Journal of Food Engineering 142 (2014) 1–8

Contents lists available at ScienceDirect

Journal of Food Engineering

journal homepage: www.elsevier.com/locate/jfoodeng

New methods for measuring surface area, seed coat separation, and 'chip and scratch' damage in almonds

Megan Fisklements, Diane M. Barrett *

Department of Food Science and Technology, University of California, One Shields Avenue, Davis, CA 95616-8598, USA

ARTICLE INFO

Article history:
Received 29 January 2014
Received in revised form 2 April 2014
Accepted 24 May 2014
Available online 2 June 2014

Keywords:
Almonds
Seed coat
Surface area
Multivariate modeling
Skin slip

ABSTRACT

This study quantified partial seed coat loss from almond kernels. Basic physical properties of Nonpareil, Monterey, and Butte–Padre kernels were measured to determine which correlate to surface area, and could therefore predict it. Rehydrated almonds were manually peeled and images of seed coats were digitally analyzed. Surprisingly, individual dimensions (length, width, thickness) did not increase with increasing surface area, nor they did scale in proportion to one other. Almond surface area is often estimated from an equivalent sphere, but the sphere-based estimate only predicted 60% of the variation in measured surface area. An empirical model was created to predict surface area ($r^2 = 0.74$), based on the almond variety, as well as length, width, and mass after rehydration. By comparing the predicted total surface area and the measured surface area of any remaining seed coat, a quantitative percentage of lost seed coat can be calculated.

1. Introduction

California produced almost 1.7 billion pounds of almonds [*Prunus dulcis* L., syn. *Prunus amygdalus*, *Amygdalus communis* L.] in 2010, earning nearly $2.7 billion (Boriss, 2011). The price per pound depends strongly on the USDA grade, which sorts almonds according to incidence of defects. 'Chip and scratch' is an important defect defined as the aggregate exposure of nut meat greater than ¼ in. in diameter (Almond Board of California, 2009). This defect often results from mechanical damage accrued during removal of the almond hull and shell.

Chip and scratch counts rose in the 2008 California crop, with some lots reportedly having as high as 60–70% of almonds exhibiting chip and scratch after roasting. Mechanical damage incurred during removal of the hull and shell would typically affect both the seed coat and the kernel. However, unlike the huller–sheller mechanical damage seen in previous years, the almond nutmeat appeared intact in this new type of chip and scratch; damage was limited to the seed coat, which slipped away from the kernel. Members of the almond industry call this phenomenon 'skin slip' to differentiate it from the almond kernel damage typical of the chip and scratch defect, although it grades the same way, since a portion of the seed coat (skin) is missing in both cases. Better understanding of seed coat adherence could therefore lead to improved product quality and reduced economic losses, since almonds with extensive chip and scratch (with or without kernel damage) may sell for 20–30% less than almonds with intact seed coats.

The USDA 'chip and scratch' defect rating is currently determined using qualitative methods; almonds are considered defective if they have 0.25 in. in diameter (32.15 mm^2) aggregate seed coat missing, irrespective of the source of the damage: chips, scratches, skin slip, or a combination thereof (Almond Board of California, 2009). Rather than using the USDA's qualitative threshold, a method for quantifying the degree of seed coat missing from individual almonds would assist with more accurately measuring and characterizing the chip and scratch defect, no matter the cause.

There is currently no validated protocol for measurement of the actual surface area of an almond. Mohsenin's review of surface area determination (Mohsenin, 1978) included work by Baten and Marshall, showing that the surface area of unpicked apples could be adequately but sub-optimally estimated based on their transverse diameters, i.e. perpendicular to the core in several planes (Baten and Marshall, 1943). Mohsenin also reported on Frechett's research into heat transfer in apples that correlated actual apple surface area (measured by peeling) to the surface area of an equivalent sphere of identical mass and density (Frechette, 1966). Calculating density requires determination of volume, which was measured through volumetric displacement. Building from these, others have more recently posited that the surface area of a pine nut or almond is

* Corresponding author. Tel.: +1 530 752 4800.
E-mail addresses: maclements@ucdavis.edu (M. Fisklements), dmbarrett@ucdavis.edu (D.M. Barrett).

http://dx.doi.org/10.1016/j.jfoodeng.2014.05.021
0260-8774/

adequately modeled by a sphere of equivalent diameter as that of a nut, using a geometric mean diameter (length × width × height)$^{1/3}$ (Ozguven and Vursavus, 2005; Ledbetter and Sisterson, 2010).

This change from modeling based on a sphere of equivalent volume to one of equivalent diameter presumably occurred because diameters can be measured faster and with less difficulty than volumes. The substitution of average diameter for volume relies on the intuitive yet untested assumption that almond volume and dimensional properties scale together. Surface area estimation based on a sphere with equivalent diameter reportedly correlates to almond mass with an $r^2 = 0.971$ (Ledbetter and Sisterson, 2010), but this only verifies that mean diameter and mass scale together. No studies validating the assumed correlation between diameters and surface area are found in the literature.

An accurate surface area estimate for almonds is vital to mass transfer calculations such as dehydration and rehydration rates, heat transfer, and drag force (Bayram, 2011). Another application for surface area was proposed by Ledbetter and Sisterson, who suggested that the regression slope of mass versus surface area could potentially be used as an objective varietal standard in order to quantitatively determine the variety of any given almond kernel (Ledbetter and Sisterson, 2010). Such objective varietal standards do not yet exist, so they would be of significant interest to the almond growing and processing industries. Furthermore, the sphere-based surface area estimate does not provide a way to measure how much of the seed coat is missing. Therefore a validated method of quantifying total and missing almond surface area is needed. The aim of this research is to evaluate surface area estimation based on a sphere of equivalent diameter, and to propose a method for quantifying both the total surface area, and the missing seed coat that separates from almonds, as occurs during the defect of chip and scratch.

