Weed Technology. 2006. Volume 20:41–51

Between-Observer Differences in Relative Corn Yield vs. Rated Weed Control[1]

WILLIAM W. DONALD[2]

Abstract: Crop yield and weed control rating have been used to measure weed and crop response to weed management treatments, eliminate unacceptable weed management treatments, and select ''best'' treatments for recommendation to farmers. However, the mathematical relationship between crop yield and rated weed control has not been reported before from such treated screening experiments. Likewise, differences have not been reported before in rated weed control among experienced observers (i.e., reliability) when rating the same experiments and for an experienced observer over time (i.e., repeatability). Data from published experiments on zone herbicide application in field corn in which weeds reduced yield to various amounts were reanalyzed to examine these issues. For this study, relative corn yield was calculated as a percentage of the 1× broadcast herbicide rate for two observers and either three experimental site-years or their average. For observer A, relative corn yield (%) increased linearly as rated total weed control (%) increased for all 3 site-yr and their average. For observer B, equations were curvilinear in 2 of 3 site-yr. For both observers, equations accounted for little data variability in relative corn yield (r^2 = 0.25 and 0.25 in site-year 1, respectively, 0.38 and 0.36 in site-year 2, 0.58 and 0.57 in site-year 3, and 0.43 and 0.42 for their average). When rated total weed control by observer A was graphed against that of observer B, the relationship was a nearly ideal 1:1 linear relationship in only 1 of 3 site-yr. In two other site-years, equations were nonlinear, indicating that one observer distinguished smaller differences between treatments at lower rated control than the other observer. Between-row total weed cover and in-row total weed height influenced observer weed control rating.

Nomenclature: Corn, *Zea mays* (L.) #[3] ZEAMX 'Pioneer 33G28'.
Additional index words: Rating; weed, *Setaria faberii* (L.) Beauv. # SETFA, *Amaranthus rudis* Sauer. # AMATA.
Abbreviations: IR, in-row; BR, between-row; ZHA, zone herbicide application.

INTRODUCTION

For over 50 yr, weed scientists have relied on relatively few measurements for evaluating weed control treatments in agronomic field crops, usually crop grain yield and visually evaluated weed control (i.e., ''rated weed control'' or ''weed control rating,'' hereafter). These measured, dependent response variables were used to screen the relative efficacy of many different weed control treatments, optimize treatment performance in many agronomic crops and environments, and pick the ''best'' weed management treatments or systems for recommendation to farmers. In practice, weed control rating and crop yield are used to either (1) rank the effectiveness of alternative treatments and ''pick a winner'' or (2) quantitatively describe a graded response to treatments that vary continuously to find the lowest, effective, consistent rate (e.g., herbicide dose–response research). In case 1, rated weed control is used to rank treatments as discrete entities, whereas in case 2, yield and rated weed control are assumed to be correlated functions of treatment rate. Other estimates of response to treatment, including weed density, height, cover, volume, aboveground biomass (e.g., dry or wet weight), grain dockage, grain moisture content, crop grain yield components, and net returns, have been reported less often. In agricultural sciences other than agronomy, additional measures of response to treatment might be of interest: crop quality or appearance (horticulture), harvesting efficiency (agricultural engineering), and weed seed production and soil seed banks (plant ecology), among others. These latter measurements are used to address other research questions. However, such measurements are seldom used by extension researchers, commercial firms, or agronomists

[1] Received for publication October 27, 2004, and in revised form May 18, 2005.
[2] Former Research Agronomist, USDA-ARS, 269 Agricultural Engineering Building, UMC, Columbia, MO 65211.
[3] Letters following this symbol are a WSSA-approved computer code from *Composite List of Weeds*, Revised 1989. Available only on computer disk from WSSA, 810 East 10th Street, Lawrence, KS 66044-8897.

to quickly, inexpensively, or efficiently screen many treatments for effectiveness in a short time.

A Brief Literature Review of Rated Weed Control. About 20 yr ago, Frans et al. (1986) reviewed the limited information on rating weed control. Rated weed control is the chief measurement used to evaluate weed response to management treatments and has been applied to either individual weed species; weed genera; or weeds grouped by similar growth habit (grass weeds, broadleaf weeds, sedges, etc.), similar life histories (annuals, biennials, perennials), or as mixed species populations (total weed control). Usually, rated weed control values are reported as percentages with values between 0 (i.e., weedy checks or control plots in which weeds are not treated or controlled) and 100% (i.e., weed free or a ground bare of weeds). Although Frans et al. (1986) suggested using a weed control rating system subdivided into 10% intervals, 1% intervals are commonly reported. Other suggested scales and systems for rating weed control have not been widely adopted (e.g., Darwent et al. 1997; Hamill et al. 1977).

In general, rated weed control does not simply or directly measure weed kill, survivorship, population density, biomass, seed production, or fitness. Perhaps rated weed control can best be described as a combined estimate of weed ground cover, height, biomass, and density. Many who report rated weed control in scientific publications interpret it as a judgment of whether a treatment would be acceptable to farmers or land managers who believe that the ground should be bare of live weeds.

Because rated weed control uses human vision and requires no specialized equipment, it is physically undemanding, inexpensive, and rapid. Consequently, rated control is probably the quickest and cheapest method for estimating weed response to treatment. However, rated weed control also requires subjective human judgment, which is based on training and experience. Because rated weed control depends on both human vision and judgment, it cannot easily be automated or scaled up for use on whole fields.

When rated weed control was first adopted, no other field methods were available for quickly and inexpensively measuring or screening many weed control treatments within 1 or 2 d of work. Other more objective, quantitative measurements, such as weed density and biomass, required specialized equipment and were physically difficult, laborious, slow, and costly. These latter limitations prevented many treatments from being screened daily.

Despite being relatively inexpensive, quick, and successful, is weed control rating a valid scientific measurement? Scientific measurements that use the human senses or even judgment or opinion are made in many scientific disciplines for various purposes and can even be the object of study. However, objective, semiquantitative (i.e., nominal or ordinal scale) or quantitative (i.e., continuous scale) scientific measurements require comparison to standard reference scales for defining fundamental units and error in terms of statistical accuracy (i.e., closeness of measurements to true values) and precision (i.e., closeness of measurements of the same quantity) over time (Sokal and Rohlf 1981). Rated weed control has no such commonly accepted, absolute, calibrated scale or standard of comparison. Although the upper limit for rated weed control is taken as 100% for bare ground that is free of live weeds, the lower limit (i.e., 0% for weedy treatments) is poorly defined, and its appearance changes as weed and crop populations grow and compete over time. Consequently, even if the same observer rated the same treated research plots, the meaning of 0% rated weed control can change over a growing season. Because the basis of comparison (i.e., 0% rated weed control of the weedy check) is ill defined and changeable, the meaning of rated weed control is ambiguous. To reduce this ambiguity, weed control rating can be supplemented with additional weed measurements to better characterize the weedy check, at a minimum. The author could find no refereed publications concerning the factors that control the statistical accuracy and precision of rated weed control.

