**Crop-Weed Relationship Studies in Additive Design: Selecting the Top Model**

Maxwel Coura Oliveira1, Gustavo Antonio Mendes Pereira2, Rodrigo Werle3, and Stevan Knezevic1

1Department of Agronomy and Horticulture, University of Nebraska-Lincoln, Concord, NE, USA.

2Department of Phytotechnic, Federal University of Viçosa, Viçosa, MG, Brazil

3Department of Agronomy and Horticulture, University of Nebraska-Lincoln, North Platte, NE, USA.

**Corresponding author:** Maxwel Coura Oliveira, E-mail: maxwelco@gmail.com

**Category:** Biometry, Modeling, and Statistics.

**ABSTRACT**: Crop-weed competition is extensively studied in weed science. The additive design, where weed density varies, and the crop density is kept constant, is the most used design for competition studies. However, most crop-weed competition research is conducted by non-statisticians, which sometimes report misleading results because of lack of knowledge with this type of data analysis. The objective of this study is to demonstrate a selection of a top model for describing the crop-weed relationship in additive design to non-statisticians. We evaluated three models routinely used to interpret competition studies, including polynomial quadratic, logistic, and rectangular hyperbola. Based on statistical criteria and meaningfulness of parameters, we demonstrated the rectangular hyperbola to be the top model to describe crop-weed competition studies in additive design. Moreover, it was showed that at low densities *C. benghalensis* is more competitive than *Richardia brasiliensis* in corn, but both weed species compete similarly at higher densities. In this paper, it is proposed the use of the rectangular hyperbola as a standardized model for crop-weed competition in additive design.

**Keywords**: AIC criterion, model selection, crop-weed competition, rectangular hyperbola

**Nomenclature:** *Commelina benghalensis*, *Richardia brasiliensis*

**Introduction**

Studies have described the relationship function of crop yield loss in response to weed density using additive design. Despite several review papers recommending the use of rectangular hyperbola in the weed science literature (Knezevic et al. 1994; Ritz et al. 2015; Swanton et al. 2015), there is no common sense to which model is the most appropriate (Silva et al. 2015; Strieder et al. 2007; Trezzi et al. 2015; Voll et al. 2002). Four major regressions curves are frequently used: linear (Figure 1A), polynomial quadratic (Figure 1B), sigmoid (Figure 1C), and rectangular hyperbola (Figure 1D).

The commonly used criteria for selection of linear and nonlinear regression models is the equation with higher R-squared (R2). The R2 tests the goodness of fit and is statistically inadequate for nonlinear model selection (Archontoulis and Miguez 2015; Zuur et al. 2007). There are several appropriate statistical criteria for selecting the best nonlinear model for datasets: Alkaike’s information criterion (AIC), Bayesian information criterion (BIC), F-test, and likelihood ratio (Anderson 2008; Lewis et al. 2011; Zucchini 2000). Non-nested models are models with different structure and parameters, in this case, the AIC and BIC are indicated for top model selection. However, F-test or any of the model selection techniques described above are applicable for nested models. Nested are models that are a special case of each other and have identical terms whereas one must have one or more extra terms.

From a practical standpoint, the top model should be selected upon a balance between statistics and biological relevance, which will help scientists answer their research questions (Archontoulis and Miguez 2015). Therefore, in crop-weed relationship studies in additive design, the model that provides a good fit and meaningful biological parameters are considered a strong candidate model.

The advances in statistical software should facilitate the use of standardized nonlinear regression analysis that could perform by non-statisticians (Knezevic et al. 2007). Here, there is a comparison of three non-nested candidate models (polynomial quadratic, logistic, and a rectangular hyperbola) for describing the crop-weed relationship. Data from an experiment of corn (*Zea mays* L.) in competition with two weed species, *Richardia brasiliensis,* and *Commelina benghalensis* was used. First, it demonstrates the selection process for the top model to describe the crop-weed relationship for each weed species. Second, it is tested the hypothesis that *C. benghalensis* and *R. brasiliensis* compete similarly with corn.

