Viability of Aedes aegypti eggs in Manaus (AM)

Accepted 14th May, 2020

ABSTRACT

This study aimed to estimate the period in which the eggs of Aedes aegypti are able to hatch and produce viable larvae (viability) after being stored for a certain time in different types of containers. To do this, the methodology of generalized linear models that use data in the form of proportion was used. As link function, it was opted for the logistics and probit functions, as linear predictors, parallel and concurrent lines. An experiment was conducted with the aim to simulate as much as possible the effect of environmental conditions that the eggs of A. aegypti are subjected. The eggs were stored over eight times in three containers (plastic cup, plastic bag and paper envelope) and treated with water at the end of each period established to observe the number of hatching. By adjusting the binomial generalized linear model with logistic function and linear predictor parallel lines, it was possible to predict the median egg viability of A. aegypti. In setting quality measures, it was used the residual deviation with p-value (descriptive level), calculated on the percentiles of the chi-square distribution and the normal probability plot with simulation envelope. The median viability estimated for the plastic cup container was 104 days, 87 days for plastic bag and 83 days for paper envelope. There was similar behavior between plastic bags and paper envelopes with respect to estimated viability of A. aegypti eggs, both differing from the viability of the plastic cup container, with the highest viability.

Key words: Binomial GLM, Data as proportion, A. aegypti de eggs, viability.

INTRODUCTION

Among the recurrent diseases, dengue sets is the most important arbovirus that affects humans and constitutes a serious public health problem. Dengue is a viral disease of tropical and subtropical countries, generally of benign evolution. It is transmitted to humans by the bite of the female when infected by the A. aegypti mosquito, vector of the four serotypes of the dengue virus. The current trend in dengue vector control is to restrict the use of chemical pesticides and instead it is encouraged the use of biological control and environmental management (Alves, 1998; Pamplona et al., 2004; Polanczyk et al., 2003). Entomopathogenic bacteria forming spores have great potential for biological control of insects. Bacillus thuringiensis is constituted with active ingredient used commercially in biopesticides (Habil and Andrade, 1998). However, chemicals are still used to combat the larvae, by applying directly to potential breeding sites and to fight the winged, directly in the air, as unfortunately there is no vaccine available in the market to prevent the disease.

The study of the characteristics of A. aegypti, as well as its patterns of behavior and development in different stages of its life cycle, is an important tool for understanding the population dynamics of this species. Moreover, it enables improving actions to combat the vector. Knowing that the egg stage represents the increased resistance of the biological cycle of the mosquito and this being the main factor favoring the spread of the vector by large geographical areas, studies on the viability of the eggs in environmental conditions are sought and become necessary, in order to obtain information that can improve the targeting of control measures.

In previous studies on the oviposition, it was shown the hatching of eggs kept dry for periods of two months. Other series of studies were conducted in the sequence, and it was
found that the viability of the eggs is much higher. The results of Silva et al. (1993) and Silva and Silva (1999) showed that 85% of them may break after three days of quiescence, when in contact with water, and less than 10% can survive for upto 492 days. Pinheiro (2005) observed hatchings of eggs that had been kept dry for a period of upto 186 days after contact with water in the city of Manaus (less than 1%). In the study, the author also showed that similar behavior to the viability period occurred in the plastic bag and paper envelope containers, using the statistical analysis tools the ANOVA methodology and Tukey test.

In this study, tests on Aedes aegypti mosquito eggs in the malaria and dengue laboratory of the National Institute of Research of Amazonas (INPA) in Manaus (AM) were carried out in order to quantify the time that eggs remain in conditions of hatching and producing viable larvae, after the embryo enters the diapause state. The eggs used were stored over eight times in three containers: eggs stored in cups, envelopes and plastic bags. It was taken as variable response for statistical analysis the proportion of viable larvae from the total of stored eggs.

The objective of this study was to determine a central estimate for the period in which eggs of A. aegypti are able to hatch and produce viable larvae, after being inducted into the state of diapause by the absence of contact with water.