In other agricultural disciplines, scientists who use rating systems have tried to minimize or overcome these concerns by creating standardized scales with well-defined upper and lower limits and fewer, well-defined subcategories (i.e., 5 to 10) (Horsfall and Cowling 1978; Horst et al. 1984; James 1971). However, in entomology and plant pathology, rating scales are often limited to damage by one pest on one crop and can even be limited to particular plant parts at specific growth stages. One attempt to create a standardized scale for rating weed control has not been widely adopted (Harvey 1993). Nevertheless, this publication helps explain rated weed control to farmers who use treatment recommendations that are based on it.

Several limitations and biases of human vision and subjective judgment in making visual ratings for quantitative scientific measurement were mentioned in other agricultural sciences (Hebert 1982; Horsfall and Cowling 1978; Horst et al. 1984; James 1971; James and Teng 1979; Nilsson 1995a, 1995b). For example, plant pathologists found that human vision made rating subjective, inconsistent, and inaccurate (Spomer and Smith 1988). Human vision was subject to fatigue, visual habituation, lack of concentration, and differences among individual observers (Hebert 1982; Nilsson 1995a, 1995b; Nutter and Schultz 1995; Nutter et al. 1993). Weed control ratings varied because observers perceived visual stimuli differently, were sensitive to different wavelengths of visible light (e.g., color blindness), and varied in their ability to differentiate among plants having different shapes, sizes, and patterns. Observer rating also was influenced by shading or lighting contrast, the background, and neighboring plants. For example, tall-growing plants can obscure or hide weeds from observers. In one study, rating error exceeded 20% and depended on the observer's familiarity with the morphology, distribution, and identity of the plant species present (Kennedy and Addison 1987). Rating error also is not uniform across the entire rating scale (Muir and McCune 1987). Observers rated more accurately and best distinguished small differences among treatments at low and high rating values. Thus, reporting a 1% resolution for rated weed control might only be reasonable above 90% and below 10% rated control, but this has not been researched. Use of 1% rating values in the middle 50% of the weed control rating scale is questionable. In practice, finer distinctions between treatments can be discerned in rated weed control than in crop yield, which is usually less responsive to treatment. In addition, rated control of individual weed species cannot be summed to estimate total weed control or control of groups of weeds of similar growth habit (i.e., grasses, broadleaf weeds, etc.). Total weed control and weed control for groups of weeds must be rated separately. Consequently, rated weed control lacks some of the mathematical characteristics of quantitative scientific measurements (i.e., addition and subtraction).

Models of rated weed control. Researchers organize measurements into explicitly defined or implicitly assumed ''models'' of reality to communicate their research results. When models are implicit, their assumptions and structure can only be inferred from the way in which researchers discuss their published results. In refereed publications, results of weed management research frequently are tabulated as columns of independent variables (i.e., treatments) followed by columns of measured dependent response variables, such as rated weed control by species, visually rated crop damage (i.e., phytotoxicity by the treatment), and crop yield. The reported dependent variables represent several different measurements on the same treated plots. For example, 60% of 102 articles in *Weed Technology* in 2003 tabulated results in this way. By presenting such tables, authors might imply and readers might assume that treatments directly caused the rated weed control and crop yield (Figure 1, scenario 1). The assumption that treatment is the only factor to cause a response is often stated as such in published results and discussions.

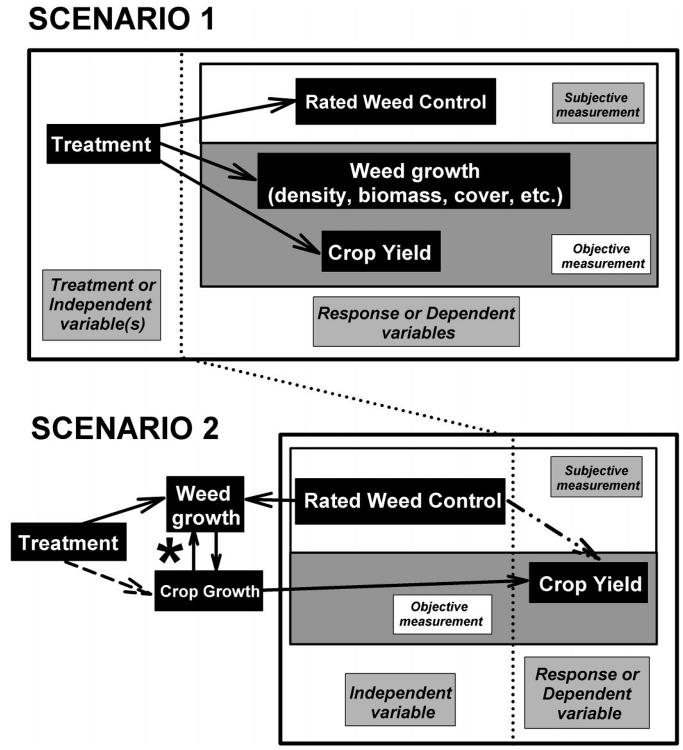

Figure 1. In scenario 1, dependent response variables (*Y*), such as subjective rated weed control, and objectively measured weed growth and crop yield are related to weed control treatment, the independent variable (*X*). In scenario 2 the frame of reference (box) for analyzing the data of the same experiment has been changed, and the independent and dependent variables are redefined and partitioned differently (dotted line) than in scenario 1. * Interference between the crop and weeds.

Scenario 2 is a logical alternative model to scenario 1 for describing the relationships between these independent and dependent variables (Figure 1). In scenario 2, treatments can directly and indirectly influence both crop growth (i.e., phytotoxicity and occasionally stimulation) and weed growth (i.e., kill or growth suppression), which interact with each another (Figure 1, scenario 2, left side). Consequently, rated weed control and yield are not simply or directly related to treatment alone. Treatment efficacy also could depend on crop suppression of weeds and interference among neighboring weeds. For example, competitive crops, such as corn (*Zea mays*), reduced the cover of weeds growing in rows more than between rows (Donald et al. 2004b). Consequently, lower soil residual herbicide rates were required in-row (IR) than between-row (BR) to reduce weed cover to the same extent. In scenario 2, rated weed control integrates the combined effects of treatment, weed growth, crop growth, and their interaction. Thus, it integrates more interactions than simply the effect of the imposed treatment alone.

