Accepted Manuscript

Prediction of cooking times of freshly harvested common beans and their susceptibility to develop the Hard-To-Cook defect using near infrared spectroscopy

Elizabeth N. Wafula, Irene N. Wainaina, Carolien Buvé, Peter K. Kinyanjui, Wouter Saeys, Daniel N. Sila and Marc E.G. Hendrickx

PII: S0260-8774(21)00020-0

DOI: [10.1016/j.jfoodeng.2021.110495](https://doi.org/10.1016/j.jfoodeng.2021.110495)

Reference: JFOE_110495

To appear in: *Journal of Food engineering*

Please cite this article as: Elizabeth N. Wafula, Irene N. Wainaina, Carolien Buvé, Peter K. Kinyanjui, Wouter Saeys, Daniel N. Sila and Marc E.G. Hendrickx (2021). Prediction of cooking times of freshly harvested common beans and their susceptibility to develop the Hard-To-Cook defect using near infrared spectroscopy. *Journal of Food Engineering*, (110495). <https://doi.org/10.1016/j.jfoodeng.2021.110495>.

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

**Prediction of cooking times of freshly harvested common beans and their
susceptibility to develop the Hard-To-Cook defect using near infrared
spectroscopy**

**WAFULA Elizabeth Nakhungu^{1,2*}, WAINAINA Irene Njoki^{1,2}, BUVÉ Carolien¹,
KINYANJUI Peter Kahenya², SAEYS Wouter³, SILA Daniel Ndaka² and HENDRICKX
Marc^{1**}**

Current affiliations:

¹*KU Leuven, Department of Microbial and Molecular Systems (M²S), Laboratory of Food
Technology, Kasteelpark Arenberg 22, Box 2457, 3001 Leuven, Belgium*

²*Jomo Kenyatta University of Agriculture and Technology, College of Agriculture and Natural
Resources, School of Food and Nutritional Sciences, Department of Food Science and Technology,
P.O. Box 62,000-00200, Nairobi, Kenya*

³*KU Leuven, Department of Biosystems (BIOSYST), MeBioS Division, Kasteelpark Arenberg 30,
Box 2456, 3001 Leuven, Belgium*

***Author to whom correspondence should be addressed during submission process**

Telephone: +254 717 557 850

Email: elizabethnakhungu.wafula@kuleuven.be

****Author to whom correspondence should be addressed on post publication**

Telephone: +32 16 32 15 72

Email: marceg.hendrickx@kuleuven.be

Abstract

The cooking time of common beans is influenced by genotype and storage conditions. This study aimed to use near-infrared (NIR) spectra of milled, freshly harvested (fresh) beans to predict their cooking times and their susceptibility to develop the storage-induced, hard-to-cook (HTC) defect. The physical characteristics of bean accessions, from two different seasons, were evaluated. The cooking times and susceptibility to HTC (determined by aging under standard adverse conditions) of the aforesaid beans were correlated to NIR spectra to develop calibrations using variable selection and partial least squares regression. The beans exhibited diverse physical characteristics, cooking times and susceptibility to HTC. The models predicting cooking times of fresh beans sufficiently overcame genotype and seasonal differences ($R^2_p=0.73$, RMSEP=4 minutes). The susceptibility of fresh beans to HTC was also successfully predicted ($R^2_{cv}=0.8$, RSECV=64%). NIR spectroscopy has high potential to rapidly identify beans with short cooking times and low susceptibility to HTC at harvest.

Key Words: Near-Infrared spectroscopy, Susceptibility, Hard-To-Cook, Common beans, Aging, Cooking times

List of Abbreviations and Symbols

HTC -Hard-To-Cook

iPLS - Interval partial least squares

PLS-R - Partial least squares regression

RMSEC - Root mean square error of calibration

RMSECV - Root mean square error of cross validation

- 43 **RMSEP** - Root mean square error of prediction
- 44 **R²_C** - coefficient of determination of calibration
- 45 **R²_{CV}** - coefficient of determination of cross validation
- 46 **R²_P** - coefficient of determination of prediction
- 47 **LV** - latent variables
- 48 **MC** - mean centering
- 49 **SNV** - standard normal variate
- 50 **OSC** - orthogonal signal correction
- 51 **DH** - de-hulled
- 52 **WSC** - with seed coat

53 **1. Introduction**

54 Common beans (*Phaseolus vulgaris* L.) are relatively affordable legumes that provide essential
 55 nutrients and when consumed regularly, render health benefits (Rebello et al., 2014; Guevara-
 56 González et al., 2006; Broughton et al., 2003). In addition, common beans are widely accepted
 57 (Broughton et al., 2003) due to their “multiple means of preparation”, “global adaptability” and
 58 “genotypic and phenotypic diversity” (Uebersax et al., 1991). The wide range of common bean
 59 seed characteristics in terms of weight, size and color influences farmers’ and consumers’
 60 preferences, and eventually their commercial viability. Thus, the assessment of physical
 61 characteristics has been an essential part of bean quality evaluation and/or development (breeding)

(Sinkovic et al., 2019; Wang et al., 2017; Blair et al., 2010; Piergiovanni et al., 2000; Wortmann et al., 1998).

As raw beans are unpalatable, cooking is necessary for reducing/eliminating (toxic) anti-nutrients, gelatinizing starch, denaturing proteins and solubilizing polysaccharides in the middle lamella and cell wall (Shiga et al., 2004; Stanley, 1992). Pectin solubilization has been identified as one of the rate limiting steps during bean softening (Chigwedere et al., 2018). The structure of pectin, which affects its solubility, is influenced by both variety and storage conditions (Njoroge et al., 2015). The latter dictate the overall cooking times of beans (Kinyanjui et al., 2017; Siqueira et al., 2016; Njoroge et al., 2015; Coelho et al., 2007). The conditions under which beans are grown also affect their cooking times (Wang et al., 2017). Accordingly, processors and consumers prefer beans with short cooking times at harvest and good conservation of cooking quality during storage (Katungi et al., 2011; Bressani, 1993).

The deterioration of cooking quality usually occurs when beans are stored for extensive periods at high temperatures (above 25 °C) and high relative humidity (above 65%). These conditions catalyze enzymatic and non-enzymatic reactions which lead to physico-chemical deterioration (Hohlberg and Stanley, 1987) altogether known as aging (Chigwedere et al., 2019). One undesirable outcome of aging on the texture of beans is the hard-to-cook (HTC) defect. The pathway leading to the development of HTC has not yet been fully elucidated. Different mechanisms have been suggested to explain the HTC phenomenon: cell membrane disruption, pectin-cation-phytate, lignification like and, mixed reaction theories and protein denaturation which impedes starch gelatinization (Galiotou-panayotou et al., 2008; Garcia et al., 1998; Kilmer et al., 1994; Aguilera and Rivera, 1992; Hentges et al., 1991; Garcia-Vela and Stanley, 1989; Hohlberg and Stanley, 1987; Jones and Boulter, 1983). The HTC defect leads to long cooking

times even after prolonged soaking of beans which results in undesirable high resource consumption (Galiotou-panayotou et al., 2008; Garcia et al., 1998; Stanley, 1992).

Consequently, there is a need to not only identify beans with short cooking times at harvest, but also to evaluate the susceptibility of beans to develop the HTC defect. This calls for a rapid and accurate protocol which can expeditiously assess a large number of bean varieties/genotypes in research and industry contexts (Wood, 2017; Siqueira et al., 2016). Near-infrared (NIR) spectroscopy might fulfill these needs as it has been widely investigated for postharvest quality evaluation of agrofood products such as fruit and vegetables (Nicolai et al., 2007; Walsh et al., 2020) and is a rapid and non-destructive analytical tool (Pasquini, 2018; Williams, 2004). The NIR spectra are characterized by broad absorption bands related to overtones and combinations of covalent bonds (C-H, O-H and N-H) which can be correlated to the quality attributes of interest using multivariate data analysis techniques such as partial least squares regression (PLS-R) to predict quality traits in future samples (Saeys et al., 2019).

Macro- and micro-nutrients, anti-nutrients, anti-oxidants, sensory properties and canning quality of beans have been previously predicted from the spectra of milled beans with their seed coats, milled cotyledons only, milled seed coats only, intact beans, cooked mashed beans and, cooked beans dried at 55°C (Carbas et al., 2020; Mendoza et al., 2014; Wang et al., 2014; Plans et al., 2013b; 2013a, 2012). Additionally, the work of Cichy et al., (2015) attempted to predict the cooking times of fresh beans using VIS/NIR spectra of intact beans, but it was put forward that the calibration obtained needed further improvement. The cooking times of soaked and unsoaked beans have also been predicted using hyperspectral imaging (Mendoza et al., 2018). However, the HTC phenomenon has so far been given limited attention.

