SPATIAL STATISTICAL ANALYSIS FOR THE AREA-OF-INFLUENCE EXPERIMENTS

Bahman Shafii

William J. Price

Don W. Morishita

Recommended Citation

SPATIAL STATISTICAL ANALYSIS
FOR THE AREA-OF-INFLUENCE EXPERIMENTS

Bahman Shafii, William J. Price, Statistical Programs,
Don W. Morishita, Division of Plant Sciences
College of Agriculture
University of Idaho
Moscow, ID 83843

ABSTRACT

The area-of-influence (AOI) approach to quantifying crop/weed competition involves measuring the effect of individual weed plants on crop growth and yield at specified distances away from the weed plant. AOI experiments are often analyzed using classical statistical techniques based on the assumption that successive observations on crop response are independent in spite of their distribution in space. However, as the distance varies along the row, the competitive ability will vary spatially so that observations located nearby are expected to be more alike than those separated by large distances. Analyses based on spatial dependencies will therefore provide a more comprehensive understanding of factors influencing crop yield reductions. A spatial statistical approach for analyzing AOI experiments is presented and applications are demonstrated using data from a field experiment in South Central Idaho designed to determine the interference of three broadleaf weed species in sugarbeets.

Key Words: competition, spatial dependence, modeling, variogram, sugarbeet.

I. INTRODUCTION

Estimation of weed-crop competition relationships is an important consideration in weed science research. Suggested methods for quantifying competition effects are based upon i) weed densities (population), and ii) area (distance) weeds influence. While the former approach attempts to identify the minimum density (economic threshold) to justify weed control expenditures, the latter is designed to determine the distance within which an individual weed plant can affect crop growth (Aldrich, 1987). The result of such research studies can provide the basis for developing guidelines in herbicide applications and other methods of weed control.

The relationship between crop response and weed population may be characterized by $y = f(density)$, where y represents crop yield reductions of a species as a function of increasing weed densities. Many useful models of yield loss-weed density relationship have been suggested including linear and nonlinear specifications [e.g., Harper (1961), Dew (1971), Zimdahl (1980), Cousens (1985)].

Models for area-of-influence (AOI) experiments may be represented by $y = f(distance)$, where the measured crop response (yield, biomass) of a species is related to distances away from the individual weed plant [e.g. Weatherspoon and Schweizer (1971), Shurtleff and Coble (1985), Aldrich (1987), Henry and Bauman (1989)]. Such models may also be used to estimate the effect of environmental or agronomic

factors on AOI of weed species.

There are two distinct statistical approaches to analyzing data from AOI experiments: classical and spatial. The classical approach involves modeling the crop response assuming independent and identically distributed observations. Specifically, it is assumed that the individual observations on crop plants taken along the row are independent (i.e. that data form a random sample). Hence, standard statistical techniques are applied to build a model and estimate its parameters. Previous researchers have used various statistical methods including analysis of variance [Mercer, et. al. (1990)], regression [Henry and Bauman (1989)], and multivariate techniques [Jordan (1989)] to report the results of AOI experiments.

Independence is a convenient statistical assumption which makes much of the subsequent classical inferences tractable. However, it is more likely that in AOI experiments successive sampling points along the row are spatially correlated. As the distance varies between points along the row, the spatial distribution of competitive effects will vary so that adjacent crop plants are expected to be more alike, with regard to competitive ability, than those far apart. The spatial statistical analysis will therefore involve quantifying crop response after spatial dependence has been modeled. This will in turn exploit the correlation structure to obtain more precise estimators of the specified model parameters.

The purpose of this study is to introduce a spatial statistical approach for analyzing AOI experiments. Empirical applications are demonstrated with reference to a field experiment in South Central Idaho.

II. METHODS
Modeling the Spatial Dependence

Detrending/Transformation

Exploratory data analytic techniques (i.e. stem-leaf-plot, box plot, pocket plot) are valuable tools in understanding and exploiting the spatial relationship among data points. Aside from their ability to detect trends and outliers, these techniques can provide a basis for decisions concerning possible transformation of data. A common stochastic model assumed for spatial data (and hence the basis for most subsequent geostatistical inferences) is the Gaussian (normal) process. Thus, it is important to initially examine the distribution of data points and determine the potential for resistant (mean or median-based removal) detrending as well as variance stabilizing transformations.

Variogram Estimation and Fitting

Let $\{x_1, \ldots, x_n\}$ represent the set of coordinate positions (spatial locations) in the field (index set) D with corresponding observed data $\{z(x_1), \ldots, z(x_n)\}$. Then the regionalized variable (Matheron. 1971) $z(x_i)$ can be considered a realization of the set of random variables $Z(X)$ for all x_i in D,

$$\text{i.e.:} \ \{Z(X) : X \in D\}, D \subset \mathbb{R}^d. \tag{1}$$

The classical variogram estimator is defined as

$$2\gamma(h) \doteq E\{[z(x_i) - z(x_i+h)]^2\} \tag{2}$$

which is estimated from the sample by

$$2\hat{\gamma}(h) = 1/N(h) \, \Sigma[z(x_i) - z(x_i+h)]^2, \ h \in \mathbb{R}^d. \tag{3}$$

The quantity $2\hat{\gamma}(.)$ [conceptually, mean-squared difference or variance] is called an estimated variogram (Matheron, 1962), and $\hat{\gamma}(.)$ is called an estimate of the semivariogram; $N(h)$ is the number of distinct pairs of observations $[z(x_i)-z(x_i+h)]$ separated by the distance/lag vector h. A robust variogram estimator due to Cressie and Hawkins (1980) is given by

$$2\tilde{\gamma}(h) = \frac{\{1/N(h) \, \Sigma |z(x_i) - z(x_i+h)|^{1/2}\}^4}{(.457 + .494/|N(h)|)} \tag{4}$$

Both variogram estimators, $2\hat{\gamma}(h)$, and $2\tilde{\gamma}(h)$, were considered for subsequent analyses.

Fitting a theoretical model to the estimated variogram is an important aspect in quantifying the spatial dependence. While a number of parametric variogram models are available [Journal and Huijbregts (1978)], only three isotropic [i.e., when $2\gamma(h)$ depends only on the magnitude and not the direction of vector h] models were considered in this study. These were:

i) Linear

$$2\gamma(h) = \begin{cases} c_0 + \beta h, & 0 < h \le r_L \\ \\ c_0 + c_L, & h \ge r_L \end{cases} \tag{5}$$

Where $c_0 \ge 0$ is the nugget effect, c_0+c_L is the sill, and r_L is the range of the variogram;

ii) Spherical

$$2\gamma(h) = \begin{cases} 0, & h=0 \\ \\ c_0 + c_s[(3/2)h/r_s - (1/2)(h/r_s)^3], & 0<h\le r_s \\ \\ c_0 + c_s, & h\ge r_s \end{cases} \tag{6}$$

where $c_0 \ge 0$, $c_s \ge 0$, $r_s \ge 0$; and

iii. Wave (hole-effect) (Cressie, 1991)

$$2\gamma(h) = \begin{cases} 0, & h=0 \\ \\ c_0 + c_w[1 - r_w\sin(h/r_w)/h], & h\ne 0 \end{cases} \tag{7}$$

where $c_0 \ge 0$, $c_w \ge 0$, $r_w \ge 0$.