Mathematical relationships between variables. The regression equations relating absolute or relative (%) crop yield to either crop density or growth, weed density, or both have been reviewed and are usually hyperbolic or sigmoidal functions (Lindquist and Knezevic 2001). In almost all competition or interference experiments, the crop and weed populations were not treated with herbicides. To the author's knowledge, such regression equations have not been published after herbicides have been imposed, as in herbicide screening research. When untreated weed densities are created for competition or interference experiments, weed control was never rated or reported relative to the highest weed density present. To the author's knowledge, neither competition research (i.e., untreated weeds) nor weed management research (i.e., usually herbicide-treated weeds) have published reports on the functional equations relating absolute or relative crop yield (%) to rated weed control (%) (Figure 1, scenario 2, dashed arrow from rated weed control to yield). However, in scenario 2, rated weed control is indirectly related to crop yield through weed growth, crop growth, and their interaction (Figure 1, scenario 2, solid arrows between rated weed control and crop yield). Because crop density or growth and weed density or growth are related to crop yield via nonlinear functions (Lindquist and Knezevic 2001), absolute or relative crop yield can be a nonlinear function of rated weed control (e.g., hyperbolic or sigmoidal function).

One objective of this research was to determine the best regression equation from published data for relating relative crop yield (%) to rated weed control (%) (Figure 2) (Donald et al. 2004a). Relative crop yield (%) was used to facilitate comparisons across site-years on a common basis. The traditional null hypothesis that relative corn yield and rated weed control are independent or unrelated was not explored (Figure 2a). The simplest hypothesis was that relative corn yield increased linearly as rated total weed control (%) increased, starting from a rated weed control (X) value of 0% (Figure 2b). Two other possible equations are that (1) yield increases linearly, but only above a threshold rated weed control value greater than 0% (Figure 2c), and (2) yield increases nonlinearly as rated weed control increases without a threshold for rated weed control (two possible nonlinear equations pictured, but others are possible) (Figure 2d). However, even when rated weed control is 0%, few weed infestations reduce absolute or relative crop yield to 0%. Consequently, the null hypotheses (Figures 2b–d) were re-expressed with a minimum threshold above 0% for relative yield (Y) at a rated weed control value of 0% (Figures 2e–g). After re-expressing the research hypotheses, the null hypothesis was that relative yield increased linearly as rated weed control increased (Figure 2e). Alternative hypotheses (Figures 2f and 2g) are analogous to previously presented cases (Figures 2c and 2d).

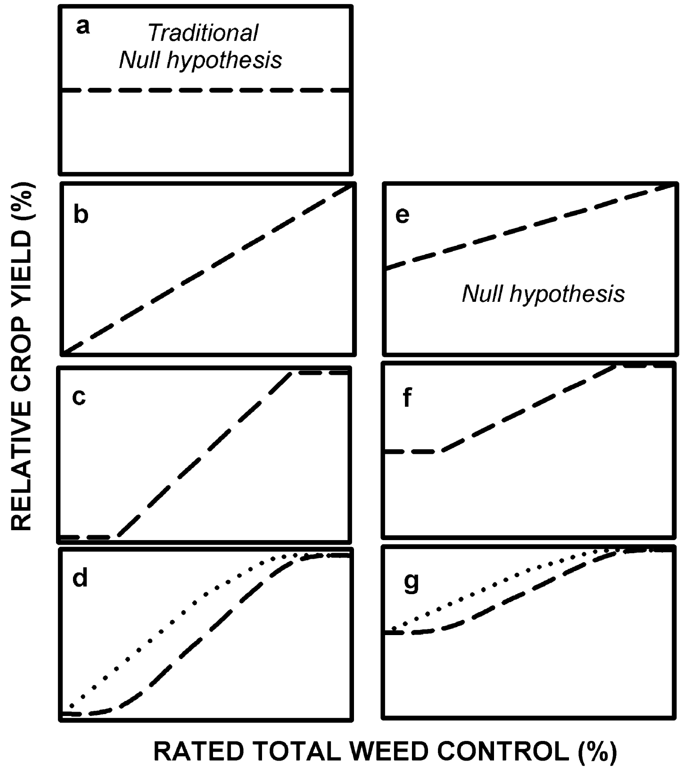

Figure 2. Alternative hypothetical regression relationship relating relative crop yield (Y, %), the dependent response variable, to rated weed control (X, %), the independent variable from scenario 2 of Figure 1 (dashed arrow). The seven alternatives are (a) Y is independent of X (traditional null hypothesis); (b) Y increases linearly with increasing X, where $X \geq 0\%$, and with $Y = 0\%$ at $X = 0\%$ and $Y = 100\%$ at $X = 100\%$; (c) Y increases linearly with increasing X, above a threshold of $X > 0$, and with $Y = 0\%$ at $X = 0\%$ and $Y = 100\%$ at $X = 100\%$; (d) Y increases nonlinearly as X increases, and with $Y = 0\%$ at $X = 0\%$ and $Y = 100\%$ at $X = 100\%$; (e) Y increases linearly with increasing X, where $X \geq 0\%$, and with $Y > 0\%$ at $X = 0\%$ and $Y = 100\%$ at $X = 100\%$ (null hypothesis); (f) Y increases linearly with increasing X, above a threshold of $X > 0$, and with $Y > 0\%$ at $X = 0\%$ and $Y = 100\%$ at $X = 100\%$; (g) Y increases nonlinearly as X increases, with $Y > 0\%$ at $X = 0\%$ and $Y = 100\%$ at $X = 100\%$.

It is widely recognized that observers sometimes rate weed control of the same research plots differently, and their rating can change over time, but the size of this variation has never been reported for rated weed control. Consequently, a second research objective was to deter-

Table 1. Dates for field operations, treatments, or measurements at Bradford or Greenley.[a]

Field operation or measurement	Bradford 2001			Greenley 2001			Greenley 2002		
	Date	DAP	DAE	Date	DAP	DAE	Date	DAP	DAE
Disk soil	11/10/01	—	—	11/12/01	—	—	11/23/02	—	—
Apply glyphosate	—	—	—	—	—	—	5/20/02	—	—
Plant crop	4/26/01	0	—	4/20/01	0	—	5/22/02	0	—
Apply PRE atrazine + s-metolachlor + clopyralid + flumetsulam	4/30/01	4	—	4/24/01	4	—	6/4/02	13	2
Crop emergence	5/3/01	7	0	4/29/01	9	0	6/2/02	11	0
Measure crop stand	5/23/01	27	20	5/16/01	26	17	6/14/02	23	12
Weed-free check only									
Apply glufosinate	5/29/01	33	26	5/16/01	26	17	—	—	—
Hoe and hand-pull weeds	6/13/01	44	41	7/24/01	95	86	6/14/02	23	12
	6/14/01	49	42	—	—	—	6/25/02	34	23
	7/6/01	71	64	—	—	—	7/16/02	55	44
	—	—	—	—	—	—	8/9/02	79	68
Rate weed control, observer A	7/20/01	85	78	7/18/01	89	80	7/16/02	55	44
Rate weed control, observer B	8/1/01	97	90	8/7/01	109	100	7/3/02	42	31
Photograph weed cover	7/16–17/01	81	74	7/11/01	82	73	7/2/02	41	30
	7/30/01	95	88	8/7/01	109	100	—	—	—
Harvest corn	9/27/01	154	147	11/6/01	200	191	9/25/02	126	115

[a] Abbreviations: DAP, days after planting; DAE, days after corn emergence.

mine whether equations relating relative corn yield to rated total weed control were the same for different experienced observers rating weed control of identical treatments in the same experiments over several site-years. A third objective was to determine the relative contribution of weed appearance (i.e., measured weed height or cover) growing in rows or between rows to rated total weed control.

