WEED RESEARCH An International Journal of Weed Biology, Ecology and Vegetation Management

DOI: 10.1111/wre.12317

Additive design: the concept and data analysis

M C OLIVEIRA*, G A M PEREIRA†, E A FERREIRA‡, J B SANTOS‡, S Z KNEZEVIC§ & R WERLE*

*Department of Agronomy, University of Wisconsin-Madison, Madison, WI, USA, †Departamento de Fitotecnia, Universidade Federal de Viçosa, Viçosa, MG, Brazil, ‡Departamento de Agronomia, Universidade Federal dos Vales do Jequitinhonha e Mucuri, Diamantina, MG, Brazil, and §Department of Agronomy and Horticulture, University of Nebraska-Lincoln, Concord, NE, USA

Received 26 January 2018
Revised version accepted 30 April 2018
Subject Editor: Clarence Swanton, Guelph, ON, Canada

Summary

Crop–weed competition is extensively studied in weed science. The additive design, in which weed density varies and the crop density is kept constant, is the most commonly utilised design in plant competition studies. The additive design is important to calculate economic weed thresholds and improve weed control decision-making. Crop–weed competition studies are usually conducted by weed scientists, who sometimes report misleading conclusions because of lack of statistical knowledge needed for data analysis of such studies. Therefore, the objective of this manuscript is to provide the concept of additive design and demonstrate the model selection approach for describing crop–weed density relationship to non-statisticians. We evaluated three models routinely used in the literature to interpret data from additive designs, including polynomial quadratic, sigmoid and rectangular hyperbola curves. Based on the described statistical criteria, we demonstrated the rectangular hyperbola to be the most appropriate model to describe data from an additive design study looking at *Richardia brasiliensis* and *Commelina benghalensis* competition with maize (*Zea mays*). Moreover, we describe step-by-step how to perform the statistical analysis in R software and interpret the results of crop–weed competition studies. We suggest the use of the rectangular hyperbola as a standardised model for crop–weed competition in additive design.

Keywords: crop–weed competition, model selection, rectangular hyperbola, weed interference, weed threshold.

Oliveira MC, Pereira GAM, Ferreira EA, Santos JB, Knezevic SZ & Werle R (2018). Additive design: the concept and data analysis. *Weed Research*. https://doi.org/10.1111/wre.12317.

Introduction

One of the most common dilemmas that farmers and practitioners face is how to decide on the timing of a weed control operation, or when to spray an herbicide. Before initiating weed control procedures, the following are some general guidelines to consider: field scouting and mapping weed patches and utilising the concepts of the critical period of weed control, weed thresholds, and decision support with computer models. Field scouting typically involves assessing the type and number of weeds, to determine whether a spray operation is necessary. Mapping and monitoring weed patches over time will also help to determine the effectiveness of the control programme.

Studies of crop–weed competition show that yield loss (YL) is sensitive to differences in the period between crop and weed emergence (Knezevic *et al.*, 1997; Hock *et al.*, 2006). It brings to light the importance of the concepts of a critical period of weed

Correspondence: Maxwel Coura Oliveira, Department of Agronomy, University of Wisconsin-Madison, Madison, WI 53506, USA. Tel: (+1) 608 262 7130; Fax: (+1) 608 262 5217; E-mail: max.oliveira@wisc.edu

control (Knezevic *et al.*, 2002; Knezevic & Datta, 2015) and economic thresholds (Coble & Mortensen, 1992; Wilkerson *et al.*, 2002). A weed threshold is described as 'a point at which weed density causes important crop losses' (Knezevic *et al.*, 2017). Knowledge of thresholds can help agriculturists make decisions on the need for herbicide applications, in deciding whether remedial weed control efforts are necessary or economically justified.

An economic weed threshold has been defined as 'the weed density at which the cost of weed control equals the increased return on yield in the current year' (Knezevic *et al.*, 2017). Because they account for crop losses only in the current cropping season, economic thresholds are single-year measures of weed effects. Also, economic thresholds are based on factors such as the price of the crop at harvest, herbicide and application cost, anticipated crop yield, and the YL–weed density relationships which are a function of environmental factors (e.g. soil types and climate). Since the primary cause of yield reductions by weeds is through competition for growth-limiting resources (light, water and nutrients), the economic threshold is not constant for particular weed–crop combinations and can differ within and across geographic regions.

In crop–weed competition, the additive design study is a primary step for calculating thresholds. In additive design, the weed density varies while crop density is kept constant (Swanton *et al.*, 2015). Several review papers recommend the use of rectangular hyperbola for crop–weed competition studies in the weed science literature (Cousens, 1985; Knezevic & Horak, 1998; Ritz *et al.*, 2015; Swanton *et al.*, 2015). However, there is still a distinct number of empirical models used for additive design studies (Voll *et al.*, 2002; Strieder *et al.*, 2007; Silva *et al.*, 2015; Trezzi *et al.*, 2015). Four regressions curves are frequently used: linear (Fig. 1A), polynomial quadratic (Fig. 1B), sigmoid (Fig. 1C) and rectangular hyperbola (Fig. 1D). The commonly used criteria for selection of linear and non-linear regression models is the equation with highest R-squared (R^2; Archontoulis & Miguez, 2015). The R^2 tests the goodness-of-fit for linear models and is statistically inadequate for non-linear model selection (Zuur *et al.*, 2007; Archontoulis & Miguez, 2015). There are several appropriate statistical criteria for selecting a non-linear model for datasets: Akaike's information criterion (AIC), Bayesian information criterion (BIC), *F*-test and likelihood ratio (Zucchini, 2000; Anderson, 2007; Lewis *et al.*, 2011). These statistical criteria are used according to the model structure, which can be non-nested or nested. Non-nested are models with different structure and parameters, such as an exponential decay and a rectangular hyperbola model. The AIC and BIC are indicated for non-nested model selection. In contrast, nested are models that are a particular case of each other and have identical terms, whereas one must have at least one additional term (e.g. three- and four-parameter log-logistic models). The

