Academic Journal of Plant Sciences 4 (4): 118-123, 2011
ISSN 1995-8986

Prediction of Apple Mass Based on Some Geometrical Properties Using Linear Regression Models

Fereydoun Keshavarzpour and Majid Rashidi

Department of Agriculture, Shahr-e-Rey Branch, Islamic Azad University, Shahr-e-Rey, Iran

Abstract: In this study, nine linear regression models for predicting apple mass from some geometrical properties of apple such as length (L), diameter (D), geometrical mean diameter (GMD), first projected area (PA_1), second projected area (PA_2), criteria area (CAE) and estimated volume based on an oblate spheroid assumed shape (V_{Sp}) were suggested. In order to predict apple mass based on outer dimensions, the mass model based on GMD as M = - 168.5 + 47.01 GMD with R^2 = 0.77 was preferred. In addition, to predict apple mass based on projected areas, the mass model based on CAE as M = - 26.82 + 4.948 CAE with R^2 = 0.77 was selected. Moreover, to predict apple mass based on estimated volume, the mass model based on V_{Sp} as M = 20.68 + 0.814 V_{Sp} with R^2 = 0.76 was chosen. However, the statistical results of study indicated that the liner regression models underpredict mass of apple and can not be appropriately used.

Key words: Apple · Mass · Geometrical properties · Prediction · Linear models · Iran

INTRODUCTION

The apple is the pomaceous fruit of the apple tree, species *Malus domestica* in the rose family (*Rosaceae*). It is one of the most widely cultivated tree fruits. There are more than 7500 known cultivars of apples [1]. At least 55 million tones of apples were grown worldwide in 2005, with a value of about $10 billion. China produced about 35% of this total. The United States is the second-leading producer, with more than 7.5% of world production. Iran is third, followed by Turkey, Russia, Italy and India [2]. Iranian apple are not exported because of variance in size and shape and lack of proper packaging [3].

Similar to other fruits, apple size is one of the most important quality parameters for evaluation by consumer preference. Consumers prefer fruits of equal size and shape [4, 5]. Sorting can increase uniformity in size and shape, reduce packaging and transportation costs and also may provide an optimum packaging configuration [6]. Moreover, sorting is important in meeting quality standards, increasing market value and marketing operations [7, 8]. Sorting manually is associated with high labor costs in addition to subjectivity, tediousness and inconsistency which lower the quality of sorting [9]. However, replacing human with a machine may still be questionable where the labor cost is comparable with the sorting equipment [10]. Studies on sorting in recent years have focused on automated sorting strategies and eliminating human efforts to provide more efficient and accurate sorting systems which improve the classification success or speed up the classification process [11, 12].

Physical and geometrical characteristics of products are the most important parameters in design of sorting systems. Among these characteristics, mass, outer dimensions, projected areas and volume are the most important ones in sizing systems [13]. The size of produce is frequently represented by its mass because it is relatively simple to measure. However, sorting based on some geometrical attributes may provide a more efficient method than mass sorting. Moreover, the mass of produce can be easily estimated from geometrical attributes if the mass model of the produce is known [14-17]. Therefore, modeling of apple mass based on some geometrical attributes may be useful and applicable. The main objectives of this research were to determine suitable mass models based on some geometrical properties of apple and to verify selected mass models by comparing their results with those of the measuring method.

MATERIALS AND METHODS

Experimental Procedure: One of the commercial varieties of apple in Iran, i.e. Damavandi (Fig. 1) was considered for

Corresponding Author: Dr. Majid Rashidi, Ph.D., Department of Agriculture, Shahr-e-Rey Branch, Islamic Azad University, Shahr-e-Rey, Iran.