Applied Statistics in Agriculture

Variogram models were fitted to the specified estimators using linear and nonlinear least square methods.

Response Model Estimation

The basic linear model for the spatial data may be written as

$$Z = X\beta + \delta \qquad (8)$$

Where

$Z \equiv [z(x_1), \ldots, z(x_n)]'$,
$X \equiv n \times p$ matrix of explanatory variables,
$\beta \equiv (\beta_1, \ldots, \beta_p)'$,
$\delta \equiv [\delta(x_1), \ldots, \delta(x_n)]'$.

Under classical (nonspatial) assumption of independent and identically distributed errors, $\delta(.)$, the ordinary least squares estimator of β is given by

$$\hat{\beta}_{OLS} = (X'X)^{-1}X'Z$$

which is equivalent to the maximum likelihood estimator if $\delta(.)$ is a Gaussian process. However, when the error process exhibits spatial correlation, the generalized least square (GLS) estimator given by

$$\hat{\beta}_{GLS} = (X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}Z \qquad (9)$$

provides the appropriate estimators which is often more efficient (Searle, 1971). In the spatial context, $\Sigma = \text{Var}(\delta)$ is an $n \times n$ symmetric, nonnegative definite matrix whose elements are determined according to the underlying spatial structure of the scientific problem (see the empirical section).

Validation and inference

AOI experiments are customarily analyzed via regression of crop response on distance away from the weed plant (e.g. Weatherspoon and Schweizer, 1971). The general spatial model (8) implies different regression models (possessing different variance-covariance structures) depending on the choice of Σ. This will in turn affect the resulting variance estimate for the GLS estimator (9) given by

$$\text{Var}(\hat{\beta}_{GLS}) = (X'\Sigma^{-1}X)^{-1} \qquad (10)$$

Let W be an $n \times n$ positive definite weight matrix. Then it follows that $W = \Sigma$, $W = \text{diag}(\Sigma)$, and $W = I$ represent respectively, the full spatial model, the heteroskedastic model, and the classical model. Note that the classical model is a special case of the spatial model for which $\Sigma = \sigma^2 I$.

For a given regression model, the GLS residuals defined as

$$\epsilon_{GLS} \equiv Z - X\hat{\beta}_{GLS} = [W - (X'W^{-1}X)^{-1}X']W^{-1}Z \qquad (11)$$

should then be used for the purpose of model validation. Once a suitable response model has been identified, statistical inferences concerning different aspects of AOI experiments may be carried out. These usually include determination of the extent of AOI for a particular weed species or testing the equality of AOI's (in terms of size and shape of competitive effects) for two or more weed species.

III. EMPIRICAL RESULTS

Experimental Specifications

Data used in this study are based on field experiments conducted during 1991 and 1992 at the University of Idaho Research and Extension Center located near Kimberly, ID. The crop investigated was sugarbeet (*Beta vulgaris L.*), variety WS-88, which was thinned to obtain a 20 cm spacing between sugarbeet plants (88066 plants/ha). The field was irrigated by sprinkler as needed (approximately weekly) and a total of 80 lb/a preplant nitrogen and 80 lbs/a foliar nitrogen were applied for the 1992 season. Plots were 3.4 m wide (6 rows) by 10 m long. One center row in each plot was used for treatment, where the sugarbeet plant in the center of the row was replaced by one of three weed species: kochia (*Kochia scoparia* (L.) Schrad.), common lambsquarters (*Chenopodium album* L.), and redroot pigweed (*Amaranthus retroflexus L.*). A fourth treatment which retained the center sugarbeet was used as a check. These treatments were arranged as a randomized complete block design with 10 replications. In order to mitigate the effects of shading, the rows were aligned in an east-west direction. Individual plants within the treated row were harvested at 20 cm intervals up to two meters from the center plant. Root weight along with distance and direction from the center plant were noted for each sample. Three harvests at different times during each growing season were taken in the same manner, however only data from the final harvest (October, 1992) are considered in this study.

All subsequent statistical computations were carried out using SAS STAT (1991) or SAS IML (1990).

Variogram Estimation and Modeling

Initial examination of the data indicated that differences between east and west sides of each plot were negligible. The data was, therefore, combined across sides by absolute distance from the center plant. The distributions of the data were unimodal and reasonably symmetric (Figure 1). Plots of mean and median sugarbeet yields versus distances away from each specified weed species are given in Figure 2. No meaningful differences were indicated by either mean or median-based trend removal. In all, the untransformed data were determined to be sufficiently well behaved for variogram estimation.

Both forms of the sample variogram estimators, (3) and (4), were computed for the data. The robust version, (4), showed little advantage over the classical form and therefore, the latter was chosen (Figure 3). Kochia exhibited the strongest spatial effect, while common lambsquarters and redroot pigweed had relatively constant variability within the specified range of the observed variograms. The check showed a low level of constant variability.

The isotropic models given in (5), (6), and (7) were fitted to each sample

variogram using linear and nonlinear least squares. Figure 4 gives the fitted models for each species up to a lag distance of 1.4 m. The last two lags were dropped because the amount of data (spatial information) contributing to these points was limited. Kochia was determined to be best modeled by the Wave (hole-effect) model (7) while common lambsquarters and redroot pigweed were best estimated with the linear model (5). The weed-free (check) treatment would serve as a reference value to which other treatments would be compared and therefore, it was not modeled.

Construction of the Weight Matrix

The experimental design and the existence of several border rows (lack of dependence) between treatments (weed species) indicated the following covariance model for the data:

$$
\text{Var(z)} \equiv \sigma^2 W = \sigma^2 \begin{bmatrix} \Sigma_1 & 0 & 0 \\ 0 & \Sigma_2 & 0 \\ 0 & 0 & \Sigma_3 \end{bmatrix}
$$

Here, the weight matrix, W, is a 300x300 block diagonal matrix where each block, Σ_i, i=1, 2, 3, is 100x100. The respective elements of each block were based on the estimated variogram models as follows:

Kochia:
$$
\sigma^2 \Sigma_1 = \{\text{cov}(z(x_i), z(x_{i+h}))\}
$$
$$
= \{\sigma^2(.188 - \hat{\gamma}_1(\|x_i - x_j\|))\};
$$

C. Lambsquarters:
$$
\sigma^2 \Sigma_2 = \{\sigma^2(.086 - \hat{\gamma}_2(\|x_i - x_j\|))\};
$$

R. Pigweed:
$$
\sigma^2 \Sigma_3 = \{\sigma^2(.110 - \hat{\gamma}_3(\|x_i - x_j\|))\}.
$$

Modeling the Crop Response

Several polynomial and nonlinear response models were considered for each species, but the simple linear model was found to be the most appropriate one for all three species. The fitted equations along with 95% prediction intervals are presented in Figure 5. Initial inspection of Figure 5 may indicate some problems with the linear response model, however a subsequent examination of the residual plots obtained from GLS fitting showed an acceptable pattern, magnitude and distribution for the residuals (Figure 6).

