



HAL
open science

Mid-infrared technique to forecast cooked puree properties from raw apples: A potential strategy towards sustainability and precision processing

Weijie Lan, Catherine M.G.C. Renard, Benoit Jaillais, Alexandra Buergy, Alexandre Leca, Songchao Chen, Sylvie Bureau

► To cite this version:

Weijie Lan, Catherine M.G.C. Renard, Benoit Jaillais, Alexandra Buergy, Alexandre Leca, et al.. Mid-infrared technique to forecast cooked puree properties from raw apples: A potential strategy towards sustainability and precision processing. *Food Chemistry*, 2021, 355 (2), pp.129636. 10.1016/j.foodchem.2021.129636 . hal-03250179

HAL Id: hal-03250179

<https://hal.inrae.fr/hal-03250179v1>

Submitted on 24 Apr 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

1 **Mid-infrared technique to forecast cooked puree properties from raw apples: a**
2 **potential strategy towards sustainability and precision processing**

3

4 Weijie Lan^a, Catherine M.G.C. Renard^{a,b}, Benoit Jaillais^c, Alexandra Buergy^a,
5 Alexandre Leca^a, Songchao Chen^d, Sylvie Bureau^{a*}

6

7 ^a INRAE, Avignon University, UMR408 Sécurité et Qualité des Produits d'Origine
8 Végétale, F-84000 Avignon, France.

9 ^b INRAE, TRANSFORM Division, F-44000 Nantes, France.

10 ^c INRAE, ONIRIS, Unité Statistiques, Sensométrie, Chimiométrie (StatSC), F-44322
11 Nantes, France.

12 ^d INRAE, Unité InforSol, F-45075 Orléans, France.

13

14 **Corresponding author***

15 Sylvie Bureau (E-mail: sylvie.bureau@inrae.fr).

16 INRAE, UMR408 SQPOV « Sécurité et Qualité des Produits d'Origine Végétale »

17 228 route de l'Aérodrome

18 CS 40509

19 F-84914 Avignon cedex 9

20 Tel: +33 432722509

21 **Other authors**

22 Catherine M. G. C Renard: catherine.renard@inrae.fr

23 Songchao Chen: Songchao.Chen@inrae.fr

24 Benoit Jaillais: benoit.jaillais@inrae.fr

25 Alexandre Leca: alexandre.leca@inrae.fr

26 Alexandra Buergy: alexandra.burgy@inrae.fr

27 Weijie Lan: Weijie.Lan@inrae.fr

28

29 **Highlights**

30 MIRS discriminated purees cooked from different apples and processing conditions.

31 MIRS on purees gave robust predictions of soluble solids and acidity ($RPD \geq 3.1$).

32 Spectra of purees could be calculated from spectra of homogenized raw apples.

33 The calculated spectra allowed acceptable predictions of puree viscosity ($RPD \geq 2.5$).

34

35 **Abstract**

36 The potential of MIRS was investigated to: i) differentiate cooked purees issued from
37 different apples and process conditions, and ii) predict the puree quality characteristics
38 from the spectra of homogenized raw apples. Partial least squares (PLS) regression was
39 tested both, on the real spectra of cooked purees and their reconstructed spectra
40 calculated from the spectra of homogenized raw apples by direct standardization. The
41 cooked purees were well-classified according to apple thinning practices and cold
42 storage durations, and to different heating and grinding conditions. PLS models using
43 the spectra of homogenized raw apples can anticipate the titratable acidity (the residual
44 predictive deviation (RPD) = 2.9), soluble solid content (RPD = 2.8), particle averaged
45 size (RPD = 2.6) and viscosity (RPD \geq 2.5) of cooked purees. MIR technique can
46 provide sustainable evaluations of puree quality, and even forecast texture and taste of
47 purees based on the prior information of raw materials.

48

49 **Key words:** *Malus x domestica* Borkh.; Mid infrared spectroscopy; PLS models; Direct
50 standardization; Discriminant analysis.

51

52 **Introduction**

53 Apple puree is one of the major industrially processed fruit products (over 0.3
54 million tons consumed per year in France and the world market value about 2000
55 million USD annually) (FranceAgriMer, 2017; Market Research Future, 2019), and can
56 be used as the basic ingredient of jams, preserves or compotes (Defernez, Kemsley, &
57 Wilson, 1995). The quality of apple purees is systematically influenced by both raw
58 material characteristics (Buergy, Rolland-Sabaté, Leca, & Renard, 2020; Lan, Bureau,
59 Chen, Leca, Renard, & Jaillais, 2021; Lan, Jaillais, Leca, Renard, & Bureau, 2020;
60 Rembiałkowska, Hallmann, & Rusaczonok, 2007) and cooking strategies (heating,
61 grinding speed and refining levels etc.) (Espinosa, To, Symoneaux, Renard, Biau, &
62 Cuvelier, 2011; Oszmiański, Wolniak, Wojdyło, & Wawer, 2008; Picouet, Landl,
63 Abadias, Castellari, & Viñas, 2009). In practical apple processing, industrial
64 manufactures have to face the variability and heterogeneity of raw apples, optimize
65 their processing strategies to maintain the sustainable and expected quality level of final
66 puree products. Thus, developing rapid, efficient and integrated methods is needed to
67 guide suitable fruit processing procedures, even to design innovative foods by using the
68 raw material variability, and to reduce fruit wastes all along the processing chain.

69 Mid infrared spectroscopy (MIRS) is one of the main candidates for both the
70 quantification and qualification of agricultural commodities and processed food
71 (Bureau, Cozzolino, & Clark, 2019; Downey, 1998). Although MIRS presents a
72 relatively lower ability for quantification than that of chromatography or mass
73 spectrometry, it has the advantage of a rapid data acquisition and can provide
74 informative fundamental vibrations of molecular bonds (Karoui, Mazerolles, & Dufour,
75 2003). It does require a minimal sample preparation as the measurement must be done
76 on homogeneous samples as liquid, puree or powder due to the very low penetration of
77 radiation into the samples. Direct MIR characterizations of raw and processed fruits
78 have shown considerable aptitudes to evaluate soluble solids content (SSC), dry matter
79 content (DMC), titratable acidity (TA), some individual sugars, organic acids,
80 rheological (viscosity and viscoelasticity) and structural (particle averaged size and
81 volume) properties (Ayvaz, Sierra-Cadavid, Aykas, Mulqueeney, Sullivan, &

82 [Rodriguez-Saona, 2016; Lan, Renard, Jaillais, Leca, & Bureau, 2020](#)). These studied
83 parameters are related to the taste (SSC, DMC, TA, malic acid), the texture (viscosity,
84 viscoelasticity, particles and cell wall contents) and the basic nutrients (fructose,
85 sucrose and glucose) impacting in a large amount puree quality ([Bureau, Ścibisz, Le
86 Bourvellec, & Renard, 2012; Espinosa-Muñoz, Renard, Symoneaux, Biau, & Cuvelier,
87 2013; Fügél, Carle, & Schieber, 2005](#)).

88 Currently, our interest is to determine the possibility of using this technique to
89 anticipate the characteristics of processed materials from the data acquired on
90 homogenized raw fruit. According to our previous studies, strong correlations of
91 spectral, chemical and textural properties between raw apples and the corresponding
92 purees have been pointed out ([Lan, Jaillais, Leca, Renard, & Bureau, 2020; Lan, Renard,
93 Jaillais, Leca, & Bureau, 2020](#)). Based on that, the quality of final processed purees
94 could be predicted by the infrared spectral information acquired on raw apples using
95 partial least square (PLS) regression ([Lan, Jaillais, Leca, Renard, & Bureau, 2020](#)). The
96 main drawback of this strategy is the need, for modelling, to systematically acquire the
97 corresponding spectra on both raw and processed materials with a large number of
98 conditions representative of the variability, giving often only rough assessments. In
99 addition, the internal correlations of quality traits during puree processing were only
100 confirmed under one of the most commonly used processing conditions ([Lan, Jaillais,
101 Leca, Renard, & Bureau, 2020](#)).

102 Direct standardization (DS) is a simple and efficient chemometric tool for the
103 calibration transfer between spectral measurements or between two different sets of
104 conditions, such as the spectral calibration from the off-line to on-line spectra of olive
105 fruits ([Salguero-Chaparro, Palagos, Peña-Rodríguez, & Roger, 2013](#)). As far as we
106 know, this method has never been considered to bridge the spectral variations along the
107 fruit processing chain. Our interest of this method is thus to find the spectra
108 relationships of all processed purees and their corresponding spectral information
109 acquired on homogenized apples, and to calculate the reconstructed processed puree
110 spectra according to their relative spectral information acquired on apples by DS, taking
111 into account both the variability of raw materials and of commonly used processing

112 conditions. If so, the predictive models of puree quality traits (biochemical and physical)
113 using their reconstructed spectra dataset open the possibility to i) predict the properties
114 of processed apples based on the prior information of raw materials; ii) provide
115 sustainable and precise processing strategies to estimate the quality potential of final
116 products, and iii) to compare *in silico* the prediction results of different processing
117 approaches to better control the quality of fruit products.

118 Accordingly, this work aimed to assess the potential of MIRS to: i) detect the
119 variability of the cooked apple purees according to the pre- and post-harvest conditions
120 (fruit thinning and storage periods) and the main processing conditions (heating
121 temperature and grinding speed); ii) calculate reconstructed spectra of purees taking
122 into account the variability of raw fruits and processing conditions; and iii) characterize
123 and anticipate the biochemical (SSC, DMC, TA, individual sugars and malic acid),
124 rheological (viscosity and viscoelasticity) and textural (particle size and volume)
125 properties of the processed purees.

126 **Materials and methods**

127 **2.1 Apples and purees**

128 **2.1.1 Apples**

129 The experiment was conducted on the cultivar ‘Golden Delicious’ in 2017 and 2019.
130 All apples were harvested at commercial maturity from La Pugère experimental orchard
131 (Mallemort, Bouches du Rhône, France) (**Figure 1**).

132 - In 2017, half of the ‘Golden Delicious’ apples were subjected to a commercial
133 chemical fruit thinning (Th+) with standard fruit load (50-100 fruits/tree), the other
134 half was not thinned (Th-), resulting in a high fruit load (150-200 fruits/tree). After
135 harvesting, apples were processed into purees the day after harvest (T0), and after
136 one (T1), three (T3) and six months (T6) of cold storage at 4°C.

137 - In 2019, the commercially thinned ‘Golden Delicious’ apples (Th+) were stored
138 for up to one month (T1) at 4 °C until starch regression, then processed into purees
139 under different processing conditions.

140 **2.1.2 Puree processing**

141 Before processing, and for each condition, around 2 kg apples were homogenized

142 at 11000 rpm with an Ultraturrax T-25 (IKA, Labortechnik, GmbH, Staufen, Germany)
143 as raw apple homogenates. Three batches of apples (3 kg each) were used to produce
144 three puree lots for each condition. After sorting and washing, Golden Delicious apples
145 were cored and cut in 8 portions, then processed in a multi-functional processing system
146 (Roboqbo, Qb8-3, Bentivoglio, Italy) by different conditions (**Figure 1**):

147 - In 2017, all apple groups (2 thinning practices \times 4 storage periods) were
148 cooked with a standard Hot Break recipe of 95°C for 5 min at a 1500 rpm grinding
149 speed, then cooled down to 65°C while maintaining the grinding speed. Finally, 24
150 different cooked purees (2 thinning practices \times 4 storage periods \times 3 lots) were prepared.

