Acta Scientiarum. Technology
ISSN: 1806-2563
eduem@uem.br
Universidade Estadual de Maringá
Brasil

Guimarães Brighenti, Carla; Serpa Miranda, Yola; Resende, Mariana; Cirillo, Marcelo Angelo; Magno Brighenti, Deodoro

Generalized estimating equations to evaluate the dependence of lifetime in confined Africanized bees

Acta Scientiarum. Technology, vol. 38, núm. 1, enero-marzo, 2016, pp. 99-105
Universidade Estadual de Maringá
Maringá, Brasil

Acta Scientiarum
http://www.uem.br/acta
ISSN printed: 1806-2563
ISSN on-line: 1807-8664
Doi: 10.4025/actascitechnol.v38i1.26633

Generalized estimating equations to evaluate the dependence of lifetime in confined Africanized bees

Carla Guimarães Brighenti[1*], Yola Serpa Miranda[1], Mariana Resende[2], Marcelo Angelo Cirillo[2] and Deodoro Magno Brighenti[1]

*[1]Universidade Federal de São João del Rei, Av. Visconde do Rio Preto, s/n, 36301-360, São João del Rei, Minas Gerais, Brazil. [2]Universidade Federal de Lavras, Lavras, Minas Gerais, Brazil. *Author for correspondence. E-mail: carlabrighenti@ufsj.edu.br*

ABSTRACT. Research with honeybee needs experimentation in cages. In laboratory bioassays to assess the lifetime of *Apis mellifera*, each sampling unit contained a group of bees confined in a cage. Since current is a longitudinal analysis to assess the effect of time dependence on the survival of bees fed on different diets, generalized estimating equations applied to the logit model were used to assess the viability of incorporating correlation structures that related the survival of bees observed in each cage over time, so that the mortality of a given number of bees might influence the mortality of the others. Results showed that there was time dependence between the observations of bees inside the cages, and the working correlation matrix adjusted by longitudinal analysis was 'Non-stationary-M-dependent' in which the number of periods of dependence must be specified. Correlations are not the same for all observations of all periods. The solution with granulated sugar and the solution with granulated sugar enriched with lemon juice proved to have the lowest level of mortality during the assessed periods. Assay showed that bees tended to die during daytime.

Keywords: *Apis*, dependence, longevity, logit.

Equações de estimação generalizada para avaliação da dependência do tempo de vida entre abelhas africanizadas confinadas

RESUMO. Os bioensaios laboratoriais para avaliar o tempo de vida de *Apis mellifera* têm algumas particularidades, pois cada unidade amostral contém um grupo de abelhas confinadas em uma gaiola. O presente estudo teve como objetivo realizar uma análise longitudinal para avaliar o efeito da dependência do tempo de vida entre abelhas confinadas alimentadas com diferentes dietas. Foi utilizado o método de estimação de equações generalizadas aplicadas ao modelo logit para avaliar a viabilidade de incorporar estruturas de correlação ao longo do tempo, à sobrevivência de abelhas observadas em cada gaiola, de modo a verificar como a mortalidade de um determinado número de abelhas pode influenciar a mortalidade das outras. Concluiu-se que há dependência do tempo entre as mortalidades de abelhas no interior das gaiolas, e a matriz de correlação de trabalho ajustada em análise longitudinal foi 'Não-estacionária-M-dependente', em que se deve especificar o número de períodos de dependência. Assim, as correlações não são as mesmas para todas as observações. Dentre as dietas testadas, a solução composta por açúcar granulado e a composta por açúcar granulado adicionado com suco de limão mostraram o menor nível de mortalidade durante os períodos avaliados. Observou-se também que as abelhas tendem a morrer durante o dia.

Palavras-chave: *Apis*, dependência, longevidade, logit.

