Mixed models for repeated multicategorical response: modelling the time effect of physical treatments on strawberry sepal quality

Peer-reviewed author version

Lammertyn, Jeroen; De Ketelaere, Bart; Marquenie, David; MOLENBERGHS, Geert & Nicolai, Bart M. (2003) Mixed models for repeated multicategorical response: modelling the time effect of physical treatments on strawberry sepal quality. In: Postharvest Biology and Technology, 30(2). p. 195-207.

DOI: 10.1016/S0925-5214(03)00100-5
Handle: http://hdl.handle.net/1942/428

MIXED MODELS FOR MULTICATEGORICAL REPEATED RESPONSE: MODELLING THE TIME EFFECT OF PHYSICAL TREATMENTS ON STRAWBERRY SEPAL QUALITY

J. Lammertyn[1,+], B. De Ketelaere[2], D. Marquenie[1], G. Molenberghs[3], B.M. Nicolaï[1]

[1]Flanders Centre/Laboratory of Postharvest Technology, Katholieke Universiteit Leuven, Leuven, Belgium

[2]Laboratory for Agro-Machinery and Processing, Katholieke Universiteit Leuven, Leuven, Belgium

[3]Biostatistics, Center for Statistics, Limburgs Universitair Centrum, Diepenbeek, Belgium

[+] Corresponding author

ABSTRACT

Generalised linear mixed models for multicategorical responses (GLMM) were applied to study the effect of a UV-C and light pulse treatment on the visual quality of strawberries sepals. It is illustrated that GLMM works well to analyse repeated quality measures over time on the same subject. The concept of random intercepts and slopes, allowed describing the biological variability inherently present in batches of fruits. The linear mixed models were adapted for multicategorical response according to the threshold concept described in literature. It was found that UV-C treatment slows down the quality decay of the sepals for doses up to 0.1 J/cm^2, but when higher doses were applied the treatment became destructive and the sepals dehydrated and discoloured brown. Since the fungal growth on the strawberry fruit flesh is inhibited significantly starting from a dose of 0.05 J/cm^2, an optimal dose of 0.1 J/cm^2 is recommended to improve the quality retention of the strawberry including the visual aspect of the sepals. The pulsed light dose appeared not to influence the sepal quality.

Keywords: biological variability, fruit quality, fungal growth, longitudinal data, mixed models, UV-C treatment, pulsed light treatment

INTRODUCTION

Postharvest and horticultural research often involves measuring quality evolution during storage and shelf life of horticultural commodities. In this type of experiments, quality characteristics of the same subject are often repeatedly and non-destructively assessed over time, typically resulting in a longitudinal data structure. The repeated measures over time are strongly correlated within subjects, but independent between the subjects. Ordinary regression techniques assume independence between the repeated measures and therefore are not appropriate to analyse the data (Verbeke and Molenberghs, 2000). Mixed-effects models allow to compensate for the within-subject correlation structure, and even more, allow to include between-subject variability in the model. The latter is of great benefit in describing the biological variability in a batch of fruits. A mixed-effects model generally contains random effects in addition to the fixed effects. For longitudinal data, the random component of the mixed model approach allows for subject-specific intercepts and slopes across time. Moreover, it allows for the presence of missing data and time-varying or invariant experimental variables (Verbeke and Molenberghs, 2000; Pinheiro and Bates, 2000).

Until now the quality measure followed over time was assumed to be continuous (e.g., firmness), but in postharvest research the quality measure often takes on discrete values (e.g., fruits exhibiting strong, mild or no symptomatology of a given disorder). Consequently, mixed-effects models for multicategorical or ordinal outcomes are recommended to analyse these data. Recently, an increasing amount of work has focused on generalised linear mixed models for non-continuous or multicategorical response data (GLMM) (Hedeker and Gibbons, 1994; Sheu, 2002). Although these statistical models are already applied in biomedical and psychological studies, no literature was found on their application in postharvest research. In postharvest research the quality measurements are often characterised by a large inherent biological variability, and, therefore GLMM is perfectly suited to extract more information from the repeated non-destructive quality measurements than has been done until now.

The objective of this paper is to illustrate, by means of a case study in postharvest research, the usefulness of generalised linear mixed models to analyse repeated quality measurements with a multicategorical response. GLMM will be applied to study the effect of physical treatments - UV-C treatment and pulsed light treatment - on the visual quality of strawberry sepals. These treatments have proven to be beneficial in reducing microbial spoilage during storage of strawberries (Marquenie et al., 2002b, Marquenie et al. 2003). However, when the physical treatment is too intense it may affect the visual quality of the fruit. Fresh green sepals make the strawberries look much more attractive than brown dehydrated sepals.

MATERIALS AND METHODS

Fruit material

The strawberry cultivar used for all experiments was *Fragaria ananassa* cv. Elsanta, since this is the most widely cultured variety in Belgium. The fruit was harvested at commercial harvesting time at the research centre *Proeftuin Aardbeien en Houtig Kleinfruit, PCF Tongeren* (Belgium). To enable almost year round experiments with strawberries, berries produced according to different culture methods (substrate in greenhouse, soil under plastic tunnels and field grown berries) were used.

Inoculation of the fruit

The conidia suspension of *Botrytis cinerea* MUCL 18864 from the BCCM^{TM}/MUCL collection (Louvain-La-Neuve collection, Belgium) was prepared as described by Marquenie et al. (2002a). Before centrifugation of the suspension, a conidia titer of the suspension was determined using a Thoma counting chamber. Therefore, 200 µl conidia suspension was transferred to the counting chamber and conidia were counted under the microscope (Wild, Switzerland, magnification x 300). The resuspension of the conidia in buffer was so that a final concentration of at least 4.5×10^5 conidia/ml was obtained. For inoculation, the berries were immersed in the spore suspension during approximately 5 s and left to dry on a grid at ambient temperature. The berries were used for the experiments when totally dry.

UV-C treatment

For the UV-C treatment, strawberries were exposed to different UV-C doses (λ=254 nm) in a Bio-Link UV chamber (Vilber Lourmat, France). The built-in UV dosimeter ensured that the programmed dose was applied to the samples. The reflecting inner walls of the treatment chamber enhanced homogeneous distribution of the emitted light. The berries were placed on a grid to enable indirect illumination of the underside of the fruit. The applied doses were 0.00 (control), 0.05, 0.10, 0.50, 1.00 and 1.50 J/cm^2. Twenty fruits were randomly assigned to each UV-C treatment. The experiment was also repeated on strawberries inoculated with *Botrytis cinerea*. The experiment was carried out on three batches of strawberries harvested within a period of one month (27/10/1999, 4/11/1999 and 23/11/1999).

Pulsed light treatment

For the pulsed white light treatment, a 100 W stroboscopic xenon lamp was used (ST100-IE, Sysmat Industrie, France), with a frequency of 15 Hz and a pulse duration of 30 μs. The emitted spectrum ranged from UV-C to infrared, with 50 % of the light in the UV region (λ=200-400 nm). As was the case for the UV light device, the inner walls of the treatment cabinet were of reflecting stainless steel. The doses of the pulsed light treatments used for the experiments were 40, 80, 120, 160, 200 and 250 s. For each condition 20 berries were treated. All strawberries were inoculated and the experiment was carried out on two batches harvested at 30/4/2001 and 7/3/2001.

