Turkish Journal of Agriculture - Food Science and Technology, 5(10): 1228-1236, 2017

Turkish Journal of Agriculture - Food Science and Technology

Available online, ISSN: 2148-127X
www.agrifoodscience.com,
Turkish Science and Technology

A Post-Harvest Prediction Mass Loss Model for Tomato Fruit Using A Numerical Methodology Centered on Approximation Error Minimization

Javier Bucio[1], Cesar Isaza[1], Amilcar Rizzo[1], Jonny Zavala[1], Karina Anaya[1], Enrique González[2]

[1]*Universidad Politécnica de Querétaro, Ingeniería en Telemática, Carretera Estatal 420, sn. El Marqués, Querétaro de Santiago, CP. 76420 México*
[2]*Universidad Politécnica de Tulancingo, Postgrado Doctorado en Optomecatronica, Ingenierias No. 100, Colonia Huapalcalco, Tulancingo, Hidalgo, México*

A R T I C L E I N F O

Research Article

Received 25 June 2017
Accepted 26 August 2017

Keywords:
Post-harvest
Tomato fruit
Mass loss
Least squares
Cross validation

*Corresponding Author:
E-mail: francisco.bucio@upq.edu.mx

A B S T R A C T

Due to its nutritional and economic value, the tomato is considered one of the main vegetables in terms of production and consumption in the world. For this reason, an important case study is the fruit maturation parametrized by its mass loss in this study. This process develops in the fruit mainly after harvest. Since that parameter affects the economic value of the crop, the scientific community has been progressively approaching the issue. However, there is no a state-of-the-art practical model allowing the prediction of the tomato fruit mass loss yet. This study proposes a prediction model for tomato mass loss in a continuous and definite time-frame using regression methods. The model is based on a combination of adjustment methods such as least squares polynomial regression leading to error estimation, and cross validation techniques. Experimental results from a 50 fruit of tomato sample studied over a 54 days period were compared to results from the model using a second-order polynomial approach found to provide optimal data fit with a resulting efficiency of ~97%. The model also allows the design of precise logistic strategies centered on post-harvest tomato mass loss prediction usable by producers, distributors, and consumers.

DOI: https://doi.org/10.24925/turjaf.v5i10.1228-1236.1395

Introduction

The ripening process of the tomato fruit during post-harvest implies dramatic physical changes in color, shape, mass, area and volume. It also classifies it as a climacteric fruit, meaning that post-harvest tomatoes continue to ripe (Alexander and Grierson, 2002). At a harvest point, tomatoes are green and hard to the touch. But manipulation leading to the end consumer causes deterioration conducing to what is considered their maximum desirable ripening condition (Liu et al., 2015). Tomatoes are one of the most important climacteric fruits produced in Mexico, representing 20% of the world's export volume (Mapes and Basurto, 2016). In the last few years, Mexico has increased its tomato production by 34%, going from 2.086 million tons in 2000 to 2.8 million tons in 2015 (Domis and Papadopoulos, 2002). This is mainly due to the increasing use of protected agricultural systems such as greenhouses. These have made all around seasonal tomato production possible (Putra and Yuliando, 2015).

During post-harvest, a great amount of product is lost to physical, chemical and mechanical processes. Companies can then experience limitations in market supply, resulting in economic losses for them and negative affectation to the end consumer. Furthermore, in developing countries like Mexico there is a gap between market and infrastructure, and product losses can fluctuate from 25 to 50% of the harvest. It has been found that tomato losses are due to two main factors: total mass loss and undesirable short ripening process time. These cause extreme fruits softening, leading to it being considered waste (Payasi and Sanwal, 2010). These factors represent significant damage for producers, distributors and final consumers. Consequently, both agricultural industry and market could benefit with better understanding and predictive knowledge on relevant timelines and other aspects of the tomato fruit ripening process.

Due to its economic importance and fast ripening process, tomato is biochemically and physically studied by scientists. In this context, the need arises to better analyze and predict its mass loss and ripening process times during post-harvest. Better management of ripening times will reward producers, distributors and end consumers economically (cash and product savings) and in terms of improved quality (fruit freshness) at the shelf. There is a significant number of studies related to biochemical processes taking place in post-harvest fruits. For example, Ponce-Valadez et al., (2016), evaluated by biochemical analysis the effects of refrigerated storage over tomato fruit. Kant et al., (2016), studied the biochemical characteristics of the tomato fruit when treated with different concentrations of salicylic acid (SA). He noted that the treatment delayed ripening and deterioration of the fruit. Wabali et al., (2016), biochemically evaluated the effects of potassium permanganate treatment at different concentrations in chilled tomatoes, under a color and texture analysis of fruit. He observed that the potassium permanganate had a preservative effect on the color and texture of refrigerated tomatoes. Bhandari and Lee, (2016), biochemically studied changes tomato fruit ripening, based on the dependence of antioxidants, color attributes and antioxidant activity of the tomato fruit in 45 post-harvest.