Introduction

A set of experimental protocols is extant for detailed studies of honey bee behavior, including investigation on pesticide and insecticide effects, proposal by Scheiner et al. (2013), so that the behavior of a characteristic of insects throughout an assessment condition or ordered scale, e.g. age or time, could be analyzed. The experimenter uses a nested experimental design in which the 'individual bee' is nested within a random effect (cage)(Pirk et al., 2013). Longitudinal data may be defined as observations collected from one experimental unit throughout repeated measures (Diggle, Heagerty, Liang, & Zeger, 2002).

Longitudinal studies are particularly useful when global or individual variations over time are assessed. Their distinctive characteristic is that repeated observations on the same individual tend to be correlated (Liang & Zeger, 1986). Therefore specific statistical analyses, Generalized Estimating Equations are applied in this case.

In the case of the social insect *A. mellifera*, each experimental unit of a bioassay is generally formed by a group of bees. However, it is reasonable to presume that the mortality of a given number of bees living in the same cage may influence the mortality of the others. Thus, we suggest the use of a statistical analysis model to correlate the observations collected from one experimental unit (cage) over time (Zeger & Liang, 1986). Discrete longitudinal data were modeled by such methods as generalized estimating equations (GEE). Therefore, several structures of time correlation may be assessed to provide a more accurate analysis, so that the inferences related to the model parameters become a proper correlation structure to determine the most appropriate diet.

The study of the lifetime of confined *Apis mellifera* Linnaeus, 1758, is a very complex analysis since it involves several factors of which feeding is one of the most important (Manning, 2001; Brighenti, Brighenti, Cirillo, & Santos, 2010). Due to scanty information on the effect of mortality per cage by proper diet of the *A. mellifera* species from tropical regions, such as Brazil, the diet most strongly associated to increased longevity for the studied species was determined, under laboratory conditions. The diet was a variable affecting the longevity of the bees. Hence, the mortality rates of bees fed on different carbohydrate diets will be assessed by means of generalized estimating equation models (Diggle et al., 2002) to investigate the effect of time dependence associated to average mortality rate of *A. mellifera* bees in each cage over time.

Material and methods

Africanized honey bee workers were obtained from a beehive from the Central Apiary at the Federal University of Lavras, Lavras, Minas Gerais State, Brazil. The possible variation in developmental time between different lineages should be taken into account when designing experiments in which the age of immature specimens or adult is relevant (Williams et al., 2013). Although the age of the bee workers is unknown, all were foragers. In fact, they were captured when they were returning to the hive with the pollen baskets full of pollen or propolis, which evidenced foraging activities and their age bracket. This factor is enough to demonstrate the validity of the statistical method.

Each group of 10 bees was placed in a 15 cm high x 10 wide PVC cage, with a tulle net over the cages and an organza fabric on the bottom. The bees were maintained in a climate controlled chamber at 26 ± 1°C, and 70 ± 10% RH and 12-h photophase. The experiments were conducted in a completely randomized design, with 10 replicates per diet represented by the cages. Thus, the repeated measures were characterized by the number of living bees observed inside each cage during the assessed periods of time. The assessment started 12 hours after the setting up of the experiment and a new count was performed until reaching 312 hours, totaling 26 periods of time. The treatments (diets) with energy supplements comprised: honey, granulated sugar, granulated sugar + tartaric acid 1%, granulated sugar + lemon juice [*Citrus limonia* (L.) Osbeck] 2%, Gludex® (natural sweetener), fructose, candy sugar, molasses (product of sugarcane refining) and sugarcane juice. Control treatment consisted of water. The candy was prepared with honey placed in a sugar well and mixed; candy was then kneaded for some minutes.

The number of living bees is represented by Y_{ijh}, or rather, the response observed in i-cage, i = 1,...k (k = 10), at the j- time observed, j = 1, ..., n (n = 26), for the h-diet, h = 1,2,...,p (p = 7).

So that the global variations found in the survival rate at each time interval of living bees at every time could be assessed, the behavior of the average response was evaluated by a marginal model μ_{ijh}, which, in synthesis, refers to the generalized linear model

$$E(Y_{ijh}) = \frac{\exp(\eta_{ijh})}{1+\exp(\eta_{ijh})},$$

where

$$\log\left(\frac{\mu_{ijh}}{1-\mu_{ijh}}\right) = \eta_{ijh}.$$

which incorporates a correlation structure, where η_{ijh} is a linear predictor.

