Article

Spatial Analysis of Digital Imagery of Weeds in a Maize Crop

Carolina San Martín [1],* [iD], Alice E. Milne [2], Richard Webster [2], Jonathan Storkey [2], Dionisio Andújar [3], Cesar Fernández-Quintanilla [3] and José Dorado [3]

[1] Columbia Basin Agricultural Research Center, Oregon State University, 48037 Tubbs Ranch Rd., Adams, OR 97810, USA

[2] Rothamsted Research, Harpenden, Hertfordshire AL5 2JQ, UK; alice.milne@rothamsted.ac.uk (A.E.M.); richard.webster@rothamsted.ac.uk (R.W.); jonathan.storkey@rothamsted.ac.uk (J.S.)

[3] Department of Crop Protection, Instituto de Ciencias Agrarias, CSIC, Serrano 115 B, 28006 Madrid, Spain; dionisioandujar@hotmail.com (D.A.); cesar@ica.csic.es (C.F.-Q.); jose.dorado@ica.csic.es (J.D.)

* Correspondence: sanmartc@oregonstate.edu

Abstract: Modern photographic imaging of agricultural crops can pin-point individual weeds, the patterns of which can be analyzed statistically to reveal how they are affected by variation in soil, by competition from other species and by agricultural operations. This contrasts with previous research on the patchiness of weeds that has generally used grid sampling and ignored processes operating at a fine scale. Nevertheless, an understanding of the interaction of biology, environment and management at all scales will be required to underpin robust precise control of weeds. We studied the spatial distributions of six common weed species in a maize field in central Spain. We obtained digital imagery of a rectangular plot 41.0 m by 10.5 m (= 430.5 m^2) and from it recorded the exact coordinates of every seedling: more than 82,000 individuals in all. We analyzed the resulting body of data using three techniques: an aggregation analysis of the punctual distributions, a geostatistical analysis of quadrat counts and wavelet analysis of quadrat counts. We found that all species were aggregated with average distances across patches ranging from 3 cm–18 cm. Species with small seeds tended to occur in larger patches than those with large seeds. Several species had aggregation patterns that repeated periodically at right angles to the direction of the crop rows. Wheel tracks favored some species (e.g., thornapple), whereas other species (e.g., johnsongrass) were denser elsewhere. Interactions between species at finer scales (<1 m) were negligible, although a negative correlation between thornapple and cocklebur was evident. We infer that the spatial distributions of weeds at the fine scales are products both of their biology and local environment caused by cultivation, with interactions between species playing a minor role. Spatial analysis of such high-resolution imagery can reveal patterns that are not immediately evident from sampling at coarser scales and aid our understanding of how and why weeds aggregate in patches.

Keywords: weeds; spatial distribution; aggregation; variance:mean ratio; geostatistics; variogram; wavelet analysis; plant competition

---

1. Introduction

Weeds tend to be distributed unevenly in arable fields. In many instances, their patterns and their spatial scales are obvious; in some, the farmers see patterns that are not obvious to the lay observer; in yet others, there are patterns that require detailed analysis to discern and quantify. Identification of the patterns is becoming increasingly important as farmers try to economize on the use of herbicides, spraying only where weeds are dense ('precision farming'), and as environmental agencies try to prevent excess herbicide reaching water courses and ground water [1–3]. Farmers and their advisors

may also wish to know how cultivation affects weed patterns, and biologists have a further interest in the interactions between species in the field and how species compete for nutrients, water and light. Recent advances in computing power and image analysis enable us to identify patterns and their spatial scales that are not immediately discernible to the naked eye or from grid sampling of quadrats at a coarse scale [4–7]. While an increased understanding of the processes that explain the aggregation of weeds at multiple scales is, in the first instance, of theoretical interest, we believe it will be necessary for the implementation of sensor technology in the precise control of weeds on farms.

