Qualitative and Quantitative Models Based on Handheld NIR Spectroscopy to Monitor the Tomato Fruit Development During Early and Full Season

C. Camps*, L. Deltheil and C. Gilli

Agroscope, Institute for Plant Production Sciences (IPS), Route Des Vergers 18, CH-1964 Conthey, Switzerland

Abstract: The present study aimed at to follow the tomato fruit development and quality by hand-held near-infrared spectroscopy. Tomato quality were followed from few days after fruit setting until harvest at commercial maturity during two seasons (spring and summer).

Results showed that in both seasons, fruit can be classified from fruit setting to harvest at maturity by using qualitative models (factorial discriminant analyses).

Quantitatives models based on PLS regressions allowed the prediction of soluble solids content (R=0.9, RMSE=0.1%Brix), titrable acidity (R=0.9, RMSE=0.6méq.100g^{-1}) and color (a*, R=0.9, RMSE=5) of fruit. The accuracy of the predictions depend on the season and also on the maturity stage.

the results are promising in the context of developing a tool to assist in fruit phenotyping on site. Other experiment are now necessary to improve the accuracy and the robustness of the models with including additional varieties growing under variable climatic conditions in our greenhouses.

Keywords: Tomato, handheld NIRs, Fruit development, season.

1. INTRODUCTION

Study the effect of agricultural practices on the quality of tomato needs hundreds analysis of fruit quality during their development. Such analyses are time consuming, expensive and need the destruction of fruits. The destruction of fruit make it impossible to follow the quality construction on the same fruit. Thereby, the development of a methodology allowing the monitoring of fruit quality on plant (without picking-up the fruit) is suitable.

Near-infrared spectroscopy is widely used to develop predictive models aiming at to non-destructively determine various fruit quality [1-5] and in particular tomato [6-13].

Recently, a study has been carried out on the development of NIR-based models allowing a rapid phenotyping of some quality traits of cherry tomato cv. Micro-Tom [14]. In this study the authors analysed tomato fruit at three maturity stages close to the harvest date (from mature green to fully ripe). The prediction of quality traits was promising for SSC, pH, color and firmness.

In most studies, the variability of tomato due to growing season and fruit maturity are not taken into account. The aim of the present study is to take into account the variability due to growing season and the fruit maturity to build chemometric models based on handheld near infrared spectroscopy. The quality of tomatoes was monitored from fruit setting until harvest at commercial maturity. Because the climatic conditions strongly influence the construction of the fruit quality, monitoring the quality of tomatoes grown in greenhouse was carried out two times in the growing season, (1) early in the season (March-May) and (2) in full summer season (June-July). Then, because the tomato growth is strongly marked by maturation stage, chemometric models will be built with fruit before maturation and fruit at maturation.

2. MATERIAL AND METHODS

2.1. Monitoring of Fruit Quality

Tomatoes (*Solanum lycopersicum*) endeavour variety, were grown in a glasshouse under controlled climatic conditions. The plantation took place the 14th February 2013 at a density of 3,5 shoot/m^2. Climate instructions were as follows: a humidity (Dx) lower than 3g/kg, temperatures of 17-19-21°C (day, night, aeration).

A first batch of tomato fruit (n=162) was picked during early season (March to May), and a second batch (n=99) during summer (June - July). Fruits were picked every 3-4 days during their development until harvest. During spring period, 18 picking dates were carried out while 11 were necessary during summer period.

*Address correspondence to this author at the Agroscope, Institute for Plant Production Sciences (IPS), Route Des Vergers 18, CH-1964, Switzerland; Tel: +41 58 481 35 30; E-mail: cedric.camps@agroscope.admin.ch

E-ISSN: 2409-9813/14

2.2. Fruit Quality

Tomatoes were ground using a robot ("Electric tomato sauce sieve mod. Testarossa") and the homogenate-free skin and seed is recovered for further operations. The resulting homogenate is centrifuged 2 minutes at 10'000 rpm. The supernatant part is used to soluble solids content and titrable acidity measurements.

The soluble solids content of fruit has been performed using a digital refractometer (Reichert r2mini Digital Pocket Refractometer, USA). Total Acidity (TA) has been measured by using a titrimeter (Metrohm, 719S, Titrino). 5 g of juice was titrated with NaOH (0.1 mol/L) and the results were expressed in meq $100g^{-1}$. Spectro-colorimetry (Minolta C.O., LTD, Chroma-meter CR-400) was used to characterize the background color of whole tomato. Three parameters were taken into account, L* (lightness), a* (red to green) and b* (blue to yellow) components.

2.3. Near-Infrared Spectroscopy and Chemometry

Spectra of tomato fruit were acquired in reflectance mode and direct contact analysis (DCA) using a MEMS based PHAZIR (NIR PHAZIR 1018, Anatec, Eke, Belgium). Spectral acquisition was carried out by placing the fruit in contact with the extremity of the NIR pistol. Absorbance spectra (average of 30 scans) were recorded at a resolution of 8 nm from 950 to 1800 nm. Before analyzing the set of samples, a white reference scan was carried out using a piece of Spectralon®. 3 measurements per fruit were carried out. A total collection of 702 spectra has been collected, 432 spectra for spring period and 270 for summer period. NIR measurements have been performed on the last 16 picking dates before harvest of spring period and the 10 last picking dates before harvest of summer period. Before these dates, fruits were too small to be analysed with the NIR device. Spectra were pre-treated by standard normal variate (SNV) in order to reduce the effects of uncontrolled baseline and the intensity variations of absorption bands [15, 16] .

2.4. Qualitative Models: Factorial Discriminant Analysis

Factorial discriminant analyses (FDA) were carried out on spectral data. A given spectrum curve forms a vector x_i of p wavelengths. The n spectra were gathered into a matrix $\mathbf{X}$ dimensioned $n \times p$. Due to the collinear nature of the wavelength absorbances, it was impossible to directly perform FDA. In order to cope with this collinearity, a modified version of FDA was applied [17]. In FDA, the qualitative groups to be discriminated were the picking dates expressed as the number of days before the harvest at commercial maturity. Therefore, during spring 16 picking dates covered a period from 48 days before harvest to harvest at commercial maturity. During summer, 10 picking dates covered a period from 41 days before harvest to harvest at commercial maturity.

A criterion of efficiency of the FDA is the proportion of correctly classified observations in validation sets. These validation tests were carried out by dividing the data matrices $\mathbf{X}$ into a training and a validation set. The FDA model was computed on the calibration set. The observations of the validation set were then classified using the established model.

The observations correctly classified were then counted and expressed in percentages. Such validation tests were independently carried out ten times, placing two thirds ($2n/3$) of the observations in the calibration set and the remaining ones ($n/3$) in the validation set.

