CONTROL EVALUATION OF *CONYZA SPP*. WITH GAMLSS MODELS: AN EFFICIENT ALTERNATIVE TO THE TRADITIONAL ANALYSIS OF VARIANCE MODEL

10.56238/edimpacto2025.022-005

Elias Silva De Medeiros[1], Paulo Vinicius Da Silva[2], Bruna Ferrari Schedenffeldt[3],
Munir Mauad[4], Pedro Antônio Vougoudo Salmazo[5], Carolina Cristina Bicalho[6]

ABSTRACT

The study aimed to present a framework on the fit of the Generalized Additive Models for Location, Scale, and Shape (GAMLSS) with the support of R software to evaluate the control of *Conyza* spp. in pre-sowing desiccation of soybean during an advanced phenological stage through herbicide associations and/or sequential applications. GAMLSS is an alternative to the traditional model, which allows modeling the response variable using a large class of probability distributions. The dataset used in this study refers to the control of *Conyza* spp., ranging from 0 to 100%. Visual evaluations were performed at 7, 14, 21, 28, and 35 days after application of different herbicides. The results of this study indicate that GAMLSS may be an efficient alternative to the traditional analysis of the variance model, especially when the data cannot be represented by a normal distribution or exhibit heteroscedasticity.

Keywords: ANOVA. Beta Distribution. Conyza spp. Herbicides. Software R.

---

[1] Federal University of Grande Dourados, Faculty of Agricultural Sciences, Dourados/Itahum Highway, Km 12, 79.804-970, Dourados, Mato Grosso do Sul, Brazil

[2] Federal University of Grande Dourados, Faculty of Agricultural Sciences, Dourados/Itahum Highway, Km 12, 79.804-970, Dourados, Mato Grosso do Sul, Brazil

[3] Federal University of São Carlos, Graduate program in agriculture and environment, Rodovia Anhanguera km 174, 13.600-970, Araras, São Paulo, Brazil

[4] Federal University of Grande Dourados, Faculty of Agricultural Sciences, Dourados/Itahum Highway, Km 12, 79.804-970, Dourados, Mato Grosso do Sul, Brazil

[5] Federal University of Grande Dourados, Faculty of Agricultural Sciences, Dourados/Itahum Highway, Km 12, 79.804-970, Dourados, Mato Grosso do Sul, Brazil

[6] Mato Grosso do Sul State University, Department of Mathematics, Dourados/Itahum Highway, Km 12, 79.804-970, Dourados, Mato Grosso do Sul, Brazil

INTRODUCTION

Weed control is essential to reduce weed competition because these plants compete for water, light, and nutrients with crops of agronomic interest, reducing their productivity and harvest quality [1]. Among the weeds, the species belonging to the genus *Conyza* spp. stand out for infesting mainly the soybean crop [2,3]. These plants exhibit aspects of ecophysiology, such as germination at temperatures below 20°C, which favors their establishment at the time of pre-sowing soybean desiccation [4]. In addition to these aspects, plants of the genus *Conyza* spp. have a high competitive ability due to high seed production, reaching more than 800,000 seeds per plant [5]. In addition to the high adaptability of these weed species to the production of soybean/corn succession, there is also difficulty in control due to the report of biotypes resistant to the herbicides glyphosate, chlorimuron, paraquat, and 2,4-D in Brazil [6,7].

In this scenario, aiming at data analysis, the traditional model of analysis of variance has been widely used in the evaluation of different herbicides for weed control. A study on the efficacy of different herbicides in controlling *Conyza* spp. during pre-sowing soybean management, found that the treatments sulfentrazone/diuron, imazethapyr/flumioxazin, and diclosulam, when used in association with the herbicides diquat, paraquat, or glufosinate, were effective in controlling *Conyza* spp. [6]. An experiment was conducted to evaluate the efficacy and selectivity of herbicides applied under preemergent conditions for the management of weeds in coffee (*Coffea arabica* L.) even after 120 days of treatment application. Zobiole et al. (2018)[8] concluded that the combination of diclosulam + halauxifen-methyl + glyphosate with ammonium-glufosinate, paraquat, or saflufenacil, in sequential applications presented satisfactory results in the control of *Conyza* spp. at different stages of development. In all the articles mentioned, the F test obtained from the analysis of variance traditionally was used. However, this traditional model assumes the adequacy of the normal distribution of the data since the F statistic has this distribution in its theoretical construction. In addition, the assumptions of normality and homogeneity for the residuals of this model, if not verified, make the hypothesis tests invalid, causing bias in the conclusions of the experimental results.