MATERIALS AND METHODS

Agronomic Practices. For this research, published data were reanalyzed (Donald et al. 2004a). Field corn was planted following soybeans [*Glycine max* (L.) Merr.] for 3 site-yr at (1) the University of Missouri's Bradford Research and Extension Center in north-central Missouri near Columbia (38°53′43.5″N, 92°12′37.9″W, 269 m altitude) in 2001 and (2) the University of Missouri's Greenley Memorial Research Center in northern Missouri near Novelty (40°0′45″N, 92°12′29″W, 254 m altitude) in 2001 and 2002. The Bradford site was on a Mexico silty clay loam (fine, smectitic, mesic Aeric Vertic Epiaqualfs), whereas the Greenley site was a Putnam silt loam (fine, montmorillonitic, mesic Vertic Albaqualfs). The soil at Bradford had 18 to 20% sand, 46 to 48% silt, 34% clay, 2.9 to 3.4% organic matter, and salt pH 5.5 to 5.7, whereas the soil at Greenley had 12 to 16% sand, 52 to 54% silt, 30 to 36% clay, 3 to 3.4% organic matter, and salt pH 6. Salt pH values run approximately 0.5 units lower than the customary water pH values. At both locations, rainfall occurred soon after herbicide application (Donald et al. 2004a).

Dates for field operations, treatments, and measurements are summarized in Table 1. In spring before planting, each site was shallowly disked to about 5 cm to redistribute and facilitate residue degradation, to incorporate fertilizer, and to prepare the seedbed. Corn was fertilized with nitrogen, phosphorous, and potassium for a grain yield goal of 10,000 kg/ha on the basis of soil tests and recommendations of the University of Missouri soil testing laboratory. $N–P_20_5–K_20$ was broadcast at 160–69–93 kg/ha at Bradford in 2001 and at 180–56–112 kg/ha at Greenley in 2001 and 2002 before incorporation. Glufosinate-resistant 'Pioneer 33G28' corn seed was planted 1.3 to 1.9 cm deep in 76-cm rows at 68,000 seeds/ha. Weather data were reported (Donald et al. 2004a).

At both sites, giant foxtail (*Setaria faberii* Herrm.) was the major weed present. At Bradford, common waterhemp (*Amaranthus rudis* Sauer) was the major broadleaf weed, but scattered, sparse Pennsylvania smartweed (*Polygonum pensylvanicum* L. POLPY) and common ragweed (*Ambrosia artemisiifolia* L. AMBEL) also were present. At Greenley, common waterhemp was the predominant broadleaf weed, and common cocklebur (*Xanthium strumarium* L. XANST), ladysthumb smartweed (*Polygonum persicaria* L. POLPE), Pennsylvania smartweed, and velvetleaf (*Abutilon theophrasti* Medik. ABUTH) were very sparse.

Herbicide Treatments. The herbicide mixture was chosen to control all weeds present over a range of registered rates. Atrazine[4] + s-metolachlor + clopyralid +

[4] Mention of trade names or commercial products in this article is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the U.S. Department of Agriculture.

flumetsulam were applied PRE by either broadcast application or zone herbicide application (ZHA) (i.e., different rates between rows and in rows) to create different BR and IR weed cover. BR and IR zone widths were each 50% of the corn row width, 76 cm, and were created with even spray nozzle tips with limited spray overlap (about ⅛ swath width overlap). The 1× rate of atrazine + s-metolachlor[5] + clopyralid + flumetsulam[6] was 2.24 + 1.75 + 0.211 + 0.067 kg ai/ha, respectively. The BR + IR ZHAs were applied at 0× (i.e., weedy check), 0.25×, 0.5×, 0.75×, and 1× in all possible BR + IR zone combinations in 2001 and a subset of these treatments in 2002. Broadcast applications were made with a single-boom backpack sprayer, whereas ZHAs were made with a dual-boom ZHA sprayer. Treatments were arranged in a randomized complete block design with four or five blocks (Hoshmand 1994). Individual plots measured 3 by 13.7 m at the Bradford Research and Extension Center and 3 by 9.1 m at the Greenley Research Center.

A backpack sprayer with flat fan nozzle tips[7] spaced 76.2 cm apart on a spray boom was used for broadcast herbicide application with a spray volume of 168 L/ha with the use of compressed CO_2 at 193 kPa as a propellant and a ground speed of 1.6 km/h. A dual-boom ZHA backpack sprayer had even spray nozzle tips[8] spaced 76.2 cm apart on two separate spray booms held adjacent to one another on a frame. Adjacent dual booms were offset 38.1 cm from each other, so that BR and IR even nozzle tips were separated 38.1 cm apart. Each dual boom applied a carrier volume of 166 L/ha through separate compressed CO_2 propellant systems at the same pressure and ground speed as previously described. To maintain uniform BR and IR zone widths, the boom height above the ground was held constant by suspending the frame holding the booms from cord that ran from each end of the frame to the top of backpack ZHA sprayer. The guy lines transferred the weight of the boom to the applicator's back, rather than from the applicator's arms; this minimized applicator fatigue and variation in boom height during spraying treatments. The boom heights were about 84 and 51 cm above the soil surface for broadcast and ZHA dual-boom sprayers, respectively.

Seedbed preparation killed the weeds present before planting. Weed-free checks were created with a sequence of POST broadcast-applied glufosinate at 0.28 kg ai/ha followed by hoeing and hand-pulling several times during the growing season until corn silking to kill later emerging weeds (Table 1). Although these hand-weeded plots were not weed-free by harvest, weeds emerging after silking and canopy closure do not reduce corn grain yields (Bedmar et al. 1999; Hall et al. 1992).

Measurements. For all 3 site-yr, two experienced observers other than the author visually evaluated total weed control of the same plots at midseason on a scale of 0 (no control) to 100% (complete kill) (Table 1). Although control of individual species and weed groups (grasses and broadleaf weeds) was visually evaluated in rows, between rows, and for entire plots, only results for total weed control are presented. Note that observer A rated total weed control in 1% increments at Bradford in 2001, but in roughly 5% increments at Greenley in 2001 and 2002, whereas observer B rated total weed control in roughly 5% increments for all 3 site-yr.

In the weedy checks, projected ground cover (''cover'' hereafter) of total, grass, and broadleaf weeds (%) was measured from photographs taken in and between crop rows (see Donald et al. 2004a).

After cutting borders at either end of all plots, corn was combine-harvested from the two center rows in an area measuring 1.5 by 10.6 m at Bradford and 1.5 by 8.2 m at Greenley, and grain yields were adjusted to 15% moisture content. To better facilitate comparisons between observers across site-years on a uniform basis, absolute corn yields (kg/ha) were re-expressed as relative corn yield (% of the 1× broadcast herbicide treatment) separately by block (Figures 3 and 4).

