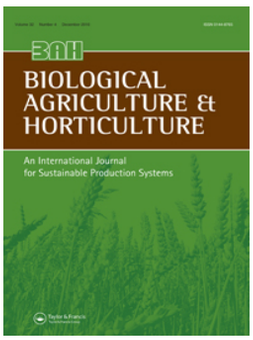

Taylor & Francis
Taylor & Francis Group

Biological Agriculture & Horticulture

An International Journal for Sustainable Production Systems

ISSN: 0144-8765 (Print) 2165-0616 (Online) Journal homepage: https://www.tandfonline.com/loi/tbah20

Application of logistic regression statistical technique to evaluate tomato quality subjected to different pre- and post-harvest treatments

Sileshi Melesse, Nafiisa Sobratee & Tilahun Workneh

To cite this article: Sileshi Melesse, Nafiisa Sobratee & Tilahun Workneh (2016) Application of logistic regression statistical technique to evaluate tomato quality subjected to different pre- and post-harvest treatments, Biological Agriculture & Horticulture, 32:4, 277-287, DOI: 10.1080/01448765.2016.1184587

To link to this article: https://doi.org/10.1080/01448765.2016.1184587

BIOLOGICAL AGRICULTURE & HORTICULTURE, 2016
VOL. 32, NO. 4, 277–287
http://dx.doi.org/10.1080/01448765.2016.1184587

Taylor & Francis
Taylor & Francis Group

Application of logistic regression statistical technique to evaluate tomato quality subjected to different pre- and post-harvest treatments

Sileshi Melesse[a], Nafiisa Sobratee[b§] and Tilahun Workneh[b]

[a]School of Mathematics, Statistics and Computer Science, University of KwaZulu-Natal, Pietermaritzburg, South Africa; [b]Department of Bioresources Engineering, School of Engineering, University of KwaZulu-Natal, Pietermaritzburg, South Africa

ABSTRACT

The study applied a combination of categorical and continuous explanatory variables via logistic regressions to model the probability of quality, defined as marketability, of tomato (*Solanum lycopersicum* L.) fruits subjected to different pre-harvest (biocatalyst application) and post-harvest treatments: (i) dipping in tap water, chlorinated water or anolyte water (ii) thin film packaging versus no packaging, and (iii) storage at 13 °C or in ambient temperature. Fungi, bacteria, coliforms, total soluble solids (TSS), glucose, fructose and ascorbic acid were determined in a 30-day full factorial experiment. The effects of each quality variable on the marketability of tomato fruits were explored. A multiple logistic regression model, consisting of both the continuous and the categorical variables, was then derived to evaluate the effectiveness of the anolyte water as dipping treatment. The simple logistic models showed that storage time, fungi, aerobic bacteria and coliform population are negatively related to the probability of marketability. In contrast, TSS, glucose and fructose are positively related to the probability of marketability. No significant relationship was observed between the probability of marketability and ascorbic acid content or pre-harvest biocatalyst application. The multiple logistic regression model showed that the probability of marketability was higher when the tomatoes were subjected to anolyte dipping treatment, packaging and storage at 13 °C. The study presented new insights into the interpretation of post-harvest quality defined as probability of marketability of the stored fresh produce.

ARTICLE HISTORY
Received 6 May 2015
Accepted 27 April 2016

KEYWORDS
Categorical variables; continuous variables; logistic regression; probability of marketability; tomato quality

Introduction

Maintaining shelf life through integrated agro-technological unit operations

Tomato is one of the most popular and versatile vegetables worldwide and the maintenance of its shelf life quality is an important concern. ComCat®, Communication Catalyzation, is a hormone containing treatment that has been introduced as an alternative agricultural input to the use of chemicals to increase production of vegetables and other crops (Pretorius et al. 2008). It is a natural biocatalyst,

CONTACT Nafiisa Sobratee nsobratee@yahoo.com
[§]Graduate School of Business and Leadership, University of KwaZulu Natal, Durban, South Africa.

which is extracted from seeds of plants and mainly consists of amino acids, gibberellin, cytokinins, auxin (indole-3-acetic acid), brassinosteroids, natural metabolites, pathogen-related PR-proteins with defence reactions, terpenoids, flavonoids, vitamins, inhibitors, other signal molecules, biocatalysts and cofactors (Pretorius & van der Watt 2011). Earlier, Schnabl et al. (2001) showed that there were incremental yields of ComCat® treated vegetables as demonstrated in cabbage (8%), tomato (16–19%), potato (9–19%), soyabeans (26–30%), eggplants (37%), cucumbers (25–32%), carrots (32%), onions (49%) and strawberries (50%). These vegetables were also shown to have better root development, improved resistance induction, less chance of deficiencies with fertilizer, higher resistance to pathogens prior to harvest and they seem to have a slightly better resistance to environmental stress, and increased protein content (Schnabl et al. 2001; Hüster 2011).