The regression model of ordinary least squares (OLS) considered in the analysis of variance, with a normal distribution, is not appropriate because it can lead to violations of the assumptions of the statistical model and affect the estimates generated by the fit [9]. These authors conducted a study with German farmers about the use of the herbicide glyphosate in conventional agriculture, discussing the choice of a suitable regression model to analyze the relationship between predictive variables and the application of glyphosate in specific crops. In addition, it should be considered that the normal distribution is not truncated only in the interval from zero to one, as is the case with proportion data, but in the entire domain of real numbers.

The Generalized Additive Models for Location, Scale, and Shape (GAMLSS) is an alternative to the traditional model, which allows modeling the probability distribution of a response variable, even when it does not follow a normal distribution [10]. Additionally, with this methodology, it is possible to model the dispersion of the data since the different types of herbicides may present different variations in the evaluation of weed control. GAMLSS with a beta distribution has been used to evaluate the performance of different post-emergent herbicides in the control of *Digitaria insularis* under desiccation before soybean sowing [11]. Similarly, now applying the beta-inflated distribution of 1's, the herbicides diquat, glufosinate ammonium, saflufenacil, and flumioxazin were evaluated for preharvest desiccation of soybean crops.[12]. However, these studies presented a model only for the location parameter, the mean, not inferring about the dispersion of the data about the different herbicides. In addition, there is the computational aspect because thus far, these models have been implemented in R software [13], through the gamlss library [10], requiring a detailed description of how to analyze the experiments using this library, which is useful information for researchers to make use of in their statistical analysis, which is similar to our study.

Thus, the present study aims to present a framework for the fit of GAMLSS models with the support of R software to evaluate the control of *Conyza* spp. in pre-sowing desiccation of soybean at an advanced phenological stage through herbicide associations and/or sequential applications.

MATERIALS AND METHODS

EXPERIMENTAL DESIGN

The experiment was conducted in the field in an area with a natural infestation of *Conyza spp.* The experimental design used was a randomized block, with four replicates for each treatment. The experimental units consisted of 2 x 4 meter plots, totaling 8 m² for each experimental unit. Herbicide applications occurred on *Conyza* spp. with up to 15 cm in height, with 14 treatments: (T1) 2.4 + glyphosate seq 1. Haloxifop seq2. Glufosinate - ammonium salt; (T2) dicamba seq 1. Haloxyfop seq2. Glufosinate - ammonium salt; (T3) fluroxypyr + clethodim+ seq 1. Haloxyfop seq2. Glufosinate - ammonium salt; (T4) triclopyr + glyphosate seq 1. Haloxyfop seq2. Glufosinate - ammonium salt; (T5) atrazine + mesotrione + glyphosate seq 1. Haloxifop seq2. Glufosinate - ammonium salt; (T6) diclosulam+ Halauxifen-methyl + glyphosate seq 1. Haloxifop seq2. Glufosinate - ammonium salt; (T7) 2,4D + glyphosate + Haloxyfop seq. Glufosinate - ammonium salt; (T8) dicamba + glyphosate + Haloxifop seq. Glufosinate - ammonium salt; (T9) triclopyr + glyphosate + Haloxyfop seq. Glufosinate - ammonium salt; (T10) atrazine + mesotrione + glyphosate + Haloxifop seq. Glufosinate - ammonium salt; (T11) diclosulan + Halauxifen-methyl + glyphosate + Haloxifop seq. Glufosinate - ammonium salt; (T12)) atrazine + mesotrione + 2,4 D + Haloxyfop seq. Glufosinate - ammonium salt; (T13) atrazine + mesotrione + Haloxifop seq. Glufosinate - ammonium salt in addition to the control (T14) without application of herbicides.

