A study to determine compound(s) that lengthens the number of days the fresh cut rose flowers stays fresh

Project: Multivariate and Hierarchical Data (3565)
Project Discovering Associations
2018-2019

1st year Master of Statistics
Hasselt University

Group 3 members:

Brigitte Umutoni (1850424)
Henry Ojo-Kolawole (1850091)
Wara Alfa Syukrilla (1747117)
Bernard Osang'ir (1849815)

Submission Date: 5 June, 2019

Lecturers:

Prof. Dr. Geert Molenberghs
Prof. Dr. Geert Verbeke
Prof. Dr. Luc Bijnens
Dr. Martin Otava
Dr. Steven Abrams
Dr. Yannick Vandendijck
Dr. Tapiwa Ganyani

Abstract

This study was conducted to explore the strength of fourteen compounds in prolonging the freshness of fresh-cut roses when compared to distilled water where three outcomes were investigated in this research, namely; count outcome (number of days a rose flower stayed fresh), binary outcome (*'yes'* when flower was still fresh or *'no'* when flower was considered dead), and Gaussian response (width of the flower measured in centimeters each day). For analysis phase, Generalized Linear Mixed Models was used to investigate the Poisson count data, Generalized Estimating Equations was the method used to investigate binary non-Gaussian outcome and Linear Mixed Model was used to investigate the Gaussian response. Irrespective of the model, four compounds namely; Zest of Zen, Oil of John's son, Essence of epiphaneia and Beerse Brew were consistently more effective in preserving the freshness of fresh cut rose flowers compared to distilled water.

Key Words: Evolution; Gaussian; GEE; Longitudinal; Random Effect

1 Introduction

Flowers are the attractive part of all flowering plants (Angiosperm) which are most often colourful and comes with variant smells (Batygina. 2019). They exist in diverse species and varieties which makes them bloom with respect to the environment they are cultivated. Flowers such as roses have long been admired and used by humans to bring beauty to their environment and can also be used as objects of romance, ritual, religion, medicine and as a source of food (Dutton, 2009). However, preserving the freshness of a cut rose is of interest to us. According (van Meeterena, 2009), fresh cut flowers will only last in the vase for about 10 days therefore, proper post-harvest care of cut flowers is essential for maintaining high quality and a longevity through oral preservatives and other additives.

This study aimed at developing a solution to Jean Baptiste who wants to marry a woman from La Rochelle France. However, depending on the weather situation, the voyage from Azores to France takes at least 8 and maximum 22 days. From his misfortune, it was then believed that water cannot preserve the tulip freshness for long until it reaches the ultimate destination. Therefore, Jean Baptiste asked the scientist to conduct an experiment to explore another compound which will maintain the freshness of the tulip flower longer than water. In this study, fifteen compounds were provided and an observational study design was conducted to assess the ability of fourteen compounds in prolonging the freshness of cut roses when compared to distilled water(the control). The fourteen compounds are Apathic Acid, Beerse Brew, Concentrate of Caduceus, Distillate of Discovery, Essence of Epiphaneia, Four in December, Granules of Geheref, KarHamel Mooh, Lucifers Liquid, Noosperol, Oil of John's Son, Powder of Perlimpinpin, Spirit of Scienza, and Zest of Zen. Due to the constraints which indicate that the availability of black tulip flowers is limited, roses were used with clear agreement from Jean Baptiste that even though the results will suggest the best compound that can be used, this may not be a good representation of the effectiveness of this compounds in preserving the tulip flowers because the tulip and rose flowers originate from different families of the flowering plants.

This report is structured in 4 sections. The first section gives introduction and background of study. Section 2 describes the experimental design, sample size calculation, data collection and description of methods that will be used in data analysis. Section 3 shows result of the analysis. Section 4 provides discussions on the study findings and recommendation.

2 Methods and Materials

2.1 Experimental Design and Data collection

Two experiments were conducted on fresh cut roses to explore the capability of 14 compounds to maintain freshness of the cut roses. In the first experiment number of days each compound kept the fresh cut rose was recorded. Two types of roses namely Floribunda and Hybrid Tea were collected from 10 subplots within two gardens. 3600 roses were cut with the help of 3 raters and the roses were put into water immediately after being cut. These roses were then assigned randomly to the 15 compounds (including water) of the same concentration and volume, with water as the control group for the experiment. The experiment was conducted in a laboratory supplied with constant humidity and temperature, it was done by critically observing of the freshness of the roses everyday until the last rose was no longer fresh. The criteria used for determining whether the rose was not-fresh or otherwise was by checking whether the rose has any form of bent stem, color change, or loose firmness. This were the definition for the stopping rule. During the experiment no missing data was observed.

The second experiment recorded the width (cm) of the roses every day from day 0 up to day 20. The two varieties of roses (the experimental units) were randomly selected from 18 subplots of 2 gardens (Southern and Northern Gardens). Within each subplot 10 rose were cut by a single operator resulting into 180 roses in total. This roses were used to assess the effect of the fifteen compounds on the rose width. Therefore, the 180 fresh cut roses were randomly assigned to compounds (water as control compound of the experiment) which were placed under homogeneous and constant environmental climate. A trained rater was in-charge of observing and recording the width of each rose everyday for 21 days including the initial measurement on day 0. Moreover, no missing data was observed during the experiment.