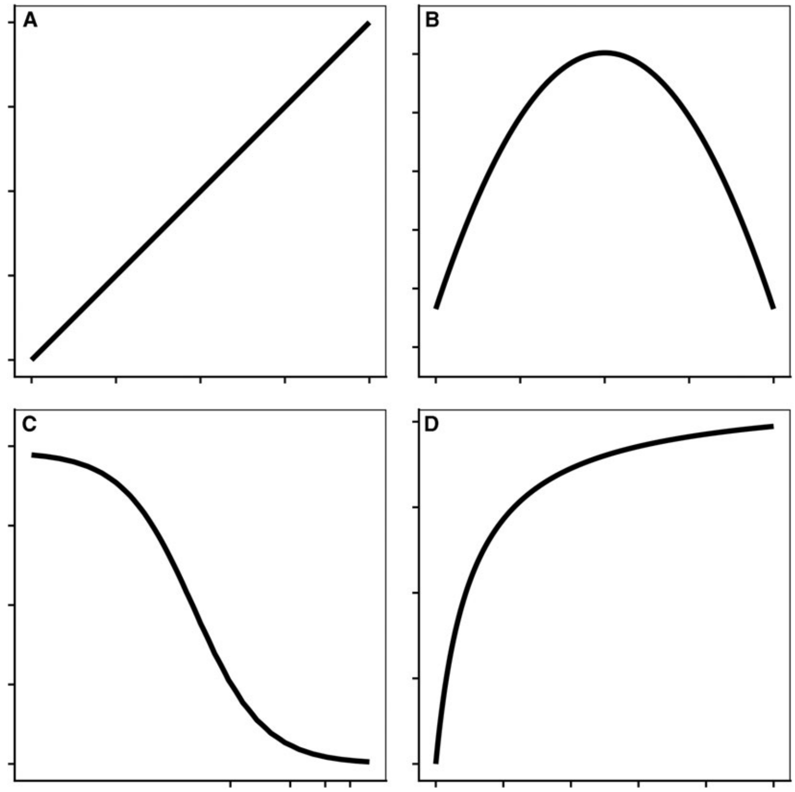

Fig. 1 Common regression curves used to describe the data from crop–weed competition studies in additive design: (A) linear; (B) polynomial quadratic; (C) sigmoid; (D) rectangular hyperbola.

AIC, BIC, F-test and likelihood ratio are appropriate selection of non-linear nested models.

From a practical standpoint, the best model should be selected based on a balance between statistics and biological relevance, which will help scientists answer their research questions (Onofri *et al.*, 2010; Werle *et al.*, 2014a,b; Archontoulis & Miguez, 2015). In additive design studies, the model that provides a good fit and essential biological parameters is considered a strong candidate model to describe the dataset. The advances in statistical software have facilitated the use of standardised non-linear regression analysis that can be performed by non-statisticians (Knezevic *et al.*, 2007). Therefore, the objectives of this manuscript were to:

1 Provide basic knowledge about the concept of crop–weed competition and additive designs.
2 Test the suitability of three non-nested candidate models (polynomial quadratic, sigmoid and a rectangular hyperbola) for describing crop–weed competition in additive design.
3 Test the null hypothesis that the weed species *Richardia brasiliensis* Gomes and *Commelina benghalensis* L. compete similarly with maize (*Zea mays* L.). This hypothesis was tested after model selection (objective 2) using the F-test.

Data from an experiment under glasshouse conditions looking at *R. brasiliensis* and *C. benghalensis* competition with maize were used for the model selection exercise. The data analysis concept presented here apply to other weed species and crops studied in glasshouse or field conditions.

Materials and methods

Plant material and growth conditions

Seed heads of *R. brasiliensis* were harvested along roadsides near Diamantina, Minas Gerais (MG), Brazil in March of 2011 and 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), stolons (vegetative propagules) of *C. benghalensis* were collected in wetlands, near Diamantina, MG. Seeds of *R. brasiliensis* and stolons of *C. benghalensis* were seeded and transplanted to separate trays (1210 cm^3) filled with red latosol soil (pH 6.1 and 1% organic matter). A single seed of glyphosate-resistant (GR) maize was sown in 8 L plastic pots filled with the aforementioned soil. This procedure was performed to maximise the competition between species. The soil was fertilised following the local recommendations and N was applied at 15 and 30 days after maize emergence (DAE) at a rate of 55 mg L^{-1} of ammonium sulphate. Glasshouse conditions were 28/19°C day/night, and pots were watered daily.

Experimental procedures

The experiment was conducted under glasshouse conditions over a period of 60 days at the Universidade Federal dos Vales do Jequitinhonha e Mucuri, Diamantina, MG, Brazil. In this study, the additive design was used (e.g. weed densities varied and maize 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 randomised design with four replications.

Maize 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 maize dry matter (g) data (shoot) were converted into YL (%) compared with the control treatment (no weeds):

$$\text{Yield loss}\,(\%) = \frac{M - B}{M} * 100 \quad (1)$$

where M is the mean maize dry mass in the absence of weed competition (g) of the control treatment, and B is the dry mass (g) of individual maize plants competing with weed(s).

Statistical analysis

Three models were fitted to YL data (%) in response to weed density (plants pot^{-1}):

Rectangular hyperbola model proposed by Cousens (1985):

$$\text{YL} = \frac{I * x}{1 + (\frac{I}{A}) * x} \quad (2)$$

where I represents YL (yield loss) per unit weed density as x (density) approaches 0, and A represents YL as x approaches ∞ (or maximum expected YL). The rectangular hyperbola model was fitted using the *nls* function of R version 3.3.1 (R Core Team, 2018).

Sigmoid model (four-parameter log-logistic curve):

$$\text{YL} = c + \frac{d - c}{1 + \exp[b(\log x - \log e)]} \quad (3)$$

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 causes 50% YL (inflection point). The parameter b is the relative slope around the parameter e, and x is the number of weeds pot^{-1}. Parameters for the sigmoidal model (four-parameter logistic) were estimated using the *drm* function of drc package in R software (Ritz & Streibig, 2005).