MATERIALS AND METHODS

The eggs of A. aegypti used during the tests were collected from six cages of the colony of INPA malaria and dengue laboratory, in which were placed daily 100 ml plastic cups containing 20 ml of water, with the sidelined with a filter paper band measuring 3 cm in height by 22 cm in length, to serve as egg laying substrate. Each cup remained in the cage for two hours, always in the afternoon. After being collected, the cards with eggs were exposed to moisture for 72 h for the development of the embryo to be finalized.

The cards were left for 24 h in the environment until completely dry. Then, it was made up counting and storing in the respective containers, paper envelopes (medium office type), plastic bags (200 ml) and opened plastic cups (100 ml).

The eggs were stored until the completion of their periods of storage, which were previously defined. Thus, the eggs remained stored for periods of 12, 19, 32, 61, 89, 118, 158, 186 days, time in which they were treated with water and the viability was measured by means of total live larvae from the total eggs used in each period and type of storage.

To stimulate the eggs, the cards were plunged in plastic basins, with 1.5 ml of artesian well water, plus 10 ml of liver powder used as food for the larvae (Scarpassa and Tadei, 1990). Randomly, two cards were selected from five, of each stored period, being sixteen basins per container, totaling 48 basins. The count of the number of larvae present in the pots was given daily over a 10 days period, checking on the sixth day a frequency of larvae almost equal to zero and no more occurrence after the tenth day. This observation time was determined on the basis of a pilot experiment in which readings were taken up to 25 days. The replenishing of water was made according to the evaporation and cleaning of basins performed twice a week. It were also performed studies at the level of scanning electron microscopy to observe changes in structural level resulting from the storage time.

The experiment consists of a completely randomized design, whose response variable $Y_{ij}^*$ is the success rate (number of larvae from total of $m_{eggs}$/total of $m_{eggs}$). As explanatory variables were considered: containers with three levels (plastic cup, paper envelope and plastic bag), classified as qualitative and storage period. This period corresponds to the time that the eggs were stored until treatment. The eggs were housed through the woods, in a screened wooden cage inside a hut with protection of sun, rain and possible predators, in order to reproduce the environmental conditions they are subjected.

Procedure

For the statistical analysis of the experiment, generalized linear models - GLMs were set (McCullagh and Nelder, 1989). A GLM is specified by three components: random component, systematic component and link function. The random component specifies the distribution assumed by the response variable, which must be a member of the exponential family of distributions. The systematic component specifies the linear structure for explanatory variables, which come in the model in the form of a linear sum of its effects, resulting in the linear predictor. The link function is a function that relates the random component to the systematic component.

Being $Y_{ij}^*$ the ratio of success, each with probability of occurrence ($\pi_i$), it follows that the natural model for this situation is the binomial GLM (McCullagh and Nelder, 1989; Dobson, 2002; Cordeiro and Demétrio, 2011). As link functions, it were opted the logistics and probit links, and as systematic part, a completely randomized design with two explanatory variables: containers (qualitative) with three levels and storage period. Linear predictors considered are lines that can be concurrent or parallel. Thus, the binomial GLM is of the form:

$$m_i Y_{ij}^* \sim bin(m_i, \pi_i), i = 1,2,\ldots, n$$

$$\mu_i = E(Y_{ij}^*) = \pi_i$$

$$\eta = X_i \beta_j$$
\[ \eta_i = \log \frac{\pi_i}{1 - \pi_i} = \log \frac{\mu_i}{m_i - \mu_i} \]

\[ \eta_i = \Phi^{-1}(\mu_i) \tag{2} \]

where \( m_i y_i \) is the number of successes; \( Y \) is the vector of observed proportions, of dimension \( n \times 1 \); \( \eta = (\eta_1, \ldots, \eta_n)^T \) is the vector of linear predictors; (1) logistic link function; (2) probit link function; \( X \) is the matrix of design, of dimensions \( n \times p \); \( \beta \) is the vector of unknown \( p \) parameters, of dimensions \( p \times 1 \) and \( m \) the number of independent assays.

