A digital elevation model to aid geostatistical mapping of weeds in sunflower crops

M. Jurado-Expósito, F. López-Granados, J.M. Peña-Barragán, L. García-Torres

Research article

A digital elevation model to aid geostatistical mapping of weeds in sunflower crops

M. Jurado-Expósito*, F. López-Granados, J.M. Peña-Barragán, L. García-Torres

Institute of Sustainable Agriculture, CSIC, PO Box 4084, 14080, Córdoba, Spain

Abstract – A major concern in landscape management and precision agriculture is the variable-rate application of herbicides in order to reduce herbicide treatment load. These applications require a correct assessment and knowledge of the density and potential spatial variability of weed species within fields. This article addresses the issue of incorporating a digital elevation model as secondary spatial information into the mapping of main weed species present in two sunflower crops in Andalusia, Spain. Two prediction methods were used and compared for mapping weed density for precision agriculture. The primary information was obtained from an intensive grid weed density sampling and the secondary spatial information, e.g., elevation from a digital elevation model. The prediction methods were two geostatistical algorithms: ordinary kriging and kriging with an external drift, which takes into account the influence of landscape. Mean squared error was used to evaluate the performance of the map prediction quality. The best prediction method for mapping most of the weed species was kriging with an external drift, with the smallest mean squared error, indicating the highest accuracy. The results showed that kriging with an external drift with elevation reduced the prediction variance compared with ordinary kriging. Maps obtained from these kriged estimates showed that the incorporation of a digital elevation model as secondary exhaustive information can improve the accuracy of predicted weed densities within fields. These results suggest that kriging with an external drift of weed density data with elevation as a secondary exhaustive variable could be used in such situations, and in this way, the accuracy of maps for precision agriculture, which is the preliminary step in a precision agricultural management program, could be improved with little or no additional cost, since a digital elevation model could be obtained as part of other analyses.

digital elevation model / kriging / precision agriculture / weed spatial variability

1. INTRODUCTION

It is well known that weeds are not uniformly distributed in arable fields, but tend to appear in patches (Heisel et al., 1996; Jurado-Expósito et al., 2003, 2004). As a consequence, the preliminary step in a precision weed management program is to map the weed variability in the area under management. The combination of a weed map, positioning devices, e.g. Differential Global Positioning Systems, DGPS, and a sprayer to differentiate herbicide application on the basis of the weed infestation makes it possible to reduce herbicide use, costs and environmental pressure.

A number of methods have been proposed for the estimation of spatial and temporal distribution and the creation of weed density maps: interpolation techniques (e.g. Ersbøll et al., 1993) such as simple inverse distance calculations (e.g. Bregt et al., 1992); linear triangulation (Gerhards et al., 1997), polynomial interpolation (Zanin et al., 1998), or geostatistics.

Geostatistics, which is based on the theory of regionalized variables (Journel and Huijbregts, 1978; Goovaerts, 1997,

1999a), is increasingly preferred because it allows one to capitalize on the spatial correlation between neighboring observations to predict attribute values at unsampled locations. The geostatistical prediction technique (kriging) weights the average of observed weed densities and is the only estimator that provides the variance of prediction error (Cressie, 1991; Isaaks and Srivastava, 1989; Kristensen and Ersbøll 1995).

Several authors have shown that the geostatistical prediction technique (ordinary kriging) provides better estimates of variability in weed spatial and temporal distribution variability than conventional methods (Heisel et al., 1996; Jurado-Expósito et al., 2003, 2004, 2005; Zanin et al., 1998).

In addition to providing a measurement of prediction error, kriging variance, a major advantage of geostatistical prediction over simpler methods is that sparsely sampled observations of the primary variable, the variable of interest, i.e. weed density, can be complemented by secondary variables that are more densely sampled. The estimation of the primary attribute is generally improved by taking into account secondary information originating from other related categorical or continuous attributes (Goovaerts, 1999b). This secondary information

* Corresponding author: montse.jurado@ias.csic.es

is said to be exhaustively sampled when it is available at all primary data locations and at all nodes of estimation grids, e.g. spectral data or digital elevation models.

Therefore, when the variable of interest is scant or poorly correlated in space, the prediction of this variable over the whole study area may be improved by accounting for secondary information exhaustively sampled over the same study area. The secondary information can be incorporated using kriging with an external drift. Several authors have reported accurate prediction maps from sparsely sampled observations of a primary attribute; for example, maps of rainfall erosivity (Goovaerts, 1999b), precipitation (Goovaerts, 2000) and evapotranspiration (Vanderlinden, 2001), which were completed by using digital elevation models as exhaustive secondary attributes that are more densely sampled, and applying different multivariate extensions of kriging. These works were based on a positive significant correlation between average annual precipitation and elevation, i.e. precipitation tends to increase with increasing elevation. Thus, when any exhaustive and usually cheaper source of secondary information is available, it could be incorporated into the mapping of weed density in order to increase the accuracy of predicted weed estimates to be used in precision farming.

Other prediction methods that also incorporate secondary information available on a large scale, such as digital data from aerial photographs, have also been developed to extend sparse and expensive soil measurements. Examples include: simple linear regression, regression trees and geostatistical methods such as co-kriging, kriging with an external drift or kriging with varying local means (Bishop and McBratney, 2001; López-Granados et al., 2005).

Field observations clearly show that the distribution of weeds across the landscape has spatial dependence, and it is well known that weed sampling on the ground is labor-intensive, time-consuming and expensive. So, if an intensive and cheaper source of secondary information, e.g. a digital elevation model, can be used to improve estimation of a sparsely sampled variable, e.g. weed density, it may be possible to reduce weed sampling. Thus, if a digital elevation model is available, it could be incorporated into the mapping of weed density in order to increase the accuracy of predicted weed density estimates to be used in precision farming. In this study, we focus on kriging with an external drift as a prediction method that accounts for an auxiliary variable in the prediction of the variable of interest, because this method has not been widely used in weed science.

There are very few applications of using the digital elevation model information for weed science. Liu et al. (2002) examined the spatial variability of herbicide efficacy in a field in which corn had been planted continuously for 10 years, concluding that weed biomass was correlated positively with elevation and the total amount of atrazine mineralized, whereas corn biomass was correlated negatively with these parameters. Shafii et al. (2004) also estimated the likelihood of yellow starthistle occurrence by incorporating elevation, slope and aspect into prediction models by using Bayesian classification.