**Material and Methods**

**Plant Material**. On March 2011, inflorescences of *R. brasiliensis* were harvest on roadsides near Diamantina, Minas Gerais (MG), Brazil. Inflorescences of *R. brasiliensis* were dried at room temperature (25 °C), cleaned, and stored at 5 °C until the onset of the experiment. Ten days before the experiment began (September 2011), stolon (vegetative propagules) of *C. benghalensis* were collected in wetlands, near Diamantina, MG. Seeds of *R. brasiliensis* and stolon of *C. benghalensis* were seeded and transplanted to separate trays (1210 cm3) filled with red latosol (pH 6.1 and 1% organic matter). A single seed of glyphosate-resistant (GR) corn (AG8088) was sown in 8 dm3 plastic pots filled with the aforementioned soil source. The soil was fertilized following the local recommendations, and N was applied at 15 and 30 DAE (days after corn emergence) at a rate of 55 mg dm-3 of ammonium sulfate. Greenhouse conditions were 26/19 °C day/night, and pots were watered daily.

**Experimental procedures**. The experiment was conducted under greenhouse conditions at the Federal University of Jequitinhonha and Mucuri, MG. In this study, the additive design for competition studies was used, whereas weed densities varied, and corn density was kept constant (Swanton et al. 2015). The treatment design was a factorial with two weed species, *R. brasiliensis* and *C. benghalensis*, and five weed densities (0, 1, 2, 3, and 4 plants pot-1), in a completely randomized design with four replications.

Corn dry matter was harvested at 60 DAE from each experimental unit. Shoot biomass was oven-dried at 65 °C until reaching constant weight, and dry weight recorded. The corn dry matter (g) data (shoot) were converted into yield loss (%) compared with the corn control treatment (no weeds):

eq. (1)

where *Μ* is the mean dry mass (g) of the untreated replicates, and is the dry mass (g) of an individual treated experimental unit.

**Statistical Analysis**

Three models were fitted to yield loss data (%) in response to weed density (plants pot-1):

*Rectangular hyperbola model* proposed by (Cousens 1985):

eq. (2)

where *I* represent YL (yield loss) per unit weed density as *D* (density) approaches 0, and *A* represents YL as D approaches ∞ (or maximum expected yield loss). The rectangular hyperbola model was fitted using the *nls* function of R version 3.3.1 (R Foundation for Statistical Computing, Vienna, Austria).

*Logistic model* (four parameter log-logistic curve):

eq. (3)

where *c* is the (lower limit or YL at low weed density), *d* is the asymptote (upper limit or YL at high weed density), and *e* represents the weed density (weeds pot-2) that cause 50% yield loss (inflection point). The parameter *b* is the relative slope around the parameter *e*, and *D* is the number of weeds pot-1. Parameters for the sigmoidal model (four-parameter logistic) were estimated using the *drm* function of drc package (Ritz and Streibig 2005) in R software.

*Polynomial quadratic model* (second order):

eq. (4)

where *α* is the intercept in the y-axis (maximum YL in the absence of weed), *a* represents the slope of the model. The parameter *b* is the quadratic term of the model, and *D* is the number of weeds plot-1. The parameters for the polynomial quadratic equation were estimated using the *lm* function of R software.

**Top model selection to describe crop-weed competition.** The AICc criterion, which is indicated by non-nested model selection (Hurvich and Tsai 1991), was calculated as:

eq. (5)

where is the likelihood function and is the number of estimated parameters in the model, and *n* is the sample size of the model. The AICc is the second-order or small sample AIC (Sugiura 1978). According to the AICc criterion, the top model has the lowest AICc value. The AICc values for each model was estimated using the *AICc* command of package AICcmodavg in R software (Mazerolle 2016).

**Model selection to evaluate weed competitiveness with the crop.** Assuming that rectangular hyperbola is the top model, the impact of *R. brasiliensis* and *C. benghalensis* on corn YL is accessed through the variance-ratio or F-test performed using equation [2] (Lindquist et al. 1996). This statistical procedure evaluates the difference of residual sum squares (RSS) of nested models (Werle et al. 2014). F-test is calculated as (Archontoulis and Miguez 2015):

eq. (6)

Where RSSFULL and RSSRED represent the minimized residual sum squares of the parameters estimated for the full (step 1) and reduced model (step 2, 3, or 4), respectively; dfFULL and dfRED represent the degrees of freedom of the full and reduced models, respectively. F-value greater than the F-critical value (P-value <0.05) indicates that two models are different. Thus, the full model should be used. F-value smaller than the F-critical value indicates that two models are not different (P-value >0.05); therefore, a model with fewer parameters (reduced model) can be used to describe the data. When P-value >0.05 we fail to reject the null hypothesis and a reduced model should be used (no difference of *I* and *A* parameter values between weed species). However, if P-value <0.05, the null hypothesis is rejected and the full model should be used (different *I* and *A* parameter values for each weed species). The F-test principle for nonlinear regression analysis was calculated for each model using *nls* *ANOVA* command in R software (Ritz and Streibig 2008).