FDA computes a set of discriminant scores, which are linear combinations of the original variables. The discriminant scores are new "synthetic variables" calculated so they can discriminate the observations. The correlation between the discriminant scores and the predictive variables will be calculated in order to highlight the relevant wavelengths implied in the FDA models.

2.5. Quantitative Models: Partial Least Square Regression

Spectra were gathered in a matrix $\mathbf{X}_{n,p}$ where n is the number of spectra (n=702) and p the number of wavelength steps (p=100). The reference-values (SSC, TA, L*, a* or b*) were gathered in a column vectors $\boldsymbol{y}_{n,1}$. The models were elaborated in three steps: 1) determination of the optimal number of latent variables (LV) to be used in the final model, 2) calibrate the model and 3) performing a cross-validation using a test-set.

The step 1 has been achieve by using two different methods. Firstly, the maximisation of the correlation coefficient (R-value) and the minimisation of the root mean square error (RMSE) in a leave-one-out procedure. Secondly, the CovSel method has been used to confirm the choice of the optimal number of LV [18]. To achieve the step 2 and 3, $\mathbf{X}_{n,p}$ has been divided into 2 subsets, placing two thirds ($2n/3$) of the

observations in the calibration set and the remaining ones (*n*/3) in the validation set.

The accuracy and goodness of models has been evaluated according to several indicators: the coefficient of determination (R^2), root mean square error corrected for bias (RMSEc), the ratio performance to deviation (RPD) [19] and the ratio performance to interquartile (RPIQ) [20]. All data analyses were performed with Matlab R2013 and partially with SAISIR Package version 1.0 (http://www.chimiometrie.fr/saisir_webpage.html).

3. RESULTS

3.1. Monitoring of Fruit Quality

Soluble solids content (SSC), total acidity (TA), calibre (CAL), fresh weight (FW) and color (L*, a* and b*) are the main quality traits currently used in practice to evaluate a batch of tomato. These traits were followed from few days after fruit setting until harvest at commercial maturity during early season (March-May) and full season (June-July) (Figure **1**). Early in the season, the fruit develops in approximately 55 days, while only 45 days are required in summer. During the 30 days before fruit maturity, the calibre and weight of fresh fruit grow in a similar way (Figure **1A**, **B**). Thus, this is before the last 30 days that the development is longer for the fruits of spring compared to summer. During the last 30 days, the development of calibre stabilizes, it is the end of the intensive phase of cell expansion [21]. Fruit maturation occurs between 10 and 14 days before harvest at commercial maturity. Then, the rapid coloring of fruits (Figure **1C**, **D**) and peaks of SSC and TA values during early maturation are observed (Figure **1E**, **F**). These increases in SSC and TA cancel before harvest. Between 40 and 30 days before harvest, the level of acidity decreased from 12 to 6 méq.$100g^{-1}$. This decrease could indirectly correspond to the influence of cycle of malate that play a role in transitory accumulation of starch allowing the increase of SSC at maturation time. Indeed, some study showed the influence of mitochondria cycle of malate on activation of AGPase and subsequent starch accumulation in plastids. Such accumulation during cell enlargement phase will be responsible of the SSC increase in cytosol in addition to glucose and fructose accumulation during maturation of fruit [22-25].

The follow-up of fruit development at these two seasons gave interesting variability of quality traits. Such variability will be used in the next parts of the present study to the development of qualitative and quantitative models using handheld near-infrared spectroscopy.

3.2. Qualitative Models

Factorial discriminant analyses have been performed on spectral data recorded from hand-held near-infrared spectroscopy. The groups to be discriminated are the picking dates expressed as the number of days before harvest.

In a first step, the optimal numbers of variables to be introduced in the models have been determined. The Figure **1A** and **1B** show the percent of correct classification by FDA according to the number of introduced variables for spring and Summer models. The optimal numbers of variables for the two models were 11 and 8 for spring and Summer models, respectively (Figure **2A**, **B**). Indeed, additional dimensions do not allow to improve the percent of correct classification and only introduce noise in the models. FDAs allowed a correct classification of 55 and 71% of fruits in validation step for spring and summer models, respectively. Factorial maps according to the first two factorial scores are presented in Figure **2C** and **D**. Concerning spring model, the ellipses of the picking dates form a parametric arc from 48 days before harvest to harvest date (0 days) (Figure **2C**). The first factorial score describe the overall variability of the parametric arc. The second factorial score describes a secondary variability where the period corresponding to 30 days before harvest would be the inflexion point. Concerning the summer model, the variability of spectral data from 41 to 30 days before harvest is described by the first factorial score while the variability of remaining data (30 days before harvest until harvest date) is explained by the second factorial score (Figure **2D**). In both models, the spectral data before and after the 30^{th} days before harvest seems to evolve differently. The Figure **3** presents the relevant wavelengths of FDA models. Most of these important wavelengths related to CH, CH_2 and CH_3 absorption bands (1^{st}, 2^{nd} or 3^{rd} overtones) and H_2O absorption band (2^{nd} overtone). The pattern of the curves presenting the relevant wavelengths is very similar for the spring and summer models.

3.3. Quantitative Models

Quantitative models aiming at to predict quality traits of tomato for both seasons (spring and summer) and global models gathering the two seasons data