Statistical Analysis. For each site-year and its average, relative corn yield (%, the dependent variable) was regressed against rated total weed control (%, the independent variable) for each observer (SPSS 2001). The following alternative least squares regression equations were compared: (a) $Y = a + bX$, (b) $Y = a + bX + cX^2$, and (c) $Y = a + bX^2$. Because F values for all equations were significant ($P \leq 0.05$), simplest parsimonious equations were selected that had both the highest adjusted r^2 and coefficients for the X terms that were different from 0%. Equation suitability was based on lack of fit statistics, adjusted r^2, and visual inspection of the distribution of residual plot scatter vs. the independent variables. Software[9] was used to prepare 2-D graphs

[5] Bicep II Magnum (atrazine + s-metolachlor), Syngenta, Greensboro, NC 27419-8300.

[6] Hornet (clopyralid + flumetsulam), Dow AgroSciences LLC, Indianapolis, IN 46268-3033.

[7] Teejet 6501 flat fan nozzle tips, Spraying Systems Co., North Avenue at Schmale Road, Wheaton, IL 60188.

[8] Teejet 4001E even nozzle tips, Spraying Systems Co., North Avenue at Schmale Road, Wheaton, IL 60188.

[9] SigmaPlot 2000 software, SPSS Inc., 444 North Michigan Avenue, Chicago, IL 60611.

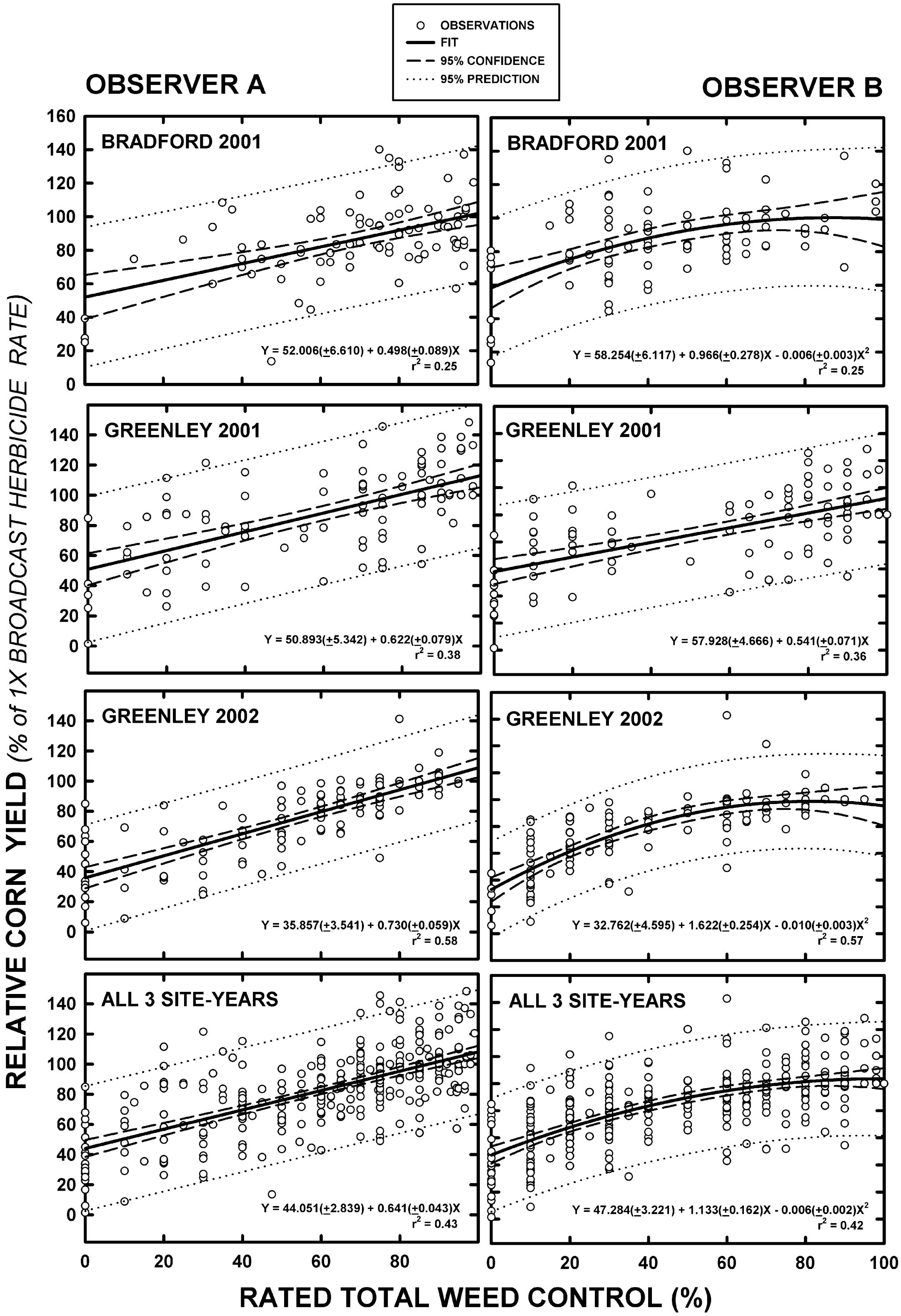

Figure 3. Linear and nonlinear regression equations for relative corn yield (%) vs. rated total weed control (%) by two different observers for 3 site-yr. Equation coefficients (± standard errors) and coefficients of determination (r^2) are presented.

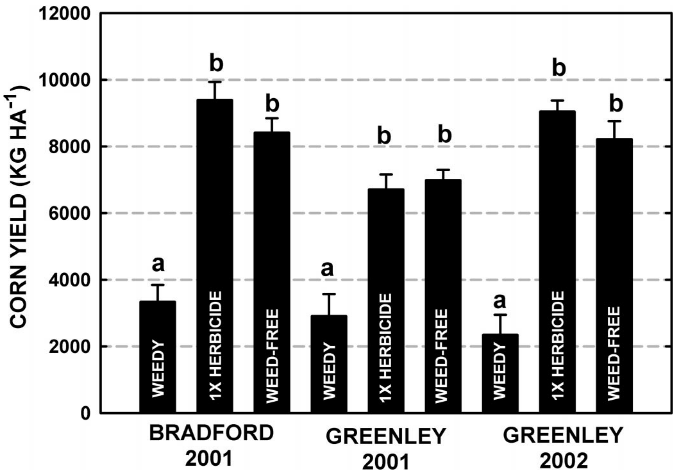

Figure 4. Absolute corn yield (kg ha^{-1}) for three treatments in 3 site-yr. Means ± standard errors are presented. Within each site-year, means followed by the same letter were not different by Duncan's multiple range test ($P \geq 0.05$).

of the best equations. In the figures, 95% confidence and prediction intervals provide different information. Confidence intervals are the range in which regression line values will occur when measurements are repeated, whereas prediction intervals are the range in which data values will occur. Thus, 95% of the time, regression lines will fall between the 95% confidence intervals pictured (inner dashed lines), whereas 95% of the time, the data values will fall within the prediction intervals (outer dotted lines).