Fig. 1: Apple (Damavandi variety)

this study. One hundred and fifteen randomly selected apples of various sizes were purchased from an orchard located in Damavand, Iran. Apples were selected for freedom from defects by careful visual inspection, transferred to the laboratory and held at 5±1°C and 90±5% relative humidity until experimental procedure. In order to obtain required parameters for determining mass models, the mass of each apple was measured to 1.0 g accuracy on a digital balance. By assuming the shape of apple as an oblate spheroid, the outer dimensions of each apple, i.e. length (L) and diameter (D) was measured to 0.1 cm accuracy by a digital caliper. The geometric mean diameter (GMD) of each apple was then calculated by equation 1.

$$GMD = (LD^2)^{1/3} \quad (1)$$

Two projected areas of each apple, i.e. first projected area (PA_1) and second projected area (PA_2) were also calculated by using equations 2 and 3, respectively. The average projected area known as criteria area (CAE) of each apple was then determined from equation 4.

$$PA_1 = \pi D^2/4 \quad (2)$$
$$PA_2 = \pi LD/4 \quad (3)$$
$$CAE = (PA_1+2PA_2)/3 \quad (4)$$

In addition, the estimated volume of each apple by assuming the shape of apple as an oblate spheroid (V_{Sp}) was calculated by using equation 5.

$$V_{sp} = \pi LD^2/6 \quad (5)$$

Table 1 shows physical and geometrical properties of the 100 apples used to determine mass models. Also, in order to verify mass models, physical and geometrical properties of fifteen randomly selected apples of various sizes were determined as described before. Table 2 shows physical and geometrical properties of the fifteen apples used to verify mass models.

Regression Model: A typical multiple variables linear regression model is shown in equation 6:

$$Y = k_0 + k_1X_1 + k_2X_2 + \ldots + k_nX_n \quad (6)$$

Where:

Y	=	Dependent variable, for example mass of apple
$X_1, X_2, \ldots, X_n$	=	Independent variables, for example geometrical properties of apple
$k_0, k_1, k_2, \ldots, k_n$	=	Regression coefficients

In order to estimate apple mass from geometrical properties, nine linear regression mass models were suggested and all the data were subjected to linear regression analysis using the Microsoft Excel 2007. Models were divided into three main classifications (Table 3), i.e. first classification (outer dimensions), second classification (projected areas) and third classification (estimated volume).

Statistical Analysis: A paired samples t-test was used to compare the apple mass values predicted using models with the apple mass values measured by digital balance. Also, to check the discrepancies between the apple mass values measured by digital balance with the apple mass values predicted by mass models, root mean squared error (RMSE) and mean relative percentage deviation (MRPD) were calculated using the equations (7) and (8), respectively [18]:

Table 1: The mean values, standard deviation (S.D.) and coefficient of variation (C.V.) of physical and geometrical properties of the 100 randomly selected apples used to determine mass models

Parameter	Minimum	Maximum	Mean	S.D.	C.V. (%)
Mass (M), g	83.00	144.0	118.5	17.07	14.41
Length (L), cm	5.000	7.000	6.042	0.365	6.044
Diameter (D), cm	5.350	6.800	6.140	0.341	5.560
Geometrical mean diameter (GMD), cm	5.300	6.698	6.105	0.317	5.198
First projected area (PA_1), cm^2	22.48	36.32	29.70	3.253	10.95
Second projected area (PA_2), cm^2	21.85	36.02	29.20	3.014	10.32
Criteria area (CAE), cm^2	22.06	35.25	29.36	3.013	10.26
Estimated volume (V_{Sp}), cm^3	77.94	157.4	120.1	18.28	15.22

Table 2: The mean values, standard deviation (S.D.) and coefficient of variation (C.V.) of physical and geometrical properties of the fifteen randomly selected apples used to verify mass models