Four major steps need to be completed to compare the parameters using this method (see supplemental file):

First, fit Equation [2] to the data of each species individually (*R. brasiliensis* and *C. benghalensis*); this represents the full model, where four parameter values (*I* and *A* for each weed species) will be estimated.

Second, pooling the data for both species (*R. brasiliensis* and *C. benghalensis*) and fit Equation [2]. This represents the reduced model, where two parameter values (*I* and *A* for both weed species combined) are estimated for the polled data. This step will allow testing the hypothesis that *I* and *A* do not vary between species, which means that both species compete similarly with corn. If the hypothesis is accepted (P-value>0.05), stop here. Otherwise, there is two more hypothesis to be tested.

Third, fit equation [2] setting a single parameter *I*, but different *A* parameter for each species. This is a reduced model, and three parameters will be estimated. This step tests the second hypothesis, that weed species compete similarly at low densities (*I*), but different at higher densities (*A*).

Fourth, fit equation [2] setting a single parameter *A*, but different *I* parameters for each species. This is a reduced model, and three parameters will be estimated. This step tests the third hypothesis, that weed species compete similarly at higher densities (*A*), but different at low densities (*I*). Additional AIC was also performed for the nested model selection for confirming the F-test model selection.

**Goodness-of-fit of the models.** Root mean squared error (RMSE), model efficiency (ME), and R2 (for polynomial quadratic model) were calculated and used to test the goodness-of-fit of non-nested and nested models (Archontoulis and Miguez 2015; Mayer and Butler 1993; Roman et al. 2000):

eq. [7]

eq. [8]

*R2* eq. [9]

where RSS and RSTare the sums of squares for the residual and total, respectively; is the number of data points; is the number of model parameters; is the observed, is the predicted, and is the mean observed value. The ME values range from -∞ and 1, with values closer to 1 indicating better predictions (Werle et al. 2014). R2 values range from 0 to 1, and it was used only for the polynomial quadratic model, which is a form of linear regression.

**Results and Discussion**

**Top model selection to describe crop-weed competition.** The retangular hyperbola model proposed (Cousens 1985) resulted in the lowest AIC (332.2), followed by a logistic model (337.6) and a polynomial quadratic model (343.1) (Table 1). The RMSE and ME resulted in similar trend for the models tested, except *R. brasiliensis* in the polynomial quadratic (Table 1), which resulted in highest ME (0.71) across models (Table 1). Therefore, MR (or R2) could potentially cause misleading model selection; thus the goodness of fit are not appropriated for selecting the top model.

According to AIC, the logistic model was the second best model to describe the data (Table 1). The maximum corn yield loss caused by the competition of *R. brasiliensis* and *C. benghalensis* (*d*) was 67.2% and 93.4%, respectively. The 50% yield loss (%) was 1.2 and 0.7 plants pot-1 of *R. brasiliensis* and *C. benghalensis*, respectively. The common concern of fitting a sigmoid curve for crop-weed competition study is the systematic lower limit (*c*) lack of fit (Table 3). The *P*-value for parameter *c* is not significant (P>0.05); therefore, this parameter is similar to zero. In this study, besides *c*, parameters *b* and *e* (*R. brasiliensis*) resulted in lack of fit as well. Moreover, the standard error is bigger than *b* (*C. benghalensis*) and *c* estimated values, which means that the sample mean is not close to the actual population mean.

Similar to the logistic model, the polynomial quadratic model provided the highest AICc (343.1), which means that this model was statistically the least appropriate for describing the data (Table 1). Similar trend was observed for RMSE. Nonetheless, the ME and R2, this model resulted in a good fit to the data.

**Model selection to evaluate weed competitiveness with the crop.** Based on AIC, the Cousens model was the top model that described the data (Table 1). Thus, the F-test to evaluate whether *R. brasiliensis* and *C. benghalensis* competed similarly with corn was conducted. The F-test of the Cousens model for corn yield loss (%) indicated that a reduced model (P=0.40) with different parameter *I* (corn yield at low weed densities) and similar parameter *A* (corn yield at higher densities) was the best model (Red. III) to describe corn competition to *R. brasiliensis* and *C. benghalensis* (Table 5). Therefore, the hypothesis of similar weed competition in corn was rejected. According to the parameter estimates in the rectangular hyperbola, at weed low densities (*I*), corn yield loss was 37.0 and 228.3 % in competition to *R. brasiliensis* and *C. benghalensis*, respectively. However, at higher densities, *R. brasiliensis* and *C. benghalensis* compete similarly, and corn yield loss was 106.1% (Figure 5 and Table 6). AICc corroborates to F-test (Table 5). The model selected by the F-test (different I, but similar A) resulted in the lowest AIC of 330.4. Despite RMSE is similar in Red. III and Full model, the highest ME (≥0.95) for *R. brasiliensis* and *C. benghalensis* would demonstrated the goodness of fit of the top model selected.