In addition, rated total weed control by observer A was regressed against that by observer B for each site-year and for the 3–site-yr average. Finally, rated total weed control (%) was regressed on IR and BR total weed cover (%) and height (cm) (Myers and Montgomery 2002; SPSS 2001). Forward, backward, and stepwise regression model selection procedures usually provided identical best equations.

RESULTS AND DISCUSSION

Absolute Corn Yield. For each of the 3 site-yr, corn yields of the 1× broadcast herbicide and weed-free treatments were indistinguishable (Figure 4). Thus, on the basis of yield or crop stand (data not presented), the applied herbicides did not damage corn. The range in absolute corn yields between the 1× broadcast herbicide and weedy (0×) treatments differed among site-years. At Bradford in 2001, the absolute yield of the weedy treatment was 36 and 40% of the broadcast 1× rate herbicide and weed-free treatments, respectively. At Greenley in 2001, it was 43 and 42%, respectively, and at Greenley in 2002, it was 26 and 29%, respectively. Year-to-year variation in weather, especially water stress during pollination and silking in July and August, likely caused yield differences among site-years (see Donald et al. 2004a). Absolute yield was re-expressed as relative yield to minimize this variation and facilitate comparison among site-years.

Weeds Present in Weedy Check. By midseason in weedy check plots, BR total weed cover exceeded IR weed cover in all 3 site-yr (Donald et al. 2004a). By midseason, BR and IR total weed cover (mean ± standard error) were 74% ± 3 and 57% ± 11 of the ground cover, respectively, in the weedy checks at Greenley in 2001 and 83% ± 7 and 59% ± 7, respectively, at Bradford in 2001. In contrast, the total BR and IR weed cover were 67% ± 6 and 60% ± 5, respectively, at Greenley in 2002. When photographs were taken in 2002, in contrast to 2001, the corn canopy had not yet closed and shaded the ground (Table 1).

By midseason, giant foxtail, the chief weed present, accounted for most BR and IR total weed cover in weedy checks at all site-years, and common waterhemp accounted for most remaining weed cover (Donald et al. 2004a). When giant foxtail cover was expressed as a percentage of total weed cover at midseason, rather than ground cover, BR and IR giant foxtail cover were similar in all 3 site-yr. At Greenley in 2001, giant foxtail was 82% of total BR weed cover and 81% of total IR weed cover in weedy checks. At Bradford in 2001, giant foxtail accounted for 63 and 61% of total BR and IR weed cover, respectively, in weedy checks. At Greenley in 2002, giant foxtail accounted for 64 and 65% of BR and IR total weed cover, respectively, in weedy checks. Although IR and BR total weed cover differed among site-years, control of all weeds, giant foxtail, and common waterhemp were rated as 0% in the weedy checks, by convention.

Relative Corn Yield vs. Rated Total Weed Control. Persistent, broad-spectrum, soil residual herbicides were applied at a range of IR and BR rates before corn and weed emergence (Donald et al. 2004a) (Table 1). Even at the lowest IR and BR rates (0.25×), weed emergence was reduced and delayed until after corn emergence. Unlike interference research with single-weed populations that are not treated with herbicide, the regression equations in this study apply to mixed weed populations, largely giant foxtail and common waterhemp, which were stunted and damaged by soil residual herbicides applied at planting.

For mixed weed populations that were treated with herbicide, relative corn yield was not a sigmoidal function of rated weed control (Figure 3). On the basis of interference research with untreated plants, sigmoidal regression equations were expected (Lindquist and Knezevic 2001). The null hypothesis verified that relative yield would be greater than 0% at a rated weed control of 0%, but the null hypothesis that relative corn yield would increase linearly as rated weed control increased was only partially verified (Figure 2e). For observer A, equations were consistent with the null hypothesis for all 3 site-yr and for the average (Figure 3). But for observer B, nonlinear equations best described data variability in 2 of 3 site-yr and for the average. Linear equations were indistinguishable for the two observers in only 1 of 3 site-yr (Greenley in 2001). Although two experienced observers, other than the author, separately rated total weed control in plots that were treated alike, the equations relating relative corn yield to rated total weed control differed between observers for 2 of 3 site-yr (Figure 3). Differences between observers would likely be greater for inexperienced observers.

In weed management research, differences among treatments often can be discerned in rated weed control, even when the yields of these treatments are statistically indistinguishable. This common observation was verified in that the range in rated weed control (i.e., 100%) exceeded the range in relative corn yield (i.e., about 50 to 60%) (Figure 3). Thus, two observers could discern finer relative differences in rated weed control than in relative corn yield.

Differences Between Observers in Rated Total Weed Control. When rated weed control by observer A was regressed against that by observer B for the same treated plots, the data were expected to fall along a 1:1 diagonal line from 0 to 100% (Figure 5, top panel). The data were never fully consistent with this expectation and partially supported this expectation (i.e., linearity) in only 1 of 3 site-yr (i.e., Greenley in 2001) (Figure 5). In two other site-years (i.e., Bradford in 2001 and Greenley in 2002), observer A was less able to distinguish rated weed control at values below about 50% than was observer B. Below about 50% for observer B, the slope of the regression line went to 0%. Differences between observers in rated weed control were more evident in Figure 5 than in Figure 3.

In other agricultural sciences, rating error was not uniform across the entire rating scale, and observers rated more accurately and best distinguished small differences among treatments at low and high rating values (Muir and McCune 1987). In 2 of 3 site-yr, the rating was least