2.2 Sample size calculation

A pilot study with a sample size 10 roses from different subplots was conducted using Distilled water as a preservative compound. The outcome recorded was the number of days the rose stayed fresh. The mean days from the pilot experimental result was used as the response variable in estimating the sample size. For the sample size determination, the null hypothesis of no difference in mean days of preserving roses between Distilled water and other compounds and the alternative hypothesis of at least one of the compounds differs in mean days of preserving roses from distilled water was formulated and tested. Since 14 multiple pairwise comparison was performed, a Bonferroni correction was used to correct the global significance level of $\alpha = 0.05$ by dividing the α by 14. The simulation procedure was then propelled while observing the effect size (δ) of 1 day and adjusting the sample size until the desired power of 80% was achieved. The appropriate sample size calculated was 240 roses per compound.

3 Statistical Methodology

The data obtained from the two experiments were explored and descriptive statistics conducted. The representation of the exploratory analysis were made using the descriptive summary tables, evolution line graphs and box plot. For further in-depth analysis, three models were fitted to the data from the two experiments.
The data with the Gaussian response which was the width of the rose observed for 21 days, was analyzed using Linear Mixed Model. For the data set with Non-Gaussian response with the number of days a particular rose stayed fresh was analysed using Generalized Linear Mixed Model(GLMM).
The binary deterministic response for freshness ($Y = 1$ if the rose was fresh and $Y = 0$ if it was no longer fresh on the observed day) was analyzed using Generalized Estimating Equations (GEE) approach.

3.1 Gaussian response

3.1.1 Linear Mixed Model

A linear mixed model is a random effects model used to describe the mean evolution of each rose separately. According to Weiss (2005) measurements in longitudinal data are collected repeatedly over time on each particular subject. Weiss also alluded that there exist a temporal ordering of the measurements closer in time within a subject and these measurements are likely to be more similar than observations farther apart from time. Usually, longitudinal data assume natural hierarchical/multilevel structure, with observations at the bottom level nested/clustered within subjects at the top level (Weiss, 2005). Roses were nested within subplots and measurements nested within roses. Therefore, width measurements within rose are correlated as well as roses from the same subplot. The model was fitted with random intercept for subplot to capture variability induced by subplot, random intercept for roses to capture the variability due to different starting points between roses and random slope for roses to accommodate the variability in evolution between roses which is not explained by the covariates included in the model (Verbeke and Molenberghs, 2009).

$$Y_{ijk} = (\beta_0 + b_{0i} + b_{0j(i)}) + (\sum_{c=1}^{15} \beta_{1c} X_c + b_{1j(i)}) * t_{ijk} + \sum_{m=1}^{2} \beta_{2m} X_m + \epsilon_{ijk}$$

Where $i = 1, ..., N$, $j = 1, ..., n_i$ and Y_{ijk} is the k^{th} measurement of width of the i^{th} rose cut from j^{th} subplot. This rose of m type ($X_m = 1$ when the rose is of Hybrid type,and $X_m = 0$ when the rose is of Floribunda type), placed in compound c is measured at t_{ijk} time.

- b_{0i} is the deviation of the i^{th} subplot from the average intercept.
- $b_{0j(i)}$ defines the deviation of the j^{th} rose, nested into i^{th} subplot, from the overall intercept
- $b_{1j(i)}$ defines the deviation of the j^{th} rose, nested into i^{th} subplot, from the overall slope.
- β_0 is a common intercept, $\beta_{1,1}, \beta_{1,2}, ..., \beta_{1,15}$ define the average time effect for each of the 15 compounds respectively.
- $b_{0i} \sim N(0, \sigma^2_{subplot})$, $\epsilon_{ijk} \sim N(0, \sigma^2 \mathbf{I})$ and $(b_{0j(i)}, b_{1j(i)})' \sim N(0, \mathbf{D})$,

D is a 2x2 variance co-variance matrix of the random effects for roses:

$$\mathbf{D} = \begin{pmatrix} d_{11} & d_{12} \\ d_{12} & d_{22} \end{pmatrix}$$

Where d_{11} is the variance of the intercepts $b_{0j(i)}$, d_{22} is the variance of the slopes $b_{1j(i)}$ and d_{12} is the co-variance between the intercepts $b_{0j(i)}$ and slopes $b_{1j(i)}$. Therefore the correlation between the intercepts and slopes which is the correlation between rate of change and the starting value of rose in population equals:

$$Corr(b_{0j(i)}, b_{1j(i)}) = \frac{d_{12}}{\sqrt{d_{11}}\sqrt{d_{22}}}$$

3.2 Non-Gaussian: Poisson data

This part of the analysis used the count response variable Y which represents the number of days the rose stayed fresh in a certain compound. This response had a Poisson distribution which in general has mean $E[Y]$ and variance $Var[Y]$ equal to (λ). Classical regression approach is not always an appropriate approach when the correlation between outcomes is suspected to be in existence therefore Generalized Linear Mixed model have been employed to account for intra-cluster correlations or corrects for over dispersion.

