Received: 3 October 2017 | Revised: 1 December 2017 | Accepted: 13 December 2017
DOI: 10.1111/jtxs.12316

RESEARCH ARTICLE

WILEY Journal of Texture Studies

Singular approach to penetrometry by preprocessing of digitized force–displacement curves and chemometry: A case study of 12 tomato varieties

Cédric Camps

Agroscope, Institute for Plant Production Sciences IPS, Conthey, Switzerland

Correspondence
Cédric Camps, Agroscope, Institute for Plant Production Sciences IPS, Route des Eterpys 18, 1964 Conthey, Switzerland.
Email: cedric.camps@agroscope.admin.ch

Abstract

This study evaluated a singular approach to the analysis of digitized force–displacement curves from penetrometry performed on tomatoes. Penetrometry is commonly used to evaluate the texture properties of fresh fruits and vegetables. Currently, the parameters are computed from force–displacement curves. The parameters are statistically analyzed to highlight differences, for instance, between fruit varieties or differentially stored fruits. In this study, digitized force–displacement curves were considered "mechanical imprints" (MIs). Twelve varieties of tomato were analyzed, and the assumption that a given variety is characterized by a singular MI was tested. The digitized curves were preprocessed through standardization and smoothing. This preprocessing enabled the classification of more than 94% of fruit according to variety compared with 45% of fruit without preprocessing. To compare this approach with a classical approach of texture analysis, 13 parameters were computed from the force–displacement curves and analyzed. This parameters-based approach enabled the classification of approximately 46% of fruit according to variety. The digitized curve enabled the correct classification of 88% more fruit than the level achieved by the parameters-based approach. Thus, penetrometry analysis presents new opportunities, particularly for breeding programs aimed at improving the texture characteristics of fruits and vegetables. Additionally, this approach could be adapted to other mechanical tests for the characterization of the textural properties of food products.

Practical applications

The force–displacement curve of penetrometry can be considered a "mechanical imprint" to improve fruit classification according to preharvest and postharvest characteristics (i.e., maturity, variety, etc.). This mechanical imprint represents a new and discriminant phenotyping criterion for tomato breeding to improve texture. The use of the digitized curve rather than computed parameters may be more easily implemented in industry.

KEYWORDS
classification, factorial discriminant analysis, mechanical imprint, preprocessing curves

1 | INTRODUCTION

Assessing texture characteristics based on the mechanical properties of fresh products from agriculture, such as tomatoes, is a key process in the overall value chain. Indeed, the textural properties of fresh fruits are important to evaluating the postharvest life of a given fruit variety and to satisfy the consumer's preferences.

Despite the recognized importance of such characteristics, texture has rarely been studied and has not been included in the value chain of fresh fruits and vegetables. The primary reasons for the lack of information concerning texture likely include (a) the lack of reference methods, (b) the limited information obtained on existing methods applied

This article was published on AA publication on: 20 December 2017.

to fruits and vegetables, and (c) the slow advance in research on the texture analysis of fresh fruits and vegetables over the last decade.

A parameter called "firmness" defines the texture properties of various fruits and remains a unique measurement of the mechanical properties of fresh products. Operators in U.S. and European markets require threshold values of firmness, making it possible to grade fruits according to their maturity.

For fresh tomatoes, textural properties play an important role in the appreciation and acceptance of these products by consumers (Causse, Buret, Robini, & Verschave, 2003; Serrano-Megías & López-Nicolás, 2006). Thus, molecular analyses have been initiated to identify chromosome regions implicated in tomato texture (Causse et al., 2002).

These properties depend on the different characteristics of fruit cells or tissue turgidity, cell size, wall composition, and integrity (Chaïb et al., 2007; Lahaye, Devaux, Poole, Seymour, & Causse, 2013). Sensory analysis is the most direct method for accessing the textural properties of fruits and vegetables and to simultaneously consider the effects of all involved cellular components. However, such analyses are difficult to implement on a large series of samples. Therefore, mechanical properties based on instrumental approaches remain the easiest method to evaluate textural properties.

Important studies using instrumental approaches for texture analysis were performed on apples in the 1990s (Chen, Duprat, Grotte, Loonis, & Pietri, 1995; Duprat, Grotte, Pietri, & Studman, 1995; Roudot, Duprat, & Wenian, 1991). In the early 2000s, several studies using the information extracted from force–displacement curves of penetrometry and compression tests to characterize and follow the texture of fruits were reported. Mehinagic, Royer, Symoneaux, Bertrand, and Jourjon (2004) computed parameters from penetrometry to link the mechanical properties of apples to sensory perception obtained using a panel trained in sensory analysis. Camps, Guillermin, Mauget, and Bertrand (2005) used similar parameters to examine apple texture changes during postharvest storage.

Information extracted from a force–displacement curve could represent another approach if we consider the whole digitized curve as potentially pertinent for discriminating fruits with different textures. Camps et al. (2005) attempted the first chemometric approach on whole apple fruit. In their study, raw data curves were analyzed by factorial discriminant analysis (FDA) to examine fruit texture changes during storage. The results showed that certain regions of the digitized curves, not exploited by classical parameter extractions, were discriminant.

In this study, we similarly considered the force–displacement curves from penetrometry as "signals" for pretreatment, followed by analysis using chemometric methods. The hypothesis of this approach is that the entire digitized curve can act as a "mechanical imprint" (MI), characterizing fruit texture at a given time. In addition, we improved the classical approach of force–displacement curve analysis by including more parameters than were examined in previous studies. The aim of this study was to compare the discriminative power of the two approaches in the case study applied to 12 tomato varieties. Indicators such as the number of factors considered were taken into account since the two approaches are characterized by data matrices of different sizes.