Polynomial quadratic model (second order):

$$\mathrm{YL} = \alpha + ax + bx^2 \tag{4}$$

where α is the intercept in the *y*-axis (no YL by weed competition), *a* represents the slope of the model. The parameter *b* is the quadratic term of the model, and *x* is the number of weeds pot^{-1}. The parameters for the polynomial quadratic equation were estimated using the *lm* function of R software.

Model selection to describe crop–weed competition

The $\mathrm{AIC_c}$ (corrected AIC for finite sample size) criterion, which is indicated for non-nested model selection (Sugiura, 1978; Hurvich & Tsai, 1991), was calculated as:

$$\mathrm{AIC_c} = -2\log(l) + 2K * (n/(n - K - 1)) \tag{5}$$

where *l* is the likelihood function and *K* is the number of estimated parameters in the model, and *n* is the sample size of the model. According to the $\mathrm{AIC_c}$ criterion, the best model has the lowest $\mathrm{AIC_c}$ value. The $\mathrm{AIC_c}$ values for each model were estimated using the *AICc* command of package *AICcmodavg* in R software.

Model selection to evaluate weed competitiveness with the crop

Assuming that rectangular hyperbola is the best model, the impact of *R. brasiliensis* and *C. benghalensis* on maize YL is accessed through the variance ratio or *F*-test performed using Eqn (2) (Lindquist *et al.*, 1996). This statistical procedure evaluates the difference of residual sum squares (RSS) of nested models. The *F*-test is calculated as:

$$F = \frac{(\mathrm{RSS_{RED}} - \mathrm{RSS_{FULL}})/(\mathrm{d.f._{RED}} - \mathrm{d.f._{FULL}})}{\mathrm{RSS_{FULL}}/\mathrm{d.f._{FULL}}} \tag{6}$$

where $\mathrm{RSS_{RED}}$ and $\mathrm{RSS_{FULL}}$ represent the minimised RSS of the parameters estimated for the full (step 1) and reduced model (step 2, 3 or 4; steps 1 through 4 are described next), respectively; $\mathrm{d.f._{RED}}$ and $\mathrm{d.f._{FULL}}$ represent the degrees of freedom of the full and reduced models respectively. In practical terms, when *P* value of the *F*-test >0.05, we fail to reject the null hypothesis and a reduced model should be used (no difference between *I* and/or *A* parameter values between weed species). However, if *P* of the *F*-test <0.05, the null hypothesis is rejected and the Full model should be used (different *I* and/or *A* parameter values for each weed species). The *F*-test principle for non-linear regression analysis was calculated for each model using *nls ANOVA* command in R software (Ritz & Streibig, 2008).

Four significant steps need to be completed to compare the parameters using this method (see Appendix S1 for statistical codes to perform these steps in R software):

Step 1: Fit Eqn (2) to the data of each species individually (*R. brasiliensis* and *C. benghalensis*); single model having separated *I* and *A* parameters for each species. This represents the Full model, where four parameter values (*I* and *A* for each weed species) are estimated.

Step 2: Pool the data for both species (*R. brasiliensis* and *C. benghalensis*) and fit Eqn (2). This represents the reduced model (Red.1), where two parameter values (*I* and *A* for both weed species combined) are estimated for the pooled 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 maize. If the hypothesis is accepted ($P > 0.05$), stop here. Otherwise, there are two additional hypotheses to be tested (steps 3 and 4).

Step 3: Fit Eqn (2) setting a single parameter *I*, but different *A* parameter for each species (Red.2). 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*).

Step 4: Fit Eqn (2) setting a single parameter *A*, but different *I* parameters for each species. This is a reduced model (Red.3), 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 $\mathrm{AIC_c}$ was also performed for the nested model selection for confirming the *F*-test model selection.

Model goodness-of-fit

Root mean squared error (RMSE) and modelling efficiency (ME) were calculated and used to test the goodness-of-fit of non-nested and nested models (Mayer & Butler, 1993; Roman *et al.*, 2000):

$$\mathrm{RMSE} = \sqrt{\frac{\mathrm{RSS}}{n - p - 1}} \tag{7}$$

$$\mathrm{ME} = 1 - \left[\frac{\sum_{i=1}^{n}(Oi - Pi)^2}{\sum_{i=1}^{n}(Oi - \bar{O}i)^2}\right] \tag{8}$$

where RSS is the residual sums of squares; *n* is the number of data points; *p* is the number of model parameters; *Oi* is the observed, *Pi* is the predicted and

$\bar{O}i$ is the mean observed value. The ME values range from $-\infty$ and 1, with values closer to 1 indicating better predictions (Werle *et al.*, 2014c).

Results

Best model selection to describe crop–weed competition

The rectangular hyperbola model resulted in the lowest AIC_c (332.2), followed by a sigmoid model (337.6) and a polynomial quadratic model (343.1). The RMSE and ME resulted in a similar trend for the models tested, except *R. brasiliensis* in the polynomial quadratic model (Table 1). The rectangular hyperbola was also the top model for describing maize leaf area, height and stem diameter reduction in response to *R. brasiliensis* and *C. benghalensis* densities (data not shown).

In the rectangular hyperbola model (Fig. 2), four parameters were estimated, which are *I* and *A* for *R. brasiliensis* and *C. benghalensis*. The parameters *I* and *A* for *R. brasiliensis* were estimated at 50.3% and 82.1% respectively (Table 2). In contrast, for *C. benghalensis,* parameter estimations were 210.2% (*I*) and 108.6% (*A*). Also, the *P* value is significant for parameters *I* and *A* of both weed species ($P < 0.05$), indicating they are different to zero.

