Multilinear Regression Approach in Predicting Osmo-Dehydration Processes of Apple, Banana and Potato

Charles Tortoe1*, John Orchard2 and Anthony Beezer3

1Council for Scientific and Industrial Research-Food Research Institute, P. O. Box M20, Accra, Ghana.
2Natural Resources Institute-FMMG, University of Greenwich, Medway Campus, Central Avenue, Chatham, Kent, ME4 4TB, U.K.
3School of Pharmacy, University of London, 29-39 Brunswick Square, London, U.K.

Abstract

The potential for improving food quality through osmo-dehydration is tremendous but limited by quantitative data and methods. A Multiple Linear Regression (MLR) approach was developed for water loss and solid gain during osmo-dehydration of apple, banana and potato taking into account the effect of temperature, concentration, time of immersion, sample size, sample type and agitation. Temperature was the most important factor influencing osmo-dehydration of the plant materials whereas agitation was the least. A regression coefficient of determination ($R^2 = 0.886$) indicating a good correlation coefficient ($r = 0.941$) between experimental and predicted data was identified for water loss. However, the regression coefficient of determination ($R^2 = 0.305$) for the solid gain did not show a good regression correlation coefficient ($r = 0.552$) between the experimental data and the predicted data. Prediction of water loss was more adequate than solid gain due to the variability of the pathways of water and solid diffusion into the different plant materials in favour of water loss.

Different models have been reported [2,23,24]. Most models are based on the assumption that mass transfer is described by a simplified unsteady state Fickian diffusion model [25,26]. According to the authors effective diffusivities are calculated by regression analysis of specific mass transport data. However, the uses of such models are largely limited to the specific experimental set up [23]. Raoult-Wack [2] reported that the fundamental knowledge for the prediction of the mass transport is still a grey area although considerable efforts have been made to improve the understanding of mass transfer in osmo-dehydration. Normally, two methods are used to determine the kinetics of osmo-dehydration. First, a continuous method that involves the measurement of weight loss of a single sample and its final moisture content at the end of the process [27]. This is rather recent but promises a lot of improvements over the second method, the discontinuous method where measurements of water loss and solid gain are carried out on separate samples supposed to be the same in terms of geometry and dimensions, weight, volume and initial moisture content. The continuous method allows a more precise determination

Keywords: Osmo-dehydration; Multiple Linear Regression; Water loss; Solid gain; Apple; Banana; Potato

Nomenclature: $m_0$: Initial mass; $m$: Mass after sample period; $s$: Mass of solid after sample period respectively; $t$: Sampling period; $W$: Water Loss; $GG$: Solid Gain; $T$: Temperature; $C$: Concentration of osmotic solution (sucrose); $t$: (min): Time of immersion; $St$: Sample type; $S$: Sample size; $Ag$: Agitation

Introduction

A Multiple Linear Regression (MLR) attempts to model the relationship between two or more explanatory variables and a response variable by fitting a linear equation to observed data. In food product development, the application of MLR among other methodologies generally leads to insights into possible mechanisms of changes in the foods, which in turn leads to new products/process development. Therefore, as a predictive application, MLR helps to explore the potential for improving existing processes without performing numerous, often expensive experiments. The potential for improving food quality through osmo-dehydration is tremendous. In general, water loss and solids gain in plant materials provided a useful tool for understanding the osmo-dehydration process for industrial applications.

Osmo-dehydration of plant materials is often applied as a pre-processing step to remove a large portion of the water content, which is accompanied by simultaneous transfer of osmoactive substances into the plant material matrix before the plant materials are subjected to further processing techniques such as air drying. The process of de-watering and direct formulation of a product is achieved by introducing further processing techniques such as air drying. The process of de-watering and direct formulation of a product is achieved by introducing further processing techniques such as air drying.

The effect of several factors on the osmo-dehydration of foods had been investigated [2,5,9,15-19]. Advances in the field were reviewed by many authors [2,4].

Different materials and forms have been applied in predicting osmo-dehydration in plant materials [5,9]. Pineapple rings of 1 cm thick had been used [20]. Potato halves had been used on unidirectional mass transfer [21]. Other authors used 4.0 cm agar gel cubes to study spatial mass distribution and 0.9 cm agar cubes for studying the effects of concentration, temperature and solute molecular weights of osmotic solutions on water loss and solid gain [22].