A complete review of model parameter *I* and *A* of the rectangular hyperbola is provided by a (Cousens 1985). Also, this model is recommended for crop-weed studies in additive design (Ritz et al. 2015; Swanton et al. 2015). Many authors have used this model to answer their research questions and improve weed control decision-making (Fischer et al. 2004a; Lindquist et al. 1996; Lindquist et al. 1999; Werle et al. 2014). For example, using the rectangular hyperbola model, it was concluded the higher competitive potential of *Amaranthus palmeri* in corn and *Kochia scoparia* in sunflower (Lewis and Gulden 2014; Massinga et al. 2001). Parameters *I* and *A* are also useful for estimating weed competition across different locations and for calculating weed thresholds ((Fischer et al. 2004b; Lindquist et al. 1996). Additionally, using parameters *I* and *A,* it was demonstrated that organic cropping systems have the potential to tolerate great abundance of weeds compared to conventional system (Ryan et al. 2009). Thus, the rectangular hyperbola proposed by Cousens (1985) and the F-test nested model selection are important and useful tools in crop-weed competition research.

Here it was demonstrated that the rectangular hyperbola proposed by Cousens (1985) was statistically and biologically the best model to describe crop-weed competition data from additive design. Potential issues, including parameter overestimation (>100%), were also addressed. Nonetheless, the Cousens (1985) model has a curve shape that fits well with the expected results from additive designs. The parameters *I* and *A* are easily interpreted and biologically meaningful. We propose Cousens (1985) as a standardized model for crop-weed competition studies in additive design. If the data trend seems linear, the experimental design needs to be adjusted by increasing the weed density to achieve the CFY. If the CFY is immediately reached at low-density treatments, density treatments need to be adjusted, or experimental units (e.g., pots) increased in size. Logistic models are recommended to another set of studies in weed research (e.g., herbicide dose-response), and polynomial quadratic curves are not recommended. This present study would aid statistical data analysis and interpretation of crop-weed competition from additive designs.

**Acknowledgments**

The authors thank CAPES (Brazilian Government Foundation) - Proc. no 9112-13-8, for financial support to the graduate student involved in this study.

References

Anderson D (2008) Model based inference in the life sciences: A primer on evidence. 2008.

Archontoulis SV and Miguez FE (2015) Nonlinear regression models and applications in agricultural research. Agron J 107(2):786-98

Cousens R (1985) A simple model relating yield loss to weed density. Ann Appl Biol 107(2):239-52

Ferreira EA, de Matos, Christiano da Conceição, Barbosa EA, Melo CAD, da Silva DV, dos Santos JB (2015) Aspectos fisiológicos de soja transgênica submetida à competição com plantas daninhas. Revista De Ciências Agrárias/Amazonian Journal of Agricultural and Environmental Sciences 58(2):115-21

Fischer DW, Harvey RG, Bauman TT, Phillips S, Hart SE, Johnson GA, Kells JJ, Westra P, Lindquist J (2004a) Common lambsquarters (chenopodium album) interference with corn across the northcentral united states. Weed Sci 52(6):1034-8

Fischer DW, Harvey RG, Bauman TT, Phillips S, Hart SE, Johnson GA, Kells JJ, Westra P, Lindquist J (2004b) Common lambsquarters (chenopodium album) interference with corn across the northcentral united states. Weed Sci 52(6):1034-8

Hurvich CM and Tsai C (1991) Bias of the corrected AIC criterion for underfitted regression and time series models. Biometrika 78(3):499-509

Knezevic SZ and Datta A (2015) The critical period for weed control: Revisiting data analysis. Weed Sci 63(sp1):188-202

Knezevic SZ, Streibig JC, Ritz C (2007) Utilizing R software package for dose-response studies: The concept and data analysis. Weed Technology 21(3):840-848