151 - In 2019, each apple group was processed with three different heating
152 temperatures of 70°C, 83°C and 95°C for 30 min, while ground at three speed levels of
153 300 rpm, 1000 rpm and 3000 rpm, respectively. Totally, 27 different cooked purees (3
154 heating temperatures \times 3 grinding speeds \times 3 lots) were prepared.

155 Finally, all cooked purees were conditioned in two hermetically sealing plastic bags:
156 one was cooled at room temperature (22.5 °C) before the next-day measurements of
157 rheological, structural and some biochemical (SSC, TA, fructose, sucrose, glucose and
158 malic acid) properties. And the other one was freeze-dried (FD) and stored at -20 °C
159 for the determination of the content of cell wall, which are known to be a major
160 contributor of rheological properties of apple purees ([Espinosa-Muñoz, Renard,](#)
161 [Symoneaux, Biau, & Cuvelier, 2013](#)).

162 **2.2 Determination of puree quality traits**

163 **2.2.1 Rheological and structural characterizations on cooked purees**

164 The cooked puree rheological measurements were carried out using a Physica
165 MCR-301 controlled stress rheometer (Anton Paar, Graz, Austria) and a 6-vane
166 geometry (FL100/6W) with a gap of 3.46 mm, at 22.5 °C. The flow curves were
167 performed after a pre-shearing period of 1 minute at a shear rate of 50 s⁻¹, followed by
168 5 minutes at rest. The viscosity was then measured at a controlled shear rate range of
169 [10; 250] s⁻¹ on a logarithmic ramp. The values of viscosity at 50 s⁻¹ and 100 s⁻¹ (η_{50}
170 and η_{100} respectively) were kept as final indicators of the puree viscosity linked to
171 sensory characteristics during consumption ([Engelen & de Wijk, 2012](#)). Amplitude

172 Sweep (AS) tests were performed at an angular frequency of 10 rad./s in the
173 deformation range of [0.01; 100] %, in order to determine the linear viscoelastic range
174 of the purees and the yield stress, defined as the crossing point between the storage
175 modulus (G') and the loss modulus (G'') curves.

176 Cooked purees were diluted in distilled water to separate particles, stained with
177 calcofluor (0.1 g/L) and highlighted with a 365 nm UV lamp (Soukup, 2014). The
178 particle sizes averaged over volume $d(4:3)$ (de Brouckere mean) and over surface area
179 $d(3:2)$ (Sauter mean) were measured by a laser granulometer (Rawle, 2003)
180 (Mastersizer 2000, Malvern Instruments, Malvern, UK).

181 **2.2.2 Biochemical analyses on cooked purees**

182 SSC was determined with a digital refractometer (PR-101 ATAGO, Norfolk, VA,
183 USA) and expressed in °Brix at 20 °C. TA was determined by titration up to pH 8.1
184 with 0.1 mol/L NaOH and expressed in mmol H^+ /kg of fresh weight (FW) using an
185 autotitrator (Methrom, Herisau, Switzerland). Individual sugars and malic acid were
186 quantified using colorimetric enzymatic kits, according to the manufacturer's
187 instructions (R-biopharm, Darmstadt, Germany). The content of glucose, fructose,
188 sucrose and malic acid were expressed in g/kg FW. These measurements were
189 performed with a SAFAS flx-Xenius XM spectrofluorimeter (SAFAS, Monaco) at 570
190 nm for the sugars and 450 nm for malic acid. The DMC was estimated from the weight
191 of freeze-dried samples upon reaching a constant weight (freeze-drier, 5 days). Cell
192 wall materials of purees were isolated using the alcohol insoluble solids method
193 proposed by Renard (2005) and the cell wall contents (AIS contents) were expressed in
194 both FW and dry matter weight (DW).

195 **2.3 Spectrum acquisition on raw apple homogenates and cooked purees**

196 The MIR spectra were acquired at 23 °C using a Tensor 27 FTIR spectrometer
197 (Bruker Optics®, Wissembourg, France) equipped with a horizontal attenuated total
198 reflectance (ATR) sampling accessory and a deuterated triglycine sulphate (DTGS)
199 detector. The samples were placed at the surface of a zinc selenide crystal (ATR-ZnSe)
200 with six internal reflections. Spectra with 32 scans for ATR-ZnSe were collected from
201 4000 cm^{-1} to 650 cm^{-1} with a 4 cm^{-1} resolution and were corrected against the

202 background spectrum of air. Three replications of spectral measurement were done for
203 each homogenized raw apples and each cooked apple puree.

204 The whole spectral dataset of MIR is described in **Figure S1**. It included i) 81
205 spectra of raw apple homogenates, of which 72 spectra acquired in 2017 (3 apple
206 batches \times 2 fruit thinning conditions \times 4 storage times \times 3 spectral replicates) and 9
207 spectra acquired in 2019 (3 apple batches \times 3 spectral replicates); and ii) 153 spectra of
208 cooked apple purees, containing 72 spectra acquired in 2017 (2 fruit thinning conditions
209 \times 4 storage times \times 3 processing lots \times 3 spectral replicates) and 81 spectra acquired in
210 2019 (3 heating temperatures \times 3 grinding levels \times 3 processing lots \times 3 spectral
211 replicates).

212 **2.4 Statistical analyses of reference data**

213 The reference data of cooked purees processed in 2017 and 2019 are presented as
214 the mean values and the data dispersion within our experimental dataset expressed as
215 standard deviation values (SD). After the Shapiro-Wilk tests, the references data of
216 processed purees affected by fruit thinning and storage times were normal distributed
217 ($\alpha=0.05$), but not for the dataset of heating temperature and grinding effects during
218 puree processing. Thus, analysis of variance (ANOVA) was carried out to determine
219 the significant differences of cooked purees due to fruit thinning and storage times
220 applied on raw apples (**Table S1**) using XLSTAT (version 2018.5.52037, Addinsoft
221 SARL, Paris, France) data analysis toolbox. Kruskal-Wallis tests were performed to
222 evaluate the effects of heating temperature and grinding levels during puree processing
223 (**Table S2**).

224 **2.5 Spectra transferred by direct standardization (DS)**

225 In this study, DS was used to find the relationship between the spectra matrices of
226 all cooked purees (***P***) and their corresponding spectra of raw apple homogenates (***F***),
227 taking into account the effects of raw material variability and processing conditions.
228 The DS transfer works were performed in R software (version 4.0.2) ([R Core Team,](#)
229 [2019](#)) following a previous report ([Ji, Viscarra Rossel, & Shi, 2015](#)):

$$230 \quad \mathbf{P} = \mathbf{FB} + \mathbf{E} \quad (1)$$

231 where ***B*** is the transfer matrix ($\lambda \times \lambda$) presenting the variations in both ***F*** and ***P***,

232 E is the residual matrix used to correct the baseline difference. F , P and E matrices
 233 have the same size $n \times \lambda$, where n presents the numbers of transfer spectra and λ is the
 234 number of wavenumbers between 1800 and 900 cm^{-1} .

235 First, to compute the transfer B and error E matrices, the whole MIR spectral
 236 dataset (P and F) was divided into: the calibration matrices, presenting the first two
 237 batches of raw apple homogenates (Fc) and the first two lots of cooked purees (Pc),
 238 and the validation matrices with the third batch of raw apple homogenates (Fv) and the
 239 third lot of cooked purees (Pv) (Figure S1).

240 In a second step, DS was performed separately on the calibration matrices of raw
 241 apple homogenates (Fc) and cooked purees (Pc) in 2017 (Fc_{2017} and Pc_{2017}) and
 242 2019 (Fc_{2019} and Pc_{2019}) (Figures S1 and 2):

243 - the calibration matrices of apples (Fc_{2017}) and purees (Pc_{2017}) were processed
 244 to obtain the B_0 and E_0 related to the effects of raw materials on the processed
 245 purees as follows:

$$246 \quad Pc_{2017} = Fc_{2017}B_0 + E_0 \quad (2)$$

247 Both (Fc_{2017}) and (Pc_{2017}) have the same size $n \times \lambda$, where $n = 48$; (2 thinning
 248 practices \times 4 storage periods \times 2 apple batches/ puree lots \times 3 spectral replicates.

249 - the calibration matrices of apples (Fc_{2019}) and purees (Pc_{2019}) were performed
 250 for each puree processing condition, as follows:

$$251 \quad Pc_{2019}^{(i)} = Fc_{2019}B_i + E_i \quad (3)$$

252 where i from 1 to 9, corresponding to 9 different processing conditions (3 heating
 253 temperatures \times 3 grinding speeds). To each spectral replicate of Fc_{2019} corresponds
 254 nine spectra according to each processing condition (Pc_{2019}). All the spectra of Pc_{2019}
 255 matrix corresponding to the same processing conditions were gathered in a specific
 256 matrix $Pc_{2019}^{(i)}$. The size of this matrix (one for each processing condition) is equal to
 257 that of raw apple homogenates Fc_{2019} ($n = 6$ spectra (2 apple batches \times 3 spectral
 258 replicates) \times λ).

259 Thirdly, once all the transfer B (B_0 and B_i) and error E (E_0 and E_i) matrices
 260 were computed, they were used to calculate the cooked puree reconstructed calibration

261 and validation spectra matrices, as follows (**Figure 2**):

$$262 \quad \mathbf{Tc}_{2017} = \mathbf{Fc}_{2017}\mathbf{B}_0 + \mathbf{E}_0 \quad (4)$$

$$263 \quad \mathbf{Tv}_{2017} = \mathbf{Fv}_{2017}\mathbf{B}_0 + \mathbf{E}_0 \quad (5)$$

$$264 \quad \mathbf{Tc}_{2019} = \mathbf{Fc}_{2019}\mathbf{B}_i + \mathbf{E}_i \quad (6)$$

$$265 \quad \mathbf{Tv}_{2019} = \mathbf{Fv}_{2019}\mathbf{B}_i + \mathbf{E}_i \quad (7)$$

266 Finally, the reconstructed calibration and validation spectral matrices of cooked
267 puree, \mathbf{Tc} ($\mathbf{Tc}_{2017} + \mathbf{Tc}_{2019}$) and \mathbf{Tv} ($\mathbf{Tv}_{2017} + \mathbf{Tv}_{2019}$) of the two years (2017
268 and 2019) were obtained with the same sizes of the real spectral matrices of cooked
269 puree, \mathbf{Pc} and \mathbf{Pv} , for the further multivariate regressions.

