Universitat Autònoma de Barcelona

Virtualization and real-time analysis of pharmaceutical and food products by near infrared spectroscopy

Dong Sun

Tesi doctoral

Programa de Doctorat de Química

Directores: Dr. Manel Alcalà Bernàrdez

Prof. Dr. Marcelo Blanco Romia

Departament de Química

Facultat de Ciències

2017

Chapter 9

Monitoring chemical and sensory parameters of tomato product with near infrared spectroscopy

9. Monitoring chemical and sensory parameters of tomato product with NIRS

Tomato is one of the most important crops in the world and its quality affects the heath of people all round the world. In this study, the NIRS was adopted to monitor the chemical and sensory parameters of tomato, such as the soluble solids, dry matter content, concentration of glucose and fructose, sweetness, taste intensity, aroma intensity, mealiness, acidity, crunchiness, juiciness and explosiveness. Quantitative models were constructed by PLSR. The validation models showed that the NIRS had a good capability for the quality control of tomato product both in chemical and sensory parameters. For the chemical parameters, RSEP was lower 5.7. For sensory parameters, RMSEP was from 0.2063 to 1.0445 with the scale between 0-10.

Key words: Quality control of tomato, Near infrared spectroscopy, Chemical and sensory parameters

9.1 Introduction

Tomato product is one of the most important crops around the world in both fresh vegetable market and food processing industry. In Spain, the production of tomato is around 4 million tons every year. [1] A main part of this crop is thermally processed and concentrated into tomato paste and the other part is purchased by consumers in raw. The major substance contained in tomato fruit is water. Then the soluble and insoluble solid consist of around 7% of this crop [2]. The quality of tomato product is controlled by several factors which can be classified into two groups, the chemical and sensory parameters. Important quality attributes such as the soluble solids, dry matter content, concentration of glucose and fructose, sweetness, taste intensity, aroma intensity, mealiness, acidity, crunchiness, skin perception, juiciness and explosiveness need to be analyzed by the manufacturer. Soluble solids and dry mater content determine the product of tomato paste, so they can indicate the factory yield. The concentration of glucose and fructose affects the sweetness of product. And sensory parameters would influence the consumer likeability of tomato directly.

In food industry, there are many methods to analyze the food quality. They can be divided into 2 parts, the subjective methods and the objective methods. The subjective methods are depended on the opinions of the analysts. They involve the physiological response of the trained analysts, the influence of personal preference and the individual perception. The objective methods are carried out with recognized standard scientific tests which exclude the sensory of analysts. These objective methods are selected according to some key factors such as speed, precision, accuracy and ruggedness [3]. Normally, the definitive methods used for nutritional labeling purposes have high precision and accuracy, for example high performance liquid chromatography (HPLC) for analyzing the sugar content. Secondary methods used for the rapid on line measurement, for example the refractive index

method, are much less accurate than the previous one. Near infrared spectroscopy (NIRS) is a powerful tool to realize the subjective and objective analysis mentioned above in food industry. Real time analysis is one of advantages of NIRS [4]. Sample spectra could be acquired in real time without or with minimal sample preparation, and the analysis time is around 1 min. And the NIR instruments have high precision, accuracy and ruggedness. Therefore, the food quality can be controlled exactly without a long time delay. Besides, NIRS is sensitive to both the chemical and physical changes in samples so it gives us an opportunity to correlate not only chemical but also sensory parameters with the NIRS spectra [5, 6]. By this way, some subjective analysis depended on personal experience would have a more stable and accurate method to obtain the result. Since in the NIRS analysis do not require the use of solvents, the technique presents a big advantage over other analytical methodologies and can be considered as a green approach [7].

In this study, NIRS has been applied to analyze process products of tomato. It is a fast, low cost and environmental friendly technique. Quantitative models of CQAs have been calculated by PLSR. The performance of models meets food industry specifications.

9.2 Materials and methods

9.2.1 Reference method

9.2.1.1 Chemical analysis

There were 316 juice samples and 270 puree samples used for the chemical analysis. For every added tomato sample, about 500 g of tomato duplicates were cut into pieces and homogenized in a blender for 3 min. Homogenates were frozen and stored in polyethylene pots at -20ºC for 2 months. Prior to analysis, each homogenate was thawed at 4ºC during 4 hours and rehomogenized.