Quality evaluation

For all treatments the visual quality of the strawberry sepals was evaluated during a period of ten days (storage at 12°C in individual glass jars). The response measurement, sepal quality, is the experimenters' assessment of the overall quality of the strawberry sepals. This variable ranges on an ordinal scale between 1 and 10, where 10 represented fresh and green looking sepals and 1 indicated brown discoloured and shrivelled sepals.

Statistical background

The linear mixed-effect model is formulated as follows (Verbeke and Molenberghs, 2000; Davidian and Giltinan, 1995)

$$\boldsymbol{y_i} = \boldsymbol{X_i}\beta + \boldsymbol{Z_i}\boldsymbol{b_i} + \boldsymbol{e_i} \tag{1}$$

where $\boldsymbol{y_i}$ is a ($n_i \times 1$) vector of continuous responses for the i^{th} experimental subject, i= 1, …, N, the total number of subjects in the study and, n_i the number of repeated measures over time for the i^{th} subject; $\boldsymbol{X_i}$ is a ($n_i \times p$) 'design matrix' that characterises the systematic part of the response, e.g., time and treatment effects, with p the number of fixed effects in the model; $\boldsymbol{\beta}$ is a ($p \times 1$) vector of parameters usually referred to as fixed effects that complete the characterisation of the systematic part of the response; $\boldsymbol{Z_i}$ is a ($n_i \times q$) 'design matrix' that characterises random variation in the response attributable to between-subject sources and with q the number of random effects in the model; $\boldsymbol{b_i}$ is a ($q \times 1$) vector of random effects that completes that characterisation of between-subject variation; $\boldsymbol{e_i}$ is a (n_i × 1) vector of within-subject errors characterising variation due to the way in which the responses are measured on the i^{th} subject; $\boldsymbol{b_i}$ and $\boldsymbol{e_i}$, the two sources of variation present in the model, describe the between-subject and within-subject variation, respectively. The distribution of the random effects and the residuals is assumed normal, with $\boldsymbol{b_i} \sim N(\mathbf{0}, \boldsymbol{D})$ and $\boldsymbol{e_i} \sim N(\mathbf{0}, \boldsymbol{\Sigma_i})$. $\boldsymbol{D}$ is a ($q \times q$) covariance matrix characterising the between-subject variation and $\boldsymbol{\Sigma_i}$ is a ($n_i \times n_i$) covariance matrix describing the variation due to within-subject sources. The structure of $\boldsymbol{\Sigma_i}$ can take on different complexities. Further, $\boldsymbol{b_i}$ is independent of $\boldsymbol{e_i}$ (Verbeke and Molenberghs, 2000).

The linear mixed models described above assume normally distributed response variables, an assumption which does not hold anymore for categorical response variables. Hedeker and Gibbons (1994) were the first to adapt this mixed model theory for categorical response variables. They formulated the statistical theory on the random effects multicategorical ordinal regression model. To motivate this regression model, it is often assumed that a continuous variable (y) exists, which is related to the actual multicategorical response (Y) through the 'threshold concept'. For a dichotomous response variable only one threshold value is assumed. However, for a multicategorical ordinal model, a series of threshold values, $\gamma_0, \ldots, \gamma_K$ where K equals the number of ordered categories, with $\gamma_0 = -\infty$ and $\gamma_K = +\infty$ is required. A response occurs in category k ($Y = k$) if the continuous response variable y exceeds the threshold value γ_{k-1}, but does not exceed the threshold value γ_k.

The response Y_{ij} represents the multicategorical response for subject i at time point t_{ij}, where $i = 1, \ldots, N$, and $j = 1, \ldots, n_i$. When all subjects are measured at the same time points, n_i can be replaced by n. The response vector for the subject i, $\mathbf{Y}_i$, is then equal to $(\mathrm{Y}_{i1}\ \mathrm{Y}_{i2}\ \ldots\ \mathrm{Y}_{in})^{\mathrm{T}}$. Using the threshold concept, this multicategorical response vector, $\mathbf{Y}_i$ is transformed into a continuous response vector, $\mathbf{y}_i$.

All model parameters (the thresholds, the fixed and the random effects) are estimated simultaneously based on the principle of maximum marginal likelihood estimation. Assuming a cumulative logistic response function, which relates the multicategorical response to the continuous response, the marginal likelihood, $ML(\boldsymbol{\theta})$, is calculated using multi-dimensional adaptive Gaussian quadrature (Abramowitz and Stegun, 1972; Pinheiro and Bates, 1995) to numerically integrate over the distribution of random effects. Subsequently, the objective function or negative log marginal likelihood function

$$f(\boldsymbol{\theta}) = -\log ML(\boldsymbol{\theta}) \tag{2}$$

is minimised numerically over $\boldsymbol{\theta}$ in order to estimate the maximum marginal likelihood parameters $\hat{\boldsymbol{\theta}}$. The latter vector contains the estimated fixed effects and variance-covariance parameters as well as the threshold values of the model. The minimisation of the non-linear objective function was carried out using the Newton-Raphson and Quasi Newton optimisation methods (SAS/STAT® User's Guide, Version 8, 1999) and was computationally expensive in terms of time and memory.

Implementing the estimated thresholds values, the fixed effects parameters and the variance parameters and accounting for a cumulative logit function, the probability for a given subject i that a response at time point j occurs in category k ($\mathrm{Y}_j = k$) can then be written as

$$P\left(\mathrm{Y}_{ij} = k \mid \boldsymbol{\beta}, \boldsymbol{b}_i\right) = \frac{1}{1+\exp\left[-\left(\gamma_k - z_{ij}\right)\right]} - \frac{1}{1+\exp\left[-\left(\gamma_{k-1} - z_{ij}\right)\right]} \tag{3}$$

where $z_{ij} = \boldsymbol{Z}_{ij}\boldsymbol{b}_i + \boldsymbol{X}_{ij}\boldsymbol{\beta}$ and $k = 1, \ldots, K\text{-}1$.

Practical implementation

The statistical analysis was carried out under the general framework for non-linear mixed models (PROC NLMIXED) with the SAS/STAT software version 8.2 (SAS Institute, Cary, NC., USA) (SAS/STAT® Users's Guide, Version 8, 1999; Sheu, 2002). In this generalised linear mixed model for multicategorical response the fixed and random effects appear as a part of the linear predictor inside of the link function (cumulative logit) (Agresti, 1996). The number of quadrature points was set at 10 and 15 for models with one and two random effects, respectively. This number was a trade-off between adequate integration accuracy and computation time.

In order to test hypotheses related to the fixed effects, the likelihood ratio chi-squared test was used for comparison of nested models:

$$G^2 = -2 \log \left[\frac{ML(\hat{\boldsymbol{\theta}}_0)}{ML(\hat{\boldsymbol{\theta}})} \right] \tag{4}$$

where *ML* denotes the marginal likelihood function and $\hat{\boldsymbol{\theta}}_0$, the parameters estimated under *ML* for a reference or null model, being a subset of the parameters of $\hat{\boldsymbol{\theta}}$. G^2 then follows, asymptotically, under the null model, a chi-squared distribution with degrees of freedom equal to the difference between the dimensions of the two parameters spaces.

