-infrared reflectance

spectroscopy (NIRS) calibrations for germplasm characterization. *Journal of Food Composition and Analysis, 22*(6), 509-516.

Bro, R. (2003). Multivariate calibration: What is in chemometrics for the analytical chemist? *Analytica Chimica Acta, 500*(1), 185-194.

Broothaerts, W., Corbisier, P., Emons, H., Emteborg, H., Linsinger, T. P., & Trapmann, S. (2007). Development of a certified reference material for genetically modified potato with altered starch composition. *Journal of Agricultural and Food Chemistry, 55*(12), 4728-4734.

Brown, C. (2005). Antioxidants in potato. *American Journal of Potato Research, 82*(2), 163-172.

Brown, C., Culley, D., Yang, C.-P., Durst, R., & Wrolstad, R. (2005). Variation of anthocyanin and carotenoid contents and associated antioxidant values in potato breeding lines. *Journal of the American Society for Horticultural Science, 130*(2), 174-180.

Brunt, K., & Drost, W. C. (2003). *Determination of potato quality by NIR.* Paper presented at the Revealing secrets of the process. Proceedings of the Fifth European Symposium on Near InfraRed (NIR) Spectroscopy.

Brunt, K., & Drost, W. C. (2010). Design, construction, and testing of an automated NIR in-line analysis system for potatoes. Part I: Off-line NIR feasibility study for the characterization of potato composition. *Potato Research, 53*(1), 25-39.

Brunt, K., Smits, B., & Holthuis, H. (2010). Design, Construction, and Testing of an Automated NIR In-line Analysis System for Potatoes. Part II. Development and Testing of the Automated Semi-industrial System with In-line NIR for the Characterization of Potatoes. *Potato Research, 53*(1), 41-60.

Büning-Pfaue, H. (2003). Analysis of water in food by near infrared spectroscopy. *Food Chemistry, 82*(1), 107-115.

Burger, J., & Geladi, P. (2007). Spectral pre-treatments of hyperspectral near infrared images: analysis of diffuse reflectance scattering. *Journal of Near Infrared Spectroscopy, 15*(1), 29-38.

Burgos, G., Salas, E., Amoros, W., Auqui, M., Muñoa, L., Kimura, M., & Bonierbale, M. (2009). Total and individual carotenoid profiles in *Solanum phureja* of cultivated potatoes: I. Concentrations and relationships as determined by spectrophotometry and HPLC. *Journal of Food Composition and Analysis, 22*(6), 503-508.

Burns, D., & Ciurczak, E. W. (1992). Handbook of near infrared spectroscopy: Dekker, New York.

Busch, J. (1999). Enzymic browning in potatoes: a simple assay for a polyphenol oxidase catalysed reaction. *Biochemical Education, 27*(3), 171-173.

CAMAG. (2012). Vitamin C in fruit juice (Quantitative determinations). Retrieved 13/01/2014, from http://www.camag.com/en/tlc_hptlc/camag_laboratory/methods.cfm

Casale, M., Casolino, C., Oliveri, P., & Forina, M. (2010). The potential of coupling information using three analytical techniques for identifying the geographical origin of Liguria extra virgin olive oil. *Food Chemistry, 118*(1), 163-170.

Commission Directive 2008/100/EC of 28 October 2008 amending Council Directive 90/496/EEC on nutrition labelling for foodstuffs as regards recommended daily allowances, energy conversion factors and definitions (2008).

Cen, H., & He, Y. (2007). Theory and application of near infrared reflectance spectroscopy in determination of food quality. *Trends in Food Science & Technology, 18*(2), 72-83.

Cevallos-Casals, B. A., & Cisneros-Zevallos, L. (2003). Stoichiometric and kinetic studies of phenolic antioxidants from Andean purple corn and red-fleshed sweetpotato. *Journal of Agricultural and Food Chemistry, 51*(11), 3313-3319.

Clark, C., Lombard, P., & Whiteman, E. F. (1940). Cooking quality of the potato as measured by specific gravity. *American Journal of Potato Research, 17*(2), 38-45.

Cozzolino, D., Fassio, A., & Gimenez, A. (2001). The use of near-infrared reflectance spectroscopy (NIRS) to predict the composition of whole maize plants. *Journal of the Science of Food and Agriculture, 81*(1), 142-146.

Cubeddu, R., D'Andrea, C., Pifferi, A., Taroni, P., Torricelli, A., Valentini, G., Ruiz-Altisent, M., Valero, C., Ortiz, C., & Dover, C. (2001). Time-resolved reflectance spectroscopy applied to the nondestructive monitoring of the internal optical properties in apples. *Applied Spectroscopy, 55*(10), 1368-1374.

Cubero, S., Aleixos, N., Moltó, E., Gómez-Sanchis, J., & Blasco, J. (2011). Advances in machine vision applications for automatic inspection and quality evaluation of fruits and vegetables. *Food and Bioprocess Technology, 4*(4), 487-504.

Changchui, H. (2008). *Welcome and introductory statement.* Paper presented at the Workshop to Commemorate The International Year of the Potato 2008, Bangkok, Thailand.

Chen, J. Y., Miao, Y., Zhang, H., & Matsunaga, R. (2004). Non-destructive determination of carbohydrate content in potatoes using near infrared spectroscopy. *Journal of Near Infrared Spectroscopy, 12*, 311-314.

Chen, J. Y., Zhang, H., Miao, Y., & Matsunaga, R. (2005). NIR measurement of specific gravity of potato. *Food Science and Technology Research, 11*(1), 26-31.

Chen, J. Y., Zhang, H., Yelian, M., & Asakura, M. (2010). Nondestructive Determination of Sugar Content in Potato Tubers Using Visible and Near Infrared Spectroscopy. *Japan Journal of Food Engineering, 11*(1), 59-64.

Chen, Y. C., & Thennadil, S. N. (2012). Insights into information contained in multiplicative scatter correction parameters and the potential for estimating particle size from these parameters. *Analytica Chimica Acta, 746*, 37-46.

Cheong, W. F., Prahl, S. A., & Welch, A. J. (1990). A review of the optical properties of biological tissues. *IEEE Journal of Quantum Electronics, 26*(12), 2166-2185.

Dacal-Nieto, A., Formella, A., Carrión, P., Vazquez-Fernandez, E., & Fernández-Delgado, M. (2011a). *Common scab detection on potatoes using an infrared hyperspectral imaging system.* Paper presented at the Image Analysis and Processing–ICIAP 2011.

Dacal-Nieto, A., Formella, A., Carrión, P., Vazquez-Fernandez, E., & Fernández-Delgado, M. (2011b). *Non–destructive Detection of Hollow Heart in Potatoes Using Hyperspectral Imaging.* Paper presented at the 14th Computer Analysis of Images and Patterns.

Dahm, D. J., & Dahm, K. D. (2007). Interpreting diffuse reflectance and transmittance. *NIR, Chichester*.

Dale, M. F. B., Griffiths, D. W., & Todd, D. T. (2003). Effects of genotype, environment, and postharvest storage on the total ascorbate content of potato (Solanum tuberosum) tubers. *Journal of Agricultural and Food Chemistry, 51*(1), 244-248.

Davies, A., & Grant, A. (1987). Review: Near infra-red analysis of food. *International Journal of Food Science & Technology, 22*(3), 191-207.

Davies, T. (1999). NIR instrumentation companies: the story so far. *NIR News, 10*(6), 14-10.

De Vries, G., Beek, J. F., Lucassen, G. W., & Van Gemert, M. J. C. (1999). The effect of light losses in double integrating spheres on optical properties estimation. *IEEE Journal of Selected Topics in Quantum Electronics, 5*(4), 944-947.

Dean, B., Jackowiak, N., Nagle, M., Pavek, J., & Corsini, D. (1993). Blackspot pigment development of resistant and susceptible *Solanum tuberosum* L. genotypes at harvest and during storage measured by three methods of evaluation. *American Potato Journal, 70*(3), 201-217.

Delwiche, S. R., & Reeves, J. B. (2004). The effect of spectral pre-treatments on the partial least squares modelling of agricultural products. *Journal of Near Infrared Spectroscopy, 12*, 177-182.

Devos, O., Ruckebusch, C., Durand, A., Duponchel, L., & Huvenne, J. P. (2009). Support vector machines (SVM) in near infrared (NIR) spectroscopy: Focus on parameters optimization and model interpretation. *Chemometrics and Intelligent Laboratory Systems, 96*(1), 27-33.

Dhanoa, M., Lister, S., Sanderson, R., & Barnes, R. (1995). The link between multiplicative scatter correction (MSC) and standard normal variate (SNV) transformations of NIR spectra. *Journal of Near Infrared Spectroscopy, 2*(1), 43-47.

Downey, G. (1996). Authentication of food and food ingredients by near infrared spectroscopy. *Journal of Near Infrared Spectroscopy, 4*, 47-62.

Du, H., Fuh, R. C. A., Li, J., Corkan, L. A., & Lindsey, J. S. (1998). PhotochemCAD: A Computer-Aided Design and Research Tool in Photochemistry. *Photochemistry and Photobiology, 68*(2), 141-142.

Dull, G. G., Birth, G. S., & Leffler, R. G. (1989). Use of near infrared analysis for the nondestructive measurement of dry matter in potatoes. *American Potato Journal, 66*(4), 215-225.

ElMasry, G., Cubero, S., Moltó, E., & Blasco, J. (2012a). In-line sorting of irregular potatoes by using automated computer-based machine vision system. *Journal of Food Engineering, 112*(1), 60-68.

ElMasry, G., Kamruzzaman, M., Sun, D. W., & Allen, P. (2012b). Principles and applications of hyperspectral imaging in quality evaluation of agro-food products: a review. *Critical Reviews in Food Science and Nutrition, 52*(11), 999-1023.

ElMasry, G., & Sun, D. W. (2010). Principles of hyperspectral imaging technology. *Hyperspectral imaging for food quality analysis and control*, 3-43.

ElMasry, G., Wang, N., Vigneault, C., Qiao, J., & ElSayed, A. (2008). Early detection of apple bruises on different background colors using hyperspectral imaging. *LWT-Food Science and Technology, 41*(2), 337-345.

Erives, H., & Fitzgerald, G. J. (2005). Automated registration of hyperspectral images for precision agriculture. *Computers and Electronics in Agriculture, 47*(2), 103-119.

Evans, M., Thai, C., & Grant, J. (1998). Development of a spectral imaging system based on a liquid crystal tunable filter. *Transactions of the ASAE, 41*(6), 1845-1852.