Sunflower is one of the most important crops in Andalusia, southern Spain, with over 320 000 ha sown annually (MAYPA, 2005). Spatial distribution of weeds in sunflower and site-specific management maps have been reported by Jurado-Expósito et al. (2003). However, the spatial distribution of weeds occurring in sunflower using a digital elevation model as a secondary information layer has not yet been assessed.

The objective of this study was to interpolate and map weed spatial variability in two sunflower fields based on an intensive weed sampling and using different prediction methods for arriving at the best approach for weed density mapping. Two different prediction techniques were applied: (a) ordinary kriging: a method that uses only the weed density sampling data recorded at the two sunflower fields, and (b) kriging with an external drift: an algorithm that combines weed density data with a digital elevation model as secondary information, and to determine the influence of field topography in the accuracy of predicted weed density to be used in a precision agricultural management program.

2. MATERIAL AND METHODS

2.1. Study area and weed sampling scheme

The study was carried out in two sunflower (*Helianthus annuus* L.) fields located in Casilla and Cruce, Córdoba, Spain. The primary information, i.e., weed density was obtained through intensive field surveys conducted in selected areas of 1.96 ha and 1.26 ha, respectively, within a larger field of around 50 ha, and their borders were at least 50 m from the main borders of the fields. The monitored fields were cropped previously with winter wheat (*Triticum aestivum* L.) shallow tillage. Sunflower was seeded at 4 kg ha^{-1} in rows 0.7 m apart, in late March.

Weed assessments took place in mid-May when the sunflower was at the 6- to 8-leaf stage and most weeds were 4 cm to 10 cm tall, using a grid of 7 m by 7 m, resulting in 449 and 256 sampled points, corresponding to 229 and 203 observations per ha, in Casilla and Cruce, respectively. The position of each node was georeferenced using a sub-meter differential DGPS TRIMBLE PRO-XRS (Trimble Navigation Limited, 749 North Mary Avenue, Sunnyvale, CA 94085, USA) provided with a TDC-1 unit. The total numbers of plants of each weed species were recorded at the grid points in 2 m^2 squares. A total of seven weed species were sampled: *Convolvulus arvensis* L., *Ridolfia segetum*, Moris., *Phalaris* spp., *Sinapis arvensis* L., *Amaranthus albus* L., *Galium aparine* L. and *Chrozophora tinctoria* (L.) A. Juss.

The analyses started with the description of weed densities. Classical descriptors, mean, maximum, standard deviation and skewness, were determined for each weed species. The descriptive statistics of weed densities suggested that they were symmetrical: skewness between –2 and 2, and therefore no transformation was used prior to estimating weed density maps, following the methodologies described in Zanin et al. (1998), López-Granados et al. (2002) and Jurado-Expósito et al. (2003, 2004, 2005).
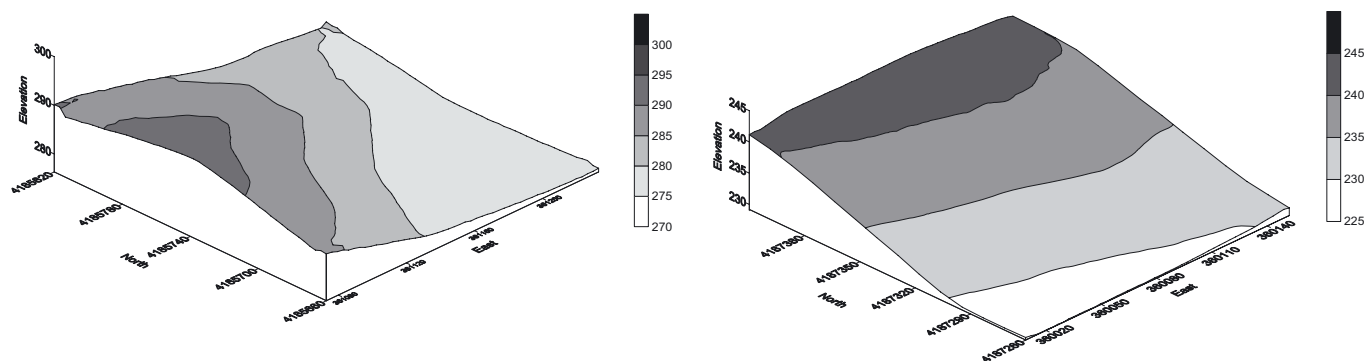
Figure 1. Digital elevation maps in sunflower fields: (a) Casilla and (b) Cruce. Axes X and Y are in Universal Transverse Mercator units (m), Z: elevation (m).

2.2. Creation of a digital elevation model

A digital elevation model is a representation of continuous elevation values over a topographic surface by a regular array of z-values, i.e., by coordinates and numerical descriptions of altitude. The digital elevation model data files are digital representations of cartographic information in a raster form. No existing digital elevation model of reasonable scale was available, so we had to create our own. The digital elevation model was derived from topographic contours of 1:10.000 maps, free from the Andalusia Government. The contours were manually digitalized at intervals of 10 m, and tagged with the elevation values, along with specific point elevations such as mountain-tops or valleys. Thus, a dxf vector file was generated with the commercial software package Autocad (Autocad 2006, Autodesk, Inc., 111 McInnis Parkway, San Rafael, CA 94903, USA). The continuous raster digital elevation model was generated with software specifically programed by using IDL language (IDL 6.3, Interactive Data Language, Research System Inc., 4990 Pearl East, Circle, Boulder, CO 80301, USA). This software extracted the x, y and z coordinates from the previous dxf vector file, and then a triangulate procedure was used to generate the digital elevation model raster file. The final result is a raster map in which each 0.5 m pixel has a unique elevation value derived from the contours. Elevation values of the digital elevation model should be very close to the known elevation values taken during the sampling of weed density (x, y and z coordinates from the DGPS receiver) (Ley, 1986). The digital elevation model was visualized with the software ENVI (ENVI 4.3, Research Systems Inc., 4990 Pearl East, Circle, Boulder, CO 80301, USA). The elevation of all estimation grid nodes can be taken directly from this database, obtaining the secondary information to be exported in the kriging process described below. Figure 1 shows the digital elevation maps of each field where weed infestation data were assessed. The elevation range within the area was 276.5 m minimum and 295.6 m maximum in Casilla and between 233.3 m and 246.9 m in Cruce, respectively.