The correlation structure or specific correlation matrix of n random variables X_1, ..., X_n is the n × n matrix whose i,j entry is the correlation between X_i and X_j time, for longitudinal data. The correlation matrix is symmetric because the correlation between X_i and X_j is the same as the correlation between X_j and X_i.

For most longitudinal data in biological applications, Wang and Carey (2003), Ziegler and Vens (2010) and Vens and Ziegler (2012) showed that the investigators should choose a working correlation structure for both statistical and biological reasons. The statistical criteria for selecting the working correlation structure may be helpful tools to decide the most reasonable structure for the investigators (Gosho, 2014).

The correlation structures considered: 'Independence', 'Exchangeable', 'Unstructured', 'Auto-Regressive (M) (AR-M)', 'Stationary-M-Dependent (stat_M_dep)' and 'Non-Stationary-M-Dependent (non_stat_M_dep)'. GEE package from R Core Team (2014) was used to obtain the structure of the appropriate correlation. Binomial distribution was adopted and the cluster was characterized by each cage.

A chart of the survival rate of living bees compared to the adjustments was prepared. The average survival rate of living bees at every time period was also assessed to detect any possible trend during a given time-period, with the initial time being the time of experiment set up. The survival rate was calculated by proportion of living bees in time n in related to time m, where n = m + 1.

Results and discussion

The analysis of the insects' average lifetime indicated that the diet composed of granulated sugar enriched with lemon juice was the most efficient, with an average survival rate of 166.44 hours. The other diets, excluding the diets composed of water, sugarcane juice and molasses, showed average survival rates between 110 and 140 hours, reported in the Box-plot (Figure 1). The diets composed of water, sugarcane juice and molasses are not recommended as energy supplements since the average survival rate of bees fed on these ingredients was very low (only 34.12 hours).

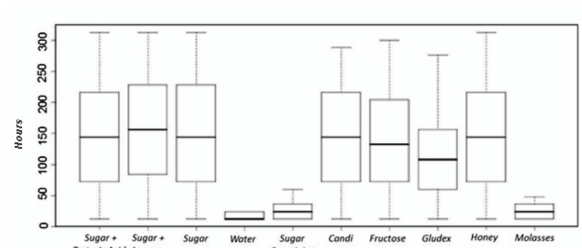

Figure 1. Box Plot by Diet.

Results for the scale factor (Table 1) demonstrated that the 'Non-stationary-M-Dependent' correlation structure was the most adequate because its result was very different from that 'independent' structure, which would indicate absence of time dependence between the observed cages.

Table 1. Estimates of parameters of logit model with different correlation structures.

Correlation Structure	Scale Factor
Independence	$\phi = 1.325$
Exchangeable	$\phi = 1.396$
Unstructured	$\phi = 1.317$
Auto-Regressive	$\phi = 1.323$
Stationary-M-Dependent	$\phi = 1.323$
Non-Stationary-M-Dependent	$\phi = 1.241$

According to this structure, it is not necessary that the correlations $cor(y_{it}, y_{it'})$ are the same for all the observations in one given time-period $|t-t'|$, and then $cor(y_{it}, y_{it'}) = \begin{cases} \alpha_{tt'} & \text{se } |t-t'| \le M \\ 0 & \text{se } |t-t'| > M \end{cases}$, where M is the number of t dependence periods. For M = 1, we have, for example : $\begin{pmatrix} 1 & \alpha_{12} & 0 & 0 \\ \alpha_{12} & 1 & \alpha_{23} & 0 \\ 0 & \alpha_{23} & 1 & \alpha_{34} \\ 0 & 0 & \alpha_{34} & 1 \end{pmatrix}$.

In other words, the time correlation observed inside each cage influenced the mortality rate of the bees in each time. The $\alpha_{tt'}$ terms of the 'Non-stationary-M-Dependent' working correlation matrix are represented in Table 2.