2 | MATERIALS AND METHODS

2.1 | Tomato varieties

Twelve round-type tomato varieties were used in this study: Admiro (De Ruiter), Cindel (Enza Zaden), Cristal (Clause), Estiva (Gauthier), Fiorentino (Enza Zaden), Gloriette (Rijk Zwaan), Megaline (Syngenta), Natyssa (Gauthier), Octydia (Gauthier), Paola (Clause), Paronset (Syngenta), and Pilu (Piluweri).

The tomatoes were grown in a Venlo-type greenhouse covering a 360 m^2 area of the Agroscope research station (CH-1964 Conthey, Switzerland). In the greenhouse, 3 replicates of 10 tomato plants (2 stems per plant) per variety were planted. A total of 360 fruits (30 tomatoes per variety) were harvested for use in this study. Tomatoes have been harvested at commercial maturity corresponding to the "light-red" stage (Sargent & Moretti, 2002).

2.2 | Penetrometry

Penetrometry was performed on the equatorial side of the tomato fruit. Two measurements were performed on each fruit by varying the orientation at intervals of approximately 90°. A TA-XTplus Texture Analyzer (Stable Microsystems, Surrey, U.K.) fitted with a needle probe 2 mm in diameter was used. The probe was moved from the surface of the fruit to a final depth of 5 mm at a speed of 10 mm/s, and the force (expressed in Newtons) was recorded for every step displacement. A trigger force of 5 g has been automatically applied by the TA-XTplus. Each digitized force–displacement curve included 250 data points (Figure 1).

2.2.1 | Computing parameters from the force–displacement curves

The force–displacement curve was divided into three parts along the probe displacement in the tomato fruit. The first part, from the beginning (0 mm) until the tomato skin failed, was used to characterize the skin strength. The second part (II) was defined from the failure point until the relative lowest force value, and the third part (III) extended from the end of the second part to the end of the probe displacement (5 mm). The third part indicated the mechanical properties of the flesh under the skin.

Thirteen numerical values, called texture parameters, were computed from the force–displacement curve (Figure 1). The maximal force F_P (N) represents the force required to puncture the tomato skin. F_P represents the skin strength. The displacement D_P (mm) indicates the probe position at F_P. This parameter allows the calculation of fruit deformation before the skin rupture. E_P (N/mm), also called stiffness, was defined as (F_P/D_P) and represents the slope of the first part of the curve (I) measured from the beginning of the acquisition until F_P is reached. W_P (N.mm) is the mechanical work required to reach the rupture point and is estimated by the area under the curve up to the skin rupture point (F_P, D_P). In the first part of the curve, a parameter called PK (no unit) was computed, and this parameter corresponds to the number of peaks (or fracture events) measured before the skin failed.

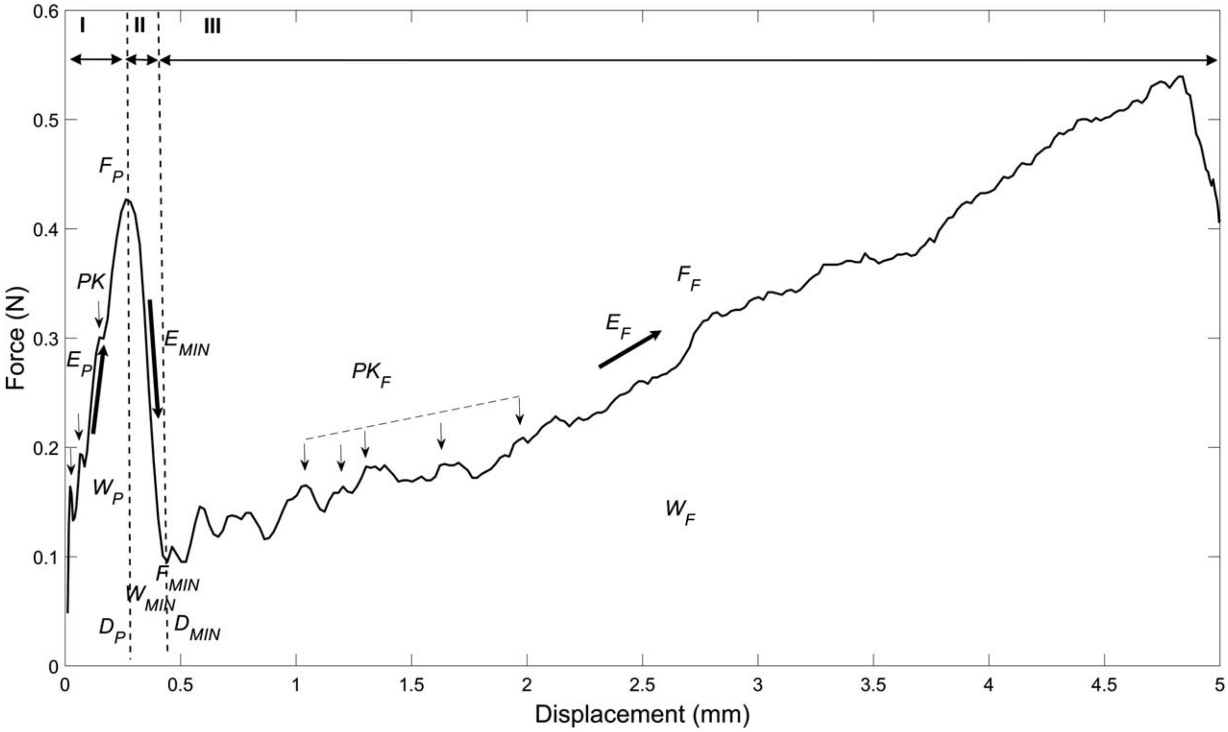

FIGURE 1 Parameters extracted from the force–displacement curves of a penetrometric test. I, II, and III represent the three steps of the test (I: compression phase, II: depression phase, III: shearing phase)

In the second part (II), the slope E_{MIN} (N/mm), mechanical work W_{MIN} (N.mm), displacement D_{MIN} (mm), and minimal force F_{MIN} (N) were computed. This part of the curve indicates the mechanical properties of exocarp (outer part of the pericarp).