The herbicide treatments were applied with a backpack sprayer pressurized with CO2, with a pressure of 2.5 bar, with a spray bar containing four Teejet 110.015 fan nozzles spaced 0.5 m apart, with an application volume of 150 L ha $^{-1}$. The control of *Conyza spp.* was determined using visual evaluation of the control through the use of a visual scale, in which scores were assigned by the percentage of control. There was no control of the weed species and 100% total control of the species. Visual evaluations were performed at 7, 14, 21, 28, and 35 days after treatment applications (DAT).

STATISTICAL ANALYSIS

The great flexibility in GAMLSS modeling is the use of different probabilistic models that adequately represent the parametric space of the response variable [10]. Therefore, as the control of *Conyza* spp. was calculated in proportion, we used the beta distribution, defined as:

$$f(y|\alpha,\beta) = \frac{1}{B(\alpha,\beta)} y^{\alpha-1}(1-y)^{\beta-1},\ 0 < y < 1.$$

The parameterizations $\mu = \frac{\alpha}{(\alpha+\beta)}$ and $\sigma = (\alpha+\beta+1)^{-\frac{1}{2}}$ were adopted in this study, and μ and σ represent the location and scale parameters, respectively [14]. In addition, $E(y) = \mu$ and $Var(y) = \sigma^2\mu(1-\mu)$. Thus, we have to $0 < \mu, \sigma < 1$.

In GAMLSS, any parameter of the probability distribution can be modeled. In this article, we present the results for modeling the location and scale parameters using the following structure:

$$g_1(\mu) = \eta_1 = u_1 + b_k + h_i + d_j + (hd)_{ij} + p_{ki}\ \ g_2(\sigma) = \eta_2 = u_2 + h_i$$

The term y_{ijk} represents the percentage of *Conyza* spp. obtained in the k-th block, the i-th herbicide and the j-th DAT. The terms u_1 and u_2 represent the effects of the intercept on the location and scale parameters, respectively. The term b_k represents the effect of the kth block $(k = 1,2,3,4)$, h_i the effect of the i-th herbicide $(i = 1,2,\ldots,13)$, d_j the effect of the j-th DAT, and $(hd)_{ij}$ is the effect of the interaction between the i-th herbicide and the j-th DAT. The random effect, p_{ki}, representing the experimental unit, was inserted into the model. In the link functions $g_1(\mu)$ and $g_2(\sigma)$ adopted, in both, the logit function, resulting in the linear predictor η expressed as:

$$\eta_1 = ln\left(\frac{\mu}{1-\mu}\right) \Rightarrow \mu = \frac{1}{1+e^{-\eta_1}} \quad \eta_2 = ln\left(\frac{\sigma}{1-\sigma}\right) \Rightarrow \sigma = \frac{1}{1+e^{-\eta_2}}$$

Building the model in the gamlss package [10] in R software [13] was performed as follows:

```
m0 <- gamlss(Buva ~ re(fixed = ~ Bloco + Trat*DAT,
          random = ~ 1|parc), data = dd,
  sigma.formula = ~ Trat,
  family = BE(mu.link = 'logit', sigma.link = 'logit'),
  trace = F)
```

In the argument `BE(mu.link = 'logit')`, the term BE indicates the Beta distribution, using the link function logit for the lease parameter. The user can consult a list of other distributions using the command `?gamlss.dist::as.family`. With the command `show.link("BE")`, it is possible to consult the list of binding functions available in the beta distribution for each of its parameters.

To obtain the deviance analysis, the function `anova(getSmo(m0))` was used, considering DAT to be a qualitative factor. The argument parc refers to the random effect of the plot, represented here by the block and treatment interaction. The residual diagnosis of the model was performed visually using the command `plot(m0)`. Graphical presentation of the results was performed using the packages gamlss.ggplots [15] and ggplot2 [16].

RESULTS

In the analysis of *deviance* (Table 1), the herbicide versus DAT interaction was statistically significant ($P < 0.05$) when the control of *Conyza* spp.

Table 1. Results of the deviance analysis about the model factors when evaluating the control of *Conyza* spp.