Application of the spatial (GLS) model resulted in little change in parameter estimates compared to the standard (OLS) model. However, the precision of the parameter estimates was influenced as implied by (10). Table 1 lists the regression summaries obtained from both methods of estimation by species. For the spatial models, a reduction in standard error of estimates is realized. This effect was most notable in kochia, which had the strongest spatial effect, and less so in common

lambsquarters and redroot pigweed. In all cases, the spatial model provided a significant reduction in residual mean square, produced more precise parameter estimates, and improved the underlying residual structure. This will in turn influence the biological interpretations of AOI experiments.

Figure 7 demonstrates the difference in the extent of AOI for kochia using both standard and spatial regression models. Here, the extent of AOI is defined as the intersection of the lower limit of a 95% confidence interval for the weed-free (check) mean and the upper limit of a 95% prediction interval of the estimated response function. The prediction interval, which is influenced by the assumed structure of Σ in (10), is wider for the standard regression model and results in a determination of extent of AOI which is smaller than that of the spatial model. This will imply a shorter distance (≈ 22 cm) within which an individual kochia plant can affect sugarbeet yields. Similar discrepancies could result when comparing response functions among species for similar size and shape of AOI.

Equality of the three species' AOI was tested using a joint hypothesis for coinciding lines and was rejected ($p < .0007$). Further examination found both the size (intercepts) and shape (slopes) of the AOI's to be significantly different among species ($p < .0001$). The extent of AOI for each species is given in Figure 8. For kochia and common lambsquarters this was determined to be about 114 cm and 69 cm, respectively. The extent of AOI for redroot pigweed could not be determined, and in fact, this species did not significantly differ from the check.

Three distinct patterns of area-of-influence were evident in this study. Kochia exhibited both a magnitude effect on the response (reduction in sugarbeet yields), as well as a spatial competitive effect, while common lambsquarters showed only the magnitude effect. Redroot pigweed demonstrated neither the magnitude nor the spatial effects. Figure 9 presents these three scenarios as predicted losses (% of control) vs distance. Both kochia and common lambsquarters resulted in losses in sugarbeet yield, with kochia having a greater effect. This was consistent with biological expectations as kochia is known to have a strong competitive ability. However, no significant losses were shown from redroot pigweed. Although further analyses (including other harvests) are required to establish a comprehensive biological interpretation, interference of above weed species in sugarbeets could also be affected by other factors not accounted for in this study. Such factors may include lack of competition for specific resources (i.e. water or nitrogen which could be nonlimiting), weed densities required to observe a competitive effect, or different root systems of the weed species.

ACKNOWLEDGEMENTS

REFERENCES

Aldrich, R. J. 1987. Predicting crop yield reductions from weeds. Weed Tech. 1:199-206.

Cousens, R. 1985. A simple model relating yield loss to weed density. Ann. Appl. Biol. 107:239-252.

Cressie, Noel A. C. 1991. Statistics for spatial data. John Wiley & Sons, Inc. New York.

Cressie, Noel A. C., and D. M. Hawkins. 1980. Robust estimation of the variogram. Journal of the International Assoc. for Mathematical Geology. 12:115-125.

Dew, D. A. 1972. An index of competition for estimating crop loss due to weeds. Can. J. Plant Sci. 52:921-927.

Harper, J. L. 1961. Approaches to the study of plant competition. Symp. Soc. Exptl. Biol. 15:1-39.

Henry, W. T., and T. T. Bauman. 1989. Interference between soybeans and common cocklebur under Indiana field conditions. Weed Sci. 37:753-760.

Jordan, N. 1989. A statistical analysis for area-of-influence experiments. Weed Tech. 3:114-121.

Journal, A. G., and C. J. Huijbregts. 1978. Mining Geostatistics. Academic Press. London.

Matheron, G. 1962. Traite de geostatistique applique, du Bureau de Richerches Geoloiques et Minieres, No. 14, Paris.

Matheron, G. 1971. The theory of regionalized variables and its applications. Cahiers du Centre de Morphologie Mathematique, No. 5. Foutainbleau, France.

Mercer, K. L., J. A. Pawalk, D. S. Murray, L. M. Verhain, M. S. Riffle, and R. W. McNew. 1990. Distance-of-influence of devil's-claw on cotton. Weed Tech. 4:87-91.

SAS Institute Inc. 1990. SAS/IML User's Guide, Version 6, First Edition, Cary, NC.

SAS Institute Inc. 1991. SAS/STAT User's Guide, Version 6, Fourth Edition, Volume 2. Cary, NC.

Searle, S. R. 1971. Linear Models. John Wiley and Sons, Inc. New York.

Shurtleff, J. L., and H. D. Coble. 1985. Interference of certain broadleaf weed species in soybeans. Weed Sci. 33:654-657.

Weatherspoon, D. M., and E. E. Schweizer. 1971. Competition between sugarbeets and five densities of kochia. Weed Sci. 19:125-128.