In contrast, there are just a few studies related to physical parametric changes in fruits during post-harvest. Examples are Kvikliene et al., (2006), who studied changes in the quality parameters of the apple fruit with the purpose of estimating the optimum harvest time. Liu et al., (2007), used an adapted 50-tomato growth model to predict carbon and water accumulation in peaches. Main adjustments were based on the decrease of the cellular wall during the fruits development and the negative influence of its initial mass/weight. Bornn et al., (2010), computationally analyzed the efficiency of mathematical methods like sensibility analysis and crossed validity, comparing them with the Monte Carlo sequential method based on Markov chains. Li et al., (2013), studied the effects of mechanical damage that tomato endures during storage. He observed that both the storage technique and time had a significant effect on the mass/water loss of the fruit. De Oliveira et al., (2014), evaluated a prediction model of the internal features of three different fruits using the NIRS (Near Infrared Spectroscopy) technique and quimio-metric methods with least squares. Kawamoto and Kabashima, (2016), applied based on the cross validation method for identifying macroscopic temporary structures, and hidden data in modular networks by predicting the minimum data spread error in disperse networks. Pila et al., (2010), studied the effects of physical-chemical treatments in the post-harvest tomato. De Ita et al., (2015), determined the dehydration curves of fruits and vegetables under equal conditions using DTA (Differential Thermal Analysis) and TGA (Thermal Gravimetric Analysis). Correia et al., (2015), evaluated temperature, time and tomato thickness effects during the adiabatic drying dehydration process in industrial conditions.

Numerical prediction models are simplified representations of reality applied to predictive processing by sets of hypotheses. They are used for explaining behavior patterns observed in the real world within different contexts. The dependent variable of interest is explained and predicted based on their own history and history of other related variables. Applying different types of functions to approximate the behavior and properties of the variable of interest requires careful study and plays a key role in the resulting predictive model capacity (Faraway, 2016). Prediction models present errors, which can be of different types: specification errors, approximation errors and estimation errors (Fuller, 2009). In this study, the framework of predictive models is used, and interest centers on the getting of one regarding the post-harvest mass loss prediction of the tomato fruit and its approximation error respect to the experimental data.

In this study, the individual mass of fifty tomatoes, taken from a protected agricultural environment located in Colón, Querétaro, México was analyzed during the entire post-harvest ripening process. Each specimen was harvested at 8:30 am on the 13th of April 2016 and its mass measured every 24 hours during a 54 days period. With the data obtained from that experimental phase, a timeline of the ripening process was established. That data was in turn used to develop a model to predict the mass loss of a tomato during the ripening process until desirable maturity. The model is based on algebraic polynomial procedures, least squares and crossed validation, all within the context of applied artificial intelligence algorithms. Its main objective is to be used as a systematic tool for producers and consumers to predict mass loss and assess the correct ripening condition of the fruit.

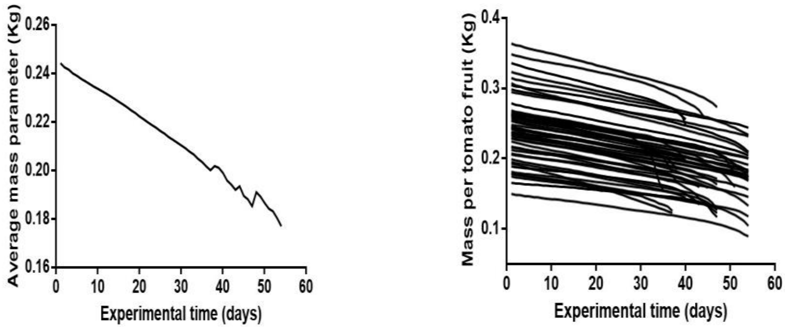
This paper is divided in six sections. After the introductory section, second section describes the analysis performed on the sample and the processes conducted to acquire the initial experimental data, as well as a description of related precedent studies using least squares and crossed validation. Third section describes measurements implied and follows up methodology developed to extract the approximation model by automatic means. Fourth section explains the mathematical formulas employed to develop the approximation model and describes the learning and evaluating methods using the experimental data extracted from it. Fifth section presents a comparison between the results of the analysis and the experimental behavior of the model. Finally, sixth section presents the conclusions of this research. Comments on prospective related research based on the here proposed are provided as well, since the proposed mass loss prediction model presented an efficiency of $\sim$97%.

Materials and Methods

For this study a sample of 50 tomatoes off the better ball variety was harvested on the 13th of April 2016 at 8:30 am. The sample was obtained from a greenhouse property of the High Group Farm Company located in the town of Colón, municipality of Ajuchitlán, state of Querétaro, México (latitude of 20°41’04.58” and longitude 100°00’23.457”). The sample was stored in a Samsung domestic refrigerator. The storage, distribution of the sample was divided in three sections: two with 24 fruits and the last one with 2 fruits. The storage temperature and relative humidity was kept controlled at 14°C and 39% respectively, while the environmental temperature and relative humidity varied between 23–29°C and 30–34% respectively. A Taylor digital precision scale (TE32C model) was used for the daily registry of each specimen’s mass, 120 carefully labeled.