Evans, S., & Muir, A. (1999). Reflectance spectrophotometry of bruising in potatoes. I. Ultraviolet to near infrared. *International Agrophysics, 13*, 203-210.

Ezekiel, R., Singh, N., Sharma, S., & Kaur, A. (2013). Beneficial phytochemicals in potato—a review. *Food Research International, 50*(2), 487-496.

Faber, N. K. M., & Bro, R. (2002). Standard error of prediction for multiway PLS: 1. Background and a simulation study. *Chemometrics and Intelligent Laboratory Systems, 61*(1), 133-149.

FAO. (2013). Food and Agricultural commodities production / Commodities by regions. Retrieved 09/11/2015, from http://faostat3.fao.org/browse/rankings/commodities_by_regions/E

FAO. (2014). Production/Crops. Retrieved 09/11/2015, from http://faostat3.fao.org/browse/Q/QC/E

Fearn, T. (2011). Limitations of Mahalanobis and H distances. *NIR News, 22*(8), 16-17.

Fernández-Ahumada, E., Garrido-Varo, A., Guerrero-Ginel, J. E., Wubbels, A., Van der Sluis, C., & Van der Meer, J. M. (2006). Understanding factors affecting near infrared analysis of potato constituents. *Journal of Near Infrared Spectroscopy, 14*(1), 27-35.

Fernández Pierna, J., Vermeulen, P., Amand, O., Tossens, A., Dardenne, P., & Baeten, V. (2012). NIR hyperspectral imaging spectroscopy and chemometrics for the detection of undesirable substances in food and feed. *Chemometrics and Intelligent Laboratory Systems, 117*, 233-239.

Floegel, A., Kim, D.-O., Chung, S.-J., Koo, S. I., & Chun, O. K. (2011). Comparison of ABTS/DPPH assays to measure antioxidant capacity in popular antioxidant-rich US foods. *Journal of Food Composition and Analysis, 24*(7), 1043-1048.

Fluck, R. C., & Ahmed, E. M. (1973). Impact testing of fruits and vegetables. *Transactions of the ASAE, 16*(4), 0660-0666.

Fraser, D., Jordan, R., Künnemeyer, R., & McGlone, V. (2003). Light distribution inside mandarin fruit during internal quality assessment by NIR spectroscopy. *Postharvest Biology and Technology, 27*(2), 185-196.

Fraser, P., & Bramley, P. (2004). The biosynthesis and nutritional uses of carotenoids. *Progress in Lipid Research, 43*(3), 228-265.

Friedman, M. (1997). Chemistry, biochemistry, and dietary role of potato polyphenols. A review. *Journal of Agricultural and Food Chemistry, 45*(5), 1523-1540.

Gao, H., Li, X., Xu, S., Huang, T., Tao, H., & Li, X. (2013a). Transmission hyperspectral detection method for weight and black heart of potato. *Transactions of the Chinese Society of Agricultural Engineering, 29*(15), 279-285.

Gao, H., Li, X., Xu, S., Tao, H., Li, X., & Sun, J. (2013b). Comparative study of transmission and reflection hyperspectral imaging technology for potato damage detection. *Guang pu xue yu guang pu fen xi= Guang pu, 33*(12), 3366-3371.

García-Segovia, P., Andrés-Bello, A., & Martínez-Monzó, J. (2008). Textural properties of potatoes (*Solanum tuberosum* L., cv. Monalisa) as affected by different cooking processes. *Journal of Food Engineering, 88*(1), 28-35.

Gat, N. (2000). *Imaging spectroscopy using tunable filters: a review.* Paper presented at the AeroSense 2000.

Goetz, A., Vane, G., Solomon, J., & Rock, B. (1985). Imaging spectrometry for earth remote sensing. *Science, 228*(4704), 1147-1152.

Gómez-Sanchis, J., Gómez-Chova, L., Aleixos, N., Camps-Valls, G., Montesinos-Herrero, C., Moltó, E., & Blasco, J. (2008a). Hyperspectral system for early detection of rottenness caused by penicillium digitatum in mandarins. *Journal of Food Engineering, 89*(1), 80-86.

Gómez-Sanchis, J., Moltó, E., Camps-Valls, G., Gómez-Chova, L., Aleixos, N., & Blasco, J. (2008b). Automatic correction of the effects of the light source on spherical objects. An application to the analysis of hyperspectral images of citrus fruits. *Journal of Food Engineering, 85*(2), 191-200.

Goñi, I., Garcia-Diz, L., Mañas, E., & Saura-Calixto, F. (1996). Analysis of resistant starch: a method for foods and food products. *Food Chemistry, 56*(4), 445-449.

Gowen, A., O'Donnell, C., Cullen, P., Downey, G., & Frias, J. (2007). Hyperspectral imaging–an emerging process analytical tool for food quality and safety control. *Trends in Food Science & Technology, 18*(12), 590-598.

Gowen, A., O'donnell, C., Taghizadeh, M., Cullen, P., Frias, J., & Downey, G. (2008). Hyperspectral imaging combined with principal component analysis for bruise damage detection on white mushrooms (Agaricus bisporus). *Journal of Chemometrics, 22*(3-4), 259-267.

Griffiths, D. W., Dale, M. F. B., Morris, W. L., & Ramsay, G. (2007). Effects of season and postharvest storage on the carotenoid content of *Solanum phureja* potato tubers. *Journal of Agricultural and Food Chemistry, 55*(2), 379-385.

Guisti, M., & Wrolstad, R. (2001). Characterization and Measurement of Anthocyanins by UV-Visible Spectroscopy *Current Protocols in Food Analytical Chemistry* (Vol. 1, pp. F1.2.1.-F1.2.13). Chichester: John Wiley and Sons, Inc.

Guthrie, J., Walsh, K. B., Reid, D., & Liebenberg, C. (2005). Assessment of internal quality attributes of mandarin fruit. 1. NIR calibration model development. *Crop and Pasture Science, 56*(4), 405-416.

Haase, N. U. (2003). Estimation of dry matter and starch concentration in potatoes by determination of under-water weight and near infrared spectroscopy. *Potato Research, 46*(3-4), 117-127.

Haase, N. U. (2006). Rapid Estimation of Potato Tuber Quality by Near-Infrared Spectroscopy. *Starch-Stärke, 58*(6), 268-273.

Haase, N. U. (2011). Prediction of potato processing quality by near infrared reflectance spectroscopy of ground raw tubers. *Journal of Near Infrared Spectroscopy, 19*(1), 37-45.

Hägg, M., Häkkinen, U., Kumpulainen, J., Ahvenainen, R., & Hurme, E. (1998). Effects of preparation procedures, packaging and storage on nutrient retention in peeled potatoes. *Journal of the Science of Food and Agriculture, 77*(4), 519-526.

Hale, G. M., & Querry, M. R. (1973). Optical constants of water in the 200-nm to 200-µm wavelength region. *Applied Optics, 12*(3), 555-563.

Hamouz, K., Lachman, J., Dvorák, P., Juzl, M., & Pivec, V. (2006). The effect of site conditions, variety and fertilization on the content of polyphenols in potato tubers. *Plant Soil and Environment, 52*(9), 407-412.

Han, J.-S., Kozukue, N., Young, K.-S., Lee, K.-R., & Friedman, M. (2004). Distribution of ascorbic acid in potato tubers and in home-processed and commercial potato foods. *Journal of Agricultural and Food Chemistry, 52*(21), 6516-6521.

Hanssen, L. M., & Snail, K. A. (2002). Integrating Spheres for Mid- and Near-Infrared Reflection Spectroscopy. *Handbook of Vibrational Spectroscopy*.

Hart, J. R., Norris, K. H., & Golumbic, C. (1962). Determination of the moisture content of seeds by near-infrared spectrophotometry of their methanol extracts. *Cereal Chem, 39*(2), 94-99.

Hartmann, R., & Büning-Pfaue, H. (1998). NIR determination of potato constituents. *Potato Research, 41*(4), 327-334.

Hegney, M. (2005). Specific gravity of potatoes. https://www.agric.wa.gov.au/potatoes/specific-gravity-potato-tubers?page=0%2C1

Heise, H. M., & Winzen, R. (2002). Chemometrics in Near-Infrared Spectroscopy. In H. W. Siesler, Y. Ozaki, S. Kawata & H. M. Heise (Eds.), *Near-infrared spectroscopy: Principles, instruments, applications* (pp. 125-162). Germany: Wiley-Vch.

Helgerud, T., Segtnan, V. H., Wold, J. P., Ballance, S., Knutsen, S. H., Rukke, E. O., & Afseth, N. K. (2012). Near-infrared Spectroscopy for Rapid Estimation of Dry

Matter Content in Whole Unpeeled Potato Tubers. *Journal of Food Research, 1*(4), 55-65.

Hesen, J., & Kroesbergen, E. (1960). Mechanical damage to potatoes I. *Potato Research, 3*(1), 30-46.

Höskuldsson, A. (1988). PLS regression methods. *Journal of Chemometrics, 2*(3), 211-228.

Hruschka, W. (2001). Data Analysis: Wavelength Selection Methods. In P. Williams & K. H. Norris (Eds.), *Near-Infrared technology in the Agricultural and Food Industries* (Second ed., pp. 39-58). Minesota, USA: American Association of Cereal Chemists, Inc.

Irudayaraj, J., & Gunasekaran, S. (2001). Optical methods: visible, NIR, and FTIR spectroscopy. In S. Gunasekaran (Ed.), *Nondestructive food evaluation techniques to analyze properties and quality* (pp. 1-2). New York: Marcel Dekker.

Jarén, C., López, A., & Arazuri, S. (2016). Advanced analytical techniques for quality evaluation of potato and its products. In J. Singh & L. Kaur (Eds.), *Advances in potato chemistry and technology* (2nd ed., pp. 563-602): Elsevier.

Jeong, J. C., Ok, H. C., Hur, O. S., & Kim, C. G. (2008). Prediction of sprouting capacity using near-infrared spectroscopy in potato tubers. *American Journal of Potato Research, 85*(5), 309-314.

Jiménez, M., Rossi, A., & Sammán, N. (2009). Phenotypic, agronomic and nutritional characteristics of seven varieties of Andean potatoes. *Journal of Food Composition and Analysis, 22*(6), 613-616.

Jin, J., Li, J., Liao, G., Yu, X., Christopher, C., & Viray, L. (2009). *Methodology for potatoes defects detection with computer vision.* Paper presented at the International Symposium on Information Processing, Huangshan.