2. Methods and materials

2.1. Raw materials

Nonpareil, Monterey, and a mixture of Butte and Padre almonds from the 2009 crop were contributed by Paramount Farms and the Almond Board of California. The Agricultural Marketing Service ruled in 2004 that Butte–Padre mixed lots can be handled, marketed, and inspected as one variety because of the very similar characteristics (Service, 2004), so mixed lots of Butte–Padres were used in these experiments. Nonpareil, Monterey, and Butte–Padre varieties were chosen as samples because together they accounted for 64% of the crop, by tonnage, during the period of this study (Almond Board of California, 2010, 2011). Almonds were obtained from three processor lots per variety, each of which contained kernels pooled from dozens of ranches.

Nonpareils from the 2010 crop were obtained from the University of California, Davis Plant Sciences Department. Additional 2010 Nonpareils were hand-harvested from commercial fields. Nonpareil almonds were also purchased from a local grocery store. Double kernels, insect damaged nuts, and almonds with damaged seed coats were excluded from these studies.

The different seed lots, whether hand-harvested or obtained from processor lots, cannot be considered statistically representative of their varieties because they were not grown adjacent to each other within the same field under the same growing conditions. Therefore any differences between these seed lots cannot be claimed to definitively typify comparative varietal differences.

2.2. Surface area modeling

A preliminary dataset ($n = 343$) was assembled in 2009–2010 in order to assess the sphere-based surface area estimation commonly used in the literature. This data set was used in preliminary modeling to screen a large number of potentially important indicators of surface area of intact almonds not exhibiting chip and scratch. It is vital to first establish the total expected surface area in order to later evaluate its absence. Once the best correlates were identified, the model was repeated with a final larger dataset ($n = 811$) of almonds with intact seed coats in order to increase precision. Statistical validity requires that an empirical model be tested on different samples than those that were used in the model creation. Therefore a separate dataset of Butte–Padre and Nonpareil almonds ($n = 510$) was used to validate the model and test its accuracy at quantifying fractional missing seed coat. This final dataset was intended to mimic missing seed coat as occurs in almonds exhibiting chip and scratch defect.

2.2.1. Data collection for preliminary modeling of intact seed coats

Length, width, thickness, mass, volume, and surface area were measured in 72 Monterey almonds, 139 Nonpareil almonds, and 132 Butte–Padre almonds. Volume was measured in 91 of the 139 Nonpareils. The total number or almonds evaluated for each variety was dictated by contributions from collaborators. Length (longest dimension), width (middle dimension), and thickness (smallest dimension) of the almonds were measured with digital calipers. Volume was determined by immersing one almond at a time into a 5 mL beaker completely full of corn oil, and weighing the amount of oil displaced when the almond was added. Displaced oil was weighed instead of measuring it volumetrically in a graduated cylinder because higher precision can be achieved by weighing the displaced liquid. Previous authors used toluene as the displaced fluid (Mohsenin, 1978; Aydin, 2003), but corn oil was utilized in this study due to its non-toxic nature, similarly low surface tension, and higher density. The masses of displaced oil were converted to volume using a corn oil density value of 0.922 g/cm^2.

Geometric mean diameter, density and sphericity were calculated from the following equations (Mohsenin, 1978):

$$\text{Geometric mean diameter} = (\text{Length} \times \text{Width} \times \text{Thickness})^{1/3}$$
$$\text{Density} = \text{mass/volume}$$
$$\text{Sphericity} = [(\text{geometric mean diameter})/\text{length}] \times 100$$

Surface area could hypothetically be quantified by removing and measuring the seed coat of an almond. Attempts to remove the seed coats of dry almonds were unsuccessful. However, almonds could be manually peeled after soaking in water (imbibition) overnight (Gradziel, 2009). Almond lengths, widths, thicknesses, masses, and volumes were measured again after imbibition in cold (4 °C) distilled water and before peeling. In cases where the volumes were measured before and after imbibition, corn oil was removed from the almond surface with a solution of diluted dish soap (Proctor & Gamble, San Ramon, CA) to remove any barrier to water absorption, and almonds were imbibed in 6-well plastic tissue culture trays (Corning, Corning, NY). Once removed, each seed coat was flattened onto opaque white plastic using transparent tape (Scotch packaging tape, 3M), and digitally photographed. Images of seed coats were analyzed using ImageJ software (NIH, Bethesda, MD) to obtain surface area (Reinking, 2007). The identity of each individual almond was maintained for each measurement so that trends between physical parameters could be observed.

2.2.2. Preliminary data analysis and modeling

A multivariate linear model was created with JMP software (version 10, SAS Institute Inc., Cary, NC) using a stepwise method to predict measured imbibed surface area on the 344 almonds made up of Nonpareil, Monterey, and Butte–Padre almonds from the 2009 crop. Thirty-three possible model effects were used in the preliminary model. One input was almond variety, and other inputs included length, width, thickness, geometric mean diameter, mass, volume, density, and sphericity. For each of these 8 measurements, 4 numbers were input: initial and imbibed measurements, as well as net and relative changes due to imbibition. Preliminary analysis by stepwise multivariate regression used the minimum Bayesian Information Criterion (SAS Institute Inc., 2010) to limit the number of model inputs. Surface area could be predicted to an r^2 of 0.6627 and root mean square error of 34.872 mm^2 using just three terms, i.e. variety, imbibed length and imbibed mass. Preliminary modeling analysis was repeated for individual almond varieties, and imbibed geometric mean density was also determined to be predictive. Therefore, for almonds measured in 2010–2011, subsequent to 2009 preliminary modeling, only imbibed length, width, thickness, and mass were recorded.