The advantages of ComCat® treatment are that only low doses are necessary to show measurable effects of these brassinosteroid-containing plant extracts in crop plants, their environmental safety and the possibility of reducing the amount of pesticides needed (Schnabl et al. 2001; AgraForUm 2010). Apart from studies on yield increase, little data is available on the effect of pre-harvest ComCat® treatment on quality of fruits and vegetables at harvest, as well as during storage. ComCat® was approved by, and registered with, the Federal German Biological Centre of Agriculture and Forestry (BBA), Institute for Integrated Plant Protection as a harmless plant strengthening substance of plant origin. It is also licensed for use in Ecological Farming, according to the EU-regulation 2092/91 (AgraForUm 2010).

Following harvest, the ripening process continues and tomatoes can become overripe very rapidly resulting in loss of quality and restricted shelf life (Geeson et al. 1985; Artés et al. 1999). Postharvest traits, many of which define product storability and quality, are characterised by interactions with the postharvest environment. The measurement of biochemical and physical variables such as sugars, acid, enzyme activity, microbial loads, phenolic or volatile component levels, firmness and colour measurements provide the quantitative bases for trait evaluation (Batu 2004; Balibrea et al. 2006; Luengwilai et al. 2010; Beckles 2012; Pinheiro et al. 2013).

Beckles (2012) suggested that the quality and shelf life of tomatoes can be extended by adopting multi-prong operations whereby several treatments, e.g. modified atmosphere, low temperature and ionisation radiation are used in succession or in combination. Anolyte water is electrochemically-activated water prepared from an aqueous solution of NaCl and is known to be a powerful, non-toxic, non-hazardous disinfectant. Two kinds of electrochemically activated water are produced, anolyte and catholyte, each having different physico-chemical properties (Prilutsky & Bakhir 1997). The anolyte water is described as having an oxidation-reduction potential (ORP) in the region of +1000 mV and catholyte an ORP of −800 mV, and the pH value of catholyte is in the alkaline region while the pH value of anolyte is in the acidic region. Some of the biocidal agents (free radicals) in the solutions are ClO_2, HClO, Cl_2, ClO^-, H_2O_2, HO_2^-, NaOH, O_2, O_3, H^+ and OH^-. Anolyte is thought to have the antimicrobial effect and catholyte a detergent or cleaning effect (Popova et al. 1999). The presence of the free radicals is considered of great importance with their oxidising effects in the solutions. It is the free radicals (working substances) that confer anolyte water its bactericidal and sporicidal activity. This is because higher organisms possess antioxidant defence systems whereas microorganisms generally do not and are therefore destroyed. Anolyte water attracts interest because it is an environmentally and ecologically friendly substance for use as a post-harvest fruit and vegetables disinfectant. Based on the antimicrobial action of anolyte water, its potential to be used as a postharvest disinfectant of fruit and vegetables therefore warrants exploration.

An understanding of the effect on quality of each post-harvest operation administered singly and in combination is necessary in order understand the complex interactions. Another aspect to consider in post-harvest research is the application of analytical tools to enable proactive interpretation of the research output. In the context of post-harvest research, the main rationale consists of improving shelf life and hence the marketability of the fresh produce. In the last decade, studies have focused on predicting quality and marketability of fresh agricultural produce by using non-destructive hyperspectral imaging techniques (Lu 2004; Nicolaï et al. 2007; Mendoza et al. 2011) whereby analytical techniques such as partial least squares have been applied for calibration and result interpretation.

However, there is a lack of satisfactory research that can directly relate the trends and effectiveness of parameter variables in predicting the quality and therefore, the marketability of fresh produce following integrated agro-technology unit operations. Evaluation of changes in chemical, biochemical and microbiological qualities of fresh produce requires use of a wide range of instrumentation, quality determination procedures, technical knowledge and laboratory inputs. Moreover, numerous quality parameters determined for the same sample are commonly subjected to analysis of variance to determine the effects of individual treatment and their interactions on the determined quality attributes. Based on this analysis, suitable regressions models could be developed that could allow prediction of marketability as shelf life of fresh produce. These could be used to develop a tool for farmers or processors to predict changes in quality of produces during storage or transportation. Consequently, several techniques have been developed for analysing postharvest fruit and vegetables quality data with categorical dependent variables including discriminant analysis, log-linear regression and logistic regression. However, each of these techniques requires different assumptions to be applicable (Agresti 1996; Kleinbaum & Klein 2010). For instance, log-linear regression requires all the explanatory fresh produce variables to be categorical variables. In contrast, discriminant analysis requires all the explanatory fresh produce quality variables to be continuous. Logistic regression can handle both categorical and continuous explanatory post-harvest produce quality variables.

In the present study, logistic regression analysis was applied to predict the probability of marketability of tomatoes based on biochemical and microbiological parameters, and pre-harvest and post-harvest operations. The first objective was to determine the relationship of individual parameter variables with respect to the probability of marketability. In the present context, 'marketability' is a term chosen to define quality of the tomatoes subjected to experimental conditions since the marketability or fitness for purpose is a function of the changes that have occurred in the quality as detected by biochemical, physical and microbiological parameters during post-harvest storage. The second objective was to apply a multiple logistic regression model to determine the factor(s) that can most accurately predict the probability of marketability.