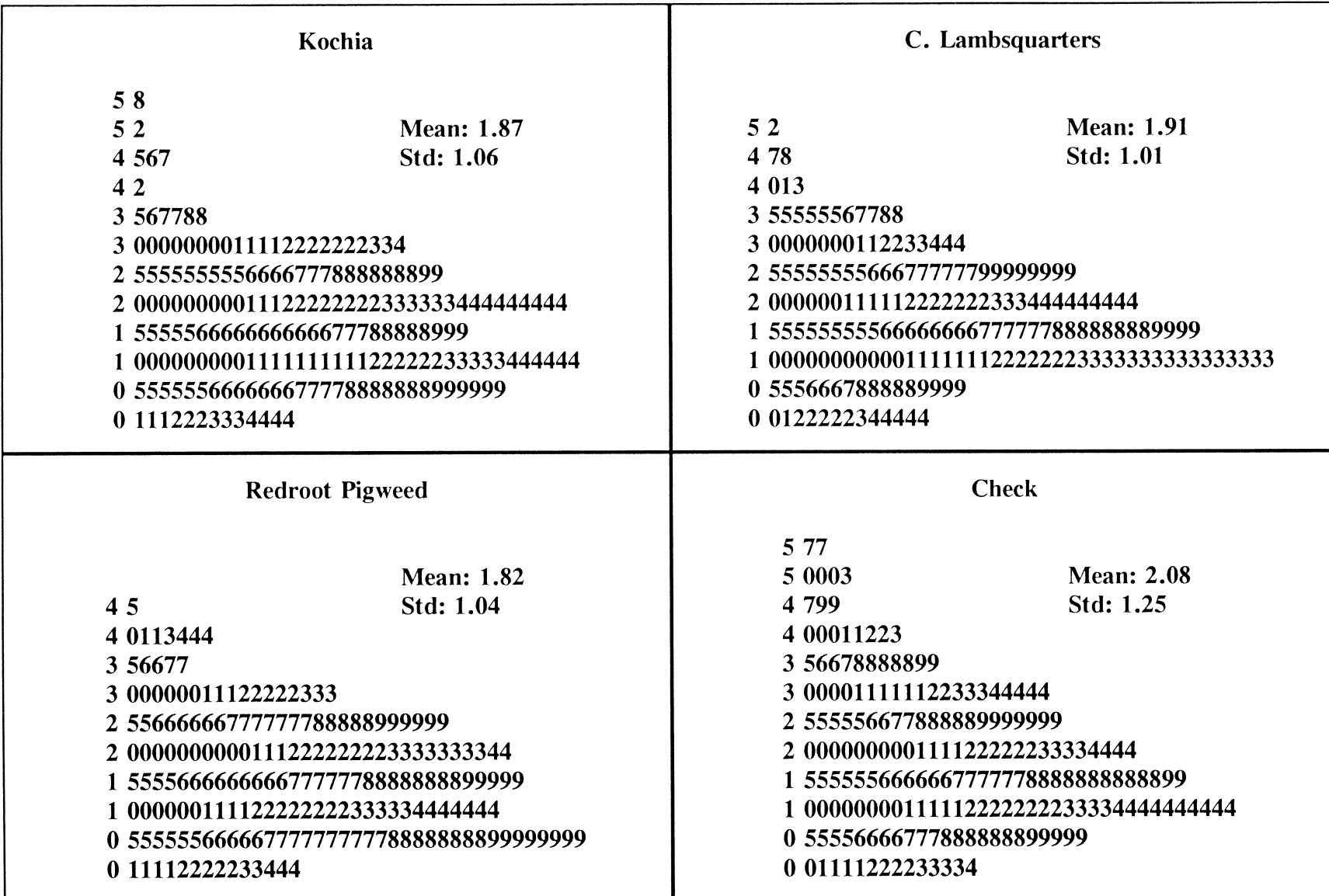Zimdahl, R. L. 1980. Weed-crop competition: a review. International Plant Protection Center, Oregon State Univ. 195p.

Kochia

```
5 8
5 2                          Mean: 1.87
4 567                        Std: 1.06
4 2
3 567788
3 00000000011112222222334
2 555555555566667777888888899
2 000000000011122222222333333444444444
1 5555566666666666777788888999
1 00000000001111111111222222233333344444444
0 555555666666677778888888999999
0 1112223334444
```

C. Lambsquarters

```
5 2                          Mean: 1.91
4 78                         Std: 1.01
4 013
3 55555567788
3 0000000112233444
2 5555555556667777777799999999
2 0000001111122222222333444444444
1 55555555556666666667777778888888889999
1 0000000000011111112222222233333333333333333
0 5556667888889999
0 0122222344444
```

Redroot Pigweed

```
                             Mean: 1.82
4 5                          Std: 1.04
4 0113444
3 56677
3 00000011122222333
2 55666666777777788888899999
2 000000000011122222222333333334
1 5555666666666777777888888899999
1 000000111222222223333334444444
0 555555666677777777777888888899999999
0 11112222233444
```

Check

```
5 77
5 0003                       Mean: 2.08
4 799                        Std: 1.25
4 00011223
3 56678888899
3 000011111112233344444
2 5555566778888899999999
2 00000000001111222222333334444
1 55555566666677777788888888888899
1 0000000011111222222222333344444444444
0 55556666777888888899999
0 01111222233334
```

Figure 1. Stem-leaf-plots with summary statistics for sugarbeet root weights (lbs) from the three weed specie and check treatments.
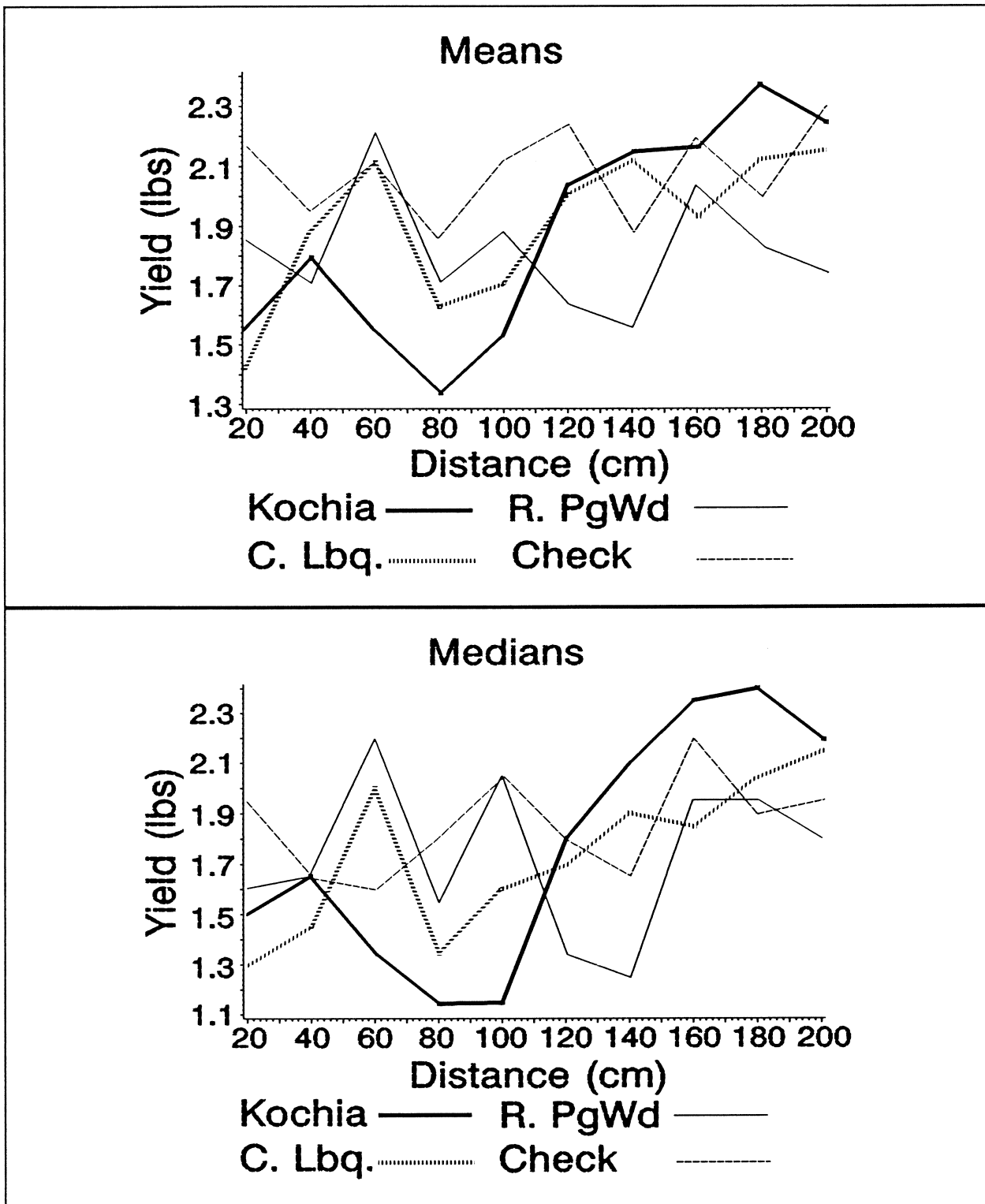
Means