Factor	Degrees of Freedom	F value	p value
Block	3	3.794	0.018
Treatment	12	9.287	0.000
DAT	4	1043.369	0.000
Treatment versus DAT	48	10.636	0.000

Figure 1 shows the results of the diagnostic analysis of the model. The package gamlss.ggplots can be used to obtain a graphical analysis of the fit of the GAMLSS model.

In this package, we have several functions that present graphical results of the fit of the model.

```
resid_qqplot(m0)
resid_wp(m0)
```

Figure 1. Results of the diagnosis of the GAMLSS models.

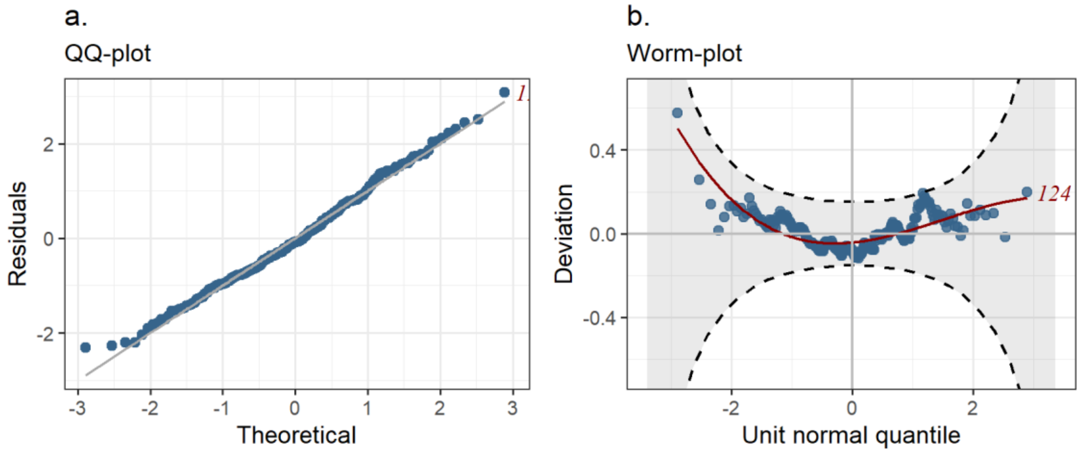


Figure 1a, constructed by the function reside_qqplot to be used to evaluate the adequacy of the residuals with the distribution that was specified in the model. To fit our data, we observed that the points are distributed along the diagonal line, indicating that the beta distribution is suitable for modeling our response variable. Figure 1b shows the Worm plot used to identify whether the model is properly specified. Note that the points are distributed within the confidence bands and that there is no cyclical or U-shaped behavior, indicating the adequacy of the model fit about its specification.

As previously mentioned, the deviance analysis indicated a significant interaction between the herbicides and DAT. Therefore, the computational procedures and the results of the splitting of treatments into each DAT are presented below.

