# The Critical Period for Weed Control:
# Revisiting Data Analysis

Stevan Z. Knezevic and Avishek Datta*

**Key words:**   Critical time for weed removal, critical weed-free period, crop–weed interactions, dose–response curves, duration of interference, experimental design, nonlinear regression, R software.

There is an ever-larger need for designing an integrated weed management (IWM) program largely because of the increase in glyphosate-resistant weeds, not only in the United States but also worldwide. An IWM program involves a combination of various methods (cultural, mechanical, biological, genetic, and chemical) for effective and economical weed control (Swanton and Weise 1991). One of the first steps in designing an IWM program is to identify the *critical period for weed control* (CPWC), defined as a period in the crop growth cycle during which weeds must be controlled to prevent crop yield losses (Zimdahl 1988). Knowing the CPWC is useful for making decisions on the need for, and timing of, weed control, depending on the specific crop (Knezevic et al. 2002). CPWC studies have been reported in a variety of crops worldwide, including corn (*Zea mays* L.) (Evans et al. 2003; Hall et al. 1992), soybean [*Glycine max* (L.) Merr.] (Knezevic et al. 2003; Van Acker et al. 1993), sunflower (*Helianthus annuus* L.) (Knezevic et al. 2013), rice (*Oryza sativa* L.) (Chauhan and Johnson 2011), cotton (*Gossypium* L. spp.) (Bukun 2004), canola (*Brassica napus* L.) (Martin et al. 2001), peanut (*Arachis hypogaea* L.) (Everman et al. 2008), carrot (*Daucus carota* L.) (Swanton et al. 2010), white bean (*Phaseolus vulgaris* L.) (Woolley et al. 1993), tomato (*Solanum lycopersicum* L.) (Weaver and Tan 1983), leek (*Allium porrum* L.) (Tursun et al. 2007), red pepper (*Capsicum annum* L.) (Tursun et al. 2012), lentil (*Lens culinaris* Medik.) (Smitchger et al. 2012), and chickpea (*Cicer arietinum* L.) (Mohammadi et al. 2005).

Several types of data analyses to determine the CPWC have been reported in the literature, including multiple-comparison techniques (Kalaher et al. 2000) and nonlinear regression models (Evans et al. 2003; Van Acker et al. 1993). A nonlinear regression procedure was suggested as a reasonable method for determining the CPWC using SAS software (Knezevic et al. 2002). SAS is powerful software for statistical analysis (SAS Institute 1999); however, because of its licensing requirements, it has not been readily available to the worldwide scientific community. Thus, there has been increased interest in statistical packages that are readily available on the Internet, such as R software (R Development Core Team 2013), which is gaining popularity worldwide. R is open-source, command-line–driven statistical software (similar to SAS) and is free (Knezevic et al. 2007). R can conduct many types of statistical analyses, including various regressions. The user only needs to fit the regression model once and then all parameter combinations of choice can be tested for significance. R also contains sets of prewritten codes (called packages) that are designed to conduct specific types of analysis. One such package is *drc* (dose–response curves) (Ritz and Streibig 2005). The package drc is an add-on package for the language and environment R and contains programmed commands for regression analysis and enables R to graph the distribution of data and regression lines. Therefore, the objectives of this article are to briefly revisit the concept of, and studies about, the CPWC and to outline a common method for CPWC data analysis based on the sets of codes from the drc package and R software. Adoption of this method of data analysis would allow easier comparison of the results among sites and among researchers.

## The CPWC: Refresh the Concept

An extensive review of the concept of the CPWC has been provided previously (Knezevic et al. 2002),
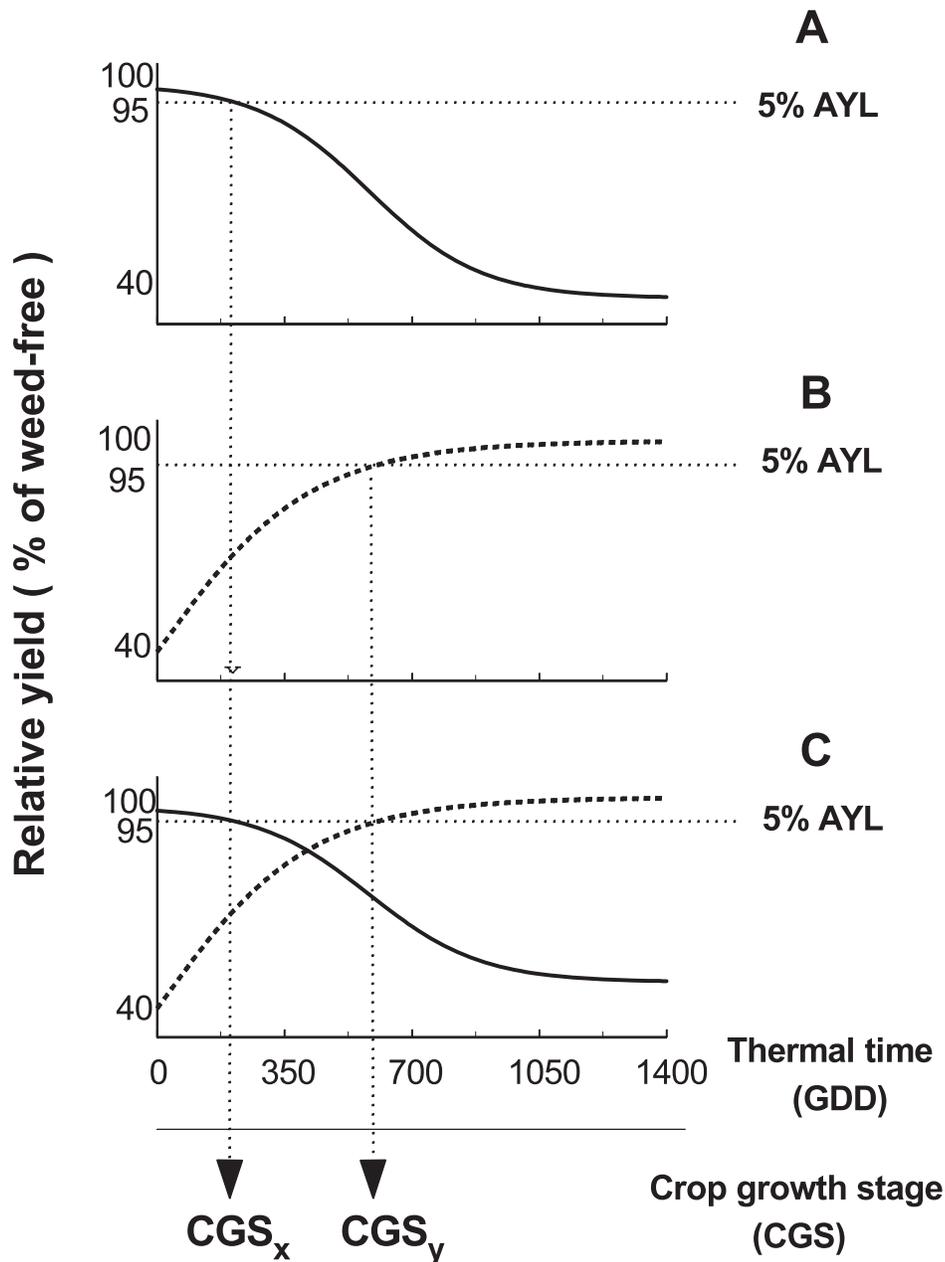
Figure 1. Functional approach used for determination of the critical period for weed control (CPWC). (A) The critical time for weed removal (CTWR) is determined from the so-called weedy curve (—; descending line), fit to data representing an increasing duration of weed interference. (B) The critical weed-free period (CWFP) is determined from the weed-free curve ($-$ $-$ $-$; ascending line), fit to data representing the increasing duration of the weed-free period. (C) The value of the x-axis that corresponds to the 95% relative yield or an acceptable yield loss (AYL) of 5% is determined for both curves and related to crop growth stage (CGS). The CPWC is then defined as the time between the two crop growth stages ($CGS_x$ to $CGS_y$) and represents the length of weed control required to protect the crop yield from more than a 5% yield loss. (Source: Knezevic et al. 2002).