Kamruzzaman, M., Barbin, D. F., ElMasry, G., Sun, D. W., & Allen, P. (2012). Potential of hyperspectral imaging and pattern recognition for categorization and authentication of red meat. *Innovative Food Science & Emerging Technologies, 16*, 316-325.

Karagiannes, J. L., & Grossweiner, L. I. (1988). Optical Measurements on Tissue Layers. In G. Moreno, R. H. Pottier & T. G. Truscott (Eds.), *Photosensitisation, Molecular, Cellular and Medical Aspects* (pp. 111-116). Berlin Heidelberg: Springer-Verlag.

Karagiannes, J. L., Zhang, Z., Grossweiner, B., & Grossweiner, L. I. (1989). Applications of the 1-D diffusion approximation to the optics of tissues and tissue phantoms. *Applied Optics, 28*(12), 2311-2317.

Kaur, L., Singh, N., Sodhi, N. S., & Gujral, H. S. (2002). Some properties of potatoes and their starches I. Cooking, textural and rheological properties of potatoes. *Food Chemistry, 79*(2), 177-181.

Kays, S. J. (1999). Preharvest factors affecting appearance. *Postharvest Biology and Technology, 15*(3), 233-247.

Kemsley, E. K., Tapp, H. S., Binns, R., Mackin, R. O., & Peyton, A. J. (2008). Feasibility study of NIR diffuse optical tomography on agricultural produce. *Postharvest Biology and Technology, 48*(2), 223-230.

Kennard, R. W., & Stone, L. A. (1969). Computer aided design of experiments. *Technometrics, 11*(1), 137-148.

Kestens, V., Charoud-Got, J., Bernreuther, A., & Emteborg, H. (2008). Online measurement of water content in candidate reference materials by acousto-optical tuneable filter near-infrared spectrometry (AOTF-NIR) using pork meat

calibrants controlled by Karl Fischer titration. *Food Chemistry, 106*(4), 1359-1365.

Kim, A., & Wilson, B. C. (2010). Measurement of ex vivo and in vivo tissue optical properties: methods and theories *Optical-Thermal Response of Laser-Irradiated Tissue* (pp. 267-319): Springer.

Kita, A., Bąkowska-Barczak, A., Hamouz, K., Kułakowska, K., & Lisińska, G. (2013). The effect of frying on anthocyanin stability and antioxidant activity of crisps from red- and purple-fleshed potatoes (*Solanum tuberosum* L.). *Journal of Food Composition and Analysis, 32*(2), 169-175.

Kolbe, H. (1997). Einflussfaktoren auf die Inhaltsstoffe der Kartoffel. Teil VII: Vitamine (Infliences on the chemical composition of potatoes. Part VII: Vitamins). *Kartoffelbau*.

Komiyama, S., Meguro, T., Yamamoto, A., Yamaguchi, A., & Yashida, M. (2002). Effect of starch content on the cooking quality of potato. *Journal of Cookery Science of Japan, 35*(4), 336-342.

Kotíková, Z., Hejtmánková, A., Lachman, J., Hamouz, K., Trnková, E., & Dvorák, P. (2007). Effect of selected factors on total carotenoid content in potato tubers (Solanum tuberosum L.). *Plant Soil and Environment, 53*(8), 355.

Krivoshiev, G., Chalucova, R., & Moukarev, M. (2000). A possibility for elimination of the interference from the peel in nondestructive determination of the internal quality of fruit and vegetables by VIS/NIR spectroscopy. *LWT-Food Science and Technology, 33*(5), 344-353.

Krzanowski, W. (2000). *Principles of multivariate analysis*: OUP Oxford.

Kumar, D., Ezekiel, R., Singh, B., & Ahmed, I. (2005). Conversion table for specific gravity, dry matter and starch content from under water weight of potatoes grown in North-Indian plains. *Potato Journal, 32*(1-2).

Lachman, J., Hamouz, K., Hejtmánková, A., Dudjak, J., Orsák, M., & Pivec, V. (2003). Effect of white fleece on the selected quality parameters of early potato (*Solanum tuberosum* L.) tubers. *Plant Soil and Environment, 49*(8), 370-377.

Larsen, F. E. (1962). External and internal (blackspot) mechanical injury of Washington Russet Burbank potatoes from field to terminal market. *American Journal of Potato Research, 39*(7), 249-260.

Lawrence, K., Park, B., Windham, W., & Mao, C. (2003). Calibration of a pushbroom hyperspectral imaging system for agricultural inspection. *Transactions of the ASAE, 46*(2), 513-521.

Lester, G. E., Lewers, K. S., Medina, M. B., & Saftner, R. A. (2012). Comparative analysis of strawberry total phenolics via Fast Blue BB vs. Folin–Ciocalteu: Assay interference by ascorbic acid. *Journal of Food Composition and Analysis, 27*(1), 102-107.

Lin, M., Rasco, B. A., Cavinato, A. G., & Al-Holy, M. (2009). Infrared (IR) Spectroscopy–Near-Infrared Spectroscopy and Mid-Infrared Spectroscopy. In D. W. Sun (Ed.), *Infrared spectroscopy for food quality analysis and control* (First ed., pp. 119-143). Burlington, USA: Elsevier.

Lino, K., Maruo, K., Arimoto, H., Hyodo, K., Nakatani, T., & Yamada, Y. (2003). Monte Carlo simulation of near infrared reflectance spectroscopy in the wavelength range from 1000 nm to 1900 nm. *Optical Review, 10*(6), 600-606.

Lister, S., Dhanoa, M., Stewart, J., & Gill, M. (2000). Classification and comparison of *Gliricidia* provenances using near infrared reflectance spectroscopy. *Animal Feed Science and Technology, 86*(3), 221-238.

Liu, Q., Donner, E., Tarn, R., Singh, J., & Chung, H. J. (2009). Advanced analytical techniques to evaluate the quality of potato and potato starch *Advances in Potato Chemistry and Technology* (pp. 221-248).

Liu, Y., Chen, Y., Wang, C., Chan, D., & Kim, M. (2006). Development of hyperspectral imaging technique for the detection of chilling injury in cucumbers; spectral and image analysis. *Applied Engineering in Agriculture, 22*(1), 101.

López-Maestresalas, A., Aernouts, B., Beers, R., Arazuri, S., Jarén, C., Baerdemaeker, J., & Saeys, W. (2015). Bulk Optical Properties of Potato Flesh in the 500–1900 nm Range. *Food and Bioprocess Technology, 9*(3), 463-470.

López, A., Arazuri, S., García, I., Mangado, J., & Jarén, C. (2013). A review of the application of near-infrared spectroscopy for the analysis of potatoes. *Journal of Agricultural and Food Chemistry, 61*(23), 5413-5424.

Lorente, D., Aleixos, N., Gómez-Sanchis, J., Cubero, S., García-Navarrete, O. L., & Blasco, J. (2012). Recent advances and applications of hyperspectral imaging for fruit and vegetable quality assessment. *Food and Bioprocess Technology, 5*(4), 1121-1142.

Love, S. L., Salaiz, T., Shafii, B., Price, W. J., Mosley, A. R., & Thornton, R. E. (2004). Stability of expression and concentration of ascorbic acid in North American potato germplasm. *HortScience, 39*(1), 156-160.

Lu, R. (2003). Detection of bruises on apples using near-infrared hyperspectral imaging. *Transactions-American Society of Agricultural Engineers, 46*(2), 523-530.

Lu, W., Haynes, K., Wiley, E., & Clevidence, B. (2001). Carotenoid content and color in diploid potatoes. *Journal of the American Society for Horticultural Science, 126*(6), 722-726.

Lunden, A. (1956). Undersokelser over forholdet mellom potetenes spesifikke vekt og deres torrstoff-og stivelsesinnhold. *Forskn. Forsok Landbruket, 7*, 81-107.

MacKerron, D. K. L., Young, M. W., & Davies, H. V. (1995). Factors influencing the calibration of near infrared reflectometry applied to the assessment of total nitrogen in potato tissues. I. Particle size, milling speed and leaf senescence. *Journal of Near Infrared Spectroscopy, 3*, 155-166.

Madiwale, G. (2012). *Effect of genotype, storage and processing on the polyphenolic content, composition, in vitro anti-cancer activity and sensory attributes of colored-flesh potatoes.* Colorado State University, Fort Collins, CO.

Magwaza, L., Opara, U., Nieuwoudt, H., Cronje, P., Saeys, W., & Nicolaï, B. M. (2012). NIR Spectroscopy Applications for Internal and External Quality Analysis of Citrus Fruit- A Review. *Food and Bioprocess Technology 5*(2), 425-444.

Manach, C., Scalbert, A., Morand, C., Rémésy, C., & Jiménez, L. (2004). Polyphenols: food sources and bioavailability. *The American Journal of Clinical Nutrition, 79*(5), 727-747.

Manley, M. (2014). Near-infrared spectroscopy and hyperspectral imaging: non-destructive analysis of biological materials. *Chemical Society Reviews, 43*(24), 8200-8214.

Manne, R. (1987). Analysis of two partial-least-squares algorithms for multivariate calibration. *Chemometrics and Intelligent Laboratory Systems, 2*(1), 187-197.

Marini, F. (2009). Artificial neural networks in foodstuff analyses: Trends and perspectives A review. *Analytica Chimica Acta, 635*(2), 121-131.

Mark, H. (1992). Qualitative discriminant analysis. In D. Burns & E. W. Ciurczak (Eds.), *Handbook of near-infrared analysis.* (Vol. 13, pp. 329-363): Marcel Dekker.

Martelli, F. (2012). Review: ABC of near infrared photon migration in tissues: the diffusive regime of propagation. *Journal of Near Infrared Spectroscopy, 20*(1), 29-42.

Martens, H., & Martens, M. (2001). Multivariate Analysis of Quality. An Introduction. *Measurement Science and Technology, 12*(10).

Martens, H., & Naes, T. (1989). *Multivariate calibration*: John Wiley & Sons.

Martens, H., & Naes, T. (2001). Multivariate calibration by data compression. In P. Williams & K. H. Norris (Eds.), *Near-infrared technology in the agricultural and food industries* (Second ed., pp. 59-100). Minesota, USA: American Association of Cereal Chemists, Inc.

Martinsen, P., & Schaare, P. (1998). Measuring soluble solids distribution in kiwifruit using near-infrared imaging spectroscopy. *Postharvest Biology and Technology, 14*(3), 271-281.