Different models have been reported [2,23,24]. Most models are based on the assumption that mass transfer is described by a simplified unsteady state Fickian diffusion model [25,26]. According to the authors effective diffusivities are calculated by regression analysis of specific mass transport data. However, the uses of such models are largely limited to the specific experimental set up [23]. Raoult-Wack [2] reported that the fundamental knowledge for the prediction of the mass transport is still a grey area although considerable efforts have been made to improve the understanding of mass transfer in osmo-dehydration. Normally, two methods are used to determine the kinetics of osmo-dehydration. First, a continuous method that involves the measurement of weight loss of a single sample and its final moisture content at the end of the process [27]. This is rather recent but promises a lot of improvements over the second method, the discontinuous method where measurements of water loss and solid gain are carried out on separate samples supposed to be the same in terms of geometry and dimensions, weight, volume and initial moisture content. The continuous method allows a more precise determination

*Corresponding author: Charles Tortoe, Council for Scientific and Industrial Research- Food Research Institute, P. O. Box M20, Accra, Ghana. Tel: +233-243241801; Fax: +233-21500111; E-mail: ctortoe@yahoo.co.uk

Received September 15, 2011; Accepted November 02, 2011; Published November 04, 2011


Copyright: © 2011 Tortoe C, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.
of experimental points and also helps in the prediction of the variations of the moisture content with respect to time.

Magee et al. [28] used a rate parameter to model osmo-dehydration of apple slices as a function of the concentration and temperature of the osmotic solution. This parameter was calculated from the slope of the straight line obtained from apple sugar concentration vs. square root of time. However this model was limited in the information that can be derived from it. Biswal et al. [29] used a similar empirical model for osmo-dehydration of sweet beans.

Developing a single empirical equation from different plant materials to describe the relationship between the dependent and independent variables had been a challenge in the past. Therefore, literature had been silent on the development of a single empirical equation from different plant materials in which sample type is a variable due to the different nature of plant materials resulting in the different rates of mass transfers. In this study, attempt was made to develop single empirical equation to identify and quantitatively predict the mass transfer processes of osmo-dehydrated apple, banana and potato.

Materials and Methods

Plant materials

Golden Delicious and Cox apple (Malus domestica Borkh) varieties with maturity levels of 150 days after full bloom (DAFB), dark-green and dark-red skin colour, respectively, were obtained from Horticultural Research International, East-Malling, Kent, England. Potato (Solanum tuberosum L.) variety Estima and banana (Musa spp.) cultivar Cavendish were purchased from a local supermarket in Chatham, Kent, England. The early stage of ripeness of the peel colour of than green. Banana fruit from a single bunch was utilized for each experiment. Banana was checked with the colour plates on a standard banana ripening chart [30]. Fruits selected were at stage four of ripeness, more yellow.

Chemicals

Sucrose and ascorbic acid were purchased from Sigma-Aldrich Chemical Company Limited, London, United Kingdom.

Experimental design

Osmo-dehydration of the plant materials were conducted to obtain two sets of data. The first set of data consisted of a short sampling period of 0 - 50 minutes with sampling intervals of 5 minutes and a longer sampling period of 0 - 10 hours with 1 hour sampling intervals conducted at three levels of temperature (32.2, 40 and 55 °C) and one level of binary osmotic solution of sucrose concentration (70%). The second set of data is a factorial design at three levels of temperature (32.2, 40 and 55 °C) and three levels of sucrose concentration solution (40, 50 and 60%) conducted at a sampling period of 0 - 3 hours with 30 minutes sampling intervals and agitation of 79 Reynolds number [31,32].

Water loss (WL) and solid gain (SG) were the mass transfer measurements and expressed in gram in relation to initial mass of sample. Process factors varied were: temperature, T (°C); concentration of osmotic solution, C (%) (sucrose); time of immersion, t (min); sample type, St; sample size, Ss (cm³) and agitation (Ag).

Determination of water loss

Apple varieties (Golden Delicious and Cox), banana and potato were peeled and cut into cylindrical segments (20.0 mm length, 12.0 mm diameter) using a metallic cork borer. Osmotic solution of sucrose concentration was 40, 50, 60 and 70 % prepared in distilled water. Ascorbic acid (2 %) was added to the osmotic solution as an anti-browning solution. From the stock osmotic solution, 6 ml was pipetted into a 30 ml Pyrex bottle and the samples were transferred into the osmotic solution. Three replicates were prepared for each treatment. The bottles were transferred into a temperature-controlled water bath at 32.2, 40 and 55 C (Grant SS40-D Shaking Bath, Grant Instruments (Cambridge) Ltd., Cambridge, England).