including a description of the types of data that need to be collected. For data collection, the CPWC represents the time interval between two separately measured crop–weed competition components (Figure 1). The first component is known as the critical time for weed removal (CTWR), based on the so-called weedy curve. The CTWR is the maximum amount of time that early season weed competition can be tolerated by the crop before it suffers from an irreversible yield reduction (Figure 1A). The CTWR is estimated to determine the "beginning" of the CPWC. The second component is known as the critical weed-free period (CWFP), based on the so-called weed-free curve. The CWFP is the minimum weed-free period required from the time of crop planting to prevent unacceptable yield

Knezevic and Datta: Critical period for weed control • 189

losses (Figure 1B). This component determines the "end" of the CPWC. Results from both components are required and are typically combined to determine the CPWC, regardless of crop species (Figure 1C) (Knezevic et al. 2002). Theoretically, weed control before and after the CPWC does not significantly contribute to the conservation of crop-yield potential.

The beginning and end of the CPWC determined using the regression approach to generate both the weedy and weed-free curves depend on the level of acceptable yield loss (AYL) used to predict its beginning and end (Figure 1). Many studies determined the beginning and end of the CPWC based on the maximum AYL level between 2.5 and 10% (Knezevic et al. 2002). The AYL can be adjusted, however, depending on various factors of interest (e.g., the cost of weed control and the anticipated financial gain). For example, the AYL for a crop can be adjusted for a particular field, depending on the market price of the crop and the cost of weed control. This is easily achieved by selecting the AYL from the regression curves according to the extent of financial risk that is acceptable in each situation.

**Experimental Approach to Determine the CPWC.** Knezevic et al. (2002) suggested that the CPWC (timing of weed control) should be based on the crop-growth stage, at least from a practical standpoint. After all, weeds are controlled because we try to protect the crop yield. Therefore, the crop should be the focus of the program. The CPWC timing ultimately depends on whether weeds emerge before or after the crop and can be adjusted accordingly based on differences between the relative emergence time of the weeds (Evans et al. 2003). Crop producers and agricultural professionals can easily follow the timing of weed control for a particular crop when the CPWC is reported and is available based on crop-growth stage with corresponding days after crop emergence. The CPWC should be related to crop-growth stage to account for environmental variation, as suggested by others (Hall et al. 1992; Martin et al. 2001; Van Acker et al. 1993). This approach suggests making a weed-control decision based on biological parameters, rather than on the overreliance of an inexpensive herbicide, which promotes herbicide-resistant weeds, especially in cropping systems using herbicide-tolerant crops (e.g., Roundup Ready [Monsanto]).

For data collection, crop yield at physiological maturity is the most important variable collected for CPWC-type studies. From the yield data, the relative yield can be calculated as a percentage of the corresponding weed-free yields. Other data should include the emergence dates of crop and weed species, weed species composition, weed density, and weekly staging of crop and weeds (Knezevic et al. 2002). In addition, weed biomass and canopy height (crop and weed) measured at regular intervals, as well as environmental and soil variables, such as daily rainfall, average daily temperature (air and soil), soil moisture, and nutrient status should also be measured. These ancillary variables provide background information that quantifies the competitive environment under which the experiment was conducted. It is important to acquire these additional data because they help extrapolate the results to other conditions (Knezevic et al. 2002).

**Number of Required Data Points.** The regression model we have suggested in this article is based on four regression coefficients; therefore, a minimum of five data points are required for fitting both the weedy and weed-free curves relating crop yield to time of weed removal. In an ideal situation, six to seven data points are perfect. The additional data points help test differences in curves (weedy and weed-free) among years and sites.

It is essential to select appropriate times for weed removal to determine the shape of the curves. The following suggestions can be used as general guidelines for selecting the appropriate times for weed removal based on our experience and the summaries of previous studies (Elezovic et al. 2012; Zimdahl 1980): (1) select two to three data points clustered around the early part of the crop-growing season (the first few weeks of the season) to best determine the suspected beginning of the CPWC, and (2) select two to three data points clustered around the later stages of crop growth (time of crop canopy closure) to best determine the end of the CPWC. One data point can be in the general area where the two curves are assumed to cross each other (Figure 1).

## General Approach for the Suggested Statistical Analysis

Treatment comparisons should be based on regression analysis where the time of weed removal (treatments, i.e., the x-axis) should be related to the

crop's actual or relative yield (the y-axis) using regression analysis. The use of growing degree days (GDD) accumulated from crop emergence or planting as the unit of time to quantify the duration of weed presence and the length of weed-free period (x-axis) is recommended (Elezovic et al. 2012; Knezevic et al. 2013). The use of GDD can also help determine the CPWC based on the respective crop-growth stage and the corresponding days after crop emergence for practical purposes because the rate of crop development is well correlated with thermal time. Overall, the use of GDD in the regression analysis is preferred because it (1) is a more biologically meaningful measure of time needed for plant growth and development, compared with some other indicators, such as days or weeks after crop emergence (Gilmore and Rogers 1958); (2) can be a valuable means for comparing data from different sites, years, and planting dates; and (3) provides a continuous and precise scale for the x-axis. For example, an x-axis based solely on the crop-growth stage is a continuous scale only during vegetative growth stages, such as leaf stages from 1 to 15. However, the scale becomes "uncertain" when attempting to define data points during reproductive growth, such as the beginning or midpoint of grain filling. It is also possible to compare the CPWC among crops using GDD. This allows for the investigation of noncrop-related reasons for the timing of the CPWC. In some cases, the end of the CPWC (the critical weed-free period) is strongly related to the emergence periodicity of the weed complex (Martin et al. 2001; Van Acker et al. 1993), which may be independent of crop species and more a function of the time of crop seeding.

The four-parameter log-logistic model is suggested for conducting nonlinear regression analyses while handling the CPWC data. This model is used in the data analysis steps outlined in this article in the subsequent section. Log-logistic equations are well suited for describing both the increasing duration of weed interference on relative yield (Figure 1A, weedy curve) and the increasing length of the weed-free period (Figure 1B, weed-free curve). The log-logistic function is given by Equation 1 (Knezevic et al. 2007):

$$Y = [C + (D - C)] / \{1 + \exp[B(\log X - \log E)]\} \quad [1]$$

where $Y$ can be actual yield (or the relative yield [percentage of season-long weed-free yield]), $C$ is the lower limit, $D$ is the upper limit, $X$ is time (the x-axis expressed in GDD or days after emergence

[DAE]), $E$ is the GDD giving a 50% response between the upper and lower limits (also known as the *inflection point* [$I_{50}$] or $ED_{50}$), and the parameter $B$ is the slope of the line at the inflection point. Interpretation of these parameters is discussed in more details in a previous article (Knezevic et al. 2007).

A few other models can also be used, such as a three-parameter log-logistic model and three- or four-parameter Weibull models (Knezevic et al. 2007). From a statistical standpoint, the model that is chosen for this article (Equation 1) should provide the best fit to the CPWC data. Although this article focuses on an analysis with the four-parameter log-logistic model, we do offer a suggestion for determining the goodness-of-fit in the following sections.

Most studies reported the use of an arbitrary AYL value of 5% (but it could be from 2.5 to 10%) to signify the beginning and end of the CPWC (Knezevic et al. 2013). That range allows a producer or crop consultant to adjust the CPWC to the extent of the risk acceptable in each situation. Such decisions should consider the economics of weed control, including the cost of herbicide and its application, crop price, and anticipated gain from control. An AYL value of 5% was used here as a threshold to determine the beginning and the end of the critical period, as presented in Figure 1C.