The linear predictor have form expression, for the model with concurrent and parallel lines, respectively:

\[ \eta_{ijk} = \alpha_j + \beta_j x_{ik} \tag{3} \]

\[ \eta_{ijk} = \alpha_i + \beta x_{ik} \tag{4} \]

with \( i = 1, 2, \ldots, 8 \), storage periods, \( k = 1, 2 \) repetitions and \( j = 1, 2, 3 \), types of containers.

Thus, the probability function of \( Y^* \) for observation \( i \) is expressed by:

\[ f(y_i^*; \pi_i, m_i) = \exp \left( y_i^* \log \frac{\pi_i}{1 - \pi_i} + \log \frac{m_i}{(m_i y_i^*)} \right) \tag{5} \]

**Exponential family and the likelihood equations**

Making in (5),

\[ \theta_i = \log \left( \frac{\pi_i}{1 - \pi_i} \right), b(\theta_i) = - \log(1 + \exp(\theta_i)), a(\phi) = \frac{1}{m_i} e^c(y_i, \phi) = \log \left( \frac{m_i}{(m_i y_i^*)} \right), \]

has:

\[ f(y_i^*; \theta_i, \phi) = \exp \left( y_i^* \theta_i - \log(1 + \exp(\theta_i)) + \log \left( \frac{m_i}{(m_i y_i^*)} \right) \right), \tag{6} \]

which shows that the distribution in (5) to data as proportion is a member of the exponential family of distributions. Being the logarithm of the likelihood function of (6) as a function of \( \beta \) expressed by:

\[ l(\beta) = \sum_{i=1}^{n} l_i = \sum_{i=1}^{n} m_i [y_i \theta_i - \log(1 + \exp(\theta_i))] + \sum_{i=1}^{n} \log \left( \frac{m_i}{(m_i y_i^*)} \right) \tag{7} \]

and the respective score vector \( U(\beta) = \partial l(\beta)/\partial \beta \) of dimension \( p \), with typical element:

\[ U_j = \frac{\partial (\beta)}{\partial \beta_j} = \sum_{i=1}^{n} \frac{\partial l_i}{\partial \beta_j} = \sum_{i=1}^{n} \frac{\partial l_i}{\partial \theta_i} \frac{\partial \theta_i}{\partial \mu_i} \frac{\partial \mu_i}{\partial \beta_j} \eta_i \tag{8} \]

in which

\[ \frac{\partial l_i}{\partial \theta_j} = m_i \left[ y_i \exp(\theta_j) \frac{1}{1 - \exp(\theta_j)} \eta_i \frac{\partial \theta_i}{\partial \mu_i} \frac{\partial \mu_i}{\partial \beta_j} \right] \]

\[ \frac{\partial \theta_i}{\partial \beta_j} = b'(\theta_i) = \frac{\text{Var}(Y_i)}{a(\phi)} \frac{1}{\text{Var}(Y_i)} \frac{1}{m_i \text{Var}(Y_i)} \tag{9} \]

\[ \frac{\partial \mu_i}{\partial \eta_i} = b'(\theta_i) = \frac{\text{Var}(Y_i)}{a(\phi)} \frac{1}{\text{Var}(Y_i)} \frac{1}{m_i \text{Var}(Y_i)} \tag{10} \]

\[ \frac{\partial \mu_i}{\partial \eta_i} \text{ depends on the link function, as } \eta_i = g(\mu_i) \tag{12} \]

so, for the case where the link function is the logistic, there is:

\[ \frac{\partial \mu_i}{\partial \eta_i} = \frac{\exp(\eta_i)}{[1 + \exp(\eta_i)]^2} \tag{13} \]

Using the results (9),(10),(11) and (13), replacing (8) and equating to zero, we arrive at the following system of equations:

\[ \sum_{i=1}^{n} m_i (y_i - \mu_i)x_{ij} = 0, \tag{14} \]

The result in (14) are the likelihood equations of binomial generalized linear model with logistic link function (Agresti, 2002).

**Models adjustment**

The process for setting the models proceeded upon settlement of the equations \( \partial(\beta)/\partial \beta = 0 \), which generally are not linear and therefore cannot be explicitly resolved (Cordeiro and Paula, 1989).

