Weed Science, 51:44–55. 2002

How good is your weed map? A comparison of spatial interpolators

J. Anita Dille
Corresponding author. Department of Agronomy, University of Nebraska, Lincoln, NE 68583. Present address: Department of Agronomy, Kansas State University, Manhattan, KS 66506; dieleman@ksu.edu

Maribeth Milner
Jeremy J. Groeteke
David A. Mortensen
Martin M. Williams II
Department of Agronomy, University of Nebraska, Lincoln, NE 68583

Recent interest in describing the spatial distribution of weeds and studying their association with site properties has increased the use of interpolation to estimate weed seedling density from spatially referenced data. In addition, farmers and consultants adopting elements of site-specific farming practices are using interpolation methods for mapping weed densities as well as soil properties. This study was conducted to compare the performance of four interpolation methods, namely inverse-distance weighting (IDW), ordinary point kriging (OPK), minimum surface curvature (MC), and multiquadric radial basis function (MUL), with respect to their ability to map weed-seedling densities. These methods were evaluated on data from four weed species, velvetleaf, hemp dogbane, common sunflower, and foxtail species, of contrasting biology and infestation levels in corn and soybean production fields in Nebraska. Mean absolute difference (MAD) and root mean square (RMS) between the observed point sample data and the estimated weed seedling density surfaces were used to evaluate the performance of the interpolation methods. Four neighborhood search types were compared within each interpolation method, and Search3 (12 to 16 neighboring locations) generated an interpolated surface with the smallest MAD and RMS indicating the highest precision. IDW with a power coefficient of $p = 4$ gave the smallest MAD and RMS, as did a test with an elliptical search and no anisotropy. The level of precision of all four interpolation methods was very poor for weed species with low infestation levels (> 75% of field weed-free; MAD ranged from 100 to 187% of the observed mean density), whereas precision was improved for weed species with high infestation levels (< 25% of field weed-free; MAD ranged from 45 to 85%). No single interpolator appears to be more precise than another. Implications of this study indicate that grid sample spacing and quadrat size are more important than the specific interpolation method chosen.

Nomenclature: Common sunflower, *Helianthus annuus* L. HELAN; foxtail species, *Setaria* spp. SETSS; hemp dogbane, *Apocynum cannabinum* L. APCCA; velvetleaf, *Abutilon theophrasti* Medicus ABUTH; corn, *Zea mays* L.; soybean, *Glycine max* (L.) Merr.

Key words: Inverse-distance weighting, kriging, minimum curvature, multiquadric, radial basis function, ABUTH, APCCA, HELAN, SETSS.

Weed species occur as aggregated or patchy populations in agricultural fields. Population densities vary greatly across individual fields, making sampling to characterize the infestation level problematic (Cardina et al. 1995; Johnson et al. 1995; Mortensen et al. 1993; Wiles et al. 1992; Wilson and Brain 1991). Research is underway to assess the use and value of maps that describe the spatial distribution of weed densities in modeling weed population dynamics (Wallinga 1995) and to guide the application of weed control practices on individual fields (Gerhards et al. 1997a; Heisel et al. 1999; Williams et al. 1999). Understanding spatial weed dynamics and using site-specific weed control could reduce input costs (Brain and Cousens 1990; Lindquist et al. 1998; Williams et al. 1999), reduce the environmental effect of control practices, and enhance control of patchy weed populations (Dieleman et al. 1999). But the accuracy of the interpolation methods from which weed seedling density maps are made has not been evaluated.

Spatial interpolation is a method of estimating variables such as weed seedling density at unobserved locations as well as reestimating variables at observed locations. There are two general groups of interpolation methods: global methods and local methods of interpolation. In global methods a model is constructed from all observations, and the estimated values at any point in the study area are dependent on all the sampling points. Small-scale heterogeneity is not described well, and only long-range variations such as gradients or major patches are modeled (Dessaint and Caussanel 1994). In local methods a model is constructed from all observations, but only a subset of data is used to estimate values. Each estimated value depends on the neighboring sampling points only. Consequently, small-scale heterogeneity may be described.

Dessaint and Caussanel (1994) used trend surface analysis, a global method of interpolation, to describe the weed community distribution in a corn field. They fit a polynomial surface to spatially referenced point sample weed density values. Once the long-range variation was removed, the residual variation was analyzed using local methods of interpolation. The local methods used in weed science studies include linear triangulation (Gerhards et al. 1997b) and kriging (Cardina et al. 1995; Donald 1994; Johnson et al. 1996). In the linear triangulation method the nearest three sampling point values are weighted equally when estimating

Table 1. Semivariogram model parameters (active lag and lag distance, selected isotropic model, nugget [C_0], sill [$C_0 + C$], range [A_0 or $3A_0$], and coefficient of determination [r^2]) as determined in GS+ and used for input in Surfer for the four weed species in two field sites in Nebraska.

Site	Weed species	Active lag	Lag distance	Semivariogram model	C_0	$C_0 + C$	A_0	$3A_0$	r^2
		m					m		
Saline County	ABUTH	110	8.5	Exponential	49.4	160.8	18.9	—	0.82
	APCCA	110	8.5	Exponential	0.38	1.90	27.0	—	0.97
Hall County	SETSS	366	15	Spherical	177.6	487.4	—	237	0.92
	HELAN	366	15	Spherical	0.424	0.870	—	409	0.97

a value at an unobserved location (Isaaks and Srivastava 1989). Kriging uses the semivariogram to incorporate weights and provides a measure of error associated with each predicted value (Gotway et al. 1996; Isaaks and Srivastava 1989). Although the above-mentioned methods have all been reported to provide adequate results, there has been no systematic comparison to determine which spatial interpolation method is the best.

Ideally, an interpolator would require a limited number of point-sampled weed density data and would accurately estimate density in unsampled locations. Such estimation is particularly challenging for weeds because, unlike soil and rock formations for which many of the interpolation methods were developed, weed density can be highly noncontinuous and very patchy. Therefore, the goal of this research was to determine if a single spatial interpolator performed consistently well across weed species and fields. The accuracy required for an interpolated weed seedling density map will depend on its intended use: modeling of weed population dynamics or implementation of site-specific weed management. Precision in estimating the entire density range is important for population dynamics modeling, whereas for site-specific management uses, it is more important to accurately interpolate weed-free and low-threshold density areas of the field. The specific objectives were to evaluate the sensitivity of interpolated outcomes to variation in input parameters and to determine which interpolator results in consistent and accurate weed seedling density estimates for population dynamics modeling.

Materials and Methods

Two on-farm study sites characteristic of corn or soybean production systems were chosen in eastern and central Nebraska. The first site, located in Saline County in eastern Nebraska, was a 4-ha portion of a field planted with soybean in 1992. A two-dimensional grid (29 rows by 28 columns) with 7-m intervals was superimposed on the field. At each grid point a 1- by 0.76-m quadrat was placed lengthwise between the crop rows. Just before postemergence (POST) weed control applications, all weeds were identified and counted at each grid point. Velvetleaf representing a high level of weed infestation (2.3% of field weed-free) and hemp dogbane representing a patchy infestation (82.8% of field weed-free) were the weed species chosen for interpolation. The second site, located in Hall County in central Nebraska, was a 15-ha portion of a field planted with corn in 1996. A two-dimensional grid (64 rows by 17 columns) with 12.1-m intervals was superimposed on the field. At each grid point a 1- by 0.38-m quadrat was placed lengthwise between the crop rows just before POST weed control applications. All weeds were identified and counted. Foxtail species (predominantly giant [*Setaria faberi*] and green foxtail [*S. viridis*]) representing a high level of infestation (23.7% of field weed-free) and common sunflower representing a patchy infestation (93.5% of field weed-free) were the weed species chosen for interpolation.

Table 2. Description of search types that were compared among spatial interpolation methods.