Soluble solids, dry matter, were directly determined in the homogenates. Soluble solids were determined by using a hand refractometer (Erma, Japan) and expressed as Brix. The dry matter content was measured by drying of the samples in air oven to constant weight (65ºC, 72 h). The sugars were extracted from tomato using deionised water. About 30 g of homogenate were three consecutive times mixed with 20-30 mL of water, shaken for 15 minutes and centrifuged. The three filtrated supernatants were brought together to a volume of 100mL. Glucose and fructose were analyzed by HPLC equipped with a pump (Beckman 110B, USA), an injector (Hewlett Packard Serie 1100, USA) and a Refractive Index Detector (Beckman 156, USA). A Luna NH2 column, 250 mm x 4.6 mm (Phenomenex, USA) was used.

9.2.1.2 Sensory analysis

Descriptive profiling was performed by a trained panel comprising 9 panelists with more than 10 years of experience in sensory analysis of tomatoes and other vegetables [8]. Prior to sensory sessions, panelists underwent 6 training sessions in order to refresh attribute descriptions and the range of variation for each attribute. All the sensory sessions took place in individual booths, meeting the standards set out by the International Organization for Standardization [9]. Panelists evaluated 4 taste related traits (sweetness, acidity, taste intensity, aroma intensity) and 3 texture related traits (skin perception, mealiness and crunchiness). For the cherry genotypes, two texture related traits (juiciness and explosiveness) were added considering their high impact on consumer acceptance. A definition of each attribute can be found in Hongsoongnern and Chambers [10], and in Sinesio et al.[11]. There were 55 puree samples used for analyzing sweetness, taste intensity, aroma intensity, mealinessm, acidity and crunchiness. And 26 puree samples were used for analyzing the juiciness and explosiveness.

In order to avoid bias in the samples presented to each panelist, owing to the high fruit-to-fruit variability existing in tomato for chemical composition [12], taste related traits were evaluated in a puree proceeding from 6 to 10 tomatoes. Chemical and NIRS analysis were further performed in aliquots of each puree, thus reinforcing the correlation ships between sensory and chemical data, as panelists evaluated strictly the same samples latter analyzed by means of instrumental analysis. Texture related traits were evaluated in longitudinal slices for fresh market type tomatoes and in fruit halves for the cherry types. Panelists scored their ratings using a 100 mm unstructured scale, left side corresponding to the lowest intensity (score=0), and right side corresponding to the higher intensity (10) of the attribute. Genotypes were evaluated per triplicate, and distributed across the sessions randomly. In each session panelists evaluated a maximum of 5 samples.

9.2.2 Spectra measurement

There were two kinds of samples, the juice and the puree. Because the physical state between them was quite different, specific accessory of the spectrometer was selected for each of them.

9.2.2.1 Tomato juice

Sample spectra were acquired by a Foss NIR spectrometer system model 5000 which equipped with a Rapid Content Analyzer at ambient temperature. Firstly, 10 ml juice of every sample was added into a

reflectance vessel. And then 0.5 mm immersion gold reflector was placed into the transflectance vessel to fix the optical path length at 1 mm. At last, the vessel was put into the box of analyzer and the spectrum was acquired in 30 seconds. The spectral range was from 1100-2498 nm, scan number were 32 and the resolution were 2 nm. For every sample, three spectra were acquired. Background air spectra were renewed every one hour using the same vessel and gold reflector.

9.2.2.2 Tomato puree

Sample spectra were acquired by the Foss NIR system model 5000 equipped with an OptiProbe Analyzer at ambient temperature. In order to fit the probe, 10 ml puree of every sample was added into a plastic tube. Then the probe was put into the tube and immerged by the 10ml puree. There is a reflect mirror in front of probe window and the optical path length is 1 mm. Finally, the spectrum was acquired in 30 seconds. The spectral range was from 1100-2498 nm, scan number were 32 and the resolution were 2 nm. For every sample, three spectra were acquired. Background air spectra were acquired every one hour.

9.2.3 Data analysis

In this study, two groups of parameters were determined by NIRS, one group consisted of four chemical parameters and the other group consisted of nine sensory parameters. With the same spectrum, these thirteen parameters can be determined by PLS models. Multivariate quantitative models were calculated with Unscrambler X (Trondheid, Norway).

For chemical parameters, the glucose content, the fructose content, the soluble solids, and the dry matter content were analyzed.

In order to determine the concentration of glucose and fructose by NIRS, two kinds of samples have been used. They were juice and puree. Because the physical changes between juice and puree have caused a lot of differences in the baseline of raw NIRS spectra, their quantitative models had to be constructed separately. In total, there were four models for the determination of glucose and fructose content.