To test whether random effects improve the model, the likelihood ratio test defined above was used. It follows asymptotically a null distribution that is a mixture of chi-squared distributions, rather than the classical single chi-squared distribution that was used to test the fixed effects (Verbeke and Molenberghs, 2000). For instance, in testing no random effect versus one random effect, the null distribution is a mixture (average) of the chi-squared distributions with 0 and 1 degrees of freedom. To test the significance of adding a second random effect the null distribution of the test statistic is the average of the chi-squared distribution with 1 and 2 degrees of freedom.

When two non-nested models need to be compared, the likelihood ratio test is not appropriate anymore and a more general approach is required. In general, the Akaike's information criterion (AIC) (Akaike, 1974) can be applied to compare two generalised linear mixed models:

$$AIC = -2 \log ML\left(\hat{\boldsymbol{\theta}}\right) + 2p \tag{5}$$

where p is the number of parameters in the model. The marginal log likelihood statistic (-2 log $ML(\hat{\boldsymbol{\theta}})$) only decreases when more explanatory variables are included in the model. It tends to select models, which over fit the available data. Therefore the AIC is often used for model selection. This is a simple penalisation of the marginal log likelihood function for the complexity of the model. Twice the number of parameters is added to the –2 log $ML(\hat{\boldsymbol{\theta}})$. A smaller AIC indicates a more preferable model in terms of the data.

The model building strategy was based on both the G^2 and these AIC values. The strategy consisted of three steps. *First*, a model with only fixed main and interaction effects was built as a preliminary screening of the dataset based on the G^2. Non-significant (p-value higher than 0.05) main and interaction effects were omitted from the model. *Second*, a mixed model including a random intercept was constructed and compared to the model with only fixed effects to test the significance of the random intercept. *Finally*, a model with random intercept and slope was constructed and the significance of the random slope was tested comparing the G^2 and the AIC values with the previous models.

RESULTS AND DISCUSSION

UV-C treatment

In Fig. 1, 3 sepal quality versus time profiles are shown for non-inoculated strawberries treated with a dose of 0.01 J/cm^2. The large biological variability between the individual strawberries at day 1 increases with time and justifies the implementation of mixed models including random intercepts and slopes. To explore the data the average quality profile of 20 strawberries is shown for each of the different UV-C treatments (Fig. 2). The strawberry quality profiles for the highest UV doses (1 and 1.5 J/cm^2) are clearly separated from the other profiles, indicating a negative effect of high UV-doses on sepal quality. This dose effect on the sepal quality will later on be reflected in the model structure.

In Table 1 an overview is given of different models constructed for three independent batches of strawberries. Model 1 includes only significant fixed effects and has an AIC of 7336. When a random intercept was included, the AIC decreased drastically from 7336 to 6055 and the likelihood ratio test statistic G^2=1282.7 (df 0:1, p<0.0001) indicated that the null hypothesis of no random intercept effect could be rejected. The random intercept describes the biological variability in sepal quality between the strawberries within one batch at harvest. A further decrease in AIC value to 5621 was observed by adding a random slope or random time effect. This random slope allows each individual strawberry to have a different sepal quality decay over time. By including this random effect, the UV-dose × inoculation and time × inoculation interactions were not significant anymore. Since model 3 had a different fixed effects structure from model 1 and 2, the likelihood ratio test statistic could not be used to assess the effect of the random effect. Since the models under investigation were not nested, the AIC was required to compare the models.

The significant effect of the random intercept illustrates that at day 1 a large biological variability exists between the 20 strawberries for a given combination of

UV-C dose and inoculation type. The random time effect (slope) shows that the quality decay over time is not equal for all strawberries within one group. Some of them have a higher decay rate than others. Intuitively it can be felt that this way of modelling is more reliable than merely fitting a model to the average sepal quality value per time point. Moreover, the analysis indicated a significant effect of inoculation, time, time^2 and UV-dose as well as the interaction between UV-dose and time (model 3). When the berry was inoculated, the sepal quality decay was faster than for non-inoculated ones. The sepal quality was quadratically related to time and for increasing UV-C dose a faster sepal quality decay was observed over ten days (Fig. 2). The significance of the interaction term between time and UV dose shows that the quality decay rate is a function of the applied dose (Fig. 2). To reduce the computation time and to assess the effect of the number of classes on the model performance, models were also built for the sepal quality graded in five classes. Hereto, two consecutive classes of the original responses were combined. The constructed models are summarized in Table 1 (model 4 to 6). Model 4 was improved considerably by including the random intercept (model 5 with AIC=3374, G^2=972.4, df 0:1, $p<0.0001$) and random time effect (model 6 with AIC=3109). The significant fixed effects were the same as for the models with ten classes.

In the aforementioned models a linear relation was assumed between UV-dose and sepal quality, resulting in a good model. In Fig. 3 this relationship is depicted for the time point equal to 4 days. A linear relation (see straight line) fits well but it can be noticed that a slight increase in sepal quality is obtained treating them with a mild UV-dose (0.01 to 0.1 J/cm^2). Five indicator variables were used to describe the relation between sepal quality and UV-dose. In Table 1 the AIC values are given for models 7 to 9, which account for this non-linear UV-dose effect. The AIC value of model 9 (AIC=3018) is considerably smaller than the AIC value of the corresponding model with the linear effect of UV-dose (model 6, AIC=3109). It means that, although more parameters are present in the latter model, the gain in –2 log *ML* is large enough to overcome the penalisation of the increase in number of parameters. Also the log likelihood ratio test indicated a significant improvement of the model (G^2=107 with 8 df, p<0.0001). The model with non-linear UV-dose effect could not be calculated for the sepal quality expressed in ten classes, since it was too computer memory intensive and time consuming. But since model 3 and model 6 have the same model structure it is plausible to assume that similar conclusion can be drawn for the model with a response graded in ten classes.

Similar results were observed for two other independent batches of strawberries (models 10 to 29). In contrast to models 3 and 23, model 13 for batch 3 contains the fixed interaction effects between UV-dose and inoculation and between time and inoculation are significant. The difference in AIC value between model 12 and 13 indicates the effect of omitting those latter two interaction terms on the model fit. With an increase of almost 30 AIC units, the most complex model is significantly better (G^2=35, 2 df, p<0.0001). Similarly, for a response variable graded in 5 classes, the model including the two interaction effects (model 17, AIC=2791) describes the data better than model 16 (AIC=2817) without these effects. The AIC value of the model with random intercept and slope and a non-linear effect of UV-dose (model 20, AIC=2800) is lower than for model 16, the corresponding model with linear dose effect, but is higher than the model including both interaction effects (model 17, AIC=2791). The obtained results for batch 1 are very similar to those described for batch 2.

All three batches had a similar value for the random intercept standard deviation, indicating similar batch variability at the beginning of the experiment. Also the random time standard deviation was similar, illustrating that the variability on the slope (the rate of quality decay) of the sepals was comparable between the batches. The measured sepal quality and the corresponding model fit are given in Fig. 4 for an untreated and treated (1.5 J/cm^2) inoculated strawberry.