In Wafula et al. (2020), we reported on the successful use of NIR spectroscopy to predict the cooking times of a limited number of aged bean varieties from the spectra of milled, aged beans. The aim of this study was to improve and extend the aforementioned protocol to develop PLS-R models to predict cooking times for a wide range of fresh bean accessions planted during two different seasons and, in addition we aimed to develop a model to predict the susceptibility of the bean accessions to develop the HTC defect during storage, using spectra acquired from milled fresh beans. To the best of our knowledge, this is the first attempt to develop such NIR spectroscopy based model.

2. Materials and methods

2.1 Source of fresh beans

The research activities undertaken in this study are schematically illustrated in **Figure 1**. Seeds of seventy (70) common bean accessions were obtained from the Genetic Research Institute (GRI) in Muguga, Kenya in 2017 and farmers around Kenya. The bean accessions belonged to twelve bean types including sugar beans (13), rose coco (13), black beans (10), carioca (9), red haricot (6), navy (5), kidney (4), pinto (4), mwezi moja (3), yellow beans (1), brown and tan beans (1) and zebra beans (1). The beans were planted in February, 2018 at Jomo Kenyatta University of Agriculture and Technology farm in Juja, Kenya and harvested in June, 2018. The harvested beans were transferred to a greenhouse where they were dried, threshed, and sorted to remove all unwanted materials. The clean beans were conditioned in a thermostatic incubator at 25 °C and 75% relative humidity (rh) maintained by a saturated sodium chloride solution, for two weeks. This batch of planting was referred to as Season 1.

The second batch (Season 2) included planting of seventeen (17) of the Season 1 bean accessions. These accessions belonged to nine bean types which were sugar beans (3), black beans (3), rose

coco (3), carioca (3), kidney beans (1), red haricot (1), zebra beans (1), pinto (1) and navy beans (1). The selected accessions had exhibited median cooking times for soaked beans within the different bean types. Three additional bean varieties comprising Nabe 4 (carioca), Nabe 15 (rose coco) and Nabe 17 (carioca) sourced from South Sudan were also planted. The planting was carried out in November 2018 at the same site as the first planting and harvesting took place in February 2019. The beans were treated following the same procedure as the beans harvested during Season 1. The bean accessions obtained from Season 2 were divided into two batches. Experiments were performed on the first batch as fresh beans, while the second batch was aged.

2.2 Aging of the bean accessions to induce HTC

Due to limited availability of the beans, only eighteen (18) of the Season 2 bean accessions were aged at 35 °C and 83% rh, maintained by a saturated potassium chloride solution, for three months in a thermostatic incubator. These conditions have been previously shown to lead to HTC development (Chigwedere et al., 2019; Kinyanjui et al., 2017; Njoroge et al., 2016). Upon completion of the aging experiment, the beans were stored at -25 °C awaiting further experiments.

2.3 Physical characterization of fresh beans

Three physical characteristics of beans were assessed. The 100-seed weight was determined by randomly selecting one hundred seeds from each bean accession and weighing the seeds in triplicate. The seed dimensions were measured on 10 seeds per accession using a vernier caliper. The length (the highest dimension parallel to the seed's hilum), the thickness (the lowest dimension parallel to the hilum) and, the width (the dimension perpendicular to the hilum) were evaluated. The L^* a^* b^* values of the seed coats of intact beans were also determined using a bench-top colorimeter with xenon flash lamp illuminant, a geometry of 45°/0° and observer angles of 2° and 10° (Model: Hunter Lab's Directional 45/0 LAV Color Spectrophotometer ColorFlex Ez, USA).

This parameter was measured in triplicate by taking a reading of one spot of three separate half-way filled measuring cups.

2.3.4 Statistical analysis of physical characteristics of beans

Statistical analysis of the physical characteristics was done using the statistical software JMP®, version 15 (SAS Institute Inc., Cary, North Carolina, USA, 1989-2019). Data were expressed as mean and standard deviation. A one-way ANOVA was used to test the significance of bean type and bean accession on the physical characteristics of Season 1 and Season 2 beans, respectively. The Tukey's HSD test was used to compare differences between the bean type and bean accession means. Whereas the Student t-test was used to compare differences between the average values of the 17 bean accessions planted in both seasons 1 and 2. The means were considered significantly different at 5% significance level.

2.4 Determination of cooking times of fresh and aged beans

The cooking kinetic data were acquired using the finger pressing method according to Kinyanjui et al., (2017, 2015). This method allowed for the determination of the number of cooked beans as a function of time. The bean accessions were divided into two batches. One batch was cooked without any soaking pre-treatment thus referred to as not soaked beans. The other batch was soaked in distilled water, in a ratio of 1:5 (w/v), at 25 °C for 16 hours before cooking and denoted soaked beans. Cooking of beans was carried out in distilled water at 95 °C at an altitude of 1416 m above sea level. Ten seeds were taken out after intervals of 5, 10 and 30 minutes for soaked fresh beans, not soaked fresh beans and soaked aged beans, respectively, without interrupting boiling. The beans were cooled in running water for one minute, pressed by exerting moderate pressure using the forefinger and thumb and, classified as cooked when the cotyledons disintegrated. These data were collected in duplicate by an individual whose finger pressure had been calibrated against a

cutting force of 11 N using a Sun-Rheometer (Compact 100 Model CR-100; Sun Scientific Company Ltd, Tokyo, Japan). This cutting force had been previously stipulated as the level of hardness/softness a bean was to obtain in order to be classified as cooked (Kinyanjui et al., 2015).

The relationship between time of cooking (T) and the percentage of cooked beans (Y) was modelled using logistic regression, as explained in Wafula et al. (2020). The logistic regression model is described by **Equations 1** and **2**. **Equation 1** represents a simple logit transform, with time of cooking (T) as the only explanatory variable. The antilog of **Equation 1** gives the probability that a bean is cooked after a given time of cooking (**Equation 2**). In this study, sufficient cooking (indicated as ' $t_{95\%}$ ') was defined as the time corresponding to a probability that 95% of the beans would be cooked. In **Equation 1**, π represents the probability that a bean will be cooked, α represents the Y intercept, β represents the regression coefficient which indicates the direction and scale to which a unit change in the time of cooking affects the log odds of the probability that the bean will be cooked; $|$ is the conditional operator and, $|T$ is the range of cooking times (0-300 minutes). While t is one specific time of cooking, which is an independent continuous variable, Y is the percentage of cooked beans, which is a dependent variable (Kutner et al., 2005; Peng et al., 2002). The R Studio (version 3.5.2, 2018) software was used to carry out the statistical analysis.

$$\text{Logit } (Y) = \text{natural log(odds)} = \ln\left(\frac{\pi}{1-\pi}\right) = \alpha + \beta T \quad \text{Equation 1}$$

$$\pi = \text{Probability } (Y = \text{outcome} | T = t, \text{ a specific value of time to cook}) = \frac{e^{\alpha + \beta t}}{1 + e^{\alpha + \beta t}} \quad \text{Equation 2}$$

The susceptibility of the beans to develop HTC was defined in two ways: (i) the absolute difference between cooking time of the aged and fresh beans ($\Delta t_{95\%, \text{ abs}}$, **Equation 3**) and (ii) the relative increase in cooking time ($\Delta t_{95\%, \text{ rel}}$) of aged beans compared to fresh beans (**Equation 4**).

198 $\Delta t_{95\%, abs} = \text{Cooking time of aged beans} - \text{Cooking time of fresh beans}$ Equation 3

199 $\Delta t_{95\%, rel} = \frac{\text{Cooking time of aged Beans} - \text{Cooking time of fresh Beans}}{\text{Cooking time of fresh Beans}} \times 100\%$

200 Equation 4

201 2.5 Acquisition of Fourier Transform (FT) NIR spectra of fresh beans

202 The collection of spectral data was analogous to and the development of PLS-R models (Section
203 2.6) was based on slight modifications of the method described in Wafula et al., (2020). Since it
204 has been demonstrated that the cotyledon and seed coat had major and minor contributions to the
205 development of HTC, respectively (Chigwedere et al., 2018), the effect of de-hulling prior to
206 spectral measurements was investigated. All Seasons 1 and 2 fresh bean accessions were divided
207 into two sets. The first set of beans were milled with their seed coats using a micro mill (IKA A10
208 Basic, IKA(r)-Werke GmbH & Co. KG, Staufen, Germany). These are further on referred to as
209 'with seed coat' (WSC) samples. The seed coats of the second set of beans were removed manually
210 before milling, resulting in the samples denoted as 'de-hulled' (DH). Afterwards, the samples were
211 sieved using a 1 mm mesh, placed in petri-dishes and transferred to airtight cabinets containing
212 desiccant to condition at room temperature for 10 days.