Many factors affect the spatial distributions of weeds. They include farming practices (such as crop rotation, herbicide application, cultivation and the kinds of farm machinery used), interactions between species and variations in the soil. Investigations of interactions between weeds and soil have tended to concentrate on soil attributes that encourage weed growth and relations at coarse scales, i.e., field scale [8,9]. A few investigators have incorporated more than one scale in their studies and found that relationships are maintained over several spatial scales [10,11]. One reason studies that compare processes at multiple scales are rare is that the surveys have been done on counts of individual weed plants in quadrats in the field, usually on regular grids. As pointed out by Rew and Cousens [10], if the grid interval (i.e., the distance between quadrat centers) is short, then the technique is very laborious. If the grid interval is large, then one obtains no information at short distances: the distance between quadrat centers limits the finest scale at which associations can be analyzed. To obtain information over a wide range of spatial scales with modest labor, Metcalfe et al. [11] designed a nested sampling scheme, whereby quadrats were placed in the field and weeds counted in them. The scheme enabled them to summarize the spatial distributions statistically, but the uneven coverage was no basis for mapping, which the farmer would need for precise control of weeds and may have missed processes operating at a fine scale (<1 m).

As well as variability in soil, farming management also influences weed distribution. However, whereas the effects of soil and interspecific competition appear to be random, those arising from cultivation are likely to be systematic, aligned in particular directions and be repeated periodically in ways determined by the scale of the cultivation equipment. Dieleman et al. [8,9] pointed out a possible effect of field practices on the annual variation of the relations between weeds and soil. Agricultural machinery locally modifies the environment of seeds, influencing dormancy and germination [12–14]. Tardif-Paradis et al. [15] and Longchamps et al. [16] have recorded the effect of wheel tracks of maize planters and tractors on weed populations. Both sets of authors found greater weed density, biomass and cover in wheel tracks and disturbed rows than in undisturbed soil between the rows. Evidently, the combination of disturbance and traffic had favored the weeds.

Other factors influencing weed distribution at fine scales depend to some degree on the mechanisms of dispersal and competition (Table 1). Perennial species (such as johnsongrass and nutsedge), which spread vegetatively, tend to occur in patches [17], whereas annual species disperse their seeds in more haphazard ways, and their patchiness is less pronounced [18]. We also expect that the spread of annual species with small seeds is less constrained by local variations in the environment. The kind of dispersal might determine the final distribution. Thus, species with seeds enclosed in capsules that fall to the ground (such as cocklebur and thornapple) are likely to have a more aggregated pattern than species with seeds that are dispersed by wind. In addition, combine harvesting can affect the dispersal of seeds [19,20] with patches becoming elongated in the direction of travel; again, this will be related to weed phenology and seed characteristics. Competition among seeds and seedlings concentrated in small areas result in them thinning themselves, and so, they are limited within micro-sites [21]. Validated hypotheses of how weed biology explains differences between species in the extent and scale of aggregation are rare in the literature, in part because of the lack of data at fine scales. We expect these biological processes to be expressed at scales finer than those detected by commonly-used protocols on grids or nested designs.

Table 1. Characteristics of the weed species *.

Species	Biotype	Fruit Type	Seed Size/mm	Means of Dispersal	1000 Seed Weight/g
Nutsedge	Perennial	Achene	1.2–1.5 × 0.4–0.7	Vegetative (rhizomes)	0.23
Johnsongrass	Perennial	Achene	4.5 × 2	Vegetative (rhizomes)	4.9
Velvetleaf	Annual	Capsule	2–3 × 2	Gravity and agriculture	9.0
Thornapple	Annual	Capsule	3.8–4.6 × 3.0–3.8	Gravity and agriculture	14.5
Nightshade	Annual	Berry	1.5–2.0 × 1.4–1.8	Zoochory	0.8
Cocklebur	Annual	Bur	5–15 × 5–7	Zoochory and agriculture	203.4

* Information from the Royal Botanic Gardens Kew Seed Information Database (SID) [22].

Our Investigation

In 2011, we obtained digital imagery of a 430.5-m^2 rectangular plot in a maize field in central Spain shortly after the maize itself had emerged. The imagery recorded precisely the positions of all the seedlings of six weed species (a total of more than 82,000 individuals). Visual inspection of the imagery suggested some weak patterning in the direction of cultivation; but were we seeing what we expected to see? Our aim was to analyze this large body of data to find out. Were seedlings aggregated? If so, then at what spatial scale? Did patchiness, if present, recur periodically in tune with the pattern of cultivation? Is there quantitative evidence that any one species inhibits or encourages others? Our hypothesis was that differences in the aggregation of weeds at fine scales could be related to their contrasting modes of dispersal (large seeds of perennial weeds versus small seeds of annuals). To answer, we built on the technique already used by San Martín et al. [23] to describe the spatial dynamics of johnsongrass in short-rotation poplar coppices, exploiting the rich detail of the digital imagery. It involved three broad groups of techniques:

1. Greig Smith's [24] test of departure from randomness,
2. standard geostatistical variography of quadrat counts and
3. wavelet analysis.