**Installation of R Software and the drc Package.** Details about the R environment are available at the uniform resource locator (URL) http://cran.at. r-project.org (R Development Core Team 2013) and in Knezevic et al. (2007). There are two free editors for use with R, including the Tinn-R and RStudio. The Tinn-R is a simple, but efficient, basic code editor that can be downloaded from URL http://sourceforge.net/projects/tinn-r/. We, however, prefer a relatively new and free editor called RStudio (http://www.rstudio.com). This editor has multiple purposes: (1) it provides a multiple-window interface for easy viewing of various functions (e.g., program codes, analysis outputs, graph); (2) it edits program codes; (3) it provides outputs from the analysis; (4) it provides graphs and figures resulting from the analysis, which can be then copied and pasted into word-processing programs for publications (e.g., Word [Microsoft]) or presentation purposes (e.g., Power-Point [Microsoft]); (5) it has a help function, which provides quick access to various manuals and

explanations of the thousands of codes available in R; and (6) it provides a list of hundreds of "precoded" packages that can be installed to conduct specific types of analyses. One such package is called drc (an add-on package) (Ritz and Streibig 2005), which is needed to analyze the regressions presented in this article. The R, RStudio, and drc, as a package, should not occupy more than 100 MB of hard drive space.

To download R, just visit the URL http://cran.us.r-project.org/ and install the executable file. To install RStudio simply visit the URL http://www.rstudio.com and download the executable file. After installing both R and RStudio, the add-on package drc must be also installed. Simply go to the lower left window of RStudio and click on "packages," select "drc," and RStudio will install it. Three additional packages are also useful to install: plotrix (plotting graphs), RODBC (to access Excel [Microsoft] files), and sciplot (plotting standard error bars). These packages will cover the basics for running the set of codes provided in this article, but there are many other add-on packages within the RStudio. At this point, the R and RStudio are ready to take the data and perform analyses.

**Data Organization and Input into R.** Data may be transferred into R from a variety of spreadsheet packages. Smaller data sets may simply be copied and pasted into R (Ritz and Streibig 2005). Medium-size files (data in a single Excel page) may be imported as comma-separated files (filename.csv). Larger files (data sets with multiple pages in a single Excel file) can be accessed via a set of codes that will allow R to read the data from the Excel spreadsheets. In general, we recommend the two most-common methods of reading data into R. First, large data sets can be read using a set of codes (Table 1, line 0.21) that provide a path from the hard drive where the Excel file is stored, which allows R to access the file. This is our preferred method for reading data into R because most scientists today generate large data sets with multiple pages of data in a single Excel spreadsheet format. The other method, for a single page data, it so access the data from a comma-separated file (.csv), depending on the character used for decimal points. Files in the continental European countries, where a comma (,) is used as a decimal point and a semicolon is used as a separator, have the filename extension *.csv2*. Files in the Anglo Saxon system, where a dot (.) is used as the decimal point and the comma as a separator, have the filename extension *.csv*.

**General Approach for the Nonlinear Regression Analysis.** Using the drc package, there is no need for initial estimates of the regression parameters because they are already provided for the common models that have potential for use in the weed science discipline and are built into the drc package (Ritz and Streibig 2005). After fitting several curves, it may be of interest to compare parameters across curves, for instance, by comparing lower limits for different curves. Treatment comparisons should be based on regression curve. The log-logistic model is the default model, presented with the code *l4* (see further text in Table 1). Other models can also be specified by appropriate commands. A summary of the parameter estimates and a test of the goodness-of-fit can be also conducted (Table 1, lines 05 and 06). The *ED* (effective dose) command calculates ED values, based on the fitted model, whereas the *SI* (selectivity index) command is commonly used to compare the relative differences in ED values among curves (Table 1, lines 07 and 08). Both the *ED* and *SI* commands are functions of the equation parameters, and their interpretation is provided in the following sections, with an example.

**Case Study Example.** This experiment was arranged as a randomized complete-block design with the timing of weed removal and duration. There were five treatments of increasing weed duration used to determine the CTWR, and five treatments of increasing length of weed-free periods used to determine the CWFP. Studies were conducted at four sites (hereafter, referred as sites 1, 2, 3, and 4) in Nebraska and four replications were maintained at each site. The treatments of increasing duration of weed interference and lengths of weed-free periods up to predetermined corn-growth stages (V3, V6, V9, V15, and R1) were compared with the season-long weed-free and season-long weedy controls established at each location. Specific details regarding the experimental site and procedures are not provided because the focus of this article is not to discuss the biological interpretation of the results.

The influence of an increasing duration of weed interference and an increase in the length of weed-free period (expressed in GDD) on corn grain yields (kg ha$^{-1}$) is graphically represented in Figure 2. The four-parameter log-logistic model fit to data representing an increasing duration of weed interference results in weedy curves and an increase in the length of weed-free period results in weed-free curves (2B) at four sites. The regression lines are plotted using equation 1. In this paper, the

Table 1. Basic commands for fitting the log-logistic curve (Equation 1) using the drc add-on package in R Statistical Software to determine (2A) the beginning (the weedy curve), and (2B) the end (the weed-free curve) of the critical period of weed control (CPWC) based on acceptable yield loss (AYL) levels of 2.5, 5, or 10%.

| Line | R program & output | Comments |
|---|---|---|
| **Step 1: Reading a data file into R** | | |
| 01 | library(drc) | This function loads and activates the drc package in R |
| 02 | dataname<-read.csv ("C:\\dir\\subdir\\ filename.csv"), skip=4) | We suggest this command for medium-size data sets (single page of data in Excel file). Assign a name to your data file (replace "dataname" with a name of your choice) and location (replace "c:\\dir\\subdir\\filename.csv" with the full path of your data file). In the Anglo Saxon system, with (.) as decimal separator you should use read.csv () whereas where a comma (,) is used as decimal point, use read.csv2(). To start reading data from a particular line in the data file, use skip=4 code, which tells R to start reading data at the indicated line number (4). |
| 02.1 | dataname<-sqlFetch(odbcConnectExcel ("C:\\dir\\subdir\\filename.xls"), sqtable = "name of worksheet page in Excel file", na.strings = "NA", as.is = TRUE) | Large data sets can be read using these sets of codes that provide a path from the hard drive where Excel file is stored, which allows R to access the file. The command "sqtable" is used to read the particular data set from multiple pages of data in the single Excel spreadsheet. Replace "name of worksheet in Excel file" according to the name of the page in the Excel file you are using. |
| 03 | head(dataname) | This line prints the first 6 lines of the data set; thus, it is useful for double checking whether data have been read correctly. An example of first 6 lines is provided. |

| | Site | Trt | Rep | GDD | WeedyYield |
|---|---|---|---|---|---|
| 1 | 1 | 7 | 1 | 0 | 6822.3 |
| 2 | 1 | 7 | 2 | 0 | 6200.6 |
| 3 | 1 | 7 | 3 | 0 | 6657.0 |
| 4 | 1 | 7 | 4 | 0 | 6078.9 |
| 5 | 1 | 7 | 1 | 185 | 6010.5 |
| 6 | 1 | 7 | 2 | 185 | 6095.3 |