The third part of the curve gathers four parameters: the average flesh firmness F_F (N), measuring the force through the mesocarp and endocarp of the fruit, the mechanical W_F (N.mm) required for the probe displacement through the flesh and the slope E_F (N/mm) during this part of the curve. Similar to the first part of the curve, the number of peaks PK_F (no unit) occurring during part III was computed.

The texture parameters were automatically computed from each curve using specific software (Texture Exponent 32, Stable Microsystems, U.K.) and a specific macro written by our research team.

2.2.2 | Analysis of raw digitized curves

Force–displacement curves collection

A given force–displacement curve is represented as a vector $\boldsymbol{y}_{(1,p)}$, where $\boldsymbol{p}$ is the number of displacement steps recorded. All $\boldsymbol{y}$ vectors were gathered in a matrix $\mathbf{X}_{(n,p)}$, where n is the number of penetrometric measurements. Analysis of raw digitized curves was performed on the **X** matrix.

Preprocessing of raw digitized curves

Three preprocessing steps were performed on the digitized force–displacement curves (standard normal variate, normalization, and smoothing).

SNV Standard normal variate (SNV) is commonly used in signal processing in spectroscopy to reduce the scatter deformation of the spectra (Barnes, Dhanoa, & Lister, 1989). Thus, SNV standardizes the rows of the **X** data matrix. SNV will be applied to raw digitized curves of penetrometry to minimize the manipulator effect. The manipulator holds the tomato in hand during the measurement of penetrometry and may move during measurement. Such movement can result in an artificial variation of force intensities at the beginning and during of the force–displacement curve recording.

Normalization The digitized curves are normalized. Normalization involves dividing each column of **X** by the corresponding standard deviation. A typical force–displacement curve of penetrometry (Figure 1) presents high values (part I of the curve) and lower values (part III of the curve). Normalizing each column of the **X** matrix enables an objective comparison of the significance of all measured forces in all displacement steps.

Smoothing Finally, the digitized curves are smoothed using "moving average filtering" to minimize the micro-events in curves, potentially reflecting actual micro-texture variations, but often result from fruit instability during needle probe displacement. The smoothing procedure requires the determination of an optimal span number. To determine the optimal span number, a FDA-based procedure is proposed in the results.

2.3 | Data analyses

The data were analyzed using modified FDAs according to Bertrand, Courcoux, Autran, and Méritan (1990). FDAs were performed on raw and preprocessed digitized curves and the texture parameters. In the

TABLE 1 Mean values of parameters computed from force-displacement curves of penetrometry.

Parameters	F_P	E_P	W_P	D_P	PK	F_{MIN}	E_{MIN}	W_{MIN}	D_{MIN}	PK_F	F_F	W_F	E_F
Varieties													
ADMIRO	0.46 a	1.67 cd	0.05 a	0.29 a	2.40 b	0.10 ab	−2.1 f	0.04 ab	0.10 ab	28.9 cd	0.38 bcd	1.72 c	0.07 bc
CINDEL	0.59 d	1.50 b	0.10 c	0.40 de	2.40 b	0.10 a	−2.9 b	0.05 de	0.10 a	27.0 bc	0.40 d	1.77 c	0.09 cd
CRISTAL	0.54 c	1.31 a	0.10 c	0.43 ef	3.00 c	0.10 ab	−2.5 cd	0.04 cde	0.10 ab	32.1 d	0.34 bc	1.49 b	0.070 b
ESTIVA	0.47 a	1.32 a	0.08 b	0.38 cd	3.17 cd	0.11 ab	−2.2 ef	0.04 a	0.11 ab	30.7 d	0.38 cd	1.71 c	0.08 c
FIORENTINO	0.49 ab	1.86 e	0.05 a	0.26 a	2.07 ab	0.12 b	−2.2 f	0.04 abc	0.12 b	21.5 a	0.47 ef	2.16 de	0.10 de
GLORIETTE	0.52 bc	1.55 bc	0.07 b	0.34 bc	2.53 b	0.12 b	−2.3 def	0.04 cde	0.12 b	24.8 ab	0.48 ef	2.15 de	0.12 e
MEGALINE	0.48 a	1.67 cd	0.06 a	0.30 ab	2.52 b	0.10 ab	−2.2 def	0.04 ab	0.10 ab	28.8 bcd	0.38 cd	1.75 c	0.08 bc
NATYSSA	0.55 c	2.03 f	0.06 a	0.26 a	1.83 a	0.12 b	−2.6 bc	0.04 cde	0.12 b	21.7 a	0.52 f	2.37 e	0.12 e
OCTYDIA	0.48 a	1.75 de	0.05 a	0.28 a	2.09 ab	0.11 ab	−2.2 ef	0.04 a	0.11 ab	26.7 bc	0.40 d	1.82 c	0.08 bc
PAOLA	0.55 c	1.310 a	0.10 c	0.43 f	3.00 c	0.09 a	−2.5 cde	0.04 cde	0.09 a	32.0 d	0.33 b	1.46 b	0.07 b
PARONSET	0.49 ab	1.73 de	0.06 a	0.29 a	2.36 ab	0.10 a	−2.2 ef	0.04 bcd	0.10 a	27.5 bcd	0.42 de	1.90 cd	0.09 cd
PILU	0.63e	1.23 a	0.14 d	0.53 g	3.50 d	0.09 a	−3.2 a	0.05 e	0.09 a	38.3 e	0.26 a	1.13 a	0.05 a
p value	< 0.0001	< 0.0001	< 0.0001	< 0.0001	< 0.0001	0.006	< 0.0001	< 0.0001	0.006	< 0.0001	< 0.0001	< 0.0001	< 0.0001
F value	19.9	25.6	38.0	32.7	9.0	2.5	11.4	4.8	2.5	11.6	14.5	16.2	13.6

Note. Analysis of variance performed at a threshold of p = .05. Mean-values have been compared by a Fisher (LSD) test at a threshold of 5%.