Table 2. Rates of $\alpha_{tt'}$ terms of the working correlation matrix.

Terms	Rates	Terms	Rates	Terms	Rates
α_{12}	0.28	α_{1011}	-0.09	α_{1819}	0.08
α_{23}	0.23	α_{1112}	0.17	α_{1920}	0.04
α_{34}	0.26	α_{1213}	0.10	α_{2021}	0.08
α_{45}	0.04	α_{1314}	-0.09	α_{2122}	0.12
α_{56}	0.06	α_{1415}	-0.11	α_{2223}	-0.14
α_{67}	0.10	α_{1516}	-0.05	α_{2324}	0.06
α_{78}	-0.25	α_{1617}	-0.05	α_{2425}	0.08
α_{89}	-0.31	α_{1718}	-0.10	α_{2526}	0.05
α_{910}	-0.13				

Since the highest absolute rates of $\alpha_{tt'}$ correlation occurred between the first 9 observation time periods, the mortality occurred in this period is more dependent on time.

Selecting an appropriate working correlation structure as an aspect of model selection may improve estimation efficiency, then, the logit model defined the effect of covariates in the marginal expectation of the response variable by:

$$E(Y_{ijh} \mid x_h, t_j) = \frac{\exp(\beta_0 + \beta_h x_h + \gamma t_j)}{1 + \exp(\beta_0 + \beta_h x_h + \gamma t_j)}$$

where β_h is the effect of h-diet (h = 1,...,7) and t_j is the j-time.

Based on these results, the estimations carried out proved to be all significant parameters ($p < 0.05$) (Table 3).

Since estimates of the parameters and robust standard errors were found in all standard errors, the parameter estimates were lower and indicated that the model does not present an over-dispersion effect. Charts were prepared in which the 'Non-stationary-1-Dependent' correlation structure was taken into consideration (Figure 2).