Type	Circular search radius: Saline County	Circular search radius: Hall County	Data points
	m	m	no.
Search1	8	13.9	2–4
Search2	11	19.2	4–9
Search3	15	26.1	12–16
Search4	22	38.3	29–32

Interpolation Methods

Four local spatial interpolation methods were selected: inverse-distance weighting (IDW), ordinary point kriging (OPK), minimum surface curvature (MC), and a multiquadric radial basis function (MUL). These four procedures are representative of a broad range of interpolation methods. IDW has become the applied industry's standard method for yield mapping and soil property interpolation. Kriging has been evaluated in research-oriented weed ecology studies (Cardina et al. 1995; Donald 1994; Johnson et al. 1996). MC was developed as an early computerized approach to generating contour maps of geophysical data (Briggs 1974), whereas MULs have been used to model the spatial distribution of rainfall over large geographic areas (Abtew et al. 1993). All interpolators were fit using Surfer,[1] whereas the initial parameter inputs for OPK were determined in GS+.[2] These interpolation methods estimate values at unobserved locations from the weighted averages of values from nearby observed locations. The methods differ in how they assign these weights.

Inverse-Distance Weighting

The weights are chosen so that the contribution of any one value at an observed location to the estimated value at an unobserved location decreases with distance. The weight for each observation is inversely proportional to a power of its distance from the location being estimated:

TABLE 3. Descriptive statistics (number of samples [*n*], mean, standard deviation [SD], maximum, and percent weed-free) for original, prediction, and validation data subsets of weed populations identified in two field sites in Nebraska.

Site	Prediction data subset	Validation data subset	Mean	SD	Max.	Weed-free	Mean	SD	Max.	Weed-free
			ABUTH[a]				APCCA[b]			
		n	no. per 0.76 m^2			%	no. per 0.76 m^2			%
Saline County	Original	812	0	15.4	13.0	92	2.3	0.46	1.41	13	82.8
	X1	406	406	19.5	13.1	81	0.5	0.40	1.20	13	82.8
	X2	406	406	11.3	11.5	92	4.2	0.53	1.59	12	82.8
	Y1	420	392	15.4	13.3	92	3.1	0.52	1.46	13	81.0
	Y2	392	420	15.5	12.7	84	1.5	0.40	1.35	10	85.2
	X1Y1	210	602	19.7	13.9	81	1.0	0.46	1.30	13	79.5
	X1Y2	196	616	19.4	12.2	63	0.0	0.33	1.07	9	86.2
	X2Y1	210	602	11.0	11.0	92	5.2	0.57	1.60	12	81.4
	X2Y2	196	616	11.6	12.0	84	3.1	0.48	1.57	10	84.2
			SETSS[c]				HELAN[d]			
			number per 0.38 m^2			%	number per 0.38 m^2			%
Hall County	Original	1,088	0	14.4	20.0	98	23.7	0.15	0.76	9	93.5
	X1	576	512	13.6	18.8	98	23.8	0.18	0.84	8	92.7
	X2	512	576	15.3	21.2	98	23.6	0.11	0.65	9	94.3
	Y1	544	544	13.7	19.6	98	25.2	0.17	0.86	9	93.0
	Y2	544	544	15.1	20.4	98	22.2	0.13	0.64	7	93.9
	X1Y1	288	800	13.4	19.0	98	26.0	0.23	1.00	8	91.7
	X1Y2	288	800	13.9	18.6	98	21.5	0.14	0.62	5	93.8
	X2Y1	256	832	14.1	20.2	94	24.2	0.10	0.65	9	94.5
	X2Y2	256	832	16.5	22.2	98	23.0	0.12	0.65	7	94.1

[a] ABUTH, velvetleaf, *Abutilon theophrasti* Medicus.
[b] APCCA, hemp dogbane, *Apocynum cannabinum* L.
[c] SETSS, foxtail species, *Setaria* species.
[d] HELAN, common sunflower, *Helianthus annuus* L.

$$\hat{Z} = \frac{\sum_{i=1}^{n} \frac{1}{d_i^p} Z_i}{\sum_{i=1}^{n} \frac{1}{d_i^p}} \qquad [1]$$

where $d_1, d_2, \ldots, d_n$ are the distances from each of the n observed locations to the location being estimated, $\hat{Z}$, and $Z_1, Z_2, \ldots, Z_n$ are the observed values at those locations (Isaaks and Srivastava 1989). The exponent, p, adjusts the weights to be inversely proportional to any power of the distance, and four values were compared (p = 0.5, 1, 2, and 4). The larger the power, the smaller is the effect of distance on the weights.

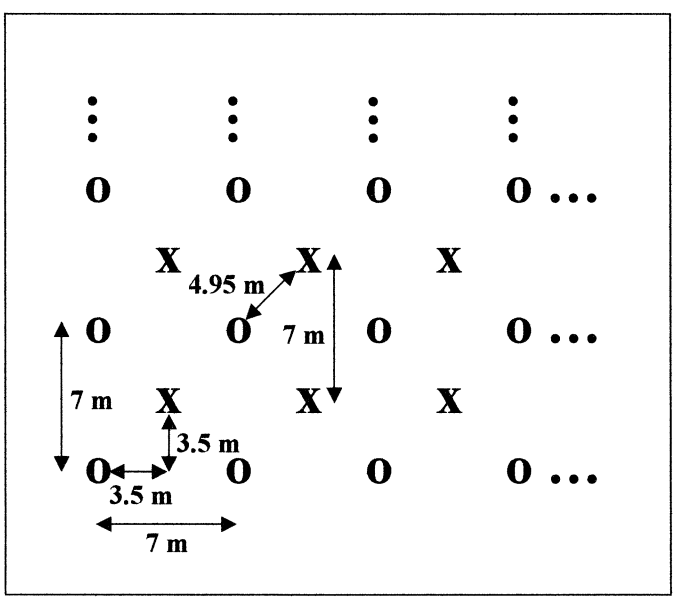

FIGURE 1. Observed locations (o) and interpolated output locations (x) that are equidistantly offset from observed locations for the Saline County field site in Nebraska.

Ordinary Point Kriging

OPK incorporates the distance between the observed locations using the semivariogram, a model of spatial autocorrelation present in the data set (Isaaks and Srivastava 1989). Weights are selected to minimize prediction mean-square error subject to an unbiased constraint and are obtained by solving kriging equations (Gotway et al. 1996). The semivariance is half the average squared difference between paired observed locations that are separated by a given lag distance:

$$\gamma_i(h) = \left(\frac{1}{2}\right) N(h) \sum_{i=1}^{N(h)} [z_i(x_i + h) - z_i(x_i)]^2 \qquad [2]$$

where $\gamma_i(h)$ is the experimental semivariance for lag distance h, and $N(h)$ is the number of pairs of observed locations separated by a given lag distance (Isaaks and Srivastava 1989).

Experimental semivariances were calculated and semivariogram models fit using GS+. A single model was selected for each weed species and was based on the largest coefficient of determination (Table 1). The spherical isotropic (independent of direction) semivariogram model is defined by