2.3. Final modeling

As with the preliminary modeling, empirical surface area modeling was performed using multivariate linear stepwise regression with the minimum Bayesian Information Criterion in order to select appropriate model inputs from imbibed length, width, thickness, mass, and geometric mean diameter. Least squares regression was then performed using inputs selected by the stepwise regression.

2.4. Quantification of missing seed coat

For this test set, Butte–Padre and Nonpareil almonds not previously included in preliminary modeling were imbibed, measured, and peeled following the same technique described in Sections 2.2 and 2.3. To mimic chip and scratch from mechanical damage or seed coat separation (skin slip), each seed coat was physically divided into two pieces of differing sizes. For some of the data, equal numbers of almonds were divided into ranges of 0–25%, ~50%, and 75–100% missing. For each divided almond seed coat, one piece was assigned to represent the 'missing' seed coat that might be removed by chipping, scratches, or skin slip, and the other was considered the 'remaining' one. Since all almonds had intact seed coats at the beginning of the experiment, the cumulative surface area of the two pieces was equal to the original total surface area of the almond. The measured percent of missing seed coat was calculated as:

$$(\text{Missing seed coat})/(\text{predicted total using empirical model}) \times 100 \\ = \text{measured percent of seed coat missing}$$

The estimated percent of missing seed coat was calculated as:

$$[(\text{Predicted total using empirical model}) \\ - (\text{remaining seed coat})]/(\text{Predicted total using empirical model}) \\ \times 100 = \text{estimated percent of missing seed coat}$$

Measured and predicted percentages of missing seed coat were compared using least squares fit analysis. To evaluate accuracy of the missing seed coat model in its lowest range, the subset of almonds missing 20% or less of their seed coat ($n = 256$) was examined separately.

3. Results and discussion

3.1. Distributions of physical properties

Almonds from seed lots of different varieties were significantly different from those in seed lots of other varieties in terms of their mean length, width, thickness, dry volume, and surface area by ANOVA and Tukey–Kramer LSD tests (Table 1). Monterey almonds had the greatest length and surface area, and the smallest width, thickness, and dry volume. Nonpareil almond length, width, thickness, dry volume, and surface area were intermediate, while Butte–Padre almonds had the smallest length and surface area. Interestingly, mean imbibed volumes from each variety were not statistically different from one another, despite pre-imbibed differences.

Fig. 1A–C illustrates that the volume and surface area distributions trend toward normal. Monterey and Nonpareil lengths, Nonpareil and Butte–Padre thicknesses and dry masses, Monterey imbibed volumes, Monterey and Nonpareil imbibed masses, and the measured surface areas of all the varieties qualify as normal distributions using the Shapiro–Wilk test for normality.

Thicknesses reported here were not significantly different from those determined by Ledbetter and Sisterson, while lengths, widths, and masses were 2 mm, 0.68 mm, and 0.13 g higher than their values, which were significantly different by Student's *t* test or Welch's *t* test for unequal variances ($p < 0.0001$) (Ledbetter and Palmquist, 2006; Ledbetter and Sisterson, 2010).

Ledbetter and Sisterson (2010) calculated that dry Nonpareil almonds had an average surface area of 569+/−58 mm^2, as estimated by the sphere-based method. This is significantly different by two-tailed *T*-test than 497.29+/−46.36 mm^2 calculated by the sphere-based method from dry measurements in the present study. Due to its multiplicative nature, the sphere-based estimation method is sensitive to small differences in measured dimensions. Both of these estimated dry surface areas are much lower than the mean measured imbibed surface area for Nonpareils, which was 672+/−51 mm^2. It is hard to say how much of this difference is due to noise within the model, and how much is due to the growth of every measured parameter of almonds during imbibition.

Previous authors estimated surface area by doubling the area of each almond's two-dimensional projection. Using that method, almonds (*cv. Taşbadem*) reportedly had a mean surface area of 178 mm^2 (Aydin, 2003), and *Fra Giulio Grande* almonds reportedly had a mean surface area of 320 mm^2 (Pliestic et al., 2008). Results using this method have not been published for Nonpareil almonds. The sphere-based surface area estimation formula can be used to compute geometric mean diameters from these reported surface area estimates in order to assess the likelihood the values. Based on these calculations, Aydin's reported surface area estimates should correspond to geometric mean diameters of 7.5 and 10.1 mm^2, which are significantly lower than the values of 11.4+/−0.15 mm^2 and 15.33+/11.28 mm^2 they report for their almonds. The apparent inaccuracy of the two-dimensional estimation is likely because it is based on only two measurable parameters.

3.2. Correlations within and between physical properties

Despite normal or normal-trending distributions of length, width, and height, these properties scale weakly, if at all, with one another (r^2 = 0.01–0.38, Table 2). Longer almonds are not more likely to be wider or thicker. Ledbetter and Sisterson also found that these physical properties scale weakly (2010).

Fig. 2A–C shows graphical representations of the low correlations between length, width, and heights. Each variety clusters together when graphed by any two properties, but within each variety, no linearity was discernible. Based on observations of this clustering effect, it is possible that performing a regression of composite variables made up of the three diameters and mass could help to develop objective varietal standards. Ledbetter and Sisterson have reported similar results in their 2013 paper, clustering

Table 1
Mean physical properties of Monterey, Nonpareil, and Butte–Padre almonds.