2.3. Prediction methods

The spatial variability of weed species was described by a semivariogram, which expresses the spatial dependence be-tween the weed density at different separation distances and directions (Isaaks and Srivastava, 1989; Webster and Oliver, 2001), using the following equation:

$$\gamma(h) = \frac{1}{2N(h)} \sum_{\alpha=1}^{N(h)} [z(u_\alpha) - z(u_\alpha + h)]^2$$

where $\gamma(h)$ is the experimental semivariogram value at distance interval h; $N(h)$ = number of sample value pairs within the distance interval h; and $z(u_\alpha), z(u_\alpha + h)$ = sample values at two points separated by the distance interval h. All pairs of points separated by the distance h were used to calculate the experimental semivariogram. Several semivariogram functions were evaluated to choose the best fit with the data. Semivariograms were calculated both isotropically and anisotropically by Variowin (Variowin 2.4 Software for Spatial Data Analysis in 2D, Spring Verlag, New York) software. The anisotropic calculations were performed in four directions (0, 45, 90 and 135°) with a tolerance of 22.5° to determine whether semivariogram functions depended on sampling direction, i.e., they were anisotropic or not. The experimental semivariograms were fitted by the least-squares procedure by Variowin software.

Spherical and exponential models were fitted to the experimental semivariograms. The parameters of the model: range, nugget and sill, were determined. Nugget is the variance at zero distance; range is the distance beyond which two measurements become statistically independent and the sill is the asymptotic plateau of the semivariogram function and is used to estimate the range.

The ratio between the nugget and the total semivariance or sill was used to define different classes of spatial dependence for every weed density variable (Cambardella and Karlen, 1999; López-Granados et al., 2002; Jurado-Expósito et al., 2003, 2004, 2005). If the ratio was less than or equal to 25%, the variable was considered as being strongly spatially dependent or strongly distributed in patches; if the ratio was between 25 and 75%, the variable was considered to be moderately spatially dependent, and if the ratio was greater than 75%, the variable was considered weakly spatially dependent.

Semivariogram models were cross-validated (trial-and-error-procedure) to check the validity of the models and to compare values estimated from the semivariogram model with

the actual values (Isaaks and Srivastava, 1989). In this procedure one weed density observation is temporarily deleted at a time from the data set and predicted from remaining data using kriging: this procedure is repeated for all observations of weed densities (Goovaerts, 2000). Differences between estimated and actual values were assessed and compared using the following cross-validation statistics: i.e. mean estimation error, not significantly different than zero; mean squared error less than the variance of the sample values (Hevesi et al., 1992) and standardized mean squared error was within the interval $1 \pm 2 \sqrt{2/n}$ (Isaaks and Srivastava, 1989; Hevesi et al., 1992; Webster and Oliver, 2001). Spatial and geostatistical analyses were performed using Variowin and WinGSlib software; cross-validation was conducted using WinGSLIB software.

Once cross-validated, the parameters of the semivariogram models described above were used to interpolate the weed density variables by using two prediction methods: ordinary point kriging and kriging with an external drift.

2.3.1. Ordinary point kriging

Weed density data were interpolated using the ordinary kriging method. Kriging is a generalized least-squares regression technique that allows one to account for the spatial dependence between observations, as revealed by the semivariogram, into spatial prediction (Goovaerts, 2000). A detailed presentation of the theory of applied geostatistics is given in Isaaks and Srivastava (1989), and Goovaerts (1997).

Ordinary kriging interpolation amounts to estimating the unknown weed density z at the unsampled location u as an unbiased linear combination of observations at neighboring locations (u_α) (Goovaerts, 1997, 2000) according to:

$$z^*_{OK}(u) = \sum_{\alpha=1}^{n(u)} \lambda^{OK}_\alpha(u) z(u_\alpha) \quad \text{with} \quad \sum_{\alpha=1}^{n(u)} \lambda^{OK}_\alpha(u) = 1.$$

The ordinary kriging weights $\lambda^{OK}_\alpha(u)$ are determined such as to minimize the estimation variance, while ensuring the unbiasedness of the estimator via the condition that the sum of kriging weights $\lambda^{OK}_\alpha(u)$ be one (Goovaerts, 2000). The only information required by the kriging system is semivariogram values for different lags, and these are readily derived once a semivariogram model has been fitted to experimental values.

The semivariogram model parameters described above were used to map weed density by ordinary kriging on a regular grid of 2 m. Spatial distribution maps of weed density were constructed using SURFER contour mapping software (SURFER 6.04, Golden Software Inc., 809, 14th Street, Golden, CO 80401-1866, USA), based on WinGSLIB ordinary point kriged estimates.

2.3.2. Kriging with an external drift

In order to improve the prediction, the weed density data were interpolated using kriging with an external drift. This kriging method is an interpolation that uses secondary exhaustive information to derive the local mean or trend m^*_{KED} of the primary attribute z (weed density), then performs simple kriging (SK) on the corresponding residuals (Goovaerts, 1997, 2000):

$$z^*_{KED}(u) - m^*_{KED}(u) = \sum_{\alpha=1}^{n(u)} \lambda^{SK}_\alpha(u) \left[z(u_\alpha) - m^*_{KED}(u_\alpha)\right]$$

where

$$m^*_{KED}(u) = a^*_0(u) + a^*_1(u)\gamma(u).$$

In the kriging with an external drift approach the regression coefficients $a^*_0(u)$ *and* $a^*_1(u)$ are implicitly estimated through the kriging system within each search neighborhood, i.e. the relation between elevation and weed density is assessed locally, which allows one to account for changes in correlation across the study area.

The regression coefficients can be computed and mapped for interpretation purposes (Goovaerts, 1997) but they are not required for estimation. So, the usual and equivalent expression for kriging with an external drift estimate is (Goovaerts, 1997, 2000; Wackernagel, 1998):

$$z^*_{KED}(u) = \sum_{\alpha=1}^{n(u)} \lambda^{KED}_\alpha(u) z(u_\alpha).$$

The kriging weights $\lambda^{KED}_\alpha(u)$ are the solution of the following system of ($n(u) + 2$) linear equations:

$$\begin{cases} \sum_{\beta=1}^{n(u)} \lambda^{KED}_\beta(u)\gamma_R(u\alpha - u\beta) + \mu^{KED}_0(u) \\ \qquad + \mu^{KED}_1(u)\gamma(u_\alpha) = \gamma_R(u_\alpha - u) \quad \alpha = 1, ..., n(u) \\ \sum_{\beta=1}^{n(u)} \lambda^{KED}_\beta(u) = 1 \\ \sum_{\beta=1}^{n(u)} \lambda^{KED}_\beta(u)\gamma(u\beta) = \gamma(u) \end{cases}$$

where $\mu^{KED}_0(u)$ and $\mu^{KED}_1(u)$ are two Lagrange parameters. In practice the residual semivariogram (γ_R) could be inferred from pairs of z values that are unaffected or slightly affected by the trend, i.e. from data pairs such that $\gamma(u_\alpha) \approx \gamma(u_\alpha + h)$. This condition is generally satisfied for small separation distances. Thus, the residual semivariogram may be identified with the corresponding z-semivariogram $\gamma(h)$ (Hudson and Wackernagel, 1994; Wackernagel, 1998; Goovaerts, 1999b, 2000).