**Step 2: Fitting multiple curves (drc)**
**The drm function is the key function in drc for fitting dose–response curves**

| Line | R program & output | Comments |
|---|---|---|
| 04 | CPWCWeedy<-drm(WeedyYield~GDD, Site, fct= l4(fixed=c(NA,NA, NA,NA)), data= dataname) | CPWCWeedy is a user-defined name that will contain all information about the fitted model generated by the drm function using data for describing weedy curve (descending line). WeedyYield is the response variable (y-axis) and GDD is the growing-degree days (x-axis). Site is the classification variable for each curve, which is basically telling R to analyze all four sites simultaneously so that we can compare ED values from all four curves using the SI function (see line 08). fct=l4() is the argument for the log-logistic curve with four parameters (Equation 1). If the fct argument is not provided, by default, a four-parameter log-logistic model is fitted to the data; data=dataname identifies the name of data file. Note that drm does not produce an output. All information on the model fit is stored in the object under the user-defined name (e.g., CPWCWeedy) to be read by further commands (e.g., lines 5–8). |
| 05 | summary(CPWCWeedy) | The summary function provides a summary of the equation parameter estimates. |
| 06 | modelFit(CPWCWeedy) | The modelFit function provides a lack-of-fit test, comparing your chosen dose–response model to the more-general ANOVA model. In this particular instance, it is a two-way ANOVA with Site and GDD as factors. |

Knezevic and Datta: Critical period for weed control • 193

Table 1. Continued.

| Line | R program & output | Comments |
|---|---|---|
| 07 | ED(CPWCWeedy,c(5)) | The *ED(effective dose)* command provides ED values of your choice based on the fitted model, e.g., the *c(5)* command will provide GDD values for the 5% AYL level to be determined from the "weedy curve" at four sites. Similarly, enter the *c(2.5)* or *c(10)* to calculate ED values for 2.5 and 10% AYL levels, respectively. |
| 08 | SI(CPWCWeedy,c(5,5)) | *SI(selectivity index)* is used to compare the relative differences of ED values (at 5%) among curves, which are the GDD values among the sites. In this case, we have compared the GDD values among four sites for the 5% AYL level. |
| 09 | plot(CPWCWeedy, log= "", xlim=c(0,1500),ylim= c(2000,7000),xlab= "Growing degree days (GDD)", ylab = expression(paste("Yield (kg ", ha-1, ")")), col = c(1,2,3,4), lwd=c(2,2,2,2), pcb=c(1,2,3,4), legendText=c("Site-1", "Site-2", "Site-3", "Site-4"), legendPos=c(600,2800), main="2A: Weedy curves at 4 sites") | This line is used to generate weedy curves for determining the beginning of the CPWC at four sites (Figure 2A; Table 2). To obtain visuals of the fitted curve, the *plot* command is used. The argument *log* specifies the x-axis without a logarithmic scale, whereas *xlim* and *ylim* specify the range on the x-axis and y-axis, respectively. The labels on the axes are provided using arguments *xlab* and *ylab*. The arguments *col*, *lwd*, and *pcb* determine the color, line width, and plot symbol used for each curve and the corresponding data points. Legend text, legend position, and figure title are provided by arguments *legendText*, *legendPos*, and *main* (Figure 2). Visit http://www.R-project.org for additional information on how to produce graphs within R environment. |
| 10 | CPWCWeedfree1 <-drm(WeedfreeYield~GDD, Site, fct=l4(fixed=c(NA,NA,NA,NA)), data=dataname) | *CPWCWeedfree1* is a user-defined name that will contain all information about the fitted model generated by the drm function using data for fitting the weed-free curve (ascending line).<br>*WeedfreeYield* is the response variable (y-axis) and *GDD* is the growing degree days (x-axis).<br>*Site* is the classification variable for each curve, which is basically telling R to analyze all four sites simultaneously so that we can compare ED values from all four curves using the *SI* function (see line 14).<br>In essence, this is similar to line 04, except this time we have used data for weed-free curve (ascending line). |
| 11 | summary(CPWCWeedfree1) | The *summary* function provides a summary of the parameter estimates. |
| 12 | modelFit (CPWCWeedfree1) | The *modelFit* function provides a lack-of-fit test, comparing your chosen dose–response model to the more-general ANOVA model. |
| 13 | ED(CPWCWeedfree1,c(95)) | The *ED* command provides ED values of your choice based on the fitted model, e.g., the *c(95)* command will provide GDD values for the 5% AYL levels to be determined from the "weedy-free curves" at four sites.<br>Note that the *c(95)* command is equivalent to 5% AYL because we are calculating ED values from the ascending line, which means $100 - 5 = 95$; thus, we use 95 in the code.<br>Similarly, enter the *c(97.5)* or *c(90)* commands to calculate ED values for 2.5 ($100 - 2.5 = 97.5$) and 10% ($100 - 10 = 90$) AYL levels, respectively. |
| 14 | SI(CPWCWeedfree1,c(95,95)) | *SI* is used to compare the relative differences of ED values among curves at the 5% AYL. |

Table 1. Continued.

| Line | R program & output | Comments |
|---|---|---|
| 15 | *plot*(CPWCWeedfree1,log= "", xlim= c(0,1500), ylim=c(2000,7000), xlab= "Growing degree days (GDD)", ylab = expression(paste("Yield (kg " , ha-1, ")")), col = c(1,2,3,4), lwd= c(2,2,2,2), pcb=c(1,2,3,4), legendText=c("Site-1", "Site-2", "Site-3", "Site-4"), legendPos= c(1000,2800), main= "2B: Weed-free curves at 4 sites") | This line is used to generate weed-free curves for determining the end of the CPWC at four sites (Figure 2B, Table 2). To obtain visuals of the fitted curve, the *plot* command is used. |

***Output from Step 2***

| Line | R program & output | Comments |
|---|---|---|
| 16 | summary (CPWCWeedy)<br>Model fitted: Log-logistic (ED50 as parameter) (4 parms)<br>Parameter estimates: | Output from *summary* command (weedy curve): parameter estimates in the four-parameter log-logistic model. The estimated *b* parameters: $b{:}1$, $b{:}2$, $b{:}3$, and $b{:}4$ of the four curves are the relative slope at $e{:}1$, $e{:}2$, $e{:}3$, and $e{:}4$, which are the $ED_{50}$ for the four curves; $c{:}1$, $c{:}2$, $c{:}3$, and $c{:}4$ are the lower limits, whereas $d{:}1$, $d{:}2$, $d{:}3$, and $d{:}4$ are the upper limits.<br>Also, the SE is provided for the parameters as is an approximate *t* test with associated P value, which is testing the hypothesis that the parameters are equal to zero.<br>Note: The four curves (Figure 2a) are significantly different from each other because at least one of their parameters is different based on their ± SE. Therefore, we have presented all four curves separately on the graph (Figure 2a).<br>For future reference, if all four parameters from each of the curves are not significantly different from each other (based on their ± SE), then the curves can be combined (e.g., data for the two or more curves can be combined and a single curve fit). |

|  | Estimate | Std. Error | t-value | P value |
|---|---|---|---|---|
| b:1 | 1.77211 | 0.72063 | 2.45912 | 0.0157 |
| b:2 | 1.98285 | 0.84090 | 2.35800 | 0.0204 |
| b:3 | 1.55263 | 1.25048 | 1.24163 | 0.2174 |
| b:4 | 2.24051 | 0.71078 | 3.15219 | 0.0022 |
| c:1 | 1355.67990 | 1503.54120 | 0.90166 | 0.3695 |
| c:2 | 2911.65631 | 784.83917 | 3.70988 | 0.0003 |
| c:3 | 3351.07567 | 2230.92430 | 1.50210 | 0.1364 |
| c:4 | 1822.08596 | 458.85616 | 3.70988 | 0.0001 |
| d:1 | 6431.19840 | 316.46918 | 20.32172 | 7.598e−37 |
| d:2 | 6905.81900 | 323.69930 | 21.33406 | 1.651e−38 |
| d:3 | 6573.18519 | 342.22185 | 19.20738 | 5.921e−35 |
| d:4 | 6150.73275 | 342.45548 | 17.96068 | 9.274e−33 |
| e:1 | 626.44313 | 241.71976 | 2.59161 | 0.0110 |
| e:2 | 446.57052 | 113.66791 | 3.92873 | 0.0002 |
| e:3 | 731.26142 | 698.44130 | 1.04699 | 0.2977 |
| e:4 | 397.74338 | 60.60495 | 6.56289 | 2.682e−09 |