FDAs, the qualitative groups for discrimination were the 12 tomato varieties. A criterion of the efficiency of the FDAs was the proportion of correctly identified observations. The confusion matrices were analyzed to evaluate the misclassified individuals. FDA computes a few sets of discriminant scores, which are linear combinations of the original variables. The correlation between the discriminant scores and the predictive variables was analyzed. For this purpose, the correlation coefficients between the discriminant scores and the original variables (13 parameters or 250 digitized data points of the digitized curve) were computed. All statistical procedures were performed using the Matlab environment (The MathWorks, Inc., 3 Apple Hill Drive, Natick, MA).

3 | RESULTS AND DISCUSSION

3.1 | Analysis of parameters approach

The values of the 13 parameters extracted from the digitized curves of penetrometry are presented in Table 1. Each parameter presented significant differences between varieties. Two parameters, F_{MIN} and D_{MIN} (phase II of the curve), were less discriminant than the other 11 parameters. To better understand the discriminant power of the 13 parameters, a FDA of the recorded values was performed.

A total of 12 factors were introduced in the analysis. These parameters were determined by counting the fruits correctly classified using FDA when the maximal number of factors is introduced (i.e., 13 factors because we have 13 parameters). An improvement in the number of correct classified fruits was observed until reaching 12 factors. Approximately 50% of tomato fruits were correctly identified for a given variety (Table 2). After the 12th factor, the misclassification rate increased. This approach was performed using a cross-validation procedure, with 2/3 of the fruit as the calibration set and the remaining 1/3 of the fruit as the validation set. Thus, the discriminant power of the parameter approach was evaluated using 12 factors in the FDA.

Figure 2 shows the factorial map according to the first two factorial scores. The projection of the factorial scores on the first two axes describes a continuous parametric arc in which the varieties are

TABLE 2 Confusion matrix from FDA performed on parameters computed from penetrometric curves.

	Predicted groups											
	ADMIRO (%)	CINDEL (%)	CRISTAL (%)	ESTIVA (%)	FIORENTINO (%)	GLORIETTE (%)	MEGALINE (%)	NATYSSA (%)	OCTYDIA (%)	PAOLA (%)	PARONSET (%)	PILU (%)
Actual groups												
ADMIRO	40	–	4	4	4	4	8	4	16	–	16	–
CINDEL	–	73	3	3	–	7	–	–	–	3	3	7
CRISTAL	7	17	30	17	–	3	–	–	–	17	–	10
ESTIVA	4	–	14	61	–	18	–	–	–	–	4	–
FIOREN	7	–	–	7	57	–	–	14	7	–	7	–
GLORIETTE	–	13	7	10	–	63	3	–	–	3	–	–
MEGALINE	18	–	6	12	6	6	12	–	24	6	12	–
NATYSSA	–	–	–	6	17	–	–	56	22	–	–	–
OCTYDIA	19	–	–	5	5	–	10	5	29	5	24	–
PAOLA	3	17	20	10	–	–	3	–	–	37	7	3
PARONSET	18	9	9	9	–	–	–	27	9	–	18	–
PILU	–	3	7	–	–	–	–	–	–	13	–	77

Note. Actual: real variety tested in the model. Predicted: prediction of correct classified tomato (in percentage) by the model FDA according to the variety.