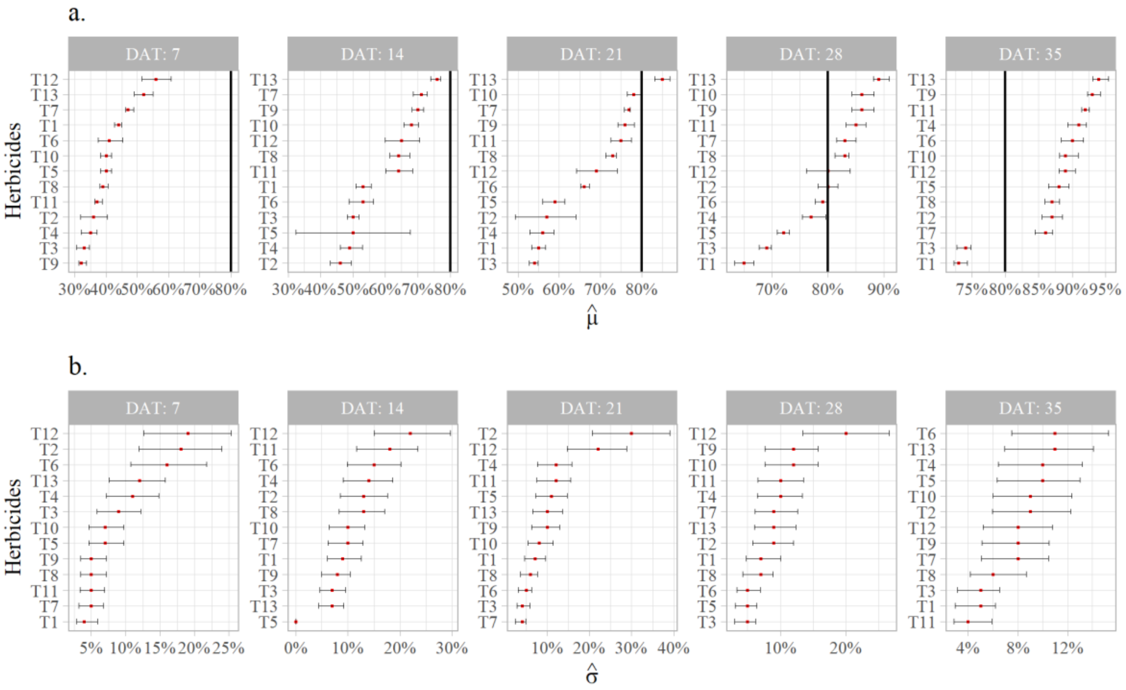
```
m1 <- gamlss(Buva ~ Trat*DAT,
  sigma.formula = ~ Trat*DAT,
  data = dd,
  family = BE(mu.link = 'logit', sigma.link = 'logit'),
  trace = F)
```

The function lpred provides the estimates for the parameter of interest, as well as the standard errors associated with each of these estimates. In the argument, we should specify which parameter we are interested in analyzing. Thus, the model provides the estimate of the response variable for each herbicide in the different DATs.

```
lp <- lpred(m1, what = "mu", type="response", se.fit = T)
```

In Figure 2a, we have the estimates for the lease parameter (μ), which represents the expected value of the percent control of *Conyza* spp., with the horizontal bars indicating the respective standard errors$(\pm SE)$ and the vertical line indicating the control of 80%.

Figure 2. Results of the diagnosis of the GAMLSS models. Estimation of location (a) and scale parameters and the respective standard errors associated with the different herbicides when evaluating the control of *Conyza* spp.



It is observed that all treatments up to 14 DAT presented control lower than 80%. At 21 DAT, only the T13 treatment presented an estimated control higher than 80%, being considered effective, even though this treatment corresponded to the best results, presenting the highest estimates from the 14th DAT. At 28 DAT, only treatments T1, T3, T4, T5 and T6 presented control means in the range of 65 to 80%, proving to be a good control but not effective.

In the last evaluation, at 35 DAT, unlike what was observed previously, only treatments T1 and T3 presented estimates lower than 80%, and on the same evaluation date, the best controls referred to treatments T13 and T9, with estimates greater than 90%, standing out as excellent alternatives in the management of *Conyza* spp. in desiccation.

Regarding the variability (Figure 2b), the modeling on the scale parameter,σ, found that the T12 treatment presented the highest estimates of variability up to 35 DAT, with values close to 20%. From the 14th DAT onward, T3 was among the treatments with the smallest variations, registering values below 10%.

DISCUSSION

The evaluation of the efficacy of different herbicides in the control of *Conyza* spp. has been the subject of studies in which the traditional model of analysis of variance has been used, assuming that the normal distribution is suitable for modeling [6,17,18]. However, these studies did not investigate the assumptions involving the residuals in the analysis of variance, such as the adequacy of the normal distribution, independence and homogeneity. Thus, the failure to verify these assumptions makes the result of this analysis, with the application of the F test for treatment effect, not valid and may lead to estimates that cause bias in the conclusions.