Theoretical background to logistic regression analysis

The logistic regression model is designed to describe a probability, which is always some number between zero and one. In the present study, such a probability refers to the chance of marketability of a given tomato sample. Logistic regression is suitable for the study of the relation between a categorical or binary response variable and one or more predictor variables. In the simplest case of one predictor X, for instance, days of storage (DOS), and one dichotomous outcome variable Y, for instance, marketability of the tomato sample, the logistic model predicts the logit of Y from the independent variable X. The logit is the natural logarithm (ln) of odds of Y.

The simple logistic model has the form:

$$\text{logit}\,(\pi(x)) = \ln\left(\frac{\pi(x)}{1-\pi(x)}\right) = \ln\,(\text{odds}) = \alpha + \beta\, x \tag{1}$$

An alternative formula for the simple logistic regression model refers directly to the probability of the outcome of interest. It is given by:

$$\pi(x) = p(Y \text{ is outcome of interest}|\, X = x) = \frac{e^{\alpha+\beta x}}{1+e^{\alpha+\beta x}} \tag{2}$$

where $\pi(x)$ = probability that a given tomato sample will be marketable when the DOS (X) takes value x, $\alpha = Y$ – intercept, β = slope parameter, X can be categorical or continuous variable.

Model (1) shows that the logistic regression model has linear form for the logit of the probability. The simple logistic regression model in Equation (1) can be extended to multiple predictors in a multiple logistic regression analysis. This relates to a single dichotomous outcome with more than one independent variable. A multiple logistic regression may be constructed as follows:

$$\text{logit}\,(\pi(X)) = \ln\left(\frac{\pi(X)}{1-\pi(X)}\right) = \alpha + \sum_{i=1}^{k}\beta_i x_i \tag{3}$$

$$\pi(x) = \frac{e^{\left(\alpha+\sum_{i=1}^{k}\beta_i x_i\right)}}{1+e^{\left(\alpha+\sum_{i=1}^{k}\beta_i x_i\right)}} \tag{4}$$

where $\pi(x)$ = the probability of the event of interest i.e. probability of tomato marketability, α = the intercept, βs = the slope parameters, Xs = a set of predictors.

The goal of the logistic regression is to estimate the unknown parameters of model (3). This is achieved by applying the maximum likelihood (ML) estimation method. It identifies the values of the parameters for which the probability of the observed data is maximum. ML computations for fitting logistic regression models are complex, but are easy to perform using statistical software (Agresti 1996).

Materials and methods

Experimental material and metabolic parameters

Tomato production

Tomatoes (cv. Marglobe), grown in the area of Bloemfontein South Africa were used. The study site is located between the latitudes 26.6 and 30.7°S and between longitudes 24.3 and 29.8°E. The minimum and maximum mean annual temperatures were 22 and 26 °C. The sample fruit were produced in an open field developed for commercial production of tomato. The tomato seedlings were raised in seedbeds and then transplanted to the field during the fourth week after sowing when they were 6.7 cm tall. No substrate treatments and no fertilizer or pesticide applications were made during the growth of the tomato plants. Drip irrigation was used to supply water to each plant during the growth period.

Biocatalyst application

During the vegetative phase, tomato plants were treated with 10 g ha^{-1} biocatalyst (ComCat®) in 350 l of water with control plants receiving 0 g ha^{-1}. Spraying was performed once at the three-leaf stage and again at the full vegetative stage. At the green mature stage, biocatalyst treated and untreated tomatoes were harvested manually from four randomized blocks early in the morning and immediately transported to the laboratory for screening and selection of blemish and defect-free tomatoes test samples. The green mature tomato maturity stage is the most standardised commercial ripeness stage for harvesting tomato fruit. In order to avoid cross contamination, all working surfaces and tools were washed and disinfected prior to use with 1% Chlorobac (Syndachem, Pty Ltd, Isando, South Africa). The selected tomatoes were washed with water at 4 °C to remove field heat and soil particles and reduce surface microbial populations. Figure 1 provides the experimental layout for the pre-harvest and post-harvest conditions to which the test tomatoes were subjected.

Postharvest treatment

Washed, biocatalyst-treated tomatoes (180 kg) were subdivided into three equal groups for dipping treatments in chlorinated water, anolyte water and tap water, the control treatment. Plastic containers were used to avoid losses of charged ions from the anolyte water. Chlorinated water, adjusted to pH 7.2, containing 100 mg ml^{-1} free chlorine was prepared from 5% sodium hypochlorite and used for a 20 min dipping time. The electrochemically-active anolyte water was prepared according to the method of Workneh et al. (2003) and used for a 5 min dipping time. Water that contained 5% NaCl was ionised with an ioniser (Radical Waters Pty Ltd, Johannesburg, South Africa) operating at a pressure of 50 Pa and adjusted to pH 6.1. The concentration of total dissolved solids after ionisation was 3.55%. As a control, tomatoes were dipped for 20 min in tap water.