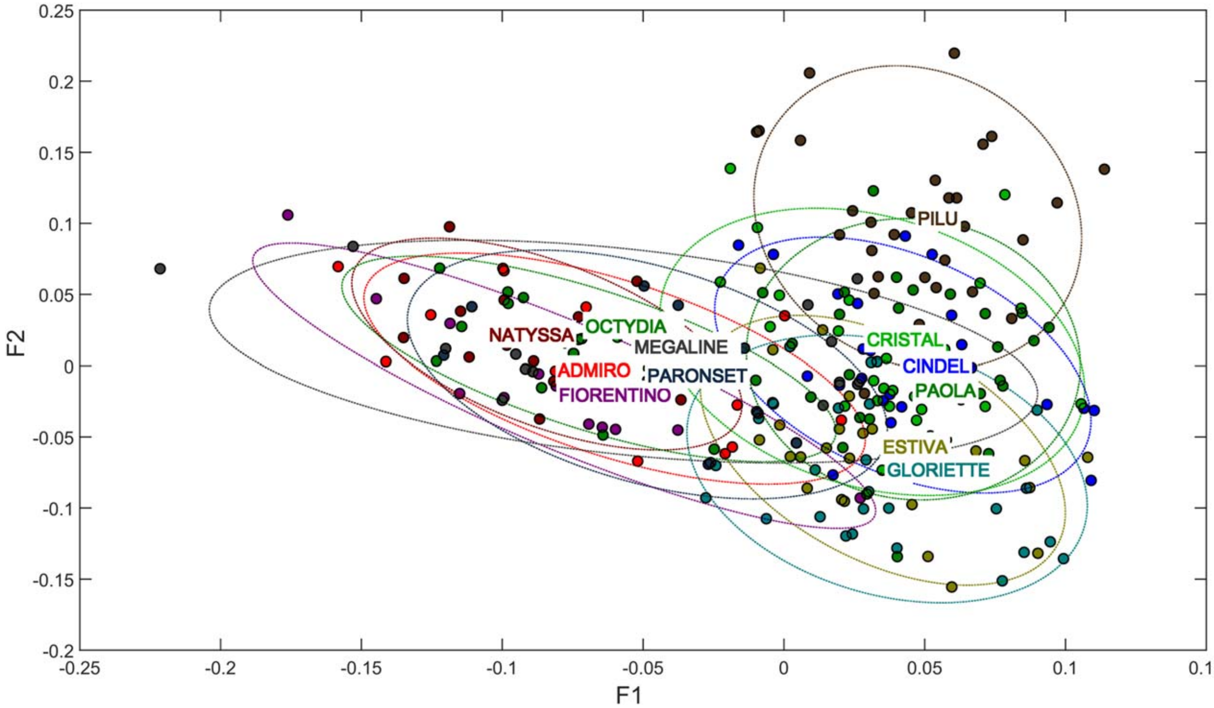

FIGURE 2 Factorial map according to the first two factorial scores of FDA. Ellipses represent the confidence intervals at a threshold of 5%

classified. In this classification, the Natyssa and Pilu varieties presented the most important differences with respect to texture. Approximately 50% of the fruits were correctly classified. However, differences among varieties were observed. More than 60% of 4 varieties (Cindel, Estiva, Gloriette, and Pilu) was correctly identified, while less than 30% of 4 other varieties (Cristal, Megaline, Octydia, and Paronset) was correctly identified. These results could highlight both a high variability of the textural properties inside a given variety but also some similarities between varieties.

The correlation between the factorial scores of the FDA and values of extracted parameters enabled the analysis of the relative significance of the effect of the parameters in the discriminant power of the FDA. In this study, 3 parameters—E_P, W_P, and D_P—showed R values higher than .75. The parameters were extracted in phase I of the force–displacement curve. With R values from −.4 to .2, the parameters from phase II of the curve were less significant. In the third phase of the curve, the most significant parameters were F_F and W_F (R value = .5 to .6).

In this approach, phase I and, to a lesser extent, phase III are the most important to classify tomatoes according to variety. In this study, phase II did not present any interest. A total of 12 tomato varieties of only one type, the round type tomato, were compared in this study. All varieties presented a similar appearance (size, form, and color). According to these results, the most significant differences between varieties appeared in the textural properties of the outer part of the fruit, that is, the peel and outer pericarp.

Studies showed that a correlation exist between firmness and tomato caliber. Serrano-Megías and López-Nicolás (2006) obtained a significant correlation (R = .72) between firmness and tomato caliber. However, in their study, the authors measured the firmness with a digital HPE (Shore HPE-A/HPE-P, DGM 93 18 389.5, Borås, Sweden). Such a device is commonly used to assess tomato firmness but is aims at measuring the fruit elasticity since is do not penetrate the fruit. With such device, a compression is applied to the fruit surface, a phenomenon that involves a larger part of the fruit tissues compared to a penetrometer that has a more local scope. To be quite precise, the correlations between the texture parameters measured in this study and fruit size have been calculated. This, to be sure that the variability of the texture parameters does not reveal a simple difference in size. The obtained R values ranged from −.32 to .25 for all texture parameters. The absence of correlation means the variability of parameter-values characterizes differences in mechanical properties of the skin of tomatoes and not differences related to the caliber. The absence of correlation is probably due to the used-probe which was a needle with pointed head, limiting the compression phenomenon at the impact point before to the penetration.

3.2 | Analyses of the digitized curves

In the following approach, the force–displacement curves from penetrometry were gathered in a rectangular matrix **X**, where the columns are the steps of displacement recorded during the mechanical test and the rows are the individuals encoded according to the tomato varieties. A total of 360 force–displacement curves were collected (Figure 3a). The means of the raw data curves highlight the differences in texture among varieties (Figure 3b).

As described in the Materials and methods, the raw data obtained from the force–displacement curves were standardized (lines of **X**) and normalized (columns of **X**) to minimize the "micro-events" resulting from (a) handling during the implementation of the