The geostatistical estimation method of kriging with an external drift requires a linear relationship between the variable of interest, i.e., primary variable, weed density and the secondary variable, elevation. If this is not the case, an appropriate transformation of the secondary variable could make that relation linear (Goovaerts, 1999b). Pearson linear correlations were established between weed density and elevation values, accepting a confidence level of 95%.

Table I. Summary of statistical data of main weed species sampled in Casilla and Cruce.

Location	Weed	Mean (plants m^{-2})	Maximum	Variance	Standard Deviation	Skew
Casilla	*Convolvulus arvensis*	2.31	51.0	59.04	7.68	1.50
(n^* = 499)	*Ridolfia segetum*	0.54	12.0	1.88	1.37	1.41
	Phalaris spp.	3.31	51.0	55.90	7.48	1.43
	Sinapis arvensis	2.36	51.0	10.05	3.17	0.84
	Amaranthus albus	16.32	51.0	479.60	21.90	0.55
Cruce	*Convolvulus arvensis*	5.29	51.0	85.90	9.27	1.60
(n^* = 256)	*Ridolfia segetum*	2.17	40.0	27.40	5.23	1.20
	Phalaris spp.	18.6	51.0	193.50	13.90	1.10
	Galium aparine	0.69	10.0	1.90	1.38	1.93
	Chrozophora tinctoria	5.82	40.0	45.50	6.75	1.05

* *n*: number of georeferenced counting units.

In this paper, the elevation data, available at all estimation grid nodes, were introduced as exhaustive secondary information into the kriging system to estimate weed density in the sunflower fields. This method was applied to the weed species showing absolute correlations $|r| > 0.2$ with elevation values. Spatial distribution maps of those weed species density were constructed using SURFER, based on WinGSLIB kriging with an external drift estimates.

2.3.3. Comparison between prediction methods

According to Goovaerts (2000), the prediction quality of the two interpolators used, ordinary kriging and kriging with an external drift, was assessed and compared using cross-validation. The comparison criterion used was the mean square error of prediction, which measures the average square difference between the true density observation $z(u_\alpha)$ and its estimate $z^*(u_\alpha)$ which is based on all other observations except $z(u_\alpha)$:

$$\text{MSE} = \frac{1}{n} \sum_{\alpha=1}^{n} \left[z(u_\alpha) - z^*(u_\alpha) \right]^2$$

where *n* is the number of weed density data (Goovaerts, 2000; Bishop and McBratney, 2001; López-Granados et al., 2005).

For each prediction method, the mean square error of each species was calculated as an overall indication of the map precision quality. Mean square error should be less than the variance of sample values, otherwise the prediction method would be worse than predicting the weed density by the overall mean equally at all locations, and smaller values of mean square error indicate a greater accuracy than larger values (Goovaerts, 2000).

3. RESULTS AND DISCUSSION

3.1. Prediction methods

Table I records mean, variance, standard deviation and skewness of weed species. Skewness ranged between +2 and –2 and therefore no transformation of data was carried out.

A. albus in Casilla, and *Phalaris* spp. in Cruce were by far the two most dominant species and had the highest density values, with a mean of 16.32 and 18.6 plants m^{-2}, respectively. *R. segetum* in Casilla, and *G. aparine* in Cruce were present at very low density: 0.54 plants m^{-2} and 0.69 plants m^{-2}, respectively. In Cruce, *C. arvensis* and *C. tinctoria* were present at a moderate density: 5.3 and 5.8, respectively.

Directional semivariograms did not show differences in spatial dependence with directions, therefore isotropic semivariograms were chosen. Spherical and exponential isotropic models were defined for all weed species studied for both fields. Figure 2 shows some of the experimental semivariograms of weed density with the model fitted. These models were used to generate weed density maps. The semivariogram parameters and associated levels of spatial dependence are listed in Table II. Weed species displayed differences in spatial dependence as determined by semivariogram analyses. Semivariogram parameters strongly varied among species and between fields. Range is the distance beyond which spatial dependence between weed density samples ceases to exist and it can be used as an indicator of the maximum cell size for a field survey in site-specific management. Thus, range is important to establish the sampling interval for future surveys. According to Kerry and Oliver (2005), the sampling interval should be less than half the semivariogram range and one should choose additional sites at half that range interval.

The sill parameter gives the variance of weed densities separated by a distance greater than the range, where spatial dependence ceases to exist. The sill strongly varied among species and between locations (Tab. II). For example, for *C. arvensis* the sill values were 28.3 and 46.4 in Casilla and Cruce, respectively. A similar variation was observed for the sill variance of *Phalaris* spp., which was also present in both locations.