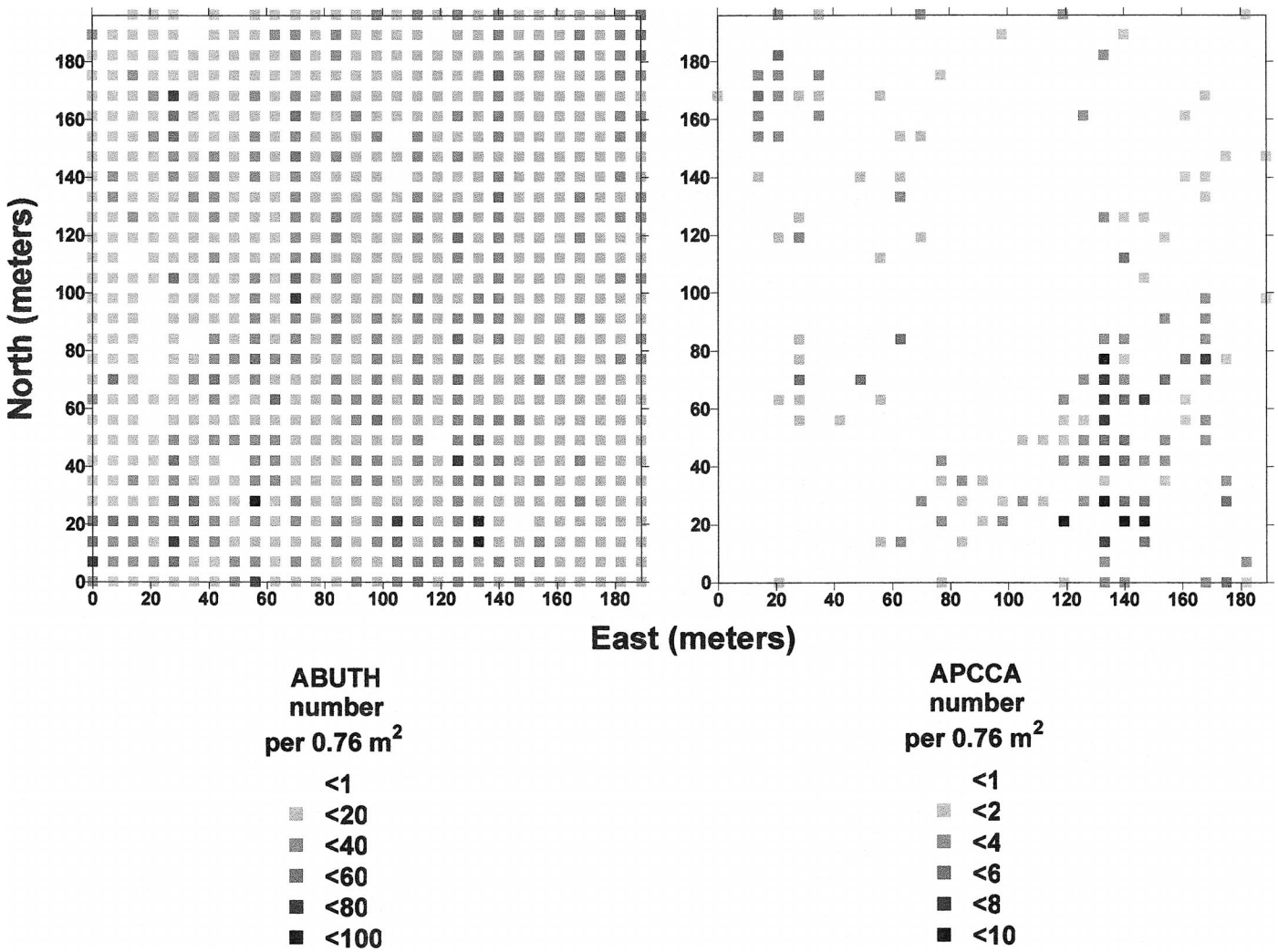

FIGURE 2. Spatial distribution of two weed species, namely, velvetleaf (*Abutilon theophrasti* Medicus ABUTH) and hemp dogbane (*Apocynum cannabinum* L. APCCA), without interpolation in the Saline County field site in Nebraska.

$$\gamma(h) = \mathrm{Co} + \mathrm{C}\left[\left(\frac{3}{2}\right)\left(\frac{h}{\mathrm{Ao}}\right) - \left(\frac{1}{2}\right)\left(\frac{h}{\mathrm{Ao}^3}\right)\right] \qquad 0 < h < \mathrm{Ao} \qquad [3a]$$

$$\gamma(h) = \mathrm{Co} + \mathrm{C} = \mathrm{sill} \qquad h \geq \mathrm{Ao} \qquad [3b]$$

where C_0, $C_0 + C$, and A_0 are nugget variance, sill (nugget plus structural variance), and range (lag distance to the sill), respectively. Nugget variance is the semivariogram value at very small lag distances, whereas sill is the value of the semivariogram at the lag distance where no more change in the variance is observed. Range is the lag distance where the semivariogram reaches the sill (Isaaks and Srivastava 1989). The exponential isotropic model is described by

$$\gamma(h) = \mathrm{Co} + \mathrm{C}\left[1 - \exp\left(-\frac{h}{a}\right)\right] \qquad [4]$$

where a is one-third of the apparent range shown on the semivariogram (Isaaks and Srivastava 1989). The parameters of the appropriate semivariogram model determined in GS+ were used in Surfer for the OPK operation.

Minimum Surface Curvature

The MC process is analogous to fitting a thin elastic plate through a set of 12 neighboring data locations (Briggs 1974; Cooke et al. 1993; Smith and Wessel 1990). This surface honors the 12 original data values and satisfies the equation

$$\frac{\partial^4 Z}{\partial x^4} + 2\left(\frac{\partial^4 Z}{\partial x^2 y^2}\right) + \frac{\partial^4 Z}{\partial y^4} = 0 \qquad [5]$$

at all observed locations (Cooke et al. 1993). On a discrete, gridded surface this corresponds to solving simultaneously the finite difference equation

$$\begin{aligned} &Z_{i+2,j} + Z_{i,j+2} + Z_{i-2,j} + Z_{i,j-2} \\ &\quad + 2(Z_{i+1,j+1} + Z_{i-1,j+1} + Z_{i+1,j-1} + Z_{i-1,j-1}) \\ &\quad - 8(Z_{i+1,j} + Z_{i-1,j} + Z_{i,j-1} + Z_{i,j+1}) + 20Z_{ij} = 0 \end{aligned} \qquad [6]$$

where Z_{ij} are the observed data values at coordinates ij away from the location being estimated (Briggs 1974; Cooke et al. 1993). The MC method requires two user-defined input

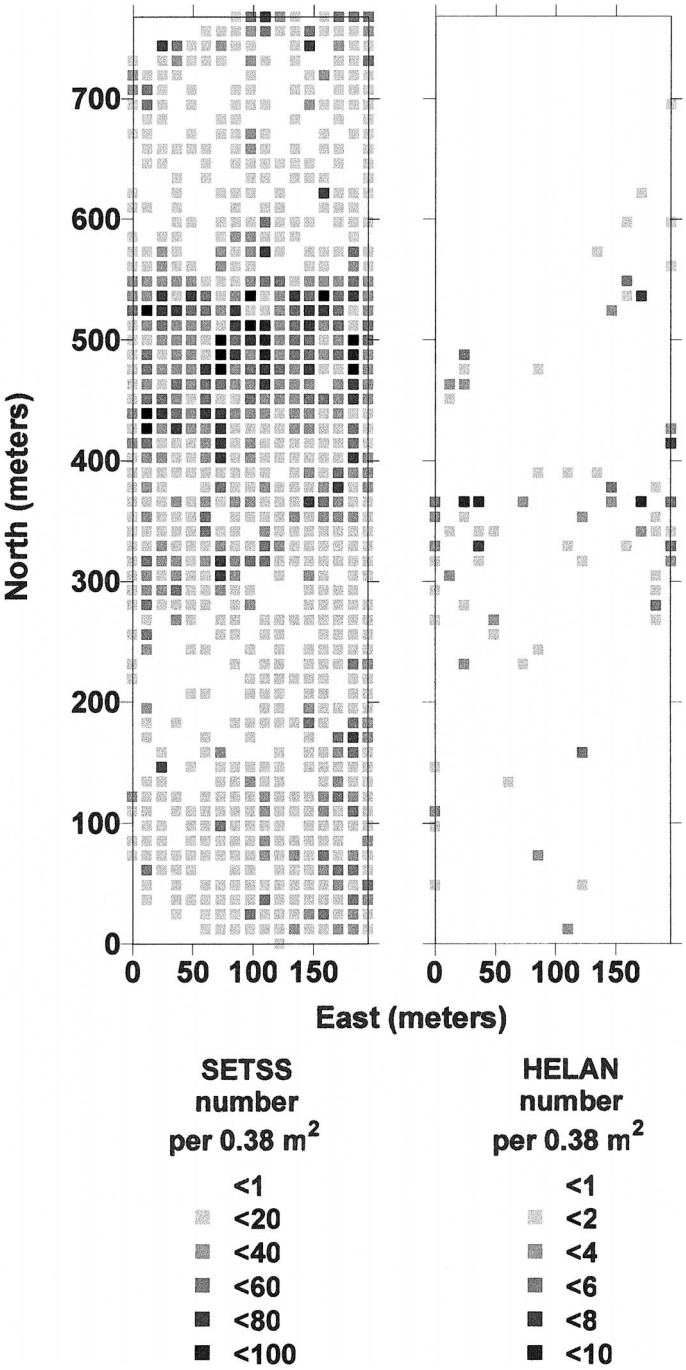

FIGURE 3. Spatial distribution of two weed species, namely, foxtail species (*Setaria* spp. SETSS) and common sunflower (*Helianthus annuus* L. HELAN), without interpolation in the Hall County field site in Nebraska.

parameters, maximum residuals, and maximum iterations to aid in the convergence of Equation 6.