213 NIR spectra were acquired with a Multi-Purpose Analyzer FT NIR spectrophotometer (MPA-FT
214 Bruker Optics, Ettlingen, Germany) as described by Wafula et al., (2020). Seven sub-samples, or
215 replicates, per bean accession were transferred into transparent glass vials of 2.2 cm diameter
216 which were filled to a height of 1 cm. They were individually scanned on the integrating sphere in
217 reflectance mode over the wavenumber range from 3,595 to 12,490 cm^{-1} with a resolution of 8 cm^{-1} .
218 ¹. Each acquired spectrum was the average of 64 scans. The data were converted from

wavenumbers to wavelengths and reduced to the 1,000 – 2,600 nm range to remove noisy regions at the edges of the spectrum (Conzen, 2014).

2.6 Development of PLS-R models for predicting cooking times

The spectra of each sample type, DH or WSC, were grouped into the following data sets; Seasons 1+2: DH beans, Seasons 1+2: WSC beans, Season 2: DH beans, and Season 2: WSC beans. Mean centering followed by principal component analysis (PCA) was carried out on all spectra/replicates of each data set. This was done to explore the data and to detect outliers based on the Hotelling T^2 and Q residual plots (Pojić et al., 2010; Saeys et al., 2019). An outlier/replicate was removed only if the correct reference value (cooking time/ $\Delta t_{95\%, abs}$ / $\Delta t_{95\%, rel}$) had been assigned and yet the spectrum was visually different from other replicates from the same sample.

In a next step, each of the Seasons 1+2 data sets was divided into a calibration set (81 samples) and a test set (9 samples, randomly selected from the initial 90 samples). The test set, for which values are italic in **Tables A.1 and A.2 in Appendix A**, represented variability in terms of bean accession, bean type, season and cooking time. This resulted in one (1) DH beans PLS-R model and one (1) WSC PLS-R model to predict cooking times of fresh soaked beans. Since only eighteen of the Season 2 bean accessions were aged, which was a low number of samples, all the bean samples were included in the Season 2 calibration sets. This resulted in two (2) DH beans and two (2) WSC beans PLS-R models to predict susceptibility to develop HTC in terms of $\Delta t_{95\%, abs}$ and $\Delta t_{95\%, rel}$.

To construct the PLS-R models, the replicate spectra for each sample were initially averaged (Williams, 2004). The reliability of the models was tested by leave 3-samples-out (4% of data) and leave-1-sample-out (6% of data) in the cross validation for the Seasons 1+2 data sets and Season 2 data sets, respectively. Forward and reverse interval partial least squares (iPLS) with window

widths of either 50 or 100 variables were explored on mean centered (MC) spectra (Andersen & Bro, 2010; Nørgaard et al., 2000). PLS-R was performed to correlate the NIR spectra (X) to the cooking time/ $\Delta t_{95\%}$, *abs*/ $\Delta t_{95\%}$, *rel* (Y) (Saeys et al., 2019; Naes et al., 2002). Subsequent pre-processing suitable for use after variable selection included standard normal variate (SNV) and orthogonal signal correction (OSC) with mean centering. The root mean square error (RMSE) of cross validation was used to select the number of latent variables (LV). The LV with the lowest RMSE with at least 2% reduction in RMSE compared to the preceding one was selected. The model performances were evaluated using the coefficient of determination (R^2) and the RMSE in cross validation and prediction. Useful models are those with high R^2 values and low RMSE values (Saeys et al., 2019; Naes et al., 2002).

Solo version 8.7, the standalone version of the PLS toolbox for Matlab (Eigenvector Research Inc., Manson, USA) was used for building the PLS-R models.

3. Results and Discussion

3.1 Physical characteristics of fresh common beans

The physical characteristics of the beans were assessed to establish their phenotypic variability. The results obtained for the different bean types/accessions planted during Seasons 1 and 2 are summarized in **Tables 1 and 2**, respectively. As it can be seen from **Table 1**, bean type had a significant effect on all the physical characteristics assessed. In terms of 100-seed weight, navy beans were the lightest (21.79 ± 6.94 g), whereas zebra beans were the heaviest (66.86 ± 0.53 g). According to Blair et al., (2010), bean accessions can be classified into three size categories depending on 100-seed weight including small (<25 g), medium (>25 g, <45 g) or large (>40 g). The small-seeded bean types were black beans, navy beans and red haricot, the medium-seeded ones were pinto and brown beans and the majority of the bean types including carioca, kidney

beans, mwezi moja, rose coco, sugar beans, yellow beans and zebra beans were large-seeded. From the results obtained for seed dimensions, the navy bean type was the shortest (8.41 ± 0.23 g) and the black beans were both the smallest in thickness (4.78 ± 0.69 g) and width (6.80 ± 1.01 g). While the zebra bean type was the longest and the yellow bean type presented the largest thickness and width. The findings on seed dimensions and 100-seed weight were on average similar to values reported for Kenyan beans (Martín-Cabrejas et al., 1997), higher and lower than the values reported for Rwandese beans (Blair et al., 2010) and, Slovenian and Italian beans (Sinkovic et al., 2019; Piergiovanni et al., 2000), respectively. The results for surface color indicated that black and navy beans had the lowest (24.39 ± 3.59) and highest (70.71 ± 3.17) L^* values, respectively. Navy beans had the lowest a^* values (-0.05 ± 0.76) whereas the kidney bean type had the highest a^* values (18.40 ± 0.85). Also, b^* values were lowest (0.69 ± 0.48) for black bean type and highest for brown beans (19.47 ± 1.52), which were significantly different from yellow, sugar, carioca and pinto bean types. The color values in this study were different from those reported in other studies for kidney beans, black beans and carioca beans (Arns et al., 2018; Parmar et al., 2016; Cichy et al., 2014). Bean accession significantly influenced all the physical characteristics of Season 2 beans. Furthermore, there were similar trends with respect to how bean types rank within the physical characteristics between Season 1 and Season 2.

Upon comparing the average values of the physical characteristics for the 17 bean accessions planted in both Seasons 1 and 2, Season 1 beans had significantly higher values for 100-seed weight, thickness and width (**Table 2**). Season 1 coincided with long rainy period with lower temperatures which likely delayed maturity allowing for the accumulation of nutrients such as starch, hence resulting in heavier and larger seeds. On the other hand Season 2 coincided with a short rainy period with higher temperatures. Syanda (2019) also reported significant differences

between the length, width and thickness of beans planted during long and short rains seasons (Syanda, 2019). The similarities and disparities observed for the morphological characteristics in the present study and the other studies cited here could be attributed to the influence of genotype/variety, growth conditions and location and, chemical composition (Wang et al., 2017; de Barros and Prudencio, 2016).

3.2 Cooking times of fresh common beans

Figure 2 visually depicts a summary of the cooking times obtained for the Seasons 1 and 2 fresh beans, which can be found in **Appendix A (Tables A.1 and A.2)**. The cooking times for not soaked beans harvested in Season 1 ranged from 62 minutes for a black bean accession to 140 minutes for a pinto bean accession. It was observed that the pinto bean type had the highest variability in cooking times for not soaked beans (**Figure 2A**). However, after soaking this variability decreased considerably as demonstrated by the cooking curves of the soaked pinto beans which are almost superimposed on each other (**Figure 2B**). This phenomenon of a reduced range of cooking times for soaked beans was observed for all the 12 bean types from 23 minutes for an accession of black beans to 57 minutes for a rose coco bean accession. The cooking times of Season 2 not soaked beans varied from 68 minutes for a black bean accession to 130 minutes for a zebra bean accession (**Figure 2C**). The shortest cooking time for soaked beans in Season 2 was 33 minutes for a navy bean accession and the longest cooking time was 50 minutes for carioca and pinto accessions. These observations are in line with the findings of Kinyanjui et al., (2015) who also found higher cooking rates for soaked beans.

The reduced variability in cooking times after soaking could be explained as follows. Upon application of heat, two steps occur when beans have not been soaked, hydration followed by softening. After soaking, the beans are already hydrated and only the softening still has to take

place. The rate of hydration of the seed coat depends on its thickness and fiber content, due to fiber's high water holding capacity, which are in turn affected by genotypic variability of the beans. Therefore, soaking beans overcomes the limitation associated with the seed coat as a barrier to hydration (Wood, 2017). In addition, the differences in cooking times exhibited by soaked beans could likely be attributed to disparities in the solubility and structure of pectin. It has been previously established that beans with shorter cooking times have higher contents of loosely bound pectin (water soluble pectin), less branched pectin, shorter pectin polymers, lower amounts of high molecular mass polymers and a higher degree of acetylation (Njoroge et al., 2014; Bernal-Lugo et al., 1997).

The correlation between the cooking times obtained for the seventeen accessions planted in both Seasons 1 and 2 was 0.67 for soaked beans (**Figure 2D**). This rather low correlation suggests a change in cooking time ranks for some of the bean accessions. Also, the average cooking times (after soaking) of the aforementioned bean accessions were 40 and 43 minutes in Season 1 and Season 2, respectively. Wang et al., (2017) reported that in addition to bean cultivar, growth conditions affected cooking times considerably as they established that warmer temperature and lesser precipitation yielded beans with longer cooking times which is similar to our findings. The cooking time outcomes of fresh beans in this study demonstrate that this quality trait is influenced by soaking pre-treatment and season of growth.

