

Estimation of Total and Kernel Mass in Walnuts Using X-ray Attenuation Mapping Method

A. Taheri^{1*}, M. Askari¹

1- Radiation Applications Research School, Nuclear Science and Technology Research Institute, Tehran, Iran

(*- Corresponding Author Email: at1361@aut.ac.ir)

<https://doi.org/10.22067/jam.2026.99948.1504>

Abstract

In this study, we propose a nondestructive method for estimating kernel mass in walnuts using X-ray inspection. X-ray radiographs of 350 walnuts were used to develop four attenuation and intensity-based features derived from segmented kernel regions. Three regression approaches namely linear, second-order polynomial, and exponential were fitted to relate these features to total walnut mass and kernel mass, using an 80/20 fit–test split ($n = 280/70$). The second-order polynomial (quadratic) model consistently outperformed the linear and exponential alternatives for both response variables. For total mass, the quadratic model achieved a test-set $R^2 = 0.87$, RMSE of 0.75 g, MAE of 0.59 g, and a relative error of 6.15%. For kernel mass, the same model type achieved $R^2 = 0.72$, RMSE of 0.73 g, MAE of 0.58 g, and a relative error of 17.13%. These results indicate that X-ray attenuation-based features, combined with a nonlinear regression framework, can provide a reliable estimate of total walnut mass and a moderately accurate estimate of kernel mass. This approach offers a nondestructive, image-based alternative to manual cracking and weighing, with potential applications in automated sorting and quality-control systems for walnut processing.

Keywords: Attenuation map, Estimation function, Kernel mass, Walnut, X-ray

Introduction

Walnuts are among the most valuable nut crops worldwide due to their high nutritional value and economic importance. The commercial value of walnuts is strongly affected by the mass of the edible kernel obtained after shell removal. Therefore, the kernel mass is considered an important quality parameter for grading, pricing, processing optimisation, and consumer acceptance. Accurate estimation of kernel mass before shelling can provide valuable information for producers, processors, and exporters by enabling improved quality assessment and more efficient sorting strategies (Laxman Kumar *et al.*, 2025; Zhang, Ning, Pu, & Ji, 2024; Zhang *et al.*, 2025).

Conventional evaluation of walnut kernel mass is mainly performed through destructive procedures, where walnuts are manually cracked, and the separated kernels are weighed. Although this approach provides direct and accurate measurements, it is labour-intensive, time-consuming, and unsuitable for evaluating

large numbers of individual walnuts in industrial environments. In addition, destructive sampling methods generally provide information at the batch level and may not adequately represent the variability among individual walnuts, resulting in uncertainty during quality grading and processing decisions. These limitations have encouraged the development of nondestructive techniques capable of rapidly evaluating internal walnut characteristics without damaging the product (Amna *et al.*, 2025; Kumar, Rahul, Arora, & Gupta, 2026).

Several nondestructive imaging techniques have been investigated for agricultural quality evaluation, including visible imaging, hyperspectral imaging, magnetic resonance imaging, and X-ray imaging (Ahmed, Yasmin, Lee, Mo, & Cho 2017; Li, Jia, & Fan, 2025; Wu & Sun, 2013). Among these methods, X-ray radiography has received considerable attention because of its ability to penetrate opaque materials and provide information about internal structures based on differences in X-ray

attenuation. In agricultural products, X-ray imaging has been successfully applied for detecting internal defects, insect damage, voids, foreign materials, and incomplete kernel development (Kotwaliwale *et al.*, 2014). In particular, the contrast between shell and kernel regions in nut radiographs provides an opportunity for extracting structural information that cannot be obtained through external inspection.

Despite these advantages, most previous X-ray-based studies on nuts have focused primarily on qualitative evaluation and classification tasks, such as identifying defective products, estimating filling levels, or separating acceptable and unacceptable samples. For example, several studies have applied X-ray imaging combined with image processing or machine learning methods for defect detection and quality classification in nuts and seeds (Fan & Zhou, 2025; Goa, Zhang, Sun, & Ren, 2022; Ma *et al.*, 2025; Zhang *et al.*, 2024). However, quantitative estimation of economically important parameters, such as total mass or kernel mass, remains relatively underexplored. Such quantitative information is particularly valuable for automated sorting systems, where decisions must be made for individual walnuts rather than average properties of a batch.

The conversion of X-ray radiographs into quantitative estimates of kernel mass involves several technical challenges. The measured X-ray intensity is influenced not only by material composition and density but also by object thickness, geometrical configuration, radiation energy, and detector response. Furthermore, walnuts have complex and irregular shapes, and the shell introduces additional attenuation variations that affect the relationship between image intensity and the amount of internal kernel material. Therefore, direct interpretation of radiographic intensity is not sufficient for reliable kernel mass estimation, and appropriate approaches are required to identify kernel regions, reduce the influence of shell attenuation, and extract image features correlated with actual kernel mass (Kastner & Heinzl, 2015; Mendoza, Dejmek, & Aguilera,

2006).

Recent developments in digital image processing and artificial intelligence have provided new opportunities for extracting meaningful information from complex biological images. Image segmentation, morphological analysis, attenuation-based mapping, and feature extraction techniques have been successfully employed for identifying internal structures and improving nondestructive evaluation accuracy in agricultural products (Liu, Zhang, & Gui, 2021; Lotufo, Rittner, Audigier, Machado, & Saúde, 2011; Tian, Wang, Liu, Qiao, & Li, 2020; Zhang, Ji, Shao, Pu, & Zhang, 2023). Moreover, deep learning approaches have demonstrated strong capabilities in automatically learning complex image features without requiring extensive manual feature engineering. These advances suggest that combining X-ray imaging with advanced image analysis methods may provide an effective solution for predicting internal quality parameters of individual walnuts at an industrial scale (Du & Sun, 2006).

However, there is a lack of reliable nondestructive methods for directly estimating walnut kernel mass from planar X-ray radiographs. Previous studies have rarely addressed the quantitative relationship between radiographic characteristics and measured kernel mass while considering the effects of walnut geometry and shell attenuation. Developing such a method requires accurate identification of walnut boundaries, extraction of relevant internal features, and establishment of predictive relationships between X-ray measurements and experimentally determined kernel mass. Therefore, the objective of this study is to develop a nondestructive method for estimating walnut kernel mass using planar X-ray radiographs. The proposed approach combines automatic walnut segmentation, identification of kernel-related regions, and extraction of quantitative image features associated with X-ray attenuation. These features are subsequently used to establish a relationship between radiographic information and experimentally measured kernel mass,

providing a potential framework for rapid quality evaluation and automated sorting of walnuts without shell removal.