Multiquadric Radial Basis Function

Rather than the weights decreasing proportionally with distance as in the IDW methods, the relationship to the neighboring observed locations with MUL is an inverted quadric surface. The relative influence of observed values increases with distance (Abtew et al. 1993; Watson 1992). Given the observation Z at location s_i, a system of simultaneous linear equations is set up and solved for the coefficients of a_i, which sum to zero. This is illustrated for four locations, but any number may be used (Watson 1992):

$$(e_1 + C(s_1 - s_2) + C(s_1 - s_3) + C(s_1 - s_4)) * a_1 = Z(s_1)$$
$$(C(s_2 - s_1) + e_2 + C(s_2 - s_3) + C(s_2 - s_4)) * a_2 = Z(s_2)$$
$$(C(s_3 - s_1) + C(s_3 - s_2) + e_3 + C(s_3 - s_4)) * a_3 = Z(s_3)$$
$$(C(s_4 - s_1) + C(s_4 - s_2) + C(s_4 - s_3) + e_4) * a_4 = Z(s_4) \quad [7]$$

where

$$C(s_i - s_j) = ((s_{ix} - s_{jx})^2 + (s_{iy} - s_{jy})^2 + e_j^2)^{0.5} \quad [8]$$

is a basis function of the distance on the x–y plane between s_i and s_j, modified by the arbitrary nonnegative constant e_j. If X is the location to be estimated with Cartesian coordinates (x, y) and the coefficients, a_i, have been determined, then the interpolated value $F(x, y)$ is

$$F(x, y) = a_1 C(X - s_1) + a_2 C(X - s_2) + a_3 C(X - s_3) + a_4 C(X - s_4) \quad [9]$$

where C is the basis function as shown in Equation 8 (Nuss and Titley 1994; Watson 1992).

Performance Evaluation of Interpolation Methods

For all interpolators except MC, search options vary and must be specified. These search options include minimum and maximum number of neighboring observed locations, shape of the search ellipse, direction to retrieve neighbors, and search radius. A series of search types were defined and compared within the IDW, OPK, and MUL interpolators (Table 2). Additionally, directional patterns were observed for weed species in the Saline County field site. Based on fitting directional semivariograms in GS+, the major axis (0, North) was three times greater than the minor axis (90, East) (data not shown). Therefore, an anisotropic ratio of 0.34 with the long axis oriented in the direction of crop rows was added to the four interpolators for velvetleaf and hemp dogbane. Additional search options tested were a circular or elliptical (no extra weight) search with or without the anisotropic ratio. Weed seedling density values were estimated for spatial locations on a 7- by 7-m (Saline County field site) or 12.1- by 12.1-m (Hall County field site) output grid that was equidistantly offset from the observed locations (Figure 1). This would allow us to potentially generate the largest interpolation errors because interpolated output locations would be the farthest away from our observed input locations. It was important to have this output grid set to equivalent spacing because small deviations from regularity exaggerated the errors among interpolators. Residual values were then calculated representing the difference between the interpolated surface and the value at the observed input location.

Four prediction data subsets with corresponding validation data subsets were created by systematically removing one-half of the data in either the x or y direction from the original data sets (Table 3) to compare among interpolation methods for a given weed species. Four additional prediction or validation data subsets were created by further systematically removing half the data, resulting in prediction data subsets that contained a quarter of the original data (Table

Table 4. MAD and RMS of residuals between interpolated surface and original data set when comparing among search types for each weed species and spatial interpolation method.[a]

		MAD			RMS		
Weed species[b]	Interpolator	Search1 and Search2	Search3	Search4	Search1 and Search2	Search3	Search4
ABUTH	IDW-2	7.08	7.66	8.24	9.34	10.15	10.83
	OPK	7.08	7.48	7.85	9.34	9.89	10.28
	MC	—[c]	9.29	—	—	12.78	—
	MUL	7.08	6.75	7.06	9.34	8.89	9.19
APCCA	IDW-2	0.44	0.49	0.55	0.91	1.00	1.08
	OPK	0.44	0.46	0.49	0.91	0.94	0.98
	MC	—	0.58	—	—	1.31	—
	MUL	0.44	0.42	0.45	0.91	0.85	0.89
SETSS	IDW-2	7.70	8.58	8.15	11.54	12.78	11.98
	OPK	7.70	9.38	5.99	11.54	13.91	10.54
	MC	—	10.97	—	—	16.74	—
	MUL	7.70	7.29	6.50	11.54	10.89	9.91
HELAN	IDW-2	0.16	0.18	0.15	0.55	0.59	0.50
	OPK	0.16	0.20	0.11	0.55	0.65	0.45
	MC	—	0.22	—	—	0.81	—
	MUL	0.16	0.17	0.13	0.55	0.53	0.44

[a] Abbreviations: MAD, mean absolute difference; RMS, root mean square; IDW-2, inverse-distance weighting squared; OPK, ordinary-point kriging; MC, minimum surface curvature; MUL, multiquadric radial basis function.
[b] ABUTH, velvetleaf, *Abutilon theophrasti* Medicus; APCCA, hemp dogbane, *Apocynum cannabinum* L.; SETSS, foxtail species, *Setaria* species; HELAN, common sunflower, *Helianthus annuus* L.
[c] Minimum surface curvature uses 12 neighboring observed locations for interpolation so that only Search3 is appropriate.

3). IDW, OPK, MC, and MUL were used to interpolate each prediction data subset, and in the process, weed density values were estimated for the missing observed input locations and compared with the corresponding validation data subset.

Two measures were used to compare search types within an interpolator and to compare across interpolators, MAD and RMS. The MAD between observed and estimated values can assess the precision of a spatial interpolator:

$$\mathrm{MAD} = \frac{\sum_{i=1}^{N} |\hat{Z}(x_i) - Z(x_i)|}{N} \qquad [10]$$

where $\hat{Z}(x_i)$ is the estimated weed density value at location x_i, $Z(x_i)$ is the observed weed density value at location x_i, and N is the total number of observed values. MAD is a criterion that incorporates both the bias and the spread of the error (Gallichand et al. 1992; Isaaks and Srivastava 1989). RMS is an estimate of the global difference between the observed and the estimated surface:

$$\mathrm{RMS} = \sqrt{\frac{\sum_{i=1}^{N} [\hat{Z}(x_i) - Z(x_i)]^2}{N}} \qquad [11]$$

The closer the fit between the two surfaces, the smaller the RMS value becomes. This measure was used by Cooke et al. (1993) and Hosseini et al. (1994) to make comparisons among interpolation methods. A relatively small value for MAD indicates high precision between interpolated

Table 5. MAD and RMS of residuals for the comparison of search options (circular or elliptical) with and without anisotropy across the four spatial interpolation methods for weed species identified in Saline County field site.[a]