At each interval recording, the dehydrated samples were blotted between two filter papers to remove surface solution, weighed using Sartorius (Portable PT 600), Goettingen, Germany and transferred into a pre-weighed stainless steel dish which was then placed in an oven (Gallenkamp Hotbox Oven, A. Gallenkamp & Co. Ltd, London, England) to dry until constant weight at 60 °C. The total moisture content was determined by placing the sample in an oven at 60 °C until constant weight and the soluble solids content of the plant materials were measured using a refractometer (ATAGO, 0-90%, Tokyo, Japan) at 20 °C.

Water loss and solid gain

Based on similar studies the initial mass differences between samples were accounted for by expressing the water loss (WL) and solid gain (SG) in g per g initial mass (g/g) as reported [33-35]. Calculations of the amount of water loss and solid gain and their rates using the gravimetric method were based on the following relations:

\[
\text{Water loss (WL) in relation to initial fresh mass of sample (g/g)} = \frac{((m_o - m) + (s - s_o))}{m_o} \times \text{time} \quad \text{...Equation (1)}
\]

\[
\text{Rate of water loss [g / g (min)]} = \frac{((m_o - m) + (s - s_o))}{m_o} \times \text{time} \quad \text{...Equation (2)}
\]

\[
\text{Solid gain (SG) in relation to initial fresh mass of sample (g/g)} = \frac{(s - s_o)}{s_o} \times \text{time} \quad \text{...Equation (3)}
\]

\[
\text{Rate of solid gain [g / g (min)]} = \frac{(s - s_o)}{s_o} \times \text{time} \quad \text{...Equation (4)}
\]

where \(m_o\), \(m\) are the initial mass and mass after sample period, \(s_o\), \(s\) are the initial mass of solids and mass of solid after sample period respectively and \(t\) = duration of osmo-dehydration treatment i.e. the sampling period.

Model data treatment

Experimental data collection from gravimetric studies on apples, banana and potato were used in a single model development. The gravimetric data used for the predictive model consisted of two sets of data. It is important to note that this is possible because the same sample size and geometry was used for each tissue type.

The two sets of data collected for each studied material were combined to form a single database of 1,250 experimental points which was subsequently used as the multiple linear regression databases. The database is the raw data of the individual studied material performed in triplicate. The total experimental data set (1,250) was subsequently divided into two equal groups (Group A and B) of 625 each. Group A data set (625) was used to train the model and assess the model...
adequacy and the least significant variables were removed from the model. Group B data set (625) was used to test the model to improve parameter estimation. The two groups were formed by numbering the total data sets systematically as 1 and 2. Every data row assigned 1 was selected and removed from the total data sets to form group A. The remaining rows assigned 2 were grouped to form group B.

Best fitted model development

Using multiple linear regressions, a single model for all materials water loss (WL) and solid gain (SG) is represented as follows:

\[ WL = W_L c + b_1T + b_2C + b_3t + b_4St + b_5Ss + b_Ag \]  \[ \text{Equation (5)} \]

\[ SG = SG c + b_1T + b_2C + b_3t + b_4St + b_5Ss + b_Ag \]  \[ \text{Equation (6)} \]

where, \( b_1, b_2, b_3, b_4 \) and \( b_5 \) = partial regression coefficient, \( W_L \) = water loss constant, \( SG \) = solid gain constant, \( T \) (°C) = temperature of osmotic dehydration, \( C \) (%) = concentration of osmotic solution, \( t \) (min) = time of immersion in osmotic solution, \( Ss \) (cm³) = sample size, \( St \) = sample type and \( Ag \) = agitation number (Re).

Single model development

Dummy variables were used to estimate one model for all the different plant material type. This enables fitting one model in which ‘type of sample’ is another variable in the model. This was done by defining the sample variables as three dummy variables: \( D_{ij} \), \( D_{ij} \), and \( D_{ij} \):

\[ D_{ij} = 0 \quad \text{if a Golden Delicious} \]

Define \( D_{ij} = 0 \), \( D_{ij} = 1 \) and \( D_{ij} = 0 \) if the fruit is a Cox apple

Define \( D_{ij} = 0 \), \( D_{ij} = 0 \) and \( D_{ij} = 1 \) if the fruit is a potato

Taking note that \( D_{ij} = 1 \), \( D_{ij} = 1 \) and \( D_{ij} = 1 \) should not occur

From the estimated model of the type:

\[ Y_i = a + b_1X_{1i} + b_2X_{2i} + \ldots \]  \[ \text{Equation (7)} \]

Fitting the model is given as:

\[ Y_i = a + b_1X_{1i} + b_2X_{2i} + \ldots + \delta_2D_{2i} + \delta_3D_{3i} + \delta_4D_{4i} \]  \[ \text{Equation (8)} \]