Parameter	Minimum	Maximum	Mean	S.D.	C.V. (%)
Mass (M), g	146.0	156.0	150.1	3.081	2.053
Length (L), cm	6.000	7.000	6.560	0.250	3.813
Diameter (D), cm	6.400	6.850	6.603	0.136	2.053
Geometrical mean diameter (GMD), cm	6.361	6.767	6.587	0.113	1.714
First projected area (PA_1), cm^2	32.17	36.86	34.26	1.410	4.116
Second projected area (PA_2), cm^2	30.87	36.31	34.02	1.364	4.008
Criteria area (CAE), cm^2	31.81	35.97	34.10	1.165	3.416
Estimated volume (V_{Sp}), cm^3	134.8	162.2	149.8	7.683	5.129

Table 3: Nine linear regression mass models and their relations in three classifications

Classification	Model No.	Model	Relation
Outer dimensions	1	$M = k_0 + k_1 L$	$M = -72.50 + 31.61 L$
	2	$M = k_0 + k_1 D$	$M = -144.4 + 42.82 D$
	3	$M = k_0 + k_1 GMD$	$M = -168.5 + 47.01 GMD$
	4	$M = k_0 + k_1 L + k_2 D$	$M = -167.3 + 11.35 L + 35.38 D$
Projected areas	5	$M = k_0 + k_1 PA_1$	$M = -14.92 + 4.491 PA_1$
	6	$M = k_0 + k_1 PA_2$	$M = -21.72 + 4.801 PA_2$
	7	$M = k_0 + k_1 CAE$	$M = -26.82 + 4.948 CAE$
	8	$M = k_0 + k_1 PA_1 + k_2 PA_2$	$M = -26.17 + 2.559 PA_1 + 2.351 PA_2$
Estimated volume	9	$M = k_0 + k_1 V_{Sp}$	$M = 20.68 + 0.814 V_{Sp}$

Table 4: Mass models, p-value of model variable(s) and coefficient of determination (R^2)

	p-value							
Model No.	L	D	GMD	PA_1	PA_2	CAE	V_{Sp}	R^2
1	1.16E-14	---	---	---	---	---	---	0.46
2	---	7.04E-30	---	---	---	---	---	0.73
3	---	---	1.79E-32	---	---	---	---	0.77
4	0.000152	8.48E-20	---	---	---	---	---	0.76
5	---	---	---	7.71E-30	---	---	---	0.73
6	---	---	---	---	9.47E-29	---	---	0.72
7	---	---	---	---	---	2.16E-32	---	0.77
8	---	---	---	1.19E-05	0.000159	---	---	0.76
9	---	---	---	---	---	---	3.93E-32	0.76

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(M_i - M_i^*)^2}{n}} \tag{7}$$

Where:

RMSE = Root mean squared error, g
M_i = Apple mass measured by digital balance, g
M^*_i = Apple mass predicted by mass model, g
n = Number of samples

$$MRPD = \frac{100 \times \sum_{i=1}^{n} \frac{\left| M_i - M_i^* \right|}{M_i}}{n} \tag{8}$$

Where:

MRPD = Mean relative percentage deviation, %

RESULTS

The p-value of the independent variable(s) and coefficient of determination (R^2) of all the linear regression mass models are shown in Table 4.

First Classification Models (Outer Dimensions): In this classification apple mass can be predicted using single variable linear regressions of length (L), diameter (D) and geometrical mean diameter (GMD) of apple, or multiple variables linear regression of apple outer dimensions (length and diameter). As indicated in Table 4, among the first classification models (models No. 1-4), model No. 3 had the highest R^2 values (0.77). Also, the p-value of independent variable (GMD) was the lowest (1.79E-32). Based on the statistical results model No. 3 was selected as the best model of first classification. Model No. 3 is given in equation 9.

$$M = -168.5 + 47.01\ GMD \quad (9)$$

Second Classification Models (Projected Areas): In this classification apple mass can be predicted using single variable linear regressions of first projected area (PA_1), second projected area (PA_2) and criteria area (CAE) of apple, or multiple variables linear regression of apple projected areas. As showed in Table 4, among the second classification models (models No. 5-8), model No. 7 had the highest R^2 values (0.77). Moreover, the p-value of independent variable (CAE) was the lowest (2.16E-32). Again, based on the statistical results model No. 7 was chosen as the best model of second classification. Model No. 7 is given in equation 10.