The non-linear UV-dose effect can be explained by looking both at the inhibitory effect of UV-C light on fungal growth and at the consequences of the exposure of plant material to UV-C light. Marquenie et al. (2002a) showed a log linear relationship between UV-C dose and inactiviation of *Botrytis* conidia under *in vitro* conditions. Plants will react to relatively low doses with several protection mechanisms. At one hand, exposure of plants to UV-C light induces synthesis and accumulation of protective components. Different kinds of UV-absorbing pigments in the epidermis of UV-C exposed plants have been described, such as anthocyanin in *Coleus* leaves (Burger and Edwards, 1996), and flavonoids in aquatic fern (Jayakumar et al., 1999). Molecules with antioxidant properties are often synthesised to protect the chloroplast enzymes. Examples are tocopherol and ascorbic acid in soybean (Kozak et al., 1999). Some epidermal pigments have also antioxidant properties, such as apigeninidin in soybean (Boveris et al., 2001). At the other hand, synthesis of enzymes required for repair mechanisms are also induced (Barka et al., 2000; Casati et al., 2001). Higher UV-C doses will cause major damages to the cellular structures, such as the PS II complex resulting in a reduced photosynthetic activity (Salter et al., 1997, Jayakumar et al., 1999), lipid peroxidation and efflux of electrolytes (Barka et al., 2000; Pennanen et al., 2002).

From the analysis it can be concluded that a mild treatment inhibits the fungal growth on the sepals and induces protective mechanisms in the plant leaves, hence improving the quality, but when the dose is increased, the destructive effect of UV-C treatment on the sepals becomes prominent.

Marquenie et al. (2002b) proposed UV-C light treatment as an alternative technique for the use of chemicals to reduce the development of *Botrytis cinerea* on strawberry fruit during storage. To investigate the effect of UV-C on microbial inactivation and on fruit quality, inoculated berries were subjected to different UV-C doses ranging from 0.05 to 1.50 J/cm^2. After the treatment, fungal growth, visual damage and fruit firmness were evaluated during a period of 10 days. The experimental data were analysed statistically using survival analysis techniques. Fungal growth on strawberries was significantly retarded using UV-C doses of 0.05 J/cm^2 and higher. Due to the physical treatment, the period before the first observation of fungal growth was increased by at least 1 or 2 days compared to the untreated inoculated strawberries. The highest doses had a negative effect on the sepals of the strawberry, causing browning and drying of the leaves. The hypothesis is thus that two processes work in the opposite direction. The optimal recommended UV-dose to slow down microbial growth and to prevent visual sepal quality decay should be taken between 0.05 and 0.5 J/cm^2. The experiments do not allow to determine more accurately the exact position of the optimal UV-dose in this range.

Pulsed light treatment

A summary of the models, which were built to analyse the effect of pulsed light treatment on the strawberry sepal quality, is given in Table 2. In model 1 (batch 1 and 10 response classes), only fixed effects were included. The pulsed light dose or the interaction of it with time had no significant effect on the sepal quality. Omitting non-significant parameters resulted in an AIC value of 4030. Subsequently, a model was constructed containing all significant fixed effects and a random intercept (model 2). By including a random intercept, which takes into account the within batch biological variability in initial sepal quality between the individual strawberries, the AIC value decreased considerably to 3011 (model 2). The likelihood ratio test indicated a significant improvement of the model (G^2=1021.3, 0:1 df, $p<0.0001$). The AIC value was further lowered to 2751 by inserting a random slope effect (G^2=263.9, 1:2 df, $p<0.0001$). This random slope allows each individual strawberry to have a slightly different sepal quality decay over time.

To assess the effect of number of classes of the response variable, the same model building strategy was applied to analyse the effect of pulsed light treatment on strawberry sepal quality for a response variable taking on only five classes. Hereto, two subsequent classes of the original response variable were combined. Again, the model is improved considerably when next to the fixed effects also a random intercept and random slope are introduced. The AIC drops from 2518 (model 4) to 1597 (model 6). The same model structure was obtained for a response variable with ten and five classes. In both model 6 and 12, no significant effect of pulsed light dose was observed and the transformed categorical variable, which is continuous, showed a quadratic trend over time.

A similar analysis was conducted for a second batch of strawberries, which were harvested 7 weeks later. For a response with ten classes, a significant effect of pulsed light dose was found in model 7 which contained only fixed effects. However when a random intercept was included this dose effect was not significant anymore (model 8). With a random slope the AIC dropped from 3745 for model 7 to 2506 for model 9. Again, in the latter model the pulsed light dose effect was not significant anymore. The models for a response of 5 classes showed a similar trend, yielding a final AIC value of 1482 for model 12 with random intercept and slope. Again, no significant effect of pulsed light dose was observed.

Table 3 shows the parameter estimates for model 3 (batch 1) and model 9 (batch 2). These parameter estimates relate the transformed multicategorical response variable, which is continuous, to the independent variables (e.g. time, pulsed light dose). A closer look at the parameter estimates for the intercept of both batches shows that the average initial (time zero) strawberry sepal quality is much higher for batch 2 than for batch 1. On the continuous response scale (Fig. 6), the intercept of batch 1 is half of that for batch 2. However, on the categorical scale this difference is limited to one class (Fig. 5), without loss of significance of the effect. Although the average initial sepal quality might be different between both batches, the magnitude of the biological variability present between strawberry sepals within a batch is similar for both batches. The standard deviation on the intercept is almost identical for both batches and equals 5.65 and 5.41 for batch 1 and 2, respectively. Both batches show a quadratic effect of the continuous response variable over time (Fig. 6). For batch 2 this quadratic curvature is more pronounced than for batch 1. The time effect on the multicategorical response variable is illustrated in Fig. 5. The sepal quality of the strawberries of batch 2 has a higher initial value and remains higher for a longer period than the strawberry sepals of batch 1. The standard deviation of the time evolution (quality decay rate) of the sepal quality is almost identical for both batches, reflecting a similar biological variability on the sepal quality decay rate (slope) for the two batches. A random slopes model implies an increase in the variance on the sepal quality with time (Verbeke and Molenberghs, 2000). The estimated covariance between the intercepts and the quality decay rates is

-2.10 and -2.14 for batch 1 and batch 2, respectively. This negative value illustrates that green fresh sepals, in general, experience a faster decay rate than sepals with a lower initial quality at harvest.

The mixed effects components in the model learn that both batches have a similar within batch biological variability but that the average sepal quality of batch 2 is clearly higher than that of batch 1. However, the sepal quality decay rate for batch 2 was higher than for batch 1. No quality extending effect on the strawberry sepals was observed for the pulsed light treatment.

The threshold values depicted in Table 3 were estimated and used to transform the multicategorical response variable into a continuous response variable (Fig. 6) on which the analyses are carried out. Afterwards, the obtained model response was then transformed back into the corresponding multicategorical response or sepal quality class (Fig. 5).

In contrast to the UV-C treatment, no effect of a pulsed light treatment on strawberry shelf-life was observed (Marquenie et al., 2003). Different doses of intense white light pulses did not affect the fungal development on strawberries during subsequent storage. Also, after the longest pulsed light treatment (250 s), no significant difference between the treated inoculated strawberries and the control group was observed. Accordingly, as reported in this paper, no detrimental effect on the strawberry sepals was noticed.