3.2.1 Generalized Linear Mixed Model

Generalized Linear Mixed Model is a subject (cluster) specific model. This modelling technique is equally the same as fitting Poisson regression model taking into account fixed and random effects b_i in the model. Moreover, the response variable (mean number of days the rose stays fresh) assumed to be conditionally independent, given the covariates (compound,rater and type of rose) and the random effects are in the model. $(Y_{ijckm}|b_i, X_1, X_2, X_3) \sim Poisson(\lambda_{ijckm})$ (Verbeke and Molenberghs, 2005). The model is given by

$$log(\lambda_{ijckm}) = \beta_0 + b_i + \sum_{c=1}^{15} \beta_{1c} X_{ijc} + \sum_{k=1}^{3} \beta_{2k} X_{ijk} + \sum_{m=1}^{2} \beta_{3m} X_{ijm}$$

Where $i = 1, .., 10$, $j = 1, ..., n_i$, $c = 1, 2, ..., 15$, $k = 1, 2, 3$, $m = 1, 2$. λ_{ijckm} is the mean number of days of j^{th} rose of m^{th} type cut from i^{th} subplot stays fresh. This particular rose is placed in compound c and observed daily by rater k. b_i is the random effect of subplots and assumed to be normally distributed $b_i \sim N(0, \sigma^2_{subplot})$

- $\sigma^2_{subplot}$ represents the variability between subplots

3.3 Non-Gaussian: Binary response

Generalized Estimating Equations is a method for doing inference if marginal population-averaged evolution inference is of interest, (Verbeke and Molenberghs, 2005). According to Molenberghs and Verbeke, 2009, GEE is frequently preferred for non-Gaussian outcomes when there is need to avoid the higher order correlations while the two way correlation is allowed to be misspecified. To note, even when the correlation structure is misspecified, GEE produces consistent and asymptotically normal regression parameters. Following this analysis with its face on the marginal probabilities, the response variable representing the number of days a certain rose stayed fresh in a particular compound was re-expressed as a binary response, where $Y = 1$ if the rose was still fresh and $Y = 0$ if the rose was no longer fresh for each specific day until the rose was considered dead. Therefore, the analysis is aimed at obtaining the marginal probability of the rose to stay fresh. The logit link was used to link the probabilities and the effects of these covariates.

$$logit[P(Y_{icklm} = 1)] = \beta_0 + \sum_{c=1}^{15} \beta_{1c} X_{ijc} * t_{ijl} + \sum_{k=1}^{3} \beta_{2k} X_{ijk} + \sum_{m=1}^{2} \beta_{3m} X_{ijm}$$

$i = 1, ...180;\ c = 1, ...15;\ k = 1, ..., 24,\ k = 1, 2, 3;\ m = 1, 2$
Where; $P(Y_{icklm} = 1)$ is the probability for i^{th} flower of m^{th} type to stays fresh after cut by k^{th} rater and placed in c^{th} compound measured in l^{th} day.
The predictive probability curves is produced to give an insight on how different compounds have different effects to influence the roses on their probabilities to stay fresh.

3.4 Multiple comparison

In order to compare the strength of different compounds with respect to water, multiple comparison was employed. In addition, multiple comparison was extended to compare the effect of the potential compounds selected. This step was performed after fitting the model for each type of data set using CONTRAST and ESTIMATE Statement in SAS.

4 Result

4.1 Explanatory Data Analysis

To get an insight from the data, descriptive summary statistics was conducted. In this trajectory, tables as well as graphs and box-plots were used. Table 1 describes the summary measures. Picking any time, that is, day 15, it can be shown that Compounds 3(Beerse

Brew), 6(Essence of Epiphaneia), 12(Oil of John's son), 15(Zest of zen) have the have smallest mean width below 10cm as compared to distilled water, that is, compound 1.

Table 1: Average width(cm) of the rose per compound at day 15

No.	Compound	N	Mean	Std Dev	Min.	Max.
1	Distilled water	12	11.3	2.6	4.2	14.1
2	Apathic acid	12	12.8	3.2	4.6	16.5
3	Beerse Brew	12	9.5	1.9	4.3	11.8
4	Concentrate of Caduceus	12	13.7	3.6	4.3	17.5
5	Distillate of discovery	12	11.0	2.6	4.4	14.1
6	Essence of epiphaneia	12	9.2	1.8	4.5	11.8
7	Four in December	12	11.8	2.81	4.4	14.6
8	Granules of Geheref	12	12.8	3.2	4.4	16.2
9	Kar Hamel Mooh	12	11.4	2.7	4.5	14.9
10	Lucifers liquid	12	13.1	3.3	4.4	16.5
11	Noospero	12	10.5	2.3	4.5	13.2
12	Oil of John's son	12	7.6	1.2	4.5	9.3
13	Powder of perlimpinpin	12	9.4	1.9	4.7	11.9
14	Spirit of scienza	12	11.4	2.6	4.4	14.2
15	Zest of zen	12	8.5	1.5	4.3	10.1

Figure 1 below, displays the individual profile of the 180 roses, it can be seen that these lines which represented the individual evolution seem to diverge from slightly different starting points, that is, at the intercept. Similarly, there was evidence of variability between the individual roses slopes. Following this manifestation, there was indeed a suggestive position to include random intercept and random slope accordingly for the roses. Furthermore, Figure 2, a plot for the average evolution for the whole roses population per compound, indicated an overall increase in the average width of the roses per compound over time. The compounds that had a wider average width of the roses at the beginning consistently stayed higher than other compounds while compounds that had a small average width of the roses at the beginning ended with the small record. However, at this point it was still difficult to decide what was significance of this difference. The widening of the lines which is a dominant trend gave a clear indication of the influence of compounds' effect on the width of the roses.