For the soluble solids and dry matter content, only spectra of tomato puree were acquired. Because puree samples contain flesh, seeds and peels of tomato, they could offer an entire view about the soluble solids and dry matter in products.

In the group of sensory parameters, sweetness, taste intensity, aroma intensity, mealiness, acidity, crunchiness, skin perception, explosiveness and juiciness of tomato have been analyzed. Spectra of

sensory parameters were acquired with puree. Because the flesh, seeds and peels in puree were important factors which have influence on human sense.

The workflow for developing models included division of data sets, spectra pretreatment, bands selection, outlier selection, calibration and validation of models. Fig. 9.1 shows that, at first, samples need to be divided into calibration set and validation set. Then, the acquired raw spectra needed to be pretreated, otherwise the useless signal and baseline drift can immerse available information of target substances and decrease the quality of calibration. SNV, MSC, Savitzky-Golay smoothing and the derivative have been applied to the raw spectra. Despite some models used the whole spectrum, by selecting bands, more simple and accurate models can be built. Specific spectral bands were selected to cover the absorbance peaks of the target substance and reduce the weight of noise and interferences in calibration models. In the selection of outlier samples, sample leverage, X variances residual and Y variances residuals were calculated to select outlier samples. Samples which show a high value of them are considered as outliers.

Full cross validation was performed to define the best factor number in calibration. The optimal LVs number is the one with the highest explained Y and offer a minimum residual in prediction value.

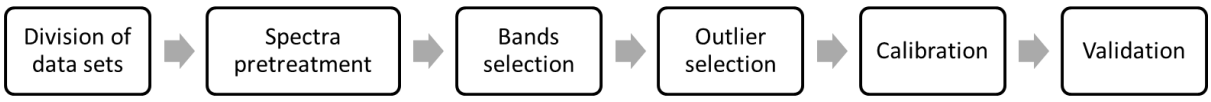


Fig. 9.1 workflow of model development

The predictive ability of the models was evaluated by slope, correlation, and RMSE of calibration and validation.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}\left(y_i^{pred} - y_i^{ref}\right)^2}{n}} \qquad \text{Equation 9.1}$$

Where n is the number of samples, y^{ref} is the reference value and y^{pred} is the predicted value of calibration or validation.

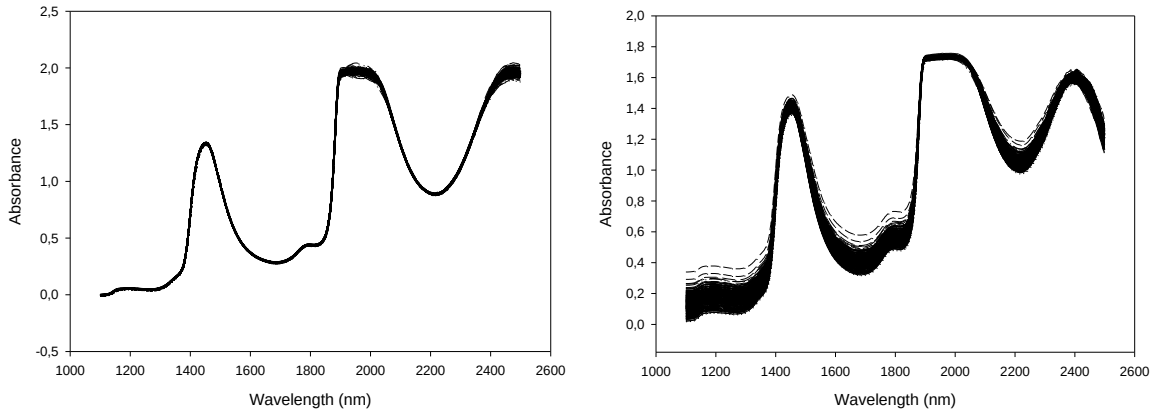
9.3 Result and discussion

9.3.1 Chemical parameters

There were 316 juice samples and 270 puree samples in total. They were divided into the calibration set and the validation set by using the PCA score plot. For juice samples, there were 256 samples in calibration set. And for puree samples, there were 210 samples in calibration set. Both of them had a prediction set with 60 samples.