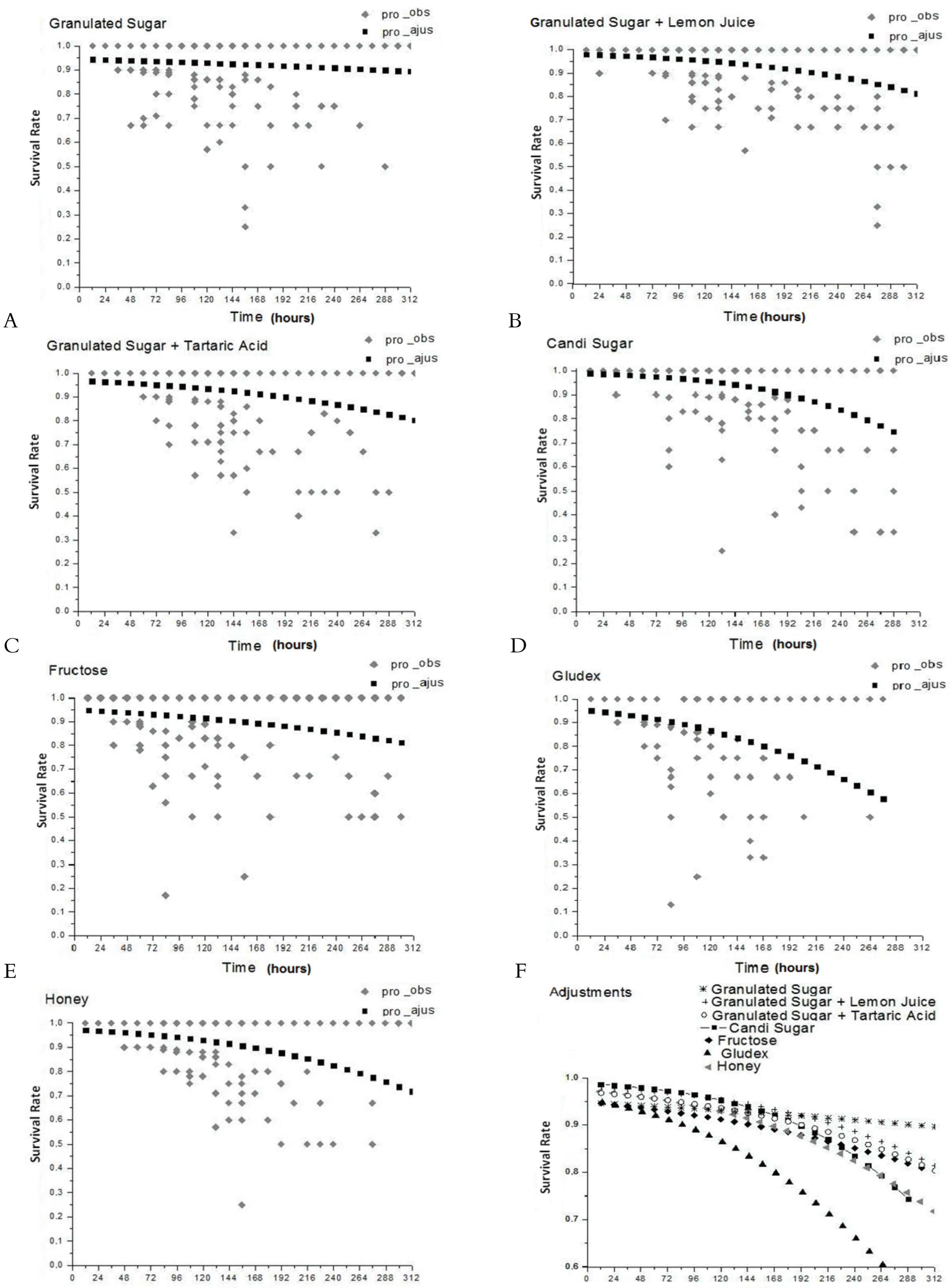

Figure 2. Survival rate of living bees observed (pro_obs) and adjusted (pro_ajus) for each diet according to the time period (A-G) taking into consideration the 'Non-stationary-1-Dependent' correlation structure.

Table 3. Maximum quasi-likelihood estimates of the model parameters logit.

Parameters	Estimates	Robust Standard Error
β_0 (Intercept)	4.403	0.45
β_1 (Candy Sugar)	-	-
β_2 (Granulated Sugar)	-1.528	0.49
β_3 (Granulated Sugar+Lemon Juice)	-0.419	0.50
β_4 (Granulated Sugar + Tartaric Acid)	-0.934	0.49
β_5 (Fructose)	-1.449	0.50
β_6 (Gludex)	-1.339	0.47
β_7 (Honey)	-0.796	0.48
γ (time)	-0.012	< 0.00

Analysis of Figure 2 reveals that, during the first 72 hours, there was practically no difference between the diets, and from this time on there was a sudden reduction in mortality rate for Gludex® diet (2-F).

In the first 120 hours the candy sugar (2-D) diet and the granulated sugar diet enriched with lemon juice (2-B) were the most promising diets. Further, from time period 136, the diet composed of granulated sugar with lemon juice (2-B) was more efficient until time period 204 when it was surpassed by the energy diet composed of granulated sugar. The proportion of living bees fed on this diet was practically stable during all the assessed time periods.

Considering the average ratio of bees surviving over each time interval, mortality peaks at every 24 hours of observation could be detected (Figure 3). The Gludex®-based diet is not appropriate since it caused mortality when compared to the other diets. Consequently, it was not included in this analysis. In the case of diets D and E, there was food rejection after 288 hours which caused a sharp drop in the last time period. However, the remaining proportion of bees was small and did not affect the analysis.