Materials and Methods

X-ray Imaging

A total of 350 walnut samples were used in this study. All samples were intact, with no visible cracks or defects on the shells, were commercially obtained from local markets, and were identified as Iranian walnuts based on supplier information. However, detailed cultivar information was not available. The

main objective of this work was to evaluate the feasibility of estimating kernel mass ratio from a single-view X-ray radiograph using attenuation-based features. Further studies involving larger datasets with multiple cultivars, geographical origins, and broader physical variations are required for developing a generalised industrial sorting system.

Table 1 reports the physical characteristics of the walnut samples used in this study, including the measured ranges and statistical distributions of walnut dimensions and mass-related parameters.

Table 1- Summary of the physical characteristics of the walnut samples, including walnut dimensions, total mass, and kernel mass

Parameter	Min	Max	Mean $\pm$ SD
Walnut length (mm)	29.28	43.40	34.78 $\pm$ 2.76
Walnut width (mm)	24.47	38.40	32.28 $\pm$ 2.38
Total mass (g)	4.20	17.23	9.83 $\pm$ 2.21
Kernel mass (g)	0.21	10.81	4.78 $\pm$ 1.47

The walnuts were scanned using an X-ray imaging system that included an industrial X-ray generator (TRACKIS-GE300) and a flat panel detector (Whale 3025 FPI-S). The active area of the detector was equal to 430×430 mm², with a pixel pitch of 140 μ m, which produced 16-bit images with a resolution of 3072×3072 pixels. The acquired images were then resized to 400×400 pixels for easier processing with a personal PC. The source-to-detector distance was fixed at 1.5 m. Walnuts were placed directly on the detector's surface using a thin A3 paper and imaged. Prior to image acquisition, standard flat-field and dark-field corrections were applied for detector calibration. The X-ray setup used a 50 kVp and 1 mA beam to ensure sufficient penetration for visualising the internal structure of the walnuts. The exposure time for each scan was about 1 s. The short-duration X-ray exposure used for image acquisition was performed only for inspection purposes and does not alter the food

product through radioactivation or irradiation effects. Also, the X-ray exposure parameters were optimised for walnut imaging based on preliminary tests to balance image quality and prevent overexposure.

For each walnut, a single X-ray radiograph was captured from a lateral view. The walnut was positioned directly on the flat panel to ensure uniform exposure. The images were saved in the TIFF format with 16-bit depth, providing a wide dynamic range of intensity values for detailed analysis. The images were then processed using MATLAB (Version 2024a) for further analysis.

For each walnut, the kernel mass was determined after shelling, and the shell mass was recorded separately. These mass measurements served as the ground truth for evaluating the accuracy of the kernel mass estimation method. Figure 1 shows a schematic view of the X-ray imaging method.

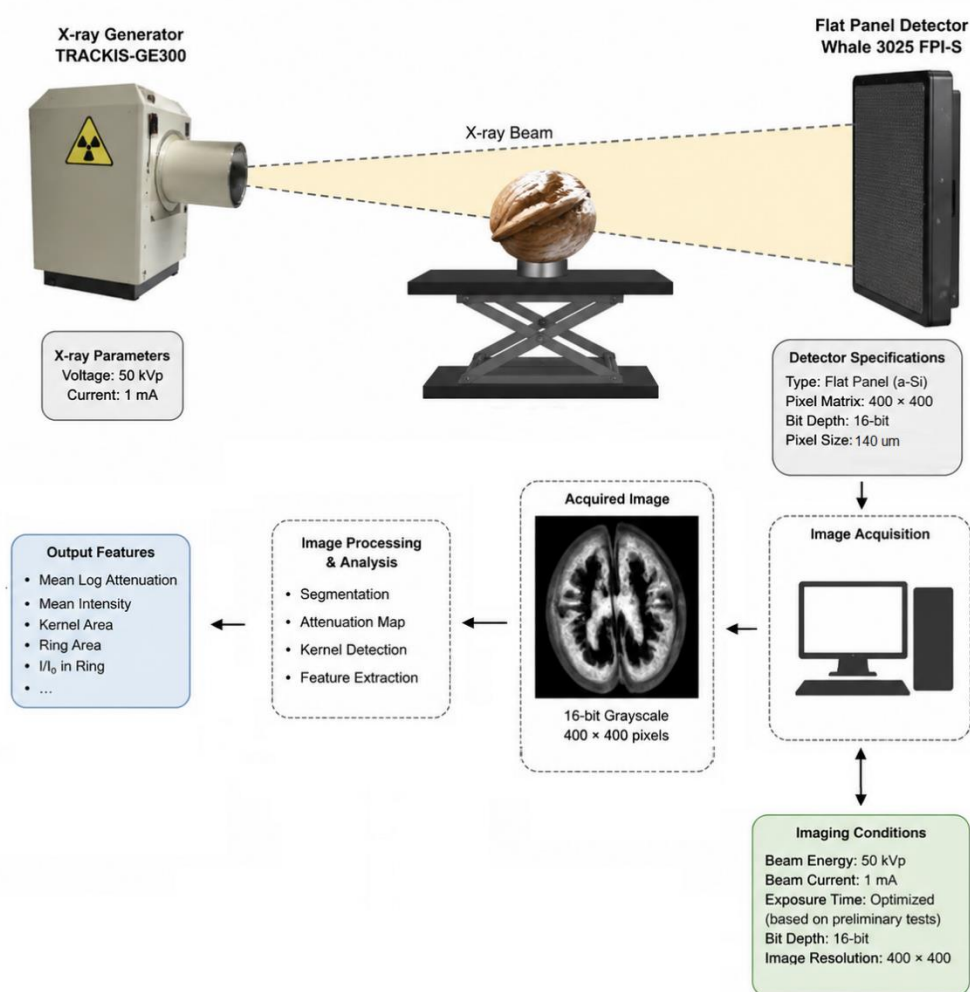


Fig. 1. Schematic representation of the X-ray imaging method used for walnut scanning. The setup includes an industrial X-ray generator (TRACKIS-GE300) and a flat panel detector (Whale 3025 FPI-S). The walnut is scanned with X-ray parameters of 50 kVp and 1 mA to capture high-resolution 16-bit grayscale images at a resolution of 400 × 400 pixels. The diagram also highlights the image acquisition and processing stages, including feature extraction for further analysis (designed with the help of ChatGPT)