Data Collections

The initial mass figure for the entire sample was about 11.9685 kg, and the average mass per tomato was about 0.2443 kg. Measurements were repeated every 24 hours for 54 consecutive days. The period for mass measurement and its registry was from 8:00 pm until 1:00 am. With the obtained figures, a database was created in order to develop a timeline describing the mass behavior of the total sample, but measurements were performed on each specimen as well, allowing for individual mass behavior time lines too. Figure 1(a), presents the total mass average sample timeline behavior during the 54 days the experiment lasted. Complementary, Figure 1(b) presents a mass loss timeline for each individual tomato constituting the sample. From Figures 1, it can be observed that the behavior of the curves is similar and independent of the number of fruits of the sample taken.



(a) (b)

Figure 1 (a) Average mass parameter of the total sample vs. experimental time in days, (b) Mass parameter per fruit vs. experimental time in days.

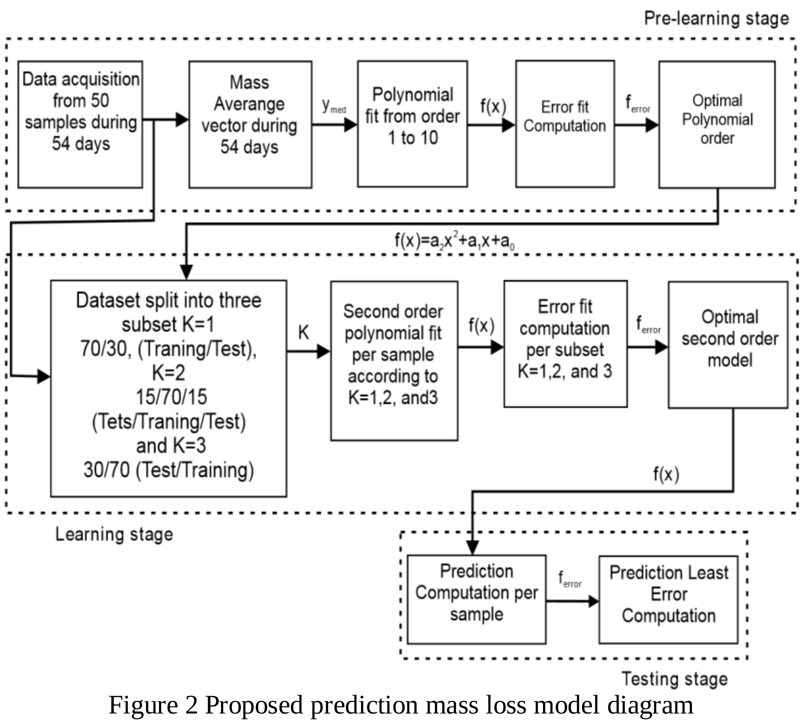


Figure 2 Proposed prediction mass loss model diagram

Proposed Methodology

The methodology proposed to develop a mass loss prediction model for the tomato fruit is divided into three stages. The first stage is the pre-learning of the data treatment model to be developed, second stage its learning, and third stage its testing in order to obtain the optimal mass loss predictive model. The general data treatment model developed in this study is described in Figure 2.

Model Pre-Learning Stage

According to the methodology in Figure 2, pre-learning stage starts with data collection of the tomato mass parameter during a determined time period. This was conducted for each individual tomato of the sample, obtaining a data set with the form $(x_0, y_0), (x_1, y_1), \ldots, (x_n, y_n)$ where x_n represents the experimental time length and y_n the registered mass of the fruit during that period. Data obtained was represented graphically in order to study mass behavior and establish timelines from fruit harvest to the end of the optimal ripening process and beginning of fruit discarding. These time lines can be observed in their both average mass of the sample and in individual form, in the figures 1(a) and 1(b) respectively. The data was then subjected to a mathematical process based on least squares linear approximation. The Weierstrass approximation theorem Burden and Faires (2011) is applied as follows:

Assume that f is defined and continuous in $[a, b]$, for every $\varepsilon > 0$, there exists a polynomial $P(x)$, with the property that $|f(x) - P(x)| < \varepsilon, \exists x \in [a, b]$. Taking into account that, we have a set of continuous functions f(x) of the type,

$$f(x)=a_n x^n+a_{n-1}x^{n-1}+a_{n-2}x^{n-2}+...+a_2x^2+a_1x^1+a_0 \quad (1)$$

where n is a non-negative integer number and represents the degree of the polynomial to be used. $a_0, a_1, \dots, a_n$ are constants that belong to a set of real numbers. Therefore, Figure 1 represents a set of functions $f(x)$, whose data (x_i, y_i) are the real mass and time values obtained from experimentation. Then we look for a polynomial of degree n that approximates the behavior of the real data of the mass of the fruit given by

$$P(x)=a_o+a_1x+...a_nx^n \quad (2)$$

where $P_n(x)$, represents the least squares polynomial approximation of degree n to be found. Since $a_0, a_1, \dots, a_{n-1}$ are unknown constants, Shalev-Shwartz and Ben-David (2014), equation (2) is applied in a dataset of degrees $n = 1,2, \dots ,10$, with which a set of polynomials of the type $P(x)$ is obtained. Finally, we obtain the least square approximation error of the set of functions $P(x)$ by means of

$$\varepsilon=\sqrt{\frac{\sum_{i=1}^{n} f_n(x_k)-(y_k)^2}{n}} \quad (3)$$

where n is the total number of data points, and x_i,y_i are the days of testing and mass values respectively. Based on the results obtained from equation (3), a comparison of the values of the minimum mean square error of the set of polynomials $P(x)$ is conducted in order to obtain an optimal second-order polynomial $P(x)$, as well as its corresponding numerical coefficients a_0, a_1, a_2.