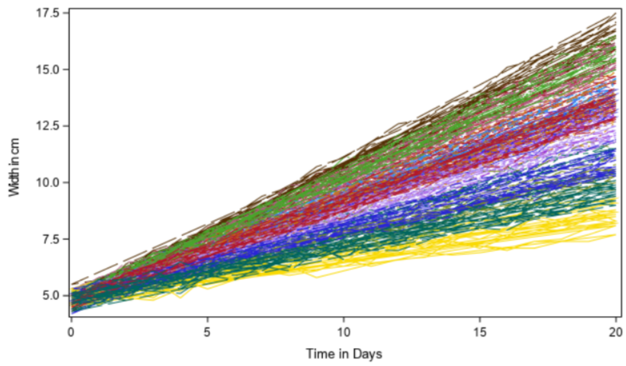


Figure 1: Evolution profile specific for each flower

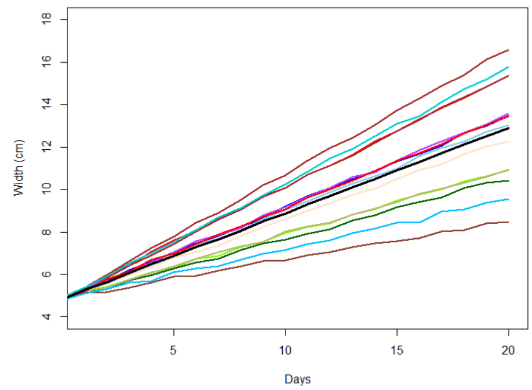


Figure 2: Average evolution of rose per compound

From the second experiment, the summary statistics was shown in Table 2. The average number of days different compounds preserved the freshness of the roses was obtained. Generally it can be seen that then minimum and the maximum number of days the rose stayed fresh was 1 day and 24 days respectively. Compounds 6 (Essence of Epiphaneia), 12 (Oil of Johns son), 13 (Powder of Perilimpinpin), 14 (Spirit of Scienza) and 15 (Zest of Zen) had a relatively high average number of days above 14 days which means on average these compounds preserved the rose longer than other compounds.

Table 2: Summary of count data

No.	Compound	N	Mean	Std Dev	Min	Max
1	Distilled water	240	12.63	5.88	2	24
2	Apathic Acid	240	12.20	5.32	1	24
3	Beerse Brew	240	14.23	5.86	2	24
4	Concentrate of Caduceus	240	10.31	5.68	1	24
5	Distillate of discovery	240	12.82	5.98	1	24
6	Essence of Epiphaneia	240	14.68	5.77	1	24
7	Four in December	240	11.74	5.72	1	24
8	Granules of Geheref	240	11.19	5.37	1	24
9	Kar Hamel Mooh	240	12.77	5.68	1	24
10	Lucifers liquid	240	10.78	5.72	1	24
11	Noosperol	240	13.33	5.97	1	24
12	Oil of John's son	240	15.90	6.10	1	24
13	Powder of Perlimpinpin	240	14.22	5.78	2	24
14	Spirit of Scienza	240	12.17	5.89	1	24
15	Zest of zen	240	16.70	5.85	2	24

The box plot of the average outcome for each compound is shown in Figure 3. The mean number of days the roses stayed fresh in all the compounds seemed not to vary from each other. Nevertheless, compounds 3 (Beerse Brew), 6 (Essence of Epiphaneia), 12 (Oil of John's son) and 15 (Zest of Zen) displayed had a distinguishable higher mean days compared to compound 1 which represents distilled water.

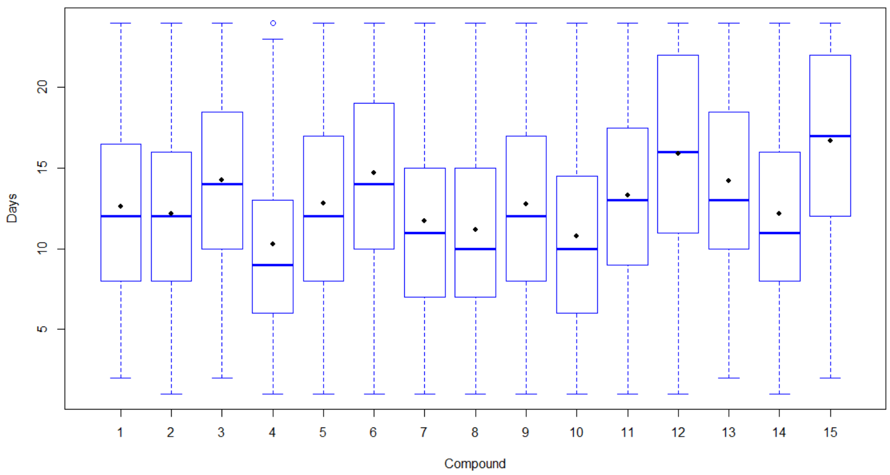


Figure 3: Box plot comparing the number of days the flowers stay fresh across the compounds

4.2 Gaussian Linear Mixed Model

In dealing with Gaussian outcome (Width), linear Mixed model (LMM) was fitted using PROC MIXED in SAS. At exploratory analysis part, there was a suggestion on the need of a random slope and random intercept. Therefore, several LMM were fitted with different combination of random slope and random intercept for the two cluster level (rose and subplot level) and a more parsimonious model would be considered based on the Akaike's Information Criterion (AIC). From Table 3, the model with random intercept and random slope for the flower (see Table) was found to be the best model with the smallest AIC (-344.3) value.

Table 3: AIC for different models

Model	AIC	Random Component
1	1018.1	Intercept
2	-353.9	Slope
3	-344.3	Intercept + Slope

Table 4 shows the type III analysis of fixed effects. There exist a significant effect of compound over time on the width of the rose flower with $p < .0001$ as well as the effect of flower type which is also significant at 5% level of significance.