		MAD				RMS			
		No anisotropy		With anisotropy		No anisotropy		With anisotropy	
Weed species[b]	Interpolator	Circular	Elliptical	Circular	Elliptical	Circular	Elliptical	Circular	Elliptical
ABUTH	IDW-2	7.66	7.39	7.77	7.75	10.15	9.67	10.27	10.16
	OPK	7.48	6.76	7.68	7.69	9.89	9.23	10.11	10.07
	MC	9.29	—[c]	9.29	—	12.78	—	12.78	—
	MUL	6.75	6.90	7.14	7.24	8.89	9.06	9.37	9.49
APCCA	IDW-2	0.49	0.46	0.51	0.49	1.00	0.95	1.03	1.02
	OPK	0.46	0.46	0.49	0.48	0.94	0.95	1.00	1.00
	MC	0.58	—	0.58	—	1.31	—	1.31	—
	MUL	0.42	0.43	0.45	0.46	0.85	0.87	0.92	0.93

[a] Abbreviations: MAD, mean absolute difference; RMS, root mean square; IDW-2, inverse-distance weighting squared; OPK, ordinary-point kriging; MC, minimum surface curvature; MUL, multiquadric radial basis function.
[b] ABUTH, velvetleaf, *Abutilon theophrasti* Medicus; APCCA, hemp dogbane, *Apocynum cannabinum* L.; SETSS, foxtail species, *Setaria* species; HELAN, common sunflower, *Helianthus annuus* L.
[c] Minimum surface curvature does not allow for an elliptical search radius to be tested.

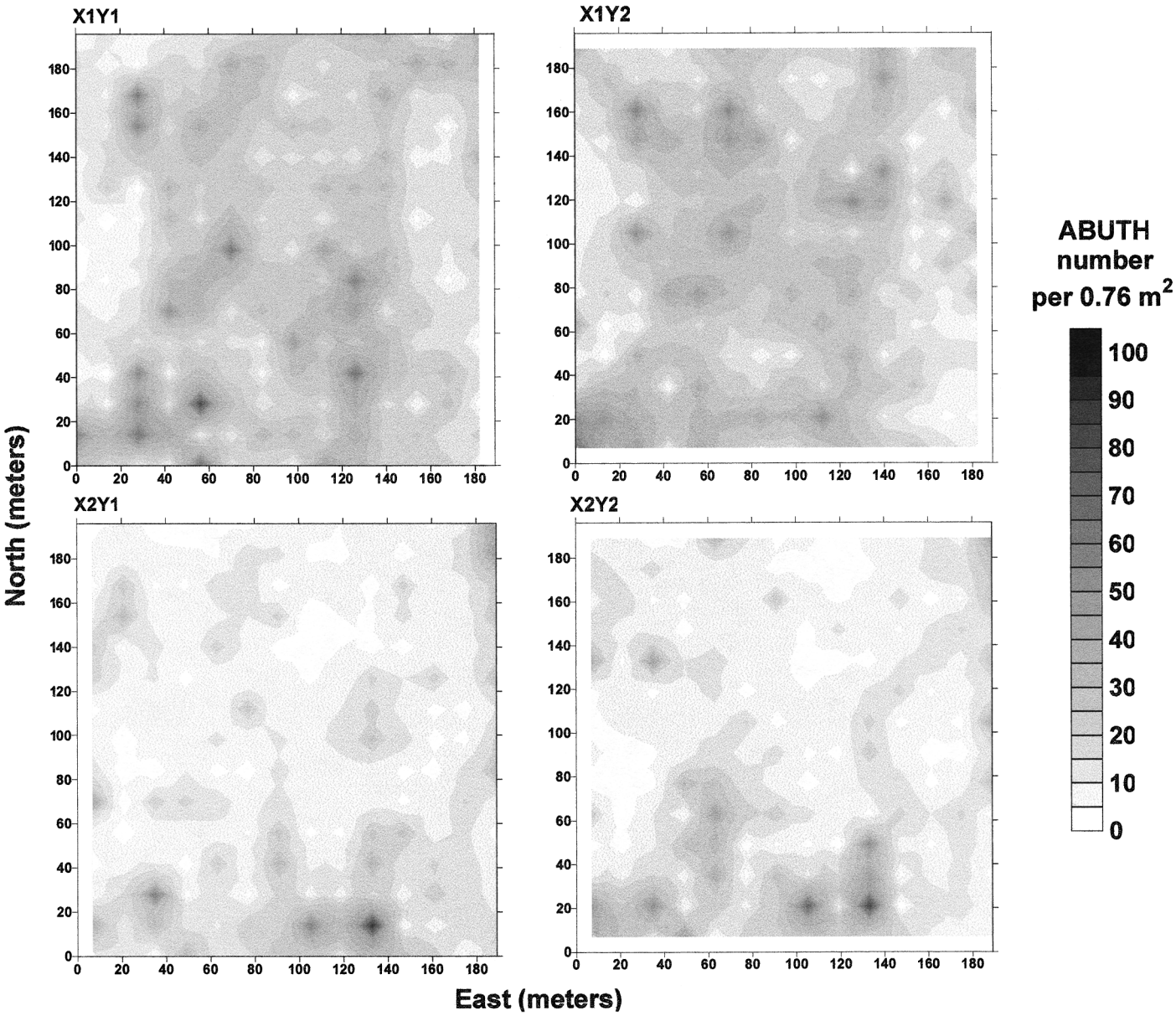

FIGURE 4. Velvetleaf (*Abutilon theophrasti* Medicus ABUTH) seedling density maps generated by inverse-distance weighting squared (IDW-2) using modified Search3 and the four quarter-data prediction subsets, X1Y1, X1Y2, X2Y1, and X2Y2 (see text for details).

values and observed sample point data, and a relatively small value for RMS indicates good global estimation of the interpolated surface compared with the original surface.

Results and Discussion

Weed Population Characteristics

The four weed species have different biological characteristics and represent different levels of infestation in the two fields. In the Saline County field site, velvetleaf was widely distributed, had a high infestation level with a mean of 15.4 seedlings per 0.76 m^2 and a maximum density of 92 seedlings per 0.76 m^2 (Figure 2; Table 3). This species has large seeds and a relatively long-lived seedbank (Burnside et al. 1996). In this same field, hemp dogbane provided an example of an isolated, patchy weed population. It had a low level of infestation with a mean density of 0.5 shoots per 0.76 m^2 and a maximum density of 13 shoots per 0.76 m^2. Hemp dogbane is a perennial weed species that predominantly reproduces from rhizomes. Foxtail species represented a high infestation in the Hall County field site, with a mean density of 14 seedlings per 0.38 m^2 and a maximum density of 98 seedlings observed per quadrat (Figure 3). In contrast to the biological characteristics of the previous two species, foxtail species have small seeds and a high germination level (Forcella et al. 1992). Common sunflower occurrence was very low in the Hall County field site, with a mean density of 0.2 seedlings per 0.38 m^2 and a maximum density of 9 seedlings observed. Tall growth habit, large seeds, and a relatively persistent seedbank characterize common sunflower (Teo-Sherrell 1996) in a similar way as velvetleaf. Infestation levels indicate differences in weed management of the two species in the two field sites.