Learning Stage

Crossed validation method is applied to validate the proposed mass loss model. This method implies dividing the experimental data set into two aleatory parts (Witten et al., 2011). The first one is used for model learning and the second is the data set upon which the model validation is conducted. This data division is based on the fact that $1/n$ is the probability that an element of the experimental data set (i.e., the mass of a specific tomato on a particular day) is considered to pertain to the validation set. Thus, in equation:

$$(1\text{-}1/n)^n \quad (4)$$

where n represents the number of independent experimental values for the masses of tomatoes as registered in the timeline of the experiment (54 days). Each of them can be considered or not to pertain to the learning set or the validation set. Accordingly, $1 - 1/n$, represents the probability that the mass of a specific tomato on a particular day will not be considered to pertain to the validation set. Then, $(1 - 1/n)^n$ is the probability that of the n available independent experimental values, approximately 36% of them will be used for validation while the rest is used for learning. It should be noted that this property is independent of the size of n. Therefore, for the present methodology, a sample of size $n = 54$ will conduce to the same data division as one of $n = 500$ or $n = 1000$. In this context, the timeline extension of the experiment ($n = 54\,days$) was determined by considering the period of time after which tomatoes start to be considered waste because their mass reduces in average $\sim$25% (see Figure 1(a)), in conjunction with other undesirable characteristics that begin to appear such as skin wrinkling. The actual size of the sample in terms of the number of fruits was in our case determined for a sample large enough to reflect the time independent variation of the mass parameter per fruit, as shown in Figure 1(b), which was of $\sim$25% as well.

Thus, one gets an estimated forecast error for all available observations and the possibility to interactively adjust the model for different combinations in the selection of learning data sets and conducted validations. With the K subsamples (K-fold) approach employed in this study, we consider subsets of the initial data set in order to obtain an estimate of the prediction error for each observation available. This is achieved by dividing the observations in $K = 3$ aleatory sub-samples of approximately equal size amounting to 70% of the original data to construct the model, test it within the remaining 30% subsample, and then repeat (iterate) the process 2 times more (for a total of 3) exhausting the possible arrangements of the sub-samples. The data used for the learning stage of the proposed model was the one gathered in the experimental phase as described in section 2. Figure 3 presents a diagram of the total data set, which consists of 54 observations in total.

Cross-validation method is applied with the option of K-subsamples, where $K = 3$ subsets were taken to develop the learning of the mass loss model. The data set was divided in 3 subsamples amounting to 70% of the total set, i.e., 38 observations of 54, as shown in Figure 4. The number of subsets is shown in Figure 5, as well as the arrangement of the data sub-sets used for the development of the learning process. For the first data sub-set taken was the one with $n = \{1, \dots ,38\}$, for the second sub-set with $n = \{9, \dots ,46\}$, and for the last subset with $n = \{17, \dots ,54\}$.

Testing Stage

For testing or validating the data obtained from the learning stage (the coefficients of the mass loss prediction model), the same procedure as in that first stage is developed. But this time the remaining data sub-sets were used, i.e., the remaining thirty percent of the total data set, equivalent in each iteration to 16 of the 54 registries originally gathered.