Processing of X-ray Images

The acquired X-ray radiographs were first converted from 16-bit to a floating-point format, preserving the full intensity range. The images were then normalised by mapping the intensity values to the [0, 1] range. This normalisation was followed by contrast enhancement using adaptive histogram equalisation, which improved the contrast of walnut features, particularly the boundary between the shell and kernel. The walnut boundary was initially identified using Otsu's thresholding method (Sezgin & Sankur, 2004) applied to the normalised image. This technique divides the image into two classes: background and object. A binary mask was then generated

to distinguish the walnut from the background. If the walnut appeared darker than the background, the binary mask was inverted to correct the segmentation. To remove noise and small irrelevant objects, the binary mask was filled to close any gaps in the walnut region and subjected to a morphological opening process with a minimum object size threshold of 1000 pixels. This ensured that only the largest connected component, representing a single walnut, was retained for further analysis. The largest connected component was then used as the final walnut mask. The kernel region inside the walnut was identified using a morphological erosion technique. A disk-shaped structuring element with a radius of 15 pixels was used to

erode the walnut mask and create a kernel region mask. The morphological erosion was used to generate an internal region of interest (ROI) rather than a precise kernel segmentation. The selected erosion radius was intended to reduce the influence of shell attenuation while retaining the central region containing most of the kernel material. This method effectively excluded the shell from the analysis, leaving the region corresponding to the kernel. However, since walnut shell thickness and kernel morphology may vary among samples, this approach represents an approximation and should be further validated using manually annotated kernel regions or adaptive segmentation methods in future studies.

The mean intensity of the background region (i.e., outside the walnut) was computed as the reference intensity, I_0 , which represents the X-ray attenuation of the air (background). This value was used to calculate the attenuation map ($\ln(I_0/I)$) for the walnut region. The intensity values within the kernel region were then normalised by computing the ratio I/I_0 , where I is the measured X-ray intensity for each pixel inside the kernel region, and I_0 is the mean intensity of the background region. Kernel edges were detected using the gradient magnitude of the attenuation map. The image gradients in both horizontal and vertical directions were computed, and the magnitude of the gradient (G_{mag}) was obtained by combining the horizontal (G_x) and vertical (G_y) gradients based on Eq. (1):

$$G_{mag} = \sqrt{G_x^2 + G_y^2} \quad (1)$$

The resulting gradient magnitude image was then smoothed using a Gaussian filter with a sigma value of 2 to reduce noise. A threshold was applied to the smoothed gradient image, and regions with strong gradients, indicating the kernel boundary, were selected as potential kernel edges. These edges were refined by ensuring they were contained within the kernel region mask and removing small isolated objects. The following features were extracted for each walnut using the mentioned procedure: (1) A_K which is the area of the kernel (internal

ROI) was calculated based on the pixel count and the pixel size; (2) H and W which are the major and minor axis lengths, respectively, were derived from the walnut's bounding box and converted to physical units; (3) I_{Mean} is the mean X-ray intensity of the kernel region; (4) $LogAtt_{Mean}$ is defined as $\ln(I_0/I)$ the mean attenuation inside the kernel region; and (5) I_{RMean} is the mean intensity ratio I/I_0 in the narrow ring region. More details about the processing procedure are shown for an example radiograph in Figure 2.

This dataset was used to develop statistical models to estimate total mass (M_{tot}) and kernel mass (M_{ker}) based on X-ray features. Variations of different proportions of these features were tested as a function of walnut and kernel masses. These proportions included $x_1 = I_{Mean} \times A_K$, $x_2 = (\exp(LogAtt_{Mean}))^{-1} \times A_K$, $x_3 = I_{Mean}/A_K$, and $x_4 = LogAtt_{Mean} \times A_K$.

It should be noted that the attenuation values calculated in this study represent relative attenuation features rather than absolute linear attenuation coefficients. The measured X-ray intensity depends on several factors, including material composition, projected thickness, beam spectrum, detector response, scattering effects, and imaging geometry. Since the objective of this study was to establish empirical relationships between X-ray-derived features and walnut kernel mass, absolute attenuation calibration was not performed. Therefore, the extracted attenuation parameters were considered image-based predictors rather than direct measurements of kernel density.

To develop predictive models for estimating the two response variables (M_{tot} and M_{ker}) from the four explanatory variables (x_1 – x_4), three regression functions with different levels of complexity were investigated: multiple linear regression, second-order polynomial regression, and exponential regression. In all cases, the four x -variables were considered as predictors, whereas M_{tot} and M_{ker} were treated as independent response variables. Thus, the modelling framework was explicitly designed to predict the M -variables from the measured x -variables.

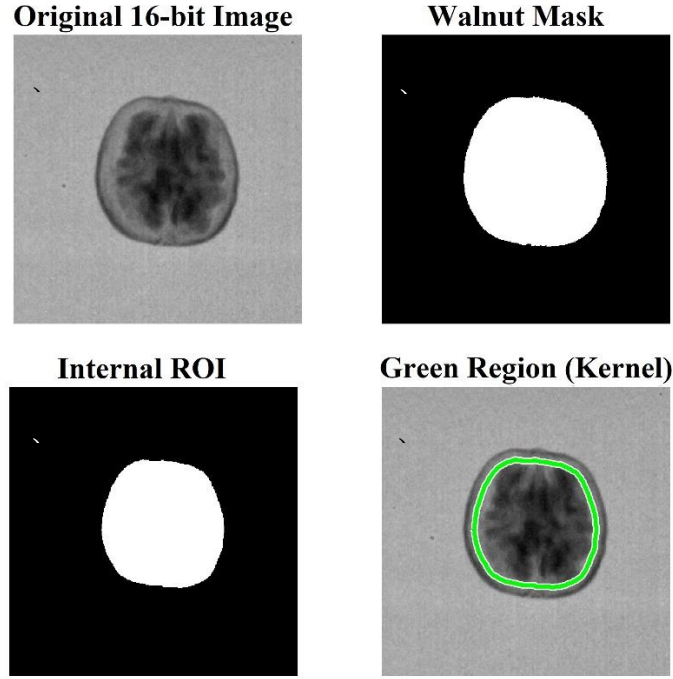


Fig. 2. Illustration of the walnut image processing workflow: (a) the original 16-bit X-ray radiograph of a walnut, (b) the segmented walnut mask, (c) the internal ROI obtained by eroding the walnut mask to remove the shell region, and (d) the green region corresponding to the kernel detected using the radial gradient method. This workflow enables the extraction of geometric and attenuation features for kernel mass estimation

The available observations were randomly divided into a fitting set (80%, $n = 280$) and an independent test set (20%, $n = 70$). The fitting data were used for parameter estimation and model selection, whereas the test data were kept completely independent and were used only for the final evaluation of predictive performance. To reduce the possibility of selecting a model based on a single favourable data split, validation was performed within the fitting dataset. The model showing the most consistent predictive performance during cross-validation was selected as the final model for each response variable.