Residual standard error:
654.0004 (96 degrees of freedom)

| Line | R program & output | Comments |
|---|---|---|
| 17 | modelFit (CPWCWeedy)<br>Lack-of-fit test | Lack-of-fit test yields a P value of 0.20, which is not significant at the 5% level of significance, indicating that there was no lack-of-fit for this nonlinear model, thus suggesting that the model provided an acceptable description of data. |

| Model | Df | RSS | Df | F value | P value |
|---|---|---|---|---|---|
| ANOVA | 84 | 31336984 |  |  |  |
| DRC model | 96 | 41060783 | 12 | 0.2172 | 0.203 |

Table 1. Continued.

| Line | R program & output | Comments |
|---|---|---|
| 18 | ED(CPWCWeedy,c(5)) <br> Estimated effective doses <br> Estimate Std. Error <br><br>      Estimate  Std.Error <br> 1 : 5  117.9269  51.0726 <br> 2 : 5  118.1537  52.3449 <br> 3 : 5  110.7665  93.4657 <br> 4 : 5  106.8712  44.6585 | <u>Output from *ED*</u>: The $ED_5$ values (5% AYL for the weedy curve) for each of the four curves and the corresponding, estimated SE. The values are indicating the beginning of CPWC in GDD. |
| 19 | SI(CPWCWeedy,c(5,5)) <br> Estimated ratios of effect doses <br><br>         Estimate  Std. Error  t–value  P value <br> 1/2:5/5  1.175705  0.790619  0.222237  0.8246 <br> 1/3:5/5  1.083454  1.033247  0.080769  0.9358 <br> 1/4:5/5  0.112806  0.666792  0.169177  0.8660 <br> 2/3:5/5  0.921536  0.918226  −0.085452  0.9321 <br> 2/4:5/5  0.946501  0.629548  −0.084980  0.9325 <br> 3/4:5/5  1.027091  0.974201  0.027808  0.9779 | <u>Output from *SI*</u>: The $ED_5$ values are compared. $ED_5$ values for curves 1 and 2 by analyzing whether the ratio $ED_5(1):ED_5(2)$ is significantly different from 1.00 (P = 0.82); $ED_5$ values for curves 1 and 3 are not significantly different (P = 0.94); $ED_5$ values for curves 1 and 4 are not significantly different (P = 0.87). <br><br> Note: Even though the curves are significantly different from each other (see line 16), the beginnings of the critical periods among four sites are statistically similar (based on P values). Thus, we can safely assume that the beginning of the critical period (weed-free period) lies around the V3–V4 stage based on the 5% AYL level (Table 2). |
| 20 | summary(CPWCWeedfree1) <br> Model fitted: Log-logistic (ED50 as parameter) (4 parms) <br> Parameter estimates: <br><br>     Estimate  Std. Error  t–value  P value <br> b:1  −3.6842  1.7028  −2.1636  0.0330 <br> b:2  −3.4391  1.3254  −2.5947  0.0110 <br> b:3  −5.8416  5.3910  −1.0836  0.2813 <br> b:4  −5.5930  1.9992  −2.7976  0.0062 <br> c:1  3181.7251  260.7719  12.2012  1.564e−21 <br> c:2  3894.6127  294.4342  13.2274  1.176e−23 <br> c:3  4371.7000  237.0179  18.4446  1.274e−33 <br> c:4  2187.9549  222.3082  9.8420  2.735e−16 <br> d:1  6337.5678  280.0041  22.6338  1.435e−40 <br> d:2  6914.0270  215.1356  32.1380  1.778e−53 <br> d:3  6445.9537  201.6157  31.9715  2.807e−53 <br><br> Residual standard error: <br> 563.9039 (96 degrees of freedom) | <u>Output from *summary* command (weed-free curve)</u>: Parameter estimates in the four-parameter log-logistic model as discussed previously for the weedy curve (line 16). <br><br> Note: The four curves (Figure 2B) are significantly different from each other because at least one of their parameters is different based on their ± SE. Therefore, we have presented all four curves separately on the graph (Figure 2B). |

Table 1. Continued.

| Line | R program & output | Comments |
|---|---|---|
| 21 | modelFit (CPWCWeedfree1)<br>Lack-of-fit test<br><br>| Model | Df | RSS | Df | Fvalue | Pvalue |<br>|---|---|---|---|---|---|<br>| ANOVA | 84 | 29140427 | | | |<br>| DRCmodel | 96 | 30526813 | 12 | 0.3330 | 0.9809 | | Lack-of-fit test yields a P value of 0.98, which is not significant at the 5% level of significance, indicating that the nonlinear model provides an acceptable description of data. |
| 22 | ED(CPWCWeedfree1,c(95))<br>Estimated effective doses<br><br>| | Estimate | Std.Error |<br>|---|---|---|<br>| 1 : 95 | 783.61 | 203.84 |<br>| 2 : 95 | 692.61 | 212.20 |<br>| 3 : 95 | 561.77 | 145.15 |<br>| 4 : 95 | 684.22 | 143.09 | | Output from ED: The $ED_{95}$ values (5% AYL for the weed-free curve) for each of the four curves and the corresponding estimated SE. |
| 23 | SI(CPWCWeedfree1,c(95,95))<br>Estimated ratios of effect doses<br><br>| | Estimate | Std. Error | t-value | P value |<br>|---|---|---|---|---|<br>| 1/2:95/95 | 1.132637 | 0.558541 | 0.237471 | 0.8128 |<br>| 1/3:95/95 | 1.400952 | 0.817726 | 0.490325 | 0.6250 |<br>| 1/4:95/95 | 1.148179 | 0.504826 | 0.293525 | 0.7698 |<br>| 2/3:95/95 | 1.236894 | 0.659988 | 0.358937 | 0.7204 |<br>| 2/4:95/95 | 1.013722 | 0.375664 | 0.036527 | 0.9709 |<br>| 3/4:95/95 | 0.819571 | 0.397172 | −0.454284 | 0.6506 | | Output from SI: The $ED_5$ values are compared. $ED_5$ values for curves 1 and 2 by analyzing whether the ratio $ED_5(1) : ED_5(2)$ is significantly different from 1.00 (P = 0.81); $ED_5$ values for curves 1 and 3 are not significantly different (P = 0.63); $ED_5$ values for curves 1 and 4 are not significantly different (P = 0.77). Even though the curves are significantly different from each other (line 20), the ends of the critical periods among four sites are statistically similar based on P values. We can safely assume that the end of the critical period (weed-free period) lies around the V13–V16 stage, based on the 5% AYL level (Table 2). |

[a] Abbreviations: ED, effective dose; GDD, growing degree days; AYL, acceptable yield loss; SE, standard error; RSS, root sum square.