Table 6. MAD and RMS of residuals when comparing the interpolated surface of the half-data prediction subset with the original data in the half-data validation subset across spatial interpolation methods for each weed species.[a]

Weed species[b]	Interpolator	MAD				RMS			
		X1	X2	Y1	Y2	X1	X2	Y1	Y2
ABUTH	IDW-2	11.67	12.22	7.82	9.02	15.62	14.85	11.06	11.21
	IDW-4	11.68	12.24	7.20	8.25	15.72	15.01	10.35	10.50
	OPK	11.68	12.22	8.18	9.44	15.61	14.81	11.45	11.63
	MC	12.09	12.39	6.98	8.14	16.10	15.63	10.20	10.62
	MUL	11.74	12.26	6.96	8.00	15.80	15.15	10.12	10.32
APCCA	IDW-2	0.62	0.61	0.48	0.59	1.23	1.37	1.14	1.15
	IDW-4	0.62	0.61	0.47	0.57	1.25	1.38	1.12	1.14
	OPK	0.62	0.61	0.48	0.58	1.23	1.37	1.15	1.15
	MC	0.68	0.66	0.51	0.61	1.34	1.41	1.15	1.20
	MUL	0.62	0.61	0.46	0.57	1.28	1.38	1.11	1.15
SETSS	IDW-2	10.20	11.49	9.92	9.85	15.82	17.12	14.29	15.34
	IDW-4	10.29	11.52	9.69	9.58	16.04	17.25	14.18	15.13
	OPK	10.26	11.54	10.27	10.22	15.82	17.15	14.72	15.78
	MC	10.51	9.19	9.79	9.90	16.56	15.52	15.58	14.76
	MUL	10.28	9.43	9.57	9.66	16.05	15.80	15.11	14.15
HELAN	IDW-2	0.21	0.23	0.23	0.24	0.75	0.72	0.85	0.72
	IDW-4	0.21	0.23	0.23	0.24	0.75	0.74	0.86	0.74
	OPK	0.22	0.23	0.23	0.25	0.76	0.71	0.85	0.72
	MC	0.23	0.25	0.26	0.28	0.77	0.77	0.89	0.81
	MUL	0.21	0.23	0.23	0.24	0.75	0.74	0.86	0.74

[a] Abbreviations: MAD, mean absolute difference; RMS, root mean square. IDW, inverse-distance weighting; OPK, ordinary-point kriging; MC, minimum surface curvature; MUL, multiquadric radial basis function.

[b] ABUTH, velvetleaf, *Abutilon theophrasti* Medicus; APCCA, hemp dogbane, *Apocynum cannabinum* L.; SETSS, foxtail species, *Setaria* species; HELAN, common sunflower, *Helianthus annuus* L.

Evaluation of Input Parameters

For IDW-2, OPK, and MUL interpolators, Search3 had the smallest MAD and RMS values for velvetleaf and hemp dogbane, with MUL having the smallest values overall (Table 4). For foxtail species and common sunflower, Search4 had the smallest MAD values for OPK and the smallest RMS values for MUL. Search1 and Search2 gave identical values across spatial interpolators for each weed species, and whereas they are small values, they do not appear to indicate any differences. Search3 was the only appropriate search for the MC interpolator and had MAD and RMS values that were the highest among interpolators and across weed species. MCs are known to have large oscillations and extraneous inflection points, which make them unsuitable for gridding a range of variables (Smith and Wessel 1990), and weed density is no exception.

As a result of this comparison, we determined that the best search type was a modified Search3, such that the search radius was 15 or 26.1 m depending on field site, the minimum number of observed locations was five, and the maximum number was 16. A concern with Search4 was the loss of data at field edges because of the limiting number of data points (29 to 32) that were brought into the search. MAD and RMS calculations did account for the loss of information at field edges, but this created an incomplete map. Typical search options selected by other authors were 4 to 24 neighboring locations to estimate clay content (Gallichand and Marcotte 1993) and six neighboring locations to estimate electrical conductivity and sodium absorption ratios (Gallichand et al. 1992). A total of 24 neighboring locations were selected to estimate nitrate-nitrogen and percent organic matter content in agricultural fields (Gotway et al. 1996) or to estimate electrical conductivity across a 16,000-ha area (Hosseini et al. 1994).

Increasing the power coefficient for the IDW interpolator from 0.5 to 4 decreased the value for MAD and RMS across weed species for Search3 (data not shown). No differences were detected for Search2 across power coefficients for each weed species. As the value of the power coefficient increases, the influence of distant observations increases (Isaaks and Srivastava 1989). As the power coefficient is decreased, the weights given to the observations become more similar. For example, with $p = 0.5$, the closest observation receives nearly 16% of the total weight and the farthest observation receives about 13%. As the value of p increases, the individual weights become more dissimilar, such that the farthest samples contribute the least and the nearest samples become more influential (Isaaks and Srivastava 1989). This appears more appropriate for interpolation of weed seedling densities because of their patchy nature and influence of nearer observations.

Directional patterns were observed for the weed populations found in the Saline County field site. An elliptical search without anisotropy had the lowest MAD and RMS values across the four spatial interpolators (Table 5). The shape of the search radius appeared to be more important than including an additional weighting through the anisotropic ratio.

Establishing the output grid spacing to which weed density estimates would be predicted was very important. For example, attempting to interpolate to an output grid spacing of equal thirds within a 7- by 7-m grid, that is, output grid points at multiples of 2.333, systematically introduced a bias from the original sample points. Because

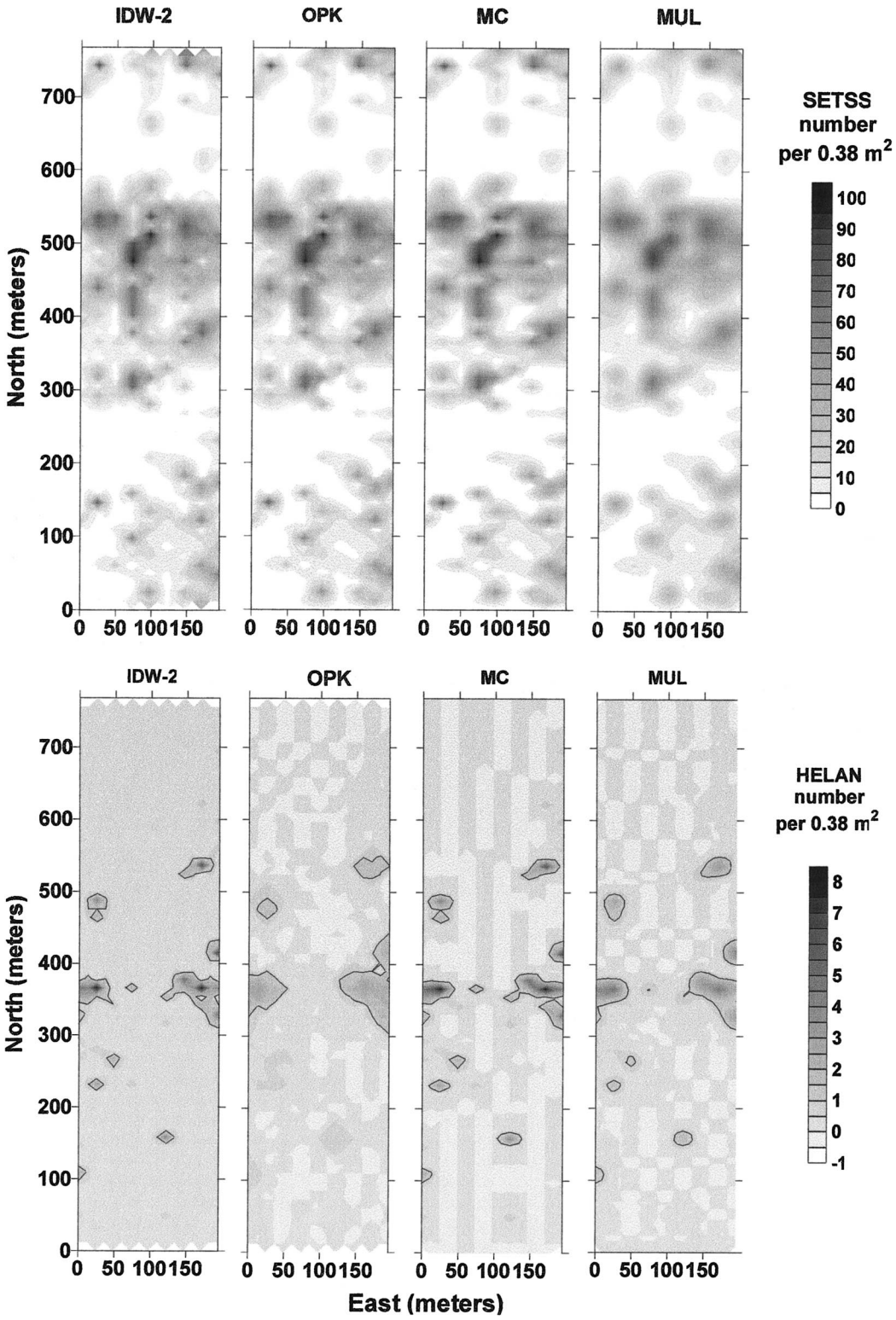

FIGURE 5. Foxtail species (*Setaria* spp. SETSS) and common sunflower (*Helianthus annuus* L. HELAN) seedling density maps generated by four interpolation methods, namely, inverse-distance weighting squared (IDW-2), ordinary point kriging (OPK), minimum surface curvature (MC), and multiquadric radial basis function (MUL), using modified Search3 and X1 prediction subset (see text for details).