Variety	Length (mm)	Width (mm)	Thickness (mm)	Mass (g)	Dry volume (mm^3)	Imbibed volume (mm^3)	Imbibed surface area
Nonpareil	**21.94**b[a] (1.21)[*]	**12.07**b (0.49)	**7.92**b (0.45)	**1.03**a (0.11)	**952.46**b (116.64)	**1428**a (16.80)	**671.91**b (51.21)
Monterey	**23.32**a (1.39)	**10.87**c (0.64)	**7.23**c (0.80)	**0.92**b (0.15)	**818.30**c (179.85)	**1461**a (23.61)	**691.21**a (75.05)
Butte–Padre	**19.91**c (1.27)	**12.41**a (0.51)	**8.82**a (0.50)	**1.04**a (0.11)	**1014.87**a (123.56)	**1394**a (17.38)	**634.36**c (57.33)

[a] Within each column, means marked with different letters were significantly different by Tukey–Kramer LSD test.
[*] Standard deviations are included in parentheses.

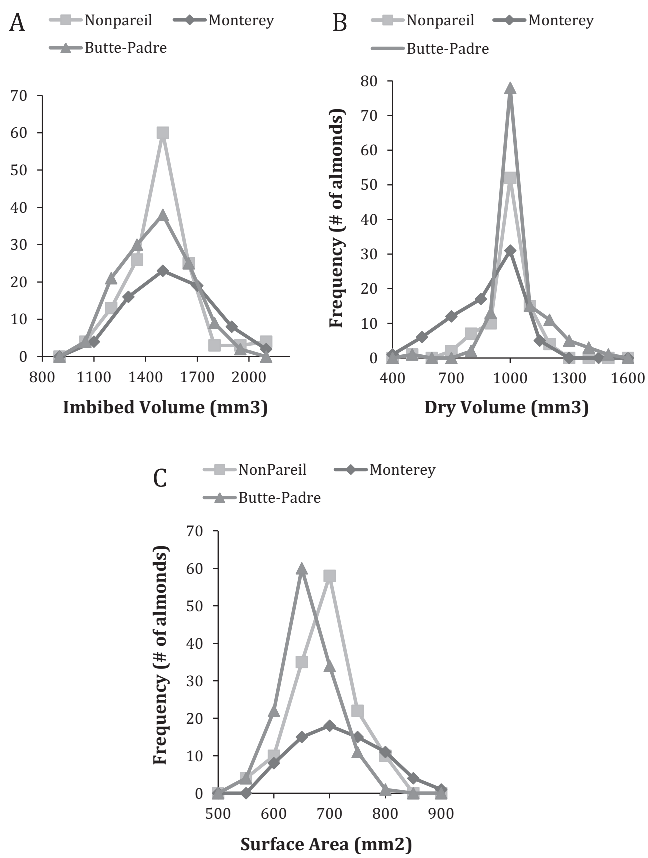

Fig. 1. (A–C) Dry volume (A), imbibed volume (B), and surface area (C) distributions of Nonpareil, Monterey, Butte–Padre almonds.

Table 2
Coefficients of determination between the three dimensions of individual almonds.

Variety	Coefficients of determination		
	Length to width	Length to thickness	Width to thickness
Nonpareil			
Dry	0.075	0.007	0.016
Imbibed	0.114	0.025	0.120
Monterey			
Dry	0.337	0.045	0.125
Imbibed	0.317	0.128	0.138
Butte–Padre			
Dry	0.380	0.011	0.006
Imbibed	0.275	0.014	0.056

Nonpareil cultivars apart from California and Mission cultivars using multivariate analysis (2013).

Inputs into the sphere-based surface area model include almond length, width, and height (or thickness), with the assumption that these physical properties scale with mass and surface area. Lengths of all varieties were shown to correlate moderately well to masses. However, this study determined that almond widths and thicknesses correlate somewhat weakly to mass and surface area, with coefficients of determination ranging between 0.222–0.580, and 0.025–0.390, respectively (Table 3). Out of 30 linear regressions tested between width or thickness and either mass or measured surface area within the different varieties, only the width and thickness of Monterey almonds explained more than half the variation in mass. It appears that width and thickness of any almond variety cannot be employed independently when estimating almond mass and surface area.

Published values for coefficients of determination between Nonpareil length and mass range from 0.656 to 0.82 (Ledbetter and Sisterson, 2010), compared to 0.526 reported here. Other investigators have reported that coefficients of determination between Nonpareil mass and width ranged from 0.750 and 0.79, compared to 0.355 reported here (Ledbetter and Palmquist, 2006; Ledbetter and Sisterson, 2010). None of the measured dimensions correlated strongly with surface area.

Samples used in these experiments were obtained from more diverse sources than those used in previously published results. Both studies previously reported in the literature sampled almonds within one orchard over different lengths of time (Ledbetter and Palmquist, 2006; Ledbetter and Sisterson, 2010). Our sampling was more comprehensive, therefore, it is expected that coefficients of determination may be slightly lower. Dissimilar kernels within almond lots could be partly responsible for the higher variation. It also may be that correlations of mass versus either length or width vary greatly by orchard, although mass versus thickness did not correlate to each other, irrespective of the sample source.

3.3. Preliminary evaluation of the sphere estimate of surface area

Ledbetter and Sisterson (2010) reported a strong correlation between dry almond mass and the sphere model of surface area, and similar results were found with our study (Fig. 3A). Estimated almond surface area based on the sphere model did indeed scale with almond mass with a coefficient of determination of $r^2 = 0.892–0.924$, depending on variety.

Unlike the relationship between mass and sphere modeled surface area, imbibed mass only somewhat correlated to actual measured surface area, as shown in Fig. 3B. Imbibed rather than dry mass was compared to measured surface area, since surface area was only measurable after imbibition. Coefficients of determination ranged from $r^2 = 0.31–0.78$, depending on variety (Table 4). Monterey almond imbibed masses exhibited a linear relationship with measured surface area ($r^2 = 0.779$), but mass accounted for less than 60% and 40% of the variation in surface area in Nonpareils and Butte–Padres, respectively.