The lack of efficiency of exposure to intense white light pulses might be caused by the relative low amount of UV-C in the emitted spectrum. As was observed by Anderson et al. (2000), the inactivation properties of white light pulses are mainly based on the presence of UV light. Removing the UV component from the emitted light resulted in an important decrease in inactivation of bacteria and fungi.

Slieman and Nicholson (2000) analysed the effect of the three UV parts on bacterial spores and showed that UV-C caused an important part of the total damage. The UV-C intensity of the emitted spectrum for the pulsed light experiments was 0.55 mW/cm², resulting in a UV-C dose varying from 0.02 to 0.14 J/cm². These doses are 10 times lower than the maximal doses applied during the UV-C experiments. Although the applied UV-C dose was similar for both light treatments, the inactivation mechanism seems to be different. A dose of 0.14 J/cm², was already sufficient to reduce the sepal quality decay rate for the continuous UV-C treatment, but was apparently not high enough to cause a significant increase in longevity of the sepals for the pulsed light treatment. This difference might be attributed to the presence of the visible part of the spectrum or to other unknown interactive effects during the pulsed light treatment.

Practical considerations of the on-line implementation of UV-C treatment

In this paragraph preliminary results on the practical implementation of UV-C treatment will be shortly discussed (Marquenie, 2002c). An experimental set-up was designed to implement surface disinfection of strawberries with UV-C. This technique was shown the most effective to reduce fungal development and, from a practical point of view, required only limited adaptations to commercial installations at a relatively low cost. Therefore, a conveyor belt similar to those used in the auctions was equipped with UV-C lamps. Pursuing commercial conditions implied also the use of non-inoculated strawberries in punnets. This aspect was the most important difference with the laboratory-scale experiments performed in the first phase of the study, and proved to be the limiting factor for the application of UV-C at a larger scale. Next to the treatment intensity, the size of the punnets had a very large influence on fungal development during the evaluation period: a significant reduction in fungal growth was only observed for the smallest punnet with 2 or 3 layers of strawberries (250 g). Since the plastic polymers used for the punnets absorb the UV-C radiation, illumination is only possible at one side. Fungal development on strawberries in punnets of 500 g with 5 or more layers was not affected by the treatment.

CONCLUSIONS

Generalised linear mixed models are very well suited to describe the time evolution of quality characteristics of individual fruits. The inclusion of random intercepts and slopes, allowed describing the biological variability inherently present in batches of fruits. Based on the threshold concept formulated in the literature, these models were adapted for multicategorical response variables, which often occur in postharvest research. This statistical technique was applied to study the sepal quality of strawberries over time as a function of UV-C dose and pulsed light treatment. These two physical treatments have shown to be effective in reducing the microbial load of strawberries. It was illustrated that a mild UV-C treatment had a beneficial effect on the sepal quality while too high doses increased the sepal quality decay rate, resulting in brown discoloured and dehydrated sepals. Since an UV-C dose of 0.05 J/cm^2 is sufficient to inhibit fungal growth on strawberry fruits, an optimal UV-C treatment of 0.1 J/cm^2 is recommended to enhance the shelf life and visual aspects of the strawberries. It was shown that the initial biological batch variability and the variability on the quality decay rate were similar for all three batches, proving the generality of the obtained results. The pulsed light treatment did not significantly reduce the sepal quality decay rate during storage.

ACKNOWLEDGEMENTS

This research is supported by the Ministry of Small Enterprises, Traders and Agriculture, Directorate of research and Development (Project S-5856). The research at the Flanders Centre of Postharvest Technology is supported by the Flemish Government. Jeroen Lammertyn is Postdoctoral Researcher of the Fund for Scientific Research Flanders (F.W.O.-Vlaanderen). Bart De Ketelaere is Doctoral Fellow with the Institute for the Promotion of Innovation by Science and Technology in Flanders (I.W.T.-Vlaanderen).

REFERENCES

Abramowitz, M. Stegun, I.A. 1972. Handbook of mathematical functions. New York: Dover Publications, Inc.

Agresti, A. 1996. An introduction to categorical data analysis. Second edition, John Wiley & Sons, New York.

Akaike, H. 1974. A new look at the statistical model identification. *IEEE Transactions on automatic control*, 19, 716-723.

Anderson, J.G., Rowan, N.J., MacGregor, S.J., Fouracre, R.A., Farish, O. 2000. Inactivation of foor-borne enteropathogenic bacteria and spoilage fungi using pulsed-light. *IEEE T. Plasma Sci.* 28, 83-88.

Barka, E.A., Kalantari, S., Makhlouf, J., Arul, J. 2000. Impact of UV-C irradiation on the cell wall-degrading enzymes during ripening of tomato (*Lycopersicon esculentum* L.) fruit. *J. Agr. Food Chem.* 48, 667-671.

Boveris, A.D., Galatro, A., Sambrotta, L., Ricco, R., Gurni, A.A, Puntarulo, S. 2001. Soybean cotyledons directly exposed to UV-C (190-280 nm) contained a colored pigment in those areas of the epidermis directly exposed to UV-C. *Photochemistry* 58, 1097-1105.

Burger, J., Edwards, G.E. 1996. Photosynthetic efficiency, and photodamage by UV and visible radiation in red- versus green-leaf *Coleus* varieties. *Plant Cell Physiol.* 37, 395-399.

Casati, P., Andreo, C.S. 2001. UV-B and UV-C induction of NADP-malic enzyme in tissues of different cultivars of *Phaseolus vulgaris* (bean). *Plant Cell Environ.* 24, 621-630.

Davidian, M., Giltinan, D. M. 1995. Nonlinear models for repeated measurement data, London: Chapman & Hall.

Hedeker, D. and Gibbons, R.D. 1994. A random-effects ordinal regression model for multilevel analysis. *Biometrics* 50, 933-944.

Jayakumar, M., Eyini, M., Selvinthangadurai, P., Lingakumar, K., Premkumar, A., Kalundaivelu G. 1999. Changes in pigment composition and photosynthetic activity of aquatic fern (*Azolla microphylla* Kaulf.). *Photosynthetica* 37, 33-38.

Kozak, R.G., Ricco, R.A., Gurni, A.A., Boveris, A.D., Puntarulo, S. Antioxidant response of soybean cotyledons (*Glycine max*) to ultraviolet irradiation. *Can. J. Plant Sci.* 79, 181-189.

Marquenie D., Lammertyn J., Geeraerd A.H., Soontjens C., Van Impe J.F., Nicolaï B.M., Michiels CW. 2002a. Inactivation of conidia of *Botrytis cinerea* and *Monilinia fructigena* using UV-C and heat treatment. *Int. J. Food Microbiol.* 74, 27-35.

Marquenie, D., Michiels, C.W., Geeraerd, A.H., Schenk, A., Soontjens, A., Van Impe, J.F. and Nicolaï, B.M. 2002b. Using survival analysis to investigate the effect of UV-C and heat treatment on storage rot of strawberry and sweet cherry. *Int. J. Food Microbiol.* 73, 191-200.