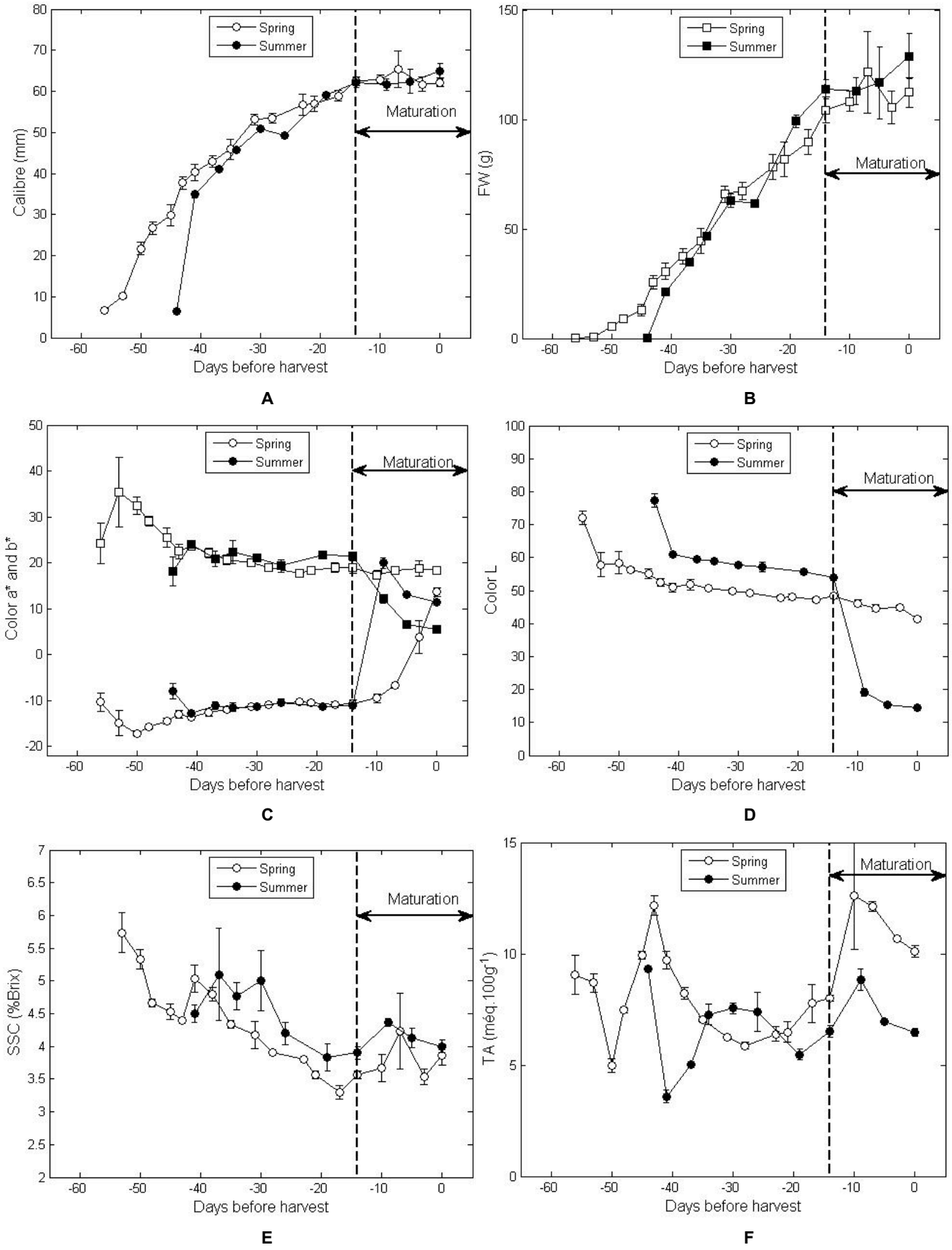

Figure 1: Fruit quality along their development in glasshouse during spring (white symbols) and summer (black symbols). Calibre (**A**), FW (**B**), Color a* and b* (**C**), Color L* (**D**), SSC (**E**), TA (**F**). The error bars represent the standard deviations.

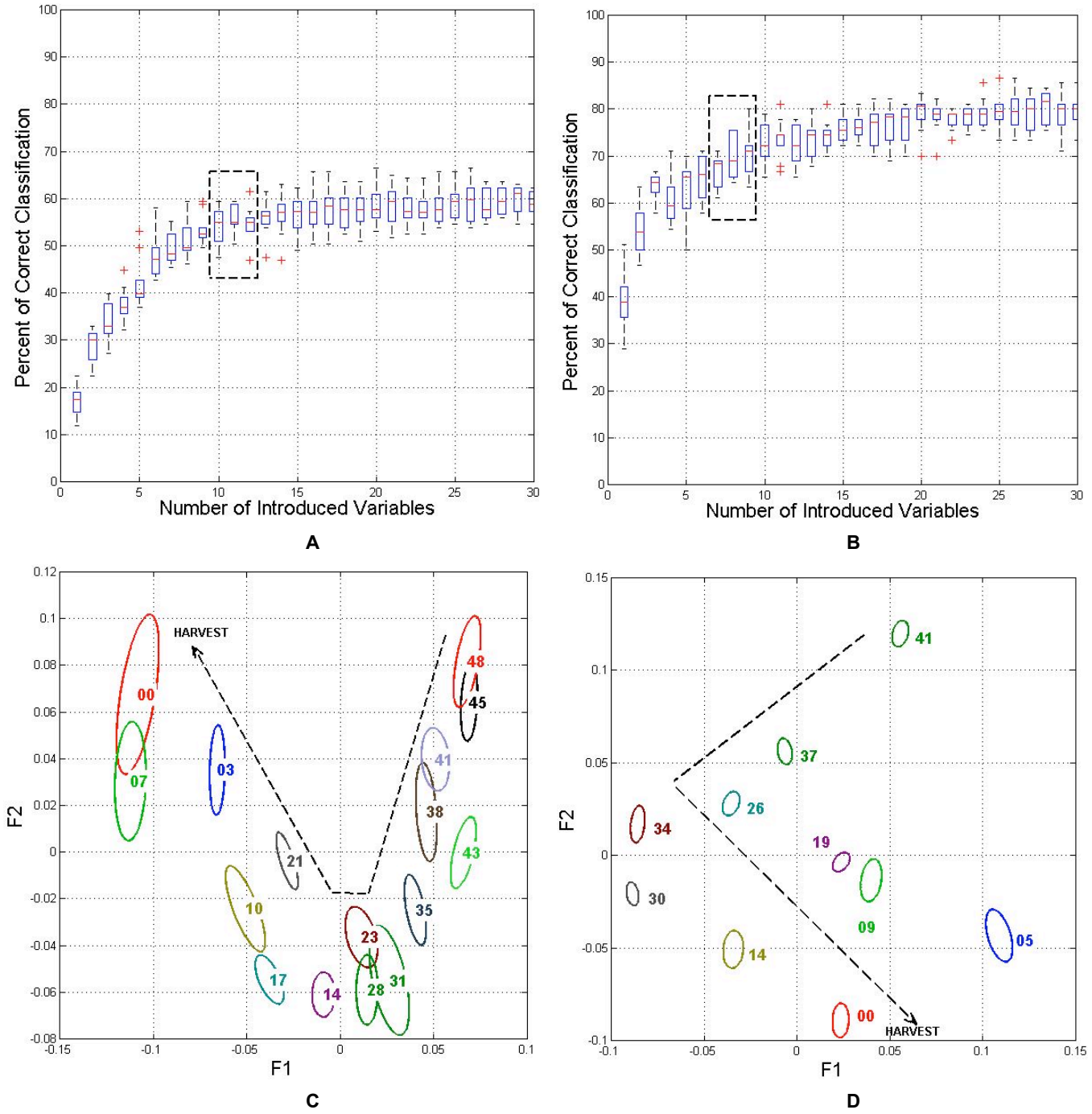

Figure 2: Percent of correct classification by FDA according to the number of introduced variables for "Spring" and "Summer" models (**A** and **B**). FDA maps of "Spring" and "Summer" models according to the first two factorial scores (**C** and **D**).

were built. In a first step, the optimal numbers of latent variables (LV) to be introduced in the models have been determined. This step is crucial to cope with the over-fitting effect due to a too large amount of introduced LV in models. The minimization of RMSE and maximization of R in a leave-one-out procedure and the maximization of the co-variance (CovSel) were the two methods used in the present study. The results of PLS regressions are summarized in the Tables **1** and **2**.