According to AIC_c, the sigmoid model was the second best model to describe the data (Table 1). The maximum maize YL caused by the competition of *R. brasiliensis* and *C. benghalensis* (*d*) was 67.2% and 93.4% respectively. The 50% maize YL (%) was achieved at 1.2 and 0.7 plants pot^{-1} of *R. brasiliensis* and *C. benghalensis* respectively. However, the sigmoid model had non-significant parameters (values not different to zero, $P > 0.05$) for both weed species, including *b*, *c* and *e* (*R. brasiliensis* only) (Table 3). The lack of data points around *e* can be observed in Fig. 3.

Table 1 Non-nested model selection (rectangular hyperbola, sigmoid and polynomial quadratic) for an additive design study of maize in competition with *Richardia brasiliensis* and *Commelina benghalensis* under glasshouse conditions

		Model selection*	Goodness-of-fit†	
Model	Species	AIC_c	RMSE	ME
Rectangular hyperbola	*R. brasiliensis*	332.2	2.2	0.64
	C. benghalensis			0.92
Sigmoid	*R. brasiliensis*	337.6	2.3	0.58
	C. benghalensis			0.85
Polynomial quadratic	*R. brasiliensis*	343.1	3.3	0.71
	C. benghalensis			0.90

*AIC_c, corrected Akaike's information criterion.
†RMSE, Root mean square error; ME, modelling efficiency.

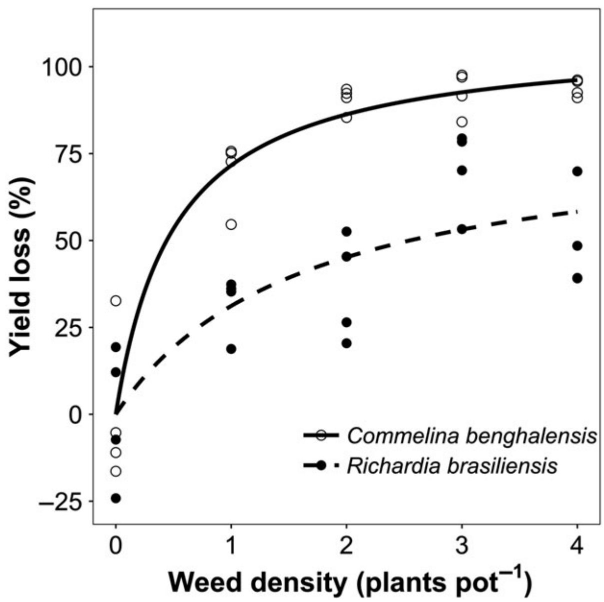

Fig. 2 The relationship between maize yield loss (%) and weed density (plants pot^{-1}) described with a rectangular hyperbola model (Full model).

Also, the standard error in *b* and *c* parameters is bigger than the estimated values (Table 3), indicating the lack of fit of a sigmoid curve. The RMSE for the sigmoid model was 2.3, and ME was 0.58 and 0.85 for *R. brasiliensis* and *C. benghalensis* respectively.

The polynomial quadratic model had the highest AIC_c (Table 1 and Fig. 4). A similar trend was observed for RMSE. However, ME of *R. brasiliensis* was highest (0.71) across the three models tested (Table 1). The α was non-significant ($P > 0.05$) for both weed species. The *a* and *b* parameters were significant ($P < 0.05$). The parameter α was 35.5% and 65.5%, and *b* −5.4 and −11.1 for *R. brasiliensis* and *C. benghalensis* respectively (Table 4).

Model selection to evaluate weed competitiveness with the crop

Based on AIC_c, the rectangular hyperbola was the best model to describe the data (Table 1). The *F*-test of the rectangular hyperbola (Full model) indicated that a reduced model with different parameter *I* (maize yield at low weed densities) and similar parameter *A* (maize yield at higher densities) was the best model (Red.3) to describe competition of *R. brasiliensis* and *C. benghalensis* with maize (Table 5). As demonstrated in the steps in the Appendix S1, the Red.1 and Red.2 models were different from the Full model ($P < 0.05$); thus, the hypotheses tested in those models were rejected (Table 5).

Table 2 Rectangular hyperbola (Full model) parameter estimates for maize yield loss (%) in competition with *Richardia brasiliensis* and *Commelina benghalensis* under glasshouse conditions

Parameters*	Species	Estimate (±SE)[†], %	95% CI (lower-upper)[‡], %	*t* value	*P* value
I	*R. brasiliensis*	50.3 (±22.6)	4.3–96.1	2.2	0.032
	C. benghalensis	210.2 (±88.6)	30.7–389.8	2.4	0.023
A	*R. brasiliensis*	82.1 (±23.1)	35.3–128.8	3.6	0.001
	C. benghalensis	108.6 (±11.1)	85.3–131.2	9.7	<0.001

*I, represents maize yield loss (%) per unit weed density as density approaches 0; A, represents maize yield loss (%) as density approaches ∞ (or maximum expected yield loss).
[†]SE, Standard Error.
[‡]95% CI, Confidence Interval.

Table 3 Sigmoid parameter estimates for maize yield loss (%) in competition with *Richardia brasiliensis* and *Commelina benghalensis* under glasshouse conditions

Parameters*	Species	Estimate (±SE)[†], %	*t* value	*P* value
b	*R. brasiliensis*	−1.5 (±1.4)	−1.1	0.291
	C. benghalensis	−3.2 (±5.1)	−0.6	0.541
c	*R. brasiliensis*	0.2 (±7.4)	0.0	0.980
	C. benghalensis	−5.3 (±7.4)	0.0	0.999
d	*R. brasiliensis*	67.2 (±26.9)	2.5	0.017
	C. benghalensis	93.4 (±8.4)	11.1	<0.001
e	*R. brasiliensis*	1.2 (±0.7)	1.6	0.047
	C. benghalensis	0.7 (±0.3)	2.1	0.124

*b, slope; c, lower limit (weed competition at low densities); d, upper limit (maximum expected maize yield loss, %); e, inflection point (weed density which maize yield loss is 50% relative to d).
[†]SE, Standard Error.