distance from the southwest origin increased, the output coordinates were systematically displaced by (−0.001), (−0.002), (−0.003), etc. units away from the original input coordinates. This resulted in exaggerated errors for each spatial interpolator.

Evaluation of Spatial Interpolators

The systematic removal of data to create the prediction or validation data subsets highlighted an important spatial pattern of velvetleaf density, a pattern oriented in the direc-

Table 7. MAD and RMS of residuals when comparing the interpolated surface of the quarter-data prediction subset with the three-quarters original data in the validation data subset across spatial interpolation methods for each weed species.[a]

Weed species[b]	Interpolator	MAD				RMS			
		X1Y1	X1Y2	X2Y1	X2Y2	X1Y1	X1Y2	X2Y1	X2Y2
ABUTH	IDW-2	10.77	11.03	9.89	9.61	13.40	13.69	13.89	13.58
	IDW-4	10.99	11.16	10.05	9.69	13.94	14.03	14.02	13.79
	OPK	10.83	11.04	9.89	9.64	13.39	13.66	13.92	13.59
	MC	11.29	11.38	10.45	10.10	14.49	14.56	14.40	14.18
	MUL	11.18	11.30	10.32	9.98	14.30	14.41	14.27	14.08
APCCA	IDW-2	0.62	0.58	0.66	0.59	1.25	1.36	1.26	1.18
	IDW-4	0.60	0.55	0.65	0.57	1.28	1.34	1.34	1.24
	OPK	0.62	0.58	0.66	0.60	1.24	1.37	1.27	1.19
	MC	0.65	0.58	0.70	0.64	1.33	1.36	1.44	1.34
	MUL	0.63	0.57	0.69	0.62	1.30	1.36	1.42	1.31
SETSS	IDW-2	10.34	11.06	10.54	10.65	15.54	16.22	16.24	15.28
	IDW-4	10.40	11.15	10.52	10.82	16.08	16.77	16.70	16.05
	OPK	10.56	11.21	10.82	10.92	15.67	16.31	16.37	15.41
	MC	10.96	11.69	10.91	11.06	16.86	17.65	17.31	16.58
	MUL	10.70	11.41	10.76	10.91	16.46	17.16	17.11	16.24
HELAN	IDW-2	0.27	0.24	0.20	0.21	0.73	0.80	0.71	0.75
	IDW-4	0.27	0.24	0.20	0.21	0.83	0.84	0.74	0.79
	OPK	0.27	0.24	0.21	0.21	0.68	0.78	0.72	0.72
	MC	0.31	0.27	0.22	0.24	0.91	0.87	0.78	0.83
	MUL	0.30	0.26	0.21	0.23	0.86	0.85	0.76	0.81

[a] Abbreviations: MAD, mean absolute difference; RMS, root mean square. IDW, inverse-distance weighting; OPK, ordinary-point kriging; MC, minimum surface curvature; MUL, multiquadric radial basis function.

[b] ABUTH, velvetleaf, *Abutilon theophrasti* Medicus; APCCA, hemp dogbane, *Apocynum cannabinum* L.; SETSS, foxtail species, *Setaria* species; HELAN, common sunflower, *Helianthus annuus* L.

tion of crop rows and field traffic movement (Figure 2). Through systematic removal in the *x* direction, the X1 prediction data subset had a higher mean velvetleaf density (19.5 seedlings per 0.76 m^2) compared with the X2 prediction data subset (11.3 seedlings per 0.76 m^2). This was in contrast to the similar velvetleaf means for the Y1 and Y2 prediction data subsets (Table 3). This pattern was also observed in the quarter-data prediction subsets (Figure 4). Four very different velvetleaf seedling density maps were generated depending on the chosen quarter-data prediction subset. In general, the variation in velvetleaf mean densities increased with fewer data points in each subset. Such a dramatic influence of systematic data removal was not observed for the other three weed species.

The levels of precision (MAD) and overall estimation (RMS) of hemp dogbane– and common sunflower–interpolated surfaces from half-data prediction subsets were similar for the IDW-2, IDW-4, OPK, and MUL methods, whereas MC had low precision and poor global estimation (Table 6). IDW-2, IDW-4, and OPK methods resulted in the best precision and overall estimation of the X1 and X2 velvetleaf prediction subsets, whereas MUL was the best method for the Y1 and Y2 prediction subsets. For foxtail species the interpolation method that gave the best level of precision depended on the prediction data subset: IDW-2 for X1 subset, MC for X2 subset, IDW-4 for Y1 subset, and MUL for both X2 and Y2 subsets (Table 6). Figure 5 highlights the subtle differences among interpolation methods for a given prediction data subset (X1) for both foxtail species and common sunflower.

For low–infestation level populations, IDW-2, IDW-4, and OPK were similar with the lowest levels of precision (MAD) and global estimation (RMS) of hemp dogbane– and common sunflower–interpolated surfaces for the quarter-data prediction subsets (Table 7). For foxtail species, IDW-2 resulted in the smallest MAD and RMS values across interpolators and prediction data subsets. For velvetleaf, IDW-2 and OPK methods resulted in the smallest MAD and RMS values across interpolators and quarter-data prediction subsets (Table 7).

Interpolation precision of all methods was low for the high-infestation weed populations, velvetleaf and foxtail species. Mean absolute difference between estimated and observed half-data prediction subset values ranged from 45 to 80% of the observed mean velvetleaf density and from 66 to 80% of the observed mean density of foxtail species (Table 7). Interpolation precision was very poor for the low-infestation populations of hemp dogbane and common sunflower. Mean absolute difference between estimated and observed half-data prediction subset values ranged from 100 to 148% of the observed mean hemp dogbane shoot density and from 140 to 187% of the observed mean common sunflower density. This poor precision is likely because of the high variability inherent in spatial weed population data (Hosseini et al. 1994; Johnson et al. 1996).

Implications

There is tremendous interest in using spatial distribution maps of weed seedling densities to study and model the spatial dynamics of weed populations and in guiding the application of weed control practices on individual fields. The results from this study indicate two key findings with implications on our ability to depict accurately the spatial distribution of weed densities in the field.

First, among the spatial interpolation methods compared, no single method appeared to outperform or be more accurate in generating weed seedling density maps. Clearly, MC was not appropriate, whereas the other three methods performed equally well. Consistently across the interpolation methods, density was underestimated in high-density portions of the field. This is of concern for accurate population dynamics modeling of density-dependent responses to weed management applications (Dieleman and Mortensen 1999; Dieleman et al. 1999; Neeser et al. 2002). The overestimation of densities in low-density areas of a field is of less concern for site-specific management because most practitioners want to ensure that no weeds are overlooked. For patchy, low–infestation level populations, IDW-2, IDW-4, OPK, and MUL interpolation methods were appropriate. For high–infestation level populations, because the available data becomes less, IDW-2 and OPK gave the best precision among interpolation methods.