SSC

Models depending on the season (spring and summer) and fruit maturity stages (before 'BM' and during maturation 'M') have been carried out. During spring, the best model was achieved with fruit before maturation (R=0.9, RMSE = 0.2%Brix). During summer, best models was built with fruit during maturation (R=0.9, RMSE = 0.1%Brix). Models gathering the fruit from both seasons were slightly less

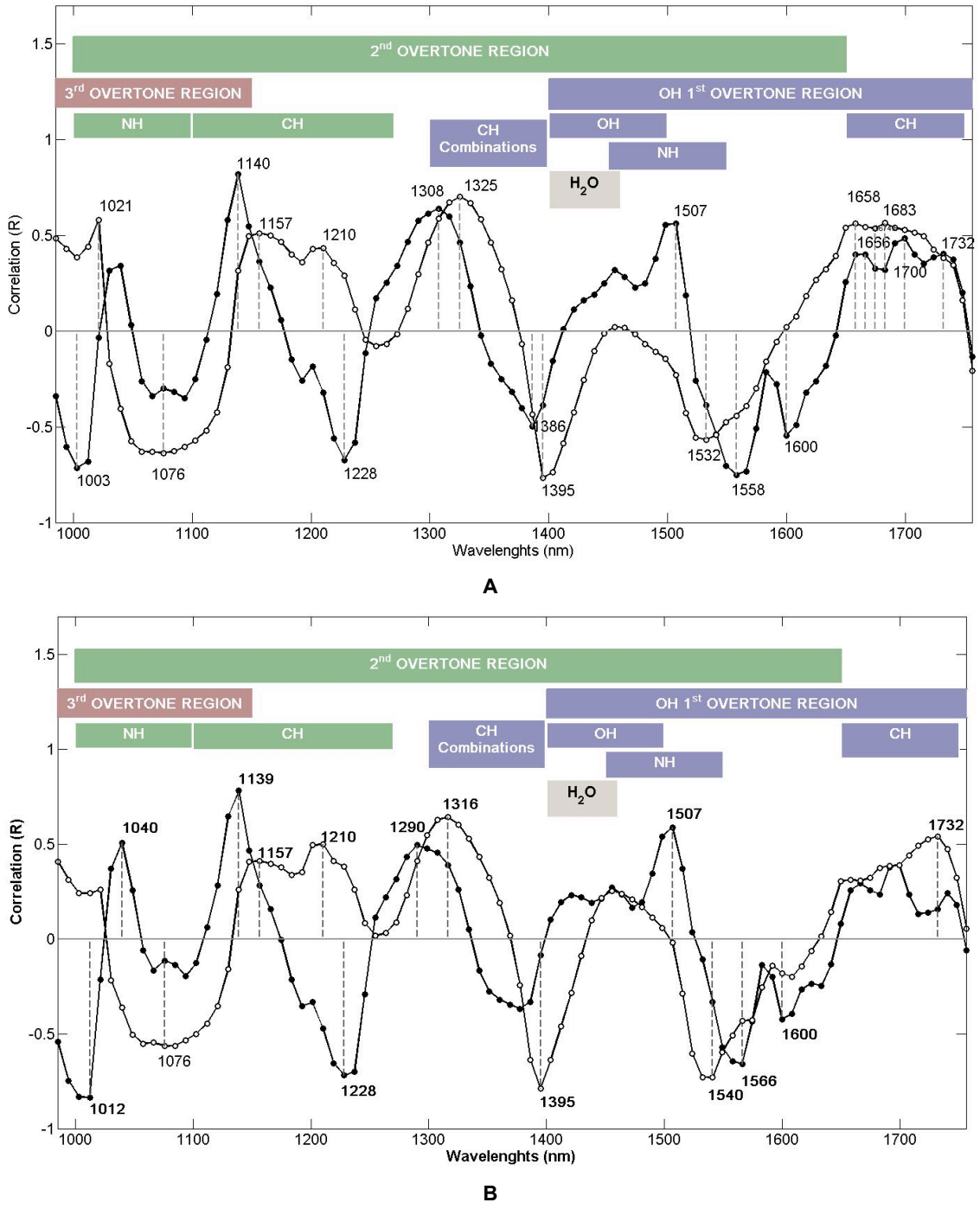

Figure 3: "Relevant Wavelengths in FDA" Correlation between the first two factorial scores of the FDA with the raw spectral data. "Spring model" (**A**) and "Summer model" (**B**). Correlation with the first factorial score F1 (●) and the second factorial score F2 (○).

accurate with fruit picked before maturation but models performed with fruit during maturation remained correct (R= 0.9, RMSE = 0.1%Brix).

Models previously described as correct presented RPD-values ranged from 1.8-2.2 (spring), 1.9-2.2 (summer) and 1.7-2.5 (spring + summer). These values must be improved to reach valuable models. The accuracy of our SSC model is comparable to those reported in the literature and based on tomato cultivars used by producers [7, 10] Actual vs. predicted values of SSC are presented Figure **4**.

TA

TA is one of the most important quality trait of tomato and one of the most difficult to predict by near-infrared spectroscopy. In the present study, follow the

Table 1: PLS-values of prediction of SSC and TA. Season: spectral and reference data used to build the PLS model according to the season, data from early season (spring), full season (summer) and gathered data (spring + summer). Subset: data used to build the PLS models according to the development stage of fruit, before maturation (BM), maturation (M) and gathered data (BM + M). Step: Calibration (C) and Validation (V). LV: number of introduced latent variables in the PLS models. R: coefficient of correlation. RMSEc: root mean square error corrected for bias. RPD: ratio performance to deviation. RPIQ: ratio performance to interquartile. CV: coefficient of variation of reference values