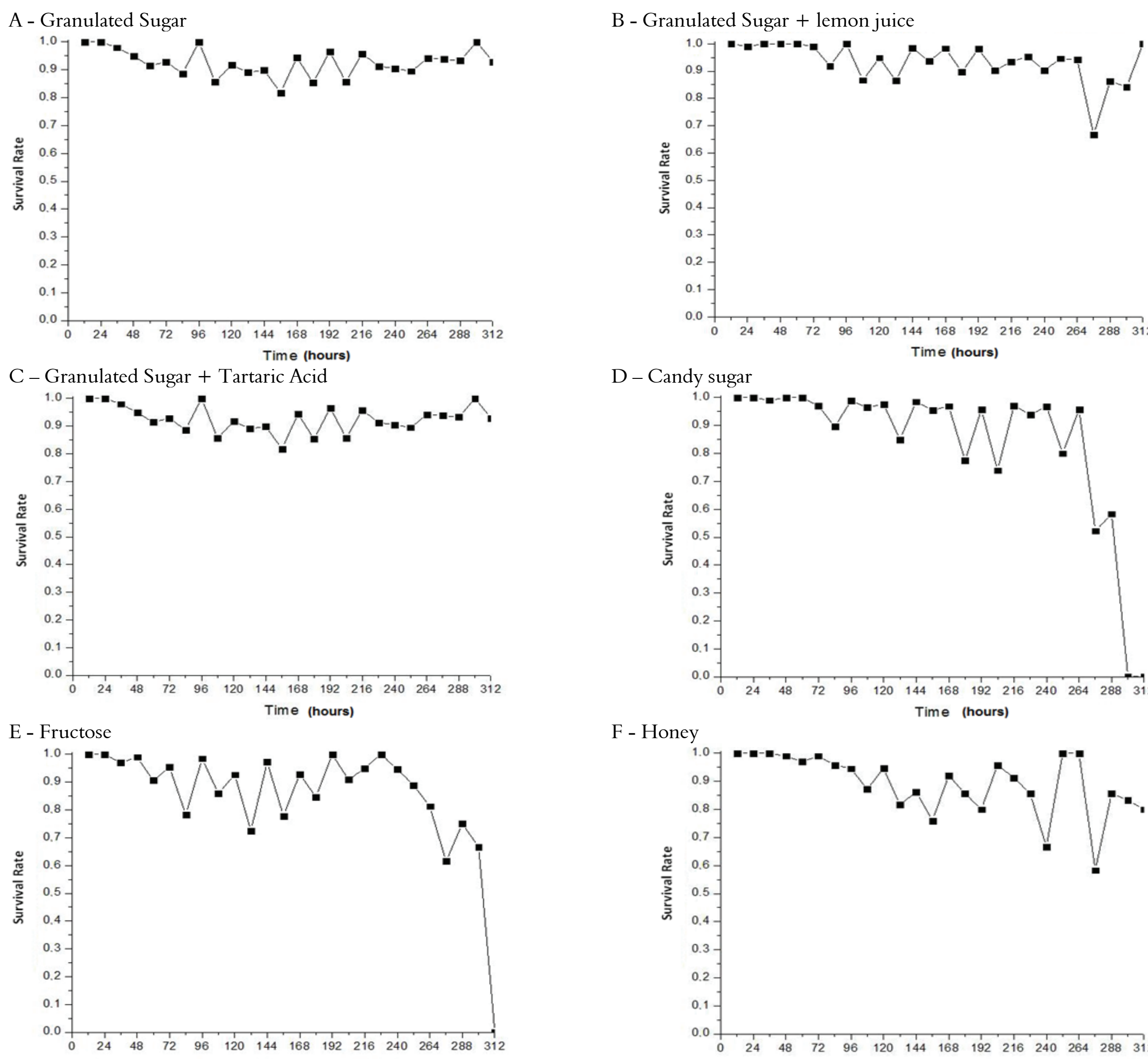

Figure 3. Survival rate at each time interval (hours) for each diet.

Results revealed that mortality peaks adjusted themselves to a circadian rhythm, with the highest mortality rates occurring in daytime periods (Figure 4). According to time in all treatments, the average profile showed the same peak pattern and discards the possibility of including a random component in the model. The highest mortality rate was observed in the collection performed at 7h00 pm, indicating that deaths were more frequent between 7h00 am and 7h00 pm rather than during the night period (7h00 pm to 7h00 am).