Knezevic SZ, Weise SF, Swanton CJ (1994) Interference of redroot pigweed (amaranthus retroflexus) in corn (zea mays). Weed Sci :568-73

Lewis DW and Gulden RH (2014) Effect of kochia (kochia scoparia) interference on sunflower (helianthus annuus) yield. Weed Sci 62(1):158-65

Lewis F, Butler A, Gilbert L (2011) A unified approach to model selection using the likelihood ratio test. Methods in Ecology and Evolution 2(2):155-62

Lindquist JL, Mortensen DA, Clay SA, Schmenk R, Kells JJ, Howatt K, Westra P (1996) Stability of corn (zea mays)-velvetleaf (abutilon theophrasti) interference relationships. Weed Sci :309-13

Lindquist JL, Mortensen DA, Westra P, Lambert W, Bauman TT, Fausey JC, Kells JJ, Langton SJ, Harvey RG, Bussler BH (1999) Stability of corn (zea mays)-foxtail (setaria spp.) interference relationships. Weed Sci :195-200

Massinga RA, Currie RS, Horak MJ, Boyer Jr J (2001) Interference of palmer amaranth in corn. Weed Sci 49(2):202-8

Mayer D and Butler D (1993) Statistical validation. Ecol Model 68(1):21-32

Mazerolle M (2016) AICcmodavg: Model Selection and Multimodel Inference Based on (Q) AIC (C)[Software]

Monquero PA, Orzari I, Silva PVd, Penha AdS (2015) Interferência de plantas daninhas em mudas de quatro espécies arbóreas neotropicais. Acta Scientiarum.Agronomy 37(2):219-32

Oliveira MC, Jhala AJ, Gaines T, Irmak S, Amundsen K, Scott JE, Knezevic SZ (2017) Confirmation and control of HPPD-inhibiting Herbicide–Resistant waterhemp (amaranthus tuberculatus) in nebraska. Weed Technol :1-13

Ritz C and Streibig JC (2008) Nonlinear regression with R. Springer Science & Business Media

Ritz C and Streibig JC (2005) Bioassay analysis using R. Journal of Statistical Software 12(5):1-22

Ritz C, Kniss AR, Streibig JC (2015) Research methods in weed science: Statistics. Weed Sci 63(sp1):166-87

Roman ES, Murphy SD, Swanton CJ (2000) Simulation of chenopodium album seedling emergence. Weed Science 48(2):217-224

Ryan M, Smith R, Mortensen D, Teasdale J, Curran W, Seidel R, Shumway D (2009) Weed–crop competition relationships differ between organic and conventional cropping systems. Weed Res 49(6):572-80

Silva DV, Pereira GAM, de Freitas, Marco Antônio Moreira, da Silva AA, Sediyama T, Silva GS, Ferreira LR, Cecon PR (2015) Produtividade e teor de nutrientes do milho em consórcio com braquiária. Ciencia Rural 45(8):1394-400

Strieder ML, Silva PR, Argenta G, Rambo L, Sangoi L, Silva AA, Endrigo PC (2007) A resposta do milho irrigado ao espaçamento entrelinhas depende do híbrido e da densidade de plantas. Ciência Rural 37(3):634

Sugiura N (1978) Further analysts of the data by akaike's information criterion and the finite corrections: Further analysts of the data by akaike's. Communications in Statistics-Theory and Methods 7(1):13-26

Swanton CJ, Nkoa R, Blackshaw RE (2015) Experimental methods for crop-weed competition studies. Weed Sci 63(sp1):2-11

Trezzi M, Vidal R, Patel F, Miotto E, Debastiani F, Balbinot A, Mosquen R (2015) Impact of conyza bonariensis density and establishment period on soyabean grain yield, yield components and economic threshold. Weed Res 55(1):34-41

Voll E, Gazziero D, Brighenti A, Adegas F (2002) Competição relativa de espécies de plantas daninhas com dois cultivares de soja. Planta Daninha 20(1):17-24

Weiner J and Freckleton RP (2010) Constant final yield. Annual Review of Ecology, Evolution and Systematics 41:173-92

Werle R, Schmidt JJ, Laborde J, Tran A, Creech CF, Lindquist JL (2014) Shattercane X ALS-tolerant sorghum F1 hybrid and shattercane interference in ALS-tolerant sorghum. J Agric Sci 6(4):p159

Zucchini W (2000) An introduction to model selection. J Math Psychol 44(1):41-61

Zuur A, Ieno EN, Smith GM (2007) Analysing ecological data. Springer Science & Business Media