3.3 The susceptibility of common beans to develop HTC

To visualize the influence of aging on the bean accessions, their cooking profiles when fresh and aged are shown in **Figure 3**. The actual cooking times of the Season 2 aged beans are shown in **Table A.3 of Appendix A**. All the bean accessions aged in this study developed the HTC defect, albeit to different extents. This is in line with the findings reported by Bento et al., (2020), Nyakuni

et al., (2008) and, Martin-Cabrejas et al., (1997). The cooking times of aged beans ranged from 112 to 273 minutes, which was 2.2 to 8.3 times larger than the corresponding cooking times for the fresh beans. The aged bean accessions grouped into three classes based on $\Delta t_{95\%, abs}$: low susceptibility to HTC (L-HTC, $\Delta t_{95\%, abs} < 90$ minutes), intermediate susceptibility to HTC (I-HTC, $\Delta t_{95\%, abs} > 90$ minutes, < 180 minutes) and, high susceptibility to HTC (H-HTC, $\Delta t_{95\%, abs} > 180$ minutes). The L-HTC class included two accessions of the carioca bean type which comprised 11% of the accessions aged. The I-HTC class encompassed the majority (61%) of bean accessions aged which were carioca, sugar, kidney, pinto and, rose coco bean types. The H-HTC class consisted of black, red haricot, navy and zebra bean types and made up 28% of bean accessions aged. The cooking time rank of bean accessions changed after aging, resulting in a low negative correlation (-0.28) between the cooking times of fresh and aged beans. This is in line with the observation that the initial hardness of fresh beans did not correlate well to their final hardness after developing the HTC defect (Michaels and Stanley, 1991). From a biochemical viewpoint, a higher susceptibility to develop HTC has been associated with higher water absorption capacity and peroxidase activity in one study (Siqueira et al., 2016) and phytic acid interaction with proteins and carbohydrates in another (Nyakuni et al., 2008).

Additionally, the negative correlations between $\Delta t_{95\%, rel}$ and seed weight/dimensions; 100-seed weight (-0.5), length (-0.5), thickness (-0.6), width (-0.6) suggest that smaller bean types, except for zebra, had a higher susceptibility to HTC, similarly put forward by the study Nyakuni et al., (2008). In contrast, Wiesenger et al. (2016) reported a positive correlation between the seed size and the percent increase in cooking time during storage. However, unlike the current study and that of Nyakuni et al., (2008), they did not include smaller bean types such as black and navy

beans. These contradicting observations perhaps minimize the role of physical characteristics in the susceptibility of beans to develop HTC.

3.4 Prediction of the cooking times of Seasons 1 and 2 fresh soaked beans

The PLS-R modelling protocol applied to predict the cooking times from NIR spectra of milled, fresh DH or WSC beans is discussed here. For the DH beans, Forward iPLS with 50 variables as the interval widths gave the best results by selecting the following wavelength regions: 1335-1370, 1738-1798, 1863-2086, 2175-2486 nm. On the other hand, the best results for the WSC beans were obtained with Reverse iPLS with intervals of 100 variables. The wavelength regions selected in the latter case were 1000-1084, 1132-1183, 1239-1370, 1446-1737, 1862-2006, and 2175-2600 nm. After variable selection, SNV followed by mean centering was selected as pre-processing combination for the DH beans, whereas only mean centering was selected for the WSC beans.

The resultant PLS-R model of the DH beans was less complex than that of the WSC beans with 7 LVs versus 11 LVs, respectively. The prediction performance of the PLS-R models developed for the Seasons 1+2 fresh soaked beans is illustrated in **Figures 4A, 4B, 4C and 4D**. The predictive power for both WSC and DH beans models were quite similar as the R^2_P and RMSEP values were in the same range. Thus, the de-hulling step during sample preparation could be eliminated with a limited loss in prediction accuracy. For both the DH and WSC beans the shorter cooking times were better predicted than longer cooking times (>50 minutes). Similar to the approach of the current study, an earlier report demonstrated that by adding at least 20 bean samples of a new season into the calibration of a former season, errors could be lowered when predicting the new season's beans (Mendoza et al., 2018). As a result, the models obtained in this study were better able to overcome not only seasonal variation but also biological diversity of the bean accessions compared to the results we obtained in our earlier study on the prediction of cooking times from

the NIR spectra of aged beans (Wafula et al., 2020). This could as well be attributed to the slight modifications in this protocol such as correlating cooking times to the mean spectra rather than individual spectrum (Williams, 2004), which introduced additional variability in the data set.

The wavelengths associated with the cooking times of the beans in this study are illustrated in **Figure B.1 in Appendix B**. These wavelengths represented the following biomolecules: polysaccharides (1929 nm), water (1938 nm), proteins (2081 nm), cellulose (2053 nm) and lipids/oils (2310 nm) (Workmann and Weyer, 2012), which are the major macromolecules in beans. However, no clear link could be identified between these wavebands and the suggested pathways leading to the development of HTC (cfr. Introduction). Cichy et al. (2015) also found important features for the prediction of the cooking time for fresh beans in the 1012 to 2436 nm. However, they also identified the waveband around 520nm in the visible range, which was attributed to anthocyanins.

3.5 Prediction of common beans susceptibility to develop HTC

In this part of the study, we related the spectra taken on fresh beans with the cooking time changes obtained after aging under adverse standard conditions that lead to HTC development (as explained in **section 3.3**). For fresh DH beans spectra, forward iPLS with 50 variables window widths and OSC together with mean centering gave the best results for both the prediction of $\Delta t_{95\%, abs}$ and $\Delta t_{95\%, rel}$ from the spectra. In the case of fresh WSC beans, reverse iPLS with interval widths of 50 variables and mean centering gave the best results for $\Delta t_{95\%, abs}$, whereas Forward iPLS with 50 variables interval widths and mean centering worked better for predicting $\Delta t_{95\%, rel}$ from the spectra. The wavelength regions selected are summarized in **Table 3**.

Based on the number of latent variables chosen, the PLS-R models of WSC beans predicting the $\Delta t_{95\%, abs}$ and $\Delta t_{95\%, rel}$ were less complex than the models of the DH beans with 9 LVs vs. 11 LVs and 9 LVs vs. 10 LVs, respectively. Additionally, the prediction of lower and higher values, was better for the WSC beans than for the DH beans, as shown in **Figures 5A, 5B, 5C and 5D**, resulting in a reduction of the RMSECV by 11 minutes for $\Delta t_{95\%, abs}$ and by 11% for $\Delta t_{95\%, rel}$. This suggests that the de-hulling removed some relevant information for predicting the susceptibility of the bean accessions to develop HTC. This could be attributed to the minor contribution of the chemical components present in the seed coat to the susceptibility of beans to develop HTC (Chigwedere et al., 2018). Therefore, it is recommended to measure the NIR spectra of milled whole fresh beans when rapidly screening for their susceptibility to develop HTC.

The functional groups that were identified as important for predicting the $\Delta t_{95\%, rel}$ of the beans are shown in **Figure B.1 in Appendix B** and were primary amides (2030 nm) and aliphatic hydrocarbons (2347 nm, 2363 nm) (Workmann and Weyer, 2012). Again, these do not support any of the HTC hypotheses. As the wavelengths representing water were not prominently associated with susceptibility to develop HTC, conditioning the beans before spectra acquisition may not be necessary. This could allow to shorten the sample preparation time by 10 days.

4. Conclusions

In this study, a protocol for the prediction of cooking times of fresh bean accessions and their susceptibility to develop HTC from the NIR spectra of fresh beans was developed. The physical characteristics of beans were initially studied and bean type was found to have a significant influence on the 100-seed weight, seed dimensions and color. Soaking pre-treatment, season of growth and, aging were identified as factors likely to change the ranking of bean accessions according to their cooking times. The aged bean accessions exhibited varying susceptibilities (low,

intermediate and high) to the HTC defect based on the absolute change in cooking time from fresh to aged. On the one hand models adequately predicted the cooking times of fresh beans below 50 minutes. On the other hand the susceptibility of fresh beans to develop HTC was also sufficiently predicted. There was evidence that de-hulling of the beans prior to milling and measurement of the NIR spectra did not result in major improvements in the predictive performance. Pre-conditioning (drying) samples was also unnecessary. Therefore, the sample preparation time could be reduced by removing the de-hulling and conditioning steps. Unfortunately, the NIR wavebands selected for predicting the cooking time and the susceptibility to develop HTC did not provide additional insights for unravelling the biochemical pathway(s) leading to the development of HTC. To gain more insight into common beans susceptibility to develop HTC, this work could be expanded to include newly bred beans, beans from different seasons and geographic locations.