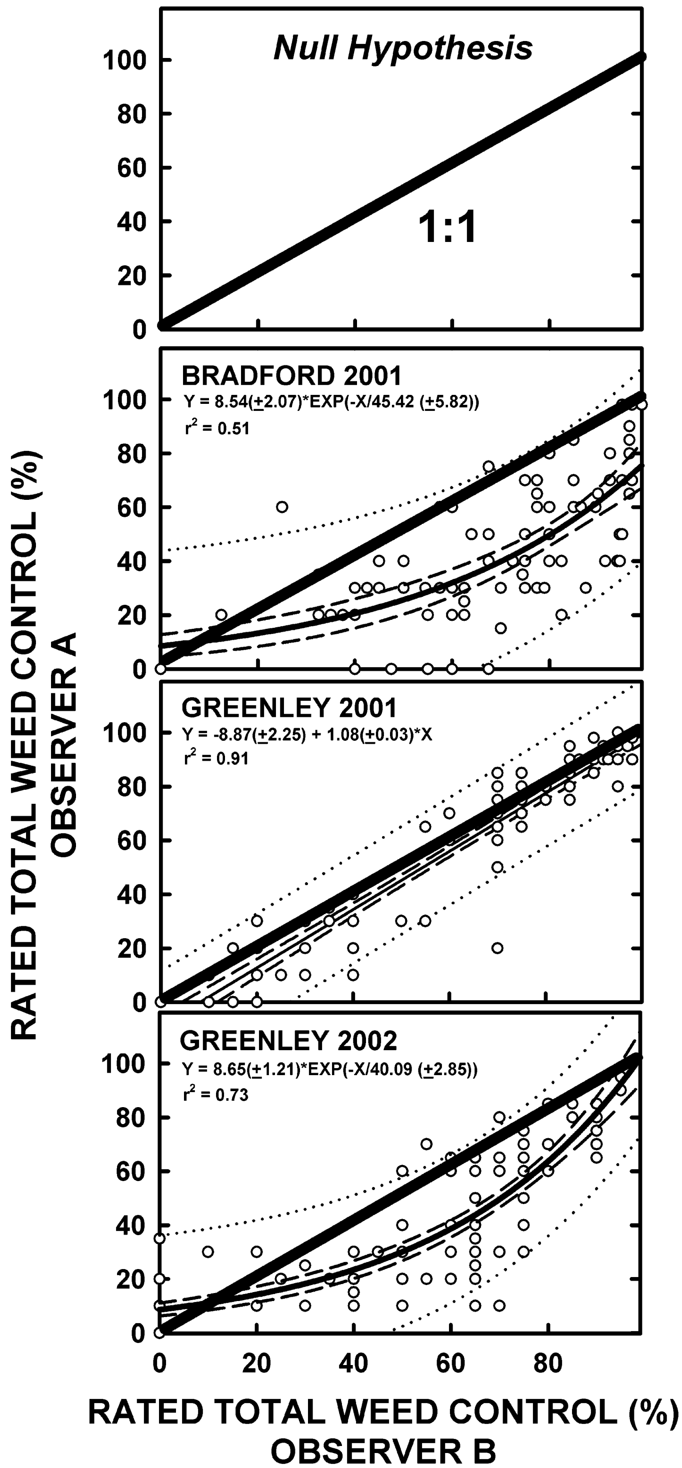

Figure 5. Linear and nonlinear regression equations related rated total weed control (%) by observer A to that of observer B for 3 site-yr. Equation data values (open circles), regression lines (——), 95% confidence interval (– – –), and 95% prediction intervals (····) are presented with regression equations, equation coefficients (± standard errors), and coefficients of determination (r^2). The null hypothesis that observers rated total weed control of the same treated plots alike is presented in the top panel (i.e., data would have fallen along the thick diagonal 1:1 line running from 0 to 100%).

Table 2. Best equations for rated total weed control (Y) as a function of between-row (BR) cover, in-row (IR) cover, and/or height.[a]

Observer	Site-year	Equation	Adjusted r^2
1	Bradford 2001	Y (%) = 105.425 (±4.476) − 0.475 (±0.114)·BR Cover (%) − 0.188 (±0.095)·IR height (cm)	0.47
	Greenley 2001	Y (%) = 94.250 (±2.260) − 0.600 (±0.064)·BR Cover (%) − 0.340 (±0.048)·IR height (cm)	0.80
	Greenley 2002	Y (%) = 90.752 (±2.623) − 0.625 (±0.090)·BR Cover (%) − 0.108 (±0.049)·IR height (cm) − 0.428 (±0.114)·IR Cover (%)	0.78
2	Bradford 2001	Y (%) = 87.766 (±3.224) − 0.860 (±0.058)·BR Cover (%)	0.72
	Greenley 2001	Y (%) = 94.338 (±2.364) − 0.652 (±0.067)·BR Cover (%) − 0.267 (±0.082)·IR Height (%) − 0.307 (±0.154)·IR Cover (%)	0.84
	Greenley 2002	Y (%) = 73.724 (±3.351) − 0.471 (±0.115)·BR Cover (%) − 0.145 (±0.063)·IR Height (%) − 0.354 (±0.146)·IR Cover (%)	0.61

[a] Coefficients (± standard error) are presented.

consistent between observers in the middle 50% of the rating scale (Figure 5). The observers rated similarly near weed control ratings of 0% in all 3 yr, and near 100% in 2 of 3 site-yr.

The current research results partially verify the expectation that relative corn yield is positively linearly related to rated total weed control (Figure 2e), at least for some observers (Figure 3). The expectation that experienced observers would rate total weed control alike was not confirmed across site-years (Figures 3 and 5). Differences between inexperienced observers would be expected to be greater, but this expectation awaits further research. In 3 site-yr, rated weed control also accounted for 25 and 25%, 38 and 36%, and 58 and 57% of data variability in relative corn yield for 2 observers (A and B), respectively. Regression equations likely accounted for little data variation, as reflected in low r^2s because of the study of mixed weed populations, variation in response to different herbicide treatments, and differences in weather and management across site-years on corn yield and competitiveness (Figure 1, scenario 2) (Donald et al. 2004a).

In other agricultural disciplines, differences among observers (i.e., reliability) and differences over time for one observer (i.e., repeatability) also reduced the accuracy and precision of rating (Hebert 1982; Horsfall and Cowling 1978; Horst et al. 1984; James 1971; James and Teng 1979; Nilsson 1995a, 1995b; Nutter and Schultz 1995; Nutter et al. 1993). However, reported differences between observers for rated weed control (Figures 3 and 5) greatly exceeded those reported for rating disease or insect damage (see literature review section). Differences among scientific disciplines in the size of rating error might be due to differences in rating methodology, such as the use or absence of standard reference scales.

Because herbicides are uniformly applied to entire plots in most herbicide screening trials, some observers might assume that weeds respond uniformly across plots without regard to row position. However, recent research demonstrates that competitive crops, such as field corn, can suppress weed growth more in rows than between rows (Donald et al. 2004b). Following application of different IR and BR PRE herbicide rates shortly after planting, differences in IR and BR weed cover became most obvious late in the growing season before harvest. When rating weed control, the human eye can be drawn to greater weed growth between rows than in rows, and this might influence judgment when rating weed control. Differences between weed species in size or height could also bias the weed control rating for specific weeds (i.e., greater between than in rows), but this issue was not addressed in this research. If such visual biasing is occurring, then rated total weed control should be more closely related to BR weed growth (i.e., height, cover, or both) than IR weed growth for all 3 site-yr for both observers. Indeed, this response occurred (Table 2). Forward, backward, and stepwise regression procedures were used to examine the relative contribution of BR and IR total weed height and cover terms to equations accounting for data variability in rated weed control. These selection procedures for regression equations included BR total weed cover in equations for all 3 site-yr and for both observers, whereas IR height was included in five of six equations (Table 2). Equations including BR cover and both IR height and cover best accounted for rated total weed control in half of all equations. More importantly, coefficient values for the BR cover term exceeded coefficient values for either IR cover or height terms or both. Consequently, these equations were consistent with the expectation that the human eye was drawn to the BR zone when rating weed control.