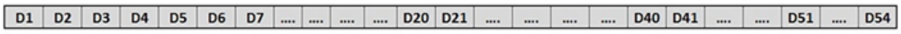


Figure 3 Total data-set

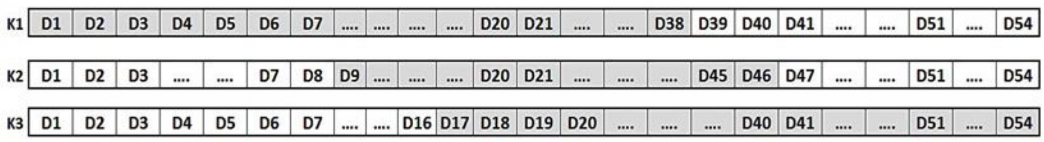


Figure 4 $K = 3$ subsets employed

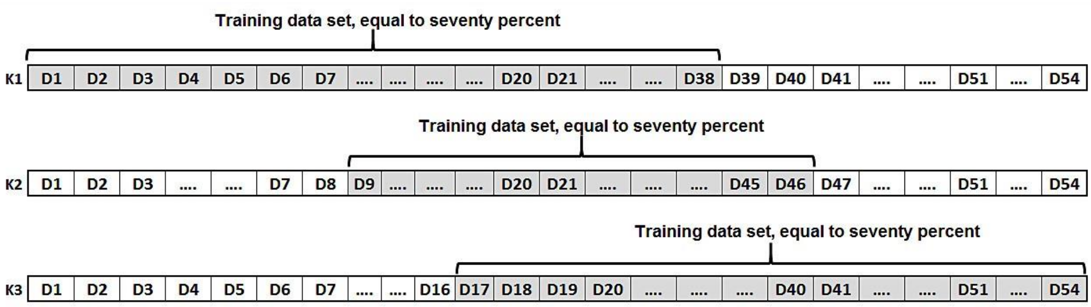


Figure 5 Learning data subsets used for mass loss model development

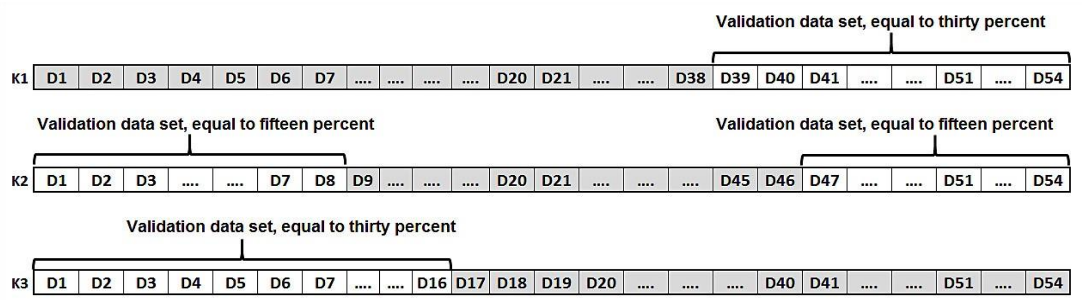
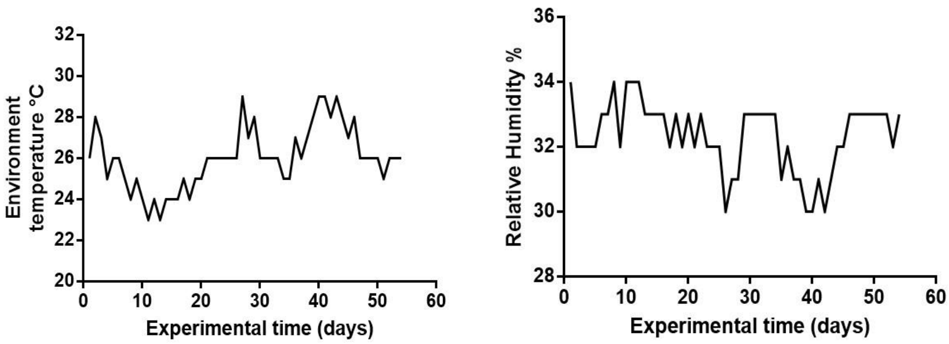


Figure 6 Data subsets used to test/validate the mass loss model



(a) (b)

Figure 7 (a) Environment temperature, (b) Environment relative humidity

Figure 6 presents the distribution of the data used for the testing stage in $K = 3$ subsets. For the first data sub-set used was the one with $n = \{39, \ldots, 54\}$, for the second sub-set with $n = \{1, \ldots, 8\} \cup \{47, \ldots, 54\}$, and for the last sub-set with $n = \{1, \ldots, 16\}$.

Results and Discussion

During the experimentation phase, mass loss was observed accompanied by other changes which are beyond the scope of this paper. Results obtained in 230 accordance with the data treatment process described in Figure 2 are presented.

Pre-Learning Stage Results

For the experiment conducted, temperatures and relative humidities were between 22–29°C and 31–34% respectively. These are typical for tomato harvest and post-harvest management within the central region of México and specific to a farm located in the vicinity of the city of Querétaro, México. Please, see Figures 7(a) and 7(b).

Figure 1(b) shows the individual mass loss of each specimen. That data was used to calculate the percentage of daily mass loss for the entire sample, as reported in Figure 8(a). At the beginning of the experiment (day 1) no mass loss 240 was considered, and this value also represents the harvest date of the specimens. At the end of the experiment (day 54) the mass of the sample presented a 64.51% change respect to its baseline.

With the data described in section 2, the numerical coefficients of the polynomial functions of degree $i = 1, \ldots, 10$ from equation (1) were calculated. These are presented in Table 1. With them, applying the equation (3), the mean square error between experimental data and model results was calculated. Figure 8(b) shows the error behavior in respect to the polynomial degree. In it, one can also notice how error values reach a stable plateau from polynomials of the second to the fourth degree. Based on these results, and preferring those lower polynomial degrees instead of the eight to the tenth degree where there's another stable plateau for the mean square error, the second degree polynomial approach is chosen as the basis of the prediction mass loss model:

$$f(x)=a_2x^2+a_1x+a_0 \quad (5)$$

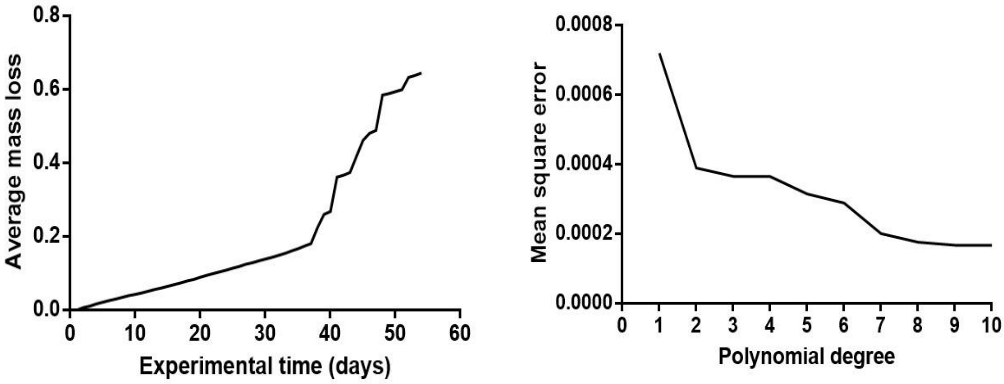
where the unoptimized numerical coefficients for the entire sample are:

$$a_2\text{=-2.7867E-6},\ a_1\text{=-1.347E-3},\ a_0\text{=0.2615} \quad (6)$$

Table 1 Numerical coefficients

D	a_0	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}
10	0.2634	-2.0661 E-3	-1.2553 E-5	1.9789 E-6	-2.2096 E-5	1.0547 E-7	-2.3405 E-9	1.3507 E-11	3.8791 E-13	-7.3922 E-15	3.8269 E-17
9	0.2633	-0.1915 E-3	-7.5516 E-5	3.1944 E-5	-3.4988 E-6	1.874 E-7	-5.591 E-9	9.4628 E-11	-8.4906 E-13	3.1319 E-15	
8	0.2639	-2.5568 E-3	1.4925 E-4	-3.758 E-6	-4.4893 E-7	3.5436 E-8	-1.0427 E-9	1.4158 E-11	-7.3924 E-14		
7	0.2647	-3.3300 E-3	3.7103 E-4	-3.1889 E-5	1.4186 E-6	-3.41686 E-8	4.2266 E-10	-2.1057 E-12			
6	0.2633	-2.2314 E-3	1.2164 E-4	-7.5866 E-6	2.2516 E-7	-3.1941 E-9	1.7313 E-11				
5	0.2626	-1.7960 E-3	4.6779 E-5	-2.2719 E-6	4.5915 E-8	-3.3742 E-10					
4	0.2618	-1.4072 E-3	-8.9999 E-7	7.2048 E-9	-4.80715 E-10						
3	0.2619	-1.4307 E-3	9.813 E-7	-4.5673 E-8							
2	0.2615	-1.347 E-3	-2.7867 E-6								
1	0.2629	-1.5003 E-3									

D: Degree polynomial



(a) (b)

Figure 8 (a) Average mass loss vs experimental time in days, (b) Mean square error vs polynomial degree

Learning Stage Results

Having chosen a second degree polynomial as the one to build the mass loss prediction model upon (equation 5), we conduct the learning stage to obtain optimized numerical coefficients for each of the K subsets. Coefficients obtained are shown in Table 2.

Testing Stage Result

Once defined the numerical coefficients of the learning stage shown in Table 2, these apply to the testing stage for validation of the model data. In sequence: the normalization of the average weight of the tomato sample was performed, as presented in Figure 9(a). Then, that data is compared to the cross-validation results in both Table 3 and Figure 9(b) where we see a comparison of our model against the data obtained from experimentation.

Mass Loss Model Results

With the results obtained from the testing stage and summarized on Table 3 and Figure 9(b), we can appreciate that the best subset to approach the experimental data is the one with $K = 2$ because it has the least approximation error (2.45%) of the global exercise. Therefore, the optimal mass loss prediction model in the form of equation 5 is given by the following optimal coefficients for an approximate data acceptance rate of ~97%:

$$a_2=-4.11907E-5,\ a_1=-3.28077E-3,\ a_0=0.99515521 \quad (7)$$

Conclusions

The marketing process of tomato fruit in Mexico requires several logistical steps in order to bring fresh fruits from the harvest to the consumer. Depending on the market to which it is directed (national or international), the post-harvest handling of the tomato will include: loading, transportation, reception, storage, sorting and sale. There are few cases in which the agricultural producer has a direct connection with supermarkets, food processing companies or end consumers. In the case of the internal or national market there are two main marketing channels for tomatoes: in the first one, the producer destines its packaged production to supply centers, self-service stores and processing companies to finally distribute it to consumers. In the second one, intermediaries collect the farmer's output and send it to local markets and warehouses. On the other hand, in the international market, the Mexican tomato producers send their packed production to a broker who is in charge of channeling it into self-service stores and distributors which make it arrive to the foreign consumer. The packaging is usually made in cardboard boxes according to size, color and quality of the fruit. The most common presentation of packaging for the international market, is a cardboard box with capacity for 33 kg of fruit. In the internal market, the fruits are packed in cartons with a capacity for 13 kg of fruit. The boxes packed product are arranged in stowage using wooden pallets and plastic strap to hold them and support the transportation, usually in refrigerated trailer. These are sent to different distribution points such as: central city markets, commercial chains, local markets, supermarkets, among others. Subsequently, the fruit is stored in refrigerated containers pending their sale or final consumption. In a tomato-consuming country like Mexico, the purchase and sale of the fruit by wholesale or retail at any point of distribution is offered based on the mass of the fruit (kilograms). During cooled post-harvest storage time, the tomato fruit suffers a loss of mass, as well as undesirable loosening and softening of the fruit. This causes waste of product, leading to considerable economic losses. In order to find an angle to this problem, a model of prediction of loss of mass of the tomato fruit in post-harvest was developed. The model provides a useful tool for both producers and consumers, because it allow them to establish optimization parameters for the purchase, sales, processing and consumption of the fruit. That is, the model allows producers to estimate the cost-benefit of product marketing tomato based on processing time. On the other hand, the model also allows the final consumer to estimate the maximum durability time of the product and make a better use of it, and supermarket chains to set time limits on product durability for sale. Near this limit, if the product has not been marketed yet, it could be sent to food processing companies, avoiding fruit waste. In a word, the model helps to optimize the post-harvest logistical managing of the fruit throughout the marketing chain.