Knezevic and Datta: Critical period for weed control • 197

Table 2. The critical period of weed control (CPWC) for corn based on three acceptable yield loss (AYL) levels (2.5, 5, and 10%) at four Nebraska sites (reported as sites 1, 2, 3, and 4), expressed in growing degree days (GDD), corresponding crop growth stage (CGS), and days after crop emergence (DAE). The four-parameter log-logistic model (Equation 1) is fit to data representing (A) an increasing duration of weed interference, resulting in weedy curves (for determining the beginning of the CPWC) (based on the 5% AYL), and (B) an increase in the length of the weed-free period, resulting in weed-free curves (for determining the end of the CPWC) (based on the 5% AYL) at four sites.

| AYL level | Site | Beginning of the CPWC (weedy curves) | | | End of the CPWC (weed-free curves) | | |
|---|---|---|---|---|---|---|---|
| | | GDD (± SE) | CGS[a] | DAE | GDD (± SE) | CGS[a] | DAE |
| % | | Degree days | | d | Degree days | | d |
| 2.5 | 1 | 79 (22) | V2 | 8 | 953 (253) | R1 | 73 |
| | 2 | 79 (23) | V2 | 9 | 854 (225) | R1 | 70 |
| | 3 | 69 (20) | V3 | 12 | 636 (147) | V14 | 54 |
| | 4 | 77 (21) | V2 | 8 | 778 (197) | V16 | 63 |
| 5 | 1 | 118 (51) | V3 | 9 | 784 (203) | V16 | 60 |
| | 2 | 118 (52) | V4 | 16 | 693 (212) | V15 | 56 |
| | 3 | 111 (93) | V3 | 10 | 562 (145) | V10 | 41 |
| | 4 | 107 (44) | V3 | 9 | 684 (143) | V13 | 49 |
| 10 | 1 | 181 (54) | V4 | 11 | 640 (190) | V10 | 43 |
| | 2 | 214 (55) | V5 | 17 | 558 (129) | V10 | 41 |
| | 3 | 177 (51) | V4 | 14 | 494 (159) | V9 | 36 |
| | 4 | 149 (47) | V3 | 13 | 598 (97) | V11 | 43 |

[a] CGS (based on number of leaves): VE, crop emergence; V2, two visible collars; V3, three visible collars; V4, four visible colors; R1, corn anthesis.

regression parameters are not presented; however, they should be presented in a table and discussed for a regular journal publication to obtain greater insights about the experimental results.

The first step in the procedure is to load the drc package, which is done using the *library* function (Table 1, line 01). Lines 02 and 02.1 assign the name for your data file and the location of the file. With a specific command, data from a single Excel spreadsheet with several pages can be read and accessed through R (line 02.1). Line 03 prints the first 6 observations providing a visual for checking whether the data file has been correctly loaded into R.

The second step is the curve-fitting procedure (Table 1, line 04) using the *drm* function. Line 05 summarizes the parameter values for model fit of the curves (four curves in the present example). The results from line 05 are shown in the output section (Table 1, line 16). Line 06 executes a lack-of-fit test. In this particular instance, the test for lack-of-fit (Table 1, line 17) is not significant (P = 0.20), which indicates that the data are well described by the selected model.

The *ED* command, line 07, calculates the user's choice of ED values. The ED values can be calculated based on user's interest and values of 2.5, 5, and 10% are widely accepted AYL levels. In this example, the GDD values for the 5% AYL at four sites were determined from the curves to provide a target range for measuring the effects of increased duration of weed interference on corn grain yield (Table 1, line 07 for the command, and line 18 for the output). The 2.5, 5, and 10% AYL levels are then expressed in GDD characterizing the influence of the duration of weed interference with corresponding crop growth stage (CGS) and days after crop emergence (DAE) (Table 2). A yield loss of more than 5% is generally considered an unacceptable loss from an economic crop-production standpoint (Knezevic et al. 2013). The ED values for each AYL level of interest and site need to be calculated two times: first for obtaining the beginning of the CPWC (weedy curves in Figure 2A) and second for obtaining the end of the CPWC (weed-free curves in Figure 2B) as shown in Table 2.

Line 08 provides estimated relative yields between the curves (using the *SI* function), which compares ED values of interest among curves to detect significant differences at that particular point (e.g., 5%) between the two curves. The output of the *SI* function is presented in line 19 of Table 1. Line 09 illustrates the *plot* command for displaying the visual of fitted curves (Figure 2).

In the present example, lines 01 to 03 (Table 1) are common commands for reading a data file into R. Lines 04 to 09 provide the visual of fitted curves along with the ED values for obtaining the beginning of the CPWC. The outputs of lines 04 to 08 (Table 1) are presented in lines 16 to 19
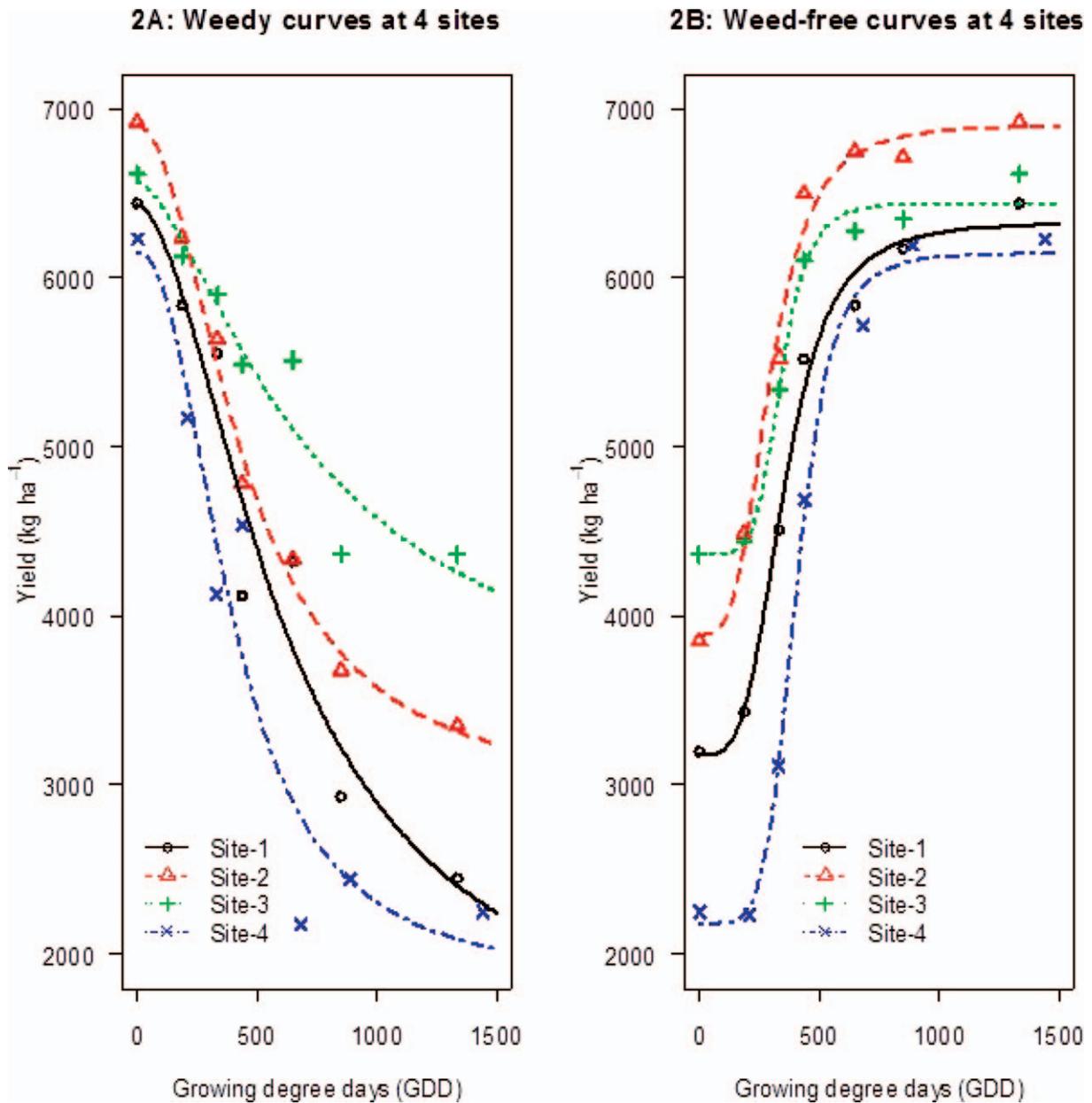
**Figure 2.** Corn grain yields (kg ha$^{-1}$) as influenced by an increasing duration of weed interference and an increase in the length of weed-free period (expressed in growing degree days [GDD]). The four-parameter log-logistic model (Equation 1) (based on the 5% AYL) fit to data representing (A) an increasing duration of weed interference resulting in weedy curves (the beginning of the CPWC), and (B) an increase in the length of weed-free period resulting in weed-free curves (end of the CPWC) at four sites.