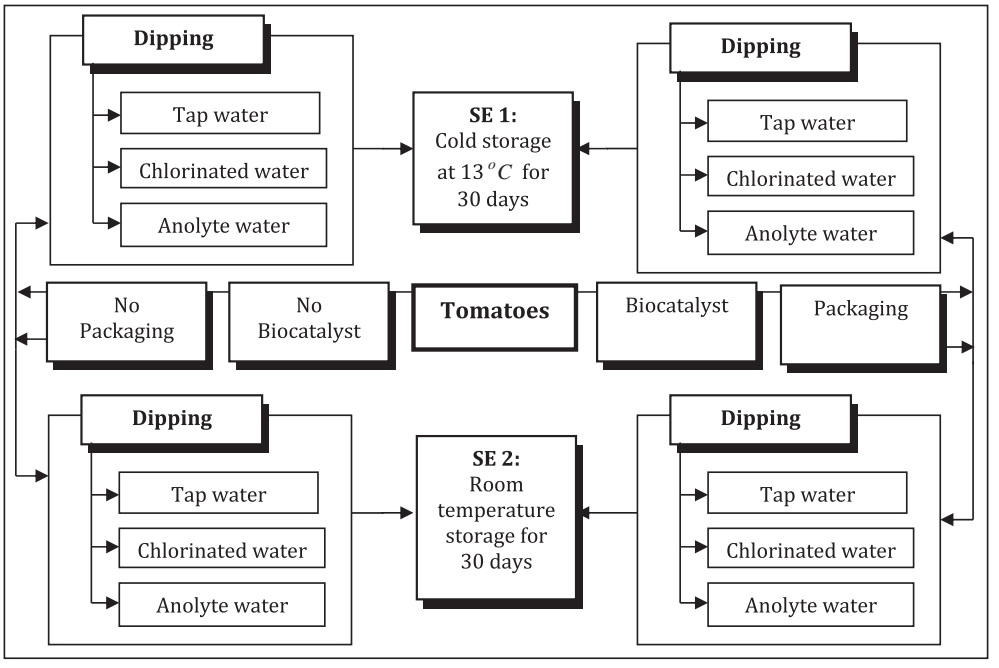

Figure 1. Experimental design comprising of three types of water used for post-harvest washing of carrots prior to storage at 1° C or room temperature, averaging 22° C, for 30 days.
Note: EU: experimental unit.

Packaging

Tomatoes were subdivided into 1 kg samples and packed in commercial micro-perforated bags (XtendR Film, Patent No. 6190710, StePac L.A., Ltd, Pretoria, South Africa), which are specifically designed for 1 kg tomato packaging and for storage at 13 °C. The permeability of the bags for O_2 and CO_2 was 9000 and 55000 $cm^3\ m^{-2}\ day^{-1}$ at 1 atm, respectively, with 17.2 g $m^{-2}\ day^{-1}$ H_2O vapour transmission rate. Unpackaged 1 kg samples for each treatment combination were placed on perforated plastic bags and left open. Tomatoes were stored at 13 °C and RH (relative humidity) of 34–76%, and at room temperature (16.9–25.2 °C) and RH 34–76%. On each sampling date, triplicate 1 kg packages of tomato were randomly taken from each treatment for quality analyses.

Biochemical analysis

Clear tomato juice was prepared according to the method of Nunes and Emond (1999) and used for all further analyses. The total soluble solid (TSS) was determined using the procedures described by Tefera et al. (2007) with an Atago N1 hand refractometer, (Bellevue, WA, USA) with a range of 0–32 °Brix, and resolution of 0.2 °Brix. Sucrose, glucose and fructose were determined by the method of Riaz and Bushway (1996). HPLC was carried out on a Waters system (501 pump) and Biorad Aminex column (Bio-Rad Laboratories, St. Albans, UK (7.8 × 300 mm) with a differential refractive index detector (R401) (Midland, ON, Canada) operated at 42 °C and a mobile phase of de-ionised water at a flow speed of 0.6 ml min^{-1} and temperature of 85 °C. The ascorbic acid content was determined by the 2,6–dichloro-phenolindophenol method (AOAC 1970) on tomato juice following extraction in 3% metaphosphoric acid. Titration was performed with standard dye to a pink end-point persisting for 15 s following centrifugation at 10,000 × *g* for 15 min. The ascorbic acid content (%) was obtained from the titration value, dye factor, dilution and volume of the test sample (10 ml).

Microbiological analysis

Microbial populations were estimated by the poured plate methods of Brackett (1990). Total aerobic microorganisms were determined on plate count agar (Oxoid CM463), *Eschericia coli* and coliform population on violet red bile agar (VBRA with MUG, Oxoid CM978) and moulds and yeasts on Rose-Bengal chloramphenicol agar base (Oxoid CM549).

Results and discussion

Probability of marketability based on continuous variables

Table 1 shows the results obtained from fitting simple logistic regression models to the different continuous variables. A detailed explanation of the output generated is provided for the parameter DOS. The fitted model considers DOS as a variable that affects the probability of marketability for tomato ($p < 0.001$), where probability of marketability refers to the overall post-harvest quality changes which occurred during the particular experimental conditions provided, since the quality or fitness for purpose of the commodity after storage is a function of the pre-harvest and post-harvest conditions the fruits were subjected to. There was a negative relationship between marketability and DOS. The probability of marketability as a function of DOS is given by:

$$\pi(x) = \frac{\text{Exp}^{3.121-0.203\,x}}{1+\text{Exp}^{3.121-0.203x}} \tag{5}$$

where $\pi(x)$ = the probability of marketability when the DOS assumes a particular value equal to x.