The nugget effect was greater than zero in all cases, meaning that observations separated by small distances were dissimilar (Isaaks and Srivastava, 1989). The nugget variance expressed as a percentage of the sill was used to define different classes of spatial dependence for weed densities (Tab. II). Low nugget ratios, from 9.5–16%, indicating a strong spatial dependence, were found for all weed species sampled in Casilla
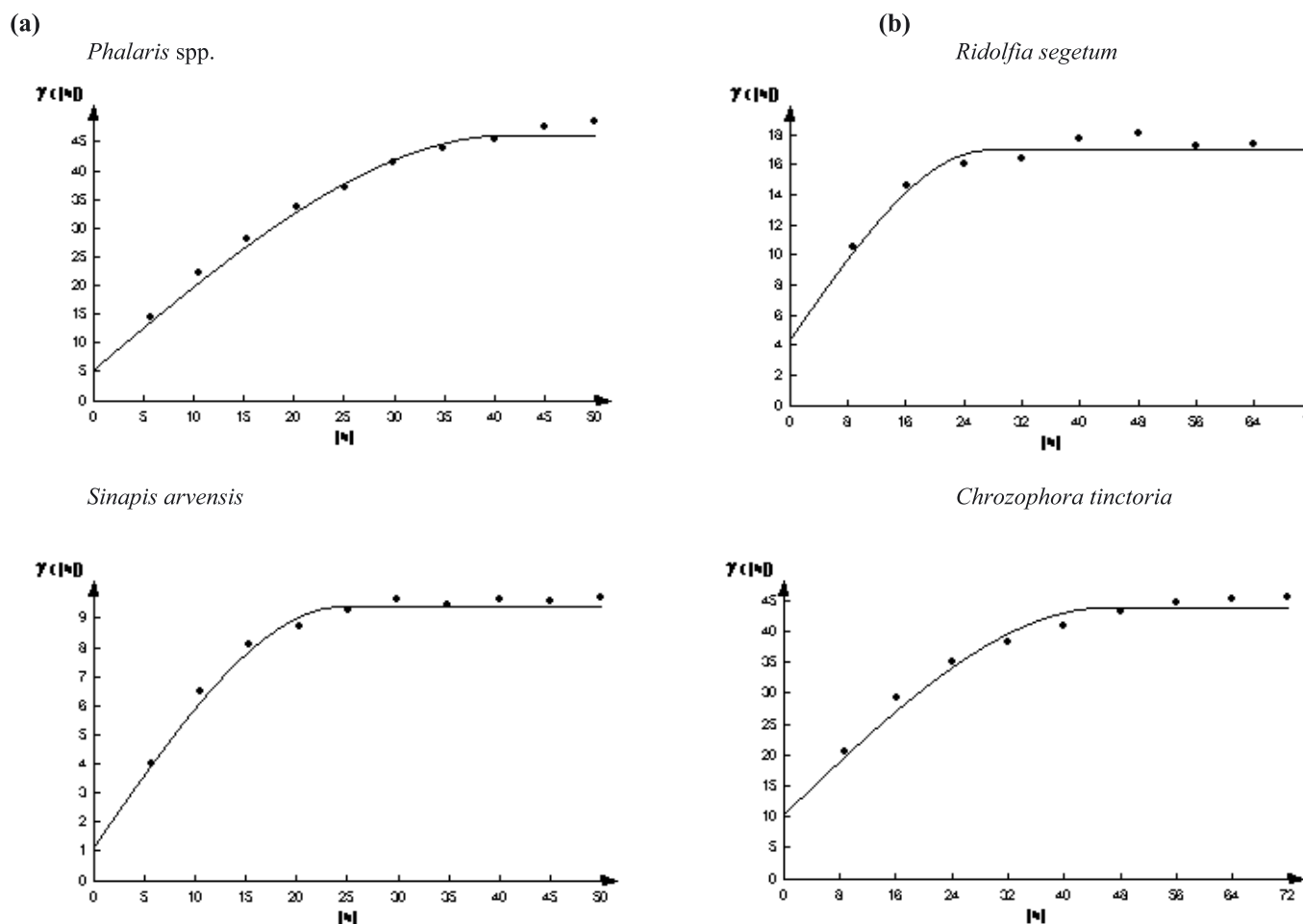
(a)

Phalaris spp.

Sinapis arvensis

(b)

Ridolfia segetum

Chrozophora tinctoria

Figure 2. Experimental (●) and modeled (-) semivariograms of weed density (plants m^{-2}) corresponding to sampling in (a) Casilla and (b) Cruce. Axis X is lag distance (m).

Table II. Spatial distribution characteristics of main weed species.

Location	Weed	Model	Range (m)	Nugget	Sill	Nugget ratio[*] (%)	Spatial class[†]
Casilla	*Convolvulus arvensis*	Spherical	22.7	27.10	28.3	95.75	W
(n[§] = 499)	*Ridolfia segetum*	Exponential	28.8	0.24	1.5	16.00	S
	Phalaris spp.	Spherical	41.5	5.60	40.9	13.70	S
	Sinapis arvensis	Spherical	24.5	1.21	8.2	14.70	S
	Amaranthus albus	Spherical	44.0	33.60	355.2	9.50	S
Cruce	*Convolvulus arvensis*	Spherical	30.1	31.80	46.4	66.80	M
(n[§] = 256)	*Ridolfia segetum*	Spherical	27.3	4.50	12.6	35.70	M
	Phalaris spp.	Exponential	49.7	44.00	126.0	34.90	M
	Galium aparine	Spherical	27.5	0.80	0.8	94.10	W
	Chrozophora tinctoria	Spherical	46.1	10.60	33.6	31.50	M

[*] Nugget ratio = (Nugget semivariance/total semivariance) × 100.
[†] Spatial class: S = strong spatial dependence; M = moderate spatial dependence; W = weak spatial dependence.
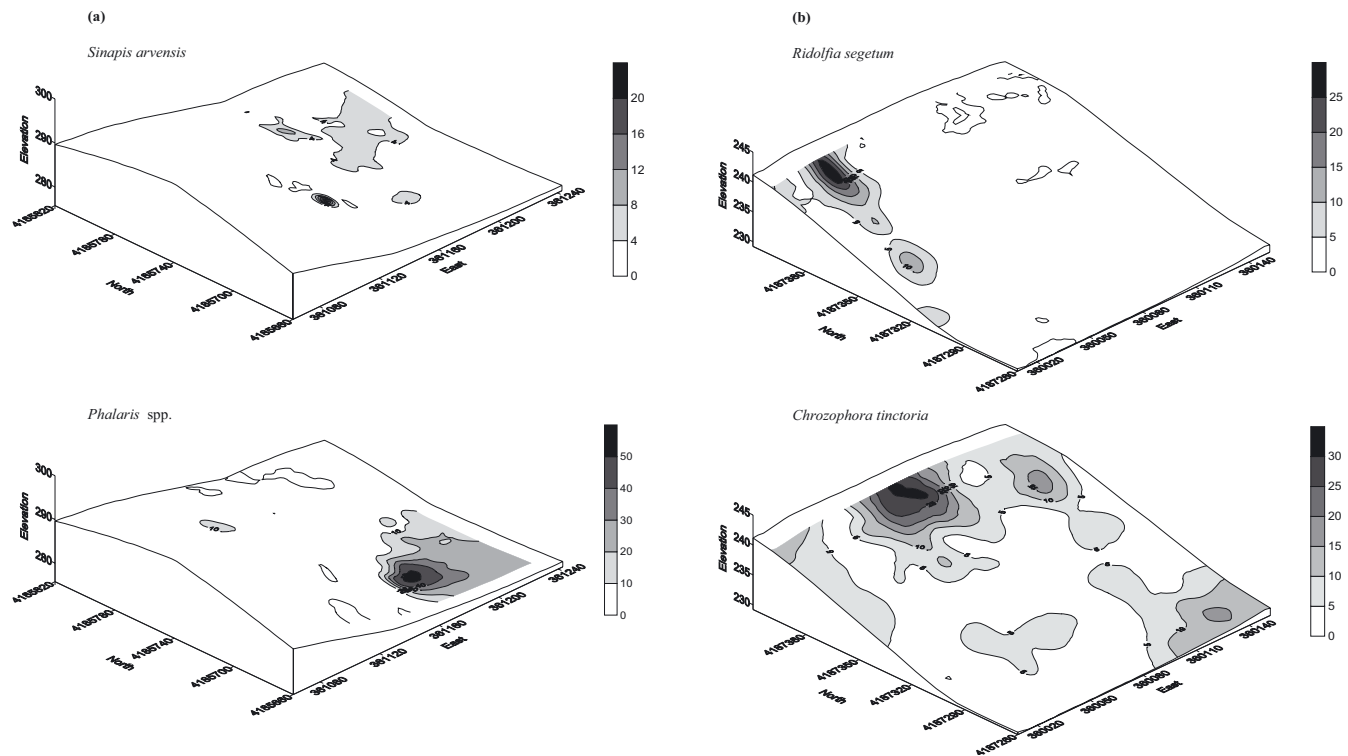[§] n: number of georeferenced counting units.