Massart, D. L., Vandeginste, B., Buydens, L., De Jong, S., Lewi, P., & Smeyers-Verbeke, J. (1997). *Handbook of chemometrics and qualimetrics. P. B* (Vol. 20A). Amsterdam: Elsevier.

Massart, D. L., Vandeginste, B., Deming, S., Michotte, Y., & Kaufman, L. (1988). *Chemometrics: a textbook* (Vol. 2). Amsterdam: Elsevier.

Massey, P. H. (1952). *Field and storage experiments on internal black spot of potatoes*: Cornell Univ., June.

Massie, D. R., & Norris, K. H. (1965). Spectral reflectance and transmittance properties of grain in the visible and near infrared. *Transactions of the ASAE, 8*(4), 0598-0600.

Mathew, R., & Hyde, G. (1997). Potato impact damage thresholds. *Transactions of the ASAE, 40*(3), 705-709.

McClure, W. (2001). Near-infrared instrumentation. In P. Willams & K. H. Norris (Eds.), *Near-infrared technology in the agricultural and food industries* (Second ed., pp. 109-127). Minesota, USA: American Association of Cereal Chemists, Inc.

McClure, W., Moody, D., Stanfield, D., & Kinoshita, O. (2002). Hand-held NIR spectrometry. Part II: An economical no-moving parts spectrometer for measuring chlorophyll and moisture. *Applied Spectroscopy, 56*(6), 720-724.

McDermott, L. (1988). Near-infrared reflectance analysis of processed foods. *Cereal Foods World, 33*(6), 498-502.

Medina, M. B. (2011). Simple and rapid method for the analysis of phenolic compounds in beverages and grains. *Journal of Agricultural and Food Chemistry, 59*(5), 1565-1571.

Mehrubeoglu, M., & Cote, G. (1997). Determination of total reducing sugars in potato samples using near-infrared spectroscopy. *Cereal Foods World, 42*(5), 409.

Miller, C. E. (2001). Chemical principles of near-infrared technology. In P. Williams & K. H. Norris (Eds.), *Near-infrared Technology in the Agricultural and Food Industries* (Second ed., pp. 17-34). Minesota, USA: American Association of Cereal Chemists, Inc.

Miller, N. J., & Rice-Evans, C. A. (1997). Factors influencing the antioxidant activity determined by the ABTS•+ radical cation assay. *Free Radical Research, 26*(3), 195-199.

Mohsenin, N. N. (1986). *Physical Properties of Plant and Animal Materials*. New York: Gordon & Breach Science.

Muhammed, H. H. (2005). Hyperspectral crop reflectance data for characterising and estimating fungal disease severity in wheat. *Biosystems Engineering, 91*(1), 9-20.

Murray, I. (1986). Near infrared reflectance analysis of forages. In W. Haresign & D. J. A. Cole (Eds.), *Recent Advances in Animal Nutrition* (First ed., pp. 141-156): Butterworths.

Naes, T., Isaksson, T., Fearn, T., & Davies, T. (2002). *A user friendly guide to multivariate calibration and classification*: NIR publications.

Navarre, D. A., Goyer, A., & Shakya, R. (2009). Nutritional value of potatoes: vitamin, phytonutrient, and mineral content. *Advances in potato chemistry and technology*, 395-424.

Nguyen Do Trong, N., Tsuta, M., Nicolaï, B. M., De Baerdemaeker, J., & Saeys, W. (2011). Prediction of optimal cooking time for boiled potatoes by hyperspectral imaging. *Journal of Food Engineering, 105*(4), 617-624.

Ni, Y., Mei, M., & Kokot, S. (2011). Analysis of complex, processed substances with the use of NIR spectroscopy and chemometrics: Classification and prediction of properties—The potato crisps example. *Chemometrics and Intelligent Laboratory Systems, 105*(2), 147-156.

Nicolaï, B. M., Beullens, K., Bobelyn, E., Peirs, A., Saeys, W., Theron, K. I., & Lammertyn, J. (2007). Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: A review. *Postharvest Biology and Technology, 46*(2), 99-118.

Nicolaï, B. M., Defraeye, T., De Ketelaere, B., Herremans, E., Hertog, M. L., Saeys, W., Torricelli, A., Vandendriessche, T., & Verboven, P. (2014). Nondestructive measurement of fruit and vegetable quality. *Annual review of Food Science and Technology, 5*, 285-312.

Noordam, J. C., Van den Broek, W. H., & Buydens, L. (2005). Detection and classification of latent defects and diseases on raw French fries with multispectral imaging. *Journal of the Science of Food and Agriculture, 85*(13), 2249-2259.

Norris, K. H., Barnes, R. F., Moore, J. E., & Shenk, J. S. (1976). Predicting Forage Quality by Infrared Replectance Spectroscopy. *Journal of Animal Science, 43*(4), 889-897.

Norris, K. H., & Hart, J. R. (1965). Direct spectrophotometric determination of moisture content of grain and seeds. *Principles and Methods of Measuring Moisture Content in Liquids and Solids, 4*, 19-25.

Oba, K., Yamamoto, A., Ohara, A., Ishii, G., & Umemura, Y. (1998). Effect of storage temperature on vitamin C levels and L-galactonolactone dehydrogenase activity of potato tubers. *Japanese Society of Food Science and Technology, 45*, 510-513.

Osborne, B. (2000). Near-infrared spectroscopy in food analysis. *Encyclopedia of analytical Chemistry*.

Osborne, B., & Fearn, T. (1986). *Near infrared spectroscopy in food analysis*: Longman.

Osborne, B., Fearn, T., & Hindle, P. (1993). *Practical NIR spectroscopy with applications in food and beverage analysis*: Longman scientific and technical.

Otto, M. (1999). *Chemometrics. Statistics and Computer Application in Analytical Chemistry* (M. Otto Ed.). Germany: John Wiley & Sons.

Park, B., Lawrence, K. C., Windham, W. R., & Buhr, R. J. (1998). *Hyperspectral imaging for detecting fecal and ingesta contamination on poultry carcasses.* Paper presented at the 2001 ASAE Annual Meeting.

Pavia, D., Lampman, G., Kriz, G., & Vyvyan, J. (2008). *Introduction to spectroscopy*: Cengage Learning.

Pedreschi, F., Segtnan, V., & Knutsen, S. (2010). On-line monitoring of fat, dry matter and acrylamide contents in potato chips using near infrared interactance and visual reflectance imaging. *Food Chemistry, 121*(2), 616-620.

Peiris, K. H. S., Dull, G. G., Leffler, R. G., & Kays, S. J. (1999). Spatial variability of soluble solids or dry-matter content within individual fruits, bulbs, or tubers: Implications for the development and use of NIR spectrometric techniques. *HortScience, 34*(1), 114-118.

Peirs, A., Lammertyn, J., Ooms, K., & Nicolaï, B. M. (2001). Prediction of the optimal picking date of different apple cultivars by means of VIS/NIR-spectroscopy. *Postharvest Biology and Technology, 21*(2), 189-199.

Penner, M. H. (2010). Ultraviolet, visible, and fluorescence spectroscopy. In S. S. Nielsen (Ed.), *Food Analysis* (pp. 387-405): Springer.

Peña, R., García, S., Iglesias, R., Barro, S., & Herrero, C. (2001). Authentication of Galician (NW Spain) quality brand potatoes using metal analysis. Classical pattern recognition techniques versus a new vector quantization-based classification procedure. *Analyst, 126*(12), 2186-2193.

Pérez-Marín, D., Garrido-Varo, A., & Guerrero, J. E. (2007). Non-linear regression methods in NIRS quantitative analysis. *Talanta, 72*(1), 28-42.

Peters, R. (1996). Damage of potato tubers, a review. *Potato Research, 39*(4), 479-484.

Pholpho, T., Pathaveerat, S., & Sirisomboon, P. (2011). Classification of longan fruit bruising using visible spectroscopy. *Journal of Food Engineering, 104*(1), 169-172.

Pickering, J. W., Prahl, S. A., Van Wieringen, N., Beek, J. F., Sterenborg, H. J., & Van Gemert, M. J. (1993). Double-integrating-sphere system for measuring the optical properties of tissue. *Applied Optics, 32*(4), 399-410.

Porteous, R., Muir, A., & Wastie, R. (1981). The identification of diseases and defects in potato tubers from measurements of optical spectral reflectance. *Journal of Agricultural Engineering Research, 26*(2), 151-160.

Prahl, S. A. (1995). The diffusion approximation in three dimensions *Optical-Thermal Response of Laser-Irradiated Tissue* (pp. 207-231): Springer.

Prahl, S. A. (2011). Everything I think you should know about Inverse Adding-Doubling. *Oregon Medical Laser Center, St. Vincent Hospital*.

Prahl, S. A., Van Gemert, M. J. C., & Welch, A. J. (1993). Determining the optical properties of turbid media by using the adding–doubling method. *Applied Optics, 32*(4), 559-568.

Prokop, S., & Albert, J. (2008). Potatoes, nutrition and diet. Retrieved 11/01/2016, from http://www.fao.org/potato-2008/en/potato/factsheets.html

Qin, J. (2010). Hyperspectral imaging instruments. In D. W. Sun (Ed.), *Hyperspectral imaging for food quality analysis and control* (First ed., pp. 129-172). London, UK: Elsevier.

Rady, A., & Guyer, D. (2015a). Evaluation of sugar content in potatoes using NIR reflectance and wavelength selection techniques. *Postharvest Biology and Technology, 103*, 17-26.

Rady, A., & Guyer, D. (2015b). Rapid and/or nondestructive quality evaluation methods for potatoes: A review. *Computers and Electronics in Agriculture, 117*, 31-48.

Rady, A., & Guyer, D. (2015c). Utilization of visible/near-infrared spectroscopic and wavelength selection methods in sugar prediction and potatoes classification. *Journal of Food Measurement and Characterization, 9*(1), 20-34.

Rady, A., Guyer, D., Kirk, W., & Donis-González, I. (2014). The potential use of visible/near infrared spectroscopy and hyperspectral imaging to predict

processing-related constituents of potatoes. *Journal of Food Engineering, 135*, 11-25.

Rao, A. V., & Rao, L. G. (2007). Carotenoids and human health. *Pharmacological Research, 55*(3), 207-216.

Razmjooy, N., Mousavi, B. S., & Soleymani, F. (2012). A real-time mathematical computer method for potato inspection using machine vision. *Computers & Mathematics with Applications, 63*(1), 268-279.