The fitted model predicts the probability that a particular tomato sample can be marketable at a particular value of DOS (Figure 2). For instance, when date of storage is 0, the probability of marketability is 0.9578. However, when the date of storage is 7, the probability of marketability of a given tomato sample is 0.8455. The chance of marketability decreases as the time of storage increases. The Exp(B) column shows the odds ratio. The odds ratio of 0.816 for DOS indicates that for one day increase in storage, the odds of marketability decreases by $(1 - 0.8116) \times 100 = 18.4\%$.

The output for the other biochemical parameters is shown in Table 1. The positive coefficients for glucose, fructose and TSS indicate their respective positive relationship with the probability of marketability. The corresponding relationship between bacteria, fungi and coliforms, and the chance of marketability is negative indicating that marketability decreases as these microbial population levels increase. The β-coefficients obtained for ascorbic acid content ($p = 0.807$) and sucrose ($p = 0.172$) were not significantly different from zero. This indicates that there was no evidence that these two parameters affected marketability of tomato.

Table 1. Output for simple logistic model using continuous variables.

				95% CI for Exp(B)	
Variables	Regression coefficient	Wald χ^2	Exp(B)	Lower	Upper
Days of storage	−0.203	70.268	0.816	0.779	0.856
Constant	3.121	59.052	22.666		
Fungi	−0.821	48.167	0.440	0.349	0.555
Constant	1.989	41.571	7.306		
Bacteria	−0.586	40.619	0.556	0.465	0.666
Constant	2.676	39.179	14.530		
Coliform	−0.556	45.417	0.574	0.488	0.674
Constant	2.327	41.999	10.244		
Glucose	5.025	22.742	152.191	19.295	1200.424
Constant	−6.396	21.504	0.002		
Fructose	7.975	56.457	2908.287	363.201	23287.75
Constant	−8.637	54.973	0		
TSS	1.861	9.120	6.428	1.922	21.505
Constant	−7.969	8.782	0		

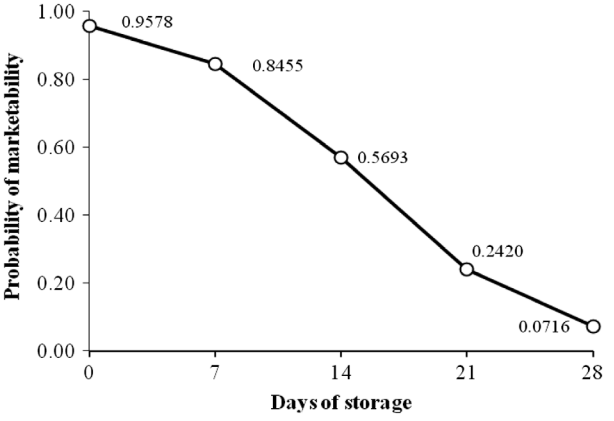

Figure 2. Predicted probability of marketability of tomato based on date of storage.

Probability of marketability based on continuous and categorical variables

Initially, DOS and storage environment (SE) were used to assess their combined effects on marketability. With regard to SE, there is a statistically significant differences between SE1 (cold temperature storage) and SE2 (room temperature storage) ($p < 0.001$) (Figure 3). The estimated model for SE1 is:

$$\hat{p} = \frac{e^{5.48-0.27\,\mathrm{DOS}}}{1 + e^{5.48-0.27\,\mathrm{DOS}}} \tag{6}$$

Table 2 displays the output for a simple logistic model using categorical variables. Here SE2 is used as reference category for the logistic regression .That is the reason why the information for SE2 cannot be seen in Table 2. For room temperature storage (SE2), the estimated model is given by:

$$\hat{p} = \frac{e^{2.697-0.27\,\mathrm{DOS}}}{1 + e^{2.697-0.27\,\mathrm{DOS}}} \tag{7}$$

The same procedure was performed in order to assess the chance of marketability by using the pre-harvest biocatalyst treatment as a categorical variable. No significant relationship was observed between biocatalyst treatment and marketability ($p = 0.365$). This categorical variable is therefore not suitable for predicting tomato marketability under the conditions studied.

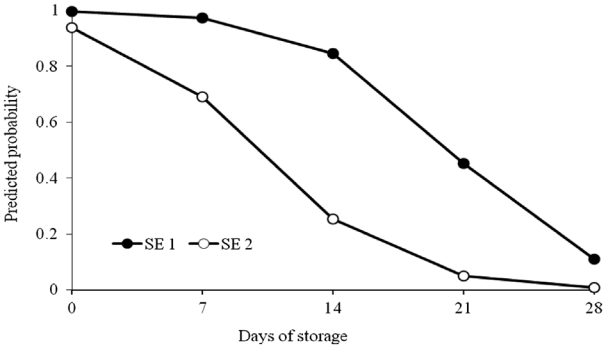

Figure 3. Predicted probabilities for storage environment based on the days of storage.
Note: SE 1 = 13° C, SE 2 = 16.9 – 25.2° C.