In summary, the main finding of this study is that NIR spectra can be used to predict not only the cooking time of fresh beans, but also the susceptibility of fresh beans to develop the HTC defect under adverse storage conditions. This opens prospects for using such protocols in future breeding programs as well as in domestic and industrial uses.

Acknowledgments

The work was supported by VLIR-UOS [grant number KE2017IUC037A101]. Opinions of the author(s) do not automatically reflect those of either the Belgian government or VLIR-UOS, and can neither bind the Belgian Government nor VLIR-UOS. The funding source played no role in the study design, collection, analysis and interpretation of data, the writing of this manuscript or in the decision to submit the manuscript for publication.

References

- Aguilera, J. M., & Rivera, J. M. (1992). Hard-to-cook defect in black beans : hardening rates, water imbibition and multiple mechanism hypothesis. *Food Research International*, 25, 101–108. [http://dx.doi.org/10.1016/0963-9969\(92\)90150-4](http://dx.doi.org/10.1016/0963-9969(92)90150-4)
- Andersen, C. M., & Bro, R. (2010). Variable selection in regression - a tutorial. *Journal of Chemometrics*, 24, 728–737. <https://doi.org/10.1002/cem.1360>
- Arns, F. D., Ribeiro, N. D., Mezzomo, H. C., De Marco Steckling, S., Kläsener, G. R., & Casagrande, C. R. (2018). Combined selection in carioca beans for grain size, slow darkening and fast-cooking after storage times. *Euphytica*, 214(66), 1–12. <https://doi.org/10.1007/s10681-018-2149-8>
- Bento, J. A. C., Lanna, A. C., Bassinello, P. Z., Oomah, B. D., Pimenta, M. E. B., Corvalho, R. N., & Moreira, A. S. (2020). Aging indicators for stored carioca beans. *Food Research International*, 134(109249). <https://doi.org/10.1016/j.foodres.2020.109249>
- Bernal-Lugo, I., Parra, C., Portilla, M., Peña-Valdivia, C. B., & Moreno, E. (1997). Cotyledon thermal behavior and pectic solubility as related to cooking quality in common beans. *Plant Foods for Human Nutrition*, 50(2), 141–150. <https://doi.org/10.1007/BF02436033>
- Blair, M. W., Gonza, L. F., Kimani, P. M., & Butare, L. (2010). Genetic diversity , inter-gene pool introgression and nutritional quality of common beans (*Phaseolus vulgaris* L .) from Central Africa. *Theory of Applied Genetics*, 121, 237–248. <https://doi.org/10.1007/s00122-010-1305-x>
- Bressani, R. (1993). Grain quality of common beans. *Food Reviews International*, 9(2), 237-297.
- Broughton, W. J., Blair, M., Beebe, S., Gepts, P., & Vanderleyden, J. (2003). Spp.) – Model Food

- 465 Legumes, 55–128. *Plant and Soil*, 252, 55–128.
- 466 Carbas, B., Machado, N., Oppolzer, D., Ferreira, L., Brites, C., Rosa, E. A. S., & Barros, A. I. R.
 467 N. A. (2020). Comparison of near-infrared (NIR) and mid-infrared (MIR) spectroscopy for the
 468 determination of nutritional and antinutritional parameters in common beans. *Food Chemistry*, 306
 469 (125509). <https://doi.org/10.1016/j.foodchem.2019.125509>
- 470 Chigwedere, C. M., Flores, J. H. M., Panozzo, A., Van Loey, A. M., & Hendrickx, M. E. (2019).
 471 Instability of common beans during storage causes hardening: The role of glass transition
 472 phenomena. *Food Research International*, 121, 506–513.
 473 <https://doi.org/10.1016/j.foodres.2018.12.006>
- 474 Chigwedere, C. M., Olaoye, T. F., Kyomugasho, C., Kermani, Z. J., Pallares, A. P., Van Loey, A.
 475 M., Hendrickx, M. E. (2018). Mechanistic insight into softening of Canadian wonder common
 476 beans (*Phaseolus vulgaris*) during cooking. *Food Research International*, 106, 522–531.
 477 <https://doi.org/10.1016/j.foodres.2018.01.016>
- 478 Cichy, K. A., Fernandez, A., Kilian, A., Kelly, J. D., Galeano, C. H., Shaw, S., Troxtell, E. (2014).
 479 QTL analysis of canning quality and color retention in black beans (*Phaseolus vulgaris* L .).
 480 *Molecular Breeding*, 33, 139–154. <https://doi.org/10.1007/s11032-013-9940-y>
- 481 Cichy, K. A., Wiesinger, J. A., & Mendoza, F. A. (2015). Genetic diversity and genome-wide
 482 association analysis of cooking time in dry bean (*Phaseolus vulgaris* L.). *Theoretical and Applied*
 483 *Genetics*, 128(8), 1555–1567. <https://doi.org/10.1007/s00122-015-2531-z>
- 484 Coelho, C. M. M., Bellato, C. de M., Santos, J. C. P., Ortega, E. M. M., & Tsai, S. M. (2007).
 485 Effect of phytate and storage conditions on the development of the ‘ hard-to-cook .’ *Journal of the*
 486 *Science of Food and Agriculture*, 87, 1237–1243. <https://doi.org/10.1002/jsfa>

- 487 Conzen, J. (2014). *Multivariate Calibration: A practical guide for developing methods in the*
 488 *quantitative analytical chemistry* (3rd ed. pp 23-49). Bruker Optik GmbH
- 489 Galiotou-panayotou, M., Kyriakidis, N. B., & Margaritis, I. (2008). Phytase – phytate – pectin
 490 hypothesis and quality of legumes cooked in calcium solutions. *Journal of the Science of Food*
 491 *and Agriculture*, 88, 355–361. <https://doi.org/10.1002/jsfa>
- 492 Garcia, E., Filisetti, T. M. C. C., Udaeta, J. E. M., & Lajolo, F. M. (1998). Hard-To-Cook Beans
 493 (*Phaseolus vulgaris*): Involvement of Phenolic Compounds and Pectates. *Journal of Agricultural*
 494 *and Food Chemistry*, 46(6), 2110–2116. <https://doi.org/10.1021/jf970848f>
- 495 Garcia-Vela, L. A., & Stanley, D. W. (1989). Water-Holding Capacity in Hard-to-Cook Beans (*Phaseolus vulgaris*): Effect of pH and Ionic Strength. *Journal of Food Science*, 54(4), 1080–1081.
- 497 Guevara-González, R. G., Torres-Pacheco, I., Reynoso-Camacho, R., Ramos-Gomez, M., &
 498 Loarca-Pina, G. (2006). Bioactive components in common beans (*Phaseolus vulgaris* L.).
 499 *Advances in Agricultural and Food Biotechnology*, 37661(2), 217–236. Retrieved from
 500 [http://www.trnres.com/ebook/uploads/gonzalezfinal/T_1231138943gonzalez 10.pdf](http://www.trnres.com/ebook/uploads/gonzalezfinal/T_1231138943gonzalez%20.pdf)
- 501 Hentges, D. L., Weaver, C. M., and Nielsen, S. S. (1991). Changes of Selected Physical &
 502 Chemical Components the Development of the Hard-to-Cook Bean Defect. *Journal of Food*
 503 *Science*, 56(2), 436–442.
- 504 Hohlberg, A. I., & Stanley, D. W. (1987). Hard-To-Cook Defect in Black Beans. Protein and
 505 Starch Considerations. *Journal of Agricultural and Food Chemistry*, 35, 571–576.
 506 <https://doi.org/10.1094/CCHEM-04-16-0126-FI>
- 507 Jones, P. M. B., & Boulter, D. (1983). The Cause of Reduced Cooking Rate in *Phaseolus vulgaris*