Coefficients of determination between measured surface area and dry mass were even weaker than with imbibed mass, ranging from $r^2 = 0.32–0.60$. It is likely that the low correlation between measured surface area and mass can be related back to the low correlations between diameters. Natural variation in almond shape, size, and dimensional proportions obfuscate intuitive theoretical

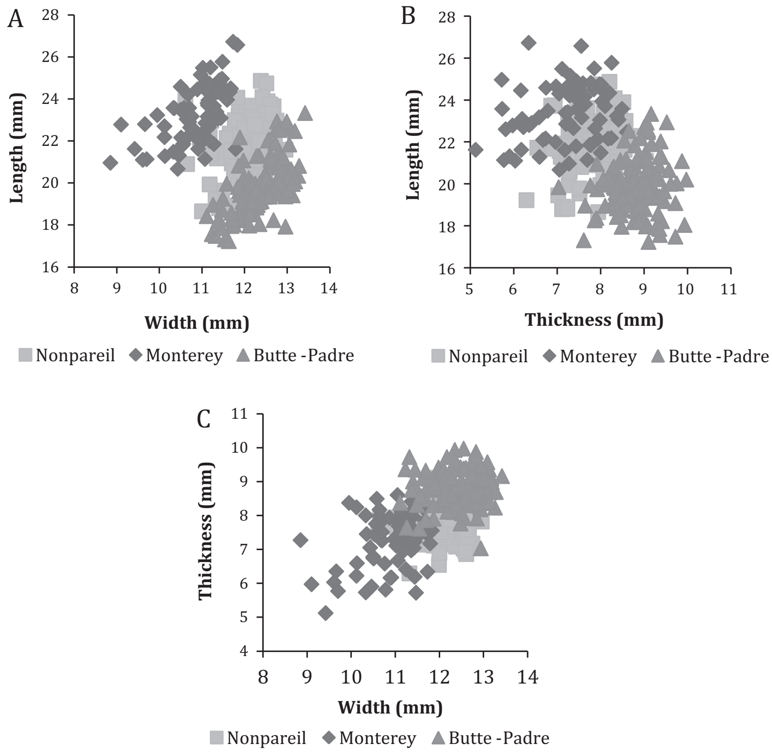

Fig. 2. (A–C) Dimensional properties of individual almonds cluster non-linearly by variety.

Table 3
Coefficients of determination between almond dimensions and either mass or surface area.

			Coefficients of determination (r^2)		
			To mass		To measured surface area
			Dry	Imbibed	
Nonpareil	Length	Dry	0.526	0.543	
		Imbibed	0.582	0.352	0.319
	Width	Dry	0.355	0.277	
		Imbibed	0.466	0.464	0.137
	Thickness	Dry	0.393	0.183	
		Imbibed	0.397	0.397	0.025
Monterey	Length	Dry	0.485	0.551	
		Imbibed	0.642	0.64	0.587
	Width	Dry	0.522	0.222	
		Imbibed	0.529	0.529	0.272
	Thickness	Dry	0.580	0.001	
		Imbibed	0.198	0.198	0.022
Butte–Padre	Length	Dry	0.690	0.513	
		Imbibed	0.534	0.578	0.546
	Width	Dry	0.487	0.451	
		Imbibed	0.447	0.447	0.390
	Thickness	Dry	0.222	0.091	
		Imbibed	0.285	0.285	0.026

relationships between many of the measurable characteristics of almonds.

It was previously hypothesized that regression slopes could potentially be used to differentiate between almond varieties (Ledbetter and Sisterson, 2010). However, when comparing regressions of dry mass and dry sphere-modeled surface area for each variety, the mean 95% confidence intervals of these slopes, based on the sum of standard errors, overlap with one another. Homogeneity of these slopes was tested by ANCOVA, and the interaction between sphere modeled surface area and variety is significant ($p > 0.0401$). Pairwise comparison using the Holm adjustment for multiplicity found that none of the slopes were significantly different from one another: i.e. Nonpareils and Montereys ($p > 0.0602$), Butte–Padres and Montereys ($p > 0.7082$), and Butte–Padres and Nonpareils ($p > 0.0813$).

Intercepts were not significantly different from one another, with the exception of Nonpareil versus Monterey almonds (Table 5a). The regression equation slopes and intercepts of imbibed mass versus actual measured surface area for all three varieties of almonds therefore cannot be effectively used to tell almond varieties apart. Multivariate analysis incorporating several measures will likely be the best way to establish objective varietal standards.

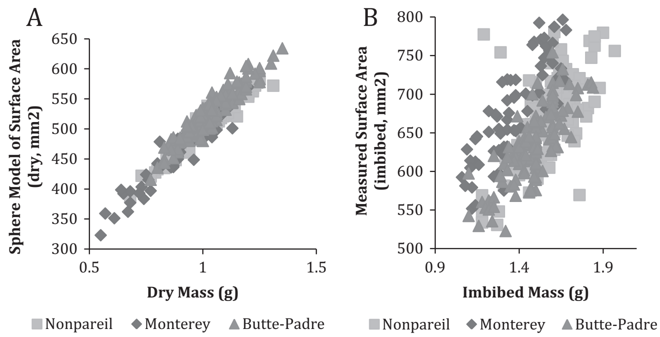

Fig. 3. (A and B) Predictive ability of mass for sphere-modeled surface area (A), and measured surface area (B). Coefficients of determination are in Table 4.

Table 4
Regression coefficients of determination (r^2) for mass versus measured surface area for Nonpareil, Monterey, and Butte–Padre almonds.