Table 2. Output for simple logistic model using categorical variables.

Variables	Regression coefficient	Wald χ^2	Exp(*B*)	95% CI for Exp(*B*)	
				Lower	Upper
Constant	2.697	35.85			
SE 1	2.78	29.75	16.16	5.95	43.91
Days of storage	−0.266	63.17	0.77	0.72	0.82

Fitting a multiple logistic model based on all variables

When all the variables (DOS, sucrose, glucose, ascorbic acid, glucose, fructose, TSSs, fungi, aerobic bacteria, coliforms, pre-harvest biocatalyst application, dipping treatment and SE) were subjected to multiple logistic regression analysis, only a few variables were found to be statistically significant in the ability to predict the probability of marketability of the test tomatoes. Among the continuous variables, only DOS and, among the categorical variables, SE and dipping treatment were significant independent variables that can predict the chance of marketability (Table 3). In this case, each odds ratio is the ratio of the odds of an event occurring in one group to the odds of the same event occurring in another group. For instance, holding SE, and disinfection and packaging (DP) at fixed value, say SE = 1 and DP = 1, will show a 32% decrease in the odds of marketability for one day increase in storage (Table 3). Once again, referring to Table 3, the categories SE = 2 and DP = 4 are not included in the output mainly because they are used as reference categories. The odds ratios are computed with reference to these categories. For instance the odds of marketability of tomatoes for SE = 1 is 51.66 times the odds of marketability for SE = 2 keeping DOS and DP at fixed values. The odds ratio was highest for DP = 2 which consisted of the combined treatments of biocatalyst + anolyte + packaging. Here keeping SE and DOS at fixed values the odds of marketability for DP = 2 is about 115 times that of DP = 4.

For the same model (Table 3), the plot of the predicted probability versus DOS for different levels of SE and DP is indicated in Figure 4. It is clear from the figure that the predicted probability of marketability for SE = 1 and DP = 2 is higher than any other combination of SE and DP for all DOS. The minimum predicted probability is observed for SE = 2 and DP = 4 combination.

Hence, the essence of the present work lies in predicting the chance of marketability, following integrated agro-technological unit operations, through multiple logistic regression, which can extract the most significant experimental components, including continuous as well as categorical variables that influence the biochemical and microbial quality of the tomatoes.

Validation of the multiple logistic regression model

The value of the likelihood ratio test statistic for the overall significance of the logistic model is 217.72 with a $p < 0.001$ (Table 4), which implies that the overall fitted logistic model is significant. This result indicates that at least one of the parameters is significantly different from zero. The score and the Wald tests also support the overall significance the fitted logistic model ($p < 0.001$). The fitted logistic model is more effective than the null (intercept only) model. The statistical significance of individual

Table 3. Storage environment, dipping treatments with packaging and date of storage as independent variables affecting marketability.

Factors	Regression coefficient	Wald χ^2	*p*	Odds ratio	(95% CI)	
					Lower	Upper
Constant	1.11	3.76	0.052			
SE1: Storage at 13 °C	3.95	30.45	0.0001	51.66	12.73	209.73
DP1: No bio + anolyte + P	3.60	18.92	0.0001	36.79	7.25	186.69
DP2: Bio + anolyte + P	4.75	26.31	0.0001	115.11	18.77	705.92
DP3: Bio + anolyte + No P	2.71	12.53	0.0001	14.97	3.35	67.01
DOS: Days of storage	−0.38	50.23	0.0001	0.68	0.614	0.759

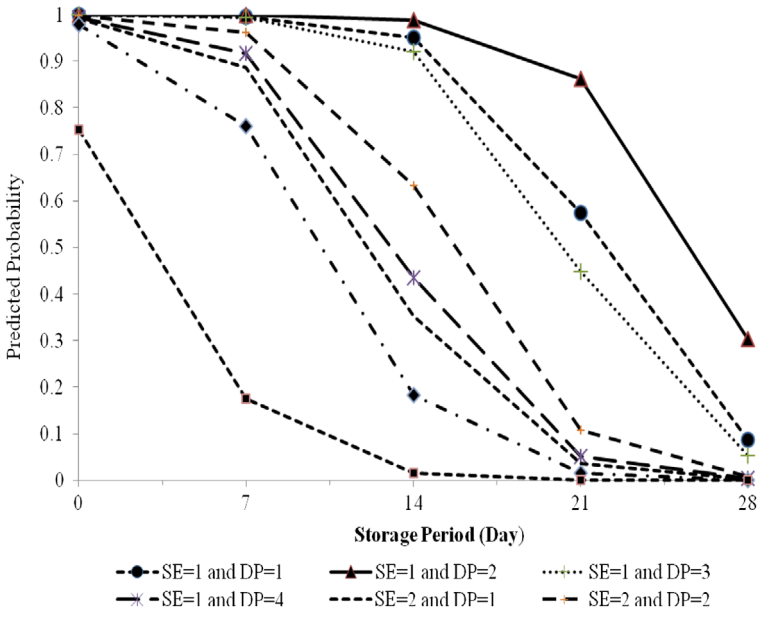

Figure 4. Predicted probability versus days of storage for different levels of SE and DP.
Notes: DP1: no biocatalyst + anolyte + packaging, DP2: biocatalyst + anolyte + packaging, DP3: biocatalyst + anolyte + no packaging, DP4: no biocatalyst + anolyte + no packaging.