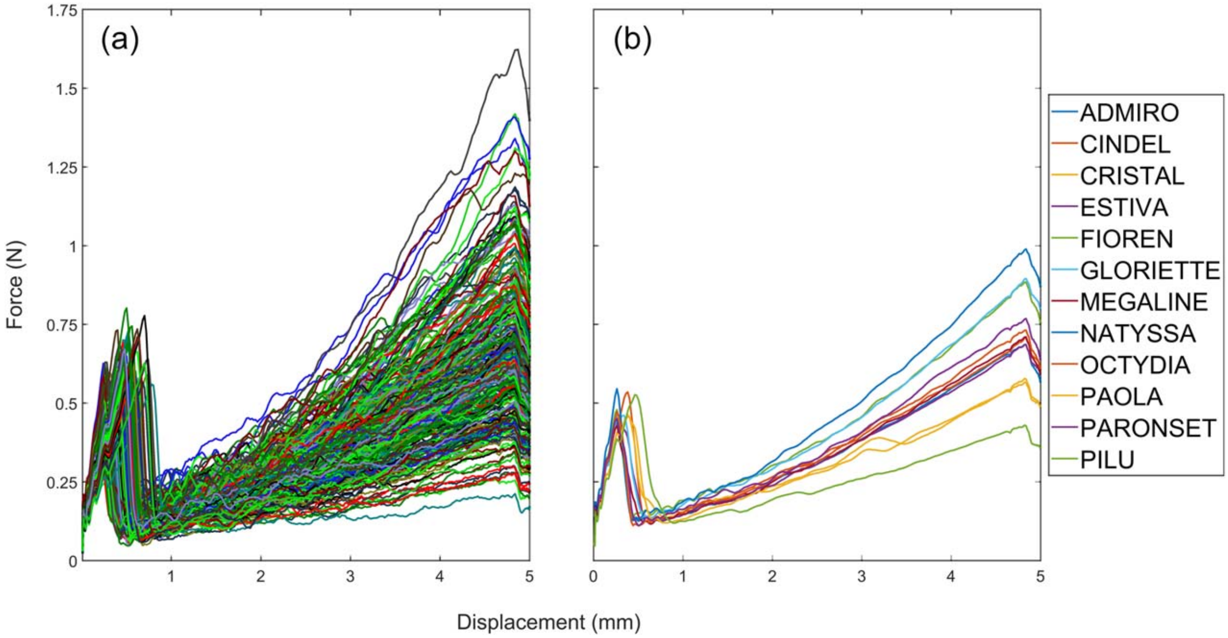

FIGURE 3 Force–displacement curves from penetrometry of the 12 tomato varieties (a) and average curves per variety (b)

mechanical test and (b) the difference in scale values among the three phases of the force–displacement curve. Indeed, the forces recorded during the first phase and at the end of the third phase are characterized by high values compared to the values recorded in other parts of the curve.

3.3 | Smoothing by moving average filtering

A third preprocessing step was attempted to optimize the discrimination power of the digitized force–displacement curves. This procedure involved smoothing the curves (Equation 1) using an optimal span value.

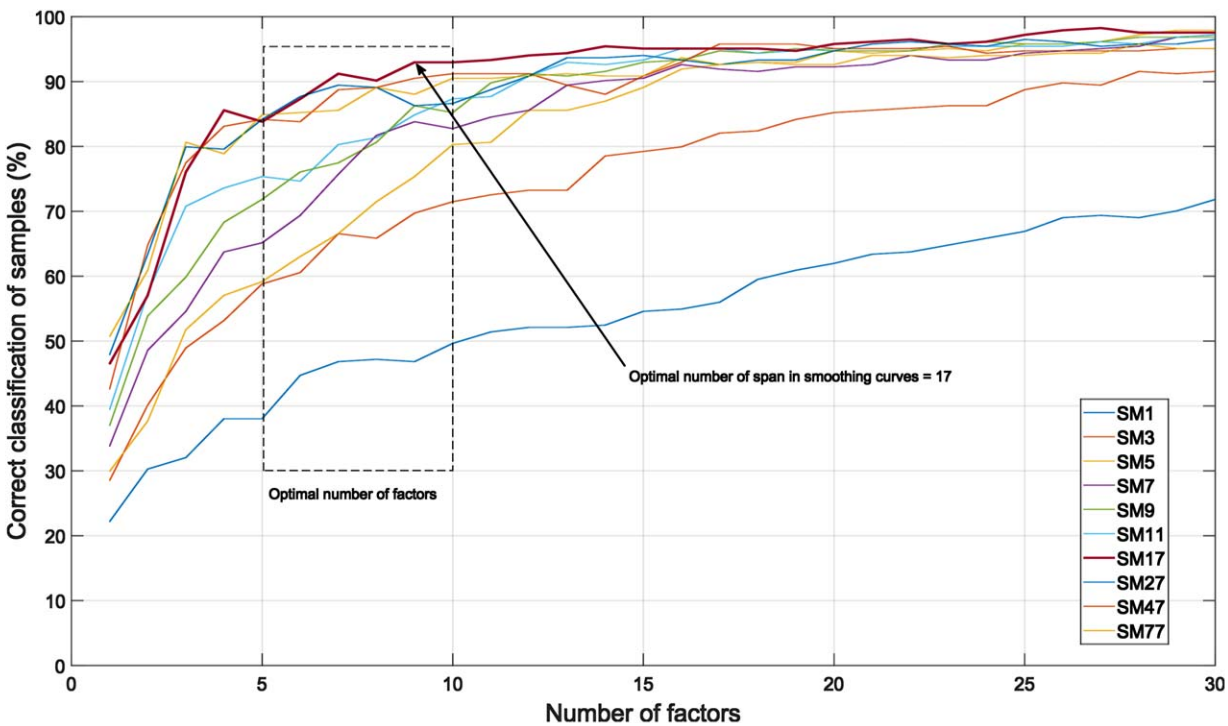

FIGURE 4 Determination of the optimal span value for smoothing force–displacement curves of penetrometry. The plot shows the percentage of samples correctly classified by FDA using as a function of the factors introduced (30 factors). SM is the span number in smoothing curves