The solution of these took place through the function-GLM of the statistical program R (R Development Core Team, 2013) using as routine the reweighted least squares iterative process, whose expression is of the form:

\[ \beta^{(m+1)} = (X^T W^{(m)} X)^{-1} X^T W^{(m)} Z^{(m)}, \tag{15} \]

in which \( \beta^{m+1} \) is the vector estimated in the \( (m+1) \) iteration, of dimension \( p \times 1 \); \( X \) is the matrix of design of dimension \( n \times p \);

\[ z^{(m)} = n m + G^{(m)} (y - \mu^{(m)}) \]

\( \eta = (\eta_1, \ldots, \eta_n)^T \) is the vector of linear predictors;

\[ G = diag \left( \frac{\partial \eta_1}{\partial \mu_1}, \ldots, \frac{\partial \eta_n}{\partial \mu_n} \right) = diag \left( g'(\mu_1), \ldots, g'(\mu_n) \right) \]

\( e W = diag (w_1, \ldots, w_n) \) with element:

\[ w_i = \left( \frac{\partial \mu_i}{\partial \eta_i} \right)^2 / \text{Var}(Y_i), \tag{16} \]
Adjustment of quality measures

To check the suitability of the adjusted models, it were taken as measures the deviation and normal probability graphs with simulated envelopes, built on deviation component residuals, which are considered to have approximately normal distribution (Lee and Nelder, 1998). It is concluded by the suitability of the probabilistic model if: i) $S_p = \varphi^{-1}D_p \leq \chi^2_{n-p, \alpha}$ (Anderson-Anderson and Quiquadrado) and yet, ii) if in the normal probability plot with envelope all the plotted points are predominantly within the limits with the possible exception of some points. In which are considered to have probability graphs with simulated envelopes, built on deviation component residuals.

In which graphs with simulated envelopes, built on deviation component residuals, which are considered to have approximately normal distribution (Lee and Nelder, 1998). It is

and in $\left(\frac{\hat{\beta}_0^2}{\hat{\beta}_1}; 0.12\right)$ and $\left(\frac{\hat{\beta}_0^2}{\hat{\beta}_1}; 0.88\right)$ there are the points of stability of the GLM curve with logistic link.

for probit link function,

$$DL_{50} = \frac{1}{\hat{\beta}_1} [\Phi^{-1}(\pi) - \beta_0]$$

which for $\pi=0.5$, the median viability is estimated by:

$$DL_{50} = \frac{\hat{\beta}_0}{-\hat{\beta}_1}$$

In which $\hat{\beta} = (\hat{\beta}_0, \hat{\beta}_1)^T$ is the estimator of maximum likelihood of $\beta = (\beta_0, \beta_1)^T$ and $\Phi$ quantiles of the cumulative standard normal distribution.

RESULTS AND DISCUSSION

In Figure 1, it is shown the dispersion of the shelf life of eggs of A. aegypti versus proportions of larva hatchings. It is observed for the plastic cups container, in the first five storage periods, greater prominence of the proportions of hatchings in relation to the containers of paper envelopes and plastic bags. Still, a sharp drop in the proportion of hatching for the period of 19 days of the plastic bag container (note 19), apparently a typical point. According to Demétrio and Zocchi (2011), a typical points (isolated failures) can arise due to gross errors in the response variable or in the explanatory variables, by wrong measures, or recall of observation, or yet transcription errors, and others. A possible explanation for this sharp drop is that it may have been a number of inviable eggs and that somehow went unnoticed as the count was performed. It is still observed a sigmoidal behavior.

Adjusted models

The standard generalized linear models initially proposed to estimate the viability showed problems of lack of fit. This inadequacy may have been due to variations greater than the stipulated by the model. For models that considered the probit link function and linear predictors in (3) and (4), the values of residual deviations are 263.60 (with 42 degrees of freedom) and 380.31 (44 degrees of freedom), respectively demonstrating lack of fit. For settings in which it was considered, the logistic link function and linear predictors in (3) and (4), the values of residual deviations are 244.31 (42 degrees of freedom) and 378.08 (44 degrees of freedom), showing evidence of lack of fit for the two models.