The linear model was expressed as:

$$y = \beta_0 + \sum_{i=1}^4 \beta_i x_i \quad (2)$$

where y represents the Mass values. Also, the second-order polynomial model included both quadratic and pairwise interaction terms:

$$y = \beta_0 + \sum_{i=1}^4 \beta_i x_i + \sum_{i=1}^4 \beta_{ii} x_i^2 + \sum_{i < j} \beta_{ij} x_i x_j \quad (3)$$

The exponential model was formulated as:

$$y = \exp(\beta_0 + \sum_{i=1}^4 \beta_i x_i) \quad (4)$$

Model performance was evaluated using the coefficient of determination (R^2), root mean square error (RMSE), mean absolute error (MAE), bias, and relative error. These complementary metrics were used because R^2 alone does not provide sufficient information about the magnitude or direction of prediction errors. RMSE was used to quantify the overall prediction error while giving greater weight to large deviations, whereas MAE represented the average absolute deviation between measured and predicted values. Bias was calculated to identify systematic over- or under-prediction, and relative error was used to express the prediction error relative to the measured response.

For the independent test set, the metrics were calculated as:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (5)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (6)$$

$$Bias = \frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i) \quad (7)$$

and

$$RE(\%) = \frac{1}{n} \sum_{i=1}^n \left| \frac{\hat{y}_i - y_i}{y_i} \right| \times 100 \quad (8)$$

where y_i and $\hat{y}_i$ denote the measured and predicted values, respectively. The final model was then refitted using the complete fitting dataset and evaluated on the independent test dataset.

Results

In the first step, the X-ray radiographs of all 350 walnuts were captured and stored for the next analyses. Figure 3 shows a part of the walnuts alongside their acquired X-ray radiographs.

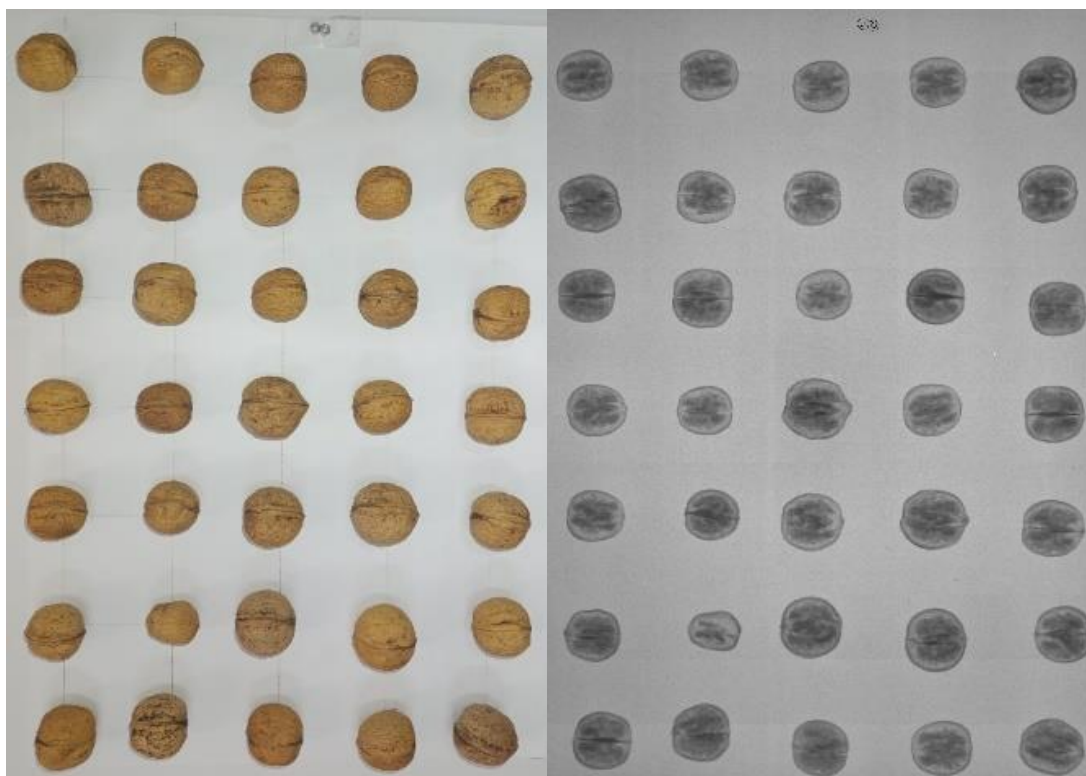


Fig. 3. A part of the selected walnuts and corresponding X-ray radiographs

The comparison of the three regression approaches indicated that the second-order polynomial model provided the most favourable predictive performance for both response variables. The obtained metrics are reported in Table 2.

For M_{tot} , the quadratic model achieved an R^2 of 0.8745 on the independent test set, with an RMSE of 0.7460, MAE of 0.5859, bias of -0.0784, and a relative error of 6.15%. The corresponding validation R^2 obtained using the fitting data was 0.8311, indicating that the

superior performance of the quadratic model was not limited to the particular Fit-test split.

Also, for M_{ker} , the quadratic model also outperformed the linear and exponential alternatives, although its predictive performance was lower than that obtained for M_{tot} . The independent test set resulted in an R^2 of 0.7225, RMSE of 0.7328, MAE of 0.5782, bias of -0.0418, and relative error of 17.13%. The validation R^2 was 0.6591. The relatively small negative bias for both responses indicate a slight tendency of the selected model to underestimate the measured values on average.

Table 2- Predictive performance of linear, second-order polynomial, and exponential regression models for estimating M_{tot} and M_{ker} from the four explanatory variables (x_1 – x_4) using the fitting and independent test datasets.