To investigate the behavior of different herbicides when evaluating the percentage of *Conyza* spp. control, we chose a regression model that incorporated the characteristics of a variable measured in proportion, with its domain defined in the open interval (0, 1). The use of ordinary least squares (OLS) regression models admitted to the normal distribution is not appropriate because it may lead to violations of the assumptions of a statistical model, directly impacting the estimates produced by the fit, without considering that the normal distribution is not truncated only in the range from zero to one but in the entire domain of real numbers. A study of 2,026 German farmers on glyphosate applications in conventional agriculture discussed the choice of an appropriate regression model to investigate the relationship between the predictive variables and glyphosate applications in specific crops [9]. The researchers state that the frequency distributions of these variables are U-shaped, with many structural zeros and ones, and that the use of an OLS regression would not be appropriate for these variables due to the limited property and nonmorality of the fractional variables that would violate the underlying assumptions of an OLS.

The GAMLSS models with beta distribution prove to be an excellent alternative for modeling proportion data. In the literature, although with a low frequency, we can find studies that made use of this methodology and achieved efficient results. A study on the incidence of citrus canker in leaves of different genotypes of pear (*Citrus sinensis*) used the zero-inflated beta distribution to model these incidence rates. were able to model the incidence rates of citrus canker more accurately and obtain valuable information about the disease in different genotypes of pear orange.[19] ... In a study to evaluate the performance of different herbicides as alternatives to ACCase inhibitors in the desiccation of *Digitaria insularis* before soybean sowing, GAMLSS models were used to compare the efficacy of the different herbicides, thus concluding that the GAMLSS model with beta distribution proved to be an efficient tool to analyze the percentage control of the species, which is a valuable approach for this type of research [11].

Therefore, it is evident that the GAMLSS methodology with beta distribution has been applied in the analysis of experiments in which proportion data were evaluated [19], although we found few references on the application of this method in agronomic experimentation, especially in the evaluation of weed control [11,12,20]. However, the aforementioned articles do not present a brief description of how these analyses were performed in statistical software since the GAMLSS models are implemented in R software. Thus, the great novelty of our study is a detailed presentation of how to implement these models in R, and this software is one of the most commonly used in academia for statistical analysis.

An important advantage of the methodology we propose in this study is the fact that we can model, in addition to the location parameter related to the mean, the scale parameter, which is related to the dispersion of the data in the experiment. Therefore, it is possible to infer which herbicides have a higher mean estimate of control and indicates which treatments have lower estimates of variability. Additionally, the inclusion of the scale parameter as a function of one or more explanatory variables can help to capture the variation in the dispersion of the data and lead to more precise estimates and more reliable predictions in the model. A study on regression with GAMLSS showed that this model is particularly effective in situations where the modeling of a response variable using the generalized linear model (GLM) or generalized additive model (GAM) is insufficient [21]. The authors state that the GLM and GAM models only model the mean and consider the dispersion parameter constant. In turn, GAMLSS is able, in principle, to model all the parameters present in the probability distribution.

CONCLUSIONS

The GAMLSS model can be an efficient alternative to the traditional analysis of variance model for evaluating the control of weeds, especially when the data cannot be adequately presented by a normal distribution or show heteroscedasticity. The use of this model may lead to more precise and reliable results in the evaluation of control effectiveness, which may have important practical implications for decision-making regarding the management of invasive plants in agriculture.

Readers interested in replicating the analyses can access the following electronic address http://dx.doi.org/10.13140/RG.2.2.29018.72647, which allows the exact reproduction of the statistical analyses that were presented in this study. The inclusion of this electronic address allows for greater transparency and reproducibility of the analyses, ensuring that other researchers can accurately examine and validate the results and replicate them in similar experiments. Thus, our article demonstrates the authors' commitment to sharing knowledge and promoting advances in the area of Conyza spp. control, in addition to highlighting the importance and efficiency of GAMLSS models in this context.

ACKNOWLEDGEMENTS

To all the employees of the Federal University of Grande Dourados who directly and indirectly contributed to this research.

COMPLIANCE WITH ETHICS GUIDELINES

Elias Silva de Medeiros, Paulo Vinicius da Silva, Bruna Ferrari Schedenffeldt, Munir Mauad, Lucas Santana, Pedro Salmazo and Carolina Cristina Bicalho declare that they have no conflict of interest or financial conflicts to disclose. This article does not contain any studies with human or animal subjects performed by any of the authors.