- 508 Following Adverse Storage Conditions. *Journal of Food Science*, 48, 623–649.
- 509 Katungi, E., Sperling, L., Karanja, D., Farrow, A., & Beebe, S. (2011). Relative importance of
 510 common bean attributes and variety demand in the drought areas of Kenya. *Journal of*
 511 *Development and Agricultural Economics*, 3(8), 411–422.
- 512 Kilmer, O. L., Seib, P. A., & Hosney, R. C. (1994). Effects of minerals and apparent phytase
 513 activity in the development of the hard-to-cook state of beans. *Cereal Chemistry*, 71(5), 476–482.
- 514 Kinyanjui, P. K., Njoroge, D. M., Makokha, A. O., Christiaens, S., Ndaka, D. S., & Hendrickx, M.
 515 (2015). *Hydration properties and texture fingerprints of easy and hard-to-cook bean varieties.*
 516 *Food Science and Nutrition* (Vol. 3). <https://doi.org/10.1002/fsn3.188>
- 517 Kinyanjui, P. K., Njoroge, D. M., Makokha, A. O., Christiaens, S., Sila, D. N., & Hendrickx, M.
 518 (2017). Quantifying the effects of postharvest storage and soaking pretreatments on the cooking
 519 quality of common beans (*Phaseolus vulgaris* L.). *Journal of Food Processing and Preservation*,
 520 41, 1–10. <https://doi.org/10.1111/jfpp.13036>
- 521 Kutner, M.H., Nachtsheim, C.J., Neter, J., & Li, W. (2005). *Applied Linear Statistical Models* (5th
 522 ed., pp 555-603). McGraw-Hill/Irwin, Boston MA.
- 523 Martin-Cabrejas, M. A., Esteban, R. M., Perez, P., Maina, G., & Waldron, K. W. (1997). Changes
 524 in Physicochemical Properties of Dry Beans (*Phaseolus vulgaris* L.) during Long-Term Storage.
 525 *Journal of Agricultural and Food Chemistry*, 45, 3223–3227.
- 526 Mendoza, F. A., Cichy, K., Lu, R., & Kelly, J. D. (2014). Evaluation of Canning Quality Traits in
 527 Black Beans (*Phaseolus vulgaris* L.) by Visible / Near-Infrared Spectroscopy. *Food Bioprocess*
 528 *Technology*, 7, 2666–2678. <https://doi.org/10.1007/s11947-014-1285-y>

- 529 Mendoza, F. A., Wiesinger, J. A., Lu, R., Nchimbi-Msolla, S., Miklas, P. N., Kelly, J. D., & Cichy,
 530 K. A. (2018). Prediction of Cooking Time for Soaked and Unsoaked Dry Beans (*Phaseolus*
 531 *vulgaris* L.) Using Hyperspectral Imaging Technology. *The Plant Phenome Journal*, 1–9.
 532 <https://doi.org/10.2135/tppj2018.01.0001>
- 533 Michaels, T. E., & Stanley, D. W. (1991). Stability and inheritance of storage induced hardening
 534 in 20 common bean cultivars. *Canadian Journal of Plant Science*, 71, 641–647.
- 535 Naes, T., Isaksson, T., Fearn, T., & Davies, T. (2002). *A user friendly guide to multivariate*
 536 *calibration and classification*. NIR Publications, Chichester, United Kingdom.
- 537 Nicolai, B., Beullens, K., Bobelyn, E., Peirs, A., Saeys, W., Theron, K.I., Lammertyn, J. (2007).
 538 Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: A
 539 review. *Postharvest biology and technology*, 46 (2), 99–118.
- 540 Njoroge, D. M., Kinyanjui, P. K., Chigwedere, C. M., Christiaens, S., Makokha, A. O., Sila, D.
 541 N., & Hendrickx, M. E. (2016). Mechanistic insight into common bean pectic polysaccharide
 542 changes during storage, soaking and thermal treatment in relation to the hard-to-cook defect. *Food*
 543 *Research International*, 81, 39–49. <https://doi.org/10.1016/j.foodres.2015.12.024>
- 544 Njoroge, D. M., Kinyanjui, P. K., Christiaens, S., Shpigelman, A., Makokha, A. O., Sila, D. N., &
 545 Hendrickx, M. E. (2015). Effect of storage conditions on pectic polysaccharides in common beans
 546 (*Phaseolus vulgaris*) in relation to the hard-to-cook defect. *Food Research International*, 76, 105–
 547 113. <https://doi.org/10.1016/j.foodres.2015.03.005>
- 548 Njoroge, D. M., Kinyanjui, P. K., Makokha, A. O., Christiaens, S., Shpigelman, A., Sila, D. N., &
 549 Hendrickx, M. E. (2014). Extraction and characterization of pectic polysaccharides from easy- and
 550 hard-to-cook common beans (*Phaseolus vulgaris*). *Food Research International*, 64, 314–322.

- 551 <https://doi.org/10.1016/j.foodres.2014.06.044>
- 552 Nørgaard, L. N., Saudland, A., Wagner, J., Nielsen, J. P., Unck, L. M., & Engelsen, S. B. (2000).
553 Interval Partial Least-Squares Regression (i PLS): A Comparative Chemometric Study with an
554 Example from Near-Infrared Spectroscopy. *Applied Spectroscopy*, 54(3), 413–419.
- 555 Nyakuni, G. A., Kikafunda, J. K., Muyonga, J. H., Kyamuhangire, W. M., Nakimbugwe, D., &
556 Ugen, M. (2008). Chemical and nutritional changes associated with the development of the hard-
557 to-cook defect in common. *International Journal of Food Sciences and Nutrition*, 59(7–8), 652–
558 659. <https://doi.org/10.1080/09637480701602886>
- 559 Parmar, N., Singh, N., Kaur, A., Viridi, A. S., & Thakur, S. (2016). Effect of canning on color,
560 protein and phenolic profile of grains from kidney bean, field pea and chickpea. *Food Research*
561 *International*, 89(1), 526–532. <https://doi.org/10.1016/j.foodres.2016.07.022>
- 562 Pasquini, C. (2018). Near infrared spectroscopy: A mature analytical technique with new
563 perspectives – A review. *Analytica Chimica Acta*, 1026, 8–36.
564 <https://doi.org/10.1016/j.aca.2018.04.004>
- 565 Peng, C. J., Lee, K. U. K. L., & Ingersoll, G. M. (2002). An Introduction to Logistic Regression
566 Analysis and Reporting. *The Journal of Educational Research*, 96(1), 3–14.
- 567 Piergiovanni, A. R., Cerbino, D., & Gatta, C. D. (2000). Diversity in seed quality traits of common
568 bean populations from Basilicata (Southern Italy). *Plant Breeding*, 119, 513–516.
- 569 Plans, M., Simó, J., Casañas, F., & Sabaté, J. (2012). Near-infrared spectroscopy analysis of seed
570 coats of common beans (*Phaseolus vulgaris* L.): A potential tool for breeding and quality
571 evaluation. *Journal of Agricultural and Food Chemistry*, 60(3), 706–712.

- 572 <https://doi.org/10.1021/jf204110k>
- 573 Plans, M., Simó, J., Casañas, F., Sabaté, J., & Rodriguez-Saona, L. (2013a). Characterization of
574 common beans (*Phaseolus vulgaris* L.) by infrared spectroscopy: Comparison of MIR, FT-NIR
575 and dispersive NIR using portable and benchtop instruments. *Food Research International*, 54(2),
576 1643–1651. <https://doi.org/10.1016/j.foodres.2013.09.003>
- 577 Plans, M., Simó, J., Casañas, F., Sabaté, J., & Rodriguez-Saona, L. (2013b). Characterization of
578 common beans (*Phaseolus vulgaris* L.) by infrared spectroscopy: Comparison of MIR, FT-NIR
579 and dispersive NIR using portable and benchtop instruments. *Food Research International*, 54(2),
580 1643–1651. <https://doi.org/10.1016/j.foodres.2013.09.003>
- 581 Pojić, M., Mastilović, J., Palić, D., & Pestorić, M. (2010). The development of near-infrared
582 spectroscopy (NIRS) calibration for prediction of ash content in legumes on the basis of two
583 different reference methods. *Food Chemistry*, 123(3), 800–805.
584 <https://doi.org/10.1016/j.foodchem.2010.05.013>
- 585 Rebello, C. J., Greenway, F. L., & Finley, J. W. (2014). Whole grains and pulses: A comparison
586 of the nutritional and health benefits. *Journal of Agricultural and Food Chemistry*, 62(29), 7029–
587 7049. <https://doi.org/10.1021/jf500932z>
- 588 Saeys, W., Nguyen, D., Van Beers, R., & Nicolai, B. (2019). Multivariate calibration of
589 spectroscopic sensors for postharvest quality evaluation: A review. *Postharvest Biology and*
590 *Technology*, 158. doi:10.1016/j.postharvbio.2019.110981
- 591 Shiga, T. M., Lajolo, F. M., & Filisetti, T. M. C. C. (2004). Changes in the cell wall
592 polysaccharides during storage and hardening of beans. *Food Chemistry*, 84(1), 53–64.
593 [https://doi.org/10.1016/S0308-8146\(03\)00166-3](https://doi.org/10.1016/S0308-8146(03)00166-3)