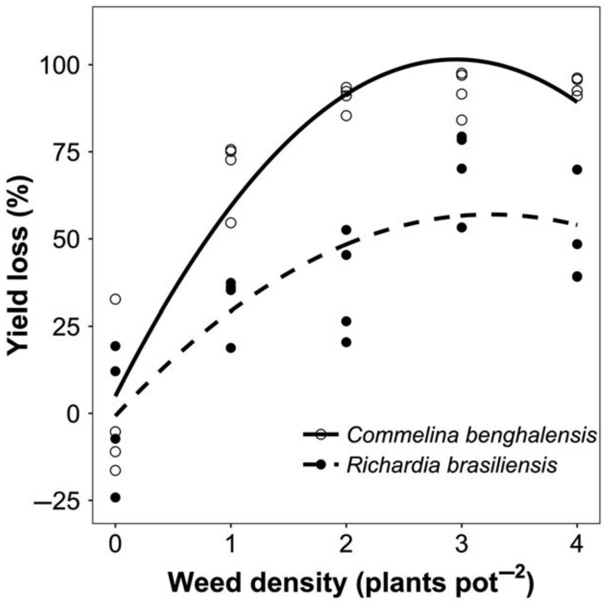

Fig. 4 The relationship between maize yield loss (%) and weed density (plants pot$^{-1)}$ described with a polynomial quadratic model.

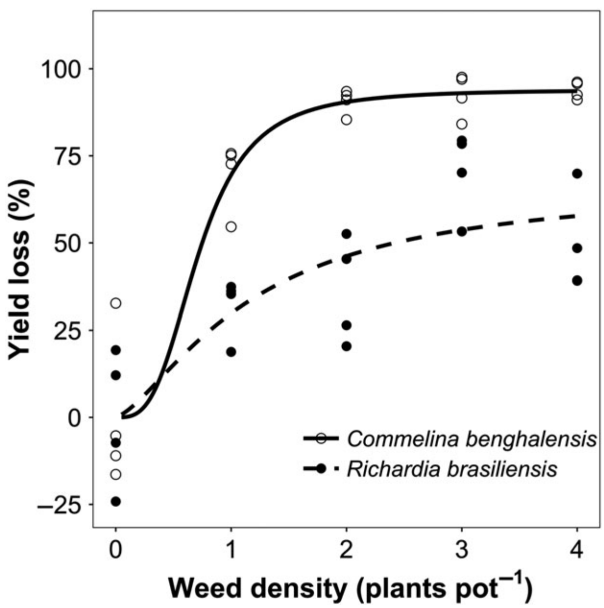

Fig. 3 The relationship between maize yield loss (%) and weed density (plants pot^{-1}) described with a sigmoid model.

Table 4 Polynomial quadratic parameters estimates for maize yield loss (%) in competition with *Richardia brasiliensis* and *Commelina benghalensis* under glasshouse conditions

Parameters*	Species	Estimate (±SE)[†], %	*t* value	*P* value
α	*R. brasiliensis*	−0.7 (±7.7)	−0.1	0.926
	C. benghalensis	4.9 (±6.1)	0.8	0.431
a	*R. brasiliensis*	35.5 (±9.1)	3.8	0.001
	C. benghalensis	65.5 (±7.3)	9.0	<0.001
b	*R. brasiliensis*	−5.4 (±2.2)	−2.5	0.024
	C. benghalensis	−11.1 (±1.7)	−6.4	<0.001

*α, intercept at Y value when density equals zero; a, is the slope of the equation; b, is the quadratic term of the equation.
[†]SE, Standard Error.

According to the parameter estimates in the rectangular hyperbola Red.3 model (Fig. 5), at low weed densities (I), maize YL was 37.0 and >100% in competition to *R. brasiliensis* and *C. benghalensis* respectively. However, at higher densities, *R. brasiliensis* and *C. benghalensis* compete similarly, and maize YL was 106.1% (Table 6). The AIC_c results corroborate the *F*-test, as the model selected by the *F*-test (different *I*, but similar *A*) resulted in the lowest AIC_c of 330.4. The RMSE was similar in Red.3 and Full model, but the highest ME for *R. brasiliensis* (0.95) and *C. benghalensis* (0.98) demonstrated the goodness-of-fit of the best model selected (Red.3).

Discussion

Model selection to describe crop–weed competition

Amongst the non-nested models tested, the rectangular hyperbola was the best model to describe maize YL (%) in response to both *R. brasiliensis* and *C. benghalensis* competition (Table 1). The model with the smallest value of AIC_c was considered the best model or the best descriptor of the full reality given the set of candidate models and the data (Anderson, 2007).

The *I* parameter of *C. benghalensis* was not biologically meaningful (>100%; Table 2). The curve has a steep inclination (Fig. 2), which is likely due to the relatively small pot size used in this study, indicating that *C. benghalensis* is very competitive with maize. Therefore, bigger pots and lower *C. benghalensis* densities would have been necessary for biologically meaningful estimation of *I* parameter. Also, the parameter *A* (*C. benghalensis*) is slightly above 100%, which means that at higher densities *C. benghalensis* is likely to suppress maize completely. Nonetheless, the maximum biological maize YL (100%) is within the confidence intervals for parameter *I* and *A* of *C. benghalensis* (Table 2). The parameters *I* and *A* were also not biologically meaningful (>100%) in a field study of *Amaranthus retroflexus* in competition with *Sorghum bicolor* (Knezevic & Horak, 1998). Herein, we highlight potential issues with experimental design and data analysis commonly occurring to either greenhouse or field studies.