Medians

Figure 2. Mean and median sugarbeet yields as a function of distance from center weed. Means and medians are based on 10 replicates per distance.
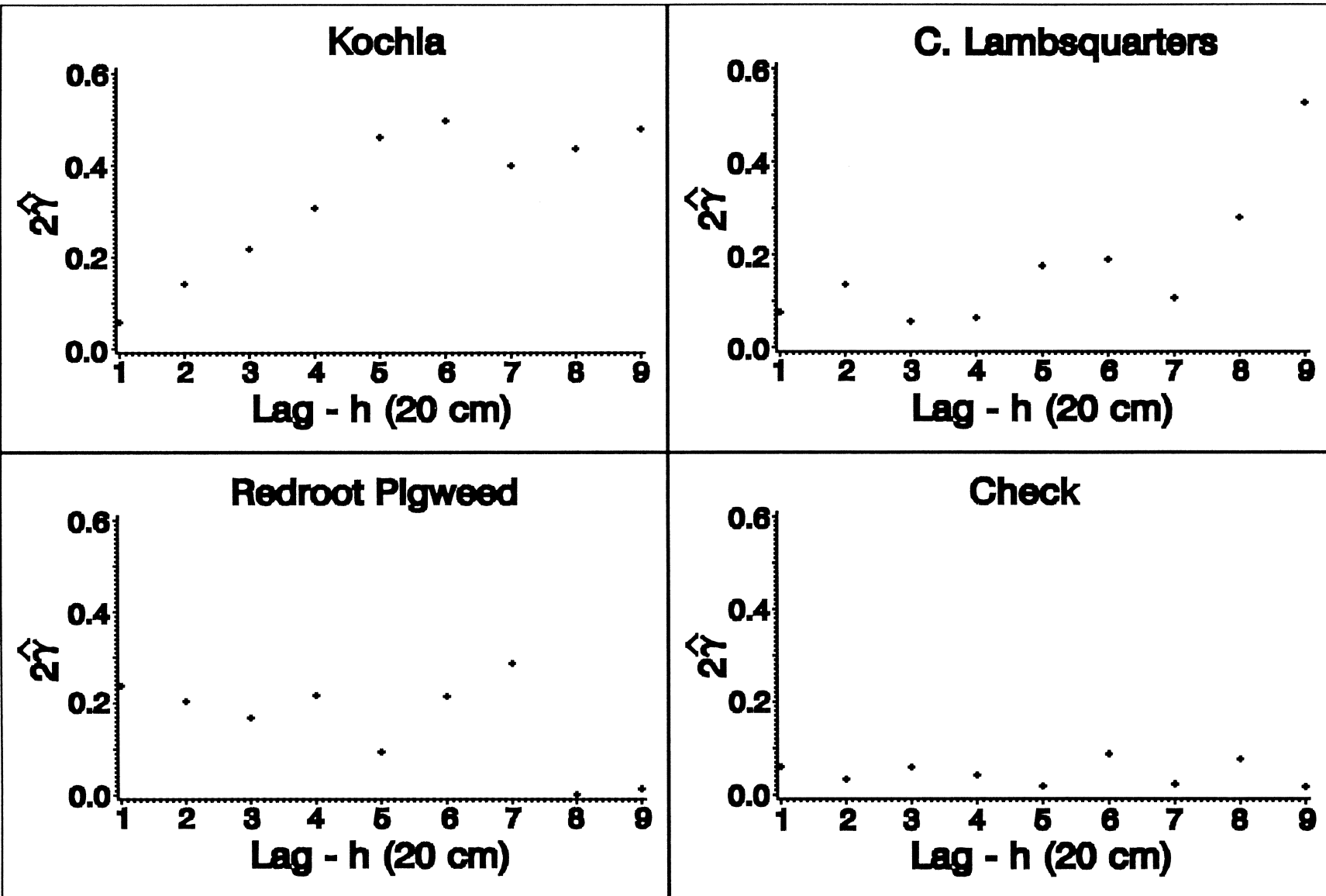
Figure 3. Observed variograms for each treatment computed using the classical variogram formula.
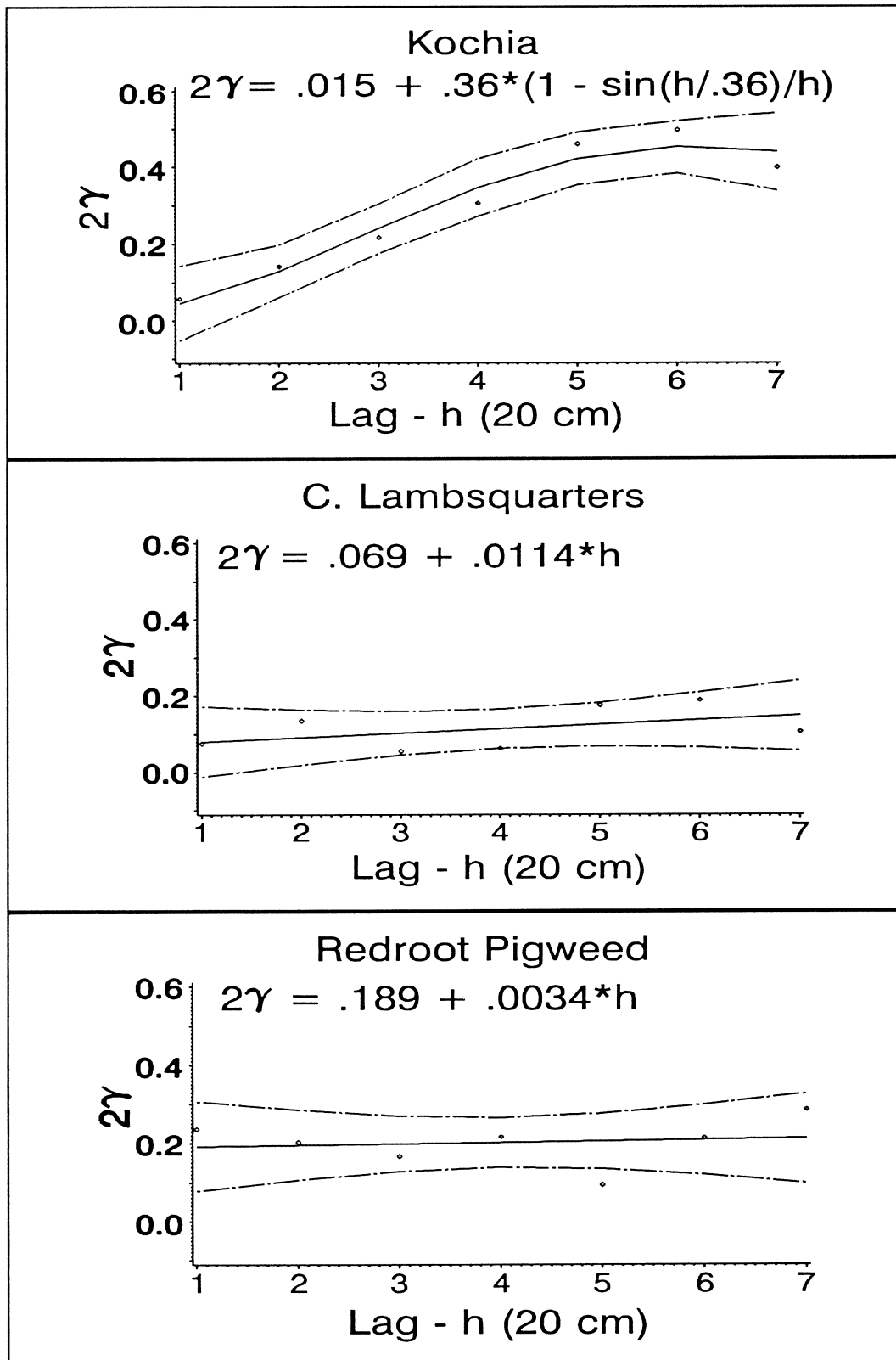
Figure 4. Estimated variogram models along with 95% prediction intervals for kochia, c. lambsquarters and redroot pigweed.