Figure 3. Weed density maps (plants m^{-2}) obtained using ordinary point kriging in (a) Casilla and (b) Cruce. Axes are in Universal Transverse Mercator units (m), Z: elevation (m).

with the exception of *C. arvensis,* which presented a high nugget ratio, 95.7%, indicating its weak spatial dependence. In Cruce, medium nugget ratios, from 31.5–66.8%, were found for most of the weeds sampled, indicating moderate spatial dependence.

3.1.1. Ordinary point kriging

Some of the maps from the ordinary kriging estimates for each weed species are shown in Figure 3a for Casilla and Figure 3b for Cruce, respectively. A visual assessment reveals a moderate to strong spatial dependence; all the maps generally indicated a patchy distribution, which is also supported by the semivariogram analyses.

3.1.2. Kriging with an external drift

The Pearson linear correlations, r, between weed density data and elevation data ranged from –0.590 to 0.290 (Tab. III), not far from 0. However, the sample sizes were very high, so these empirical correlation values sufficed to prove statistically ($\alpha = 0.05$) that the correlations between these two variables were different from zero, except the species *C. arvensis* in Casilla and *C. arvensis*, *R. segetum* and *G. aparine* in Cruce, where we could not prove this.

Figure 4 shows experimental semivariograms of elevation computed from the digital elevation model of Figure 1, with

Table III. Pearson linear correlation coefficients between weed density (plants m^{-2}) and elevation (m).

Location	Weeds	Coefficients
Casilla	*Convolvulus arvensis*	0.020^{ns}
($n^{\dagger} = 499$)	*Ridolfia segetum*	0.296**
	Phalaris spp.	–0.235**
	Sinapis arvensis	–0.243**
	Amaranthus albus	–0.590**
Cruce	*Convolvulus arvensis*	0.115^{ns}
($n^{\dagger} = 256$)	*Ridolfia segetum*	0.071^{ns}
	Phalaris spp.	–0.357**
	Galium aparine	0.033^{ns}
	Chrozophora tinctoria	0.290**

† n: number of georeferenced counting units.
* , ** Significant at the 0.05 and 0.01 levels, respectively; ns: not significant.

the model fitted. They show that there are consistencies in the semivariogram ranges of the weed density sampled and elevation. In general, to perform kriging with an external drift for any two properties they should be at least moderately correlated and have semivariograms with similar ranges.

Weeds showing absolute correlations $|r| > 0.2$ with elevation were considered for kriging with an external drift, i.e., *R. segetum*, *Phalaris* spp., *S. arvensis* and *A. albus* in Casilla
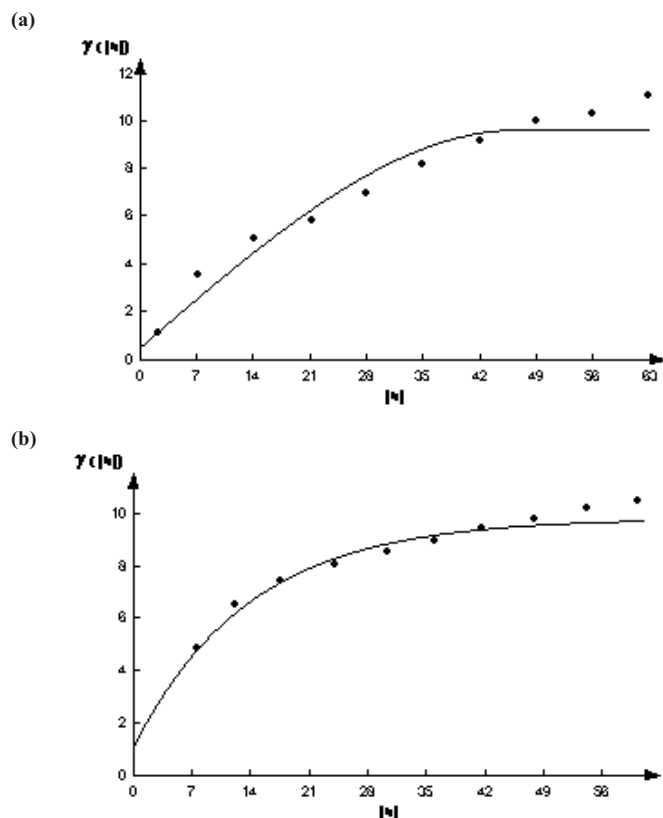
(a)

(b)

Figure 4. Experimental (●) and modeled (-) semivariograms of elevation computed from the Digital Elevation Model in (a) Casilla and (b) Cruce. Axis *X* is lag distance (m).

and *Phalaris* spp. and *C. tinctoria* in Cruce. Figures 5a and 5b show the maps of weed density estimates obtained using KED for Casilla and Cruce, respectively. There is a similarity in the pattern of weed density as produced by ordinary kriging (Figs. 3a, b). However, ordinary kriging over-smoothed the spatial variability of weed density.