Constraining parameter values is a common practice for non-linear modelling (Fox & Weisber, 2011). For example, the function *nls* in R software allows boundaries, such as setting the *A* parameter of the rectangular hyperbola = 100%, which was done by

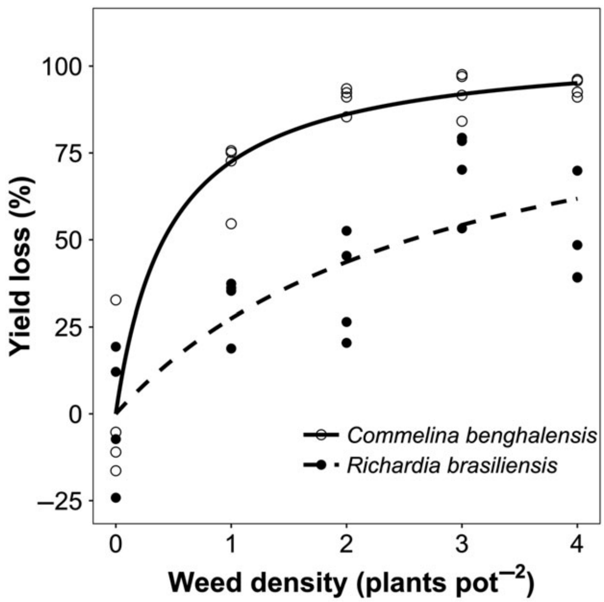

Fig. 5 The relationship between maize yield loss (%) and weed density (plants pot^{-1}) described with a rectangular hyperbola model (Red.3).

Table 5 Hypothesis testing, nested model selection criteria and goodness-of-fit of the rectangular hyperbola model parameters *I* and *A* of maize yield loss (%) in competition with *Richardia brasiliensis* and *Commelina benghalensis* under glasshouse conditions

		Model selection*			Goodness-of-fit†	
		F-test		AIC_c		
Hypothesis	Species	*F* value	*P* value		RMSE	ME
Different *I* and *A* (Full)	*R. brasiliensis*	-	-	332.2	2.2	0.64
	C. benghalensis					0.92
Similar *I* and *A* (Red.1)	*R. brasiliensis*	32.3	0.00	368.2	3.6	0.84
	C. benghalensis					
Similar *I* but different *A* (Red.2)	*R. brasiliensis*	4.1	0.04	333.9	2.3	0.69
	C. benghalensis					0.97
Similar *A* but different *I* (Red.3)	*R. brasiliensis*	0.7	0.39	330.4	2.2	0.95
	C. benghalensis					0.98

**F*-test model selection; if *P* value<0.05: significantly different models; if *P* value > 0.05: non-significantly different models; AICc, Akaike's information criterion; †RMSE, Root mean square error; ME, modeling efficiency.

Parameters*	Species	Estimate (±SE)[†], %	95% CI (lower-upper)[‡], %	t value	P value
I	*R. brasiliensis*	37.0 (±6.2)	24.4–49.6	5.9	<0.001
	C. benghalensis	228.3 (±100.2)	25.4–431.3	2.3	0.028
A	*R. brasiliensis*	106.1 (±10.3)	85.3–127.1	10.3	<0.001
	C. benghalensis				

Table 6 Rectangular hyperbola (Red.3 model) parameters estimates for maize yield loss (%) in competition with *Richardia brasiliensis* and *Commelina benghalensis*

**I*, represents maize yield loss (%) per unit weed density as density approaches 0; *A*, represents maize yield loss (%) as density approaches ∞ (or maximum expected yield loss).
[†]SE, Standard Error.
[‡]95% CI, Confidence Interval.

Barnes *et al*. (2018) in their study evaluating competition of *Ambrosia artemisiifolia* with soybeans. Constraining an upper limit (e.g. maximum $A = 100\%$) might be a valuable tool to improve the biological meaning of parameter *A* (no >100% YL), but not for parameter *I*. Setting limits to parameter *I* would impact in the inclination of the curve (slope), potentially misleading research findings. In our analysis, for the exercise proposes, we have not set an upper limit to parameter *A*. But the Appendix S1 shows an example of how to set an upper limit to parameter *A*.

To understand the nature of crop–weed competition modelling, one needs to comprehend the concept of constant final yield (CFY). The CFY is described from low to high densities, whereas the relationship between total biomass per unit area and density is initially linear, but eventually reaches a plateau (e.g. YL remains constant despite the increase in density; Weiner & Freckleton, 2010). In our study, the CFY was reached at low density of *C. benghalensis*. As a result, the estimation of parameter *I* and *A* from *C. benghalensis* was above 100% (Table 2). In contrast, for *R. brasiliensis*, *I* and *A* were reached below 100%. Thus, the weed density for attaining CFY can vary amongst species. Other studies showed that CFY was reached with the estimation of *I* and *A* under 100%, indicating that some weed species may not lead to total crop YL (Knezevic *et al*., 1997; Knezevic & Horak, 1998). Consequently, for proper additive design studies, different weed densities based on the competitive potential of each species might be necessary. Additionally, a competition study that reports a linear relationship trend between crop YL and weed density has not reached CFY (Fig. 1A). It is likely that either the appropriated weed density for the study was not selected or plants were harvested before significant competition occurred (Weiner & Freckleton, 2010). Therefore, the crop–weed competition experiments need to be properly designed so CFY is achieved and model parameter estimates are statistically accurate and biologically meaningful.

According to AIC_c, the sigmoid was the second model to best describe the data (Table 1). The sigmoid model does not seem to be appropriate to describe the data from additive design studies (Fig. 3). The symmetric shape of sigmoid models is related to the rate of change. One of the assumptions when using the sigmoid (logistic) model is that the inflection point (*e*) is always at 50% of the asymptotic size (Knezevic *et al*., 2007; Ritz *et al*., 2015). Therefore, sigmoid curves have no biological meaning for competition studies in additive design (rectangular hyperbola pattern). Although the sigmoid model is not recommended for additive design, it is one of the most commonly used and appropriate models in other weed research topics. Sigmoid curves are extensively utilised for predicting weed emergence (Werle *et al*., 2014a,b), herbicide dose–response (Ritz *et al*., 2015) and critical time for weed removal (Knezevic & Datta, 2015). For example, in herbicide dose–response studies, the parameter *e* is meaningful and important for comparison of herbicides doses that control 50% of a weed population (Oliveira *et al*., 2017).