270 **2.6 Multivariate regression**

271 Spectral pre-processing and multivariate data analysis were performed with Matlab
272 7.5 (Mathworks Inc. Natick, MA, USA) software using the SAISIR package ([Cordella](#)
273 [& Bertrand, 2014](#)). After pretests of several pre-processing treatments (baseline
274 correction, standard normal variate (SNV) and a derivative transform calculation using
275 Savitzky–Golay method (window size = 11, 21, 31) applied on several different spectral
276 regions, the best results of prediction and discrimination were obtained on the range
277 1800-900 cm^{-1} , which has been already highlighted ([Lan, Renard, Jaillais, Leca, &](#)
278 [Bureau, 2020](#)). Principal Component Analysis (PCA) and Factorial Discriminant
279 Analysis (FDA) were applied on SNV pre-treated spectra of cooked purees to detect
280 their differences related to the variability of both, raw apples and processing conditions.
281 The specificity and sensitivity values of FDA discriminations were calculated by the
282 already reported method of Nargis et al. ([2019](#)).

283 PLS models were developed using the SNV pre-processed puree spectra (1800-900
284 cm^{-1}) of the calibration set \mathbf{Pc} and the DS transferred spectra of purees (\mathbf{Tc}),
285 corresponding to the same reference dataset. The two calibration matrices of cooked
286 purees included a total of 102 spectra (48 spectra in 2017: 2 thinning practices \times 4
287 storage periods \times 2 lots \times 3 spectral replicates, and 54 spectra in 2019: 3 heating
288 temperatures \times 3 grinding speeds \times 2 lots \times 3 spectral replicates). Then, the developed
289 PLS models were applied on their corresponding validation spectra sets of \mathbf{Pv} and
290 \mathbf{Tv} , with a total of 51 spectra in 2017: 2 thinning practices \times 4 storage periods \times 1 lot

291 × 3 spectral replicates and in 2019: 3 heating temperatures × 3 grinding speeds × 1 lot
292 × 3 spectral replicates. PLS model performance was assessed using the determination
293 coefficient of calibration (R_c^2) and validation (R_v^2), the root-mean-square error of
294 validation (RMSEV), the number of latent variables for calibration (LVs), the residual
295 predictive deviation of validation set (RPD), as described by Nicolai et al. (2007). The
296 linkable spectral regions of the acceptable PLS models presenting RPD values higher
297 than 2.5 (Nicolai, Beullens, Bobelyn, Peirs, Saeys, Theron, et al., 2007) were displayed
298 based on their β - coefficients (Tables 1 and 2).

299 3. Results and discussion

300 3.1 Variability of cooked purees based on their MIR Spectra

301 3.1.1 Variability induced by the raw materials

302 According to ANOVA (F-values), fruit thinning applied on apples during their
303 growth in orchard resulted in a significant variation ($p < 0.001$) of viscosity (η_{50} and
304 η_{100}), viscoelasticity (G' , G'' , yield stress), particle sizes (d4:3 and d3:2) and
305 biochemical compositions (DMC, SSC, TA, pH, malic acid, sucrose, fructose and AIS)
306 of the cooked purees. Particularly, the impact of thinning on the viscosity, DMC and
307 SSC of purees was higher than the effect of post-harvest storage at 4°C (Table S1).
308 Purees processed from thinned apples (Th+) had higher viscosity values (η_{50} and η_{100})
309 and bigger particle sizes (d4:3) than those from the non-thinned apples (Th-), observed
310 after the three months of cold storage (T3 and T6) (Buegy, Rolland-Sabaté, Leca, &
311 Renard, 2020). Moreover, an intensive decrease of average particle sizes (d 4:3) was
312 observed in the purees cooked with the apples stored one month at 4°C (T1) for both
313 ‘thinning’ (Th+) and ‘non-thinning’ (Th-) treatments. PCA applied on the spectra of
314 cooked purees in 2017 showed a good ability to detect the effects of treatments applied
315 on raw apples (fruit thinning and storage periods) (Figure 3a and 3b). The effect of
316 thinning on the first principal component (PC1 90.1%) was much higher than that of
317 storage on the second principal component (PC2 6.9%), which was in line with our
318 previous results (Lan, Renard, Jaillais, Leca, & Bureau, 2020). In addition, the increase
319 of the band at 1022 cm^{-1} and the decrease of the bands at 1061-1065 cm^{-1} , attributed to
320 sucrose and fructose respectively (Bureau, Cozzolino, & Clark, 2019), were the major

321 contributors of the observed discriminations on the two PCs (**Figure 3c and 3d**).

322 3.1.2 Variability induced by processing conditions

323 The different grinding speeds affected significantly ($p < 0.05$) the viscosity (η_{50}
324 and η_{100}), viscoelasticity (G' and G'') and particle size (d4:3 and d3:2) of the cooked
325 purees (**Table S2**). Particularly, the increase of grinding speed significantly ($p < 0.001$)
326 decreased viscosity (η_{50} and η_{100}), viscoelasticity (yield stress, G' and G'') and particle
327 sizes (d4:3) which was observed at each tested temperature. From macroscopic images
328 of purees (data not shown), the larger particles disappeared with increasing grinding
329 speeds, which was enough to cause a decrease in the puree viscosity ([Espinosa-Muñoz,](#)
330 [Renard, Symoneaux, Biau, & Cuvelier, 2013](#)). Inversely, the increasing heating
331 temperatures induced no significant ($p > 0.05$) changes of puree viscosity (η_{50} and η_{100})
332 and viscoelasticity (yield stress, G' and G''). The highest heating temperature (95°C)
333 resulted in a significant ($p < 0.05$) increase of DMC and SSC and a decrease of TA and
334 malic acid. Consequently, the changes of grinding speed during the puree processing
335 significantly modified the structural properties and viscoelastic behaviors of purees,
336 whereas heating temperature affected strongly the biochemical composition of purees.

337 FDA performed on the cooked puree spectra in 2019 successfully classified the
338 processing changes induced at the different heating temperatures (**Figure 4a**) and
339 grinding speeds (**Figure 4b**). The samples cooked at 95 °C were well-separated from
340 the other two conditions (4 factors, 100% of sensitivity and specificity in **Table S3a**),
341 according to the first factorial component (F1) (**Figure 4a**). The specific bands at 1745
342 cm^{-1} and 1539 cm^{-1} were attributed to the increase of soluble pectins, probably in
343 relationship with their solubilization in puree serum from apple cell walls, enhanced
344 with the increasing heating temperature ([Liu, Renard, Rolland-Sabaté, Bureau, & Le](#)
345 [Bourvellec, 2020](#)). Moreover, the negative peaks at 1057 cm^{-1} and 998 cm^{-1} could be
346 due to the hydrolysis of sucrose during thermal processing, thus resulting in the increase
347 of fructose (1022 cm^{-1}) and glucose (1107 cm^{-1}) contents.

348 The three different grinding levels could be discriminated according to the first two
349 factorial components (F1 and F2) (**Figure 4b**), especially for the highest grinding speed
350 at 3000 rpm ('G3' in **Figure 4b**) (4 factors, over 85.19% of specificity and sensitivity

351 in **Table S3b**). The intensive negative spectral peaks at 1558, 1539-1541, 1508 and
352 1458 cm^{-1} along both the two discriminant factors (F1 and F2 in **Figure 4d and 4e**)
353 were all located in the region between 1450 cm^{-1} and 1600 cm^{-1} , which has been already
354 attributed to the changes of particle size and rheological behavior after apple puree
355 mechanical refining in a previous experiment (Lan, Renard, Jaillais, Leca, & Bureau,
356 2020). These peaks indicated the decrease of particle size (d4:3 and d3:2) and viscosity
357 of purees with increasing grinding speeds, which was in line with our reference
358 measurements (**Table S2**) and Espinosa et al. (2011).

359 Briefly, MIR technique could detect several kinds of variability sources such as
360 thinning practices during fruit cultivation, cold storage and processing conditions
361 (temperature and grinding) in the cooked purees. In addition, the spectral region 1450-
362 1750 cm^{-1} was validated as being a reliable analytical signal of processing linked to the
363 textural and rheological changes in purees.

364 **3.2 Prediction of quality traits of cooked purees by MIRS**

365 For all developed PLS models, as expected the decreases of determination
366 coefficients between the calibration set (R_c^2) and the validation set (R_v^2) were observed
367 in **Table 1**. According to RPD values over 2.5 (Nicolai, Beullens, Bobelyn, Peirs, Saeys,
368 Theron, et al., 2007), prediction was acceptable to good (RPD from 2.6 to 3.3) for
369 viscosity (η_{50} and η_{100}), average particle sizes (d4:3), SSC, TA, pH values and malic
370 acid were content in cooked purees by MIRS, taking into account a large variability of
371 raw apple materials and processing conditions.

372 Apparent puree viscosity at a shear rate value of 50 s^{-1} (η_{50}), which has been
373 described to be the highly correlated with the in-mouth texture perception of fluid food
374 (Chen & Engelen, 2012), could be predicted by MIRS with a R_v^2 of 0.87 and a RPD of
375 3.2. MIR prediction of apparent puree viscosity at a single shear rate value of 100 s^{-1}
376 (η_{100}) ($R_v^2 = 0.85$, RPD= 3.0) observed here were much better than its prediction by
377 NIRS (RPD = 1.3) (Lan, Jaillais, Leca, Renard, & Bureau, 2020), These results
378 evidenced the possibility of MIRS to estimate puree viscosity. For the two apparent
379 puree viscosity values measured at η_{50} and η_{100} , the main wavenumber regions at 1718-
380 1730 cm^{-1} and 1618-1678 cm^{-1} were still observed in our previous work (Lan, Renard,

381 [Jaillais, Leca, & Bureau, 2020](#)). This could validate the application of MIRS to predict
382 puree viscosity by taking into account not only the raw fruit variability but also the
383 complex effects of processing conditions. However, the predictions ($R_v^2 < 0.81$, RPD
384 < 2.1) of the viscoelastic parameters of purees (G' , G'' and yield stress) were not precise
385 enough to estimate the viscoelastic behaviors and the moment when puree starts to flow.
386 Indeed, heating and grinding affected puree viscoelasticity (SD values of 2362 Pa for
387 G' , 595 Pa for G'' , 27.1 Pa for yield stress) and resulted in more than twice higher
388 variations of these parameters than those induced by thinning and cold storage on raw
389 materials (SD values of 1001 Pa for G' , 234 Pa for G'' and 12.9 Pa for yield stress).
390 These new prediction accuracies of the viscoelastic parameters of purees (G' , G'' and
391 yield stress) were not as good as our previous ones by MIRS ([Lan, Renard, Jaillais,](#)
392 [Leca, & Bureau, 2020](#)), but could be more robust to be considered for future
393 applications. MIR coupled with the linear regression (PLS) showed a good performance
394 ($R_v^2 = 0.87$, RPD= 3.1) to evaluate the volume average particle size of purees (d4:3),
395 but not the surface average particle size (d3:2). Particularly, the most informative
396 wavenumbers to evaluate puree particle size at 1701-1713 cm^{-1} , 1655-1668 cm^{-1} and
397 1537-1541 cm^{-1} have been already observed previously to predict puree viscosity, to
398 discriminate the purees prepared with different grinding speeds (mentioned in **part 3.1**)
399 and with different refining levels ([Lan, Renard, Jaillais, Leca, & Bureau, 2020](#)). Such
400 good prediction of puree average particle size (d4:3) could not come from internal
401 correlations with puree composition such as SSC, DMC or AIS contents because of
402 their poor correlation ($R^2 < 0.48$), but probably some specific signals needing to be
403 identified and confirmed. Moreover, we confirmed here the impossibility to predict the
404 cell wall content directly in puree by MIRS, without any preparation such as freeze-
405 drying ([Lan, Renard, Jaillais, Leca, & Bureau, 2020](#))