We describe these in turn after introducing the data.

2. Materials and Methods

2.1. Field of Study

The region studied was a rectangular plot 41.0 m by 10.5 m (= 430.5 m^2) in a 4-ha maize field at La Poveda Experimental Farm, Arganda del Rey, Madrid, Spain (40.31 ° N, 3.49 ° W). The climate is Mediterranean Continental with cold winters, hot summers and a mean annual rainfall of about 400 mm. The soil is a sandy loam (Eutric Fluvisol [25]) with 1.6% organic matter in the plow layer. In 2011, the field was conventionally tilled, and maize was planted on 1 April in rows 0.75 m apart to give a population of 85,000 plants per hectare. Maize had been grown continuously on this field for the previous nine years, with conventional tillage and sprinkler irrigation. The field was tilled in straight lines in the direction of the longest side of the plot (i.e., in the 41.0-m direction). Although herbicides had been applied in in previous years, none were applied in 2011. The field was heavily infested with four annual weeds, namely common cocklebur (*Xanthium strumarium* L.), fierce thornapple (*Datura ferox* L.), velvetleaf (*Abutilon theophrasti* Medik.) and black nightshade (*Solanum nigrum* L.), and two perennial weeds (johnsongrass (*Sorghum halepense* L. Pers.) and purple nutsedge (*Cyperus rotundus* L.).

2.2. Weed Imaging

We obtained digital imagery of the plot by recording images of contiguous rectangles (1148 of them 0.75 m × 0.50 m) and assembling them into a complete image. For this, we used a D70 Nikon digital camera held 1.5 m above the ground pointed vertically down. The longest side of the rectangles fit between the rows of maize (0.75 m) with sides lying in the rows. Images were obtained

on 1 May at the 2–4-leaf stage of the crop, which is the optimal time for post-emergence weed control. The weeds themselves were at various stages of growth having from 1–4 true leaves. The species of each and every individual weed plant was identified and its geographic coordinates recorded. Using ArcGIS® 10.3 [26], we built a file containing the coordinates for all six weed species. These data were the source for the analyses we describe below.

2.3. Statistical Analysis

2.3.1. Aggregation Analysis

One way of quantifying patchiness and its spatial scale is that due to Greig-Smith [24]. Greig-Smith pointed out that if individual plants are distributed at random, their counts in quadrats have a Poisson distribution with the mean equal to the variance. A patchy distribution has a variance greater than its mean. Greig-Smith built on this fact to identify the average size of patches by varying the sizes of quadrats, counting the individuals in each and calculating the variance:mean ratio. The size of quadrat with the largest ratio corresponds roughly to the average size of the patches. This method is one way by which one can discover the spatial scale(s) at which plants are aggregated.

We used the method to see whether the distribution of each species differed significantly from random. We counted the number of plants of each species in quadrats ranging in size from 1 cm × 1 cm–150 cm × 150 cm in steps of 1 cm and from 150 cm × 150 cm–10 m × 10 m in steps of 50 cm. The side of the last was almost equal to the smaller side of the plot.

We calculated the variance:mean ratio for each size of quadrat and tested the departure of this ratio from 1; and hence, the departure of the spatial distribution from pure randomness. The statistical significance of a departure from random can be obtained from the index of dispersion:

$$I \; = \; (N-1)\frac{s^2}{\overline{z}} \, , \tag{1}$$

where s^2 is the variance of the counts, $\overline{z}$ is the mean of the counts and N is the number of observations. This value is compared with χ^2 with $N-1$ degrees of freedom.

To investigate whether the wheel tracks and crop rows influenced the density of weeds present in each area, we defined three zones (Figure 1):

(1)   inter-rows without wheel tracks (IR);
(2)   inter-rows with wheel tracks (WT); and
(3)   crop rows (CR).