The polynomial quadratic model was statistically the least appropriate for describing the data. The α and *a* parameters estimated from a polynomial quadratic model possibly have biological meaning. However, the *b* parameter does not. Nonetheless, this model does not provide biological parameters that would improve the discussion, test hypothesis and help researchers understand the results from crop–weed competition studies. Also, the polynomial quadratic curve is symmetric around the *x*-axis (Fig. 1B), which makes such response biologically unlikely in an additive design study. For example, the predicted maize YL (%) is lower at four plants pot^{-1} than at three plants pot^{-1} (Fig. 4). The highest ME for *R. brasiliensis* could potentially mislead model selection; thus, ME should be used as a goodness-of-fit indicator and not for model selection. Therefore, a polynomial quadratic curve is not recommended for describing the results of additive design studies.

In additive design studies, because of misleading model selection using goodness-of-fit (usually R^2), it is common to find multiple equations describing response variables (Ferreira *et al*., 2015; Silva *et al*., 2015). For example, over six models were used to describe the competition of two weed species [*Urochloa decumbens* (Stapf) R.D. Webster and *Ipomoea grandifolia* (Dammer) O'Donell] with three neotropical trees [*Senegalia polyphylla* (DC.) Britton & Rose, *Ceiba speciosa* (A.St.-Hil.) Ravenna, and *Luehea divaricata* Mart] (Monquero *et al*., 2015). It becomes difficult to evaluate and compare weed competitiveness when different equations with non-related parameters are used.

Model selection to evaluate weed competitiveness with the crop

It was statistically and biologically demonstrated that the rectangular hyperbola model was the best model to describe crop–weed competition in an additive design. The *F*-test nested model selection showed that at high densities (*A*), competition of *R. brasiliensis* and *C. benghalensis* in maize YL is similar, but different at low densities (*I*; Table 5). Therefore, the model Red. 3 is selected and the hypothesis that competition of *R. brasiliensis* is similar to *C. benghalensis* in maize was partially rejected. The practical implication of this model selection exercise is that, at low densities, *C. benghalensis* is more competitive than *R. brasiliensis* in maize (under glasshouse conditions).

A complete review of parameter *I* and *A* of the rectangular hyperbola model is provided by Cousens (1985). Many authors have used this model to answer their research questions and improve weed control decision-making (Lindquist *et al*., 1999; Cathcart & Swanton, 2003; Fischer *et al*., 2004; Werle *et al*., 2014c). For example, using parameters *I* and *A*, it was demonstrated that organic cropping systems have the potential to tolerate greater abundance of weeds compared with conventional systems (Ryan *et al*., 2009). Additionally, using the rectangular hyperbola model, the high competitive potential of *Amaranthus palmeri* in maize (Massinga *et al*., 2001), *Kochia scoparia* in sunflower (Lewis & Gulden, 2014) and interference of *A. artemisiifolia* in soybeans (Barnes *et al*., 2018) was demonstrated. Parameters *I* and *A* are also useful for estimating weed competition across different locations and appropriate for calculating economic weed thresholds (Lindquist *et al*., 1996; Lindquist & Mortensen, 1998). Thus, the rectangular hyperbola proposed by Cousens (1985) and the *F*-test nested model selection are important and useful tools of crop–weed competition in additive design research.

Here, we demonstrate that the rectangular hyperbola was statistically and biologically the best model to describe crop–weed competition using data from this additive design study. Potential issues, including not biologically meaningful parameters (>100%), were also addressed. Nonetheless, the rectangular hyperbola model has an asymptote curve shape that fits well with the expected results from most additive design studies. The *I* and *A* parameters are easily interpreted and biologically meaningful. We suggest the rectangular hyperbola as a standard model for crop–weed competition studies in additive design. Sigmoid models are adequate for some other studies (e.g. herbicide dose–response), and the use of polynomial quadratic curves is not encouraged in weed research. Also, a step-by-step statistical analysis with R codes was performed for a competition study in additive design; we believe that this review will assist data analysis performed by non-statisticians studying crop–weed competition.

Acknowledgements

The authors thank CAPES (Brazilian Government Foundation), Proc. n° 9112-13-8, for financial support first author of this manuscript.

References

Anderson DR (2007) *Model Based Inference in the Life Sciences: A Primer on Evidence*. Springer, New York, NY, USA.

Archontoulis SV & Miguez FE (2015) Nonlinear regression models and applications in agricultural research. *Agronomy Journal*, **107**, 786–798.

Barnes ER, Jhala AJ, Knezevic SZ, Sikkema PH & Lindquist JL (2018) Common ragweed (*Ambrosia artemisiifolia* L.) interference with soybean in Nebraska. *Agronomy Journal*, **110**, 1–8.

Cathcart RJ & Swanton CJ (2003) Nitrogen management will influence threshold values of green foxtail (*Setaria viridis*) in corn. *Weed Science*, **51**, 975–986.

Coble HD & Mortensen DA (1992) The threshold concept and its application to weed science. *Weed Technology*, **6**, 191–195.

Cousens R (1985) A simple model relating yield loss to weed density. *Annual Applied Biology*, **107**, 239–252.

Ferreira EA, Matos CC, Barbosa EA, Melo CAD, Silva DV & Santos JB (2015) Physiology aspects of transgenic soybean submitted to competition with weed. *Amazonian Journal of Agricultural and Environmental Sciences*, **58**, 115–121.

Fischer DW, Harvey RG, Bauman TT *et al*. (2004) Common lambsquarters (*Chenopodium album*) interference with corn across the northcentral United States. *Weed Science*, **52**, 1034–1038.

Fox J & Weisber S (2011) *An R Companion to Applied Regression*. Sage, Thousand Oaks, CA, USA.

Hock SM, Knezevic SZ, Martin AR & Lindquist JL (2006) Soybean row spacing and weed emergence time influence weed competitiveness and competitive indices. *Weed Science*, **54**, 36–46.

Hurvich CM & Tsai C (1991) Bias of the corrected AIC criterion for underfitted regression and time series models. *Biometrika*, **78**, 499–509.

Knezevic SZ & Datta A (2015) The critical period for weed control: revisiting data analysis. *Weed Science*, **63**, 188–202.