Variety	Mass	Coefficient of determination between mass and spherical model of surface area (mm^2)		And between mass and measured surface area (mm^2)
		Dry	Imbibed	Imbibed
Nonpareil	Dry	0.906	0.795	0.314
	Imbibed	0.717	0.933	0.388
Monterey	Dry	0.924	0.708	0.389
	Imbibed	0.182	0.708	0.779
Butte–Padre	Dry	0.892	0.816	0.598
	Imbibed	0.617	0.799	0.599

3.4. Large-scale evaluation of the sphere-estimate of surface area

The sphere-based model for surface area was evaluated by comparing via linear regression its predictions for 811 almonds to their measured surface areas, as described in Section 2.3. Since measured surface area is taken from imbibed almonds, the inputs into the sphere-based model were also taken after imbibition. The predictive accuracies for each variety, as well as over all the varieties together, were evaluated separately.

The average measured surface area of 756 Nonpareil almonds was 755.20 mm^2, 120 Monterey almonds averaged 672.63 mm^2, and 167 Butte–Padre almonds averaged 634.57 mm^2 (Table 6). The average surface area predicted by the sphere-based model was 4.0% lower than the average measured value for Monterey almonds, 9.1% lower for Nonpareil almonds, and 3.6% lower in Butte–Padre almonds.

The sphere model predicted about half the variation in measured surface area; coefficients of determination were 0.55, 0.62, and 0.44 for Nonpareil, Monterey, and Butte–Padre almonds. Regression equation slopes for Nonpareil and Monterey were close to 1.0 (1.03 and 0.98, respectively) but for Butte–Padre almonds, it was 0.74. Root mean square errors for Monterey and Butte–Padre almonds were 43.50 mm^2 and 42.88 mm^2 but for Nonpareils, it was nearly twice that: 73.25 mm^2. Based on these indicators, the sphere-based model for surface area does correlate strongly to measured surface area.

3.5. Empirical modeling of almond surface area

In order to model surface area empirically, stepwise multivariate linear regression was performed on 811 almonds with intact seed coats, as described in Section 2.3. The following empirical model resulted:

$$\begin{aligned}\text{Surface area} = {} & -278.741 + (\text{variety factor}) + 15.363 \\ & \times (\text{imbibed length}) + 34.117 \times (\text{imbibed width}) \\ & + 99.109 \times (\text{imbibed mass})\end{aligned}$$

The variety factor for Nonpareil almonds was −21.93, 23.61 for Monterey almonds, and −1.68 for Butte–Padre. This regression is depicted in Fig. 4.

Compared to the sphere-based model, the empirical model more closely predicts measured surface area of each variety separately, as well as predicting actual surface area of all the varieties combined. The empirical surface area model yielded

Table 5
(A and B) Regression slopes and intercept of dry (A) and imbibed (B) mass versus measured surface area for Nonpareil, Monterey, and Butte–Padre almonds. Standard errors are parenthetical.

Variety	Slope	+/−95% Confidence interval	Intercept	+/−95% Confidence interval
Panel A				
Nonpareil	313.10 (8.22)	16.25	188.42 (8.40)	16.61
Monterey	345.57 (11.82)	23.57	151.27 (11.00)	21.93
Butte–Padre	340.21 (10.44)	20.65	172.68 (15.86)	31.38
Panel B				
Nonpareil	217.3 (23.42)	46.31	326.62 (36.74)	72.65
Monterey	306.66 (19.55)	38.98	257.11 (27.00)	55.81
Butte–Padre	250.67 (17.94)	35.49	264.03 (26.75)	52.92

Table 6
Predictive accuracy of the sphere-based surface area model, as evaluated by linear regression to measured surface area.

Model parameter	Nonpareil	Monterey	Butte–Padre	Over all varieties
n	756	120	167	1043
Mean measured surface area in mm^2 (standard deviation)	**755.19** (112.09)	**672.63** (69.84)	**634.57** (57.23)	**726.41** (111.67)
Mean sphere model estimate of surface area (standard deviation)	**724.76** (78.88)	**611.17** (55.88)	**643.27** (51.32)	**698.68** (84.58)
Intercept (standard error)	**9.36** (24.67)	**73.40** (43.79)	**157.83** (41.86)	**20.90** (17.84)
Slope (standard error)	**1.03** (0.03)	**0.98** (0.07)	**0.74** (0.07)	**1.01** (0.03)
Coefficient of determination (r^2)	**0.55**	**0.61**	**0.44**	**0.60**
Root mean square error in mm^2	73.24	43.49	42.88	69.12

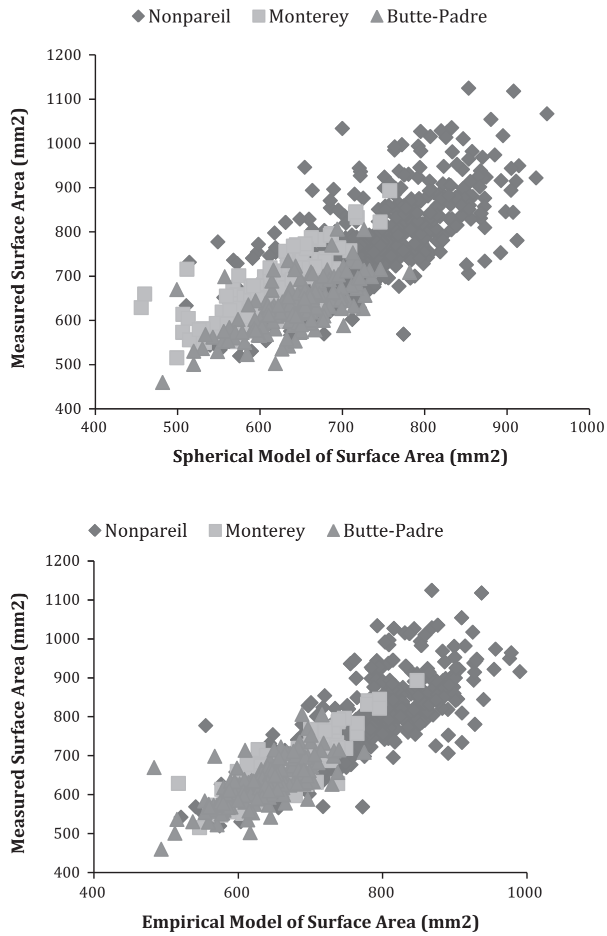

Fig. 4. Predictive accuracy of the empirical surface area model to measured surface area.

average predictions of 752.83 mm^2 for Nonpareil almonds, 669.342 mm^2 for Monterey almonds, and 639.204 mm^2 for Butte–Padre almonds. These averages were 0.31%, 0.49%, and −0.73% lower than actual measured values. Empirical model predictions were an order of magnitude more accurate than predictions based on the sphere model.