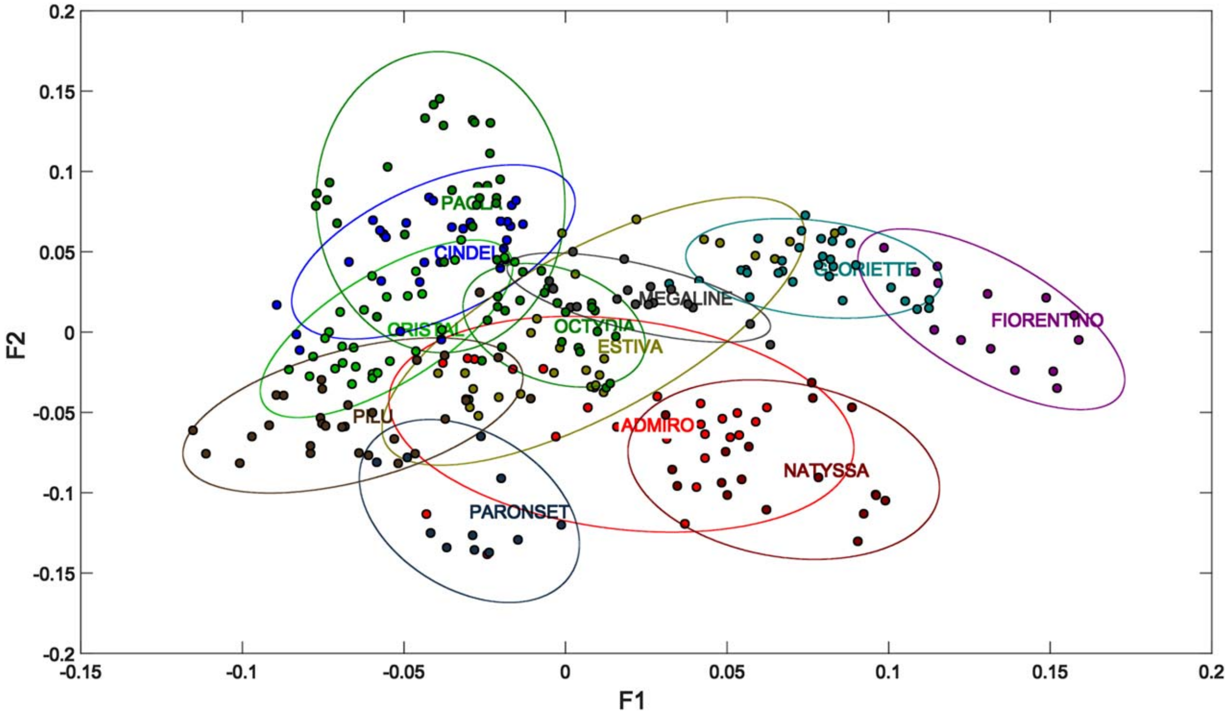

FIGURE 5 Factorial map according the first two factorial scores of the FDA. Ellipses represent the confidence interval at a threshold of 5%

$$y_s(i) = \frac{1}{2K+1}\left(y(i+K)+y(i+K-1)+\ldots+y(i-K)\right) \quad (1)$$

where $y_s(i)$ is the smoothed value for the ith displacement step of the curve, K is the number of neighboring data points on either side of $y_s(i)$, and $2K+1$ is the span.

To determine the optimal span value, a FDA was performed on **X** to discriminate the tomato varieties. The discriminant analysis was voluntarily performed using all factors available (i.e., 250), leading to an over-fitted model. This model provided the information for the optimal span number to smooth the curves and the optimal number of factors introduced in the final discriminant model. Figure 4 presents the correct classification obtained by the factorial analysis (y axis) as a function of the number of factors introduced (x axis) and the span number (plotted curves). First, we determined the optimal number of factors in the vicinity of 10, where an inflexion of the curve occurred. Next, an observation of the number of correctly classified individuals according to the span number in this region (approximately 10 factors) enabled the determination of the optimal span number. An increase of correct classification occurred until a span value of 17 and subsequently decreased for higher values of span (i.e., 27, 47, and 77).

Thus, the discriminant power of the digitized curve approach was evaluated using force–displacement curves standardized and smoothed with a span number of 17 and processed in a FDA using 10 factors.

TABLE 3 Confusion matrix from FDA performed on pre-processed force-displacement digitized curves.

Predicted groups												
	ADMIRO (%)	CINDEL (%)	CRISTAL (%)	ESTIVA (%)	FIORENTINO (%)	GLORIETTE (%)	MEGALINE (%)	NATYSSA (%)	OCTYDIA (%)	PAOLA (%)	PARONSET (%)	PILU (%)
Actual groups												
ADMIRO	84	–	–	–	–	–	4	4	8	–	–	–
CINDEL	–	87	–	7	–	–	–	–	7	–	–	–
CRISTAL	–	–	93	3	–	–	–	–	3	–	–	–
ESTIVA	–	–	–	96	4	–	–	–	–	–	–	–
FIOREN	–	–	–	–	100	–	–	–	–	–	–	–
GLORIETTE	–	–	–	–	7	93	–	–	–	–	–	–
MEGALINE	–	–	–	–	–	–	100	–	–	–	–	–
NATYSSA	–	–	–	–	–	–	–	100	–	–	–	–
OCTYDIA	–	–	–	–	–	–	–	–	100	–	–	–
PAOLA	–	–	–	–	–	–	–	–	3	80	7	–
PARONSET	–	–	–	–	–	–	–	–	–	–	91	9
PILU												100

Note. Actual: real variety tested in the model. Predicted: prediction of correct classified tomato (in percentage) by the model FDA according to the variety.

Figure 5 shows the factorial map according to the first two factorial scores.

The factorial map indicates, consistent with the parameters-based approach, that Natyssa and Pilu varieties are significantly different from respect to texture. However, the factorial map highlights other differences not as clearly observed using the parameters-based approach. Furthermore, the discrimination among varieties was more accurate since less overlapping was observed between confidence ellipses. The confusion matrix (Table 3) confirms the visual observations. Indeed, approximately 94% correct classification was achieved (versus 46% using the parameters-based approach). To evaluate the effect of preprocessing on the discriminant power of the digitized curve, a FDA was performed on non-preprocessed curves using 10 factors. The result showed that only 45% of the tomatoes were correctly classified, a result similar to that found with the parameters-based approach.