Knezevic SZ & Horak MJ (1998) Influence of emergence time and density on redroot pigweed (*Amaranthus retroflexus*). *Weed Science*, **46**, 665–672.

Knezevic SZ, Horak MJ & Vanderlip RL (1997) Relative time of redroot pigweed (*Amaranthus retroflexus* L.) emergence is critical in pigweed-sorghum [*Sorghum bicolor* (L.) Moench] competition. *Weed Science*, **45**, 502–508.

Knezevic SZ, Evans SP, Blankenship EE, Van Acker RC & Lindquist JL (2002) Critical period for weed control: the concept and data analysis. *Weed Science*, **50**, 773–786.

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

Knezevic SZ, Jhala A & Datta A (2017) Integrated weed management. In: *Encyclopedia of Applied Plant Sciences*, 2nd edn (eds B Thomas, BG Murray & DJ Murphy), 459–462. Academic Press, Waltham, MA.

Lewis DW & Gulden RH (2014) Effect of kochia (*Kochia scoparia*) interference on sunflower (*Helianthus annuus*) yield. *Weed Science*, **62**, 158–165.

Lewis F, Butler A & Gilbert L (2011) A unified approach to model selection using the likelihood ratio test. *Methods in Ecology and Evolution*, **2**, 155–162.

Lindquist JL & Mortensen DA (1998) Tolerance and velvetleaf (*Abutilon theophrasti*) suppressive ability of two old and two modern corn (*Zea mays*) hybrids. *Weed Science*, **46**, 569–574.

Lindquist JL, Mortensen DA, Clay SA *et al.* (1996) Stability of corn (*Zea mays*)-velvetleaf (*Abutilon theophrasti*) interference relationships. *Weed Science*, **44**, 309–313.

Lindquist JL, Mortensen DA, Westra P *et al.* (1999) Stability of corn (*Zea mays*)-foxtail (*Setaria* spp.) interference relationships. *Weed Science*, **47**, 195–200.

Massinga RA, Currie RS, Horak MJ & Boyer J (2001) Interference of Palmer amaranth in corn. *Weed Science*, **49**, 202–208.

Mayer DG & Butler DG (1993) Statistical validation. *Ecological Modelling*, **68**, 21–32.

Monquero PA, Orzari I, Silva PV & Penha ADS (2015) Interference of weeds on seedlings of four neotropical tree species. *Acta Scientiarum Agronomy*, **37**, 219–232.

Oliveira MC, Jhala AJ, Gaines T *et al.* (2017) Confirmation and control of HPPD-inhibiting herbicide–resistant waterhemp (*Amaranthus tuberculatus*) in Nebraska. *Weed Technology*, **31**, 67–79.

Onofri A, Carbonell EA, Piepho H-P, Mortimer AM & Cousens RD (2010) Current statistical issues in Weed Research. *Weed Research* **50**, 5–24.

R Core Team (2018) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. Available at: https://www.R-project.org/ (last accessed 15 January 2018).

Ritz C & Streibig JC (2005) Bioassay analysis using R. *Journal of Statistical Software*, **12**, 1–18.

Ritz C & Streibig JC (2008) *Nonlinear Regression with R*. Springer, New York, NY, USA.

Ritz C, Baty F, Streibig JC, Gerhard D, Baun A & Nyholm N (2015) Dose-response analysis using R. *PLoS ONE*, **10**, e0146021.

Roman ES, Murphy SD & Swanton CJ (2000) Simulation of *Chenopodium album* seedling emergence. *Weed Science*, **48**, 217–224.

Ryan MR, Smith RG, Mortensen DA *et al.* (2009) Weed – crop competition relationships differ between organic and conventional cropping systems. *Weed Research*, **49**, 572–580.

Silva DV, Pereira GAM, Freitas MAM *et al.* (2015) Nutritional efficiency of maize in intercropping with signalgrass. *Ciência Rural*, **45**, 1394–1400.

Strieder ML, Silva PRF, Argenta G *et al.* (2007) The response of irrigated corn to row spacing depends on hybrid and plant density. *Ciência Rural*, **37**, 634–642.

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

Swanton CJ, Nkoa R & Blackshaw RE (2015) Experimental methods for crop–weed competition studies. *Weed Science*, **63**, 2–11.

Trezzi MM, Vidal RA, Patel F *et al.* (2015) Impact of *Conyza bonariensis* density and establishment period on soyabean grain yield, yield components and economic threshold. *Weed Research*, **55**, 34–41.

Voll E, Gazziero DLP, Brighenti AAM & Adegas FS (2002) Relative weed species competition with two soybean cultivars. *Planta Daninha*, **20**, 17–24.

Weiner J & Freckleton RP (2010) Constant final yield. *Annual Review of Ecology Evolution, and Systematics*, **41**, 173–192.

Werle R, Bernards ML, Arkebauer TJ & Lindquist JL (2014a) Environmental triggers of winter annual weed emergence in the midwestern United States. *Weed Science*, **62**, 83–96.

Werle R, Sandell LD, Buhler DD, Hartzler RG & Lindquist JL (2014b) Predicting emergence of 23 summer annual weed species. *Weed Science*, **62**, 267–279.

Werle R, Schmidt JJ, Laborde J, Tran A, Creech CF & Lindquist JL (2014c) Shattercane X ALS-tolerant sorghum F1 hybrid and shattercane interference in ALS-tolerant sorghum. *Journal of Agricultural Science*, **6**, 159–165.

Wilkerson GG, Wiles LJ & Bennett AC (2002) Weed management decision models: pitfalls, perceptions, and possibilities of the economic threshold approach. *Weed Science*, **50**, 411–524.

Zucchini W (2000) An introduction to model selection. *Journal of Mathematical Psychology*, **44**, 41–61.

Zuur AF, Ieno EN & Smith GM (2007) *Analyzing Ecological Data*. Springer, New York, NY, USA.

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Appendix S1 Using R to analyse additive designs