3.1.3. *Comparison between prediction methods*

The cross-validation statistics for the prediction methods compared are listed in Table IV. Mean square error was used to compare the prediction performances of the two geostatistical interpolation algorithms. Generally, the estimation method using elevation gave more favorable mean square error results than the interpolation method using only weed density data, ordinary kriging, indicating that the correlation of weed density with elevation data is very important for mapping weed density. Thus, the least accurate estimation for ordinary kriging with the highest mean square error values was because elevation data were ignored and only the spatial component of weeds was considered.

R. segetum and *C. tinctoria* showed positive correlations with elevation, which means that these weeds were dominant in areas with high elevation values. On the other hand,

Phalaris spp., *S. arvensis* and *A. albus* showed negative correlations with elevation, so they were predominant in areas with low values of elevation. The results of this work indicate that kriging with an external drift may become useful for describing weed infestation maps according to the elevation of the field, which could be used to design intermittent spraying strategies.

Bishop and McBratney (2001) also found that KED was the best prediction method for mapping soil properties using bare soil aerial photographs as secondary information. So, when secondary information is available, it should be used because generic geostatistical techniques that only use the primary variable, such as ordinary kriging, do not obtain the prediction performance of methods incorporating that secondary information.

The mean square errors for the different approaches (Tab. IV) show that the generic geostatistical technique, ordinary kriging, exhibited the highest mean square error values because it ignores the secondary information and only used the primary weed density variable. The kriging with an external drift mean square error results for *R. segetum*, *Phalaris* spp., *S. arvensis* and *A. albus* of 0.93, 13.9, 8.2 and 99.1 in Casilla indicated 16%, 11%, 17% and 9% improvements in performance relative to ordinary kriging, respectively, and with regards to *Phalaris* spp. and *C. tinctoria* in Cruce, mean square error values of 82.06 and 21.6 indicated 13% and 8% improvements in performance relative to ordinary kriging, respectively. These improvements are due to the relationship between the primary and secondary variables being evaluated locally, which permits taking into account changes in the correlation across the sampling area. (Hevesi et al., 1992).

According to López-Granados et al. (2005), our results indicated that even when a moderately correlated ($r < 0.5$) secondary attribute is available, the methods incorporating this secondary variable into the map of the primary variable, i.e., kriging with an external drift, perform better than other approaches which only incorporate the target weed density variable, such as ordinary kriging.

4. CONCLUSION

In conclusion, the estimation method using a digital elevation model gave more favorable mean square error results than prediction methods using only a primary variable. These results suggest that kriging with an external drift of weed density data with elevation as a secondary exhaustive variable could be used in such situations, and in this way, the accuracy of maps for precision agriculture could be improved with little or no additional cost, since the digital elevation model could be obtained as part of other analyses.

The growing interest of weed scientists in multivariate geostatistics has arisen because they realize more and more that quantitative spatial prediction should incorporate the spatial correlation among observations. Also, geostatistics offers an increasingly wide range of techniques well suited to the diversity of problems and information that weed scientists have to deal with. Multivariate geostatistical interpolation such as
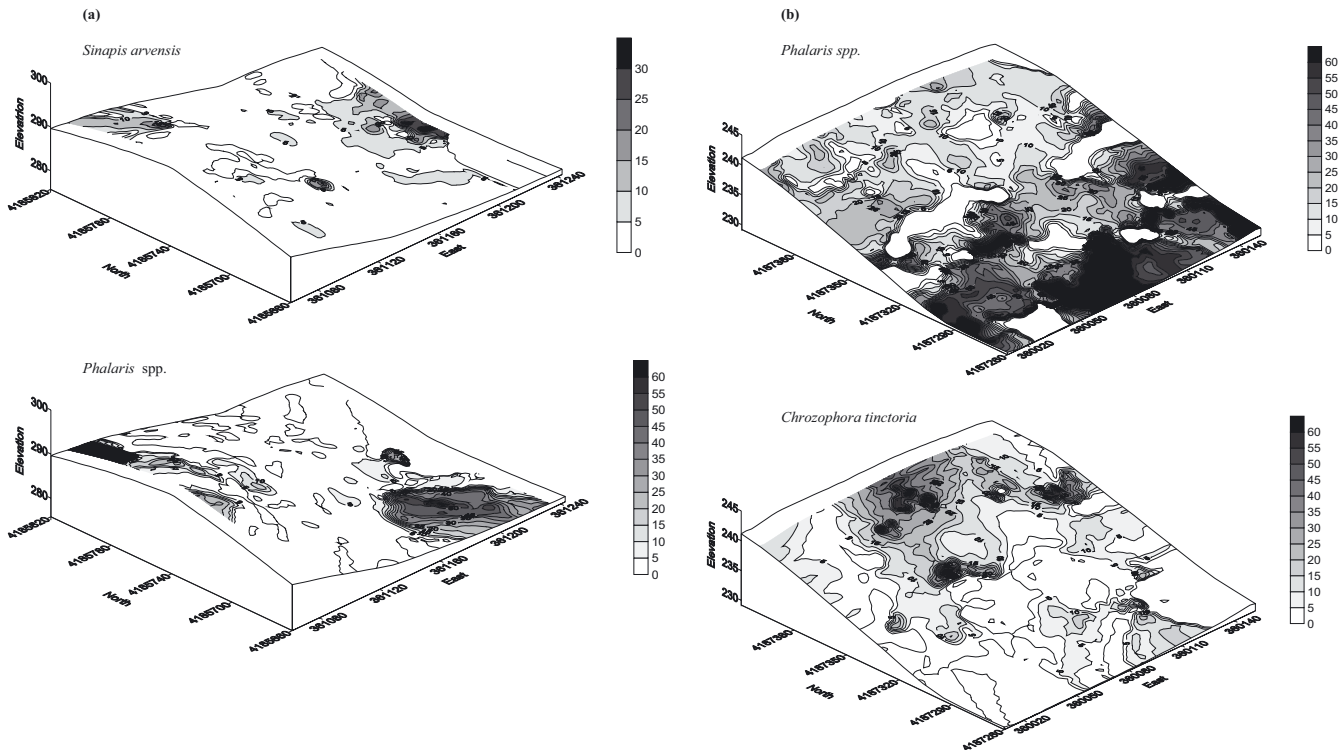
Figure 5. Weed density maps (plants m^{-2}) obtained using kriging with an external drift in (a) Casilla and (b) Cruce. Axes *X* and *Y* are in Universal Transverse Mercator units (m), *Z*: elevation (m).

Table IV. Cross-validation statistics for the prediction methods compared when estimating weed density.