Season		Spring						Summer						Spring + Summer					
Subset		BM+M		BM		M		BM+M		BM		M		BM+M		BM		M	
Step		C	V	C	V	C	V	C	V	C	V	C	V	C	V	C	V	C	V
SSC	LV	9	9	6	6	6	6	7	7	8	8	6	6	9	9	8	8	8	8
	R	0.87	0.79	0.90	0.89	0.71	0.55	0.77	0.53	0.88	0.79	0.91	0.86	0.72	0.66	0.77	0.74	0.93	0.85
	RMSEc	0.26	0.32	0.23	0.24	0.28	0.30	0.30	0.43	0.25	0.34	0.07	0.09	0.32	0.40	0.32	0.40	0.10	0.14
	RPD	1.73	1.45	2.08	1.84	1.00	0.97	1.20	0.88	1.84	1.41	2.24	1.90	1.04	0.89	1.23	0.83	2.54	1.70
	RPIQ	3.13	3.11	4.13	3.74	1.07	1.00	2.00	1.40	2.40	2.06	4.29	3.33	1.88	2.00	2.34	2.00	3.00	2.14
	CV	12.6	12.7	12.8	12.7	10.2	8.9	10.8	11.1	12.1	12.3	4.3	4.0	11.6	13.1	12.4	14.2	6.7	6.5
TA	LV	5	5	5	5	7	7	8	8	7	7	8	8	8	8	3	3	6	6
	R	0.51	0.49	0.61	0.58	0.84	0.57	0.92	0.76	0.97	0.94	0.98	0.84	0.79	0.69	0.73	0.62	0.92	0.88
	RMSEc	1.99	1.91	1.42	1.49	0.69	1.43	0.63	1.15	0.35	0.57	0.22	0.52	1.41	1.7	1.46	1.71	0.42	0.53
	RPD	0.59	0.61	0.77	0.71	1.51	0.87	2.34	1.47	4.46	2.67	4.78	1.46	1.3	1.08	1.06	0.83	2.28	1.77
	RPIQ	1.78	1.84	1.57	1.4	1.96	1.11	1.57	1.37	7.11	4.46	9.77	2.77	2.08	1.39	1.22	0.92	2.05	1.7
	CV	26.5	24.7	22.6	23.3	11.1	15.0	23.3	24.0	25.4	26.4	13.7	12.9	29.3	28.9	28.8	29.3	15.4	16.0

Table 2: PLS-values of prediction of L*, A* and B* colour indicators. Season: spectral and reference data used to build the PLS model according to the season, data from early season (spring), full season (summer) and gathered data (spring + summer). Subset: data used to build the PLS models according to the development stage of fruit, before maturation (BM), maturation (M) and gathered data (BM + M). Step: Calibration (C) and Validation (V). LV: number of introduced latent variables in the PLS models. R: coefficient of correlation. RMSEc: root mean square error corrected for bias. RPD: ratio performance to deviation. RPIQ: ratio performance to interquartile. CV: coefficient of variation of reference values

Season		Spring						Summer						Spring + Summer					
Subset		BM+M		BM		M		BM+M		BM		M		BM+M		BM		M	
Step		C	V	C	V	C	V	C	V	C	V	C	V	C	V	C	V	C	V
L*	LV	6	6	7	7	7	7	8	8	5	5	8	8	8	8	8	8	4	4
	R	0.88	0.89	0.89	0.79	0.92	0.66	0.93	0.91	0.88	0.81	0.96	0.75	0.84	0.82	0.94	0.93	0.97	0.97
	RMSEc	2	1.83	1.49	1.9	0.99	2.07	7.47	7.3	1.33	1.52	0.73	1.7	7.1	7.95	1.96	2.01	3.93	3.9
	RPD	1.84	1.94	2	1.4	2.27	1.23	2.47	2.36	1.9	1.45	3.78	1.36	1.55	1.41	2.67	2.68	3.99	4
	RPIQ	2.74	2.67	2.69	2.21	3.76	1.64	5.46	0.97	3.12	2.67	5.13	1.44	1.36	1.3	4.56	4.72	8.16	8.39
	CV	8.6	8.3	6.6	6.2	5.6	5.6	46.9	37.7	5.0	4.6	17.6	15.6	28.2	31.2	10.8	10.3	50.7	49.6
A*	LV	9	9	7	7	6	6	7	7	6	6	7	7	8	8	5	5	9	9
	R	0.91	0.76	0.91	0.79	0.89	0.90	0.86	0.85	0.70	0.50	0.95	0.90	0.84	0.77	0.53	0.40	0.93	0.90
	RMSEc	3.43	4.56	0.81	1.16	4.59	4.71	6.37	6.64	0.85	0.98	1.51	2.01	5.84	7.01	0.95	0.93	4.12	5.02
	RPD	2.18	1.33	2.17	1.37	1.94	2.18	1.68	1.73	0.97	0.84	2.96	2.06	1.51	1.23	0.62	0.58	2.54	1.94
	RPIQ	0.9	0.68	3.46	2.41	3.97	4.97	3.51	3.46	1.41	1.38	2.9	2.67	1.15	0.48	1.45	1.51	5.29	4.78
	CV	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
B*	LV	7	7	6	6	2	2	8	8	6	6	7	7	8	8	6	6	3	3
	R	0.88	0.84	0.89	0.86	0.48	0.17	0.94	0.86	0.67	0.52	0.92	0.87	0.80	0.74	0.69	0.58	0.84	0.77
	RMSEc	1.71	2.02	1.71	1.8	1.87	2.09	2.52	3.21	1.84	2.2	1.35	2.29	2.93	3.43	1.78	2.02	3.17	3.71
	RPD	1.82	1.56	1.94	1.56	0.55	0.36	2.65	1.57	0.91	0.83	2.26	1.84	1.34	1.19	0.94	0.84	1.54	1.32
	RPIQ	2.69	1.88	2.68	2.39	1.87	1.41	5.74	1.95	1.52	1.52	1.55	2.6	1.38	1.21	1.67	1.39	3.87	3.1
	CV	17.2	18.3	17.5	16.5	11.9	11.4	43.0	35.4	11.6	12.0	44.4	52.4	27.1	28.8	12.1	12.3	43.5	39.8