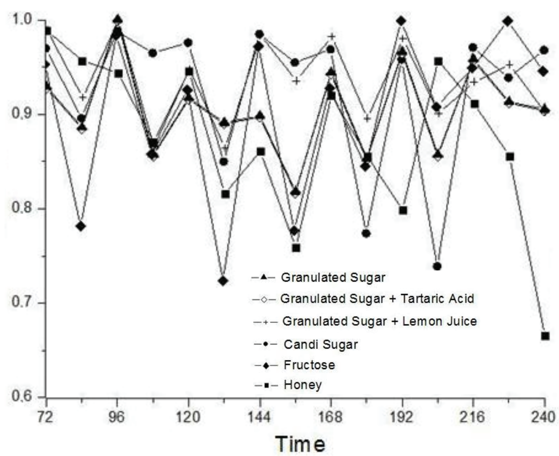

Figure 4. Survival rate at each time interval (hours) for each diet.

The determination of the best diet was inferred by means of longitudinal analysis since the responses observed over time were repeated in each cage. Therefore, the incorporation of a correlation structure evaluated by parameter estimates was justified and aimed at assessing the effect of time dependence on the survival of *A. mellifera* bees in each cage (Brighenti, Cirillo, & Brighenti, 2008).

Based on the above results, it may be observed that the technique of generalized estimating equations has proven to be adequate for assessing the existence of time dependence for entomological experiments, even stressing the daytime mortality peaks. Further, in longitudinal analysis in current study, the 'Non-stationary-1-Dependent' correlation structure was identified and might explain a possible time dependence caused by the mortality of *A. mellifera* confined in the same cages over time. Two different diets according to the time period of the experiment are recommended in laboratory experiments with adult bees: the diet composed of granulated sugar enriched with lemon juice is recommended during the first 192 hours. From this time on, the granulated sugar diet is recommended. The use of these two diets ensures a higher survival rates for bees.

Proni & Macieira (2004) stressed that the organisms have physiological and behavioral differences according to time and the day, and may respond differently to stimuli.

Withrow (1949) introduced the term circadian to characterize rhythms with 24 hour periods, synchronized in 24-hour light-dark cycles. It was also found that the mortality rate of *A. mellifera* is affected by their circadian rhythm.

The most interest aspect concerning the study of the 24-hour rhythm would not be the fact that given activities occur at this interval, but rather that these repetitions might persist in the absence of normal environmental changes (Harker, 1958).

In the case of diets composed of granulated sugar with tartaric acid, fructose and honey, the circadian rhythm does not influence the mortality rate after 240 hours, although the number of surviving bees at that time is already much reduced.

Nevertheless, this is also consistent with the research by Frisch and Koeniger (1994), who observed the social synchronization of the activity rhythms of individual bees in isolation and of colonies of *Apis mellifera capensis*, and reported that, after 12 days of isolation, all previous coincidences between the phases of these two rhythms (individual bees and colony) disappeared, which led to the conclusion that the common periods of activity and rest resulted from mutual (social) synchronization in contrast to the rhythm of individual bees.

Conclusion

The technique of generalized estimating equations has proved to be adequate for assessing the existence of time dependence for entomological experiments, even stressing the daytime mortality peaks.

The 'Non-stationary-1-Dependent' correlation structure was identified and may explain a possible time dependence caused by the mortality of Africanized bees confined in the same cages over time.

Two different diets according to the time period of the experiment are recommended in laboratory experiments with adult bees: granulated sugar enriched with lemon juice is recommended during the first 192 hours and granulated sugar diet as from this time.

Acknowledgements

The authors would like to thank the Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG) for its financial support.

References

Brighenti, C. R. G., Cirillo, A. M., & Brighenti, D. M. (2008). Análise longitudinal na determinação do fotoperíodo adequado para criação de abelhas em laboratório. *Revista Brasileira de Biometria, 26*(3), 111-124.