Reyes, L., Miller, J., & Cisneros-Zevallos, L. (2005). Antioxidant capacity, anthocyanins and total phenolics in purple- and red-fleshed potato (*Solanum tuberosum* L.) genotypes. *American Journal of Potato Research, 82*(4), 271-277.

Riccioli, C., Pérez-Marín, D., Guerrero-Ginel, J. E., Saeys, W., & Garrido-Varo, A. (2011). Pixel selection for near-infrared chemical imaging (NIR-CI) discrimination between fish and terrestrial animal species in animal protein by-product meals. *Applied Spectroscopy, 65*(7), 771-781.

Rice-Evans, C. A., Miller, J. M., & Paganga, G. (1996). Structure—antioxidant activity relationship of flavonoids and phenolic acids. *Free Radical Biology and Medicine, 20*, 933-956.

Rinnan, Å., Berg, F. v. d., & Engelsen, S. B. (2009). Review of the most common pre-processing techniques for near-infrared spectra. *TrAC Trends in Analytical Chemistry, 28*(10), 1201-1222.

Robinson, J. W., Frame, E. S., & Frame II, G. M. (2005). *Undergraduate instrumental analysis*. New York: Marcel Dekker.

Rodriguez-Amaya, D. B. (2001). *A guide to carotenoid analysis in foods*. Washington, DC: ILSI press

Rodriguez-Amaya, D. B., & Kimura, M. (2004). *HarvestPlus Handbook for Carotenoid Analysis* (Vol. 2). Washington, DC, MD, USA, Cali, Colombia.

Rosén, J., & Hellenäs, K. E. (2002). Analysis of acrylamide in cooked foods by liquid chromatography tandem mass spectrometry. *Analyst, 127*(7), 880-882.

Saeys, W., Mouazen, A. M., & Ramon, H. (2005). Potential for onsite and online analysis of pig manure using visible and near infrared reflectance spectroscopy. *Biosystems Engineering, 91*(4), 393-402.

Saeys, W., Velazco-Roa, M. A., Thennadil, S. N., Ramon, H., & Nicolaï, B. M. (2008). Optical properties of apple skin and flesh in the wavelength range from 350 to 2200 nm. *Applied Optics, 47*(7), 908-919.

Salar, M. (2009). *Determinación de daños durante la recolección mecanizada y la manipulación de patata.* Universidad Pública de Navarra, Pamplona.

Samanta, D., Chaudhury, P. P., & Ghosh, A. (2012). Scab Diseases Detection of Potato using Image Processing. *International Journal of Computer Trends and Technology, 3*(1).

SASA. (2015). The European cultivated potato database. Retrieved 02/03/2015, from http://www.europotato.org/menu.php

Savitzky, A., & Golay, M. J. (1964). Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry, 36*(8), 1627-1639.

Sawyer, R., & Collin, G. (1960). Black spot of potatoes. *American Journal of Potato Research, 37*(4), 115-126.

Scanlon, M., Adam, L., & Pritchard, M. (1997). *NIR analysis of dry matter and sugars in whole potato tubers.* Paper presented at the Sensors for nondestructive testing. Measuring the quality of fresh fruits and vegetables, Orlando, Florida.

Scanlon, M., Pritchard, M., & Adam, L. (1999). Quality evaluation of processing potatoes by near infrared reflectance. *Journal of the Science of Food and Agriculture, 79*(5), 763-771.

Scudder, W. T. (1951). *Black spot of potatoes.* Cornell Univ., Ithaca, N. Y.

Schmilovitch, Z. e., Mizrach, A., Hoffman, A., Egozi, H., & Fuchs, Y. (2000). Determination of mango physiological indices by near-infrared spectrometry. *Postharvest Biology and Technology, 19*(3), 245-252.

Schulz, R. B., & Semmler, W. (2008). Fundamentals of optical imaging *Molecular Imaging I* (pp. 3-22): Springer.

Seddon, J. M., Ajani, U. A., Sperduto, R. D., Hiller, R., Blair, N., Burton, T. C., Farber, M. D., Gragoudas, E. S., Haller, J., & Miller, D. T. (1994). Dietary carotenoids, vitamins A, C, and E, and advanced age-related macular degeneration. *Jama, 272*(18), 1413-1420.

Segelstein, D. J. (1981). *The complex refractive index of water.* University of Missouri, Kansas City.

Segtnan, V. H., Kita, A., Mielnik, M., Jørgensen, K., & Knutsen, S. H. (2006). Screening of acrylamide contents in potato crisps using process variable settings and near-infrared spectroscopy. *Molecular Nutrition & Food Research, 50*(9), 811-817.

Shenk, J. S., & Westerhaus, M. O. (1991). Population definition, sample selection, and calibration procedures for near infrared reflectance spectroscopy. *Crop Science, 31*(2), 469-474.

Shenk, J. S., & Westerhaus, M. O. (1993). Analysis of agriculture and food products by near infrared reflectance spectroscopy. *Infrasoft International, Port Matilda, PA*.

Shiroma-Kian, C., Tay, D., Manrique, I. n., Giusti, M. M., & Rodriguez-Saona, L. E. (2008). Improving the screening process for the selection of potato breeding lines with enhanced polyphenolics content. *Journal of Agricultural and Food Chemistry, 56*(21), 9835-9842.

Shiroma, C., & Rodriguez-Saona, L. (2009). Application of NIR and MIR spectroscopy in quality control of potato chips. *Journal of Food Composition and Analysis, 22*(6), 596-605.

Siesler, H. W., Ozaki, Y., Kawata, S., & Heise, H. M. (2008). *Near-infrared spectroscopy: principles, instruments, applications*: John Wiley & Sons.

Skibsted, E., Boelens, H., Westerhuis, J., Witte, D., & Smilde, A. (2004). New indicator for optimal preprocessing and wavelength selection of near-infrared spectra. *Applied Spectroscopy, 58*(3), 264-271.

Smith, O., & Nash, L. (1940). Potato quality. I. Relation of fertilizers and rotation systems to specific gravity and cooking quality. *American Journal of Potato Research, 17*(7), 163-169.

Stahl, W., & Sies, H. (2005). Bioactivity and protective effects of natural carotenoids. *Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease, 1740*(2), 101-107.

Steinberger, J., & Tesar, R. (2007). *Knowledge-poor multilingual sentence compression.* Paper presented at the 7th Conference on Language Engineering (SOLE 2007), Le Caire, Egypte.

Steponavičius, R., & Thennadil, S. N. (2009). Extraction of chemical information of suspensions using radiative transfer theory to remove multiple scattering effects: application to a model two-component system. *Analytical Chemistry, 81*(18), 7713-7723.

Steponavičius, R., & Thennadil, S. N. (2011). Extraction of chemical information of suspensions using radiative transfer theory to remove multiple scattering

effects: application to a model multicomponent system. *Analytical Chemistry, 83*(6), 1931-1937.

Steponavičius, R., & Thennadil, S. N. (2013). Full Correction of scattering effects by using the radiative transfer theory for improved quantitative analysis of absorbing species in suspensions. *Applied Spectroscopy, 67*(5), 526-535.

Sternberg, J. C., Stillo, H., & Schwendeman, R. (1960). Spectrophotometric analysis of multicomponent systems using least squares method in matrix form. Ergosterol irradiation system. *Analytical Chemistry, 32*(1), 84-90.

Storey, R. (2011). The harvested crop. *Potato Biology and Biotechnology Advanced Perspectives*, 441-470.

Storey, R., & Davies, H. (1978). Tuber quality. In P. Harris (Ed.), *The potato crop. The scientific basis for improvement* (First ed., pp. 504-543): Springer.

Stuffins, C. (1967). The determination of phosphate and calcium in feeding stuffs. *Analyst, 92*(1091), 107-111.

Stushnoff, C., Holm, D., Thompson, M., Jiang, W., Thompson, H., Joyce, N., & Wilson, P. (2008). Antioxidant properties of cultivars and selections from the Colorado potato breeding program. *American Journal of Potato Research, 85*(4), 267-276.

Subedi, P., & Walsh, K. B. (2009). Assessment of potato dry matter concentration using short-wave near-infrared spectroscopy. *Potato Research, 52*(1), 67-77.

Sumner, J. B. (1924). The estimation of sugar in diabetic urine, using dinitrosalicylic acid. *Journal of Biological Chemistry, 62*(2), 287-290.

Sun, D. W. (2010). *Hyperspectral imaging for food quality analysis and control*: Elsevier.

Sun, D. W. (2011). *Computer vision technology for food quality evaluation*. London: Academic Press.

Swinehart, D. (1962). The beer-lambert law. *Journal of Chemical Education, 39*(7), 333.

Tao, Y., Heinemann, P., Varghese, Z., Morrow, C., & Sommer Iii, H. (1995). Machine vision for color inspection of potatoes and apples. *Transactions of the ASAE, 38*(5), 1555-1561.

Teow, C. C., Truong, V.-D., McFeeters, R. F., Thompson, R. L., Pecota, K. V., & Yencho, G. C. (2007). Antioxidant activities, phenolic and β-carotene contents of sweet potato genotypes with varying flesh colours. *Food Chemistry, 103*(3), 829-838.

Thybo, A. K., Bechmann, I., Martens, M., & Engelsen, S. (2000). Prediction of sensory texture of cooked potatoes using uniaxial compression, near infrared spectroscopy and low field ^{1}H NMR spectroscopy. *LWT-Food Science and Technology, 33*(2), 103-111.

Thybo, A. K., Jespersen, S. N., Lærke, P. E., & Stødkilde-Jørgensen, H. J. (2004). Nondestructive detection of internal bruise and spraing disease symptoms in potatoes using magnetic resonance imaging. *Magnetic Resonance Imaging, 22*(9), 1311-1317.

Thygesen, L., Engelsen, S., Madsen, M. H., & Sorensen, O. (2001). NIR spectroscopy and partial least squares regression for the determination of phosphate content and viscosity behaviour of potato starch. *Journal of Near Infrared Spectroscopy, 9*(2), 133-140.

Tuchin, V. V. (2007). Methods and algorithms for the measurements of optical parameters of tissues *Tissue optics: light scattering methods and instruments for medical diagnosis* (Second ed., Vol. 642, pp. 143-256). Washington, USA: SPIE press Bellingham.

Tudela, J. A., Espın, J., & Gil, M. (2002). Vitamin C retention in fresh-cut potatoes. *Postharvest Biology and Technology, 26*(1), 75-84.