Table 4: Type III Tests of Fixed Effects LMM

Effect	Num DF	Den DF	F Value	p-value
Time*Compound	15	3420	3227.36	<.0001
Flower type	1	3420	82.37	<.0001
Significance level 0.05				

The variance co-variance matrix (the **D** matrix) for random intercept and random slope for the rose as cluster level. The resulting estimate for the random effects **D** matrix are reported as

$$D = \begin{pmatrix} 0.003043 & -0.00062 \\ -0.00062 & 0.000612 \end{pmatrix}$$

Table 5 contains the parameter estimates, estimated standard errors, and approximate t-tests for all the fixed effects in the model as well as the co-variance parameter estimates. It can be observed that the effects of compounds over time are highly significant (p<.0001). Compounds that had smaller parameter estimate β compared to the parameter estimate for distilled water were considered as potential compounds. Therefore, from Table 5, it can be observed that Beerse brew, Distillate of discovery, Essence of epiphaneia, Noosperol, Oil of John's son, Powder of perlimpinpin, spirit of scienza and zest of zen had the desired characteristics.

The results further showed some impacting variability within rose $\sigma^2_{res} = 0.04231$ and the variability between subplot ($\sigma^2_{subplot} = 0.0111$) which was low. Moreover, a negative correlation between the random slope and the random intercept for the rose nested in the subplot was evidenced ($\rho = -0.4531$) which suggested that the flower which started with a bigger width tended to have a slower rate of increase in width while the flower which started with smaller width tended to have faster rate of increase in width.

Table 5: Compound with time parameter estimates, standard errors

Effect	Parameter	Estimate	Std Error	Pr > $\|t\|$
Intercept	β_0	4.9060	0.02688	<.0001
Time effects:				
Distilled water	β_1	0.4313	0.006770	<.0001
Apathic Acid	β_2	0.5309	0.006770	<.0001
Beerse Brew	β_3	0.3039	0.006770	<.0001
Concentrate of Caduceus	β_4	0.5890	0.006770	<.0001
Distillate of Discovery	β_5	0.4129	0.006770	<.0001
Essence of Epiphaneia	β_6	0.2828	0.006770	<.0001
Four in December	β_7	0.4662	0.006770	<.0001
Granules of Geheref	β_8	0.5259	0.006770	<.0001
Kar Hamel Mooh	β_9	0.4346	0.006770	<.0001
Lucifers liquid	β_{10}	0.5457	0.006770	<.0001
Noosperol	β_{11}	0.3757	0.006770	<.0001
Oil of John's son	β_{12}	0.1863	0.006770	<.0001
Powder of Perlimpinpin	β_{13}	0.3069	0.006770	<.0001
Spirit of Scienza	β_{14}	0.4307	0.006770	<.0001
Zest of Zen	β_{15}	0.2344	0.006770	<.0001
Hybrid Tea (Type2)		-0.1248	0.0138	<.0001
Random subplot variance	$\sigma^2_{subplot}$	0.0111	0.0040	
Covariance of b_i				
Var ($b_{0j(i)}$)	d_{11}	0.003043	0.0012	
Var ($b_{1j(i)}$)	d_{22}	0.000612	0.0001	
Cov($b_{0j(i)}$,($b_{1j(i)}$)	$d_{12} = d_{21}$	-0.00062	0.0002	
Residual variance				
var(ε_{ijk})	σ^2_{res}	0.04231	0.0010	

Significance level 0.05

The contrast estimates between 14 compound and distilled water was performed, there was an indication that the 14 parameter estimates were indeed having different evolution patterns as compared to water. The compounds that had a negative contrast with respect to water were desirable, which indicated that it had less effect on the width of the roses over time. The estimated contrast value, estimated standard errors, approximate ttests were reported in Table 6.

Table 6: Contrast estimates

Compound	Estimate	Std. Error	Pr > $\|t\|$
Contrast estimates between the Compounds with Distilled			
Apathic Acid	0.0995	0.0095	<.0001
Beerse Brew	-0.1275	0.0095	<.0001
Concentrate of Caduceus	0.1577	0.0095	<.0001
Distillate of Discovery	-0.0184	0.0095	0.0531
Essence of Epiphaneia	-0.1485	0.0095	<.0001
Four in December	0.0349	0.0095	0.0002
Granules of Geheref	0.0945	0.0095	<.0001
KarHamel Mooh	0.0033	0.0095	0.7285
Lucifers Liquid	0.1144	0.0095	<.0001
Noosperol	-0.0557	0.0095	<.0001
Oil of Johns son	-0.2451	0.0095	<.0001
Powder of perlimpinpin	-0.1245	0.0095	<.0001
Spirit of Scienza	-0.0006	0.0095	0.9490
Zest of zen	-0.1969	0.0095	<.0001
Contrast estimates between the potential compounds			
Beerse Brew vs Essence of Epiphaneia	0.0210	0.0095	0.0263
Beerse Brew vs Oil of John's son	0.1176	0.0095	<.0001
Beerse Brew vs Zent of Zen	0.0695	0.0095	<0001
Essence of Epiphaneia vs Oil of John's son	0.0965	0.0095	<0001
Essence of Epiphaneia vs Zent of zen	0.0484	0.0095	<0001
Oil of John's son vs Zent of Zen	-0.0481	0.0095	<0001

From Table 6 above, multiple comparison between the potential compounds exhibiting the desired property with respect to water was conducted. The results suggested that Zest of Zen as the best compound followed by Essence of Epithaneia, Beerse Brew and Oil of John's son.