As strategies to decrease the extra variation present in...
the data and consequently improve the quality of fit of the models, it was decided to consider the binomial variable with weight \((w=1/\phi)\). Thus, the residual deviation values, case in which it was considered the probit link function and linear predictors in (3) and (4) are 37.788 (with 42 degrees of freedom) and 40.817 (with 44 degrees of freedom), with respective values of the descriptive level \((p=0.6563\) and \(p=0.6088)\). For GLMs with logistic link function and linear predictors (3) and (4), the values of residual deviations are 39.25 (with 42 degrees of freedom) and 40.14 (with 44 degrees of freedom) with the respective descriptive levels \((p=0.592\) and \(p=0.637)\).

Although the residual deviations present evidence of appropriate adjustments of the normal probability plot with simulated envelopes, Figure 2a, presents several points out of the envelope and the distribution of these in the envelope quite irregular. Figure 2b presents most prominently two points (19 and 26) and in Figure 3a, and 3b, both, several points outside the envelopes and the distribution of points in very irregular envelopes, thus

Figure 1: Dispersal plot, storage of A. aegypti eggs versus proportion of larvae hatching.

Figure 2: Normal probability plots with simulated envelopes for the binomial GLM with probit link function and linear predictors concurrent lines (a) and parallel lines (b)
Figure 3: Normal probability plots with simulated envelopes for the binomial GLM with logistic link function and linear predictors concurrent lines (a) and parallel lines (b).

The weights added in the binomial variable to GLMs with probit link were \( wc = \frac{1}{6.9757}\) (concurrent lines) and \( wp = \frac{1}{9.3175}\) (parallel lines) and for GLMs with logistic link \( wc = \frac{1}{6.2245}\) (concurrent lines) and \( wp = \frac{1}{9.4193}\) (parallel lines).

The strategy to include weight in the binomial variable greatly improved the fit of GLMs, however, it was not enough to ensure a satisfactory adjustment. Paula (2013) and Fox (2002) discuss the possible alternative of lack of adjustment caused by the presence of aberrant/influential points to the withdrawal of this in the data set. In Figure 3, there is waste from the observations number 19 and number 26 quite outstanding, which may be due to compromised adjustments. These observations were removed and GLMs were adjusted. Table 1 shows the residual deviation values and their respective numbers of degrees of freedom (df).

Figure 4a, which shows the normal probability plot with simulated envelopes, has at least four points off limits and
Figure 5: Graphics of larvae proportions of *A. aegypti* versus the age at which the eggs were treated and curves set by GLM with logistic link, using as linear predictors concurrent lines (a) and parallel lines (b).

Table 2: Estimates of binomial GLM parameters with logistic connection and linear predictor parallel lines considering all the observations and inclusion of weight.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimates</th>
<th>Standard.E</th>
<th>T-Value</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_{01}$</td>
<td>4.291</td>
<td>0.316</td>
<td>13.56</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>$\beta_{02}$</td>
<td>3.398</td>
<td>0.2779</td>
<td>12.23</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>$\beta_{03}$</td>
<td>3.522</td>
<td>0.2816</td>
<td>12.51</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>-0.041</td>
<td>0.003</td>
<td>-15.89</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

Table 3: Estimates of binomial GLM parameters with logistic connection and linear predictor parallel lines eliminating two observations.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimates</th>
<th>Standard.E</th>
<th>T-Value</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_{01}$</td>
<td>4.4612</td>
<td>0.2712</td>
<td>16.45</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>$\beta_{02}$</td>
<td>3.4577</td>
<td>0.2523</td>
<td>13.70</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>$\beta_{03}$</td>
<td>3.6731</td>
<td>0.2413</td>
<td>15.22</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>-0.0423</td>
<td>0.002</td>
<td>-19.07</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

Waste spread somewhat irregular in the envelope that highlights the lack of model fit with concurrent lines. Yet from Figure 4b showing normal probability plot with simulated envelopes, it can observed every point in the envelope and the much better distribution, indicating that the model with parallel lines fits the data well.

In Figure 5, there are the graphs of larvae proportions *versus* the age at which the eggs were treated and the curves fitted. Similar results of diagnosis by envelopes can be found in (Paula, 2013). The author also discusses results in which an observation or two appear out of the envelope, but very close to the limits, and certifies a good fit.
of the investigated model.