406 A good prediction of global puree quality traits, SSC (RPD= 3.1) and DMC (RPD=
407 2.9), was obtained with 5 LVs (**Table 1**). The variation of SSC and DMC in purees were
408 highly correlated ($R^2 = 0.76$), which explained the good estimations of these two
409 parameters. Their respective fingerprint wavenumbers of SSC and DMC prediction
410 were similar and detected at 996-1001 cm^{-1} , 1048-1057 cm^{-1} and 1109-1112 cm^{-1} ,

411 corresponding to the variations of sucrose, fructose and glucose in purees (Bureau,
412 Cozzolino, & Clark, 2019). Moreover, the prediction of TA was excellent (RPD = 3.2),
413 with a limited RMSEV of 7.6 mmol H⁺/kg FW. The typical wavenumber region at
414 1709-1720 cm⁻¹ in TA prediction has already been attributed to the C=O vibration of
415 acid group (Bureau, Quilot-Turion, Signoret, Renaud, Maucourt, Bancel, et al., 2013;
416 Clark, 2016). Depending on the good correlations between the different contributors of
417 apple acidity (R² = 0.81 between TA and malic acid, R² = 0.76 between TA and pH),
418 MIRS provided an acceptable prediction of pH (R_v² = 0.83 and RPD = 2.5) and malic
419 acid content (R_v² = 0.85 and RPD= 2.7). Despite the similar typical fingerprints
420 observed in the β-coefficients of PLS models of malic acid and pH, the relative lower
421 RPD values and R_v² of pH compared to malic acid were probably due to the low pH
422 variation. Concerning the main individual sugars, acceptable prediction was obtained
423 only for fructose (RPD = 2.6), but neither for sucrose (RPD= 1.3) nor for glucose (RPD
424 = 1.5), which was in line with our previous results (Lan, Renard, Jaillais, Leca, &
425 Bureau, 2020). The lower concentration of glucose (10.4-25.4 g/kg FW) than the other
426 individual sugars (34.9-98.7 g/kg FW of fructose, 39.1-118.5 g/kg FW of sucrose) led
427 to its worse prediction by MIR results. A higher internal biochemical correlation
428 between the major compounds (SSC, TA) and fructose (R² = 0.79 for SSC and fructose,
429 R² = 0.76 for TA and fructose) in apple purees might explain the better prediction of
430 fructose than the one of sucrose (R² = 0.58 for SSC and sucrose, R² = 0.44 for sucrose
431 and TA).

432 Briefly, MIR technique can provide a simultaneous and robust estimation of
433 biochemical compositions (dry matter, soluble solids, titratable acidity, pH, malic acid
434 and fructose), rheological behaviors (viscosity at η₅₀ and η₁₀₀) and particle size (d_{4:3})
435 of apple purees, taking into account the large variability along the apple puree
436 production chain (agricultural practices, post-harvest storage and processing
437 conditions).

438 **3.3 Reconstructed spectra for prediction of puree quality traits**

439 In this part, MIR prediction models were developed using the reconstructed spectra
440 of the calibration set of cooked purees (**Tc**), only done for the well-predicted parameters

441 mentioned in **part 3.2**, which were η_{50} and η_{100} , d4:3, SSC, DMC, TA, pH, malic acid
442 and fructose. Then, these models were applied on the validation reconstructed spectra
443 of the cooked purees (***Tv***).

444 Overall, based on the PLS regression applied on the puree reconstructed spectra,
445 acceptable predictions were obtained ($RPD > 2.5$) for rheological (η_{50} , η_{100}), structural
446 (d4:3) and global biochemical (SSC, DMC, TA) parameters (**Table 2**). In contrast,
447 predictions appeared not acceptable for malic acid ($RPD = 2.3$), fructose ($RPD = 1.7$)
448 and pH ($RPD = 2.1$). Compared to the previous prediction from the real puree spectra
449 (**Table 1**), lower R_v^2 and higher LVs have been generally obtained for all parameters
450 giving some lower prediction performance (**Table 2**). Particularly, the use of the
451 reconstructed spectra showed a good ability to predict puree viscosity parameters (η_{50}
452 and η_{100}) with $R_v^2 > 0.82$, $RPD > 2.5$ and prediction errors (RMSEV) of 0.21 Pa.s and
453 0.10 Pa.s for η_{50} and η_{100} , respectively. These results were close to those from the real
454 spectra of purees (**Table 1**). The fingerprint wavenumbers used in the PLS models were
455 similar for both reconstructed and real spectra, mainly 1718-1734 cm^{-1} , 1616-1336 cm^{-1}
456 and 1547-1553 cm^{-1} as described in **Part 3.2**. Although a relative lower R_v^2 and RPD
457 ($R_v^2 = 0.84$ and $RPD = 2.6$) were obtained for particle size (d4:3) compared to the results
458 on the real puree spectra (**Table 1**), the consistent fingerprints were highly related to
459 the puree texture such as 1701-1715 cm^{-1} , 1537-1541 cm^{-1} and 1101-1107 cm^{-1} . These
460 prediction performances revealed for the first time the possibility to evaluate the
461 variation of averaged particle sizes in the cooked purees based on the MIR information
462 of the corresponding raw apple homogenates. Considering the other global quality
463 parameters, acceptable predictions were obtained for SSC ($R_v^2 = 0.85$ and $RPD = 2.8$)
464 and DMC ($R_v^2 = 0.84$ and $RPD = 2.6$) contents. The specific wavenumbers in the ranges
465 997-1001 cm^{-1} and 1048-1057 cm^{-1} for sucrose and in the ranges 1009-1112 cm^{-1} for
466 fructose mainly contributed to the PLS models for both reconstructed and real spectra.
467 These ranges have been already mentioned to be linked to these sugars ([Bureau,
468 Cozzolino, & Clark, 2019](#)), which are the main ones in apples. For acidity, the
469 reconstructed spectra gave an excellent prediction of TA ($R_v^2 = 0.86$ and $RPD = 2.9$),
470 using the spectral regions between 1709-1720 cm^{-1} of the typical C=O absorption

471 (Clark, 2016). Consequently, PLS applied on the reconstructed MIR spectra calculated
472 from the spectra of raw apple homogenates showed the possibility to directly predict
473 the viscosity, averaged particle sizes, SSC, DMC and TA of cooked purees.

474 Several initial attempts have been tested to monitor the quality of cooked food from
475 infrared information of their raw materials with the objectives to predict the texture of
476 cooked poultry pectoralis major muscles (Meullenet, Jonville, Grezes, & Owens, 2004),
477 of cooked rice (Windham, Lyon, Champagne, Barton, Webb, McClung, et al., 1997)
478 and of apple purees (Lan, Jaillais, Leca, Renard, & Bureau, 2020). In these studies, the
479 spectra matrix of the raw materials and the reference data of the corresponding
480 processed materials were used to calibrate models. The predictions thus obtained are
481 mainly due to the strong internal correlations of quality traits between materials before
482 and after processing, which could provide semi-quantitative prediction accuracy for
483 practical uses. However, the internal correlations of quality traits during fruit processing
484 still remain unreliable when using a large variability of raw materials and various
485 industrial processing systems (Lan, Jaillais, Leca, Renard, & Bureau, 2020). Further,
486 such a direct modelling method requires a necessary step of acquisition of the infrared
487 information on raw material batches for each processing condition, in order to obtain
488 the same matrix sizes of spectra and reference data for calibration.

489 Here, a potential strategy has been firstly proposed to build reconstructed MIR
490 spectra of processed purees from the spectra of raw apple homogenates using a spectral
491 transfer method. The high consistency of the specific fingerprints used in the PLS
492 models obtained for both the real spectra and the reconstructed spectra, confirmed our
493 choice for this modelling strategy. Compared to the direct modelling method, a great
494 advantage of using spectral transfer strategy is that the calibration dataset only needs
495 the infrared information and reference data of several processed purees and just a
496 limited number of spectra of corresponding raw apples. For example, in our dataset of
497 2019, the reconstructed spectra of 27 different processed purees could be transferred
498 from only 3 corresponding spectra of the same apple batches.

499 After a simple scanning of raw apple homogenates by MIRS, our models revealed
500 the possibility to i) predict the quality of apple purees, such as viscosity, SSC and TA

501 using a standard processing recipe (95 °C for 5 mins and grinding at 1500 rpm), even
502 though a large variability of raw apples was used (different fruit thinning and cold
503 storage periods); and ii) to monitor and anticipate the organoleptic properties of cooked
504 purees under different processing strategies, which is relevant for the processors and
505 market. For example, a higher viscosity and acidity in-mouth feeling (predicted $\eta_{50} =$
506 1.42 ± 0.09 Pa.s, predicted TA = 65.8 ± 3.5 meq/kg FW) were predicted with the recipe
507 at 83 °C for 30 min and grinding speed of 1000 rpm than with the recipe at 95 °C for
508 30 min and grinding speed of 3000 rpm (predicted $\eta_{50} = 0.98 \pm 0.14$ Pa.s, predicted
509 TA = 56.4 ± 4.5 meq/kg FW).

510 **Conclusion**

511 As far as we know, this is the first study that shows the ability of MIRS to estimate
512 the quality of processed fruit products taking into account a large variability coming
513 from agricultural practices, post-harvest storage and processing conditions along the
514 whole processing chain. MIR technique provided reliable assessment of viscosity,
515 averaged particle sizes and major compositions (SSC, DMC, TA and malic acid) of
516 apple purees.

517 Further, a simple spectroscopic transfer method (direct standardization) was
518 applied for the first time to develop the reconstructed spectra of purees from their
519 corresponding spectra of raw apple homogenates. MIRS coupled with PLS regression
520 obtained acceptable predictions of TA, DMC, SSC, viscosity (η_{50} and η_{100}) and
521 averaged particle sizes of the final puree based on their reconstructed spectra. With a
522 simple scanning of raw apple homogenates, MIR technique opens the possibility to i)
523 predict the quality of final purees under a standard processing procedure, which is
524 beneficial for fruit processing sustainability; and even ii) to monitor the texture and
525 tastes of purees under different processing conditions for a better management.

526

527 **Acknowledgements**

528 The authors thank Patrice Reling, Barbara Gouble, Marielle Boge, Caroline Garcia and
529 Gisèle Riqueau (INRAE, SQPOV unit) for their technical help. The ‘Interfaces’ project
530 is an Agropolis Fondation Flashship project publicly funded through the ANR (French
531 Research Agency) under “Investissements d’Avenir” programme (ANR-10-LABX-01-
532 001 Labex Agro, coordinated by Agropolis Fondation). Weijie Lan was supported by a
533 doctoral grant from Chinese Scholarship Council.