Table 2 Optimized coefficients obtained from learning stage

Description	a_2	a_1	a_0
Subset K = 1	-3.80182E-05	-0.003261324	0.995263397
Subset K = 2	-4.11907E-05	-0.00328077	0.99515421
Subset K = 3	-5.61685E-05	-0.002789021	0.993191135

Table 3 Approximation error obtained from testing stage

Description	a_2	a_1	a_0	Approximation Error
Total approximation model	-4.87504E-05	-0.00308615	0.994152511	2.66%
Subset K = 1	-3.80182E-05	-0.003261324	0.995263397	2.79%
Subset K = 2	-4.11907E-05	-0.00328077	0.99515421	2.45%
Subset K = 3	-5.61685E-05	-0.002789021	0.993191135	2.75%

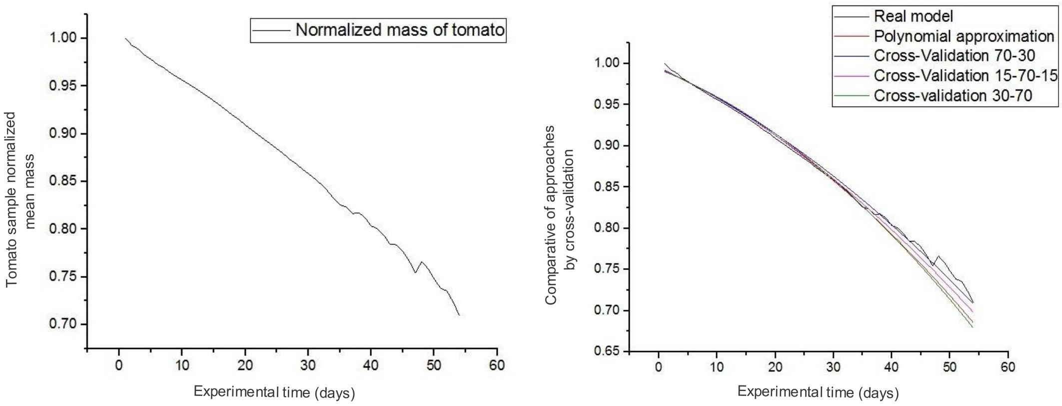


Figure 9 (a) Tomato sample normalized mean mass vs experimental time in days, (b) Comparison of approaches by cross-validation

In summary, the post-harvest mass loss prediction model for the tomato fruit developed in this paper presented a high accuracy efficiency in terms of relative percentage total sample mass loss (97.55%) and low approximation error (2.45%) when compared to experimental data. That was achieved by means of a second order polynomial approximation to the experimental data obtained as described in previous sections. Therefore, the post-harvest tomato mass loss prediction model developed here can be a useful tool for both producers and consumers to better manage post-harvest market logistics challenges. In this context, the model could allow producers and consumers to reduce the economic losses that affect the tomato during the ripening process and marketing. Moreover, this basic methodology to construct the mass loss model can also be applicable to other fruits with similar climacteric characteristics such as mango, banana, pear, peach, apricot, apple, avocado, etc.

Future Work

Among the side results of this study there is a database with dimensional parameters such as area, diameter (longitudinal, signal and transverse), complementing the acquisition of digital images during the tomato fruit ripening process. This information will allow further studies related to the physical structure of the fruit related to parameters such as volume, surface characteristics, color and their changes while ripening.

Acknowledgements

The present study was carried out, thanks to the facilities provided by the company High Tech Farm Mexico for the development of the model of prediction of mass of tomato fruit.

References

Alexander L, Grierson D. 2002. Ethylene biosynthesis and action in tomato: a model for climacteric fruit ripening. Journal of Experimental Botany, 53: 2039–2055. DOI: https://doi.org/10.1093/jxb/erf072

Bhandari SR, Lee JG. 2016. Ripening-dependent changes in antioxidants, color attributes, and antioxidant activity of seven tomato (*Solanum lycopersicum L.*) cultivars. Journal of Analytical Methods in Chemistry, Volume 2016, Article ID 5498618, 13 pages. DOI: 10.1155/2016/5498618

Bornn L, Doucet A, Gottardo R. 2010. An efficient computational approach for prior sensitivity analysis and cross-validation. Canadian Journal of Statistics, 38: 47–64. DOI: 10.1002/cjs.10045

Burden R, Faires JD. 2011. Numerical Analysis. 9th Edition, Boston, MA, USA, Brooks/Cole, Cengage Learning, ISBN-13: 978-0-538-73351-9.