$$M = -26.82 + 4.948\ CAE \quad (10)$$

Third Classification Model (Estimated Volume): In this classification apple mass can be predicted using single variable linear regression of estimated volume of apple (V_{Sp}). As indicated in Table 4, model No. 9 had high R^2 value (0.76). In addition, the p-value of independent variable (V_{Sp}) was (3.93E-32). Once more, based on the statistical results model No. 9 was chosen as a suitable model. Model No. 9 is given in equation 11.

$$M = 20.68 + 0.814\ V_{sp} \quad (11)$$

DISCUSSION

Among the linear regression models (models No. 1-9), models No. 3, 7 and 9 were chosen based on the statistical results and a paired samples t-test was used to compare the apple mass values predicted using models No. 3, 7 and 9 with the apple mass values measured by digital balance. Also, to check the discrepancies between the apple mass values predicted by the models with the apple mass values measured by digital balance, RMSE and MRPD were calculated.

Comparison of Model No. 3 with Measuring Method: The apple mass values predicted by model No. 3 were compared with the apple mass values measured by digital balance and are shown in Table 5. A plot of the apple mass values predicted by model No. 3 and measured by digital balance with the line of equality (1.0: 1.0) is shown in Fig. 2. The paired samples t-test results indicated that the apple mass values predicted with model No. 3 were significantly less than the apple mass values measured by digital balance (Table 6). The mean apple mass difference between two methods was -8.89 g (95% confidence interval: -11.8 g and -5.97 g; P = 1.000). The standard deviation of the apple mass difference was 5.28 g. RMSE and MRPD were also used to check the discrepancies between the two methods. The amounts of RMSE and MRPD were 10.2 g and 6.15%, respectively. Thus, apple mass predicted by model No. 3 may be 10.2 g or 6.15% less than apple mass measured by a digital balance.

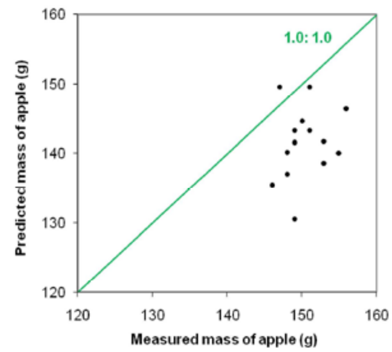

Fig. 2: Measured mass of apple and predicted mass of apple using model No. 3 with the line of equality (1.0: 1.0)

Comparison of Model No. 7 with Measuring Method: The apple mass values predicted by model No. 7 were compared with the apple mass values measured by digital balance and are shown in Table 5. A plot of the apple mass values predicted by model No. 7 and measured by digital balance with the line of equality (1.0: 1.0) is shown in Fig. 3. The paired samples t-test results indicated that the apple mass values predicted with model No. 7 were significantly less than the apple mass values measured by digital balance (Table 6). The mean apple mass difference between two methods was -8.15 g (95% confidence interval: -11.3 g and -5.02 g; P = 1.000). The standard deviation of the apple mass difference was 5.65 g. RMSE and MRPD were also used to check the discrepancies between the two methods. The amounts of RMSE and MRPD were 9.8 g and 5.80%, respectively. Thus, apple mass predicted by model No. 7 may be 9.8 g or 5.80% less than apple mass measured by a digital balance.