4.3 Non-Gaussian: Poisson data

4.3.1 Generalized Linear Mixed Model

After fitting the Generalized Linear mixed model (GLMM), the significance of covariates were checked first where Compound, Rater and Rose Type were found to be statistically significant. Table 8 below illustrated the significance test of the covariates.

Table 7: Type III Tests of Fixed Effects

Effect	Num DF	Den DF	F Value	Pr > F
Compound	14	3573	58.63	$< .0001$
Rater	2	3573	201.27	$< .0001$
RoseType	1	3573	20.20	$< .0001$

Table 8, summarizes the parameter estimates for the compounds. As observed, the average number of days a rose can stay fresh in Distilled water have been estimated to be $exp(2.4199) = 11.2$ days. However, since the interest was on the compounds which preserve the rose more days than Distilled water, compounds with larger estimated mean days were selected, these include; Zest of Zen, Oil of John's son, Essence of Epiphaneia, Beerse Brew, Powder of Perlimpinpin with the estimated mean days of $exp(2.6925) = 14.8$ days, $exp(2.6588) = 14.3$ days, $exp(2.5644) = 12.99$ days, $exp(2.5443) = 12.7$ days and $exp(2.5360) = 12.6$ days Therefore multiple pairwise comparison of each compound with Distilled water was performed and compounds that had significant positive difference in mean days were selected for further comparison. In addition, this model also considered random intercepts with the variance $\sigma^2_{subplot} = 0.03419$. This parameter represents variability or heterogeneity between the subplots.

Table 8: Fixed effects estimates

Effect	Estimate	Std. Error	$Pr > \|t\|$
Distilled water	2.4199	0.06180	$< .0001$
Apathic Acid	2.3962	0.06186	$< .0001$
Beerse Brew	2.5443	0.06147	$< .0001$
Conc of Caducues	2.2272	0.06236	$< .0001$
Distillate of Discovry	2.4383	0.06175	$< .0001$
Essence of Epiphaneia	2.5644	0.06144	$< .0001$
Four in December	2.3467	0.06205	$< .0001$
Granules of Geheref	2.3090	0.06214	$< .0001$
Kar-Hamel Mooh	2.4250	0.06182	$< .0001$
Lucifer's Liquid	2.2596	0.06231	$< .0001$
Noospherol	2.4781	0.06169	$< .0001$
Oil of John's son	2.6588	0.06120	$< .0001$
Powder of Perlimpinpin	2.5360	0.06150	$< .0001$
Spirit of Scienza	2.3898	0.06193	$< .0001$
Zest of Zen	2.6925	0.06119	$< .0001$
Rater 2 vs 1	0.01972	0.01169	0.0918
Rater 3 vs 1	0.2018	0.01126	$< .0001$
RoseType 2 vs 1	0.04150	0.009233	$< .0001$
Variance(Subplot)	0.0342	0.0154	

From Table 9 below, compounds which had significant positive difference in the mean days a rose stayed fresh were selected as the best compounds that can preserve the roses longer than distilled water: Zest of Zen, Oil of John's son, Essence of Epiphaneia, Beerse-Brew and Powder of Perlimpinpin with $\alpha = 0.0036$. But since the idea was to come up with the best compound among the selected ones, a multiple comparison of all combinations of those compounds was done. The multiple comparisons with 10 possible combinations using $\alpha = 0.005$, indicated that Zest of Zen and Oil of John's son were significantly the best compounds which preserved the rose longer than distilled water.

Table 9: Contrast estimates

Effect	Estimate	Standard Error	$Pr > \|t\|$
Contrast Estimates between Compounds with Distilled water			
Apathic Acid	-0.02366	0.02592	0.3615
Beerse Brew	0.1244	0.02496	$< .0001$
Concentrate of Caduceus	-0.1926	0.02710	$< .0001$
Distillate of Discovery	0.01848	0.02559	0.4703
Essence of Epiphaneia	0.1446	0.02478	$< .0001$
Four in December	-0.07316	0.02618	0.0052
Granules of Geheref	-0.1108	0.02651	$< .0001$
Kar-Hamel Mooh	0.005145	0.02562	0.8409
Lucifer's Liquid	-0.1602	0.02677	$< .0001$
Noosperol	0.05826	0.02535	0.0216
Oil of John's son	0.2389	0.02434	$< .0001$
Powder of Perlimpinpin	0.1161	0.02496	$< .0001$
Spirit of Scienza	-0.03006	0.02594	0.2466
Zest of Zen	0.2726	0.02408	$< .0001$
Contrast estimates between the potential compounds			
Zest of Zen vs Oil of John's son	0.03370	0.02264	0.1367
Zest of Zen vs Essence of Epiphaneia	0.1280	0.02310	$< .0001$
Zest of Zen vs Beerse Brew	0.1482	0.02330	$< .0001$
Zest of Zen vs Powder of Perlimpinpin	0.1565	0.02330	$< .0001$
Oil of John's son vs Essence of Epiphaneia	0.09432	0.02337	$< .0001$
Oil of John's son vs Beerse Brew	0.1145	0.02355	$< .0001$
Oil of John's son vs Powder of Perlimpinpin	0.1228	0.02357	$< .0001$
Essence of Epiphaneia vs Beerse Brew	0.02019	0.02402	0.4007
Essence of Epiphaneia vs Powder of Perlimpinpin	0.02844	0.02402	0.2365
Beerse Brew vs Powder of Perlimpinpin	0.008251	0.02421	0.7332

4.4 Binary GEE (Generalize Estimating Equation)

GEE approach with Auto-regressive(AR) working correlation structure was fitted. An AR working structure was chosen since the study involves time effect on subjects and measurements close in time were expected to be correlated than those more apart in time. Each rose was observed over time until all the roses were not fresh. This experiment lasted for 24^{th} day. Table 10 illustrated that all variables in the model; compounds over time, rater, and rose type were statistically significant. However, main interest in this study is the effect of compound over time on the rose probability being fresh.

