Real-Time Monitoring System for Shelf Life Estimation of Fruit and Vegetables

Methodology

A critical measure that should be immediately taken is to reduce the huge number of postharvest losses, which are reported to be 1.3 billion tons a year, which represents 33% of the production according to. Shelf life is usually defined as the time during which a food product remains safe according to microbiological standards and retaining a desired sensory, physico-chemical and nutritional quality. The patented monitoring system is based on the real-time control of the most influencing environmental variables to estimate the shelf life of the products during the whole supply chain.

The materials and methods described in this paper have been divided into three sections: the equipment designed and used for the continuous monitoring of variables during transport, the physico-chemical and sensory quality tests performed on the commodity to determine its shelf life according to storage temperature, and the mathematical methods used to represent the estimation model.

Component Used

- **Selected Commodity:** Iceberg lettuce (Lactuca sativa')
  Why picked Lettuce as a model? This is due to the great trade importance of this vegetable and also to the physiological disorders that occur when stored under improper postharvest conditions of temperature, relative humidity, ethylene and CO_2.
- Statgraphics Plus software
- R GNU statistical software
-

Algorithms and Equations

Two regression models have been tested, multiple linear regression (MLR) and multiple nonlinear regression (MNLR), with n predictor variables**.**

The model are developed from a training set D = {X, Y} of S samples, which is composed of the predictor matrix, or also called design matrix, $X = [x_1, \ldots, x_i, \ldots, x_S]^T$ and the response matrix $Y = [y_1, \ldots, y_i, \ldots, y_S]^T$. x_i is a column vector of K elements, that can contain all features at a given trial i.

For the MLR Model:

$$x_i = [SP;A;\Delta T;AT]^T$$

and y_i is a column vector of M elements, containing the corresponding days to be estimated at that trial i. Since in our application this is only the days that the commodity keeps the score above 3, y_i is reduced to a scalar and M = 1:

$$y_i = \textit{days with score 3}$$

We consider MLR as:

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \beta_4 x_{4i} + \epsilon_i \; i = 1, \ldots, n$$

In the MNLR model analyzed, we consider that the design matrix X is given by:

$$x_i = [SP;A;\Delta T;AT, SP^2, A^2, SP^*A]^T$$

to determine the influence of quadratic terms, as well as the possible interaction between the set-point and the area above the set-point.
In this case, the MNLR model is:

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_{11} x_{1i}^2 + \beta_{22} x_{2i}^2 + \beta_{12} x_{1i} x_{2i} + \cdots + \epsilon_i \; i = 1, \ldots, n$$

Both models can be expressed in the matrix form as:

$$y = X\beta + \epsilon$$

applying the least-squares criterion, we must estimate those β values that minimize the mean square error, i.e.,:

$$\sum_{i=1}^{n} \epsilon_i^2 = \epsilon' \epsilon = (y - X\beta)'(y - X\beta)$$

This criterion determines those models that maintain a high explanatory level and contain regressors with statistically significant active influence on the response variable without collinearity issues.

Findings

Temp. (°C)	Shelf Life (Days)
20	5
15	8
10	12
5	19
2	25

Table 1: The Shelf Life of Iceberg Lettuces According to the Fixed Storage Temperatures

Model 1 is the model that best correlates with the results. It predicts very well the duration in days and the errors are always in a conservative sense (it proposes a duration lower than that observed) with a maximum error of 1 day. However, it has a complicated interpretation.
Model 2 is easily interpreted. Each increased degree causes a loss of 1.17 days of duration and a unit increase of the area also produces a decrease of 0.025 days. Makes mistakes in both directions, in a range of 3 to 2 days. The model may have some problems since there is a decrease in the R^2 predictor relating to the others.
In model 3, the quadratic term of the area is not significant. This model conservatively makes mistakes, in a range of 0 to 2 days.
Although the model 1 and 3 have similar behaviour, model 3 is much easier to interpret than model 1, since an increase in the value of predictor A (Area) leads to a reduction in shelf life. Thus, in model 3, the coefficient corresponding to predictor A is negative (−0.02063), while in model 1 it is positive (0.1654), in the latter case compensating the duration with the negative coefficient of the interaction (−0.03721). In summary, applying the criterion of statistical parsimony, model 3 is the most appropriate since it presents an explanatory level very similar to that obtained with model 1 using a smaller number of predictors.

Novelty

Concerning the regression models, the best results were obtained using the temperature of the set-point and the area above. The average temperature is not decisive in the prediction. The temperature variation observed in the tests corresponds to certain phases of loading or unloading the lettuce into the truck or, in the case of the simulated tests, the transportation of the goods from the handling industry to the cold storage rooms at university.

Analysis

Physicochemical Analysis

Weight loss was determined in % as:

$$\% = [(\text{Initial weight} - \text{Final weight}) \times 100]/\text{Initial weight}$$

The results were expressed in N mm^{-1}:

$$\Delta E = [\Delta L^*2 + \Delta a^*2 + \Delta b^*2]^{1/2}$$

'Human' Sensory Analysis

Visual appearance, colour, compactness, flavour and overall quality were assessed using a 5-point hedonic scale of acceptability:

- Score: 5: excellent
- Score 4: good
- Score 3: fair – limit of retail (LR)
- Score 2: poor
- Score 1: extremely bad.

Statistical Analysis

The experiment was a two–factor (temperature × storage time) design subjected to analysis of variance (ANOVA) using Statgraphics Plus software. Models were generated using R GNU statistical software. Statistical significance was assessed at the level $p = 0.05$, and Tukey's multiple range test was used to separate means

Research Gap

Even though the nodes can measure many atmospheric variables, this paper is just focused on temperature as it is the most significant factor by far in the quality degradation kinetics of perishable products, intending to propose a prediction methodology which will apply to the other parameters in future research. The size and flexibility of the nodes allow them to be placed anywhere, i.e., on the walls of the truck, on a pallet or inside the cardboard of commodities.

These observations allow using the system presented as a 'Shelf life estimator sensor'. The more data tested, the better the model can predict quality and remaining shelf life.

Future Work

The purpose of this work is not to make an exhaustive study of the shelf life, but to provide an estimation model using the technical means currently used in the land transportation of lettuce, such as air temperature measurement. The presented system allows, to the different intermediaries, to know the conditions in which the load have been stored and/or transported and could contribute to reducing the food waste due to bad conditions of conservation or transportation. However, in order to make a more accurate quality loss prediction model, it is necessary to consistently extend the training and validation work throughout the time, not only after transportation but also in the different phases of the cold chain. This target should be evaluated in future works.

In future researches, this tool should be implemented. It should be able to predict the remaining shelf life for all kinds of commodities according to the initial quality and the environmental factors affecting the location where they have been stored or transported from farm to fork. It should also include non-sensory quality attributes such as, i.e. nutritional compounds, like vitamin C content loss, which is being increasingly demanded by consumers for fresh food.