The maize was sown with a four-row planter with rows 75 cm apart. The planter was pulled by a tractor with tires 38 cm wide and 1.44 m between the centers of the wheels. According to these dimensions, we considered 30-cm wide transects 75 cm apart for CR zones and 30-cm wide transects 150 cm apart for IR and WT zones (see Figure 1). In Figure 2, one can see faint stripes running the length of the plot, and it seems likely that these correspond to the positions of wheel tracks and crop rows. In the Supplementary Material, we provide a map indicating the exact position of wheel tracks in the experimental plot (Figure S1).

The proportions of the plot allocated to each zone were 20% for IR (i.e., seven strips), 20% for WT (i.e., seven strips) and 40% for CR (i.e., 15 strips). For each species, we constructed bar charts showing the densities of each species within each area. Note that these densities are the 'population' densities of the species; they are not estimates based on samples. There is no error associated with the counts, and statistical tests of significance are not relevant.
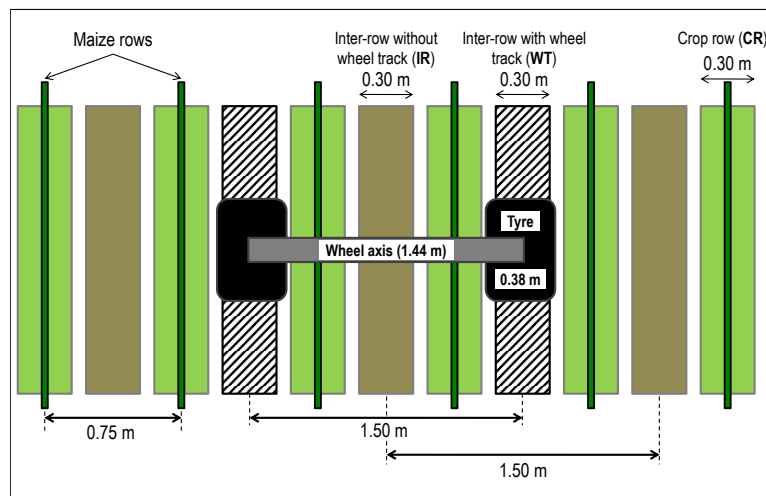
Figure 1. Diagram showing the three areas of influence of cultivation practices. The light green indicates the crop rows; the brown indicates the inter-rows without wheel tracks; and the hatched shading indicates the inter-rows with wheel tracks. The narrow dark green strips show the rows of maize.
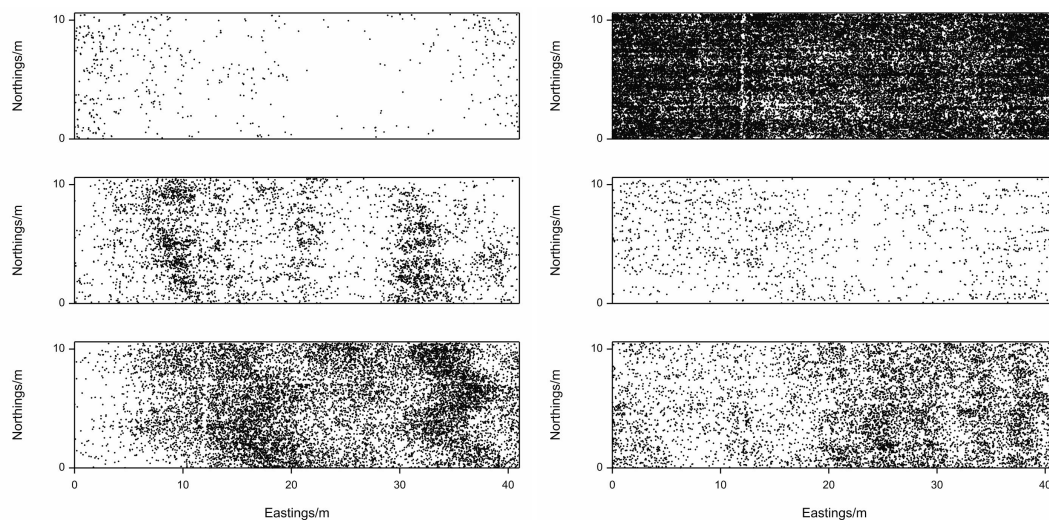

Figure 2. Positions of all seedlings of the six weed species. Left side from top to bottom: nutsedge, johnsongrass and velvetleaf; right side from top to bottom: thornapple, nightshade and cocklebur.