Table 10: Type III Tests of Fixed Effects Binary GEE

Source	DF	Chi-Square	Pr > ChiSq
Time*Compound	15	3296.73	<.0001
Rater	2	159.73	<.0001
RoseType	1	8.92	0.0028

The parameter estimates and empirical standard errors were summarized in Table 11. Looking at the effect of each compound over time, all compounds showed significant effect in preserving the rose freshness. Compound(s) with bigger estimate were preferred since they suggested to be good preservative compounds for the roses. Hence, compound 15 (Zest of zen), 12 (Oil of John's son), 3 (Beerse Brew) and 6 (Essence of Ephiphaneia), were considered as the potential compounds

Table 11: Parameter Estimates Binary GEE

Parameter	Estimates	SE	P-Value
Intercept	3.4107	0.068	<.0001
Time effects:			
Distilled water	-0.2954	0.01	<.0001
Apathic acid	-0.3074	0.009	<.0001
Beerse Brew	-0.2596	0.008	<.0001
Concentrate of Caduceus	-0.3492	0.011	<.0001
Distillate of discovery	-0.2862	0.009	<.0001
Esence of Epiphaneia	-0.2552	0.008	<.0001
Four in December	-0.3158	0.01	<.0001
Granules of Geheref	-0.3328	0.011	<.0001
Kar Hamel Mooh	-0.292	0.009	<.0001
Lucifers liquid	-0.3377	0.011	<.0001
Noosperol	-0.2761	0.009	<.0001
Oil of John's son	-0.2262	0.007	<.0001
Powder of Perlimpinpin	-0.2649	0.008	<.0001
Spirit of Scienza	-0.2991	0.01	<.0001
Zest of Zen	-0.2196	0.007	<.0001
Rater 2	0.0375	0.069	<.0001
Rater 3	0.877	0.074	<.0001
Rose Type 2	0.1772	0.058	0.0019

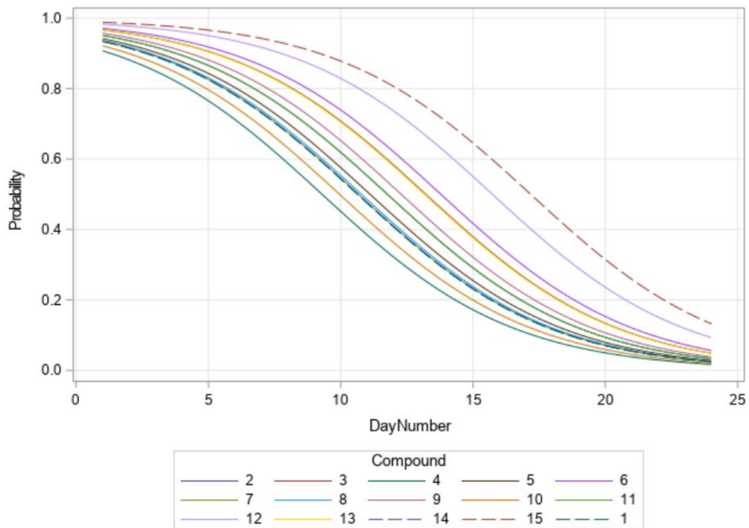
A multiple pairwise comparison test was done to see how compounds differ from each others in prolonging the rose freshness. Based on Table 12, compound 3(Beerse Brew), 6(Essence of Epiphaneia), 12(Oil of Johns sons), and 15(Zest of zen) were significantly different from distilled water with positive estimates. However since goal is to find the best compound among the selected ones, a multiple comparison of all combinations of those compounds was done. The multiple pairwaise comparisons with 6 possible combinations using $\alpha = 0.0083$, indicated that compound 12 (Oil of John's sons) and compound 15 (Zest of zen) were selected as the best compounds to preserve the rose longer than water. The estimates for Oil of John son compound is -0.2263 which implies that as time increases by one day, the probability of a rose to stay fresh compared to not fresh when placed in Oil of Johns son decreases by 20%. The probability of a rose to stay fresh (against not fresh) when placed in Oil of John son compared to Zest of zen at a certain time is accounted for $exp(-0.0066 * time)$.

Table 12: Estimates of Contrast using GEE

Compound	Estimate	SE	P-Value
Estimates of contrast between Compound and Distilled Water			
Apathic acid	-0.012	0.0124	0.3332
Beerse Brew	0.0358	0.0115	0.0019
Concentrate of Caduceus	-0.0538	0.0139	0.0001
Distillate of discovery	0.0092	0.0124	0.4577
Essence of epiphaneia	0.0401	0.0114	0.0004
Four in December	-0.0205	0.0129	0.1123
Granules of Geheref	-0.0374	0.0133	0.0051
Kar Hamel Mooh	0.0034	0.012	0.7795
Lucifers liquid	-0.0423	0.0136	0.0018
Noosperol	0.0192	0.012	0.11
Oil of John's son	0.0692	0.011	<.0001
Powder of perlimpinpin	0.0305	0.0117	0.0093
Spirit of scienza	-0.0038	0.0126	0.7655
Zest of zen	0.0758	0.0108	<.0001
Contrast estimates between the potential compounds			
Oil of John's son vs Zest of zen	-0.0066	0.0088	0.4501
Beerse Brew vs Zest of zen	-0.04	0.0094	<.0001
Essence of epiphaneia vs Zest of zen	-0.0356	0.0093	0.0001
Beerse Brew vs Oil of John's son	-0.0334	0.0096	0.0005
Essence of epiphaneia vs Oil of John's son	-0.3052	0.0129	<.0001
Beerse Brew vs Essence of epiphaneia	-0.0044	0.01	0.663

The predictive probability plot is shown in Figure 4. From the plot, Compound 12 (Oil of John's son) and Compound 15 (Zest of zen) respectively were the compounds with first and second slowest rate of change in probability of a rose to stay fresh.