534

535 **References**

- 536 Ayvaz, H., Sierra-Cadavid, A., Aykas, D. P., Mulqueeney, B., Sullivan, S., & Rodriguez-Saona, L. E.
537 (2016). Monitoring multicomponent quality traits in tomato juice using portable mid-infrared
538 (MIR) spectroscopy and multivariate analysis. *Food Control*, 66, 79-86.
- 539 Buergy, A., Rolland-Sabaté, A., Leca, A., & Renard, C. M. G. C. (2020). Pectin modifications in raw
540 fruits alter texture of plant cell dispersions. *Food Hydrocolloids*, 107, 105962.
- 541 Bureau, S., Ścibisz, I., Le Bourvellec, C., & Renard, C. M. G. C. (2012). Effect of Sample Preparation
542 on the Measurement of Sugars, Organic Acids, and Polyphenols in Apple Fruit by Mid-infrared
543 Spectroscopy. *Journal of Agricultural and Food Chemistry*, 60(14), 3551-3563.
- 544 Bureau, S., Cozzolino, D., & Clark, C. J. (2019). Contributions of Fourier-transform mid infrared (FT-
545 MIR) spectroscopy to the study of fruit and vegetables: A review. *Postharvest Biology and
546 Technology*, 148, 1-14.
- 547 Bureau, S., Quilot-Turion, B., Signoret, V., Renaud, C., Maucourt, M., Bancel, D., & Renard, C. M. G.
548 C. (2013). Determination of the Composition in Sugars and Organic Acids in Peach Using Mid
549 Infrared Spectroscopy: Comparison of Prediction Results According to Data Sets and Different
550 Reference Methods. *Analytical Chemistry*, 85(23), 11312-11318.
- 551 Chen, J., & Engelen, L. (2012). *Food oral processing: fundamentals of eating and sensory perception*:
552 John Wiley & Sons.
- 553 Clark, C. J. (2016). Fast determination by Fourier-transform infrared spectroscopy of sugar–acid
554 composition of citrus juices for determination of industry maturity standards. *New Zealand
555 Journal of Crop and Horticultural Science*, 44(1), 69-82.
- 556 Cordella, C. B. Y., & Bertrand, D. (2014). SAISIR: A new general chemometric toolbox. *TRAC Trends
557 in Analytical Chemistry*, 54, 75-82.
- 558 Defernez, M., Kemsley, E. K., & Wilson, R. H. (1995). Use of infrared spectroscopy and chemometrics
559 for the authentication of fruit purees. *Journal of Agricultural and Food Chemistry*, 43(1), 109-
560 113.
- 561 Downey, G. (1998). Food and food ingredient authentication by mid-infrared spectroscopy and
562 chemometrics. *TRAC Trends in Analytical Chemistry*, 17(7), 418-424.
- 563 Engelen, L., & de Wijk, R. A. (2012). Oral Processing and Texture Perception. In J. Chen & L. Engelen
564 (Eds.), *Food Oral Processing*, (pp. 157-176): Wiley-Blackwell.
- 565 Espinosa-Muñoz, L., Renard, C. M. G. C., Symoneaux, R., Biau, N., & Cuvelier, G. (2013). Structural
566 parameters that determine the rheological properties of apple puree. *Journal of Food
567 Engineering*, 119(3), 619-626.
- 568 Espinosa, L., To, N., Symoneaux, R., Renard, C. M. G. C., Biau, N., & Cuvelier, G. (2011). Effect of
569 processing on rheological, structural and sensory properties of apple puree. *Procedia Food
570 Science*, 1, 513-520.
- 571 FranceAgriMer. (2017). La Pomme en 2016-2017. Accessed October 2020, from
572 <https://www.rnm.franceagrimer.fr>.
- 573 Market Research Future (2019). Fruit Puree Market information: by fruit type (apple puree, banana puree,
574 plum puree, strawberry puree, and others), by application (baby food, bakery, beverages, and
575 others), by category (conventional and organic) and by Region - Global Forecast till 2023.
576 Accessed March 2019, from [https://www.marketresearchfuture.com/reports/fruit-puree-
577 market-5281](https://www.marketresearchfuture.com/reports/fruit-puree-market-5281).
- 578 Fügél, R., Carle, R., & Schieber, A. (2005). Quality and authenticity control of fruit purées, fruit

579 preparations and jams—a review. *Trends in Food Science & Technology*, 16(10), 433-441.

580 Ji, W., Viscarra Rossel, R., & Shi, Z. (2015). Accounting for the effects of water and the environment on
581 proximally sensed vis–NIR soil spectra and their calibrations. *European Journal of Soil Science*,
582 66(3), 555-565.

583 Karoui, R., Mazerolles, G., & Dufour, É. (2003). Spectroscopic techniques coupled with chemometric
584 tools for structure and texture determinations in dairy products. *International Dairy Journal*,
585 13(8), 607-620.

586 Lan, W., Bureau, S., Chen, S., Leca, A., Renard, C. M. G. C., & Jaillais, B. (2021). Visible, near- and
587 mid-infrared spectroscopy coupled with an innovative chemometric strategy to control apple
588 puree quality. *Food Control*, 120, 107546.

589 Lan, W., Jaillais, B., Leca, A., Renard, C. M. G. C., & Bureau, S. (2020). A new application of NIR
590 spectroscopy to describe and predict purees quality from the non-destructive apple
591 measurements. *Food Chemistry*, 310, 125944.

592 Lan, W., Renard, C. M. G. C., Jaillais, B., Leca, A., & Bureau, S. (2020). Fresh, freeze-dried or cell wall
593 samples: Which is the most appropriate to determine chemical, structural and rheological
594 variations during apple processing using ATR-FTIR spectroscopy? *Food Chemistry*, 330,
595 127357.

596 Liu, X., Renard, C. M. G. C., Rolland-Sabaté, A., Bureau, S., & Le Bourvellec, C. (2020). Modification
597 of apple, beet and kiwifruit cell walls by boiling in acid conditions: Common and specific
598 responses. *Food Hydrocolloids*, 106266.

599 Meullenet, J. F., Jonville, E., Grezes, D., & Owens, C. M. (2004). Prediction of the texture of cooked
600 poultry pectoralis major muscles by near-infrared reflectance analysis of raw meat. *Journal of*
601 *Texture Studies*, 35(6), 573-585.

602 Nargis, H., Nawaz, H., Ditta, A., Mahmood, T., Majeed, M., Rashid, N., Muddassar, M., Bhatti, H.,
603 Saleem, M., & Jilani, K. (2019). Raman spectroscopy of blood plasma samples from breast
604 cancer patients at different stages. *Spectrochimica Acta Part A: Molecular and Biomolecular*
605 *Spectroscopy*, 222, 117210.

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

609 Oszmiański, J., Wolniak, M., Wojdyło, A., & Wawer, I. (2008). Influence of apple purée preparation and
610 storage on polyphenol contents and antioxidant activity. *Food Chemistry*, 107(4), 1473-1484.

611 Picouet, P. A., Landl, A., Abadias, M., Castellari, M., & Viñas, I. (2009). Minimal processing of a Granny
612 Smith apple purée by microwave heating. *Innovative Food Science & Emerging Technologies*,
613 10(4), 545-550.

614 R Core Team, R. C. (2019). R: A Language and Environment for Statistical Computing. 1(1358), 34.

615 Rawle, A. (2003). Basic of principles of particle-size analysis. *Surface coatings international. Part A*,
616 *Coatings journal*, 86(2), 58-65.

617 Rembiałkowska, Ewa; Hallmann, Ewelina and Rusaczonek, Anna (2007) Influence of Processing on
618 Bioactive Substances Content and Antioxidant Properties of Apple Purée from Organic and
619 Conventional Production in Poland. Poster at: 3rd QLIF Congress: Improving Sustainability in
620 Organic and Low Input Food Production Systems, University of Hohenheim, Germany, May
621 31-June 3 2005.

622 Renard, C. M. G. C. (2005). Variability in cell wall preparations: quantification and comparison of

- 623 common methods. *Carbohydrate Polymers*, 60(4), 515-522.
- 624 Salguero-Chaparro, L., Palagos, B., Peña-Rodríguez, F., & Roger, J. M. (2013). Calibration transfer of
625 intact olive NIR spectra between a pre-dispersive instrument and a portable spectrometer.
626 *Computers and Electronics in Agriculture*, 96, 202-208.
- 627 Soukup, A. (2014). Selected Simple Methods of Plant Cell Wall Histochemistry and Staining for Light
628 Microscopy. In V. Žárský & F. Cvrčková (Eds.), *Plant Cell Morphogenesis: Methods and*
629 *Protocols*, (pp. 25-40). Totowa, NJ: Humana Press.
- 630 Windham, W. R., Lyon, B. G., Champagne, E. T., Barton, F. E., Webb, B. D., McClung, A. M.,
631 Moldenhauer, K. A., Linscombe, S., & McKenzie, K. S. (1997). Prediction of cooked rice
632 texture quality using near-infrared reflectance analysis of whole-grain milled samples. *Cereal*
633 *Chemistry*, 74(5), 626-632.
- 634

635 **Figure captions:**

636 **Figure 1.** Experimental scheme for apple production, puree preparation and the sample
637 characterization by infrared spectroscopy and reference measurements.