(weedy curve). The end of the CPWC is obtained by lines 10 to 15 of Table 1, which provide the visual of fitted curves along with the ED values. The outputs of lines 11 to 14 of Table 1 are presented in lines 20 to 23 (weed-free curve). By following this procedure of data analysis, the data from two separately measured crop–weed competitions can be easily analyzed to obtain the beginning and end of the CPWC (Table 2).

Once users are familiar with the codes and R software for handling data analysis, they can produce the same graph (Figure 2) in many different ways to suit their interests. In the present example, both the weedy and weed-free curves in Figure 2 (Figure 2A: weedy curves at four sites; Figure 2B: weed-free curves at four sites) is presented by site. However, the users can also produce the weedy and weed-free curves by each site as presented in Figure 3 for greater clarity. The codes for producing Figures 3A–D from the plotting standpoint and a brief explanation of the codes are provided in Table 3.
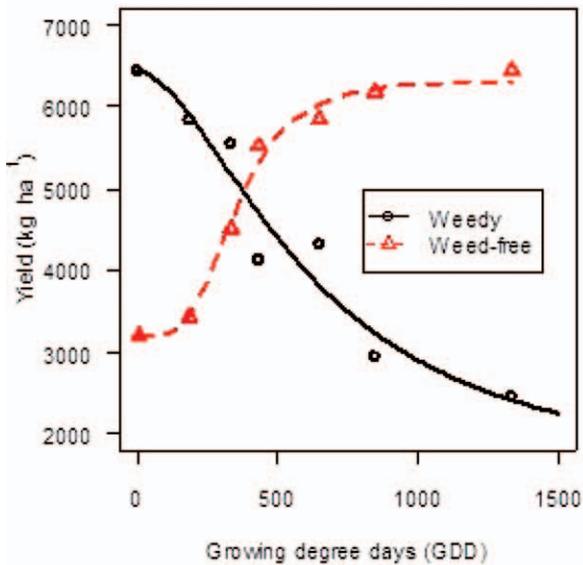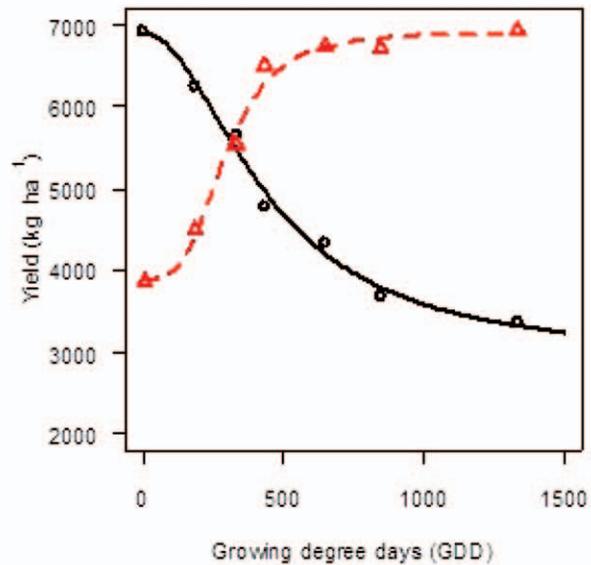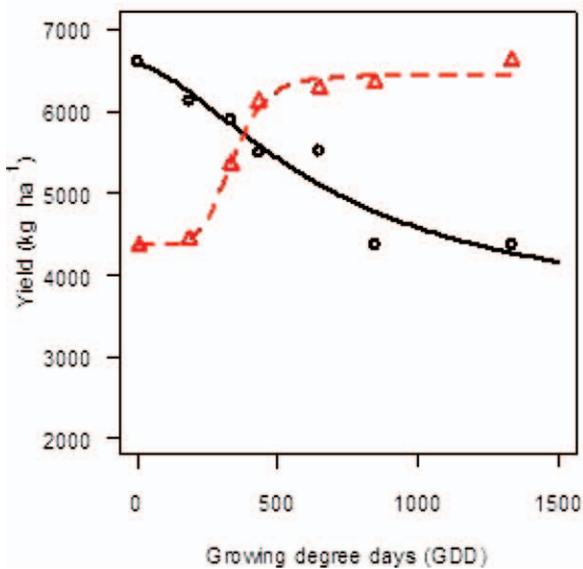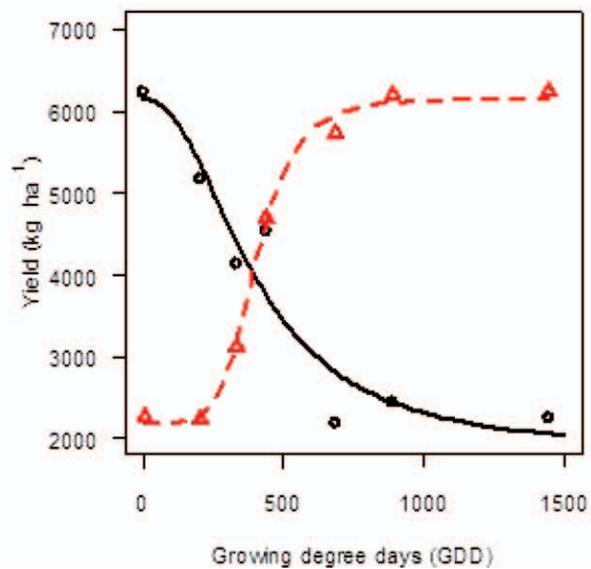
Knezevic and Datta: Critical period for weed control • 199

Figure 3. Corn grain yields (kg ha$^{-1}$) as influenced by an increasing duration of weed interference and an increase in the length of the weed-free period (expressed in growing degree days [GDD]). The four-parameter log-logistic model (Equation 1) fit to data representing an increasing duration of weed interference resulting in weedy curves and an increase in the length of weed-free period resulting in weed-free curves at (A) site 1, (B) site 2, (C) site 3, and (D) site 4.

## Potential Value to Weed Scientists and Weed Control Practitioners

In this article, a simple method for determining the CPWC is illustrated using the open-source and free statistical software R. The statistical analyses of one experimental design are demonstrated. This method allows researchers to make much easier comparisons of results among experiments, sites, and years. The additional value of this data analysis approach is that it can be easily adopted for other types of studies using nonlinear regression (e.g., plant growth analysis).