This research verifies and expands on several problems with subjective, vision-based observations requiring human judgment, such as rated weed control, as quantitative scientific measurements. Differences between observers in rated total weed control (Figures 3 and 5) and indirect evidence of visual bias (Table 2) are troubling, as is the variability of equations relating relative crop yield to rated total weed control (Figures 3 and 5).

Although rated weed control has been widely used by weed scientists for over 50 yr, its validity as a quantitative scientific measurement has been largely unquestioned and unexamined. Additional research is warranted. The usefulness of scientific measurements depends on whether they provide data, information, and, eventually, knowledge that answer users' questions and meet their needs. As noted in the introduction, rated weed control and yield have been used to screen the relative efficacy of different weed control treatments, optimize treatment performance in many agronomic crops and environments, and pick the best weed management treatments or systems for recommendation to farmers. With the use of rated weed control and yield, successful, practical, commercial weed management systems have been developed for these purposes.

As the research questions that weed scientists raise change and the research scale increases to the field scale, is rated weed control still a useful method for evaluating weed management? More importantly, does the research technique of rating weed control limit the type of questions that weed scientists raise? Does rated weed control have additional utility for nonscientific purposes, such as picking a ''best'' treatment in decision analysis? In business, economics, and industry, decision analysis and multiple criteria analysis are used to choose the best management option, approach, or course of action among several conflicting alternative choices (Clement 1996; Romero and Rehman 2003). Perhaps visual weed control rating, yield, yield variation, input costs, net returns in partial budget analysis, etc. could be combined with many other criteria from models (e.g., drift hazard, health hazards to applicators, water contamination by herbicides, etc.) in formal, nonstatistical decision analysis before making management recommendations to farmers. Such an inclusive approach could be more appropriate for weighing the pros and cons of different treatments than simply treating weed control rating as a scientific measurement with the use of statistical analysis.

ACKNOWLEDGMENTS

The author thanks Aaron Beshears and Jim Wait for their assistance and John Sadler for his criticism of an earlier version of this manuscript.

LITERATURE CITED

Bedmar, F., P. Manetti, and G. Monterubbianesi. 1999. Determination of the critical period of weed control in corn using a thermal basis. Pesq. Agropec. Bras. Brasilia 34:187–193.

Clement, R. T. 1996. Making Hard Decisions: An Introduction to Decision Analysis. 2nd ed. New York: Duxbury Press. 664 p.

Darwent, A. L., D. Cole, and N. Malik. 1997. Imazethapyr, alone or with other herbicides for weed control during alfalfa (*Medicago sativa*) establishment. Weed Technol. 11:346–353.

Donald, W. W., D. Archer, W. G. Johnson, and K. Nelson. 2004a. Zone herbicide application controls annual weeds and reduces residual herbicide use in corn. Weed Sci. 52:821–833.

Donald, W. W., W. G. Johnson, and K. A. Nelson. 2004b. In-row and between-row interference by corn modifies annual weed control by preemergence residual herbicides. Weed Technol. 18:497–504.

Frans, R., R. Talbert, D. Marx, and H. Crowley. 1986. Experimental design and techniques for measuring and analyzing plant response to weed control practices. *In* N. D. Camper, ed. Research Methods in Weed Science. 3rd ed. Champaign, IL: Southern Weed Science Society. Pp. 29–46.

Hall, M. R., C. J. Swanton, and G. W. Anderson. 1992. The critical period of weed control in grain corn (*Zea mays*). Weed Sci. 40:441–447.

Hamill, A. S., P. B. Marriage, and G. Friesen. 1977. A method for assessing herbicide performance in small plot experiments. Weed Sci. 25:386–389.

Harvey, R. G. 1993. A simple technique for predicting future weed problems. Madison, WI: University of Wisconsin Extension Publications A3595 (1-10-93-10M-S). 6 p.

Hebert, T. T. 1982. The rationale for the Horsfall–Barratt plant disease assessment scale. Phytopathology 72:1269.

Horsfall, J. G. and E. B. Cowling. 1978. Pathometry: the measurement of plant disease. *In* J. G. Horsfall and E. B. Cowling, eds. Plant Disease. An Advanced Treatise. Volume 2. New York: Academic Press. Pp. 119–136.

Horst, G. L., M. C. Engelke, and W. Meyers. 1984. Assessment of visual evaluation techniques. Agron. J. 76:619–622.

Hoshmand, A. R. 1994. Experimental Research Design and Analysis. A Practical Approach for Agricultural and Natural Sciences. Boca Raton, FL: CRC Press. Pp. 59–170.

James, W. C. 1971. An illustrated series of assessment keys for plant disease. Their preparation and usage. Can. Plant Dis. 51:39–65.

James, W. C. and P. S. Teng. 1979. The quantification of production constraints associated with plant diseases. Appl. Biol. 4:201–267.

Kennedy, K. A. and P. A. Addison. 1987. Some considerations for the use of visual estimates of plant cover in biomonitoring. J. Ecol. 75:151–157.

Lindquist, J. L. and S. Z. Knezevic. 2001. Quantifying crop yield response to weed populations: applications and limitations. *In* R.K.D. Peterson and L. G. Higley, eds. Biotic Stress and Yield Loss. Chapter 12. Boca Raton, FL: CRC Press. Pp. 205–232.

Muir, P. S. and B. McCune. 1987. Index construction for foliar symptoms of air pollution injury. Plant Dis. 71:558–565.

Myers, R. H. and D. C. Montgomery. 2002. Response Surface Methodology. Process and Product Optimization Using Designed Experiments. 2nd ed. New York: John Wiley & Sons. Pp. 17–155.

Nilsson, H. E. 1995a. Remote sensing and image analysis in plant pathology. Annu. Rev. Phytopathol. 15:489–527.

Nilsson, H. E. 1995b. Remote sensing and image analysis in plant pathology. Can. J. Plant Pathol. 17:154–166.

Nutter, F. W., Jr., and P. M. Schultz. 1995. Improving the accuracy and precision of disease assessments: selection of methods and use of computer-aided training programs. Can. J. Plant Pathol. 17:174–184.

Nutter, F. W., Jr., M. L. Gleason, J. H. Jenco, and N. C. Christians. 1993. Assessing the accuracy, intra-rater repeatability, and inter-rater reliability of disease assessment systems. Phytopathology 83:806–812.

Romero, C. and T. Rehman. 2003. Multiple Criteria Analysis for Agricultural Decision. New York: Elsevier. 186 p.

Sokal, R. R. and F. J. Rohlf. 1981. Biometry. 2nd ed. New York: W. H. Freeman & Co. P. 13.

Spomer, L. and M.A.L. Smith. 1988. Image analysis for biological research: camera influence on measurement accuracy. Intelligent Instrum. Comput. July/Aug:201–216.

SPSS, Inc. 2001. SPSS User's Guide. Version 11. 233 South Wacker Dr. 11th Floor, Chicago, IL 60606-6307.