Target	Model	Dataset	N	R ²	RMSE	MAE	Bias	Relative Error (%)
M_{tot}	Linear	Fit	280	0.8356	0.9054	0.7252	0.0000	7.6658
M_{tot}	Linear	Test	70	0.8621	0.7819	0.6069	-0.0784	6.4496
M_{tot}	Quadratic	Fit	280	0.8608	0.8331	0.6730	0.0000	7.1558
M_{tot}	Quadratic	Test	70	0.8745	0.7460	0.5859	-0.0784	6.1533
M_{tot}	Exponential	Fit	280	0.8307	0.9188	0.7332	-0.0409	7.7024
M_{tot}	Exponential	Test	70	0.8606	0.7863	0.6128	-0.1023	6.4590
M_{ker}	Linear	Fit	280	0.6480	0.8797	0.6538	0.0000	22.0515
M_{ker}	Linear	Test	70	0.6816	0.7849	0.5702	-0.0523	20.5043
M_{ker}	Quadratic	Fit	280	0.7298	0.7706	0.5991	0.0000	16.3705
M_{ker}	Quadratic	Test	70	0.7225	0.7328	0.5782	-0.0418	17.1342
M_{ker}	Exponential	Fit	280	0.5575	0.9862	0.7398	-0.0904	21.7930
M_{ker}	Exponential	Test	70	0.5954	0.8848	0.6419	-0.0759	20.3664

Overall, the results demonstrate that incorporating nonlinear and interaction effects through the second-order polynomial formulation substantially improved the predictive capability compared with a purely linear relationship. However, the difference in performance between M_{tot} and M_{ker} should be considered when interpreting the predictive reliability of the proposed approach. In particular, the relative error of 17.13% for M_{ker} indicates that its prediction remains associated with a considerably higher uncertainty than that of M_{tot} . Therefore, the quadratic model can be considered a relatively reliable predictor for M_{tot} , whereas predictions of M_{ker} should be interpreted with greater caution.

The best-fitted regression equations obtained from the fitting dataset are presented in Table 3. For each response variable, the equation corresponding to the selected predictor combination and functional form is reported together with the estimated regression coefficients. The equations provide the explicit mathematical relationship between the predictor variables (x_1 , x_2 , x_3 , and x_4) and the predicted response variables (M_{tot} and M_{ker}).

The final models for M_{tot} and M_{ker} were both second-order polynomial models based on x_1 , x_2 , x_3 , and x_4 . The polynomial equations include the linear terms, squared terms, and pairwise interaction terms associated with the selected predictors. The estimated coefficients were obtained using only the fitting dataset, thereby

preventing the independent test observations from influencing the fitted equations.

These equations can be directly used to calculate the predicted values of M_{tot} and M_{ker} from new measurements of the corresponding x -variables within the range of the experimental data. The fitted equations therefore provide not only a quantitative description of the observed relationships but also a practical predictive model for estimating the response variables.

Figure 4 compares the predictive performance of the three regression functions (linear, second-order polynomial, and exponential) for estimating the response variable M_{tot} using the independent test dataset. Each point represents one test sample, with the measured and predicted values plotted on the horizontal axis and the test sample number on the vertical axis.

Among the evaluated models, the second-order polynomial regression produced predictions indicating the highest prediction accuracy and the smallest systematic deviation. The linear regression model showed a wider dispersion, particularly for samples with relatively high values of M_{tot} , suggesting that a purely linear relationship could not fully describe the nonlinear behaviour of the data. The exponential regression model exhibited the largest deviations, indicating inferior predictive capability compared with the polynomial model.

These observations are consistent with the

statistical evaluation, where the quadratic model achieved the highest coefficient of determination and the lowest prediction errors on the independent test dataset. The improved

performance demonstrates that incorporating quadratic and interaction terms provides a more accurate representation of the relationship between the predictor variables and M_{tot} .

Table 3- the best-fitted regression equations and estimated coefficients of the selected predictive models for M_{tot} and M_{ker} . The models were developed using the fitting dataset, with x_1 – x_4 as candidate predictor variables. All coefficients were rounded to two decimal places

Target	Model	Equation
M_{tot}	Quadratic	$105.81 - 4387.56x_1 + 123739.26x_2 - 3.61x_3 + 0.05x_4$ $+21455.25x_1^2 + 285013.53x_1x_2 + 1e - 4x_1x_3$ $+6.27x_1x_4 + 3749.67x_2^2 - 1251.24x_2x_3$ $-219.86x_2x_4 + 0.03x_3^2 + 4e - 3x_3x_4$ $-3e - 4x_4^2$
	Quadratic	$273.81 - 4921.96x_1 + 258766.42x_2 - 7.13x_3 - 0.08x_4$ $-2797.53x_1^2 - 5782887x_1x_2 - 3e - 3x_1x_3$ $+9.78x_1x_4 - 76677.14x_2^2 - 2115.14x_2x_3$ $-141.78x_2x_4 + 0.05x_3^2 + 4e - 3x_3x_4$ $-3e - 4x_4^2$

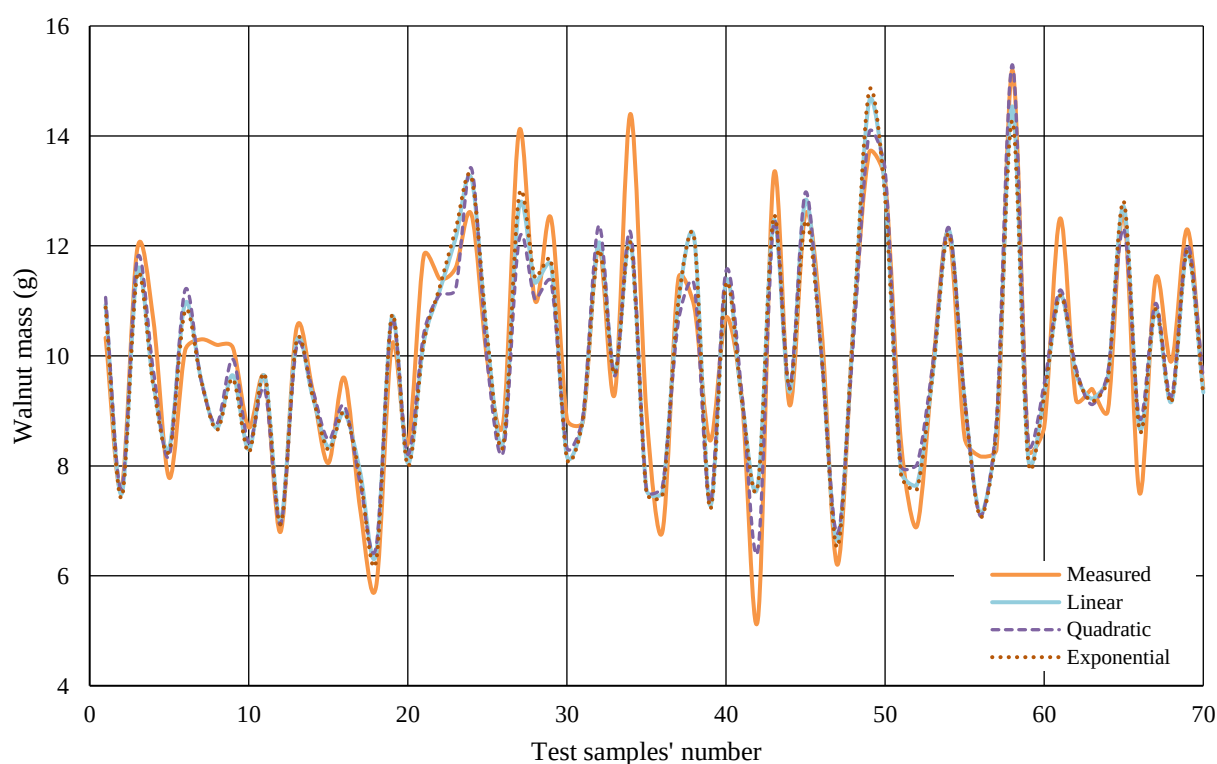


Fig. 4. Comparison of measured and predicted M_{tot} values obtained using linear, second-order polynomial, and exponential regression models on the independent test dataset. The second-order polynomial model produced the closest agreement with the experimental measurements and the smallest prediction error among the evaluated regression functions