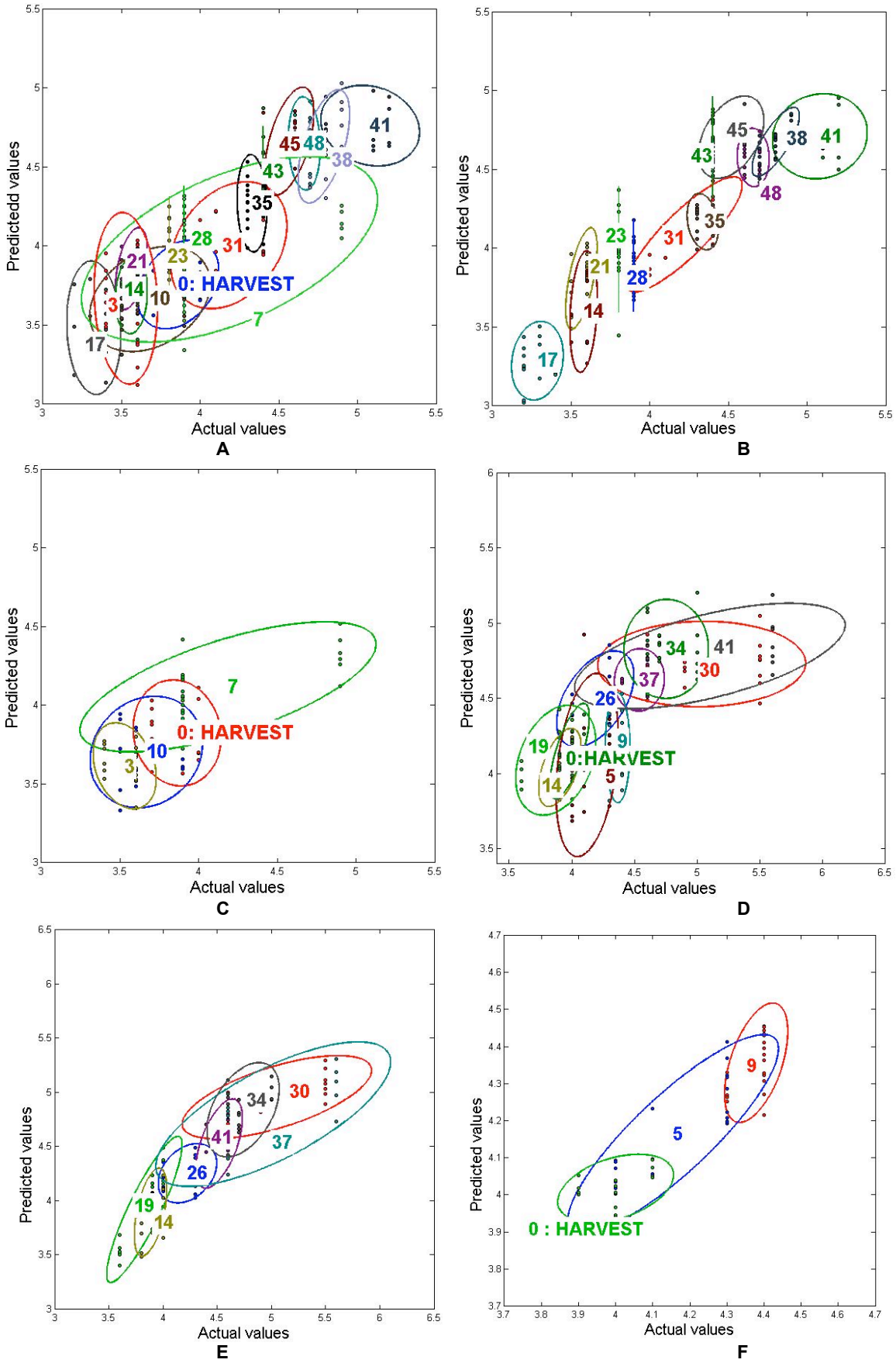

Figure 4: Actual vs. predicted values of SSC parameter with PLS models. Spring models gathering BM and M data (**A**), BM data (**B**) and M data (**C**). Summer models gathering BM and M data (**D**), BM data (**E**) and M data (**F**). Confidence ellipses (p=0.05).

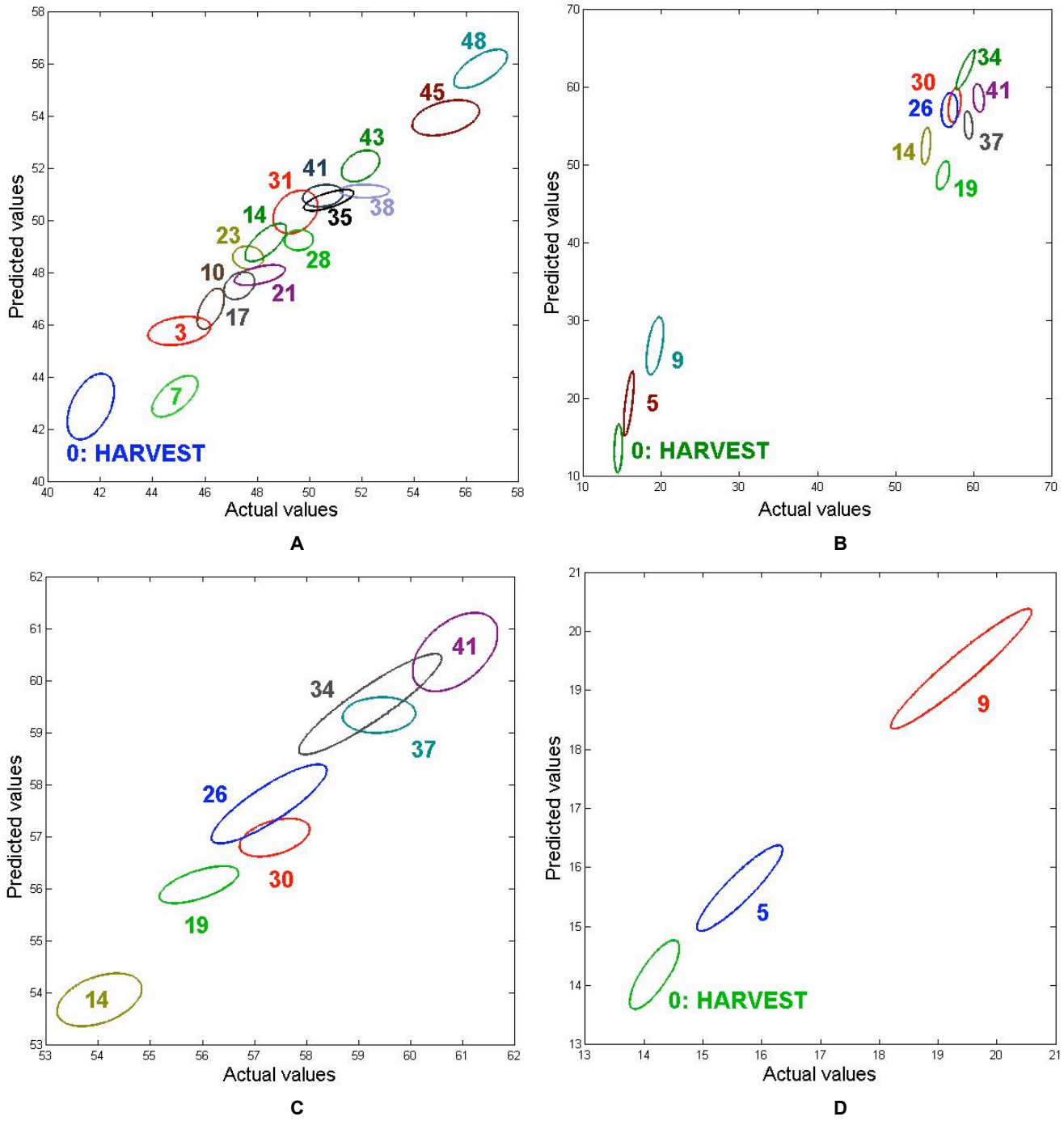

Figure 5: Actual vs. predicted values of L* parameter with PLS models. Spring (**A**) and Summer (**B**) models gathering BM and M data. Summer model of L* prediction with BM data (**C**) and M data (**D**). Centroids of confidence ellipses (p=0.05).

fruit development from early stages to harvest at maturity allowed to collect data with a large range of TA-value (3.41 to 12.42 méq.100g^{-1}). During spring, best models was built with fruit at maturation stage (R=0.6-0.8, RMSE = 0.7-1.4 méq.100g^{-1}). During summer, correct predictions were reached with both maturity stages of fruit (before maturation and fruit at maturation) (R=0.9, RMSE ≤ 0.6 méq.100g^{-1}). RPD-values were correct for models of summer (RPD=4.8) and both seasons (RPD=2.28) but low for spring model (RPD=1.5).