REFERENCES

1. Eli-Chukwu N C. Applications of Artificial Intelligence in Agriculture: A Review[J]. Engineering, Technology & Applied Science Research, 2019, 9(4): 4377–4383

2. Heap I. The International Herbicide-Resistant Weed Database[J]. WeedScience.org, 2023

3. Diniz W A, Côrrea F R, Silva N F da. Controle de buva em pré-semeadura na cultura da soja em Goiás, Brasil[J]. Brazilian Journal of Science, 2023, 2(3): 82–91

4. Wu H, Walker S, Rollin M J. Germination, persistence, and emergence of flaxleaf fleabane (Conyza bonariensis [L.] Cronquist)[J]. Weed Biology and Management, 2007, 7(3): 192–199

5. Kaspary T E, Lamego F P, Cutti L. Growth, phenology, and seed viability between glyphosate-resistant and glyphosate-susceptible hary fleabane[J]. Bragantia, 2017, 76(1): 92–101

6. Albrecht A J P, Albrecht L P, Alves S N R. Pre-sowing application of combinations of burndown and pre-emergent herbicides for Conyza spp. control in soybean[J]. Agronomía Colombiana, 2021, 39(1): 121–128

7. Queiroz A R S, Delatorre C A, Markus C. Rapid necrosis II: physiological and molecular analysis of 2,4-D resistance in Sumatran fleabane (Conyza sumatrensis)[J]. Weed Science, 2022, 70(1): 36–45

8. Zobiole L H S, Krenchinski F H, Pereira G R. Management Programs to Control Conyza spp. in Pre-Soybean Sowing Applications[J]. Planta Daninha, 2018, 36

9. Wiese A, Schulte M, Theuvsen L. Interactions of glyphosate use with farm characteristics and cropping patterns in Central Europe[J]. Pest Management Science, 2018, 74(5): 1155–1165

10. Stasinopoulos M, Rigby R A, Heller G Z. Flexible regression and smoothing using GAMLSS in R[M]. Boca Raton: CRC Press, 2017

11. Palharani W, Mauad M, Silva P V. Management of Digitaria insularis in soybean pre-sowing desiccation by ACCase alternative herbicides and its impact on soybean carryover[J]. Journal of Environmental Science and Health, Part B, 2023, 58(2): 110–119

12. Silva P V da, Mendonça Bezera M, Medeiros E S de. Pre-harvest desiccation strategies of soybean culture: a scenario without paraquat[J]. Journal of Environmental Science and Health, Part B, 2022, 57(9): 710–719

13. R Core Team. R: A language and environment for statistical computing[J]. R: A Language and Environment for Statistical Computing, 2022

14. Rigby R A, Stasinopoulos M D, Heller G Z. Distributions for Modeling Location, Scale, and Shape[M]. Distributions for Modeling Location, Scale, and Shape, Chapman and Hall/CRC, 2019

15. Stasinopoulos M, Rigby B, Bastiani F De. gamlss.ggplots: Plotting Generalised Additive Model for Location, Scale and Shape[J]. R package version 2.0-1, 2022: 15

16. Wickham H. ggplot2 -Positioning Elegant Graphics for Data Analysis[M]. Springer, 2016

17. Krenchinski F H, Pereira V G C, Zobiole L H S. Halauxifen-Methyl+Diclosulam: New Option to Control Conyza spp. Prior Soybean Sowing[J]. Planta Daninha, 2019, 37

18. Netto A G, Presoto J C ., Resende L S. Effectiveness and selectivity of herbicides applied under preemergence conditions in weed management for coffee crops[J]. Coffee Science, 2022, 16: 1–12

19. Pompeu D D S, Guedes T A, Janeiro V. Modeling the incidence of citrus canker in leaves of the sweet orange variety ‘Pera’[J]. Acta Scientiarum. Agronomy, 2017, 39(1): 1

20. Silva P V da, Medeiros E S, Schedenffeldt B F. Selectivity of post-emergence herbicides in soybean and their efficacy on the control of Conyza spp[J]. Brazilian Journal of Agricultural and Environmental Engineering, 2023, 27(8): 600–609

21. Stasinopoulos M D, Rigby R A, Bastiani F De. GAMLSS: A distributional regression approach[J]. Statistical Modelling, 2018, 18(3–4): 248–273