Figure 5 compares the predictive performance of the linear, second-order polynomial, and exponential regression models

for estimating the response variable M_{ker} using the independent test dataset. Similar to Figure 4, each point corresponds to one test sample.

The second-order polynomial regression again provided the best overall agreement with the measured values compared to those obtained from the linear and exponential models. However, the dispersion of the data was larger than that observed for M_{tot} , indicating a lower prediction accuracy for M_{ker} .

The linear regression model underestimated several high-value samples and showed greater variability across the prediction range, whereas the exponential model produced the largest prediction errors and the weakest agreement

with the measured observations. The larger scatter observed for all models suggests that the relationship between the predictor variables and M_{ker} is more complex and contains greater variability than the relationship for M_{tot} .

Overall, the quadratic model provided the most reliable predictions for M_{ker} , although its prediction uncertainty remained higher than that obtained for M_{tot} , as reflected by the lower test-set R^2 value and the higher relative prediction error.

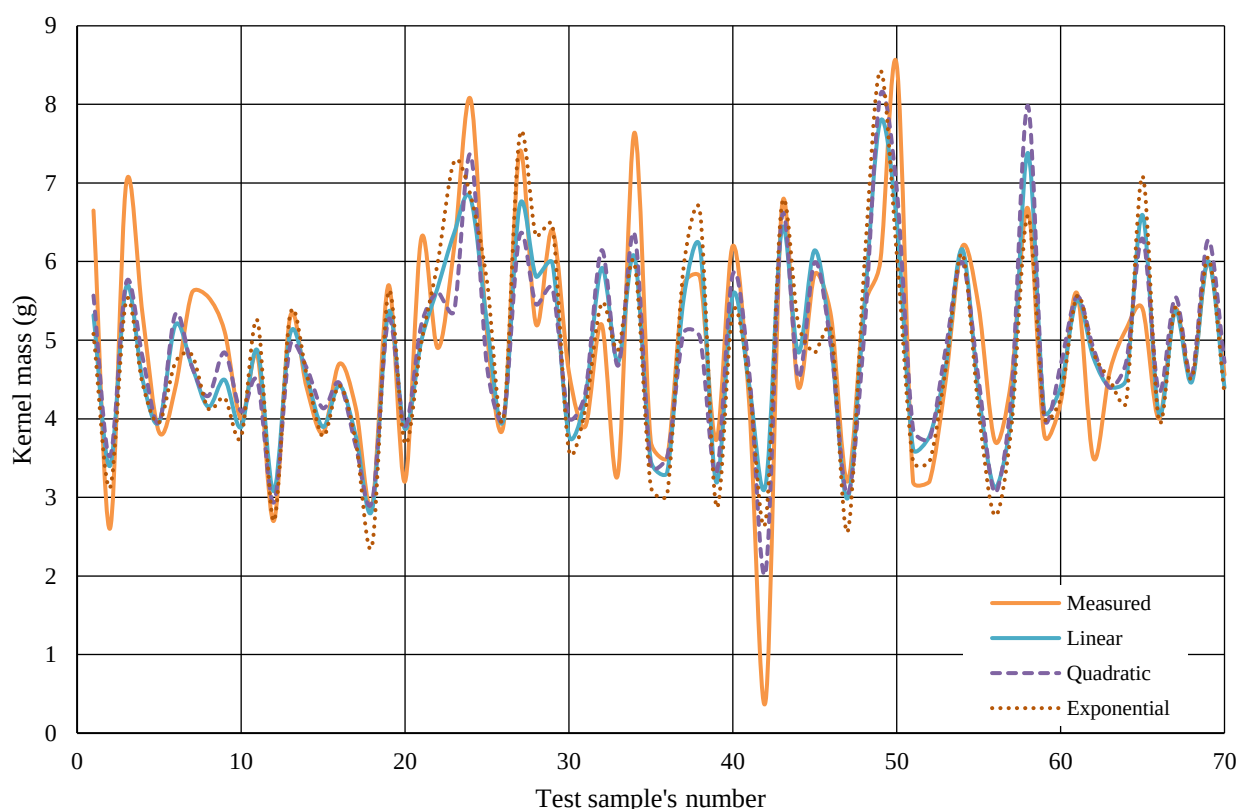


Fig. 5. Comparison of measured and predicted M_{ker} values obtained using linear, second-order polynomial, and exponential regression models on the independent test dataset. The second-order polynomial model provided the best overall predictive performance, although a larger dispersion indicates greater prediction uncertainty than that observed for M_{tot}

Discussion

The present study investigated the capability of empirical regression models to predict walnut and kernel masses (M_{tot} and M_{ker}) from four measured predictor variables (x_1-x_4) obtained from X-ray radiographs of the walnuts. Unlike approaches based on a predefined regression equation, both the

functional form and the combination of predictor variables were systematically evaluated. Linear, second-order polynomial, and exponential models were compared using all possible combinations of the four predictors, while an independent test dataset was used to assess model robustness and predictive performance.

The results showed that the second-order polynomial model provided the most consistent performance for both response variables. For the complete four-variable configuration, the cross-validation R^2 increased from 0.8175 for the linear model to 0.8311 for the quadratic model for y_1 , while a more pronounced improvement was observed for M_{ker} , from 0.5942 to 0.6591. This improvement indicates that the relationships between the predictors and the responses are not adequately described by simple linear relationships. The inclusion of quadratic terms (x_i^2) and pairwise interaction terms ($x_i x_j$) allowed the models to account for nonlinear effects and interactions among the predictors. In contrast, the exponential model did not provide a consistent improvement and showed substantially poorer performance for M_{ker} . Thus, the results suggest that a polynomial representation was more appropriate for the present dataset than either a purely linear or exponential formulation.