The empirical model correlates to measured surface area with r^2 of 0.48–0.73, depending on the almond variety (Table 7). The regression equation slopes for both measured and sphere-based surface area were close to 1.0 for measured and empirically modeled surface area (0.74–1.03 versus 0.79–1.06), but the root mean square errors of the empirical model's regression are 4–18% lower than those of the sphere based model.

Monterey almond surface area can be calculated the most accurately of the three varieties using either model, while Butte–Padres have the most unaccounted-for variation in both models. This is almost certainly because Butte–Padre samples are made up of a mixture of both Butte and Padre varieties, and the morphological differences between these two introduces variation into the modeling. Repetition of the multivariate regression analysis using only data from Nonpareil or from Butte–Padre almonds while excluding data from the other varieties did not yield a model with a higher coefficient of determination between measured and estimated surface area (data not shown). It is likely that the increased specificity of the single-variety models was outweighed by the larger number of samples made available for modeling when all three varieties were used in the statistical analysis.

Table 7
Predictive accuracy of the empirical surface area model, as evaluated by linear regression to measured almond surface area.

Model parameter	Nonpareil	Monterey	Butte–Padre	Over all varieties
n	756	120	167	1043
Mean measured surface area in mm^2 (standard deviation)	**755.20** (112.09)	**672.63** (69.84)	**634.57** (57.23)	**726.41** (111.67)
Mean sphere model estimate of surface area (standard deviation)	**724.77** (78.88)	**611.18** (55.88)	**643.28** (51.32)	**698.69** (84.58)
Mean empirical model estimate of surface area in mm^2 (standard deviation)	**752.83** (85.21)	**669.34** (56.87)	**639.20** (50.04)	**725.10** (90.19)
Intercept (standard error)	**−44.27** (19.98)	**−28.45** (39.66)	**128.01** (41.16)	**−32.08** (14.24)
Slope (standard error)	**1.06** (0.03)	**1.05** (0.06)	**0.79** (0.06)	**1.05** (0.02)
Coefficient of determination (r^2)	**0.68**	**0.73**	**0.48**	**0.74**
Root mean square error in mm^2	61.57	36.66	41.39	56.56

3.6. Quantification of missing seed coat

In order to quantify the amount of seed coat missing in almonds exhibiting the USDA defect chip and scratch, the surface area of the remaining seed coat of each almond was compared to a predicted total using an empirical surface area model, as described in Section 2.4.

Measured % seed coat removed was calculated as:

$$(\text{measured total seed coat} - \text{measured remaining seed coat})/(\text{measured total seed coat})$$

Estimated % of seed coat removed was calculated as:

$$(\text{estimated total seed coat using empirical model} - \text{measured seed coat remaining})/(\text{estimated total seed } \textit{coat using empirical model total}$$

Fig. 5 shows the prediction accuracy of the missing seed coat estimate for each almond. With greater portions of seed coat missing (i.e. as the measured remaining seed coat goes to zero), the slope calculation approaches total seed coat/total seed coat, and the variance decreases. Therefore r^2 is not a useful evaluation of this missing seed coat surface area model. However, root mean square error in the range of 0–20% missing seed coat was 5.60% (n = 256, as described in Section 2.4). When the sphere-based model for estimating surface area was used in place of the

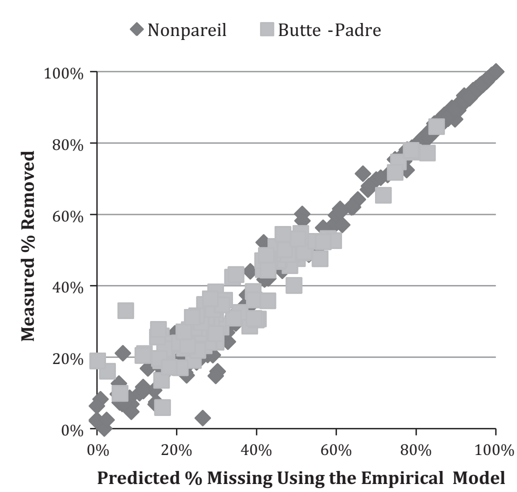

Fig. 5. Predictive accuracy of missing surface area using the empirical model, compared to the actual measured percent of surface area that was removed.

empirical model, the root mean square error in the range of 0–20% seed coat missing was 5.67%. It appears that the higher variation in the sphere-based surface area model does not cause high variation in this application.

Surface area model values are only one component of the predicted missing calculation; therefore variation in predicted surface area contributes only a small fraction of the predicted missing value. Additionally, some of the systemic error in the surface area model may cancel out, since the model surface area term appears in both the numerator and denominator of the percent missing prediction equation.

4. Conclusions

In this study, basic physical properties of almonds were measured to determine which correlate strongly to surface area and therefore could predict it. Kernels were imbibed, peeled, and the surface areas of the seed coats were quantified. The resulting measured values for surface area provided a way to evaluate the relationships between the physical properties of almonds, and several potential methods for predicting almond surface area.