USDA. (1983). *United States Standards for Grades of Processing*. Baltimore: United States Department of Agriculture.

USDA. (2016). National Nutrient Database for Standard Reference Release 28. Retrieved 11/01/2016

Van den Berg, H., Faulks, R., Granado, H. F., Hirschberg, J., Olmedilla, B., Sandmann, G., Southon, S., & Stahl, W. (2000). The potential for the improvement of carotenoid levels in foods and the likely systemic effects. *Journal of the Science of Food and Agriculture, 80*(7), 880-912.

Van Dijk, C., Fischer, M., Holm, J., Beekhuizen, J.-G., Stolle-Smits, T., & Boeriu, C. (2002). Texture of cooked potatoes (Solanum tuberosum). 1. Relationships between dry matter content, sensory-perceived texture, and near-infrared spectroscopy. *Journal of Agricultural and Food Chemistry, 50*(18), 5082-5088.

Van Gemert, M. J. C., Welch, A. J., & Star, W. M. (1995). One-Dimensional Transport Theory *Optical-Thermal Response of Laser-Irradiated Tissue* (pp. 47-72): Springer.

Van Zeebroeck, M., Tijskens, E., Van Liedekerke, P., Deli, V., De Baerdemaeker, J., & Ramon, H. (2003). Determination of the dynamical behaviour of biological materials during impact using a pendulum device. *Journal of Sound and Vibration, 266*(3), 465-480.

Vandeginste, B. G. M., Massart, D. L., Buydens, L. M. C., De Jong, S., Lewi, P. J., & Smeyers-Verbeke, J. (1998). Handbook of Chemometrics and Qualimetrics: Part B. Amsterdam: Elsevier.

Vanderslice, J. T., Higgs, D. J., Hayes, J. M., & Block, G. (1990). Ascorbic acid and dehydroascorbic acid content of foods-as-eaten. *Journal of Food Composition and Analysis, 3*(2), 105-118.

Váradi, G., Miklos, E., & Ijjasz, I. (1987). *Determination of total nitrogen content in leaf samples by near infrared (NIR) diffuse reflectance analysis.* Paper presented at the Near Infrared Diffuse Reflectance/Transmission Spectroscopy, Proceedings of the International NIR/NIT Conference.

Velleman, P. F., & Welsch, R. E. (1981). Efficient computing of regression diagnostics. *The American Statistician, 35*(4), 234-242.

Vidal, M., & Amigo, J. M. (2012). Pre-processing of hyperspectral images. Essential steps before image analysis. *Chemometrics and Intelligent Laboratory Systems, 117*, 138-148.

Vila-Francés, J., Calpe-Maravilla, J., Gómez-Chova, L., & Amorós-López, J. (2010). Analysis of acousto-optic tunable filter performance for imaging applications. *Optical Engineering, 49*(11), 113203.

Walsh, K. B., Golic, M., & Greensill, C. V. (2004). Sorting of fruit using near infrared spectroscopy: application to a range of fruit and vegetables for soluble solids and dry matter content. *Journal of Near Infrared Spectroscopy, 12*, 141-148.

Wang-Pruski, G., & Nowak, J. (2004). Potato after-cooking darkening. *American Journal of Potato Research, 81*(1), 7-16.

Wang, L., & Jacques, S. L. (1992). Monte Carlo modeling of light transport in multi-layered tissues in standard C. *The University of Texas, MD Anderson Cancer Center, Houston*.

Wang, W., & Li, C. (2013). Measurement of the light absorption and scattering properties of onion skin and flesh at 633 nm. *Postharvest Biology and Technology, 86*, 494-501.

Wang, W., Li, C., Tollner, E. W., Gitaitis, R. D., & Rains, G. C. (2012). Shortwave infrared hyperspectral imaging for detecting sour skin (Burkholderiacepacia)-infected onions. *Journal of Food Engineering, 109*(1), 38-48.

Wetzel, D. (2001). Contemporary near-infrared instrumentation. In P. Willams & K. H. Norris (Eds.), *Near-Infrared Technology in the Agricultural and Food Industries* (Second ed., pp. 129-144).

Wilson, J., & Lindsay, A. M. (1969). The relation between specific gravity and dry matter content of potato tubers. *American Potato Journal, 46*(9), 323-328.

Williams, J. D., Woodall, W. H., Birch, J. B., & Sullivan, J. H. (2006). On the Distribution of T2 Statistics Based on Successive Differences. *Journal of Quality Technology, 38*, 217-219.

Williams, P. (2003). Near-infrared technology—Getting the best out of light. *PDK Grain, Nanaimo, Canada*.

Williams, P., & Norris, K. H. (2001). *Near-infrared technology in the agricultural and food industries* (Second ed.). Minesota, USA: American Association of Cereal Chemists, Inc.

Williams, P. C. (1975). Application of near infrared reflectance spectroscopy to analysis of cereal grains and oilseeds. *Cereal Chemistry*.

Williams, P. C. (2001). Implementation of Near-Infrared Technology *Near-Infrared technology in the agricultural and food industries* (Second ed., pp. 145-169). Minesota, USA: American Association of Cereal Chemists, Inc.

Wold, H. (1966). Estimation of principal components and related models by iterative least squares. *Multivariate Analysis, 1*, 391-420.

Wold, S. (1972). Spline-funktioner-ett nytt verktyg i dataanalysen. *Kemisk Tidskrift, 3*, 34-37.

Wold, S., Antti, H., Lindgren, F., & Öhman, J. (1998). Orthogonal signal correction of near-infrared spectra. *Chemometrics and Intelligent Laboratory Systems, 44*(1), 175-185.

Wold, S., & Sjöström, M. (1998). Chemometrics, present and future success. *Chemometrics and Intelligent Laboratory Systems, 44*(1), 3-14.

Wold, S., Sjöström, M., & Eriksson, L. (2001). PLS-regression: a basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems, 58*(2), 109-130.

Woolfe, J., & Poats, S. (1987). The nutritional value of the components of the tuber. *The Potato in the Human Diet*, 19-51.

Workman, J., & Burns, D. A. (1992). Commercial NIR instrumentation. In D. Burns & E. W. Ciurczak (Eds.), *Handbook of near-infrared analysis. Practical spectroscopy series* (Vol. 13, pp. 37-51).

Workman, M., & Holm, D. (1984). Potato clone variation in blackspot and soft rot susceptibility, redox potential, ascorbic acid, dry matter and potassium. *American Potato Journal, 61*(12), 723-733.

Wu, X., Beecher, G. R., Holden, J. M., Haytowitz, D. B., Gebhardt, S. E., & Prior, R. L. (2004). Lipophilic and hydrophilic antioxidant capacities of common foods in the United States. *Journal of Agricultural and Food Chemistry, 52*(12), 4026-4037.

Xia, J., Berg, E., Lee, J., & Yao, G. (2007). Characterizing beef muscles with optical scattering and absorption coefficients in VIS-NIR region. *Meat Science, 75*(1), 78-83.

Xing, J., & De Baerdemaeker, J. (2005). Bruise detection on 'Jonagold' apples using hyperspectral imaging. *Postharvest Biology and Technology, 37*(2), 152-162.

Xing, J., Saeys, W., & De Baerdemaeker, J. (2007). Combination of chemometric tools and image processing for bruise detection on apples. *Computers and Electronics in Agriculture, 56*(1), 1-13.

Yamada, Y. (2000). Fundamental studies of photon migration in biological tissues and their application to optical tomography. *Optical Review, 7*(5), 366-374.

Young, M. W., MacKerron, D. K. L., & Davies, H. V. (1995). Factors influencing the calibration of near infrared reflectometry applied to the assessment of total nitrogen in potato. II. Operator, moisture and maturity class. *Journal of Near Infrared Spectroscopy, 3*, 167-174.

Young, M. W., MacKerron, D. K. L., & Davies, H. V. (1997). Calibration of near infrared reflectance spectroscopy to estimate nitrogen concentration in potato tissues. *Potato Research, 40*(2), 215-220.

Zamora-Rojas, E., Aernouts, B., Garrido-Varo, A., Pérez-Marín, D., Guerrero-Ginel, J. E., & Saeys, W. (2013a). Double integrating sphere measurements for estimating optical properties of pig subcutaneous adipose tissue. *Innovative Food Science & Emerging Technologies, 19*, 218-226.

Zamora-Rojas, E., Aernouts, B., Garrido-Varo, A., Saeys, W., Pérez-Marín, D., & Guerrero-Ginel, J. E. (2013b). Optical properties of pig skin epidermis and dermis estimated with double integrating spheres measurements. *Innovative Food Science & Emerging Technologies, 20*, 343-349.

Zamora-Rojas, E., Garrido-Varo, A., Aernouts, B., Pérez-Marín, D., Saeys, W., Yamada, Y., & Guerrero-Ginel, J. E. (2014). Understanding near infrared radiation propagation in pig skin reflectance measurements. *Innovative Food Science & Emerging Technologies, 22*, 137-146.

Zhao, J., Ouyang, Q., Chen, Q., & Wang, J. (2010). Detection of bruise on pear by hyperspectral imaging sensor with different classification algorithms. *Sensor Letters, 8*(4), 570-576.

Zheng, C., Sun, D. W., & Zheng, L. (2006). Recent applications of image texture for evaluation of food qualities—a review. *Trends in Food Science & Technology, 17*(3), 113-128.

Zhou, Z., Zeng, S., Li, X., & Zheng, J. (2015). Nondestructive Detection of Blackheart in Potato by Visible/Near Infrared Transmittance Spectroscopy. *Journal of Spectroscopy, 2015*.

Zhu, X., Li, S., Shan, Y., Zhang, Z., Li, G., Su, D., & Liu, F. (2010). Detection of adulterants such as sweeteners materials in honey using near-infrared spectroscopy and chemometrics. *Journal of Food Engineering, 101*(1), 92-97.