Figure 4: Probability Plot of Being Fresh for Each Compound by Time



5 Discussion and Conclusion

In this experimental analysis, all the models consistently identified four compounds which were effective in preserving the freshness of the roses than distilled water amongst other compounds regardless of the outcome and the models used in the analysis. These compounds were Zest of zen, Oils of John's son, Essence of epiphaneia, Beerse Brew. The two Non-Gaussian models GLMM for Poisson and GEE for Binary outcome ranked Zest of zen and Oil of John's son as best compounds with model considering binary response reversing the order starting with Oil of John's son as the best compound. On the other hand, the Gaussian model also ranked Zest of zen as the best. In conclusion, Zest of zen, Oil of John's son, Essence of epiphaneia and Beerse Brew were consistently more effective than distilled water in preservation of the freshness of the rose. Therefore, these compounds are recommended as best compounds in prolonging the freshness of fresh cut rose flower. Since Black tulip flowers were not utilized in this experiment the decision to choose among this compounds in transporting fresh cut black tulips from the Azores to France and present them to his love whilst still fresh stays with Jean Baptiste.

References

[1] Verbeke G. and Molenberghs G. (2009). Linear Mixed Models for Longitudinal Data,

[2] Aerts, M., Molenberghs, G., Geys, H. and Ryan, L.M., (2002). Topics in modelling of clustered data. Chapman and Hall/CRC.

[3] Agresti, Alan. (2013). Categorical Data Analysis 3rd Edition. New-Jersey: Wiley.

[4] Molenberghs G, Verbeke G. (2005). Models for Discrete Longitudinal Data, 1 edn.New York: Springer-Verlag New York.

[5] Gaecki, A., Burzykowski, T. (2013). Linear mixed-effects models using R: A step-by-step approach. Springer Science Business Media.

[6] Robert E. Weiss. (2005). Modeling Longitudinal Data-Springer-Verlag New York.

[7] Van Meeteren, U. (2009). Causes of quality loss of cut flowers-a critical analysis of postharvest treatments.

[8] Batygina, T. B. (Ed.). (2019). Embryology of Flowering Plants: Terminology and Concepts, Vol. 3: Reproductive Systems. CRC Press.

[9] Dutton, D. (2009). The art instinct: beauty, pleasure, human evolution. Oxford University Press, USA.

Appendix - R/SAS code

```
# 1. Sample size determination - R
N = 240
mean_dwater = 9.3
mean_comp = 10.3
alpha = 0.05/(14)
NumberSimulation = 1000
pval = numeric(NumberSimulation)
set.seed(1234)
for (i in 1:NumberSimulation){
Distilled_water= rpois(N, lambda = mean_dwater)
Comp= rpois(N, lambda = mean_comp)
simData <- data.frame(resp = c(water, Comp),Comp = rep(c(0,1), each = N))
pval[i] <- summary(glm(response ~ Comp, data = simData,
                         family=poisson()))$coeff["Comp", "Pr(>|z|)"]}
sum(pval/2 < alpha)/NumberSimulation # Estimate power

# 2. Generalized Linear Mixed Model for Poisson Data - SAS
proc glimmix data=count_data1 method=quad;
title 'GLMM poisson model with random intercept';
```

```
class Rater Compound RoseType BushID;
model Days =Compound Rater RoseType/ link=log dist=poisson solution noint;
random intercept / subject=BushID g ;
*Examples of contrast and estimate of compound against water applied
contrast 'C1 vs C2 Compound -1 1 0 0 0 0 0 0 0 0 0 0 0 0 0;
estimate 'C1 vs C3' Compound -1 0 1 0 0 0 0 0 0 0 0 0 0 0 0
/
 ;run;

# 2. Generalized Estimating Equation for Binary Data
proc genmod data=pda.binary ;
title 'GEE binary Prob being Fresh';
class BushID Compound Rater (ref='1') RoseType (ref='1') FlowerID;
model Outcome (ref='0')=  Compound*Time Rater RoseType
/dist=binomial link=logit type3 ;
repeated subject=FlowerID(BushID) / type=AR(1) covb corrw ;
*Examples of contrast and estimate of compound against water applied;
contrast 'Compound 2' Compound*Time -1 1 0 0 0 0 0 0 0 0 0 0 0 0 0/;
estimate 'Compound 2' Compound*Time -1 1 0 0 0 0 0 0 0 0 0 0 0 0 0
/;run;

# 3. Linear Mixed Model for Gaussian outcome
proc mixed data=m.Tulip1a;
title 'LMM for Gaussian'
class Compound Type (ref='1') Subplot Flower_index;
model Width=  Compound*Time Type/ solution residual outpm=outpred outp=pred;
random intercept/ subject=Subplot;
random intercept Time / type=un subject=Flower_index(Subplot) g gcorr;run;
```