Brighenti, D. M., Brighenti, C. R. G., Cirillo, M. A., & Santos, C. (2010). Optimization of the components of an energetic diet for Africanized bees through the modelling of mixtures. *Journal of Apicultural Research, 49*(4), 326-333.

Diggle, P. J., Heagerty, P., Liang, K. Y., & Zeger, S. L. (2002). *Analysis of longitudinal data* (2nd ed.). New York, NY: Oxford University Press.

Frisch, B., & Koeniger, N. (1994). Social synchronization of the activity rhythms of honeybees within the colony. *Behavior Ecology Sociobiology, 35*(2), 91-98.

Gosho, M. (2014). Criteria to select a working correlation structure for the generalized estimating equations method in SAS. *Journal of Statistical Software, 57*(1), 1-10.

Harker J. E. (1958). Diurnal rhythms in the animal Kingdom. *Review of Biology, 33*(1), 1-52.

Liang, K. Y., & Zeger, S. L. (1986). Longitudinal data analysis using generalized linear models. *Biometrika, 73*(1), 13-22.

Manning, R. (2001). Fatty acids in polen: a review of their importance for honey bees. *Bee World, 82*(2), 60-75.

Pirk, C. W. W., Miranda, J. R., Fries, I., Kramer, M., Paxton, R., Murray, T., … Van Dooremalen, C. (2013). Statistical guidelines for *Apis mellifera* research. In V. Dietemann, J. D. Ellis & P. Neumann (Eds.), The coloss beebook, volume I: standard methods for *Apis mellifera* research. *Journal of Apicultural Research*, *52*(4), 1-24.

Proni, E. A., & Macieira, O. J. D. Ritmo circadiano da taxa respiratória de *Tetragonisca angustula fiebrigi* (Schwarz), *T. a. angustula* (Latreille) e *Trigona spinipes* (Fabricius) (Hymenoptera: Apidae). (2004). *Revista Brasileira de Zoologia, 21*(4), 987-993.

R Core Team. R: a language and environment for statistical computing. (2014). *R Foundation for Statistical Computing*, Vienna, W. Recuperado de http://www.R-project.org/.

Scheiner, R., Abramson, C. I., Brodschneider, R., Crailsheim, K., Farina, W., Fuchs, S., … Thenius, R. (2013). Standard methods for behavioural studies of *Apis mellifera*. In V. Dietemann, J. D. Ellis, P. Neumann (Eds.), The coloss beebook, volume I: standard methods for *Apis mellifera* research. *Journal of Apicultural Research, 52*(4), 1-58.

Wang, Y, Carey, V. (2003). Working correlation structure misspecification, estimation and covariate design: Implications for generalised estimating equations performance. *Biometrika, 90*(1), 29-41.

Williams, G. R., Alaux, C., Costa, C., Csáki, T., Doublet, V., Eisenhardt, D., … Brodschneider, R. (2013). Standard methods for maintaining adult *Apis mellifera* in cages under in vitro laboratory conditions. In V. Dietemann, J. D. Ellis, P. Neumann (Eds.), The coloss beebook, volume I: standard methods for *Apis mellifera* research. *Journal of Apicultural Research, 52*(4), 1-53.

Withrow, A. P., & Withrow R. B. (1949). Photoperiodism chlorosis in tomato. *Plant Physiology, 24*(4), 657-663.

Vens, M., & Ziegler, A. (2012). Generalized estimating equations and regression diagnostics for longitudinal controlled clinical trials: a case study. *Computational Statistics & Data Analysis, 56*(5), 1232-1242.

Zeger, S. L., & Liang, K. Y. (1986). Longitudinal data analysis for discrete and continuous outcomes. *Biometrics, 42*(1), 121-130.

Ziegler, A., & Vens, M. (2010). Generalized estimating equations: notes on the choice of the working correlation matrix. *Methods of Information in Medicine, 49*(5), 421-42.

Received on February 12, 2015.
Accepted on September 22, 2015.