For M_{tot} , the quadratic model using x_1 to x_4 provided the highest validation performance R^2 of 0.8311 and a test R^2 of 0.8745. For M_{ker} , the best configuration was the quadratic model, which achieved cross-validation R^2 of 0.6591 and a test R^2 of 0.7225, with an RMSE of 0.7328 and a relative error of approximately 17.1%. The lower predictive performance compared with M_{tot} suggests that the variability in M_{ker} is more difficult to explain using the available predictors. Although the selected model captured a substantial proportion of the variation, the relatively high relative error indicates that individual predictions may still contain considerable uncertainty. Therefore, the model should be regarded as a useful empirical predictor rather than a highly precise estimator of M_{ker} .

The close agreement between validation and independent test performance provides evidence that the selected models were not simply fitting the observations. In addition, the relatively small bias values indicate that systematic overprediction or underprediction was limited, while the dispersion of the measured-versus-predicted observations provides a visual confirmation of the

differences in predictive performance between M_{tot} and M_{ker} . These findings emphasise the importance of using an independent test dataset and multiple error metrics rather than relying solely on the fitting R^2 .

Overall, the results demonstrate that systematic optimisation of both predictor selection and functional form can improve the development of interpretable predictive equations. The second-order polynomial models provided the best overall performance for both M_{tot} and M_{ker} . However, the developed equations are empirical models and should be applied primarily within the experimental range used for model development. Further validation using independent datasets collected under different experimental conditions would be necessary to confirm their generalisability and practical applicability. In addition, the extracted attenuation parameters are influenced by the walnut projection geometry. Since only a single lateral X-ray image was acquired for each sample, variations in walnut orientation, projected thickness, kernel distribution, shell thickness, and internal voids may affect the measured intensity values. Therefore, the extracted features should be considered as orientation-dependent image descriptors rather than intrinsic material properties. Future studies should investigate multi-view X-ray imaging, orientation correction methods, or larger datasets containing diverse walnut poses to improve model generalisation for industrial sorting applications.

Conclusion

This study evaluated the feasibility of estimating walnut total mass and kernel mass from single-view X-ray radiographs using four attenuation- and intensity-derived features together with three regression frameworks of increasing complexity. Among the tested approaches, the second-order polynomial (quadratic) model consistently provided the most accurate and stable predictions for both response variables, outperforming the linear and exponential models across fitting, cross-validation, and independent test evaluations.

For total walnut mass, the quadratic model

achieved a test-set R^2 of ~ 0.87 with a relative error of approximately 6%, indicating a strong and reliable predictive relationship. For kernel mass, the same modelling approach achieved a lower but still meaningful test-set R^2 of 0.72, with a relative error of approximately 17%, reflecting the greater difficulty of isolating kernel-specific attenuation signals through the shell. The slight negative bias observed for both targets indicate a consistent tendency toward underestimation rather than random error, a pattern that could be addressed through additional calibration.

These findings support X-ray attenuation mapping as a viable nondestructive alternative to manual, destructive kernel-mass assessment, and highlight the value of comparing multiple functional forms rather than assuming a fixed linear relationship. Future work should focus on multi-view imaging, larger and more cultivar-diverse datasets, and refined kernel-boundary segmentation to reduce the prediction uncertainty currently observed for kernel mass, with the goal of enabling integration into automated, high-throughput sorting lines for the walnut processing industry.

Declaration of Generative AI and AI-Assisted Technologies

During the preparation of this manuscript, the authors used artificial intelligence (AI)-based tools to assist with language editing, writing refinement, and data analysis. AI tools were used to improve the clarity and readability of the manuscript and to assist in the analysis and interpretation of the results. All AI-assisted outputs were critically reviewed and verified by the authors, who remain fully responsible for the accuracy, originality, scientific validity, and final content of the manuscript. No AI tool was used as a substitute for the authors' scientific judgment or responsibility.

Author Contributions

Ali Taheri: Conceptualisation, Funding acquisition, Investigation, Project administration, Resources, Supervision, Writing – original draft, Writing – review & editing.

Mojtaba Askari: Conceptualisation, Methodology, Software, Supervision, Validation.

References

1. Ahmed, M. R., Yasmin, J., Lee, W.-H., Mo, C., & Cho, B.-K. (2017). Imaging Technologies for Nondestructive Measurement of Internal Properties of Agricultural Products: A Review. *Journal of Biosystems Engineering*, 42(3), 199–216. <https://doi.org/10.5307/JBE.2017.42.3.199>
2. Amna, M. W. A., Li, G., Zuhair Akram, M., Faheem, M., Omar, M. M. & Hassan, M. G., (2025). Machine vision-based automatic fruit quality detection and grading. *Frontiers of Agricultural Science and Engineering*, 12(2), 274–287. <https://doi.org/10.15302/J-FASE-2023532>
3. Du, C. J., & Sun, D. W. (2006). Learning techniques used in computer vision for food quality evaluation: A review. *Journal of Food Engineering*, 72(1), 39–55. <https://doi.org/10.1016/j.jfoodeng.2004.11.017>
4. Fan, X., & Zhou, J. (2025). Nondestructive detection and quality grading system of walnut using X-ray imaging and lightweight WKNet. *Foods*, 14(13), 2346. <https://doi.org/10.3390/foods14132346>
5. Goa, T. Zhang, S., Sun, H., & Ren, R. 2022. Mass detection of walnut based on X-ray imaging technology. *Journal of Food Process Engineering*, 45(8), e14034. <https://doi.org/10.1111/jfpe.14034>
6. Kastner, J., & Heinzl, C. (2015). X-ray Computed Tomography for Non-destructive Testing and Materials Characterization. In: Liu, Z., Ukida, H., Ramuhalli, P., Niel, K. (eds) *Integrated Imaging and Vision Techniques for Industrial Inspection. Advances in Computer Vision and Pattern Recognition*. Springer, London. https://doi.org/10.1007/978-1-4471-6741-9_8
7. Kotwaliwale, N., Singh, K., Kalne, A., Jha, S. N., Seth, N., & Kar, A. (2014). X-ray imaging methods for internal quality evaluation of agricultural produce. *Journal of Food Science and*