regression coefficients is tested using the Wald χ^2 statistic. The Wald χ^2 statistic is the square of the ratio of the estimated slope parameter over its standard error. The individual parameter estimates for SE, DP, and DOS are all significant at the 5% level of significance (Table 3).The Hosmer and Lemeshow test for goodness of fit of this model is 5.15 with $p = 0.742$, which further indicates that the model fits the data efficiently. The Pearson χ^2 ($p = 0.635$) and deviance based ($p = 0.795$) goodness of fit tests also implied that the logistic model is sufficient to explain the data (Table 4).

Another important aspect of the fitted logistic regression that needs to be checked is the validation of predicted probabilities. The degree to which the predicted probabilities agree with the actual outcomes was expressed using a classification table with a cut-off points set at 0.5. Based on this, the prediction for marketability was almost as accurate as the prediction non-marketability. This observation was supported by the value of sensitivity (89.9%) compared with specificity (91%). Sensitivity measures the proportion of correctly classified outcome of interest (marketability), whereas specificity measures the proportion of correctly classified unwanted outcomes (not-marketable). The false positive and false negative rates were respectively 9 and 10.1%. The false positive rate measures the proportion of

Table 4. Overall model evaluation.

Model evaluation parameter	Wald χ^2	Degree of freedom	*p*
Overall significance			
Likelihood ratio test	217.72	5	<0.001
Score test	145.59	5	<0.001
Wald test	50.56	5	<0.001
Goodness of fit test			
Hosmer and Lemeshow	5.148	8	0.742
Pearson χ^2	30.61	30	0.635
Deviance χ^2	27.06	30	0.795
Association of predicted probabilities and observed response			
Somers's *D*	0.931		
Goodman Kruskal Gamma	0.937		
c-Statistic	0.966		

observations misclassified as outcome of interest (marketability) over all those classified as marketable. The false negative measures the proportion of observations misclassified as not marketable over all those classified as not marketable. Overall the per cent of cases for which the dependent variables were correctly predicted given the model is about 90.4%. The association of predicted probabilities and observed responses was assessed using measures of association. The value of Somers' *D* statistic is 0.931, the value of the Goodman Kruskal Gamma statistic is 0.937 and the value of the *c*-statistic is 0.966. The values of all three statistics are very close to 1 and support that there is strong association between the predicted probabilities and the observed responses.

Some previous work that applied logistic regression models analysed factors influencing core breakdown in Conference pears (Lammertyn et al. 2003) and the impact of storage temperature and maturity on avocado (Dixon et al. 2003). Ailes et al. (2008) applied a more integrated methodology by evaluating the effects of post-harvest processing, importation and season on the microbial concentrations of fresh produce. Thus, it was possible to identify the effect of categorical variables which can optimise the maintenance of the product quality under the studied experimental conditions. A similar approach has been applied in the current study to identify the changes in post-harvest quality during the combined pre-harvest bio-catalyst application, washing for disinfection, packaging and cold temperature storage whereby it has been found that electrochemically-activated water dipping treatment of the tomatoes followed by storage at 13 °C prior to packaging leads to the highest odds ratio of 115 ($p = 0.05$).

Conclusions

The logistic regression model is presented as a novel approach for calculating the probability of marketability of fresh produce subjected to multiple agro-technological practices, where marketability is defined as the overall changes in the quality of the tomato fruit subjected to post-harvest unit operations to maintain shelf life. Models for probability of marketability were developed for each postharvest tomato quality parameter. In a final multiple logistic regression model, only a few of the variables were found to significantly predict the probability of marketability. It was found that DP, storage temperature and DOS are the main determinants of marketability of the tomato. The predicted probability of marketability for storage at 13 °C and electrochemically activated water dipping treatment, when combined with packaging, had more influence than any other combination of SE and packaging throughout the storage period. The developed probability models can be used to predict the marketability of tomato subjected to packaging and storage conditions.

Acknowledgement

The authors wish to thank the College of Agriculture, Engineering and Sciences of the University of KwaZulu Natal for funding the research.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

AgraForUm. 2010. ComCat® plant strengthening agent. [Internet]. [cited 2010 Feb 12]. Available from: http://www.agraforum-sa.com/pages.aspx?pid=2910

Agresti A. 1996. An introduction to categorical data analysis. New York (NY): Wiley.

Ailes EC, Leon JS, Jaykus JLM, Clayton HA, Blanding S, Kleinbaum DG, Backer LC, Moe CL. 2008. Microbial concentrations on fresh produce are affected by postharvest processing, importation, and season. J Food Prot. 71:2389–2397.