Table 5: Geometrical properties of the fifteen apples used in evaluating selected mass models

Sample No.	Geometrical properties of apple: GMD (cm)	CAE (cm^2)	V_{sp} (cm^3)	Apple mass (g): Measured by digital balance	Predicted by model No. 3	Predicted by model No. 7	Predicted by model No. 9
1	6.5	32.8	141.6	146	135.5	135.7	135.9
2	6.5	32.8	141.6	146	135.5	135.7	135.9
3	6.8	36.0	162.2	147	149.6	151.1	152.7
4	6.6	33.9	148.2	148	140.2	140.8	141.3
5	6.5	33.2	143.8	148	137.0	137.4	137.7
6	6.4	31.8	134.8	149	130.5	130.6	130.4
7	6.6	34.6	152.8	149	143.3	144.2	145.1
8	6.6	34.2	150.1	149	141.5	142.3	142.9
9	6.7	34.9	154.8	150	144.7	145.8	146.7
10	6.8	36.0	162.2	151	149.6	151.1	152.7
11	6.6	34.6	152.8	151	143.3	144.2	145.1
12	6.6	34.2	150.6	153	141.8	142.5	143.2
13	6.5	33.5	146.0	153	138.6	139.1	139.5
14	6.6	33.9	148.1	155	140.1	140.7	141.3
15	6.7	35.3	157.5	156	146.4	147.6	148.9

Table 6: Paired samples t-test analyses on comparing apple mass determination methods

Determination methods	Average difference (g)	Standard deviation of difference (g)	p-value	95% confidence intervals for the difference in means (g)
Model No. 3 vs. measuring	-8.89	5.28	1.0000	-11.8, -5.97
Model No. 7 vs. measuring	-8.15	5.65	1.0000	-11.3, -5.02
Model No. 9 vs. measuring	-7.45	6.09	0.9998	-10.8, -4.07

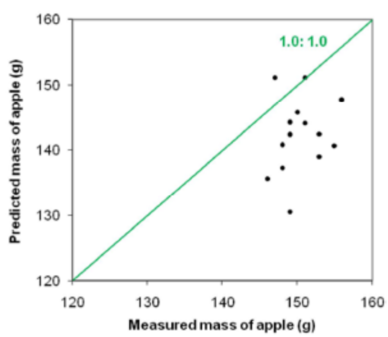

Fig. 3: Measured mass of apple and predicted mass of apple using model No. 7 with the line of equality (1.0: 1.0)

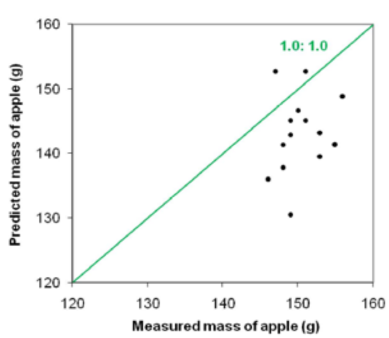

Fig. 4: Measured mass of apple and predicted mass of apple using model No. 9 with the line of equality (1.0: 1.0)

Comparison of Model No. 9 with Measuring Method: The apple mass values predicted by model No. 9 were compared with the apple mass values measured by digital balance and are shown in Table 5. A plot of the apple mass values predicted by model No. 9 and measured by digital balance with the line of equality (1.0: 1.0) is shown in Fig. 4. The paired samples t-test results indicated that the apple mass values predicted with model No. 9 were significantly less than the apple mass values measured by digital balance (Table 6). The mean apple mass difference between two methods was -7.45 g (95% confidence interval: -10.8 g and -4.07 g; P = 0.9998). The standard deviation of the apple mass difference was 6.09 g. RMSE and MRPD were also used to check the discrepancies between the two methods. The amounts of RMSE and MRPD were 9.5 g and 5.61%, respectively. Thus, apple mass predicted by model No. 9 may be 9.5 g or 5.61% less than apple mass measured by a digital balance.

CONCLUSIONS

In order to predict apple mass based on outer dimensions, the mass model based on GMD as M = - 168.5 + 47.01 GMD with R^2 = 0.77 was preferred. In addition, to predict apple mass based on projected areas, the mass model based on CAE as M = - 26.82 + 4.948 CAE with

$R^2 = 0.77$ was selected. Moreover, to predict apple mass based on estimated volume, the mass model based on V_{Sp} as $M = 20.68 + 0.814\ V_{Sp}$ with $R^2 = 0.76$ was chosen.7 However, the statistical results of study indicated that the liner regression models underpredict mass of apple and can not be appropriately used.