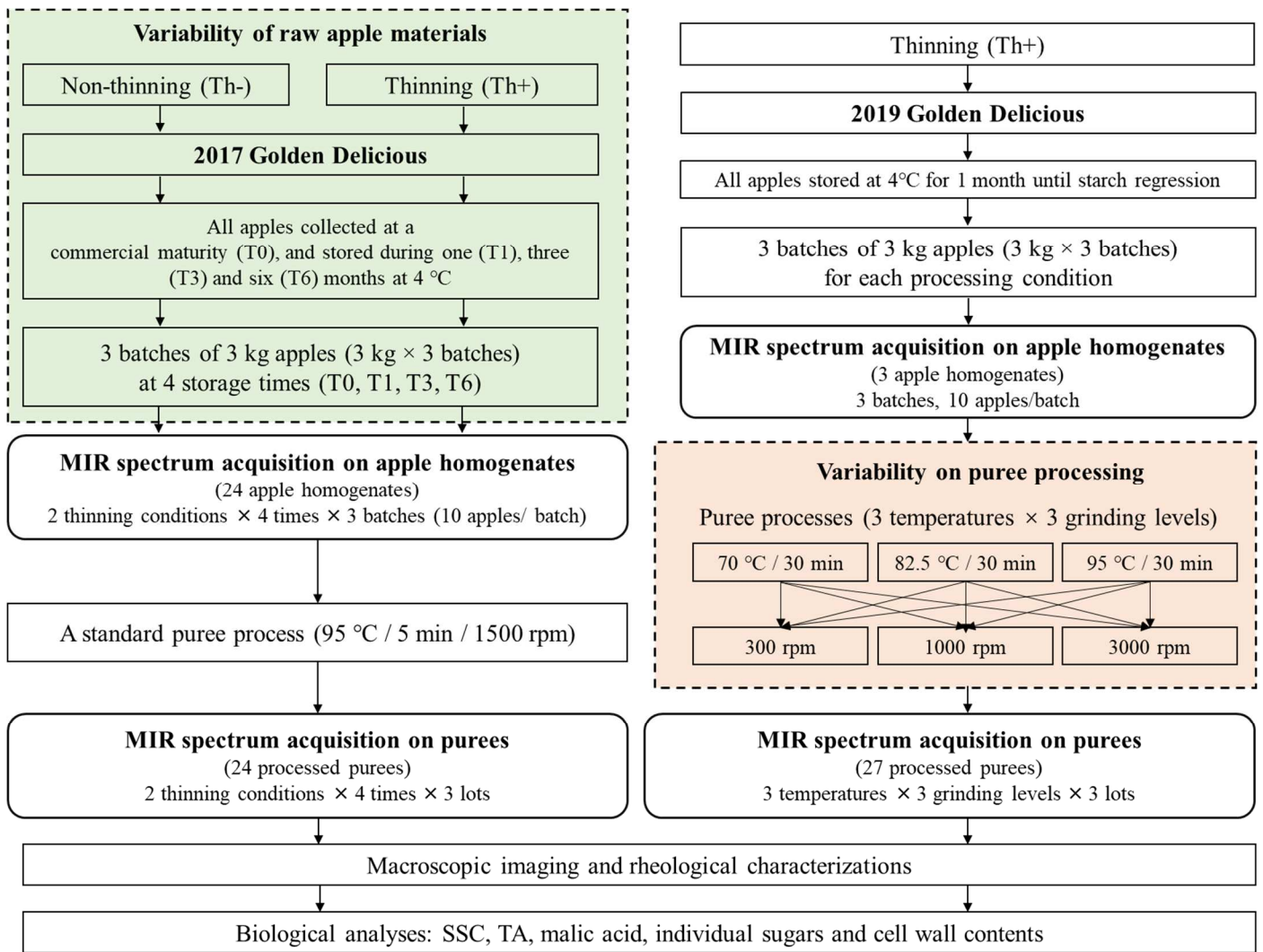
638 **Figure 2.** Overview of the applied methodology to exploit reconstructed MIR spectra
639 of purees and multivariate regression.

640 **Figure 3.** Principal Component Analysis on the SNV pre-treated MIR spectra (900-
641 1800 cm^{-1}) of purees cooked with thinned (Th+) and non-thinned (Th-) ‘Golden
642 Delicious’ apples stored at 4°C during 0, 1, 3 and 6 months (T0, T1, T3 and T6): **(a)** the
643 scores plot of the first two components (PC1 and PC2) related to fruit thinning; **(b)** the
644 scores plot of the first two components (PC1 and PC2) related to storage periods; **(c)**
645 the loading plot of PC1; **(d)** the loading plot of PC2.

646 **Figure 4.** Maps of Factorial Discriminant Analysis (FDA) performed on the SNV pre-
647 treated MIR spectra (900-1800 cm^{-1}) of purees cooked with: **(a)** three different
648 temperatures (70 °C, 83 °C and 95 °C) and **(b)** three grinding speeds (G0 at 300 rpm,
649 G1 at 1000 rpm and G3 at 3000 rpm); **(c)** the first factorial score (‘F1’) of heating
650 temperature discrimination; **(d)** the first factorial score (‘F1’) of grinding discrimination;
651 **(e)** the second factorial score (‘F2’) of grinding discrimination.

652 **Figure S1.** Overview of MIR spectra pre-processing, direct standardization (SD) and
653 multivariate regression

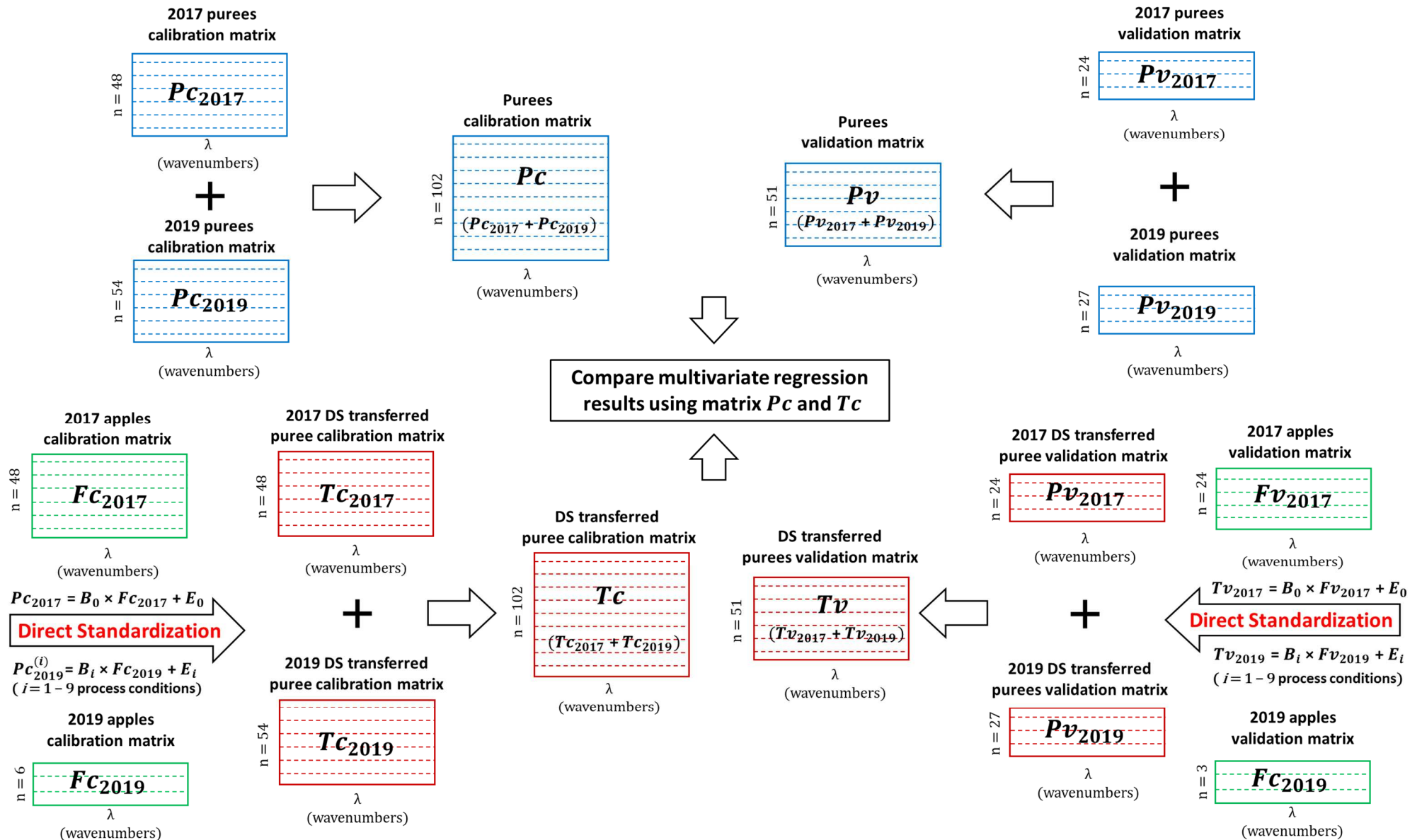
654



655

656

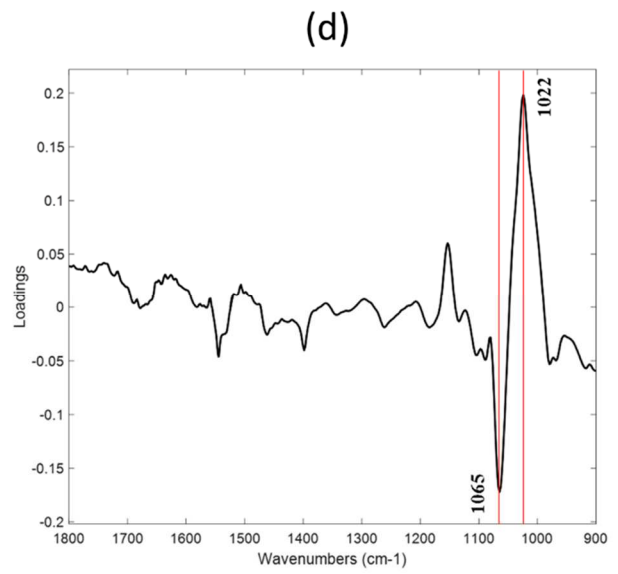
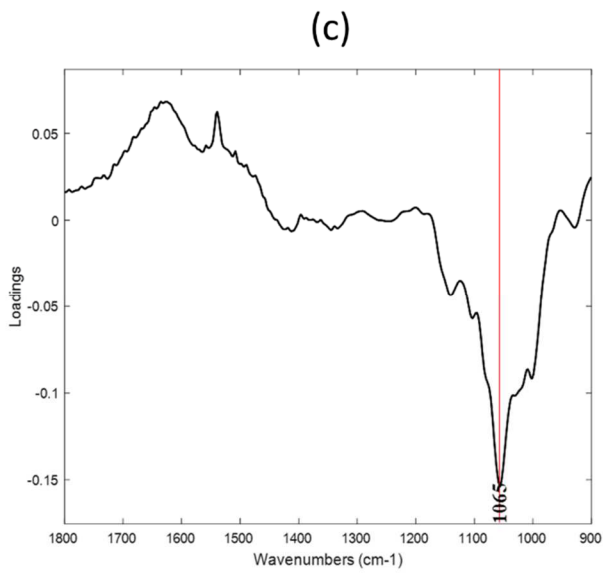
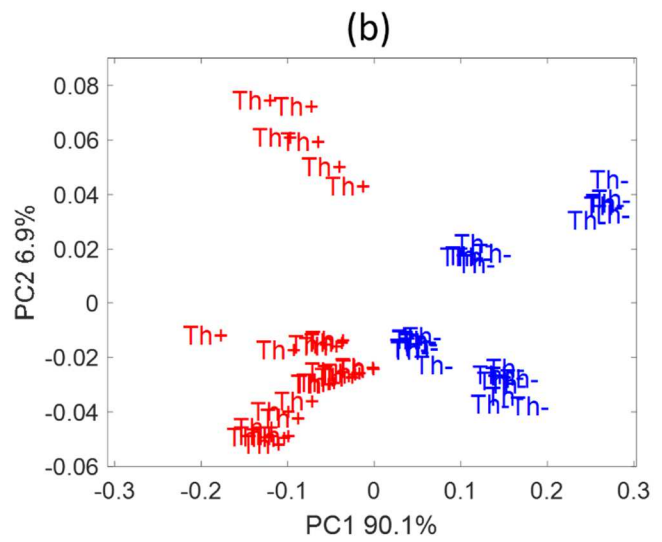
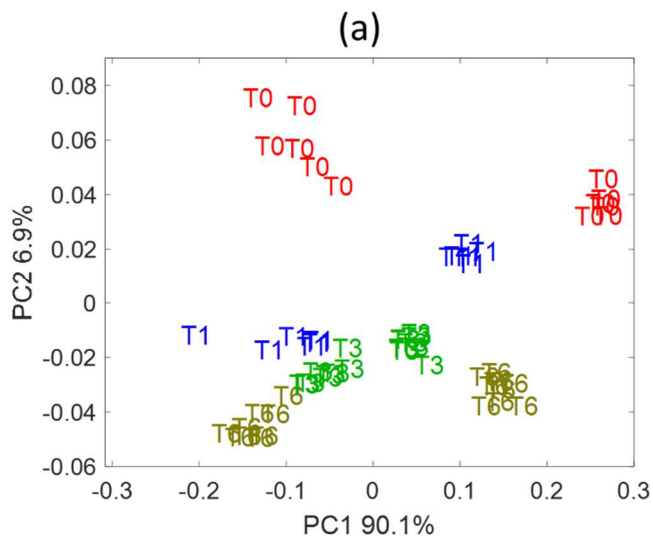
Figure 1