As previously concluded by other researchers, estimation of almond surface area based on a model sphere of equivalent average diameter correlated highly to almond mass with a coefficient of determination of $r^2 = \sim 0.90$ (Ozguven and Vursavus, 2005; Ledbetter and Sisterson, 2010). However, this correlation alone did not validate the use of this equation. Estimation of almond surface area based on a sphere of equivalent diameter only explained about half the variation in measured surface area.

Surprisingly, individual dimensions (length, width, and thickness) did not increase with increasing surface area, nor they did scale in proportion to one other. However, these measurements did cluster nonlinearly within almonds of the same variety. Based on this pattern, also observed by Ledbetter and Sisterson (2013), it is feasible that almond varieties could be defined based on composite variables of empirical measurements.

Given that measured surface area did not correlate strongly with measured parameters, the complex relationship between dimensions and mass did not lend itself to theoretical modeling. Therefore an empirical model was created, which depends on the almond variety, as well as length, width, and mass after imbibition. This proposed model for surface area is fundamentally different from previous models because it is the first to be based on physical measurements of almond skin, thereby including all features and nuances of the kernel's topography. It could be used to improve the accuracy of mass transfer and energetic transfer calculations in almond processing. Future work is needed to determine the relationship between the surface areas of imbibed and dry almonds. However, to date no method has been published for measuring the surface area of a dry almond.

Both the empirical and sphere-based surface area models can be used to quantify the amount of surface area missing in almonds exhibiting the USDA chip and scratch defect. By comparing the predicted total surface area and the measured surface area of any remaining seed coat, a quantitative percentage of seed coat lost can be obtained. This new method could potentially be used to quantify the USDA grading standard for the defect of chip and scratch.

The USDA grading standard for 'chip and scratch' defect classifies almonds as defective if an aggregate diameter of 0.25 in. (32.15 mm^2) of seed coat is missing. Ninety-five percent of all the 1545 almonds measured during this study ranged from 515.96 mm^2 to 942.24 mm^2 of total surface area, placing the critical threshold of missing seed coat at 3.41–6.23%. Therefore the current USDA standard could not be completely replaced by this model because the USDA grading threshold occurs where this model has its lowest sensitivity. However, the quantification of missing seed coat beyond the first few percent yields useful and currently uncollected information. This may be used to explore many aspects of almond physiology dependent on or resultant from seed coat integrity, such as germination, rehydration kinetics, processing damage, or blanching efficacy.

Acknowledgements

This research was supported by the Almond Board of California. The authors wish to thank Jordan Wilheim, Melanie Meinen, and Erica Kenney for technical assistance in the both the laboratory and the field. We are grateful for the editorial assistance of Drs. Thomas Gradziel and Judy Jernstedt, and the statistical assistance of Dr. Jerome Braun.

References

Almond Board of California, A., 2009. Technical Information Kit. Modesto, Almond Board of California: 16.
Almond Board of California, A., 2010. 2010 Almond Almanac. A. Almond Board of California, Modesto.
Almond Board of California, A., 2011. 2011 Almond Almanac. USDA Incoming Receipts from FV 193 certificates reported to Almond Board of California. A. Almond Board of California. Modesto, Almond Board of California: 17.
Aydin, C., 2003. Physical properties of almond nut and kernel. J. Food Eng. 60 (3), 315–320.
Baten, W.D., Marshall, R.E., 1943. Some methods for approximate prediction of surface area of fruits. J. Agric. Res. 66, 0357–0373.
Bayram, M., 2011. Comparison of unsplit inshell and shelled kernel of the pistachio nuts. J. Food Eng. 107 (3–4), 374–378.
Boriss, H.B.H., 2011, March 2011. Almond Profile. Commodities & Products Retrieved May 4th, 2011, 2011. <http://www.agmrc.org/commodities__products/nuts/almond_profile.cfm>.
Frechette, R.J.a.Z.J.W., 1966. Surface area-weight relationships for McIntosh apples. Trans. ASABE 9 (4), 526–527.
Gradziel, T., 2009. Rapid Seed Viability Assay for Monitoring Postharvest Pasteurization Strategies. Department of Plant Sciences, University of California, Davis.
Ledbetter, C.A., Palmquist, D.E., 2006. Comparing physical measures and mechanical cracking products of 'Nonpareil' almond (*Prunus dulcis* Mill. D.A. Webb.) with two advanced breeding selections. J. Food Eng. 76 (2), 232–237.
Ledbetter, C.A., Sisterson, M.S., 2010. Carpological variability of almond *Prunus dulcis* (Mill.) DA Webb cv. Nonpareil in a single orchard during seven consecutive harvests. HortScience 45 (12), 1788–1792.
Ledbetter, C.A., Sisterson, M.S., 2013. Distinguishing Non[areil marketing group almond cultivars through multivariate analysis. J. Food Sci. 78 (9), S1430–S1436.
Mohsenin, N., 1978. Physical Properties of Plant and Animal Materials. Gordon and Breach Science Publishers Ltd., New York.
Ozguven, F., Vursavus, K., 2005. Some physical, mechanical and aerodynamic properties of pine (*Pinus pinea*) nuts. J. Food Eng. 68 (2), 191–196.
Pliestic, S., Dobricevic, N., et al., 2008. Influence of moisture content on physical and mechanical properties of almond (*Prunus dulcis* cv. Fra Giulio Grande). Trans. Asabe 51 (2), 653–659.
Reinking, L., June 2007, Updated June 2007. Examples of Image Analysis Using ImageJ. <http://imagej.nih.gov/ij/docs/pdfs/examples.pdf>.
SAS Institute Inc., S. (Ed.), 2010. JMP® 9 Modeling and Multivariate Methods. Cary, NC, SAS Institute Inc.
Service, A.M., 2004. Almonds Grown in California; Revision of Quality Control Provisions. Federal Register. D. o. Agriculture: 57820-57822.