Marquenie, D. 2002c. Evaluation of physical techniques for surface disinfection of strawberry and sweet cherry. PhD. thesis nr. 542, Faculty of Agricultural and Applied Biological Sciences, Catholic University of Leuven.

Marquenie, D., Michiels, C.W., Van Impe, J.F., Schrevens, E., Nicolaï, B.M. 2003. Pulsed white light in combination with UV-C and heat to reduce storage rot of strawberries. *Postharvest Biol. Technol.* (*in press*)

Pennanen, A., Xue, T.L., Hartikainen, H. 2002. Protective role of selenium in plants subjected to severe UV irradiation stress. *J. Appl. Bot.-Angew. Bot.* 76, 66-76.

Pinheiro, J.C., Bates, D.M. 1995. Approximations to the log-likelihood function in nonlinear mixed-effects model. *J. Comput. Graph. Stat.*, 4, 12-35.

Pinheiro, J.C. and Bates, D.M. 2000. Mixed effects models in S and S-Plus, Springer Series in Statistics and Computing. New-York: Springer-Verlag.

Salter, A.H., Koivuniemi, A., Strid, A. 1997. UV-B and UV-C irradiation of spinach PSII preparations *in vitro*. Identification of different fragments of the D1 protein depending upon irradiation wavelength. *Plant Physiol. Biochem.* 35, 809-817.

SAS Institute Inc., SAS/STAT®, 1999. User's Guide, Version 8, Cary, NC: SAS Institute Inc.

Sheu, C.F. 2002. Fitting mixed-effects models for repeated ordinaloutcomes with the NLMIXED procedure. *Behav. Res. Methods Instr. Comput.* 34, 151-157.

Slieman, T.A., Nicholson, W.L. 2000. Artificial and solar UV radiation induces strand breaks and cyclobutane pyrimidine dimers in *Bacillus subtilis* spore DNA. *Appl. Environ. Microb.* 66, 199-205.

Verbeke, G. and Molenberghs, G. 2000. Linear mixed models for longitudinal data, Springer Series in Statistics. New-York: Springer-Verlag.

TABLES

Table 1. Overview of the different generalised linear mixed models with multicategorical response (5 or 10 classes) for UV-C treatment. + and – signs indicate whether the variable is entered in the model or not. The –2Log ML(***θ***) value can be calculated from the AIC according to Eq. (5).

Model	Int	Rand Int	Time	Rand Time	Time²	UV dose	Inoc	UV dose × Time	UV dose × Inoc	Time × Inoc	AIC
Batch 2: 10 response classes, linear effect of UV dose											
1	+	-	+	-	+	+	+	+	+	-	7336.2
2	+	+	+	-	+	+	+	+	+	-	6055.5
3	**+**	**+**	**+**	**+**	**+**	**+**	**+**	**+**	**-**	**-**	**5621.8**
Batch 2: 5 response classes, linear effect of UV dose											
4	+	-	+	-	+	+	+	+	+	+	4344.5
5	+	+	+	-	+	+	+	+	+	+	3374.1
6	**+**	**+**	**+**	**+**	**+**	**+**	**+**	**+**	**-**	**-**	**3109.5**
Batch 2: 5 response classes, non-linear effect of UV dose											
7	+	-	+	-	+	+	+	+	-	-	4023.9
8	+	+	+	-	+	+	+	+	-	-	3255.4
9	**+**	**+**	**+**	**+**	**+**	**+**	**+**	**+**	**-**	**-**	**3018.9**
Batch 3: 10 response classes, linear effect of UV dose											
10	+	-	+	-	+	+	+	+	+	+	7137.2
11	+	+	+	-	+	+	+	+	+	+	5166.7
12	+	+	+	+	+	+	+	+	-	-	4765.5
13	**+**	**+**	**+**	**+**	**+**	**+**	**+**	**+**	**+**	**+**	**4734.5**
Batch 3: 5 response classes, linear effect of UV dose											
14	+	-	+	-	+	+	+	+	+	+	4576.5
15	+	+	+	-	+	+	+	+	+	+	3093.7
16	+	+	+	+	+	+	+	+	-	-	2817.6
17	**+**	**+**	**+**	**+**	**+**	**+**	**+**	**+**	**+**	**+**	**2791.5**
Batch 3: 5 response classes, non-linear effect of UV dose											
18	+	-	+	+	+	+	+	+	+	+	4632.6
19	+	+	+	-	+	+	+	+	+	+	3133.5
20	**+**	**+**	**+**	**+**	**+**	**+**	**+**	**+**	**-**	**-**	**2800.2**
Batch 1: 10 response classes, linear effect of UV dose											
21	+	-	+	-	+	+	+	+	+	+	6989.3
22	+	+	+	-	-	+	+	+	+	+	4916.9
23	**+**	**+**	**+**	**+**	**+**	**+**	**+**	**+**	**-**	**-**	**4599.8**
Batch 1: 5 response classes, linear effect of UV dose											
24	+	-	+	-	+	+	+	+	+	+	3565.8
25	+	+	+	-	-	+	+	+	+	+	2857.0
26	**+**	**+**	**+**	**+**	**+**	**+**	**+**	**+**	**-**	**-**	**2675.8**
Batch 1: 5 response classes, non-linear effect of UV dose											
27	+	-	+	-	+	+	+	+	-	-	3547.2
28	+	+	+	-	+	+	+	+	-	-	2815.1
29	**+**	**+**	**+**	**+**	**+**	**+**	**+**	**+**	**-**	**-**	**2638.8**

Table 2. Overview of the different generalised linear mixed models with multicategorical response (5 or 10 classes) for the pulsed light treatment. + and – signs indicate whether the variable is entered in the model or not. The $-2\text{Log ML}(\boldsymbol{\theta})$ value can be calculated from the AIC according to Eq. (5).

Model	Int	Rand Int	Time	Rand Time	Time^2	Pulsed light dose	Pulsed light × time	AIC
Batch 1: 10 response classes								
1	+	-	+	-	+	-	-	4030.4
2	+	+	+	-	+	-	-	3011.1
3	**+**	**+**	**+**	**+**	**+**	**-**	**-**	**2751.2**
Batch 1: 5 response classes								
4	+	-	+	-	+	-	-	2518.6
5	+	+	+	-	+	-	-	1746.8
6	**+**	**+**	**+**	**+**	**+**	**-**	**-**	**1597.9**
Batch 2: 10 response classes								
7	+	-	+	-	+	+	-	3745.6
8	+	+	+	-	+	-	-	2812.2
9	**+**	**+**	**+**	**+**	**+**	**-**	**-**	**2506.3**
Batch 2: 5 response classes								
10	+	-	+	-	+	+	-	2384.8
11	+	+	+	-	+	+	-	1689.7
12	**+**	**+**	**+**	**+**	**+**	**-**	**-**	**1482.6**

Table 3. Parameter estimates for model 3 (batch 1) and model 9 (batch 2)

	Parameter estimates	
Parameter	Batch 1 (model 3)	Batch 2 (model 9)
Intercept	6.76	13.3
Time	-1.54	-0.279
Time^2	-0.0623	-0.163
Standard Dev. (intercept)	5.65	5.41
Standard Dev. (time)	0.54	0.663
Covariance (intercept, time)	-2.07	-2.14
Threshold 1	-13.3	-15.1
Threshold 2	-9.89	-9.22
Threshold 3	-7.87	-6.53
Threshold 4	-5.46	-3.29
Threshold 5	-2.77	-0.829
Threshold 6	1.03	1.44
Threshold 7	4.35	4.57
Threshold 8	7.77	9.00
Threshold 9	14.8	15.7

FIGURE CAPTIONS

Figure 1: Strawberry sepal quality versus time profiles for 3 non-inoculated strawberries treated with a UV-C dose of 0.01 J/cm^2. Each symbol indicates another strawberry.