- 594 Sinkovic, L., Pipan, B., Sinkovic, E., & Meglic, V. (2019). Morphological Seed Characterization
595 of Common (*Phaseolus vulgaris* L.) and Runner (*Phaseolus coccineus* L.) Bean Germplasm : A
596 Slovenian Gene Bank Example. *BioMed Research International*, 1–13.
- 597 Siqueira, B. S., Bassinello, P. Z., Malgaresi, G., Pereira, W. J., & Fernandes, K. F. (2016).
598 Analyses of technological and biochemical parameters related to the HTC phenomenon in carioca
599 bean genotypes by the use of PCA. *LWT - Food Science and Technology*, 65, 939–945.
600 <https://doi.org/10.1016/j.lwt.2015.09.030>
- 601 Stanley, D. W. (1992). A possible role for condensed tannins in bean hardening. *Food Research*
602 *International*, 25(3), 187–192. [https://doi.org/10.1016/0963-9969\(92\)90136-S](https://doi.org/10.1016/0963-9969(92)90136-S)
- 603 Syanda, J. S. (2019). *The effects of physical properties of common bean (Phaseolus vulgaris L.)*
604 *varieties on soaking and cooking time.* South Eastern Kenya University.
605 <http://repository.seku.ac.ke/bitstream/handle/123456789/4462/Syanda>
- 606 Uebersax, M. A., Ruengsakulrach, S. and Occena, L. G. (1991). Strategies and procedures for
607 processing dry beans. *Food Technology*, 45(9), 104-111.
- 608 Wafula, E. N., Wainaina, I. N., Buve, C., Nguyen, N. D., Kinyanjui, P. K., Saeys, W., ...
609 Hendrickx, M. (2020). Application of near-infrared spectroscopy to predict the cooking times of
610 aged common beans (*Phaseolus vulgaris* L.). *Journal of Food Engineering*, 284(110056).
611 <https://doi.org/10.1016/j.jfoodeng.2020.110056>
- 612 Walsh, K. B., Blasco, J., Zude-Sasse, M., & Sun, X. (2020). Visible-NIR ‘point’ spectroscopy in
613 postharvest fruit and vegetable assessment: The science behind three decades of commercial use.
614 *Postharvest Biology and Technology*, 168, 111246.

- 615 Wang, J., Liu, H., & Ren, G. (2014). Near-infrared spectroscopy (NIRS) evaluation and regional
616 analysis of Chinese faba bean (*Vicia faba* L.). *Crop Journal*, 2(1), 28–37.
617 <https://doi.org/10.1016/j.cj.2013.10.001>
- 618 Wang, N., Hou, A., Santos, J., & Maximiuk, L. (2017). Effects of cultivar, growing location, and
619 year on physicochemical and cooking characteristics of dry beans (*Phaseolus vulgaris*). *Cereal*
620 *Chemistry*, 94(1), 128–134. <https://doi.org/10.1094/CCHEM-04-16-0124-FI>
- 621 Wiesinger, J. A., Cichy, K. A., Glahn, R. P., Grusak, M. A., Brick, M. A., Thompson, H. J., &
622 Tako, E. (2016). Demonstrating a Nutritional Advantage to the Fast-Cooking Dry Bean (*Phaseolus*
623 *vulgaris* L.). *Journal of Agricultural and Food Chemistry*, 64, 8592–8603.
624 <https://doi.org/10.1021/acs.jafc.6b03100>
- 625 Williams, P. (2004). Near-Infrared technology: Getting the best out of light. *Near Infrared*
626 *Technology*, 132.
- 627 Wood, J. A. (2017). Evaluation of cooking time in pulses: A review. *Cereal Chemistry*, 94(1), 32–
628 48. <https://doi.org/10.1094/CCHEM-05-16-0127-FI>
- 629 Workmann, J.J., & Weyer, L. (2012). *Practical guide and spectral atlas for interpretive near-*
630 *infrared spectroscopy* (2nd ed., pp. 239-274). Boca Roca, USA.
- 631 Wortmann, C. S., Kirkby, R. A., Eledu, C. A., & Allen, D. J. (1998). *Atlas of Common Bean*
632 *Production in Africa*. Cali, Colombia.

633

634

Table 1: Physical characteristics of the Season 1 seventy bean accessions grouped according to bean typeValues are represented as mean $\pm$ standard deviationMeans followed by different letters in the same column are significantly different at $p < 0.05$

Seed Type	Number of accessions	100-seed Weight (g)	Length (mm)	Thickness (mm)	Width (mm)	L^*	a^*	b^*
Black beans	10	24.46 $\pm$ 13.31 d	10.18 $\pm$ 0.16 d	4.78 $\pm$ 0.69 e	6.80 $\pm$ 1.01 d	24.39 $\pm$ 3.59 g	0.87 $\pm$ 0.54 e	0.69 $\pm$ 0.48 f
Brown beans	1	29.42 $\pm$ 0.32 cd	9.45 $\pm$ 0.50 de	5.98 $\pm$ 0.26 bcd	7.21 $\pm$ 0.33 cd	48.40 $\pm$ 1.61 bcd	4.23 $\pm$ 1.71 de	19.47 $\pm$ 1.52 a
Carioca	9	52.41 $\pm$ 5.47 a	13.41 $\pm$ 0.18 b	7.10 $\pm$ 0.48 a	8.21 $\pm$ 0.51 a	40.54 $\pm$ 3.48 de	14.87 $\pm$ 0.60 bc	15.87 $\pm$ 0.54 a
Kidney beans	4	50.93 $\pm$ 6.25 a	15.91 $\pm$ 0.25 a	5.95 $\pm$ 0.83 bc	7.74 $\pm$ 0.59 abc	28.91 $\pm$ 4.58 fg	18.40 $\pm$ 0.85 a	5.74 $\pm$ 0.76 de
Mwezi moja	3	50.05 $\pm$ 2.07 a	16.11 $\pm$ 0.29 a	5.72 $\pm$ 0.43 cd	7.61 $\pm$ 0.52 bc	34.75 $\pm$ 3.77 ef	7.72 $\pm$ 0.98 d	1.92 $\pm$ 0.88 ef
Navy beans	5	21.79 $\pm$ 6.94 d	8.41 $\pm$ 0.23 e	5.15 $\pm$ 0.42 de	6.01 $\pm$ 0.37 e	70.71 $\pm$ 3.17 a	-0.05 $\pm$ 0.76 e	12.57 $\pm$ 0.68 b
Pinto	4	36.62 $\pm$ 9.05 bc	12.30 $\pm$ 0.25 c	5.75 $\pm$ 0.97 cd	8.20 $\pm$ 1.21 ab	49.38 $\pm$ 5.08 c	4.76 $\pm$ 0.85 d	17.28 $\pm$ 0.76 a
Rose coco	13	54.06 $\pm$ 6.57 a	15.75 $\pm$ 0.14 a	6.23 $\pm$ 0.68 b	7.69 $\pm$ 0.66 c	37.50 $\pm$ 6.17 e	16.24 $\pm$ 0.47 abc	8.22 $\pm$ 0.42 cd
Red Haricot	6	23.46 $\pm$ 4.35 d	10.72 $\pm$ 0.21 cd	5.18 $\pm$ 0.54 de	6.9 $\pm$ 0.63 cd	28.84 $\pm$ 3.40 fg	17.05 $\pm$ 0.7 ab	6.26 $\pm$ 0.62 d
Sugar beans	13	50.84 $\pm$ 10.46 a	15.96 $\pm$ 0.13 a	6.03 $\pm$ 0.73 bc	7.72 $\pm$ 0.79 c	55.31 $\pm$ 6.05 b	5.79 $\pm$ 0.46 d	16.56 $\pm$ 0.41 a
Yellow beans	1	54.18 $\pm$ 0.51 ab	15.17 $\pm$ 0.50 ab	6.36 $\pm$ 0.61 bc	8.01 $\pm$ 0.47 abc	53.07 $\pm$ 3.59 bc	5.47 $\pm$ 1.79 de	18.83 $\pm$ 1.52 a
Zebra beans	1	66.86 $\pm$ 0.53 a	17.26 $\pm$ 0.50 a	6.26 $\pm$ 0.62 bc	7.75 $\pm$ 0.44 abc	50.13 $\pm$ 1.65 bcd	10.63 $\pm$ 1.71 cd	12.47 $\pm$ 1.52 abc

Table 2: Physical characteristics of the Season 2 bean accessions

Values are represented as mean $\pm$ standard deviation

**Average values represented as mean $\pm$ standard deviation of the 17 bean accessions planted in both Seasons 1 and 2.

Means followed by different letters in the same column are significantly different at $p < 0.05$