Correia A, Loro A, Zanatta S, Spoto M, Vieira T. 2015. Effect of temperature, time, and material thickness on the dehydration process of tomato. International Journal of Food Science, Volume 2015, Article ID 970724, 7 pages. http://dx.doi.org/10.1155/2015/970724

De Ita A, Flores G, Franco F. 2015. Fruits and vegetables dehydration. VII International Congress of Engineering Physics, Mexico City, Mexico, 24-28 November 2014, IOP Publishing. Journal of Physics: Conference Series 582(2015) 012065 DOI: 10.1088/1742-6596/582/1/012065

De Oliveira GA, Bureau S, Renard CMGC, Pereira-Netto AB, De Castilhos F. 2014. Comparison of nirs approach for prediction of internal quality traits in three fruit species. Food Chemistry 143: 223–230.

Dorais M, Papadopoulos A, Gosselin A. 2001. Greenhouse tomato fruit quality. (Eds. Janick, J.). Horticultural Reviews, Volume 26, NY, john Willey & Sons, Inc., pp 239-319, ISBN: 0-471-38789-4.

Faraway J. 2016. Extending the Linear Model with R: Generalized Linear, Mixed Effects and Nonparametric Regression Models, 2nd Edition. Boca Raton, FL. USA. Chapman & Hall, ISBN: 13: 978-1-4987-2098-4.

Fuller WA. 1987. Measurement Error Models. Wiley Series in Probability and Statistics. NY, USA. John Wiley & Sons Inc., ISBN: 0-471-86187-1.

Kant K, Arora A, Singh V. 2016. Salicylic acid influences biochemical characteristics of harvested tomato (*Solanum lycopersicon L.*) during ripening. Indian Journal of Plant Physiology 21: 50–55.

Kawamoto T, Kabashima Y. 2017. Cross-validation model assessment for modular networks. Available from: https://arxiv.org/abs/1605.07915v2 [Accessed Date 01.07.2017] DOI: 10.1038/s41598-017-03623-x

Kvikliene N, Kviklys D, Viskelis P. 2006. Changes in fruit quality during ripening and storage in the apple cultivar "Auksis". Journal of Fruit and Ornamental Plant Research 14(2): 195-202.

Li Z, Li P, Liu J. 2013. Effect of mechanical damage on mass loss and water content in tomato fruits. International Agrophysics, 25: 77–83.

Liu HF, G´enard M, Guichard S, Bertin N. 2007. Model-assisted analysis of tomato fruit growth in relation to carbon and water fluxes. Journal of Experimental Botany, 58: 3567–3580. DOI: 10.1093/jxb/erm202

Liu M, Pirrello J, Chervin C, Roustan JP, Bouzayen M. 2015. Ethylene control of fruit ripening: revisiting the complex network of transcriptional regulation. Plant Physiology, 169: 2380–2390. DOI: https://doi.org/10.1104/pp.15.01361

Mapes C, Basurto F. 2016. Biodiversity and edible plants of México. (Eds. Rafel L., Casas A., Blancas J) In: Ethnobotany of Mexico: Interactions of People and Plants in Mesoamerica. Springer, NY. USA. pp. 83–131. ISBN: 978-1-4614-6669-7. DOI: 10.1007/978-1-4614-6669-7 5

Payasi A, Sanwal G. 2010. Ripening of climacteric fruits and their control. Journal of Food Biochemistry, 34: 679–710. DOI: 10.1111/j.1745-4514.2009.00307.x

Pila N, Gol NB, Ramana Rao TV. 2010. Effect of postharvest treatments on physicochemical characteristics and shelf life of tomato (Lycopersicon esculentum mill.) fruits during storage. American-Eurasian Journal Agricultural & Environmental Sciences, 345(9): 470–479.

Ponce-Valadez M, Escalona-Buendía HB, Villa-Hernández JM, de León-Sánchez FD, Rivera-Cabrera F, Alia-Tejacal I, Pérez-Flores LJ. 2016. Effect of refrigerated storage (12.5 c) on tomato (*Solanum lycopersicum*) fruit flavor: A biochemical and sensory analysis. Postharvest Biology and Technology, 111: 6–14. https://doi.org/10.1016/j.postharvbio.2015.07.010

Putra PA, Yuliando H. 2015. Soilless culture system to support water use efficiency and product quality: a review. Agriculture and Agricultural Science Procedia, 3: 283–288. https://doi.org/10.1016/j.aaspro.2015.01.054

Shalev-Shwartz S, Ben-David S. 2014. Understanding Machine Learning: From Theory to Algorithms. Understanding Machine Learning: From Theory to Algorithms. Cambridge University Press. ISBN 978-1-107-05713-5.

Wabali VC, Esiri A, Zitte L. 2016. A sensory assessment of color and textural quality of refrigerated tomatoes preserved with different concentrations of potassium permanganate. Food Science & Nutrition. 5(3): 434-438. DOI: 10.1002/fsn3.410

Witten I, Frank E, Hall M. 2011. Data Mining: Practical Machine Learning Tools and Techniques. The Morgan Kaufmann Series in Data Management Systems. 3rd Edition, Burlington, MA, USA. Elsevier Science. ISBN: 978-0-12-374856-0.