- Technology, 51(1), 1–15. <https://doi.org/10.1007/s13197-011-0485-y>
8. Kumar, Y., Rahul, K., Arora, N., & Gupta, R. P. (2026). Computer Vision Technologies for Food Quality and Safety Assurance: Evaluating Trends and Analytical Results. *Food Analytical Methods*, 19, 106. <https://doi.org/10.1007/s12161-026-03019-6>
 9. Laxman Kumar, A., Venugopal, S., Vittal, H., Maruthi, B., Ashok Kumar, D., & Priyanka, B., (2025). Production constraints and improvement strategies of temperate nut crops. In A. W. Wani, R. Rajan, & F. Ahmad (Eds.), *Temperate nut crops: Cultivation, science, and innovation* (pp. 257–286). Apple Academic Press. <https://doi.org/10.1201/9781779643513-8>
 10. Li, L. Jia, X., & Fan, K. (2025). Recent advance in nondestructive imaging technology for detecting quality of fruits and vegetables: a review. *Critical Reviews in Food Science and Nutrition*, 65(26), 5181–5199. <https://doi.org/10.1080/10408398.2024.2404639>
 11. Liu, Y., Zhang, P., & Gui, Z. (2021) An enhancement framework based on gradient domain tone mapping and fuzzy logical for X-ray image of complex workpiece. *NDT & E International*, 121, 102455. <https://doi.org/10.1016/j.ndteint.2021.102455>
 12. Lotufo, R. A., Rittner, L., Audigier, R., Machado, R. C., & Saúde, A. V. (2011). Morphological image processing applied in biomedicine. In T. M. Deserno (Ed.), *Biomedical image processing* (pp. 107–129). Berlin, Germany: Springer. https://doi.org/10.1007/978-3-642-15816-2_4
 13. Ma, M., Fu, H., Wang, Y., Lu, Z., Wang, Z., Cheng, J., Liu, X., Dai, H., Pi, F., & Wang, J. (2025). Nondestructive classification of internal defects in camellia seeds and chestnuts using X-ray imaging: Optimization and comparison of deep learning classifiers. *Food Control*, 176, 111367. <https://doi.org/10.1016/j.foodcont.2025.111367>
 14. Mendoza, F., Dejmek, P., & Aguilera, J. M. (2006). Calibrated color measurements of agricultural foods using image analysis. *Postharvest Biology and Technology*, 41(3), 285–295. <https://doi.org/10.1016/j.postharvbio.2006.04.004>
 15. Sezgin, M., & Sankur, B. (2004). Survey over image thresholding techniques and quantitative performance evaluation. *Journal of Electronic Imaging*, 13(1), 146–165. <https://doi.org/10.1117/1.1631315>
 16. Tian, H., Wang, T., Liu, Y., Qiao, X., & Li, Y. (2020). Computer vision technology in agricultural automation —A review. *Information Processing in Agriculture*, 7(1), 1-19. <https://doi.org/10.1016/j.inpa.2019.09.006>
 17. Wu, D., & Sun, D-W. (2013). Advanced applications of hyperspectral imaging technology for food quality and safety analysis and assessment: A review — Part I: Fundamentals. *Innovative Food Science & Emerging Technologies*, 19, 1-14. <https://doi.org/10.1016/j.ifset.2013.04.014>
 18. Zhang, H., Ji, S., Shao, M., Pu, H., & Zhang, L. (2023). Non-destructive Internal Defect Detection of In-Shell Walnuts by X-ray Technology Based on Improved Faster R-CNN. *Applied Sciences*, 13(12), 7311. <https://doi.org/10.3390/app13127311>
 19. Zhang, H., Ning, X., Pu, H., & Ji, S. (2024). A novel approach for non-destructive detection of shriveling degrees in walnuts using improved YOLOv5n based on X-ray images. *Postharvest Biology and Technology*, 214, 113007. <https://doi.org/10.1016/j.postharvbio.2024.113007>
 20. Zhang, T., Wang, J., Xu, D., Bai, X., Li, W., Cao, J., & Niu, X. (2025). The Study on Shell-Breaking Characteristics and Crack Formation Patterns of Walnuts (*Juglans Sigillata* D.) Under Static Loading. *Journal of Food Process Engineering*, 48(7), e70182. <https://doi.org/10.1111/jfpe.70182>

تخمین جرم کل و مغز گردو با استفاده از روش نگاشت تضعیف پرتو ایکس

علی طاهری^{۱*}، مجتبی عسکری^۱

تاریخ دریافت: ۱۴۰۵/۰۴/۳۰

تاریخ پذیرش: ۱۴۰۵/۰۷/۰۶

چکیده

در این مطالعه، روشی غیرمخرب برای تخمین جرم مغز گردو با استفاده از بازرسی پرتو ایکس پیشنهاد شده است. برای استخراج چهار ویژگی مبتنی بر تضعیف و شدت (که از نواحی تفکیک شده‌ی مغز گردو به دست آمده بودند)، از تصاویر رادیوگرافی پرتو ایکس ۳۵۰ عدد گردو استفاده شد. به منظور برقراری رابطه میان این ویژگی‌ها و جرم کل گردو و همچنین جرم مغز آن، سه رویکرد رگرسیونی شامل رگرسیون خطی، چندجمله‌ای درجه دوم و نمایی با استفاده از تقسیم‌بندی ۲۰/۸۰ داده‌ها برای برازش و آزمون (۲۸۰ نمونه برای برازش و ۷۰ نمونه برای آزمون) به کار گرفته شدند. مدل چندجمله‌ای درجه دوم (کوادراتیک) در هر دو متغیر پاسخ، عملکردی به مراتب بهتر از مدل‌های خطی و نمایی از خود نشان داد. برای جرم کل، این مدل به مقادیر $R^2=0.87$, $RMSE=0.75$ g, $MAE=0.59$ g و خطای نسبی ۶/۱۵٪ در مجموعه آزمون دست یافت. برای جرم مغز گردو نیز همین نوع مدل، مقادیر $R^2=0.72$, $RMSE=0.73$ g, $MAE=0.58$ g و خطای نسبی ۱۷/۱۳٪ را ثبت کرد. نتایج حاکی از آن است که ویژگی‌های مبتنی بر تضعیف پرتو ایکس، در ترکیب با یک چارچوب رگرسیون غیرخطی، می‌توانند برآوردی قابل اعتماد از جرم کل گردو و برآوردی با دقت متوسط از جرم مغز آن ارائه دهند. این رویکرد جایگزینی غیرمخرب و مبتنی بر تصویر برای روش‌های دستی شکستن و وزن کشی محسوب می‌شود و پتانسیل کاربرد در سیستم‌های خودکار درجه‌بندی و کنترل کیفیت در فرآوری گردو را دارد.

واژه‌های کلیدی: پرتو ایکس، تابع تخمین، جرم مغز گردو، گردو، نقشه تضعیف

۱- پژوهشکده کاربرد پرتوها، پژوهشگاه علوم و فنون هسته ای، تهران، ایران

(*) نویسنده مسئول: Email: at1361@aut.ac.ir