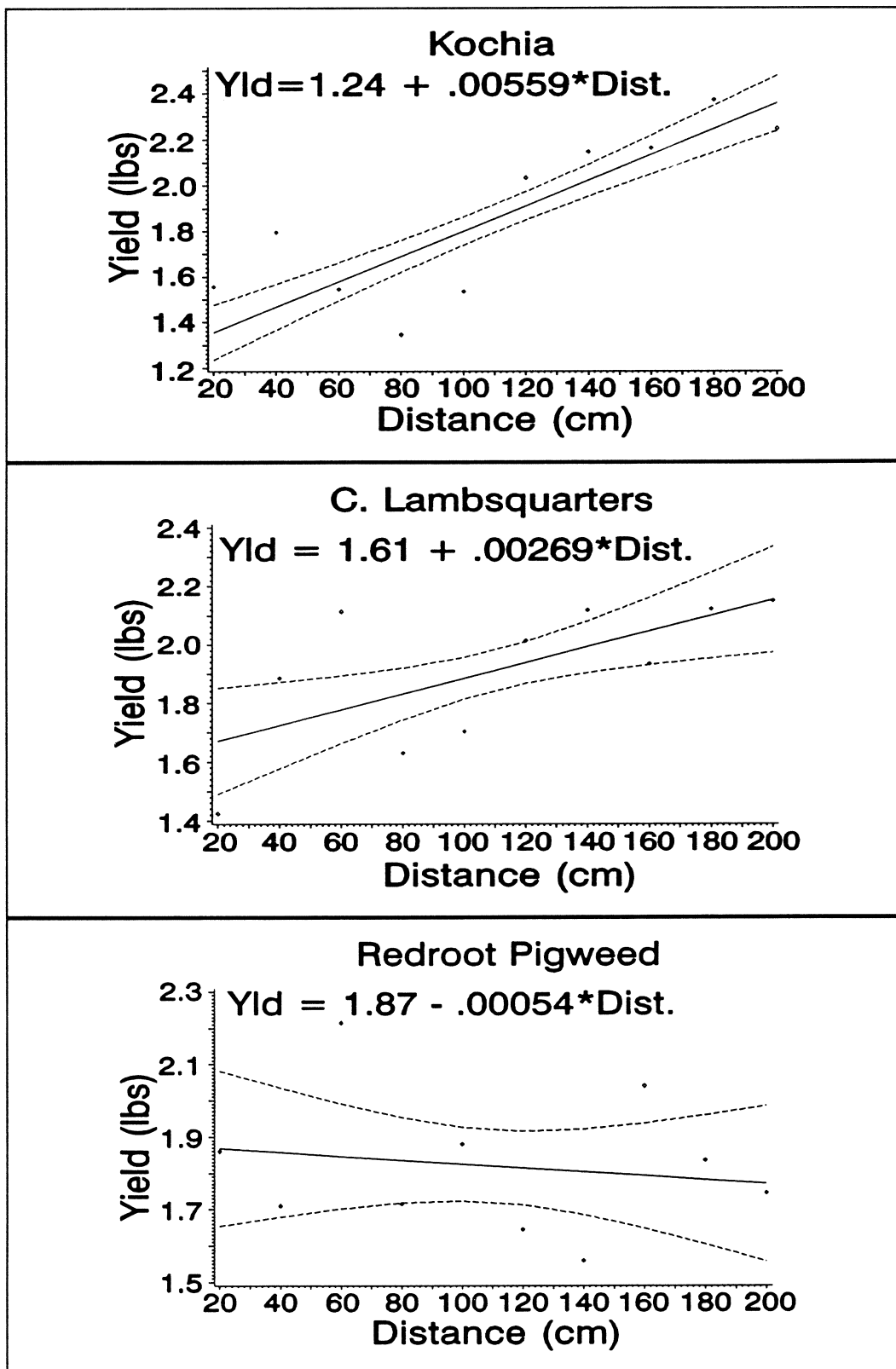
Figure 5. Estimated sugarbeet yield response functions with 95 % prediction intervals for each weed species. Individual data points represent mean of 10 blocks.
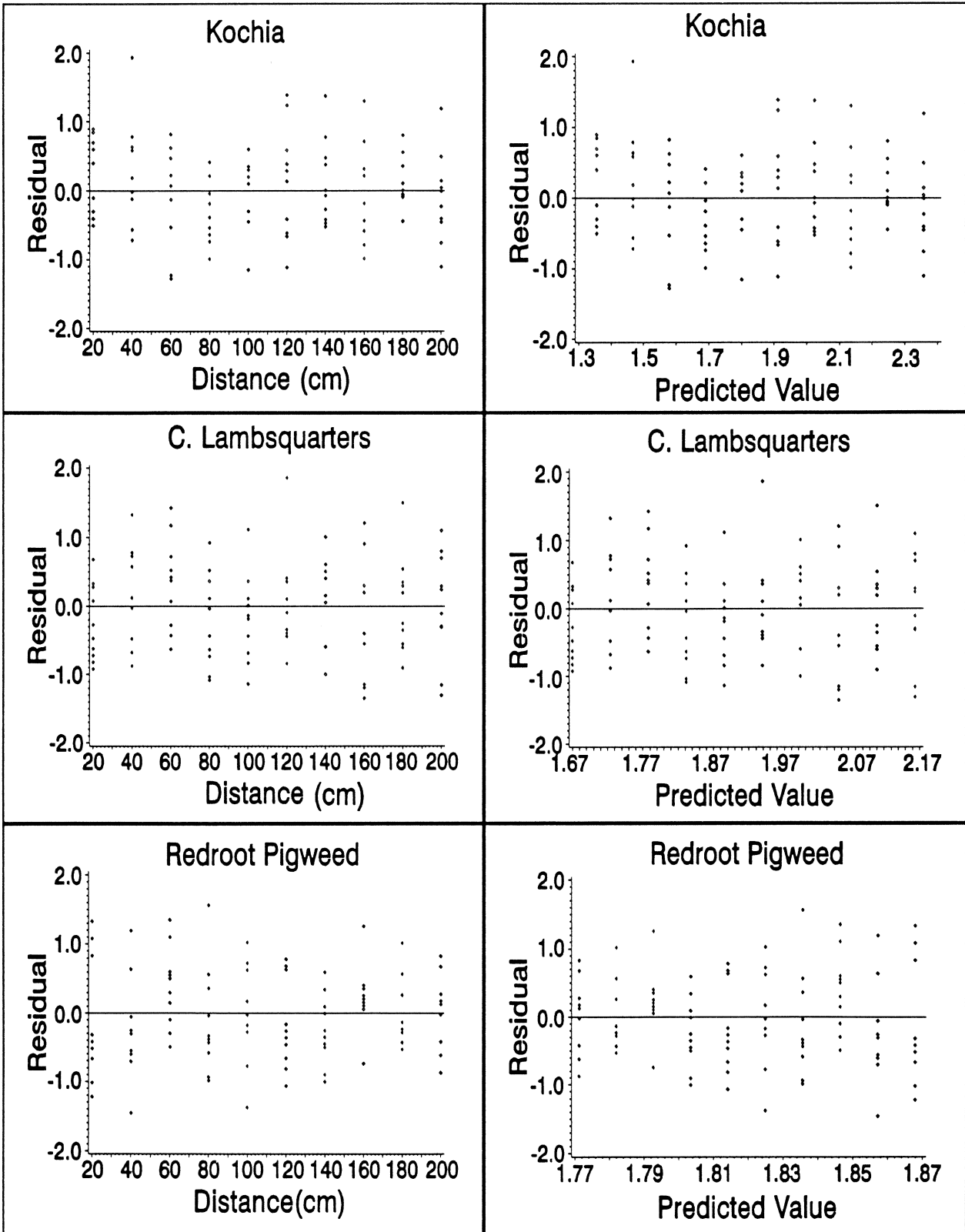
Figure 6. Residual plots from estimated sugarbeet yield response functions for each species. Plots show residuals against the regressor, distance, and predicted values.