2.3.2. Geostatistical Analysis

A second way of exploring the spatial aggregation of weeds is by a geostatistical analysis of the quadrat counts. The principal tool of geostatistics is the variogram. This function links the expected squared difference of a variable between any two places as:

$$\gamma(\mathbf{h}) \;=\; \frac{1}{2}\mathrm{E}\left[\{Z(\mathbf{x}) - Z(\mathbf{x}+\mathbf{h})\}^2\right] \,, \tag{2}$$

where $Z(\mathbf{x})$ and $Z(\mathbf{x}+\mathbf{h})$ are random variables at positions $\mathbf{x}$ and $\mathbf{x}+\mathbf{h}$. The vector $\mathbf{h}$ defines the separation between the positions in one, two or three dimensions. In our application, we have two dimensions. An intrinsically stationary random process has a variogram that is bounded with a range,

or effective range, and if the variable is generated from a point process, as in our case, then the effective range may be interpreted as the distance across patches of aggregation [27].

Values of $\gamma(\mathbf{h})$, semivariances, can be estimated by the method of moments:

$$\widehat{\gamma}(\mathbf{h}) \;=\; \frac{1}{2m(\mathbf{h})} \sum_{i=1}^{m(\mathbf{h})} \{z(\mathbf{x}_i) - z(\mathbf{x}_i + \mathbf{h})\}^2 \; . \tag{3}$$

For our application, $z(\mathbf{x}_i)$ and $z(\mathbf{x}_i + \mathbf{h})$ are the counts in quadrats centered at $\mathbf{x}_i$ and $\mathbf{x}_i + \mathbf{h}$ separated by the vector $\mathbf{h}$, and $m(\mathbf{h})$ is the number of paired comparisons at that separation. By changing the lag, $\mathbf{h}$, we obtain an ordered series, the experimental variogram, which describes the way the variance changes with the changing separation in the particular set of data.

We chose for z the counts in quadrats of size 0.25 m × 0.25 m. Quadrats of this size fit neatly between the rows (0.75 m apart) and provide data at a fairly fine resolution. Larger quadrats would have led to a relatively crude analysis; smaller ones would have resulted in too many quadrats containing no weeds of some species. Our choice seemed the best compromise.

When fitting a variogram to the data, we assume that the process is intrinsically stationary, that is to say that Equation (2) holds throughout the region and that there is no trend. In some instances, there appears to be a trend (in the data for cocklebur in this study), and this assumption does not hold (i.e., the mean is not stationary). In these cases, we may remove the trend by fitting an appropriate surface (e.g., a linear or quadratic surface) to the data and calculating the variogram of the residuals (i.e., data minus fit surface). This allows us to reveal the short-range variation in the data [28].

Equation (3) is general for any vector $\mathbf{h}$. For data on a regular grid, such as our quadrat counts, we can refine the formula to obtain semivariances in two-dimensions with the following formulae:

$$\widehat{\gamma}(p,q) \;=\; \frac{1}{2(m-p)(n-p)} \sum_{i=1}^{m-p} \sum_{j=1}^{n-q} \{z(i,j) - z(i+p,j+q)\}^2 \tag{4}$$

and:

$$\widehat{\gamma}(p,-q) \;=\; \frac{1}{(2(m-p)(n-p)} \sum_{i=1}^{m-p} \sum_{j=q+1}^{n} \{z(i,j) - z(i+p,j-q)\}^2 \; , \tag{5}$$

in which p and q are the lags along the rows (41.0 m long) and down the columns (10.5 m long) of the grid, respectively. The variogram is symmetric about its center, and the full set of semivariances is obtained by:

$$\widehat{\gamma}(-p,-q) \;=\; \widehat{\gamma}(p,q)$$
$$\text{and} \qquad \widehat{\gamma}(-p,q) \;=\; \widehat{\gamma}(p,-q) \; .$$

We programmed the above formula to compute the two-dimensional variograms in MATLAB, Release R2015b, The MathWorks, Inc., Natick, Massachusetts, USA. The code is provided in the Supplementary Material.