Color

Results of prediction of L*, a* and b* for seasonal models were as accurate as for global one. L* was predicted with an accuracy of about 2, 7.5 and less than 8 L* color units during spring, summer and when gathered fruit from both seasons, respectively (Table

2). Depending on the season, R-values are ranged from 0.75 to 0.97. RPD-values are between 1.4 and 4 for summer and global models and between 1.2 and 1.9 for spring models. Such lower RPD-value compared to RPD of summer and global models is mainly due to the smaller range of TA-value during spring model (41.21 – 51.05) compared to summer model (14.10 – 60.25). In the summer model, at maturation time (10 to 14 days before harvest) an important decrease in L-values from 50 to 14 units occurred in only 4 days. In the same maturation period, L-values of fruit from spring period remained in a steady state with a L-value at harvest in the vicinity of 40 units. Actual values as a function of the predicted values of the L* color parameter are presented in the Figure **5**.

The prediction of a* color parameter is important during the maturation stage since the red coloration of fruit skin occurs only during this stage. The accuracy of the prediction is correct for all the seasons during maturation stage ($R \geq 0.9$). RPD-values are ranged from 2 to 3 and RMSE is less than 2 and 5 during summer and spring, respectively.

Prediction of b* color parameter is depending on the maturity stage. During spring, best model was achieve with fruit before maturation while the best model during summer has been obtained with fruit during maturation. This makes sense because during the summer, the variations of b* occur during maturation while during spring (Figure **1**).

4. DISCUSSION

Since a decade, the promising results obtained with near-infrared spectroscopy in various applications of agricultural and food chain have led the manufacturers to move towards the development of portable and handheld NIR spectrometers. The development of these devices responds to a request of users of NIR spectroscopy who want to use this technique further upstream in the quality chain. Some perspectives aimed at the application at field with all the difficulties that entails.

Results about handheld NIR device remain scares in literature, often because the level of models accuracy is less promising than those reached with laboratory NIR devices. However, more results about handheld NIR technologies have to be published in order to inform the scientific community about the current development and potential of such recently developed devices.

Furthermore, the level of expected accuracy with such devices will be certainly ever lower compared to laboratory devices but the objectives and the reasons why an user could use such device are different or should be different. A handheld NIR spectrometer, when it is used at field or on-site, could be developed to give a first idea on a screening method based on quantitative or qualitative chemometric model. Of course, in some cases a precise quantification of a given quality trait should be confirm by a laboratory measurement.

In the present study, a follow-up of fruit quality during spring and summer seasons was perform on tomato growing in greenhouse. The measured variability was used to build chemometric models based on NIR spectroscopy. The models were qualitative and quantitative. The qualitative models showed that NIR spectroscopy could be an promising tool able to follow the fruit development during spring and summer seasons. The fruit development was significantly shorter during summer compared to spring, about 10 days. In this way, carried out 2 discriminant (FDA) models was necessary to follow the fruit growing at the two seasons.

The quantitative models allowed to reach an first estimation of qualitative traits of tomato. The models were particularly accurate for color prediction (L, a* and b*) and SSC but less accurate for TA.

In the present study we decided to take into account the two main development stages of tomato that correspond to the time before maturation (BM) and the maturation time (M). In most cases, PLS models built separately for BM and M fruit were more accurate than those performed with BM and M fruit gathered in a single data set. Some studies presented the promising potential of portable NIR spectroscopy to predict tomato quality during maturation stage [14]. The study of Ecarnot *et al*. allowed the phenotyping of tomato subjected to a breeding program and analyze the fruit quality in the last days of its maturation. In view of the scientific literature, no study has used the inherent variability of tomato fruit growth and development to build chemometric models based on near infrared spectroscopy.

Our results showed that the fruit growth and quality traits in early physiological stages (before maturation) could be followed by handheld NIR spectroscopy. The maturation stage was quite similar at the two seasons in terms of duration. Our study showed that the

environmental factors characterizing the two seasons mainly act during the period before maturation, certainly during cell expansion stage. In this way, it is more difficult to build a single PLS model for both seasons concerning the BM fruit.

In the present study, two parameters were calculated to evaluate the potential use of the models, the RPD and RPIQ. RPD was developed in 1987 to assess the performance of predictions models [26]. However, this parameter requires that the analysed data are normally distributed, which is not always the case. Typically, in our study some quality traits (ex. Colour) change brutally after a long period of steady state, such parameter is not normally distributed. In order to cope with this problem, the RPD could be replaced by the RPIQ parameter that take into account the variability between the 1st and 3rd quartiles to be compared to root mean square error of the model [20]. The models built in the present study reached RPIQ values allowing a first screening of tomato. Concerning SSC, RPIQ values were comprised between 2 and 4.3. RPIQ of TA were quite good during summer (more than 4.46) but very low during spring (less than 2.0). RPIQ of L* parameter were higher than 4.5 in a model gathering spring and summer fruit but with separating BM and M fruit. Finally, a* parameter is important because it characterize the red coloration of tomato skin during maturation. RPIQ of a* parameter was particularly high (between 2.67 and 5.29) for M fruit.

This study takes into account the maturity of the fruit and the season as a factor influencing the construction of predictive models using NIR spectroscopy. Other factors may be taken into account through further studies.

REFERENCES

[1] Christen D, Camps C, Summermatter A, Gabioud Rebeaud S, Baumgartner D. Prediction of the pre- and postharvest apricot quality with different VIS/NIRs devices. Acta Horticulturae 2012; 966: 149-54.

[2] Camps C, Simone C, Gilli C. Assessment of tomato quality using portable NIR spectroscopy and PLSR with wavelengths selection. Acta Horticulturae 2012; 936: 437-42.

[3] Camps C, Christen D. On-tree follow-up of apricot fruit development using a hand-held NIR instrument. Journal of Food, Agriculture and Environment 2009; 7(2): 394-400.