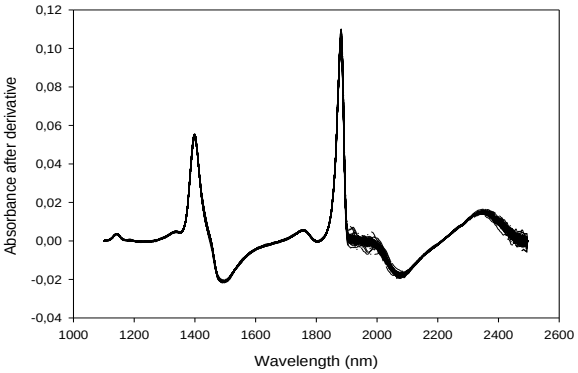
9.3.1.1 Spectra pretreatment

Fig. 9.2 shows the raw spectra of tomato samples, a) is juice and b) is puree. We can see that, at 1386-1578 nm and 1860-2212 nm, there were two absorbance bands. And from 2350 to 2498 nm, there was noise signal. These two high peaks were from the combination and 1st overtone of O-H bond and could decrease the signal weight of target substance. Compared to a), the gap between every spectrum on whole range was larger in b). This baseline drift was caused by the solid maters in puree. Therefore, raw spectra need to be pretreated. Several pretreatment methods were tested in order to find out the best way to pretreat the data. Savitzky-Golay smoothing, 1st derivative, 2nd derivative, full MSC and gap-segment derivative were good representatives of them. Fig. 9.3 shows the results of spectra after different pretreatments.

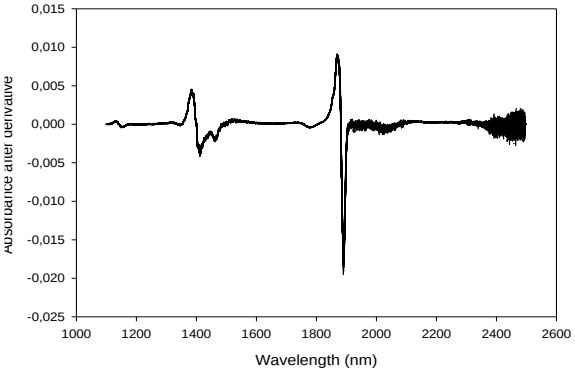


a) Tomato juice b) Tomato puree

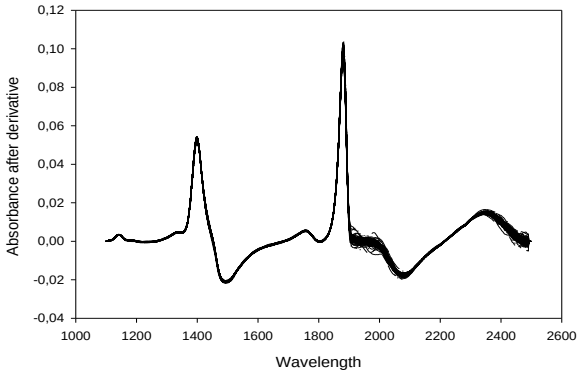
Fig.9.2 Raw spectra of juice and puree



Spectra of juice with Savitzky-Golay smoothing and 1st derivative, order 2, 5 points.

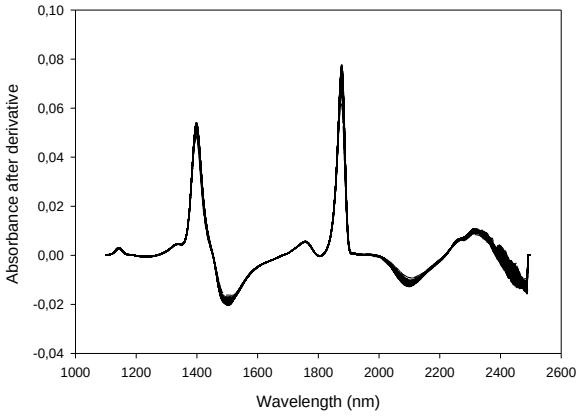


Spectra of juice with Savitzky-Golay smoothing and 2nd derivative, order 2, 5 points.