Second, the characteristics of the chosen sampling grid, that is, the distance between sampling points and quadrat size, can dramatically influence weed seedling map outcomes. Often, individual weed patches occurred between sampling points and were not recorded. Important information is missed between the sampling quadrats. Additionally, if the weed population is observed to have a systematic pattern across the field, such as the velvetleaf population in the Saline County field site, a fixed sampling grid could under- or overrepresent the population because of sampling anomalies. This indicates the need to explore alternative and innovative sampling strategies, for example, increase in the number of points sampled in areas of a field with greater uncertainty of the weed population's spatial distribution. Other ways of obtaining spatial weed population data, such as remote sensing, ground-based sensing, and directed sampling, can be integrated to assist in developing more accurate weed seedling density maps.

Weed populations exhibit a unique footprint in each agricultural field, ranging from low to high infestation levels and the location in the field where a given weed species is found. The results from this study clearly indicate that the particular method of interpolating densities is not critical as long as poor performers are not selected. The future challenge is to develop general but robust rules for sampling weed populations, to bring a diversity of data information sources together, and to strive to generate accurate weed seedling density maps.

Sources of Materials

[1] Surfer for Windows, V.6, Contouring and 3D Surface Mapping, Golden Software, Inc., 809 14th Street, Golden, CO 80401.

[2] GS$^+$: Geostatistics for the Environmental Sciences, Gamma Design Software, P.O. Box 201, Plainwell, MI 49080.

Acknowledgments

This article is published as University of Nebraska Agricultural Research Division Journal Series No. 13379.

Literature Cited

Abtew, W., J. Obeysekera, and G. Shih. 1993. Spatial analysis for monthly rainfall in south Florida. Water Res. Bull. 29:179–188.

Brain, P. and R. Cousens. 1990. The effect of weed distribution on predictions of yield loss. J. Appl. Ecol. 27:735–742.

Briggs, I. C. 1974. Machine contouring using minimum curvature. Geophysics 39:39–48.

Burnside, O. C., R. G. Wilson, S. Weisberg, and K. G. Hubbard. 1996. Seed longevity of 41 weed species buried 17 years in eastern and western Nebraska. Weed Sci. 44:74–86.

Cardina, J., D. H. Sparrow, and E. L. McCoy. 1995. Analysis of spatial distribution of common lambsquarters (*Chenopodium album*) in no-till soybean (*Glycine max*). Weed Sci. 43:258–268.

Cooke, R. A., S. Mostaghimi, and J. B. Campbell. 1993. Assessment of methods for interpolating steady-state infiltrability. Trans. ASAE 36: 1333–1341.

Dessaint, F. and J. P. Caussanel. 1994. Trend surface analysis: a simple tool for modelling spatial patterns of weeds. Crop Prot. 13:433–438.

Dieleman, J. A. and D. A. Mortensen. 1999. Characterizing the spatial pattern of *Abutilon theophrasti* seedling patches. Weed Res. 39:455–467.

Dieleman, J. A., D. A. Mortensen, A. R. Martin, and D. Y. Wyse-Pester. 1999. Influence of velvetleaf (*Abutilon theophrasti*) and common sunflower (*Helianthus annuus*) density on weed management outcomes. Weed Sci. 47:81–89. [Errata 47(2):254]

Donald, W. W. 1994. Geostatistics for mapping weeds, with a Canada thistle (*Cirsium arvense*) patch as a case study. Weed Sci. 42:648–657.

Forcella, F., R. G. Wilson, K. A. Renner, J. Dekker, R. G. Harvey, D. A. Alm, D. D. Buhler, and J. Cardina. 1992. Weed seedbanks in the U.S. corn belt: magnitude, variation, emergence, and application. Weed Sci. 40:636–644.

Gallichand, J., G. D. Buckland, D. Marcotte, and M. J. Hendry. 1992. Spatial interpolation of soil salinity and sodicity for a saline soil in southern Alberta. Can. J. Soil Sci. 72:503–516.

Gallichand, J. and D. Marcotte. 1993. Mapping clay content for subsurface drainage in the Nile Delta. Geoderma 58:165–179.

Gerhards, R., M. Sökefeld, K. Schulze-Lohne, D. A. Mortensen, and W. Kühbauch. 1997a. Site specific weed control in winter wheat. J. Agron. Crop Sci. 178:219–225.

Gerhards, R., D. Y. Wyse-Pester, D. A. Mortensen, and G. A. Johnson. 1997b. Characterizing spatial stability of weed populations using interpolated maps. Weed Sci. 45:108–119.

Gotway, C. A., R. B. Ferguson, G. W. Hergert, and T. A. Peterson. 1996. Comparison of kriging and inverse-distance methods for mapping soil parameters. Soil Sci. Soc. Am. J. 60:1237–1247.

Heisel, T., A. K. Ersbøll, and C. Andreasen. 1999. Weed mapping with co-kriging using soil properties. Prec. Agric. 1:39–52.

Hosseini, E., J. Gallichand, and D. Marcotte. 1994. Theoretical and experimental performance of spatial interpolation methods for soil salinity analysis. Trans. ASAE 37:1799–1807.

Isaaks, E. H. and R. M. Srivastava. 1989. An Introduction to Applied Geostatistics. New York: Oxford University Press. pp. 249–266, 278–322.

Johnson, G. A., D. A. Mortensen, and C. A. Gotway. 1996. Spatial and temporal analysis of weed seedling populations using geostatistics. Weed Sci. 44:704–710.

Johnson, G. A., D. A. Mortensen, L. J. Young, and A. R. Martin. 1995. The stability of weed seedling population models and parameters in eastern Nebraska corn (*Zea mays*) and soybean (*Glycine max*) fields. Weed Sci. 43:604–611.

Lindquist, J. L., J. A. Dieleman, D. A. Mortensen, G. A. Johnson, and D. Y. Wyse-Pester. 1998. Economic importance of managing spatially heterogeneous weed populations. Weed Technol. 12:7–13.

Mortensen, D. A., G. A. Johnson, and L. J. Young. 1993. Weed distribution in agricultural fields. Pages 113–124 *in* P. C. Robert and R. H. Rust, eds. Soil Specific Crop Management. Madison, WI: American Society of Agronomy.

Neeser, C., J. A. Dieleman, and D. A. Mortensen. 1999. Density dependant velvetleaf (*Abutilon theophrasti*) seedling survival in response to post-emergence herbicide applications. Page 337 *in* Proceedings of the Weed Science Society of America. San Diego, CA: Weed Science Society of America.

Nuss, W. A. and D. W. Titley. 1994. Use of multiquadric interpolation for meteorological objective analysis. Monthly Weather Rev. 122:1611–1631.

Smith, W.H.R. and P. Wessel. 1990. Gridding with continuous curvature splines in tension. Geophysics 55:293–305.

Teo-Sherrell, C.P.A. 1996. The Fates of Weed Seeds. Ph.D. dissertation. University of Nebraska, Lincoln, NE.

Wallinga, J. 1995. The role of space in plant population dynamics: annual weeds as an example. Oikos 74:377–383.

Watson, D. F. 1992. Contouring—A Guide to the Analysis and Display of Spatial Data. Tarrytown, NY: Pergamon Press. pp. 118–130.
Wiles, L. J., G. W. Oliver, A. C. York, H. J. Gold, and G. G. Wilkerson. 1992. Spatial distribution of broadleaf weeds in North Carolina soybean (*Glycine max*) fields. Weed Sci. 40:554–557.
Williams, M. M., II, R. Gerhards, and D. A. Mortensen. 1999. Spatiotemporal outcomes of site-specific weed management in maize. Pages 897–906 *in* J. V. Stafford, ed. Precision Agriculture '99. Sheffield, Great Britain: Society of Chemical Industry.
Wilson, B. J. and P. Brain. 1991. Long-term stability of distribution of *Alopecurus myosuroides* Huds. within cereal fields. Weed Res. 31:367–373.

Received June 19, 2001, and approved May 31, 2002.