Figure 2: Average sepal quality time profiles (based on 20 inoculated strawberries) for different UV-C treatments (Δ: 0.00, ◊: 0.01, ○: 0.10, +: 0.50, ∗: 1.00 and □: 1.50 J/cm^2).

Figure 3: Sepal quality as a function of the UV-C dose. The straight line indicates the linear relation between UV-C dose and the sepal quality.

Figure 4: Measured and predicted sepal quality versus time profiles for one arbitrarily chosen treated (1.5 J/cm^2) and untreated strawberry. Legend: ×: 0.0 J/cm^2, measured; ○: 0.0 J/cm^2, model; ◊: 1.5 J/cm^2, measured; +: 1.5 J/cm^2, model.

Figure 5: Multicategorical response profile of one strawberry from batch 1 (○) and batch 2 (+).

Figure 6: Continuous response profile of one strawberry from batch 1 (○) and batch 2 (+), corresponding to the multicategorical response profiles depicted in Figure 5. The threshold values given in Table 3 were used to transform the continuous into the multicategorical response and vice versa.

FIGURES

Figure 1

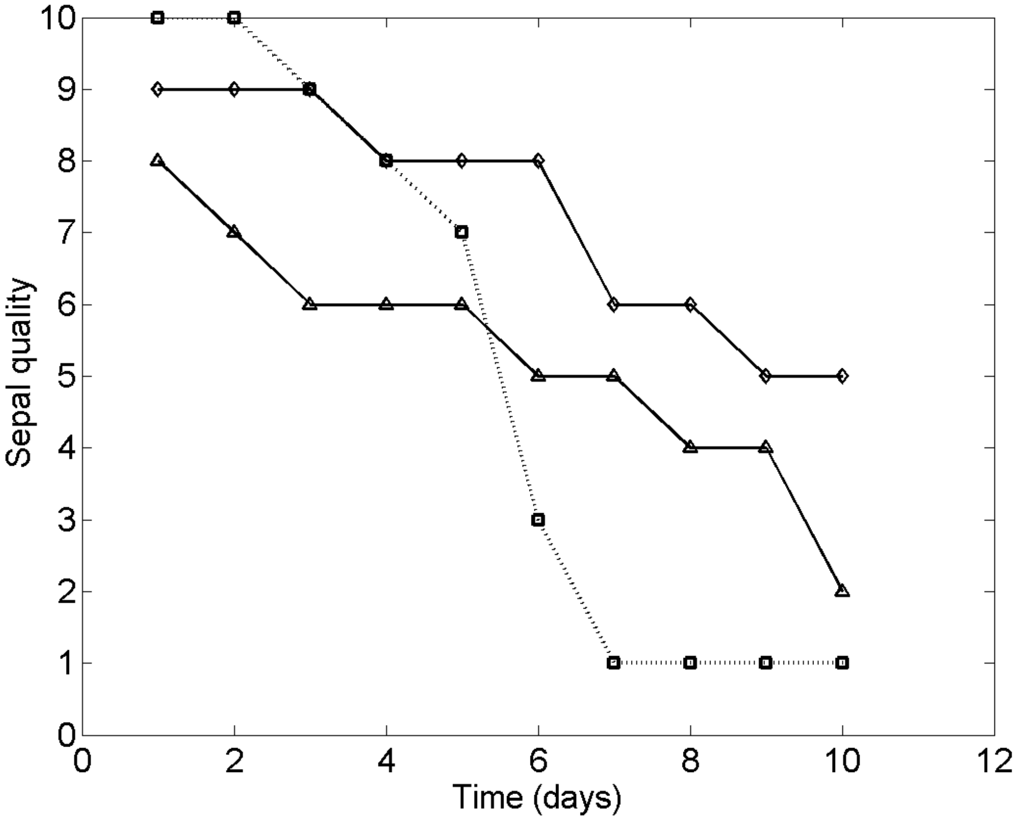

Figure 2

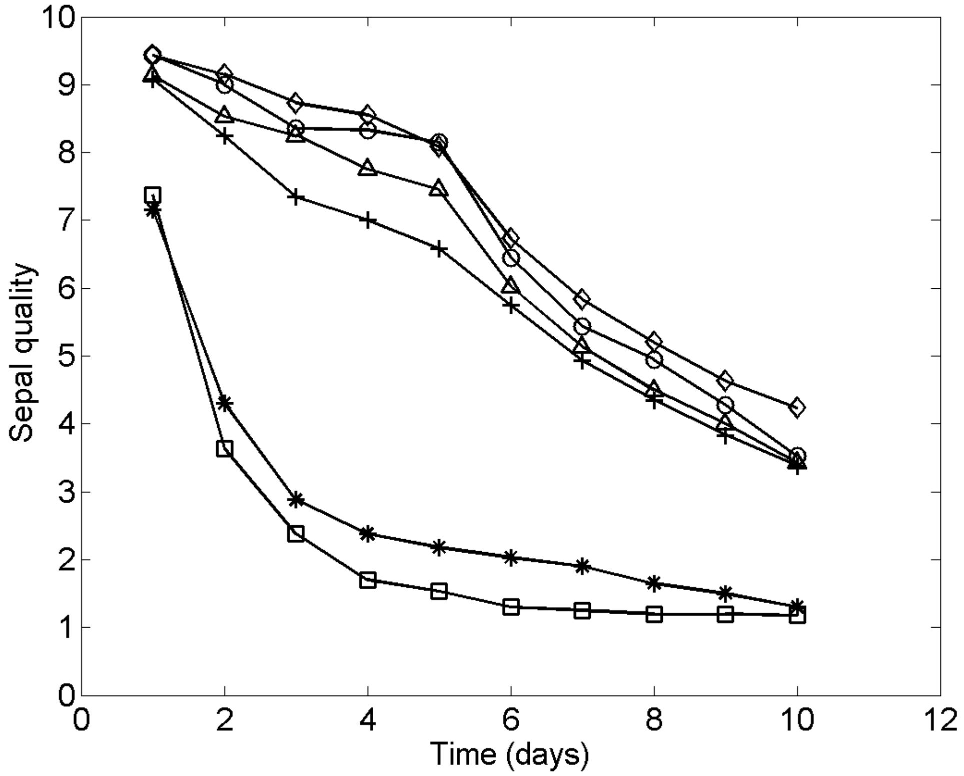

Figure 3

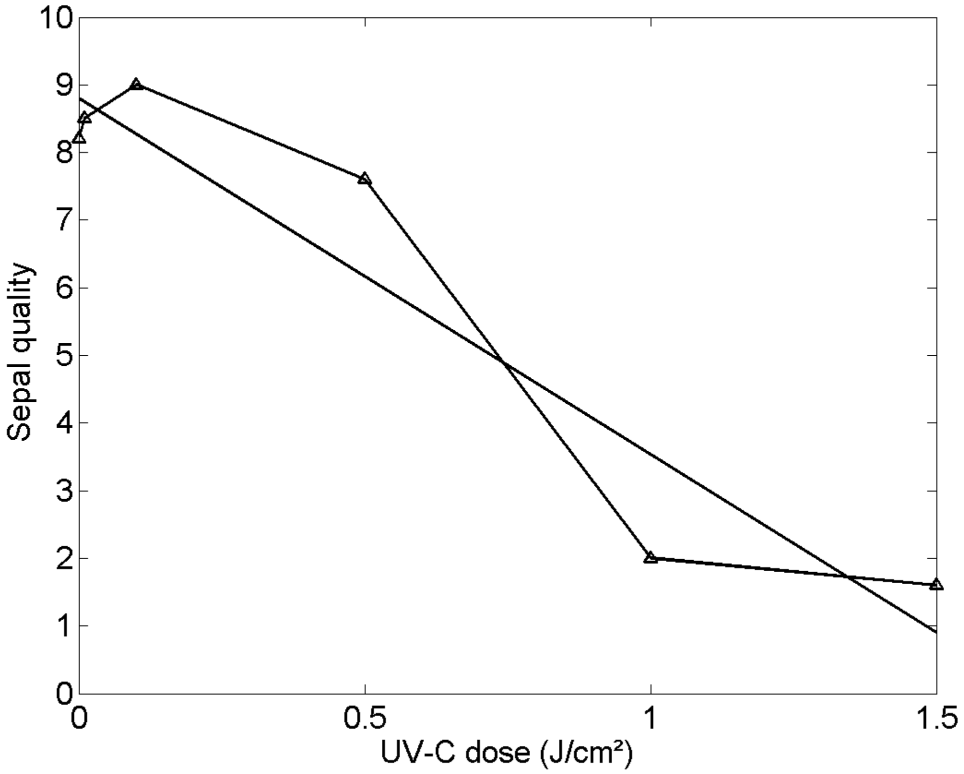

Figure 4

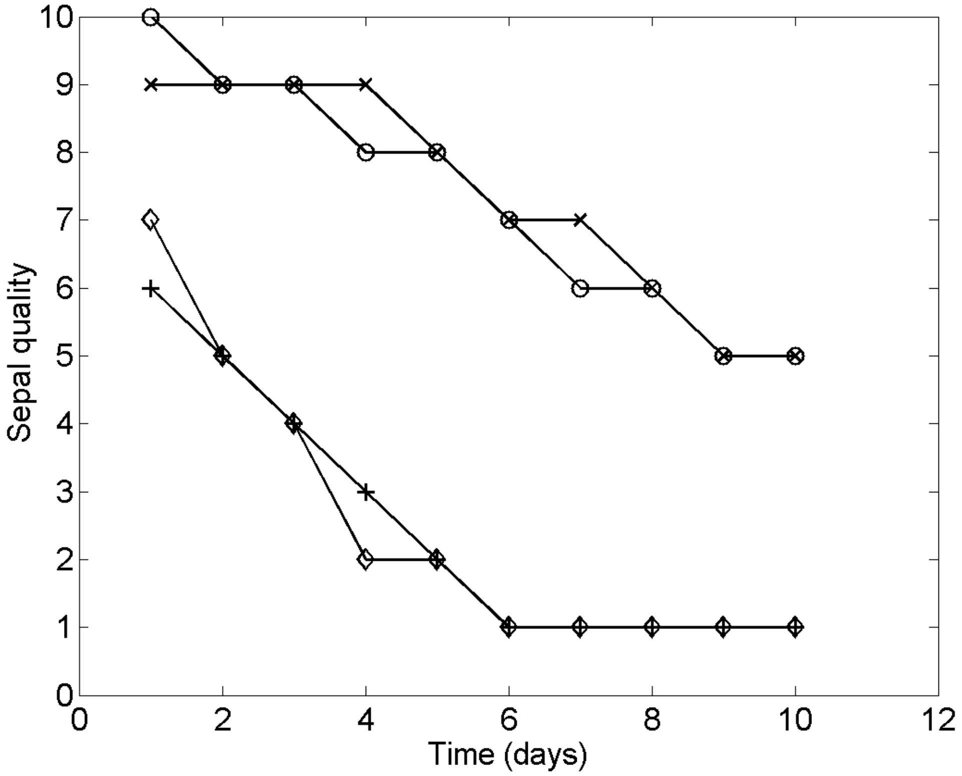

Figure 5

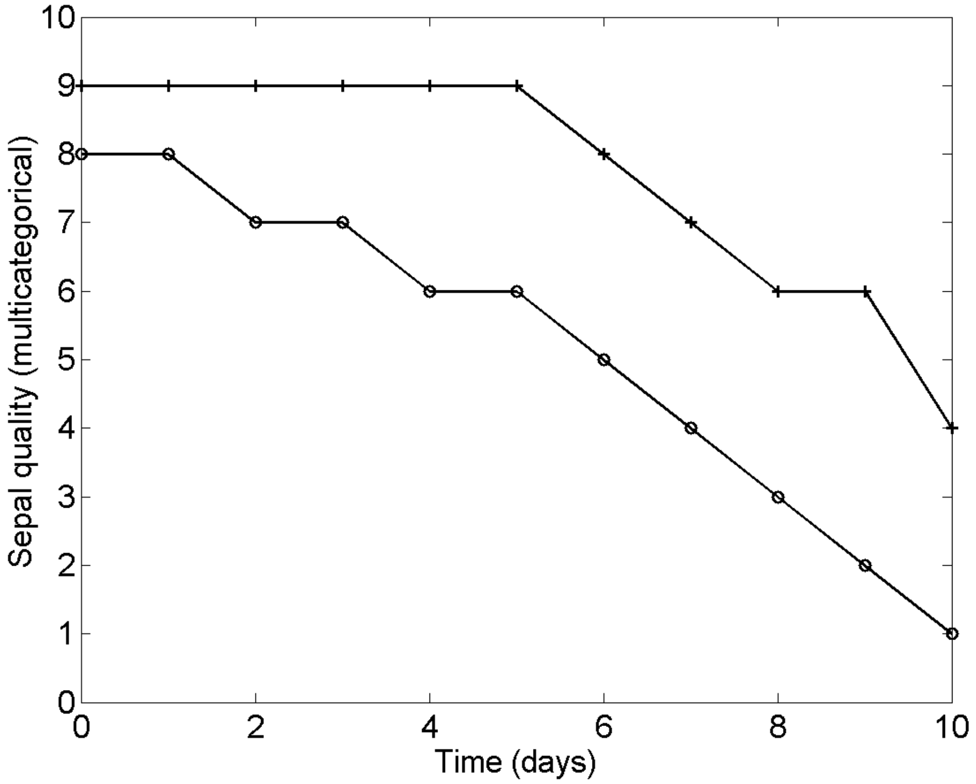
Sepal quality (multicategorical)
0
1
2
3
4
5
6
7
8
9
10
0
2
4
6
8
10
Time (days)

Figure 6