ACKNOWLEDGMENT

The authors are very grateful to the “East Tehran Branch, Islamic Azad University, Iran” for giving all type of support in publishing this study.

REFERENCES

1. Kheiralipour, K., A. Tabatabaeefar, H. Mobli, S. Rafiee, M. Sharifi, A. Jafari and A. Rajabipour, 2008. Some physical and hydrodynamics properties of two varieties of apple (*Malus domestica* Borkh L.). International Agrophysics, 22: 225-229.
2. FAO, 2007. FAOSTAT database. Available at: http://faostat.fao.org.
3. Gorji Chakespari, A., A. Rajabipour and H. Mobli, 2010. Mass modeling of two apple varieties by geometrical attributes. Australian Journal of Agricultural Engineering, 1(3): 112-118.
4. Rashidi, M. and K. Seyfi, 2007. Classification of fruit shape in cantaloupe using the analysis of geometrical attributes. World Applied Sciences Journal, 3: 735-740.
5. Rashidi, M. and M. Gholami, 2008. Classification of fruit shape in kiwifruit using the analysis of geometrical attributes. American-Eurasian Journal of Agricultural and Environmental Sciences, 3: 258-263.
6. Sadrnia, H., A. Rajabipour, A. Jafary, A. Javadi and Y. Mostofi, 2007. Classification and analysis of fruit shapes in long type watermelon using image processing. International Journal of Agriculture and Biology, 9: 68-70.
7. Tabatabaeefar, A. and A. Rajabipour, 2005. Modeling the mass of apples by geometrical attributes. Sci. Hort., 105: 373-382.
8. Wilhelm, L.R., D.A. Suter and G.H. Brusewitz, 2005. Physical Properties of Food Materials. Food and Process Engineering Technology. ASAE, St. Joseph, Michigan, USA.
9. Wen, Z. and Y. Tao, 1999. Building a rule-based machine-vision system for defect inspection on apple sorting and packing lines. Expert Systems with Application, 16: 307-713.
10. Kavdir, I. and D.E. Guyer, 2004. Comparison of artificial neural networks and statistical classifiers in apple sorting using textural features. Biosystems Engineering, 89: 331-344.
11. Kleynen, O., V. Leemans and M.F. Destain, 2003. Selection of the most effective wavelength bands for ‘Jonagold’ apple sorting. Postharvest Biology and Technology, 30: 221-232.
12. Polder, G., G.W.A.M. van der Heijden and I.T. Young, 2003. Tomato sorting using independent component analysis on spectral images. Real-Time Imaging, 9: 253-259.
13. Mohsenin, N.N., 1986. Physical Properties of Plant and Animal Materials. Gordon and Breach Science Publishers. New York. USA.
14. Khanali, M., M. Ghasemi Varnamkhasti, A. Tabatabaeefar and H. Mobli, 2007. Mass and volume modeling of tangerine (*Citrus reticulate*) fruit with some physical attributes. International Agrophysics, 21: 329-334.
15. Khoshnam, F., A. Tabatabaeefar, M. Ghasemi Varnamkhasti and A. Borghei, 2007. Mass modeling of pomegranate (*Punica granatum* L.) fruit with some physical characteristics. Sci. Hort., 114: 21-26.
16. Rashidi, M. and K. Seyfi, 2008. Modeling of kiwifruit mass based on outer dimensions and projected areas. American-Eurasian Journal of Agricultural and Environmental Sciences, 3: 14-17.
17. Taheri-Garavand, A. and A. Nassiri, 2010. Study on some morphological and physical characteristics of sweet lemon used in mass models. International Journal of Environmental Sciences, 1: 580-590.
18. Rashidi, M. and M. Gholami, 2010. Prediction of soil sinkage by multiple loadings using the finite element method. International Journal of Agriculture and Biology, 12: 911-915.