Appendix. Protocols of chemical analysis

1. Qualitative analysis

1.1. Extraction and quantitative determination of vitamin C

Tubers were homogenized in oxalic acid 4% (1:1 w/v), until a homogeneous puree was obtained (1 min) in a Palson blender at maximum speed (1200 Watts). To optimize the homogenization, 20 g of puree were weighed in a 50 ml tube and a teaspoon of steel balls was added to each tube and re-homogenized in Bullet Blender (Next Advance, Albany, NY, USA) 15 min at power 9. Then, samples were centrifuged at 3350x*g*, 10 min. The supernatant was preserved at -80ºC until analysis. Under these conditions, vitamin C was stable for at least 2 months (data not shown). On the day of analysis, samples were thawed submerging them in cold tap water for 15 min. Samples were shacked in a vortex, centrifuged for 5 min at 372xg and filtered with glass fibre filter of 1 µm. The filtered extract was oxidized with 0.5% aqueous diclorophenol-indophenol solution and subsequently mixed with 2% 2,4-dinitrophenylhydrazine in 70% sulphuric acid to allow the formation of hydrazones. The hydrazones were extracted with ethyl acetate/acetic acid 98:2, and this orange-coloured extract was applied directly onto the HPTLC plates (nano silica gel F 254, 20 x 10 cm, Fluka, Milwaukee, WI, USA) by using a semi-automatic sampler LINOMAT 5 (CAMAG, Muttenz, Switzerland). The chromatograph was developed in horizontal chamber as described in the application note A-10.5 (CAMAG, 2012). The total ascorbic acid was measured at 510 nm by a CAMAG TLC Scanner 3 and was quantified with the Wincats software (Muttenz, Switzlerland). Results were expressed as g vitamin C kg^{-1} fresh weight (FW).

1.2. Extraction and quantitative determination of total soluble phenolics

Phenolics were extracted from lyophilized powder (1 g) with 10 ml MeOH/H_2O (70:30, v/v). The solid was suspended by shaking in a vortex for 1 min. The mixture was centrifuged at 7730x*g* for 10 min at 4ºC, and the supernatant containing the extracted phenolics was collected. The extraction operation was repeated twice again with the pellet and the final volume carried to 30 ml. Phenolic compounds were quantified following the method described by Medina (2011). Results were expressed as g of Gallic Acid Equivalent (GAE) kg^{-1} FW.

1.3. Extraction and quantitative determination of total monomeric anthocyanins

Monomeric anthocyanins were extracted and quantified by the pH differential method (Guisti & Wrolstad, 2001). Anthocyanins were extracted from lyophilized powder

(0.25 g) with 10 ml MeOH/HCl (99:1, v/v). The solid was suspended by shaking in a vortex for 1 min. The mixture was centrifuged at 7730x*g* for 10 min at 4ºC, and the supernatant containing the extracted anthocyanins was collected. The extraction operation was repeated twice again with the pellet and the final volume carried to 30 ml. After the determination of the appropriate dilution factor, two solutions of each test sample were prepared, one with pH=1.0 aqueous buffer (potassium chloride, 0.025 mol L^{-1}) and the other with 4.5 pH aqueous buffer (sodium acetate, 0.4 mol L^{-1}). Total monomeric anthocyanin concentration was calculated measuring the optical density at 520 and 700 nm at two different pH values (pH=1.0 and 4.5). Absorbance was measured within 20-50 min of preparation vs MeOH/HCl (99:1, v/v) at pH 1 and 4.5 using the same dilution factor. The calculation of monomeric anthocyanin concentration (TMA), expressed as g cyanidin 3-*O*-glucoside equivalents (CGE) kg^{-1} FW, as follows in equation (12):

$$\text{TMA} = \frac{A \times DF \times MW \times V \times DM}{\varepsilon \times W \times l \times 10} \qquad (12)$$

where $A = (A_{520nm} - A_{700nm})_{pH=1.0} - (A_{520nm} - A_{700nm})_{pH=4.5}$; *DF* is the dilution factor; *MW* is the molecular weight (449.2 g mol for cyanidin 3-*O*-glucoside); *V* is the total extraction volume (in ml); DM (dry matter) is g DM kg^{-1} FW; ε is the molar extinction coefficient (26,900 cm^{-1} M^{-1} in aqueous solution); *W* is the sample weight (in g); and *l* is the path length (in cm).

1.4. Extraction and quantitative determination of total carotenoids

Total carotenoids were extracted and quantified according to Lachman *et al.* (2003) with some modifications. Total carotenoids were extracted from fresh sample (15 g) after cryogenic grinding with 10 ml of chilled acetone. Borosilicate tubes with acetonic extracts were covered with tinfoil to prevent light activity and stored 3 days at 4ºC. After this period, the borosilicate tubes were put in an ultrasound bath and sonicated for 20 min and centrifuged at 7730x*g* for 10 min at 4ºC. The supernatant was collected, 10 ml of cold acetone were added to the pellet, the solid was re-suspended by shaking in a vortex for 1 min and the mixture was centrifuged. The extraction process was repeated once again and the filtrates were made up to 25 ml. The absorbance was measured at 444 nm using acetone as blank and the total carotenoid (TC) content was expressed as g lutein equivalents (LE) kg^{-1} FW from the equation (13):

$$\text{TC} = \frac{A \times V \times 15}{0.259 \times W \times 10^3} \qquad (13)$$

where *A* is the absorbance at 444 nm; *V* is the total extraction volume (in ml); and *W* is the sample weight (in g).

1.5. Extraction and quantitative determination of hydrophilic antioxidant capacity

Hydrophilic antioxidant capacity was analysed following two methods, ABTS 2,2'-azino-bis (3-ethylbenzothiazoline-6-sulphonic acid) and DPPH (2,2-diphenylpicrylhydrazyl), according to Teow *et al.* (2007). These two indicator radicals were neutralized either by direct reduction via electron transfers or by radical quenching via H atom transfer, respectively. The ABTS assay measures the relative ability of antioxidants to scavenge the radical ABTS generated in aqueous phase compared with a Trolox standard (vitamin E analogue). The DPPH assay is based on the loss of absorption of the DPPH radical when reduced by antioxidants. Both methods are widely used to determine antioxidant capacity of fruits, vegetables and beverages (Floegel *et al.*, 2011). Fresh samples (2.5 g) were ground in a mortar in liquid nitrogen. Hydrophilic antioxidants were extracted with 10 ml MeOH/H_2O (70:30, v/v). The mixture was centrifuged at 7730x*g* for 10 min at 4ºC, and the supernatant containing the extracted antioxidant was collected. The solid was suspended by shaking in a vortex for 1 min and the extraction operation was repeated twice again with the pellet and the final volume carried to 30 ml. The $ABTS^{\bullet+}$ solution was prepared by mixing 8 mmol L^{-1} of ABTS salt with 3 mmol L^{-1} of potassium persulfate in 25 ml of DIH_2O. The solution was held at room temperature in the dark for 16 h before use. The $ABTS^{\bullet+}$ solution was diluted with MeOH/H_2O (70:30, v/v) in order to obtain an absorbance between 0.8 and 0.9 at 734 nm. Antioxidant or standard solutions, 20 µL, were mixed with 980 µL of diluted $ABTS^{\bullet+}$ solution and incubated at room temperature. A reaction of 30 min was used for all the ABTS assays. In the DPPH assay, aliquots of the hydrophilic extracts were diluted (1:10, v/v) and 0.1 ml of the diluted sample was added to 3.9 ml of $DPPH^{\bullet}$ MeOH/H_2O (70:30, v/v) solution (6 x 10^{-5} mol L^{-1}) to initiate the reaction. Absorbance was measured at 516 nm. A reaction of 3 h was used for all the DPPH assays. In both methods MeOH/H_2O (70:30, v/v) was used as blank and Trolox MeOH/H_2O (70:30, v/v) dilutions were used as standard (0, 100, 200, 300, 400 and 500 µmol L^{-1}). The antioxidant activity was reported in mol Trolox equivalents (TE) kg^{-1} FW.

2. Quantitative analysis

2.1. Determination of specific gravity, dry matter (DM) and starch content

For dry matter and starch determination, first, specific gravity of potatoes needed to be calculated. For this, about 5 kg of tubers were taken at random from the sample to be measured. These should be sound tubers free of soil because any hollow heart or adhering soil could affect the result. Tubers were weighed in air, transferred to another tared weighing basket and weighed under water. The tubers sink in the water, so their weight is heavier than an equal volume of water. The weight measured was the difference between the weight of the sample, and the weight of an equal volume of water (Kumar *et al.*, 2005; USDA, 1983). Those two weights were then applied to the equation for specific gravity calculation; as in equation (4) described in section 2.4.1. Once specific gravity of samples was estimated, DM was calculated following equation (14) (Barredo, 1993):

$$\text{DM (\%)} = 210.893904 * \text{Specific gravity}(\text{g cm}^{-3}) - 206.622201 \quad (14)$$

Similarly, starch content was estimated following equation (15) (Barredo, 1993):

$$\text{Starch (\%)} = 207.055556 * \text{Specific gravity}(\text{g cm}^{-3}) - 207.873497 \quad (15)$$

Starch content was reported as % of DW.

2.2. Estimation of reducing sugars

Estimation of reducing sugars (RS) was performed following the dinitrosalicylic acid method. This estimation was adapted by Barredo (1993) from the protocol described by Bernfeld (1955). The dinitrosalicylic acid reagent was developed by Sumner (1924) for the determination of reducing sugars. For reagent preparation, first 4.8 g of NaOH were diluted in 60 ml of distilled water. Then, 3 g of dinitrosalicylic acid and 150 ml of distilled water were added to the solution. After complete dilution, 90 g of Rochelle salt were added to a total volume of 330 ml. After that, 0.3 g of ground potato were diluted in 1 ml of distilled water in a test tube. Next, 2 ml of the dinitrosalicylic acid solution were added to the test tube. Samples were placed in a bain-marie for 5 min. Then, they were introduced in a cold water bath, shaken for 2 min and left to rest for another 10 min. Later on, 1 ml of sample solution was diluted in 5 ml of distilled water. The absorbance of samples at 546 nm was calculated with a spectrophotometer. Finally, content of reducing sugars was calculated by the relationship between absorbance and percentage in sugars, equation (16), and reported as percentage of DW:

$$RS\ (\%) = (Absorbance - 0.00385\ ) * 1.07893 \quad (16)$$

2.3. Extraction and quantitative determination of nitrogen and crude protein

For the crude protein (CP) determination, first the estimation of total nitrogen content of the samples by Kjeldahl method was carried out. The Kjeldahl method is the standard method of nitrogen determination. It consists of three basic steps: 1) digestion of the sample in sulphuric acid with a catalyst, which results in conversion of nitrogen to ammonia; 2) distillation of the ammonia into a trapping solution; and 3) quantification of the ammonia by titration with a standard solution. Once the percentage of the total nitrogen was calculated, the CP content was determined by equation (17) and reported as percentage of DW:

$$CP\ (\%) = Nitrogen\ content\ (\%) * 6.25 \quad (17)$$

2.4. Extraction and quantitative determination of total soluble phenolics

The extraction and quantitative determination of TSP was carried out in year 2 (2013) in a slightly different manner than that used in the qualitative analysis and data were given as mg GAE g^{-1} DW.