Notably, although the analyzed **X** matrix of digitized curves presented 250 variables (against 13 for the parameter matrix), the final number of factors used in the FDA was close to that in the parameters-based approach. Only 10 factors were used in the digitized approach. Thus, using the digitized approach did not lead to an overfitted and/or artificial model compared to the parameters-based approach. Moreover, a progression of more than 80% of the discriminating power was observed between the two approaches.

4 | CONCLUSION

The aim of this study was to evaluate the use of digitized curves from penetrometry to discriminate 12 varieties of tomato. The digitized curve approach proposed a preprocessing of the curves based on standardization and smoothing procedures. This approach was compared to an optimized conventional method using texture parameters computed from the force–displacement curves of penetrometry.

The conventional approach enabled the classification of 50% of the tomato varieties compared with more than a 90% success rate with the preprocessed digitized curves approach. This improvement of the discriminating power is notable since the models use a similar number of factors (between 10 and 12).

The effect of the preprocessing was significant for improving the discrimination power of the digitized curves approach. Without preprocessing, 46% of tomatoes were accurately classified compared the 94% after preprocessing.

Thus, the proposition for penetrometry test analysis presents new opportunities for analyses of food supply chains. Indeed, texture is a key factor for the breeding of various fruits and vegetables. However, the reference parameter (firmness) is often unsatisfactory to develop molecular markers and breed new varieties with significant improvements in textural properties. Thus, the use of the entire digitized curve as an MI for characterizing fruit texture appears to be a promising approach.

Moreover, this approach could be developed for other mechanical tests to compute force–displacement or force–time curves in other fruits and vegetables.

ETHICAL STATEMENTS

Conflict of Interest: The author declares that he does not have any conflict of interest.

Ethical Review: This study does not involve any human or animal testing.

Informed Consent: No co-author.

ORCID

Cédric Camps http://orcid.org/0000-0003-1086-5668

REFERENCES

Barnes, R. J., Dhanoa, M. S., & Lister, S. J. (1989). Standard normal variate transformation and de-treding of near-infrared diffuse reflectance spectra. *Applied Spectroscopy*, *43*, 772–777.

Bertrand, D., Courcoux, P., Autran, J. C., & Méritan, R. (1990). Stepwise canonical discriminant analysis of continuous digitalized signals: Application to chromatograms of wheat proteins. *Journal of Chemometrics*, *4*, 427–413.

Camps, C., Guillermin, P., Mauget, J. C., & Bertrand, D. (2005). Data analysis of penetrometric force/displacement curves for the characterization of whole apple fruits. *Journal of Texture Studies*, *36*, 387–401. https://doi.org/10.1111/j.1745-4603.2005.00023.x

Causse, M., Buret, M., Robini, K., & Verschave, P. (2003). Inheritance of nutritional and sensory quality traits in fresh market tomato and relation to consumer preferences. *Journal of Food Science*, *68*, 2342–2350. https://doi.org/10.1111/j.1365-2621.2003.tb05770.x

Causse, M., Saliba-Colombani, V., Lecomte, L., Duffe, P., Rousselle, P., & Buret, M. (2002). QTL analysis of fruit quality in fresh market tomato: A few chromosome regions control the variation of sensory and instrumental traits. *Journal of Experimental Botany*, *53*, 2089–2098.

Chaïb, J., Devaux, M.-F., Grotte, M.-G., Robini, K., Causse, M., Lahaye, M., & Marty, I. (2007). Physiological relationships among physical, sensory, and morphological attributes of texture in tomato fruits. *Journal of Experimental Botany*, *58*, 1915–1925. https://doi.org/10.1093/jxb/erm046

Chen, H., Duprat, F., Grotte, M., Loonis, D., & Pietri, E. (1995). Relationship of impact transmission wave to apple texture during ripening. *Journal of Texture Studies*, *27*, 123–141.

Duprat, F., Grotte, M. G., Pietri, E., & Studman, C. J. (1995). A multipurpose firmness tester for fruits and vegetables. *Computers and Electronics in Agriculture*, *12*, 211–223.

Lahaye, M., Devaux, M. F., Poole, M., Seymour, G. B., & Causse, M. (2013). Pericarp tissue microstructure and cell wall polysaccharide chemistry are differently affected in lines of tomato with contrasted firmness. *Postharvest Biology and Technology*, *76*, 83–90. https://doi.org/10.1016/j.postharvbio.2012.09.009

Mehinagic, E., Royer, G., Symoneaux, R., Bertrand, D., & Jourjon, F. (2004). Prediction of the sensory quality of apples by physical measurements. *Postharvest Biology and Technology*, *34*, 257–269. https://doi.org/10.1016/j.postharvbio.2004.05.017

Roudot, A. C., Duprat, F., & Wenian, C. (1991). Modelling the response of apples to loads. *Journal of Agricultural Engineering Research*, *48*, 249–259.

Sargent, S. A., & Moretti, C. L. (2002). Tomato. In K. C. Gross, Wang, C. Y., & Saltveit, M. E. (Eds.), *Agricultural handbook 66 - The commercial storage of fruits, vegetables and florist and nursery crops*. Retrieved from http://www.ba.ars.usda.gov/hb66/contents.html

Serrano-Megías, M., & López-Nicolás, J. M. (2006). Application of agglomerative hierarchical clustering to identify consumer tomato preferences: Influence of physicochemical and sensory characteristics on consumer response. *Journal of the Science of Food and Agriculture*, *86*, 493–499. https://doi.org/10.1002/jsfa.2392

How to cite this article: Camps C. Singular approach to penetrometry by preprocessing of digitized force–displacement curves and chemometry: A case study of 12 tomato varieties. *J Texture Stud*. 2017;00:1–9. https://doi.org/10.1111/jtxs.12316