[4] Camps C, Christen D. Non-destructive assessment of apricot fruit quality by portable visible-near infrared spectroscopy. LWT - Food Science and Technology 2009; 42(6): 1125-31. http://dx.doi.org/10.1016/j.lwt.2009.01.015

[5] Camps C, Guillermin P, Mauget JC, Bertrand D. Discrimination of storage duration of apples stored in a cooled room and shelf-life by visible-near infrared spectroscopy. Journal of near Infrared Spectroscopy 2007; 15(3): 169-77. http://dx.doi.org/10.1255/jnirs.726

[6] Baranska M, Schütze W, Schulz H. Determination of lycopene and beta-carotene content in tomato fruits and related products: Comparison of FT-Raman, ATR-IR, and NIR spectroscopy. Analytical Chemistry 2006; 78(24): 8456-61. http://dx.doi.org/10.1021/ac061220j

[7] Clement A, Dorais M, Vernon M. Nondestructive Measurement of Fresh Tomato Lycopene Content and Other Physicochemical Characteristics Using Visible-NIR Spectroscopy. Journal of Agricultural and Food Chemistry 2008; 56(21): 9813-8. http://dx.doi.org/10.1021/jf801299r

[8] Clement A, Dorais M, Vernon M. Multivariate approach to the measurement of tomato maturity and gustatory attributes and their rapid assessment by Vis-NIR Spectroscopy. Journal of Agricultural and Food Chemistry 2008; 56(5): 1538-44. http://dx.doi.org/10.1021/jf072182n

[9] De Nardo T, Shiroma-Kian C, Halim Y, Francis D, Rodriguez-Saona LE. Rapid and Simultaneous Determination of Lycopene and beta-Carotene Contents in Tomato Juice by Infrared Spectroscopy. Journal of Agricultural and Food Chemistry 2009; 57(4): 1105-12. http://dx.doi.org/10.1021/jf802920z

[10] Flores K, Sanchez MT, Perez-Marin D, Guerrero JE, Garrido-Varo A. Feasibility in NIRS instruments for predicting internal quality in intact tomato. J Food Eng 2009; 91(2): 311-8. http://dx.doi.org/10.1016/j.jfoodeng.2008.09.013

[11] Kusumiyati A, Akinaga T, Tanaka M, Kawasaki S. On-tree and after-harvesting evaluation of firmness, color and lycopene content of tomato fruit using portable NIR spectroscopy. Journal of Food Agriculture & Environment 2008; 6(2): 327-32.

[12] Pedro AMK, Ferreira MMC. Nondestructive determination of solids and carotenoids in tomato products by near-infrared spectroscopy and multivariate calibration. Analytical Chemistry 2005; 77(8): 2505-11. http://dx.doi.org/10.1021/ac048651r

[13] Pedro AMK, Ferreira MMC. Simultaneously calibrating solids, sugars and acidity of tomato products using PLS2 and NIR spectroscopy. Analytica Chimica Acta 2007; 595(1-2): 221-7. http://dx.doi.org/10.1016/j.aca.2007.03.036

[14] Ecarnot M, Baczyk P, Tessarotto L, Chervin C. Rapid phenotyping of the tomato fruit model, Micro-Tom, with a portable VIS-NIR spectrometer. Plant Physiol Biochem 2013; 70: 159-63. http://dx.doi.org/10.1016/j.plaphy.2013.05.019

[15] Barnes RJ, Dhanoa MS, Lister SJ. Standard normal variate transformation and de-treding of near-infrared diffuse reflectance spectra. Applied Spectroscopy 1989; 43(5): 772-7. http://dx.doi.org/10.1366/0003702894202201

[16] Moons E, Sinnaeve G. non destructive Vis and NIR spectroscopy measurement for the determination of apple internal quality. Acta Hort 2000; 517: 441-8.

[17] Bertrand D, Courcoux P, Autran JC, Meritan R, Robert P. Stepwise canonical discriminant-analysis of continuous digitized signals - application to chromatograms of wheat proteins. Journal of Chemometrics 1990; 4(6): 413-27. http://dx.doi.org/10.1002/cem.1180040605

[18] Roger JM, Palagos B, D. B, E. F-A. CovSel: Variable selection for highly multivariate and multi-response calibration: Application to IR spectroscopy. Chemometrics and Intelligent Laboratory Systems 2011; 106: 216-23. http://dx.doi.org/10.1016/j.chemolab.2010.10.003

[19] Williams P, Sobering D. Comparison of commercial near infrared transmittance and reflectance instruments for analysis of whole grains and seeds. Journal of Near Infrared Spectroscopy 1993; 1(1): 25-32. http://dx.doi.org/10.1255/jnirs.3

[20] Bellon-Maurel V, Fernandez-Ahumada E, Palagos B, Roger JM, McBratney A. 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 2010; 29(9): 1073-81. http://dx.doi.org/10.1016/j.trac.2010.05.006

[21] Giovannoni JJ. Genetic regulation of fruit development and ripening. Plant Cell 2004; 16: S170-S80. http://dx.doi.org/10.1105/tpc.019158

[22] Beckles DM, Hong N, Stamova L, Luengwilai K. Biochemical factors contributing to tomato fruit sugar content: a review. Fruits 2012; 67(1): 49-64. http://dx.doi.org/10.1051/fruits/2011066

[23] Centeno DC, Osorio S, Nunes-Nesi A, Bertolo ALF, Carneiro RT, Araujo WL, *et al*. Malate Plays a Crucial Role in Starch Metabolism, Ripening, and Soluble Solid Content of Tomato Fruit and Affects Postharvest Softening. Plant Cell 2011; 23(1): 162-84. http://dx.doi.org/10.1105/tpc.109.072231

[24] Luengwilai K, Beckles DM. Starch Granules in Tomato Fruit Show a Complex Pattern of Degradation. J Agric Food Chem 2009; 57(18): 8480-7. http://dx.doi.org/10.1021/jf901593m

[25] Petreikov M, Yeselson L, Shen S, Levin I, Schaffer AA, Efrati A, *et al*. Carbohydrate Balance and Accumulation during Development of Near-isogenic Tomato Lines Differing in the AGPase-L1 Allele. J Am Soc Hortic Sci 2009; 134(1): 134-40.

[26] Williams PC. Variables affecting near-infrared reflectance spectroscopic analysis. Near-infrared Technology in the Agricultural and Food Industries: Williams P, Norris K (eds). American Association of Cereal Chemists: St. Paul, MN; 1987. p. 143–67.

Received on 10-11-2014 Accepted on 18-11-2014 Published on 27-11-2014

DOI: http://dx.doi.org/10.15377/2409-9813.2014.01.01.4