Phenolic compounds were extracted from lyophilized samples (0.5 g) with 10 ml methanol: H_2O (70:30). The solid was re-suspended by shaking in a vortex for 5 min. The mixture was centrifuged at 8000 rpm for 10 min at 4ºC, and the supernatant containing extracted phenolics was collected. The extraction operation was repeated once again with the pellet and the mixture was centrifuged at 8000 rpm for 10 min at 4ºC once again. The methanolic extracts were pooled, transferred to 20 ml volumetric flasks, and made up to the 20 ml mark. Analyses were carried out in triplicate. Total phenolics were determined by an adapted microscale protocol for the Fast Blue BB spectrophotometric method (Lester *et al.*, 2012; Medina, 2011) using gallic acid as standard. The extract solution (1 ml) was taken in a cuvette, then 0.1 ml of 0.1% FBBB [4-benzoylamino-2,5-dimethoxybenzenediazonium chloride hemi(zinc chloride) salt] dissolved in methanol was added, mixed for 30 s, followed by 0.1 ml 5% NaOH, mixed, and the resulting mixture allowed to incubate for 90 min at room temperature. Absorbance was measured at 420 nm using methanol 70% as blank. Total phenolic were calculated by interpolating the absorbance data in a calibration curve prepared with gallic acid (concentration range: 0.0-500.0 µg/ml).

2.5. Extraction and quantitative determination of hydrophilic antioxidant capacity

Hydrophilic antioxidant capacity (HAC) was analysed in year 2 (2013) following two methods, ABTS and DPPH as in the qualitative analyses but with differences to some extent.

ABTS radical-scavenging activity of the hydrophilic fractions was determined by a procedure reported by Miller and Rice-Evans (1997). The $ABTS^{\bullet+}$ solution was prepared by mixing 8 mM of ABTS salt with 3 mM of potassium persulfate in 25 ml of distilled water. The solution was held at room temperature in the dark for 16 h before use. The $ABTS^{\bullet+}$ solution was diluted with 95% methanol (approximately 600 µl ABTS to 40 ml 95% methanol), in order to obtain an absorbance between 0.8 and 0.9 at 734 nm. Fresh $ABTS^{\bullet+}$ solution was prepared for each analysis. Antioxidant or standard solutions, 30 µl, were mixed with 1 ml of diluted $ABTS^{\bullet+}$ solution and incubated at 30ºC. The absorbance at 734 nm was read every minute for 30 min using a spectrophotometer. Methanol (70%) was used as a blank. Trolox with concentrations from 0 to 1000 µM was used as a standard. The free-radical-scavenging activity was expressed as µmoles of Trolox per gram (g) of sample (µmol TE/g DW).

Analysis of HAC by DPPH method was performed following the procedure described by Teow *et al.* (2007). For this, 30 µl of sample was pipetted into 1.17 ml of $DPPH^{\bullet}$ solution to initiate the reaction. The absorbance was read every minute at 515 nm for 180 min using a spectrophotometer. Under these conditions, the decrease in absorbance reached a plateau within the 2 h sampling period. Therefore, a reaction time of 2 h was used for all the $DPPH^{\bullet}$ assays. Methanol (70%) was used as a blank. Trolox with concentrations from 0 to 1000 µM was used as a standard. The antioxidant activity was reported in µmoles of Trolox equivalents per g sample (µmol TE/g DW).

List of publications

List of publications

International journals indexed in the JCR

López, A., Arazuri, S., García, I., Mangado, J., & Jarén, C. (2013).
A review of the application of near-infrared spectroscopy for the analysis of potatoes.
Journal of Agricultural and Food Chemistry, *61*(23), 5413-5424.

López-Maestresalas, A., Aernouts, B., Beers, R., Arazuri, S., Jarén, C., Baerdemaeker, J., & Saeys, W. (2015).
Bulk Optical Properties of Potato Flesh in the 500-1900 nm range.
Food and Bioprocess Technology, 9(3), 463-470.

Tierno, R., López, A., Riga, P., Arazuri, S., Jarén, C., Benedicto, L., & Ruiz de Galarreta, J. I. (2016).
Phytochemicals determination and classification in purple and red fleshed potato tubers by analytical methods and near infrared spectroscopy.
Journal of the Science of Food and Agriculture, 96, 1888-1899.

López-Maestresalas, A., Keresztes, J. C., Goodarzi, M., Arazuri, S., Jarén, C., & Saeys, W. (2016).
Non-destructive detection of blackspot in potatoes by VIS-NIR and SWIR hyperspectral imaging.
Food Control (Submitted).

Book chapters

Jarén, C., López, A., & Arazuri, S. (2016).
Advanced analytical techniques for quality evaluation of potato and its products. In J. Singh & L. Kaur (Eds.), Advances in Potato Chemistry and Technology (2nd ed., pp. 563-602): Elsevier.

Communications in conferences

López, A; Pérez, C; Tierno, R; Riga, P; Arazuri, S; Ruiz de Galarreta, J. I; Jarén, C (2016). Clasificación de cultivares de patata de acuerdo a su contenido en antocianinas mediante espectroscopia en el infrarrojo cercano (NIRS). In: VIII Congreso de Mejora Genética de Plantas, July 2016, Vitoria, Spain.

A. López-Mestresalas, B. Aernouts, R. Van Beers, S. Arazuri, C. Jarén, J. De Baerdemaeker, W. Saeys (2016). Optical characterization of potato tissues in the Vis-NIR range using double integrating spheres. In: International Conference of Agricultural Engineering, CIGR - AgEng 2016, June 2016, Aarhus, Denmark.

C. Pérez-Roncal, A. López-Maestresalas, C. Jarén, S. Arazuri (2016). Evaluation of impacts in a potato handling and packaging line with sensors based on triaxial accelerometers. In: International Conference of Agricultural Engineering, CIGR - AgEng 2016, June 2016, Aarhus, Denmark.

López, A.; Goodarzi, M.; Keresztes, J.; Arazuri, S.; Jarén, C.; Saeys, W. (2015). Classification of Potato Varieties using Visible and Near Infrared Hyperspectral Image Analysis coupled with Chemometrics. In: SSC14 Symposium. 14th Scandinavian Symposium on Chemometrics, June 2015, Sardegna, Italy.

Arazuri, S.; López, A.; Pérez, C.; Jarén, C. (2015). Near infrared spectroscopy combined with Partial Least Square Discriminant Analysis for bruise damage detection on potatoes. In: SSC14 Symposium. 14th Scandinavian Symposium on Chemometrics, June 2015, Sardegna, Italy.

Nguyen Do Trong, N.; López-Maestresalas, A.; Arazuri, S.; Jarén, C.; Saeys, W. (2014). Spatially resolved diffuse reflectance measurements in the wavelength range 400-1000 nm for characterization and classification of raw potato tissues. In: 7th International Conference on the Food Factory for the Future, November 2014, Uppsala, Sweden.

López-Maestresalas, A.; Arazuri, S.; Jarén, C.; Tierno, R.; Riga, P.; Ruiz de Galarreta, J.I. (2014). Ability of NIRS to estimate total phenolic contents of lyophilized potatoes.

In: XXVI Congreso asociación Latinoamericana de la papa-ALAP, September, 2014, Bogotá, Colombia.

López, A.; Jarén, C.; Arazuri, S.; Mangado, J.; Tierno, R.; Ruiz de Galarreta, J.I.; Riga, P. (2014). Estimation of the total phenolic content in potatoes by NIRS. In: International Conference of Agricultural Engineering, July, 2014, Zurich, Switzerland.

Goodarzi, M.; Keresztes, J.; Van Roy, J.; Khojastehnazhand, M.; Moscetti, R.; López, A.; Morshchavka, S.; Saeys, W. (2014). Hyperspectral imaging coupled with chemometrics in food sensing. In: XIV Chemometrics in Analytical Chemistry, June, 2014, Virginia, USA.

López, A.; Arazuri, S.; Jarén, C.; Mangado, J.; Arnal, P.; Ruiz de Galarreta, J.I.; Riga, P.; López, R. (2013). Crude protein content determination of potatoes by NIRs technology. In: 6th International Conference on Information and Communication Technologies in Agriculture, Food and Environment, September, 2013, Corfu, Greece.

López, A.; Arazuri, S.; Mangado, J.; Jarén, C. (2013). Determinación de la Penetración NIR en patata. In: VII Congreso Ibérico de Agroingeniería y Ciencias Hortícolas, Agosto, 2013, Madrid, Spain.

Arazuri, S.; Mangado, J.; Jarén, C.; López, A.; Ruiz de Galarreta, J.I.; Riga, P. (2013). Determinación del punto de adquisición de datos para la clasificación de variedades de patata mediante tecnología NIRS. In: VII Congreso Ibérico de Agroingeniería y Ciencias Hortícolas, Agosto, 2013, Madrid, Spain.

Jarén, C.; López, A.; Arazuri, S.; Ruiz de Galarreta, J.I.; Riga, P.; López, R. (2013). Determinación del contenido de almidón en patata por tecnología NIRS. In: VII Congreso Ibérico de Agroingeniería y Ciencias Hortícolas, Agosto, 2013, Madrid, Spain.

López-Maestresalas, A.; Jarén, C.; Arazuri, S.; Mangado, J.; Ruiz de Galarreta, J.I.; Riga, P. (2013). Influence of different pretreatments in the classification of potato varieties by NIRS. In: 16th International Conference on Near Infrared Spectroscopy, June, 2013, La Grande-Motte, France.

Others

López, A., Arazuri, S., Jarén, C., Mangado, J., Arnal, P., Ruiz de Galarreta, J. I., Riga, P., & López, R. (2013).
Crude Protein content determination of potatoes by NIRS technology.
Procedia Technology, *8,* 488-492.

López, A.; Arazuri, S.; Mangado, J.; Jarén, C.; Ruiz de Galarreta, J.I. (2013).
Determinación de la penetración NIR en patata.
Horticultura, 309, 38-43.

López, A.; Jarén, C.; Arazuri, S. (2014).
Analysis of the influence of the skin on the near infrared absorbance spectra of potato tubers.
NIR News, 25 (3), 6-8.