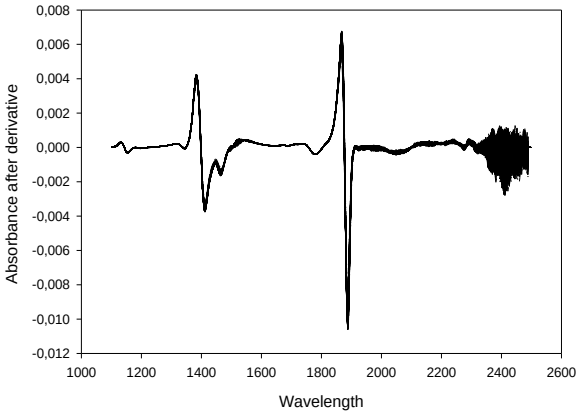


Spectra of juice with MSC and 1st derivative, gap size 5,

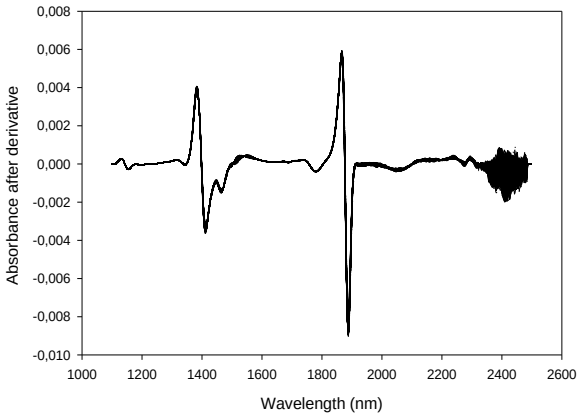
a) Pretreated spectra of juice



Spectra of pure with Savitzky-Golay smoothing and 1st derivative order 2, 11 points



Spectra of pure with Savitzky-Golay smoothing and 2nd derivative order 2, 11 points



Spectra of puree with MSC and 2nd derivative, gap size 3, segment size 3

b) Pretreated spectra of puree

Fig. 9.3 Spectra with different pretreatments, a) is the pretreated juice spectra and b) is the pretreated puree spectra.

Table 9.1 shows the performance of the models calculated with pretreated spectral range from 1100-2498 nm. Compared to the raw spectra, we can find that the baseline drift in the pretreated spectra has been reduced. The absorbance of water peaks was still very high, and from 2328 nm to the end of spectra there was noise signal. The peaks of water and noise could affect the quality of calibration model and decrease the contribution of available peaks. Therefore, it is necessary to select the spectral bands that represent information of available information to get the most efficient model.

Table 9.1 Calibration models with different pretreatment in whole spectral range. SG means Savitzky-Golay smoothing, 1st D means 1st derivative and 2nd D means 2nd derivative.

Parameter	Pretreatment	Calibration			
		Slop	Correlation	RMSEC	Factor
Fructose in juice	SG and 1st D, order 2, 5 points.	0.8194	0.9052	1.0728 (g/L)	7
	SG and 2nd D, order 2, 5 points.	0.8382	0.9155	1.0153 (g/L)	7
	MSC and 1st D, cap size 5, segment size 2	0.8402	0.9166	1.0090 (g/L)	7
Glucose in juice	SG and 1st D, order 2, 5 points.	0.8322	0.9122	1.0615 (g/L)	7
	SG and 2nd D, order 2, 5 points.	0.823	0.9072	1.0901 (g/L)	7
	MSC and 1st D, cap size 5, segment size 2	0.9003	0.9488	0.8181 (g/L)	7
Fructose in puree	SG and 1st D order 2, 11 points	0.8741	0.935	0.2153 (g/100g)	3
	SG and 2nd D order 2, 11 points	0.784	0.8854	0.2821 (g/100g)	3
	MSC and 2nd D, cap size 3, segment size 3	0.6951	0.8337	0.3352 (g/100g)	3
Glucose in puree	SG and 1st D order 2, 11 points	0.9261	0.9624	0.1758 (g/100g)	4
	SG and 2nd D order 2, 11 points	0.8566	0.9255	0.2450 (g/100g)	4
	MSC and 2nd D, cap size 3, segment size 3	0.8208	0.906	0.2738 (g/100g)	4
Brix in puree	SG and 1st D order 2, 11 points	0.9154	0.9568	0.3878 (°Bx)	4
	SG and 2nd D order 2, 11 points	0.8128	0.9016	0.6071 (°Bx)	4
	MSC and 2nd D, cap size 3, segment size 3	0.8103	0.9002	0.6112 (°Bx)	4
Dry matter in puree	SG and 1st D order 2, 11 points	0.8653	0.9302	0.5745 (g/100g)	3
	SG and 2nd D order 2, 11 points	0.7871	0.8872	0.7223 (g/100g)	3
	MSC and 2nd D, cap size 3, segment size 3	0.6762	0.8223	0.8907 (g/100g)	3