## Literature Cited

Bunkun B (2004) Critical periods for weed control in cotton in Turkey. Weed Res 44:404–412

Chauhan BS, Johnson DE (2011) Row spacing and weed control timing affect yield of aerobic rice. Field Crops Res 121:226–231

Table 3. Basic commands for fitting the log-logistic curve (Equation 1) using the drc add-on package to R Statistical Software to produce the weedy and weed-free curves for each site as presented in Figures 3 A–D based on the 5% acceptable yield loss (AYL) level.

| R program & output | Comments |
|---|---|
| **SITE 1**<br>FIT WEEDY CURVE<br>*CPWCWeedy1<-drm(WeedyYield~GDD, data=test, fct=l4(fixed=c(NA,NA,NA,NA)), subset = Site == "1", na.action=na.omit)*<br>*modelFit (CPWCWeedy1)*<br>*plot(CPWCWeedy1, log="", xlim=c(0,1500),ylim=c(2000,7000),xlab="Growing degree days (GDD)", ylab = expression(paste("Yield (kg ", hâ-1, ")")), col=1, lwd=2, pch=1, main="3A: Weedy and weed-free curves at site-1")*<br>FIT WEED-FREE CURVE<br>CPWCWeedfree1<-drm(WeedfreeYield~GDD, data=test, fct=l4(fixed= c(NA,NA, NA,NA)), subset = Site == "1", na.action=na.omit)<br>modelFit (CPWCWeedfree1)<br>plot(CPWCWeedfree1, add=T, col=2, lwd=2, lty=2, pch=2)<br>legend(800,5000, legend = c("Weedy", "Weed-free"), col=c(1,2), lty =c(1,2), pch=c(1,2)) | These codes are used to generate weedy and weed-free curves at site 1 (Figure 3A).<br>Statistical outputs are identical to the ones in Table 1, which are not shown here. |
| **SITE 2**<br>FIT WEEDY CURVE<br>*CPWCWeedy2<-drm(WeedyYield~GDD, data=test, fct=l4(fixed=c(NA,NA,NA,NA)), subset = Site == "2", na.action=na.omit)*<br>*modelFit (CPWCWeedy2)*<br>*plot(CPWCWeedy2, log="", xlim=c(0,1500),ylim=c(2000,7000),xlab="Growing degree days (GDD)", ylab = expression(paste("Yield (kg ", hâ-1, ")")), col=1, lwd=2, pch=1, main="3B: Weedy and weed-free curves at site-2")*<br>FIT WEED-FREE CURVE<br>CPWCWeedfree2<-drm(WeedfreeYield~GDD, data=test, fct=l4(fixed= c(NA,NA, NA,NA)), subset = Site == "2", na.action=na.omit)<br>modelFit (CPWCWeedfree2)<br>plot(CPWCWeedfree2, add=T, col=2, lwd=2, lty=2,pch=2) | These codes are used to generate weedy and weed-free curves at site 2 (Figure 3B). |
| **SITE 3**<br>FIT WEEDY CURVE<br>CPWCWeedy3<-drm(WeedyYield~GDD, data=test, fct=l4(fixed=c(NA,NA, NA,NA)), subset = Site == "3", na.action=na.omit)<br>modelFit (CPWCWeedy1)<br>plot(CPWCWeedy3, log="", xlim=c(0,1500),ylim=c(2000,7000),xlab="Growing degree days (GDD)", ylab = expression(paste("Yield (kg ", hâ-1, ")")), col=1, lwd=2, pch=1, main="3C: Weedy and weed-free curves at site-3")<br>FIT WEED-FREE CURVE<br>CPWCWeedfree3<-drm(WeedfreeYield~GDD, data=test, fct=l4(fixed= c(NA,NA, NA,NA)), subset = Site == "3", na.action=na.omit)<br>modelFit (CPWCWeedfree3)<br>plot(CPWCWeedfree3, add=T, col=2, lwd=2, lty=2,pch=2) | These codes are used to generate weedy and weed-free curves at site 3 (Figure 3C). |
| **SITE 4**<br>FIT WEEDY CURVE<br>CPWCWeedy4<-drm(WeedyYield~GDD, data=test, fct=l4(fixed=c(NA,NA, NA,NA)), subset = Site == "4", na.action=na.omit)<br>modelFit (CPWCWeedy4)<br>plot(CPWCWeedy4, log="", xlim=c(0,1500),ylim=c(2000,7000),xlab="Growing degree days (GDD)", ylab = expression(paste("Yield (kg ", hâ-1, ")")), col=1, lwd=2, pch=1, main="3D: Weedy and weed-free curves at site-4")<br>FIT WEED-FREE CURVE<br>CPWCWeedfree4<-drm(WeedfreeYield~GDD, data=test, fct=l4(fixed= c(NA,NA, NA,NA)), subset = Site == "4", na.action=na.omit)<br>modelFit (CPWCWeedfree4)<br>plot(CPWCWeedfree4, add=T, col=2, lwd=2, lty=2,pch=2) | These codes are used to generate weedy and weed-free curves at site 4 (Figure 3D). |

Elezovic I, Datta A, Vrbnicanin S, Glamoclija D, Simic M, Malidza G, Knezevic SZ (2012) Yield and yield components of imidazolinone-resistant sunflower (*Helianthus annuus* L.) are influenced by pre-emergence herbicide and time of post-emergence weed removal. Field Crops Res 128:137–146

Evans SP, Knezevic SZ, Shapiro CA, Lindquist JL (2003) Nitrogen application influences the critical period for weed control in corn. Weed Sci 51:408–417

Everman WJ, Clewis SB, Thomas WE, Burke IC, Wilcut JW (2008) Critical period of weed interference in peanut. Weed Technol 22:63–67

Gilmore EC, Rogers RS (1958) Heat units as a method of measuring maturity in corn. Agron J 50:611–615

Hall MR, Swanton CJ, Anderson GW (1992) The critical period of weed control in grain corn (*Zea mays*). Weed Sci 40:441–447

Kalaher CJ, Stoller E, Young B, Roskamp G (2000) Proper timing of a single post-emergence glyphosate application in three soybean row spacings. Proceedings of the North Central Weed Science Society 55:113

Knezevic SZ, Elezovic I, Datta A, Vrbnicanin S, Glamoclija D, Simic M, Malidza G (2013) Delay in the critical time for weed removal in imidazolinone-resistant sunflower (*Helianthus annuus*) caused by application of pre-emergence herbicide. Int J Pest Manag 3:229–235

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

Knezevic SZ, Evans SP, Mainz M (2003) Row spacing influences the critical timing for weed removal in soybean (*Glycine max*). Weed Technol 17:666–673

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

Martin SG, Van Acker RC, Friesen LF (2001) Critical period of weed control in spring canola. Weed Sci 49:326–333

Mohammadi G, Javanshir A, Khooie FR, Mohammadi SA, Salmasi SZ (2005) Critical period of weed interference in chickpea. Weed Res 45:57–63

R Development Core Team. 2013. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/. Accessed March 3, 2014

Ritz CJ, Streibig C (2005) Bioassay analysis using R. J Stat Softw 12:1–22

[SAS] Statistical Analysis Systems. 1999. SAS. Version 8.1. Cary, NC: Statistical Analysis Systems Institute

Smitchger JA, Burke IC, Yenish JP (2012) The critical period of weed control in lentil (*Lens culinaris*) in the Pacific Northwest. Weed Sci 60:81–85

Swanton CJ, O'Sullivan J, Robinson DE (2010) The critical weed-free period in carrot. Weed Sci 58:229–233

Swanton CJ, Weise SF (1991) Integrated weed management: the rationale and approach. Weed Technol 5:648–656

Tursun N, Akinci IE, Uludag A, Pamukoglu Z, Gozcu D (2012) Critical period for weed control in direct seeded red pepper (*Capsicum annum* L.). Weed Biol Manag 12:109–115

Tursun N, Bukun B, Karacan SC, Ngouajio M, Mennan H (2007) Critical period for weed control in leek (*Allium porrum* L.). Hortscience 42:106–109

Van Acker CR, Swanton CJ, Weise SF (1993) The critical period of weed control in soybean [*Glycine max*. (L) Merr.]. Weed Sci 41:194–200

Weaver SE, Tan CS (1983) Critical period of weed interference in transplanted tomatoes (*Lycopersicum esculentum*): growth analysis. Weed Sci 31:476–481

Woolley BL, Michaels TE, Hall MR, Swanton CJ (1993) The critical period of weed control in white bean (*Phaseolus vulgaris*). Weed Sci 41:180–184

Zimdahl RL (1980) Weed–Crop Competition—A review. Corvallis, OR: Oregon State University. 195 p

Zimdahl RL (1988) The concept and application of the critical weed-free period. Pages 145–155 *in* Altieri MA and Liebmann M, eds. Weed Management in Agroecosystems: Ecological Approaches. Boca Raton, FL: CRC Press