When \( D_{ij} = 1 \), \( D_{ij} = 0 \) and \( D_{ij} = 0 \) the model is

\[ Y_i = a + b_1X_{1i} + b_2X_{2i} + \ldots + \delta_2D_{2i} + (a + \delta_2) + b_4X_{4i} + b_5X_{5i} \]  \[ \text{Equation (9)} \]

When \( D_{ij} = 0 \), \( D_{ij} = 1 \) and \( D_{ij} = 0 \) the model is

\[ Y_i = a + b_1X_{1i} + b_2X_{2i} + \ldots + \delta_3D_{3i} + (a + \delta_3) + b_4X_{4i} + b_5X_{5i} \]  \[ \text{Equation (10)} \]

When \( D_{ij} = 0 \), \( D_{ij} = 0 \) and \( D_{ij} = 1 \) the model is

\[ Y_i = a + b_1X_{1i} + b_2X_{2i} + \ldots + \delta_4D_{4i} + (a + \delta_4) + b_4X_{4i} + b_5X_{5i} \]  \[ \text{Equation (11)} \]

When \( D_{ij} = 0 \), \( D_{ij} = 0 \) and \( D_{ij} = 0 \) the model is

\[ Y_i = a + b_1X_{1i} + b_2X_{2i} + \ldots + a + b_4X_{4i} + b_5X_{5i} \]  \[ \text{Equation (12)} \]

Equations (9), (10), (11) and (12) hold for Golden Delicious, Cox, banana and potato, respectively.

There are advantages of combining the data from all the commodities as this gives the b coefficients estimated from a larger sample size and a more effective model to describe the relationship. Notable difference between the four models above is the constant term that is different in each model. This shows a commodity effect.

Data analysis

Model building and parameter estimation were performed by multiple regressions using the software SPSS 10.0 (SPSS Science, Chicago, Ill., U.S.A). Data for the model building was transformed by logarithms (log x) to compress the dynamic range of the variables and thereby ensure that the relative accuracy is maintained even when some of the quantities have small values. ANOVA on water loss, solid gain, model parameters and rate constants were conducted. Mean separation was done using Least Significant Difference (LSD).

Results and Discussion

Model development

Different models development had been authored [36] including multiple-input and multiple-output systems [36-39]. Development of a single model for the water loss and solid gain based on equations (5) and (6) was conducted using data from all the plant materials studied to give more description of the model for the independent and the dependent relationship.

Independent variables of the proposed model were temperature of osmotic dehydration, \( T \) (°C); concentration of osmotic solution, \( C \) (%); duration of immersion in osmotic solution, \( t \) (min); sample size, \( Ss \) (cm³) and agitation level, \( Ag \) (Re). Dependent variables were water loss (WL) and solid gain (SG).

Best-fitted model for describing the experimental data are given as:

\[ \text{Log WL}_se = -1.807 + 0.469\text{log T} + 0.586\text{log C} + 0.148\text{log T} - 0.211\text{log Ss} + 1.534 \times 10^{-3}\text{log St} \]  \[ \text{Equation (13)} \]

\[ \text{Log SG}_se = -2.110 + 0.348\text{log T} - 3.528 \times 10^{-3}\text{log C} + 0.227\text{log T} - 1.994 \times 10^{-3}\text{log Ss} - 0.182\text{log St} \]  \[ \text{Equation (14)} \]

where \( \text{WL}_se \) = water loss in single model, \( \text{SG}_se \) = solid gain in single model.

Table 1 presents the parameter estimates and their standard errors for water loss and solid gain model showing a variation of -0.303 between the model constants of water loss and solid gain. The regression correlation coefficient \( r \) and the regression coefficient of determination \( R^2 \) for water loss and solid gain models are presented in Table 2. The water loss model shows a regression coefficient of determination \( R^2 = 0.886 \) indicating a good correlation coefficient \( r = 0.941 \) between experimental and predicted data (Table 2). The good prediction of the water loss model is illustrated in Figure 1 and 2.

Table 1 shows that training and testing for water loss model had similar determination (\( R^2 \)) indicating a good correlation coefficient (\( r \)) and the regression coefficient of determination (\( R^2 \)).
determination ($R^2$), which indicates a good predictability of the model. The model plots for water loss and solid gain during training shows the distribution of the experimental points to be linear for water loss (Figure 1) and non-linear for solid gain (Figure 3). Residual plots of water loss during training portray even distribution of experimental points (Figure 2) whereas that for solid gain showed uneven distribution of experimental points (Figure 4).