TABLE 1. Regression summaries obtained from spatial (GLS) and standard (OLS) fitting of the response model for the three weed species.

Spatial Regression

Kochia:

Parameter	Parameter Estimate	Standard Error	t	Pr > t
Intercept	1.2444	0.07065	17.61	.0001
Slope	0.0056	0.00058	9.65	.0001

RMS: .0675

Common Lambsquarters:

Parameter	Parameter Estimate	Standard Error	t	Pr > t
Intercept	1.6172	0.10875	14.7	.0001
Slope	0.0027	0.00106	2.55	.0062

RMS: .0433

Redroot Pigweed:

Parameter	Parameter Estimate	Standard Error	t	Pr > t
Intercept	1.8786	0.12726	14.76	.0001
Slope	-0.0005	0.00106	-0.47	.3191

RMS: .0480

Standard Regression

Kochia:

Parameter	Parameter Estimate	Standard Error	t	Pr > t
Intercept	1.3320	0.14193	9.39	.0001
Slope	0.0049	0.00114	4.31	.0001

RMS: .4317

Common Lambsquarters:

Parameter	Parameter Estimate	Standard Error	t	Pr > t
Intercept	1.6076	0.15188	10.58	.0001
Slope	0.0027	0.00122	2.25	.0267

RMS: .4943

Redroot Pigweed:

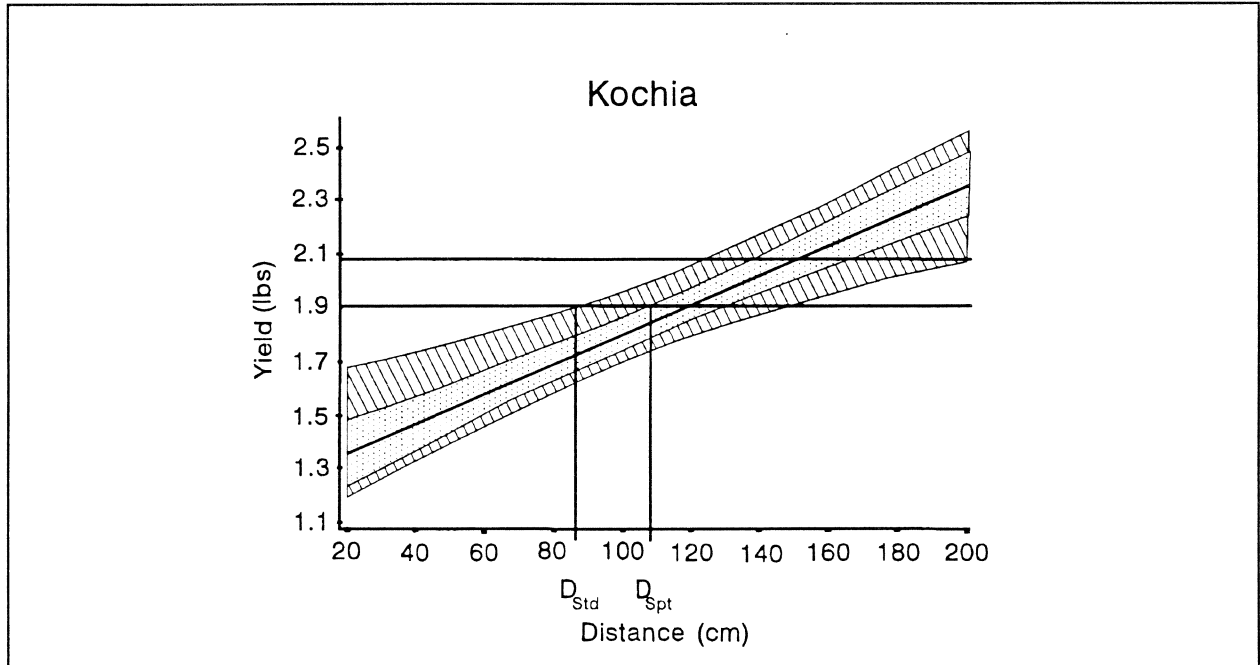| Parameter	Parameter Estimate	Standard Error	t	Pr > t
Intercept	1.8783	0.14411	13.03	.0001
Slope	-0.0005	0.00116	-0.45	.6517