Tables 2 and 3 show the parameter estimates as well as the standard errors for the model with all the observations and the model in which it was opted for the withdrawal of observations 19 and 26. It observed that the parameter estimates were quite close in both cases, standard errors were decreasing for the model in which the two observations were removed and inferential results were not changed.

From the selected binomial GLM with logistic link, it is obtained, respectively for plastic cups, paper envelopes and plastic bags, the equations:

\[
\log\left(\frac{\hat{p}_1(x_1)}{1-\hat{p}_1(x_1)}\right) = 4.461 - 0.042x_i \tag{22}
\]

\[
\log\left(\frac{\hat{p}_2(x_2)}{1-\hat{p}_2(x_2)}\right) = 3.458 - 0.042x_i \tag{23}
\]

\[
\log\left(\frac{\hat{p}_3(x_3)}{1-\hat{p}_3(x_3)}\right) = 3.673 - 0.042x_i \tag{24}
\]

and the median egg viability of *A. aegypti* for plastic cups, plastic bags and paper envelopes with their respective confidence intervals are:

- plastic cups: $\hat{DL}_{50} = \frac{4.461 - 0.042}{3.673 - 0.042} = 106 (98; 114)$ days;
- plastic bags: $\hat{DL}_{50} = \frac{3.458 - 0.042}{3.673 - 0.042} = 87 (80; 93)$ days;
- paper envelopes: $\hat{DL}_{50} = \frac{3.673 - 0.042}{3.673 - 0.042} = 83 (73; 93)$ days.

In Figure 5, it can be seen that the curves fitted to the plastic bag and paper envelope containers are close, and hence it is interesting to test whether they differ statistically. Thus, a new model has been used, considering the data of the paper envelope and plastic bag containers like a single container. Cordeiro and Demétrio (2011) present similar procedure in a study of branch buds of three varieties of apple trees, in order to verify if the varieties differ statistically. The deviation to the new model with logistic function and linear predictor in (4) is 45.31 (with 43 degrees of freedom) with descriptive level ($p = 0.3756$), showing that the model with parallel lines fits well to the data. The difference of residual deviations between the new model and the selected one (Figure 4b) is (4.5354) with ($F = 0.7264$ and $p = 0.3989$), indicating there is evidence that the paper envelope and plastic bag containers behave similarly.

The greater viability that occurred in the plastic cup container (106 days) may have been favored due to the eggs of *A. aegypti* being stored in open containers throughout the study, which may have provided to these greater contact with air. It was also found that only 12% of these survive after 153 days of storage. As for the eggs stored in paper envelope and plastic bag container, they remained closed, having little or no contact with air. For eggs stored in paper envelopes, it was found that only 12% survive after 134 days and for those stored in plastic bags, only 12% survive after 129 days.

**Conclusions**

Using the generalized linear model, weighing in binomial variable to the purpose of decreasing of this variation in the data, presented itself as a good alternative to adjustment of the generalized linear models. In this study, the binomial generalized linear model, with logistic function and linear predictor parallel lines, fitted well to the data, making it possible to satisfactorily estimate the parameter $DL_{50}$, being the main contribution of this work. Nonetheless, the good result of the model was only possible by the withdrawal (elimination) of two observations of the data set. Observations that when removed did not change the parameter estimates and inferences. Still, the adjustment quality analysis was given by the normal probability plot with simulation envelope on the deviation component residuals.

The median viability estimated for eggs stored in plastic cup containers was 106 days; plastic bag, 87 days; and paper envelope, 83 days. It was also observed significant differences in median estimates of the viability of the eggs of *A. aegypti*, and for the plastic cup container, it was obtained the highest viability. Furthermore, it was found that only 12% of the eggs stored in plastic cups survive after 153 days. From the eggs stored in plastic bags, 12% survive after 134 days and for eggs stored in paper envelopes, only 12% survive after 129 days.

**REFERENCES**


Cite this article as:


Submit your manuscript at http://www.academiapublishing.org/journals/ajar