Location	Weed	OK*			KED†		
		MEE‡	MSE§	SMSE#	MEE	MSE	SMSE
Casilla							
(n & = 499)	*Ridolfia segetum*	0.000	1.11	0.82	−0.01	0.93	0.79
	Phalaris spp.	0.020	15.60	1.08	−0.07	13.90	1.06
	Sinapis arvensis	0.010	9.90	1.67	−0.04	8.20	1.06
	Amaranthus albus	−0.004	108.90	1.08	−0.09	99.10	1.01
Cruce							
(n = 256)	*Phalaris* spp.	0.040	94.53	0.78	−0.08	82.06	0.80
	Chrozophora tinctoria	−0.040	23.61	1.06	−0.08	21.60	1.00

* OK: Ordinary Point Kriging.
† KED: Kriging with an External Drift.
‡ MEE: Mean Estimation Error.
§ MSE: Mean Squared Error.
# SMSE: Standardized Mean Squared Error.
& *n*: number of georeferenced counting units.

kriging with an external drift allows one to complement a few expensive measurements of the attribute of interest, e.g., weed density with more abundant data on correlated attributes that are cheaper to determine, e.g., elevation. There is still research to be done on the incorporation of secondary variables measured on different supports, in particular the combination of field data, e.g. yield data, weed or soil sampling, the primary variable, with reflectance data from remote-sensing images, the secondary variable.

This study demonstrated that maps from kriging with an external drift estimates combining geostatistical techniques and elevation data were accurate enough to improve the identification of weed density patches, which is the first step for site-specific weed management.

REFERENCES

Bishop T.F.A., McBratney A.B. (2001) A comparison of prediction methods for the creation of field-extent soil property maps, Geoderma 103, 149–160.

Bregt A.K., Gesing H.J., Alkasuma M. (1992) Mapping the conditional probability of soil variables, Geoderma 53, 15–29.

Cambardella C.A., Karlen D.L. (1999) Spatial analysis of soil fertility parameters, Precis. Agr. 1, 5–11.

Cressie N.A.C. (1991) Median polish kriging, in: Cressie N.A.C. (Ed.), Statistics for Spatial Data, John Wiley and Sons, New York, pp. 183–199.

Ersbøll A.K., Kristensen K., Nordbo E., Christensen S. (1993) Estimating the number of weed plants using kriging, in: Boelskifte S. (Ed.), Symposium in Anvendt Statistik, Ålborg, Denmark, UNIC, pp. 140–151.

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

Goovaerts P. (1997) Geoestatistics for Natural Resources Evaluation, Oxford University Press, New York.

Goovaerts P. (1999a) Geostatistics in soil science: state-of-art and perspectives, Geoderma 89, 1–45.

Goovaerts P. (1999b) Using elevation to aid the geostatistical mapping of rainfall erosivity, Catena 34, 227–242.

Goovaerts P. (2000) Geostatistical approches for incorporating elevation into the spatial interpolation of rainfall, J. Hydrol. 228, 113–129.

Heisel T., Andersen C., Ersbøll A.K. (1996) Annual weed can be mapped with kriging, Weed Res. 36, 325–337.

Hevesi J.A., Istok J.D., Flint A.L. (1992) Precipitation estimation in mountains terrain using multivariate geostatistics. Part I: structural analysis, J. Appl. Meteorol. 31, 661–676.

Hudson G., Wackernagel H. (1994) Mapping temperature using kriging with and external drift: theory and an example from Scotland, Int. J. Climatol.14, 77–91

Isaaks E.H., Srivastava R.M. (1989) An Introduction to Applied Geostatistics, Oxford University, New York.

Journel A.G., Huijbregts C.J. (1978) Mining Geostatistics, Academic Press, London, UK.

Jurado-Expósito M., López-Granados F., García-Torres L., García-Ferrer A., Sánchez de la Orden M. (2003) Multi-species weed spatial variability and site-specific management maps in cultivated sunflower, Weed Sci. 51, 319–328.

Jurado-Expósito M., López-Granados F., González-Andujar J.L., García-Torres L. (2004) Spatial and temporal analysis of *Convolvulus arvensis* L. populations over four growing seasons, Eur. J. Agron. 21, 287–296.

Jurado-Expósito M., López-Granados F., González-Andujar J.L., García-Torres L. (2005) Characterizing population growth rate of *Convolvulus arvensis* in wheat-sunflower no tillage systems, Crop Sci. 45, 2106–2112.

Kerry R., Oliver M. (2005) Variograms of ancillary data of aid sampling for soil surveys, Precis. Agr. 4, 261–278.

Kristensen K., Ersbøll A.K. (1995) The use of geostatistics methods in variety trials where some variety is unreplicated, in: Proceeding Fifth Working Seminar on Statistical Methods in Variety Testing, Zakopane, Poland, pp. 12–16.

Ley R.G. (1986) Accuracy assessment of Digital Terrain Models, in: Proceedings of AutoCarto, AutoCarto London, London, UK, pp. 455–464.

Liu Z., Clay S.A., Clay D.E. (2002) Spatial variability of atrazine and alachlor efficacy and mineralization in an eastern South Dakota field, Weed Sci. 50, 662–671.

López-Granados F., Jurado-Expósito M., Atenciano S., García-Ferrer A., Sanchez de la Orden M., García-Torres L. (2002) Spatial variability of agricultural soil parameters in southern Spain, Plant Soil 246, 97–105.

López-Granados F., Jurado-Expósito M., Peña-Barragán J.M., García-Torres L. (2005) Using geoestatistical and remote sensing approaches for mapping soil properties, Eur. J. Agron. 23, 279–289.

MAYPA (2005) MAYPA (Spanish Ministry of Agriculture and Fischeries), Available at http://www.maypa.es/es/estadistica/pags/superficie/superficie/.htm.

Shafii B., Price W.J., Prather T.S., Lass L.W., Thill D.C. (2004) Using landscape characteristics as prior information for Bayesian classification of yellow starthistle, Weed Sci. 52, 948–953.

Vanderlinden K. (2001) Análisis de procesos hidrológicos a diferentes escalas espacio-temporales, Tesis Doctoral, Universidad de Córdoba, Spain, pp. 65–125.

Wackernagel H. (1998) Multivariate geostatistics, 2nd ed., Springer, Berlin.

Webster R., Oliver M.A. (2001) Geostatistics for environmental scientists, John Wiley and Sons, LTD, Chichester, England.

Zanin G., Berti A., Riello L. (1998) Incorporation of weed spatial variability into the weed control decision-making process, Weed Res. 38, 107–118.