[AOAC] Association of Official Agricultural Chemists. 1970. Official methods of analysis. Washington (DC): Association of Official Agricultural Chemists.
Artés F, Conesa MA, Hernández S, Gil MI. 1999. Keeping quality of fresh-cut tomato. Postharvest Biol Technol. 17:153–162.
Balibrea ME, Martínez-Andújar C, Cuartero J, Bolarín MC, Pérez-Alfocea F. 2006. The high fruit soluble sugar content in wild *Lycopersicon* species and their hybrids with cultivars depends on sucrose import during ripening rather than on sucrose metabolism. Funct Plant Biol. 33:279–288.
Batu A. 2004. Determination of acceptable firmness and colour values of tomatoes. J Food Eng. 61:471–475.
Beckles DM. 2012. Factors affecting the postharvest soluble solids and sugar content of tomato (*Solanum lycopersicum* L.) fruit. Postharvest Biol Technol. 63:129–140.
Brackett RE. 1990. Influence of modified atmosphere packaging on microflora and quality of fresh bell peppers. J Food Prot. 53:255–257.
Dixon J, Pak HA, Smith DB, Elmsly TA. 2003. New Zealand avocado fruit quality: the impact of storage temperature and maturity. Actas V Congreso Mundial del Aguacate [Proceedings V World Avocado Congress]; Bethlehem, Tauranga, New Zealand; p. 647–652.
Geeson JD, Browne KM, Maddison KJ, Shepherd J, Guarald F. 1985. Modified atmosphere packaging to extend the shelf life of tomatoes. J Food Technol. 20:339–349.
Hüster T. 2011. Growth and yield responses of selected crops to treatment with ComCat® [PhD dissertation]. Bloemfontein: University of Free State.
Kleinbaum DG, Klein M. 2010. Logistic regression. A self-learning text. 3rd ed. New York (NY): Springer.
Lammertyn J, Scheerlinck N, Jancsók P, Verlinden BE, Nicolaï BM 2003. A respiration-diffusion model for 'Conference' pears I: model development and validation. Postharvest Biol Technol. 30:29–42.
Lu R. 2004. Multispectral imaging for predicting firmness and soluble solids content of apple fruit. Postharvest Biol Technol. 31:147–157.
Luengwilai K, Fiehn OE, Beckles DM. 2010. Comparison of leaf and fruit metabolism in two tomato (*Solanum lycopersicum* L.) genotypes varying in total soluble solids. J Agric Food Chem. 58:11790–11800.
Mendoza F, Lu R, Ariana D, Cen H, Bailey B. 2011. Integrated spectral and image analysis of hyperspectral scattering data for prediction of apple fruit firmness and soluble solids content. Postharvest Biol Technol. 62:149–160.
Nicolaï BM, Beullens K, Bobelyn E, Peirs A, Saeys W, Theron KI, Lammertyn J. 2007. Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: a review. Postharvest Biol Technol. 46:99–118.
Nunes MCN, Emond JP. 1999. Chlorinated water treatments affects postharvest quality of green bell peppers. J Food Qual. 22:353–361.
Pinheiro J, Alegria C, Abreu M, Gonçalves EM, Silva CLM. 2013. Kinetics of changes in the physical quality parameters of fresh tomato fruits (*Solanum lycopersicum*, cv. 'Zinac') during storage. J Food Eng. 114:338–345.
Popova I, Kiselev VI, Yu VI, 1999. On the nature of biological activity of water and aqueous solutions after treatment in the diaphragm electrolyser. 2nd International Alexandr Gurwitsch Conference. Non equilibrium and coherent systems in biology, biophysics and biotechnology; Москва, Россия, Moscow; p. 6–10.
Pretorius JC, Du Plessis A, van der Watt E. 2008. Bio-stimulatory properties in seeds of plants from the families Caryophyllaceae and Fabaceae with application potential in agriculture. South Afr J Plant Soil. 25:194–203.
Pretorius JC, van der Watt E. 2011. Natural products from plants: commercial prospects in terms of antimicrobial, herbicidal and bio-stimulatory activities in an integrated pest management system. In: Dubey NK, editor. Natural products in plant pest management. London: CAB International; p. 42–90.
Prilutsky VI, Bakhir VM. 1997. Electrochemically activated water: anomalous properties, mechanism of biological action. Moscow (Russia): All Russian Scientific Research and Experimental Institute of Medical Engineering (VNIIIMT).
Riaz MN, Bushway AA. 1996. Compositional analysis of four red raspberry cultivars grown in Maine. J Food Qual. 19:457–465.
Schnabl H, Roth U, Friebe A. 2001. Brassinosteroids-induced stress tolerances of plants. Recent Res Dev Phytochem. 5:169–183.
Tefera T, Seyoum T, Woldetsdik K. 2007. Effect of disinfection, packaging, and storage environment on the shelf life of mango. Biosyst Eng. 96:1537–1550.
Workneh TS, Osthoff G, Pretorius JC, Hugo CJ. 2003. Comparison of anolyte and chlorinated water as a disinfecting dipping treatment for stored carrots. J Food Qual. 26:463–474.