657

658

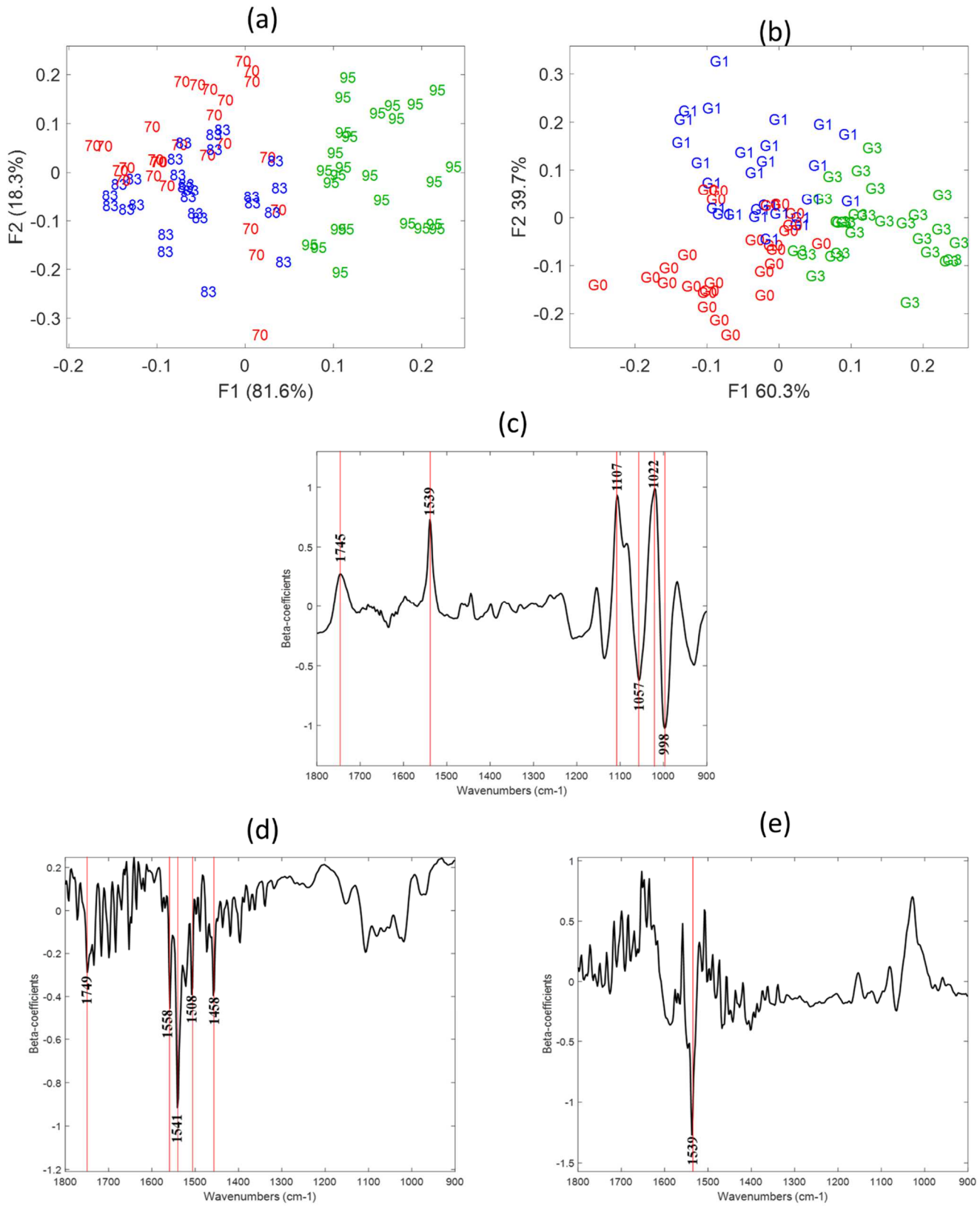
Figure 2



659

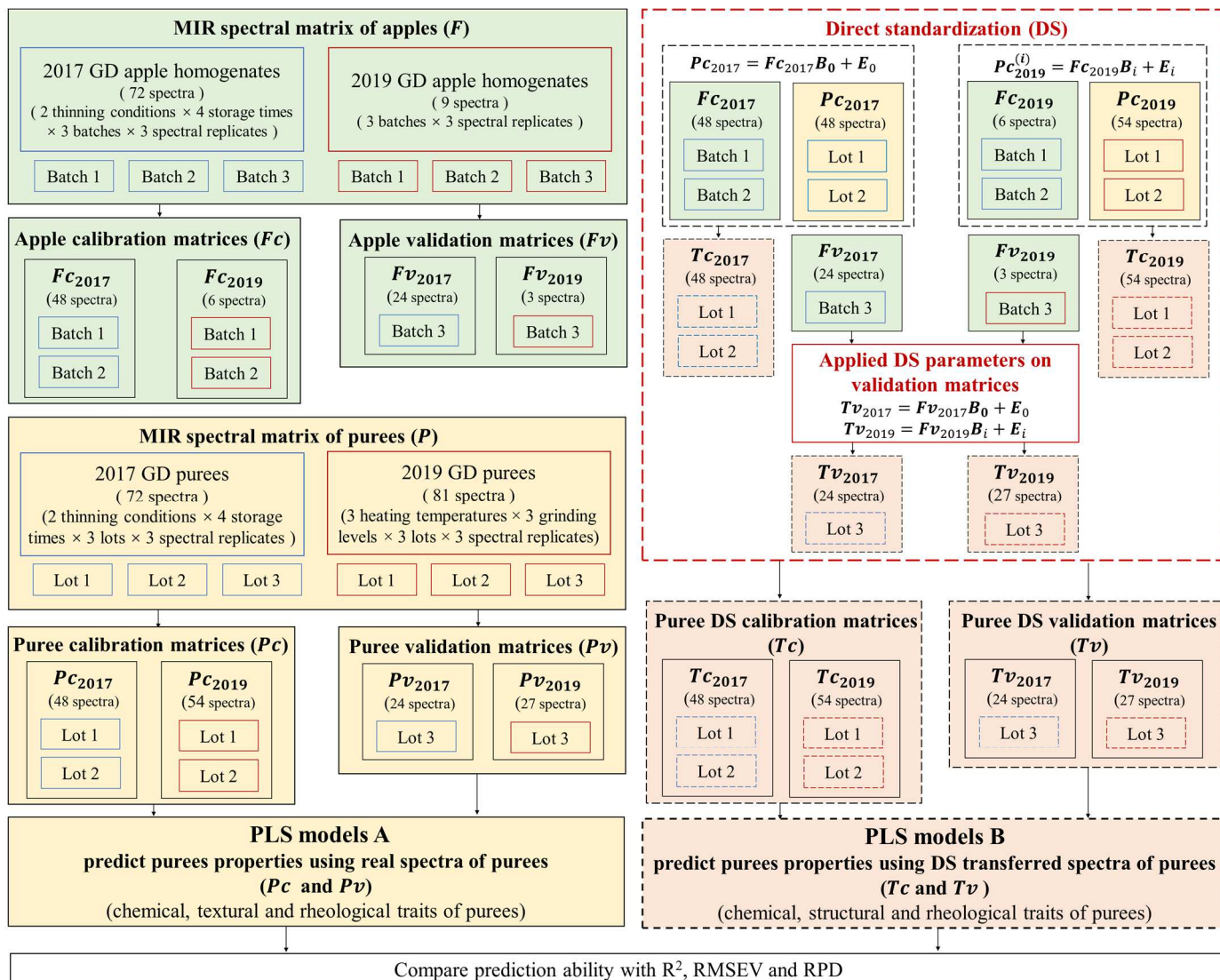
660

Figure 3



661

662 **Figure 4**



663

664

Figure S1

665 **Table 1** Prediction of biochemical, structural and rheological properties of apple purees using PLS regression based on their MIR spectra between 900-1800 cm⁻¹
 666 ¹.

Parameter	Range	SD	R _c ²	R _v ²	RMSEV	RPD	LVs	Linkable regions (cm ⁻¹)
η 50	0.57-2.28	0.39	0.91	0.87	0.19	3.2	7	1780-1788, 1718-1734, 1659-1678, 1636-1616, 1549-1558, 1157-1163
η 100	0.26-1.22	0.27	0.89	0.85	0.07	3.0	8	1780-1788, 1718-1734, 1659-1678, 1636-1616, 1549-1558, 1157-1163
AS-G' (Pa)	1069-11154	2362	0.85	0.75	854	2.0	12	/
AS-G'' (Pa)	210-2707	595	0.87	0.73	298	1.9	12	/
yield stress	8.7-131	27.1	0.86	0.81	12.8	2.1	11	/
d 4:3	196-920	197	0.90	0.87	65	3.1	8	1780-1788, 1701-1713, 1649-1653, 1537-1541, 1101-1107, 1032-1043 1011-1026
d 3:2	44-360	60.1	0.76	0.62	41.7	1.4	11	/
AIS (DM)	133.1-193.6	12.4	0.79	0.50	7.6	1.4	10	/
AIS (FW)	21.8-14.1	3.4	0.80	0.50	2.2	1.5	6	/
DMC (g/g FW)	0.15-0.23	0.02	0.87	0.84	0.01	2.9	5	997-1001,1051-1057, 1101-1109
SSC (°Brix)	12.6-18.6	1.9	0.91	0.86	0.6	3.1	5	997-1001,1051-1057, 1101-1109
TA (mmol H ⁺ /kg FW)	5.1-73.5	25.5	0.89	0.86	7.6	3.2	6	1713-1709, 1105-1109, 1016-1018, 1074-1072, 1038-1042
pH	3.6-4.4	0.2	0.89	0.83	0.1	2.6	7	1713-1709, 1105-1109, 1016-1018, 1074-1072, 1038-1042
malic (g/kg FW)	2.4-7.0	1.2	0.88	0.85	0.4	2.7	6	1721-1709, 1105-1109, 1016-1018, 1074-1072
fructose (g/kg FW)	34.9-98.7	16.0	0.89	0.84	6.1	2.6	9	1709-1713, 1259-1265, 1105-1109, 1074-1080,1038-1042, 1016-1020, 970-974
sucrose (g/kg FW)	39.1-118.5	18.5	0.79	0.64	14.0	1.3	9	/
glucose (g/kg FW)	10.4-25.4	3.6	0.74	0.68	2.4	1.5	7	/

667 Note: Puree spectra and references data from ‘Golden Delicious’ apples, including variability of two different thinning conditions, cold storage (during 0, 1, 3
 668 and 6 months), three heating temperatures (70, 83 and 95 °C) and three grinding levels (300, 1000, 3000 rpm). All results based on the SNV pre-treated MIR
 669 spectra at 900-1800 cm⁻¹. R_c²: determination coefficient of the calibration set; R_v²: determination coefficient of the validation set; RPD: the residual predictive
 670 deviation of validation set; the linkable regions based on the β-coefficients of PLS models with the RPD values higher than 2.5; “/” presented the unacceptable
 671 results with the RPD values lower than 2.5.