RMS: .4450

Kansas State University

Kochia

Figure 7. Estimated sugarbeet yield response functions and their 95% prediction intervals for kochia under standard (diagonal hatch) and spatial (stiple hatch) regression. Upper and lower horizontal lines represent the check mean and its lower 95% confidence interval. Vertical lines indicate the determined extent of AOI under standard, D_{Std}, and spatial, D_{Spt}, estimation.
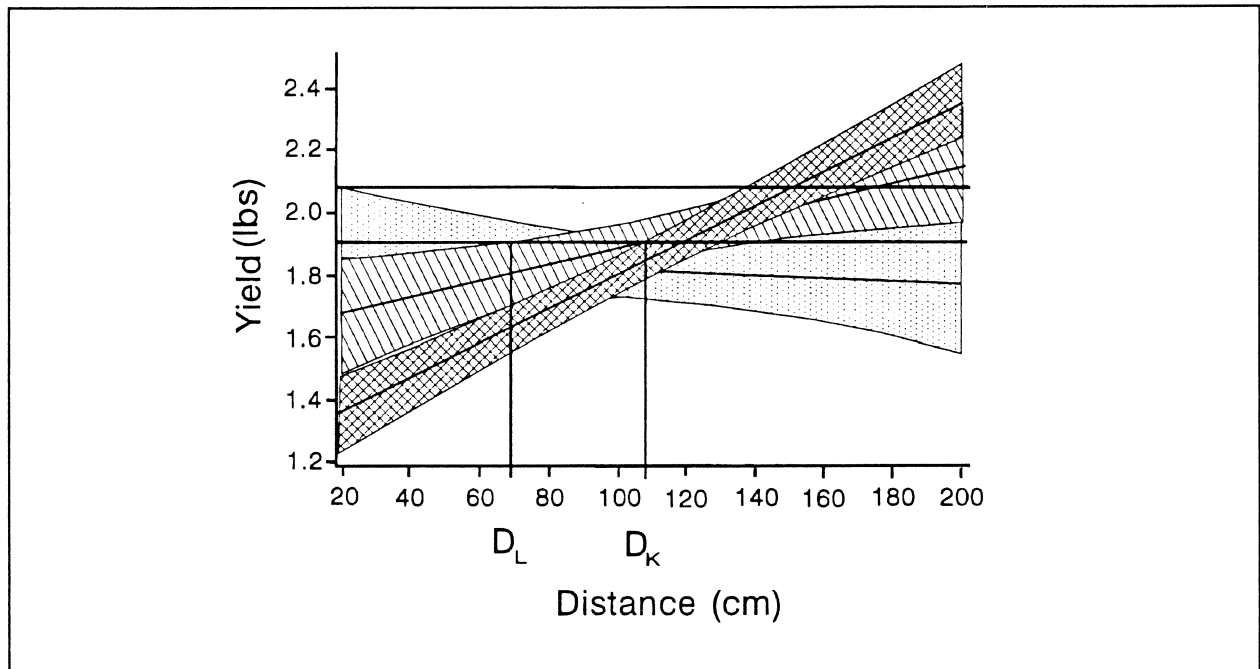

Figure 8. Estimated sugarbeet yield response functions and their 95% prediction intervals for kochia (cross hatched), c. lambsquarters (diagonal hatch), and r. pigweed (stiple hatch). Upper and lower horizontal lines represent the check mean and its lower 95% confidence interval. The extent of AOI are indicated for c. lambsquarters (D_L) and kochia (D_K). R. pigweed had no detectable extent of AOI.
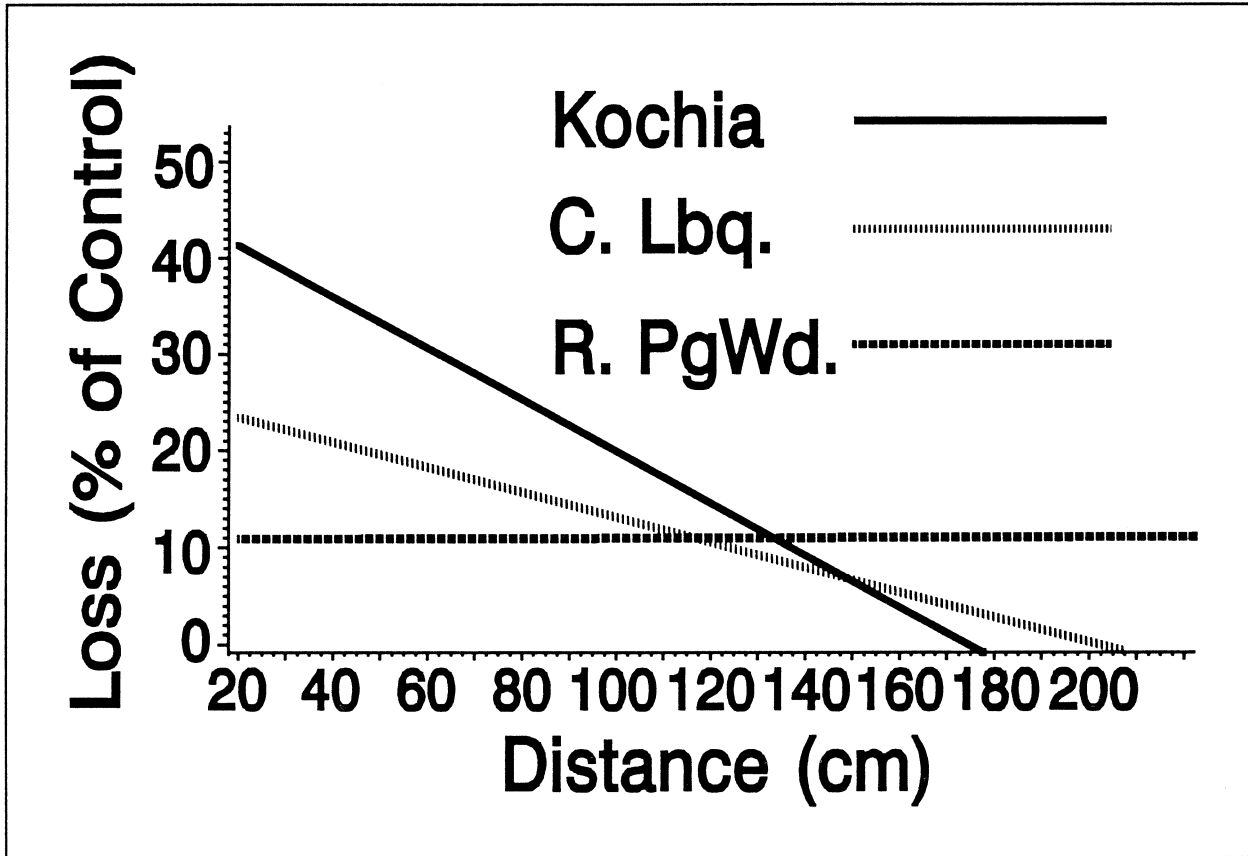
Figure 9. Predicted relative losses (% of control) in sugarbeet yields as a function of distance from one of three weed species.