### Water loss and solids gain

In similar studies the water loss and solid gain behaviour was attributed to the morphological structures particularly the intercellular spaces present in the plant tissue and the starch content of the materials involved in this case a cellulose material represented by apple, starchy material represented by banana and potato [5,31,40]. As a result of the more compact nature of the intercellular spaces of tissues of banana and potato more than apple, there is retarding of diffusion of water from banana and potato [5,6,31,40]. The coefficient of determination ($R^2$) for the training and testing of the water loss model are 0.891 and 0.886, respectively (Table 2). The regression coefficient of determination ($R^2$) for the solid gain model does not show a good regression correlation coefficient ($r = 0.552$) between the experimental data and the predicted data. This is probably due to the variability of the pathways of solid diffusion into the different plant materials [31,40]. However it may well be improved by the introduction of one or more additional variables for its determination. The coefficient of determination ($R^2$) for the training and testing of the solid gain model was 0.336 and 0.305, respectively. In the case of the solid gain model there was much departure from experimental data (Figure 3) and its residuals illustrate the uneven distribution of the experimental data (Figure 4). Therefore the solid gain model was not good in predicting the experimental data. The high standard error (±0.210) of the solid gain constant (SGc)

### Table 2: Model summary of the water loss and solid gain for the model described in equations (13) and (14).

<table>
<thead>
<tr>
<th>Model</th>
<th>R</th>
<th>R Square</th>
<th>Adjusted R Square</th>
<th>Std. Error of the Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>WL se</td>
<td>Training</td>
<td>0.944</td>
<td>0.981</td>
<td>0.890</td>
</tr>
<tr>
<td></td>
<td>Testing</td>
<td>0.941</td>
<td>0.886</td>
<td>0.880</td>
</tr>
<tr>
<td>SG se</td>
<td>Training</td>
<td>0.619</td>
<td>0.383</td>
<td>0.378</td>
</tr>
<tr>
<td></td>
<td>Testing</td>
<td>0.552</td>
<td>0.305</td>
<td>0.301</td>
</tr>
</tbody>
</table>

* : water loss in single model; : solid gain in single model ; : not available

### Figures

- **Figure 1**: Multiple Linear Regression (MLR) model training for predicted against experimental water loss.
- **Figure 2**: Multiple Linear Regression (MLR) model training for residual against experimental water loss.
- **Figure 3**: Multiple Linear Regression (MLR) model training for predicted against experimental solid gain.
- **Figure 4**: Multiple Linear Regression (MLR) model training for residual against experimental solid gain.
attests to this fact indicating the need for additional variable(s) for its prediction (Table 1).

Model behaviour

The uneven behaviour of the solids gain models was equated to the high percentages of water loss than solid gain in each material contributing to the different behaviour of the models [5,9,21,36,37,40,41]. In similar reports, the model behaviour was explained that the sucrose passes through the cell wall and accumulates between the cell wall and the cell membrane forming a hypertonic solution resulting in water diffusion out through the cell membrane. However, as a result of the large molecular weight (342 g/mol) the sucrose was unable to penetrate and accumulate in the intracellular volume resulting in lower sucrose diffusion into the entire tissue. In addition, the diffusion coefficient of water and solute in the various tissues contributed to the difference between water loss and solid gain in essences the water loss and solid gain behaviour was attributed to the morphological structures particularly the intercellular spaces present in the plant tissue and the starch content of the materials under study [2-4,8,19].

Model analysis

Statistical analyses (ANOVA) of the water loss and solid gain models are presented in Table 3, whereas the correlation matrix of the model variables is presented in Table 4. The sum of squares for water loss was 6.361 with residual of 0.781, whereas that of solid gain was 12.906 and its residual was 20.823. The correlation matrix was significant at 0.01 for variables for water loss, solid gain, temperature, concentration, time, sample type and sample size. However, for correlation matrix significant at 0.05 was for variables of sample size with temperature and time, sample type and sample size. The sum of squares for water loss model (R² = 0.886, r = 0.941) was adequate than solid gain model (R² = 0.305, r = 0.552), which was due to the variability of the pathways of solid diffusion into the different plant materials under study. In such situations, introduction of one or more additional variables in the solid gain model may improve its determination. Literature was limited on single model developed from different plant materials in which sample type was incorporated as a variable. Therefore, future studies was required using other plant materials on single models development.

**Conclusion**

Predictions of water loss model (R² = 0.886, r = 0.941) was adequate than solid gain model (R² = 0.305, r = 0.552), which was due to the variability of the pathways of solid diffusion into the different plant materials under study. In such situations, introduction of one or more additional variables in the solid gain model may improve its determination. Literature was limited on single model developed from different plant materials in which sample type was incorporated as a variable. Therefore, future studies was required using other plant materials on single models development.

Acknowledgement

Sponsorship from NRI-HEFCE funds is gratefully acknowledged.

References