672

673 **Table 2** Prediction of biochemical, structural and rheological properties of apple purees using PLS regression based on their reconstructed MIR spectra of raw
 674 apple homogenates between 900-1800 cm⁻¹.

Parameter	Range	SD	R _c ²	R _v ²	RMSEV	RPD	LVs	Linkable regions (cm ⁻¹)
η 50	0.57-2.28	0.39	0.85	0.82	0.21	2.5	8	1720-1734, 1636-1614, 1556-1560, 1547-1533, 1506-1510, 1448-1470, 1157-1169
η 100	0.26-1.22	0.27	0.86	0.83	0.10	2.6	9	1720-1734, 1661-1675, 1636-1616, 1549-1558, 1507-1512, 1445-1468, 1157-1163
d 4:3	196-920	197	0.89	0.84	76	2.6	9	1740-1745, 1701-1715, 1645-1659, 1583-1587, 1537--1541, 1508-1510, 1452-1470, 1100-1112
DMC (g/g FW)	0.15-0.23	0.02	0.87	0.84	0.01	2.6	6	1161-1165, 1101-1107, 1084-1090, 1051-1063, 989-1001
SSC (°Brix)	12.6-18.6	1.9	0.89	0.85	0.7	2.8	5	1101-1112, 1084-1090, 1051-1069, 997-1001
TA (mmol H ⁺ /kg FW)	5.1-73.5	25.5	0.88	0.86	8.9	2.9	7	1715-1710, 1107-1113, 1082-1086, 1059-1063, 1038-1042, 1001-993
pH	3.6-4.4	0.2	0.84	0.79	0.1	2.1	8	1715-1709, 1105-1110, 1016-1018, 1074-1072, 1038-1042
malic (g/kg FW)	2.4-7.0	1.2	0.87	0.82	0.6	2.3	9	1713-1709, 1105-1109, 1080-1088, 1058-1064, 1016-1018, 1001-998
fructose (g/kg FW)	34.9-98.7	15.0	0.82	0.72	8.9	1.7	11	/

675 Note: Puree spectra and references data from ‘Golden Delicious’ apples, including variability of two different thinning conditions, cold storage (during 0, 1, 3
 676 and 6 months), three heating temperatures (70, 83 and 95 °C) and three grinding levels (300, 1000, 3000 rpm). All results based on the SNV pre-treated MIR
 677 spectra at 900-1800cm⁻¹. R_c²: determination coefficient of the calibration set; R_v²: determination coefficient of the validation set; RPD: the residual predictive
 678 deviation of validation set; the linkable regions based on the β-coefficients of PLS models with the RPD values higher than 2.5; “/” presented the unacceptable
 679 results with the RPD values lower than 2.5.

680

Table S1 Biochemical, structural and rheological data of apple purees and ANOVA results.

Fruit thinning	Storage periods	η_{50} Pa.s	η_{100} Pa.s	G' Pa	G'' Pa	Yield stress Pa	d 4:3 -	d 3:2 -	DMC g/g FW	SSC °Brix	TA mmol H+/kg FW	glucose g/kg FW	fructose g/kg FW	sucrose g/kg FW	malic acid g/kg FW	pH	AIS mg/g FW	AIS mg/g DW
Th-	T0	1.28	0.77	3127.8	626.7	47.5	909.9	251.5	0.19	13.4	58.1	18.9	50.5	66.7	4.5	3.7	164.5	31.6
	T1	1.13	0.70	1960.2	466.7	21.9	694	351.9	0.19	15.0	54.4	15.4	49.4	59.1	2.8	3.8	147.2	27.6
	T3	0.87	0.55	1849	453	13.9	339.8	205.9	0.20	14.1	46.7	18.6	84.1	84.8	3.6	4.0	140.0	27.3
	T6	0.92	0.50	1816	427	14	316.1	223.6	0.19	13.8	26.8	23.0	85.1	77.3	2.7	4.4	145.7	27.6
Th+	T0	1.75	0.97	3375.1	816.4	52.1	831.6	231.6	0.21	15.5	70.9	23.5	85.3	64.4	5.5	3.6	163.3	33.9
	T1	1.54	0.94	2783.7	639.5	25.2	489	261.8	0.21	17.6	69.3	16.8	80.3	115.9	5.6	3.8	150.9	31.9
	T3	1.25	0.70	2517.6	609	22.3	405.1	228.3	0.22	16.9	59.9	13.8	88.0	102.5	4.9	3.8	143.3	31.6
	T6	1.60	0.88	3168.2	751.7	33.9	393.5	255.1	0.23	17.5	34.7	23.8	95.7	44.0	3.6	4.3	150.3	34.8
Storage time	significance	***	***	***	***	***	***	***	**	**	***	***	***	***	***	***	ns	*
	F-values	20.4	13.6	24.7	15.0	72.8	216.1	41.5	6.6	8.3	279.4	74.7	76.2	38.5	13.8	436.3	1.8	4.3
Fruit thinning	significance	***	***	***	***	***	***	*	***	***	***	ns	***	**	***	***	**	**
	F-values	138.2	61.5	67.3	91.5	29.6	47.2	5.5	157.7	115.7	176.9	1.3	187.8	13.8	57.9	48.8	15.9	10.3

Data expressed in fresh weight (FW) or dry weight (DW); values correspond to the mean of 3 puree replications (3 kg per replication). Raw apples were stored at 4°C: from harvest (T0)

and during one (T1), three (T3) and six months (T6). Two conditions of fruit load during cultivation: non-thinning with 100% number of apples (Th-) and thinning with 50% number of

apples (Th+) per tree. In grey, two way- ANOVA results obtained for Golden Delicious purees. ns, *, **, ***: Non significant or significant at P < 0.05, 0.01, 0.001 respectively.

Table S2 Biochemical, textural and rheological data of apple purees and results of Kruskal-Wallis non-parametric test.

Temperatures	Grinding speeds	η_{50}	η_{100}	G'	G''	Yield stress	d 4:3	d 3:2	DMC	SSC	TA	glucose	fructose	sucrose	malic acid	pH	AIS	AIS
°C	rpm	Pa.s	Pa.s	Pa	Pa	Pa	-	-	g/g FW	°Brix	mmol H+/kg FW	g/kg FW	g/kg FW	g/kg FW	g/kg FW		mg/g FW	mg/g DW
70	300	1.27	0.89	9629.8	2389.5	97.7	583.8	103.6	0.17	13.8	63.6	16.1	67.9	70.9	6.3	3.9	26.8	158.0
	1000	1.42	0.91	3295.4	768.8	38.8	633.7	274.8	0.16	13.8	67.8	18.0	66.2	76.6	6.4	3.9	28.0	171.4
	3000	0.64	0.40	1111.3	215.6	10.2	353.5	207.6	0.17	14.4	69.1	17.3	65.5	86.0	6.6	3.9	28.1	165.1
83	300	1.23	0.93	8437.5	2078.9	97.9	553.0	212.8	0.18	14.9	59.9	14.6	67.7	70.6	5.8	3.8	29.5	166.7
	1000	1.38	0.87	3036.9	764.4	35.6	647.8	324.8	0.16	13.4	69.8	17.4	71.9	75.2	5.2	3.9	25.6	159.6
	3000	0.91	0.54	1312.2	259.9	11.8	297.4	192.4	0.17	14.8	70.3	16.6	66.8	76.1	5.7	3.9	28.2	162.6
95	300	1.93	1.16	3708.2	1101.3	30.0	492.9	262.1	0.17	14.9	60.2	17.0	72.6	66.9	4.8	3.9	27.9	160.8
	1000	1.44	0.84	1955.4	522.2	16.6	332.9	209.8	0.17	14.9	64.1	17.8	71.3	64.5	5.1	3.9	26.2	157.3
	3000	1.07	0.63	1399.4	362.2	14.0	206.4	153.2	0.17	14.8	62.0	18.9	69.2	65.8	5.4	3.8	26.2	154.1
Temperature	significances	<i>ns</i>	<i>ns</i>	<i>ns</i>	<i>ns</i>	<i>ns</i>	*	<i>ns</i>	*	*	*	*	*	**	**	<i>ns</i>	<i>ns</i>	<i>ns</i>
	F-values	4.1	1.3	0.9	0.4	1.5	7.0	1.2	6.3	6.8	8.4	6.1	8.9	9.1	10.2	2.1	2.3	4.2
Grinding speeds	significances	**	***	***	***	***	**	**	<i>ns</i>	<i>ns</i>	<i>ns</i>	<i>ns</i>	<i>ns</i>	<i>ns</i>	<i>ns</i>	<i>ns</i>	<i>ns</i>	<i>ns</i>
	F-values	10.9	17.2	21.6	22.6	19.4	13.7	9.2	1.1	2.8	1.4	0.4	0.1	1.4	0.2	1.3	0.4	0.2

687 Data expressed in fresh weight (FW) or dry weight (DW); values correspond to the mean of 3 puree replications (3 kg per replication). Processing conditions variations were: three
688 heating temperatures at 70°C, 83°C and 95°C for 30 min, and three grinding speeds at 300, 1000 and 3000 rpm at each temperature. In grey, Kruskal-Wallis results obtained on Golden
689 Delicious purees. ns, *, **, ***: Non significant or significant at P < 0.05, 0.01, 0.001 respectively.

690 **Table S3** The results of sensitivity (in blue cells) and specificity (in yellow cells) from: **(a)** the FDA discrimination (4 factors) of three different heating
 691 temperatures; and **(b)** the FDA discrimination (4 factors) of three different grinding speeds.

692 **(a)**

Temperatures (°C)	70	83	95
70	/	76.67%	100%
83	83.33%	/	100%
95	100%	100%	/

693 **(b)**

Grinding speeds (rpm)	300	1000	3000
300	/	75.00%	85.19%
1000	76.92%	/	87.10%
3000	85.19%	100%	/

694

695