Bean accessions	100 Seed Weight (g)	Length (mm)	Thickness (mm)	Width (mm)	L*	a*	b*
3. Sugar beans	24.07 $\pm$ 0.12 k	11.82 $\pm$ 0.93 ef	4.55 $\pm$ 0.59 e	6.25 $\pm$ 0.62 ghi	55.38 $\pm$ 0.98 bc	7.75 $\pm$ 0.15 efg	13.99 $\pm$ 0.43 efg
6. Sugar beans	29.02 $\pm$ 0.75 ij	13.42 $\pm$ 1.09 bcd	4.86 $\pm$ 0.37 cde	6.85 $\pm$ 0.36 defg	57.97 $\pm$ 1.38 a	7.43 $\pm$ 0.24 fgh	14.37 $\pm$ 0.82 fgh
8. Black beans	18.70 $\pm$ 0.81 lm	9.71 $\pm$ 0.80 hi	4.70 $\pm$ 0.52 cde	6.33 $\pm$ 0.48 gh	17.54 $\pm$ 0.15 de	0.21 $\pm$ 0.27 gh	-0.77 $\pm$ 0.13 gh
11. Rose coco	41.62 $\pm$ 1.03 bc	13.97 $\pm$ 0.71 bed	5.34 $\pm$ 0.31 bc	7.31 $\pm$ 0.30 bcd	34.56 $\pm$ 1.65 bcde	11.16 $\pm$ 0.70 bcd	5.17 $\pm$ 1.10 bcd
12. Rose coco	37.57 $\pm$ 0.59 def	13.48 $\pm$ 0.63 bcd	5.57 $\pm$ 0.37 ab	6.90 $\pm$ 0.31 defg	34.26 $\pm$ 1.78 bcde	9.87 $\pm$ 0.51 cde	4.01 $\pm$ 0.48 cde
15. Sugar beans	28.33 $\pm$ 0.47 j	13.74 $\pm$ 1.25 bcd	4.53 $\pm$ 0.42 e	6.48 $\pm$ 0.51 fgh	59.46 $\pm$ 0.88 a	8.66 $\pm$ 0.20 fgh	16.92 $\pm$ 0.45 fgh
18. Carioca	30.56 $\pm$ 1.39 hij	11.90 $\pm$ 0.35 ef	5.60 $\pm$ 0.22 ab	7.11 $\pm$ 0.34 cdef	31.31 $\pm$ 0.79 bcde	18.43 $\pm$ 0.21 abc	12.92 $\pm$ 1.17 cdefgh
26. Black beans	16.73 $\pm$ 0.24 m	8.12 $\pm$ 1.53 j	4.46 $\pm$ 0.35 e	5.82 $\pm$ 0.36 hi	19.98 $\pm$ 0.70 de	0.15 $\pm$ 0.12 gh	-0.93 $\pm$ 0.11 i
35. Carioca	32.50 $\pm$ 1.29 ghi	11.39 $\pm$ 0.78 fg	5.39 $\pm$ 0.51 bc	6.74 $\pm$ 0.52 efg	31.25 $\pm$ 2.57 bcde	18.25 $\pm$ 0.20 ab	12.74 $\pm$ 1.47 abcdef
39. Kidney beans	39.69 $\pm$ 0.95 bcd	14.02 $\pm$ 0.60 bcd	5.65 $\pm$ 0.34 ab	7.42 $\pm$ 0.46 abcde	22.90 $\pm$ 0.65 cde	13.90 $\pm$ 0.12 bcd	1.68 $\pm$ 0.16 i
43. Red haricot	21.81 $\pm$ 1.65 kl	10.13 $\pm$ 0.97 gh	4.58 $\pm$ 0.57 de	6.24 $\pm$ 0.52 ghi	30.49 $\pm$ 0.86 bcde	19.73 $\pm$ 0.46 a	8.18 $\pm$ 0.50 defghi
49. Carioca	35.32 $\pm$ 1.10 efg	11.64 $\pm$ 0.97 ef	6.16 $\pm$ 0.41 a	7.50 $\pm$ 0.41 abcd	35.16 $\pm$ 0.71 bcde	17.91 $\pm$ 0.40 abc	16.98 $\pm$ 0.96 abcde
50. Zebra beans	50.21 $\pm$ 2.58 a	16.13 $\pm$ 1.13 a	5.70 $\pm$ 0.55 ab	7.95 $\pm$ 0.52 ab	47.11 $\pm$ 1.22 bcd	7.89 $\pm$ 0.55 def	8.37 $\pm$ 0.59 defghi
51. Rose coco	35.70 $\pm$ 0.52 efg	14.66 $\pm$ 0.73 b	5.24 $\pm$ 0.32 bcd	7.24 $\pm$ 0.45 bcd	37.89 $\pm$ 0.92 bcde	13.19 $\pm$ 0.33 abc	4.11 $\pm$ 0.47 fghi
52. Black beans	15.73 $\pm$ 1.12 m	8.67 $\pm$ 0.57 ij	4.79 $\pm$ 0.17 cde	6.00 $\pm$ 0.27 hi	20.30 $\pm$ 0.50 e	0.34 $\pm$ 0.06 gh	-1.36 $\pm$ 0.09 i
54. Pinto	33.66 $\pm$ 0.44 gh	11.59 $\pm$ 0.50 efg	5.12 $\pm$ 0.27 bcde	8.08 $\pm$ 0.38 a	54.77 $\pm$ 0.37 a	6.12 $\pm$ 0.39 fgh	14.77 $\pm$ 0.24 ab
62. Navy beans	15.06 $\pm$ 1.04 m	7.88 $\pm$ 0.87 j	4.59 $\pm$ 0.39 de	5.61 $\pm$ 0.45 i	70.93 $\pm$ 0.42 a	1.96 $\pm$ 0.08 h	12.16 $\pm$ 0.58 abc
Nabe15 Rose coco	34.44 $\pm$ 0.53 fg	12.71 $\pm$ 1.00 def	6.15 $\pm$ 0.49 a	7.81 $\pm$ 0.46 abc	37.75 $\pm$ 1.20 b	17.81 $\pm$ 0.59 abc	10.78 $\pm$ 0.88 abcd
Nabe17 Carioca	42.13 $\pm$ 2.35 b	12.89 $\pm$ 0.57 cde	6.15 $\pm$ 0.53 a	8.04 $\pm$ 0.48 a	26.01 $\pm$ 1.00 bcde	16.94 $\pm$ 0.69 abc	5.71 $\pm$ 0.18 efghi
Nabe 4 Carioca	38.41 $\pm$ 0.55 cde	14.23 $\pm$ 0.88 bc	5.30 $\pm$ 0.58 bc	7.25 $\pm$ 0.48 bcde	39.06 $\pm$ 1.89 bc	15.96 $\pm$ 0.84 abc	8.11 $\pm$ 0.35 bcdfg
Season 1**	42.52$\pm$16.84 b	13.30$\pm$3.13 a	5.94$\pm$0.92 b	7.45$\pm$0.92 b	41.51$\pm$14.01 a	9.53$\pm$7.08 a	10.78$\pm$6.75 a
Season 2**	29.78$\pm$10.04 a	11.90$\pm$2.39 a	5.11$\pm$0.52 a	6.81$\pm$0.72 a	38.90$\pm$15.93 a	9.58$\pm$6.63 a	8.43$\pm$6.45 a

647 **Table 3: The wavelength regions selected for prediction of the susceptibility of beans to HTC**

PLS-R models	Wavelength region selected
DH beans ($\Delta t_{95\%, abs}$)	1182-1210, 1488-1532, 2270-2373
WSC beans ($\Delta t_{95\%, abs}$)	1131-1156, 1239-1270, 1446-1488, 1862-2174, 2488-2600
DH beans ($\Delta t_{95\%, rel}$)	2006-2086, 2270-2373
WSC beans ($\Delta t_{95\%, rel}$)	1040-1062, 1107-1156, 1239-1301, 1336-1407, 1446-1532, 1737-1797, 1931-2269, 2374-2600

648

Figure 1: Flow of the research steps undertaken in this study.

Figure 2: Summarized illustration of the cooking times of the fresh beans obtained during Seasons 1 and 2. (A) Average cooking times of the 70 not soaked and soaked bean accessions according to bean type planted during Season 1. The error bars indicated in Figure 2A represent the variability in cooking times of different accessions within a bean type; (B) Example of cooking profiles of the pinto bean type: The continuous/dashed lines in black and blue are the logistic regression predictor curves for non-soaked and soaked beans, respectively, and the symbols are the finger pressing data; (C) Average cooking times of the 20 non-soaked and soaked bean accessions planted during Season 2; (D) Scatter plot of the average cooking times in Season 1 versus Season 2 for soaked beans.

Figure 3: Illustration of cooking profiles of Season 2 fresh and aged beans which had low susceptibility to HTC (L-HTC), intermediate susceptibility to HTC (I-HTC) and, high susceptibility to (H-HTC).

Figure 4: Output of PLS-R models for the prediction of cooking times of fresh Seasons 1 and 2 beans (A) DH beans model calibration performance. (B) DH beans model prediction of test set. (C) WSC beans model calibration performance. (D) WSC beans model prediction of test set.

Figure 5: Output of PLS-R for the prediction of Season 2 beans delta (Δ) $t_{95\%, abs}$ (A) DH beans model calibration performance. (B) WSC beans model calibration performance. Output of PLS-R for the prediction of Season 2 beans delta (Δ) $t_{95\%, rel}$. (C) DH beans model calibration performance (D) WSC beans model calibration performance.