There seemed to be two principal directions of distinct variation, namely along the rows of the maize and perpendicular to the rows. We therefore calculated semivariances for these directions independently by:

$$\widehat{\gamma}(h) \;=\; \frac{1}{2m(h)} \sum_{i=1}^{m(h)} \{z(x_i) - z(x_i + h)\}^2 \; , \tag{6}$$

where $z(x_i)$ and $z(x_i + h)$ are the counts in two quadrats with centers h apart and $m(h)$ is the number of paired comparisons that lag. Note that h is now a scalar value in distance only along the rows

or columns. By changing h, we again obtained an ordered set of semivariances, the experimental variogram for the particular direction.

To find the range or effective range of the variation, we fit plausible models to the experimental variograms. The simplest of these is the exponential model:

$$
\begin{aligned}
\gamma(h) &= c_0 + c_1 \{1 - \exp(-h/a)\} \quad \text{for } h > 0 \\
&= 0 \quad \text{for } h = 0 \,.
\end{aligned}
\tag{7}
$$

Here, c_1 is the correlated variance at $h = \infty$, c_0 is the intercept of the model at zero lag, known as the nugget variance, and a is a distance parameter. The sum $c_0 + c_1$ is the sill, an asymptote. The quantity $3a$ is the effective range of the spatial correlation.

The experimental variogram for nightshade in the direction parallel to the rows appeared unbounded, and for it, we fit a power model:

$$
\begin{aligned}
\gamma(h) &= c_0 + \beta h^\alpha \quad \text{for } 0 < \alpha < 2 \text{ and } h > 0 \\
&= 0 \quad \text{for } h = 0 \,.
\end{aligned}
\tag{8}
$$

Here, c_0 is nugget variance as before, β describes the intensity of the variation and α, which must lie between 0 and 2, describes the curvature.

Other experimental variograms appeared periodic, and to these, we fit a periodic model:

$$
\begin{aligned}
\gamma(h) &= c_0 + A \left\{ 1 - \frac{\omega}{2\pi h} \sin\left(\frac{2\pi h}{\omega}\right) \right\} \quad \text{for } h > 0 \\
&= 0 \quad \text{for } h = 0 \,,
\end{aligned}
\tag{9}
$$

where ω is the period, the distance parameter of the function, A is the amplitude and c_0 is the nugget variance.

We fit the models using the `FITNONLINEAR` directive in GenStat, Version 18 [29].

We also computed the cross-variograms to reveal any spatial relationships between pairs of weeds species. The method-of-moments formula for the experimental cross-variogram is:

$$
\widehat{\gamma}_{uv}(h) = \frac{1}{2m(h)} \sum_{i=1}^{m(h)} \{z_u(x_i) - z_u(x_i + h)\} \{z_v(x_i) - z_v(x_i + h)\} \,,
\tag{10}
$$

where $z_u(x_i)$ and $z_v(x_i)$ are the two regionalized variables corresponding to species u and v. Again, the computations were done in the two principal directions of the grid.

2.3.3. Wavelet Analysis

Wavelet analysis is a more recent form of spatial analysis, developed originally for image processing, but finding increasing application for quantifying variation over a range of scales. Percival and Walden [30] deal comprehensively with the subject. Here, we give a short description of the salient aspects for our application. We used the maximal overlap discrete wavelet transform (MODWT) to quantify the correlation between species at several spatial scales.

A basic (or mother) discrete wavelet function,

$$
\boldsymbol{\psi}_1 = \{\psi_1(1), \psi_1(2), \ldots, \psi_1(N)\} \,,
$$

is a linear filter. When we convolve this function with a series of regularly-spaced data $z(x)$ for which $x = 1, 2, \ldots, N$, we obtain a set of wavelet coefficients $W_1(x)$ for $x = 1, 2, \ldots, N$:

$$W_j(x) = \sum_{i=1}^{N} \psi_j(i) z(x - i) \quad \text{for } i = 1, 2, \ldots, N, \tag{11}$$

where $j = 1$. The wavelet functions are non-zero over only a limited narrow interval (in other words, they have compact support), and so, these coefficients capture local components of the variance. The coefficients obtained with the mother wavelet function are nominally associated with the 2–4 unit scale, which is the finest scale interval. To obtain information for coarser scales, one dilates the mother wavelet function. We denote the j-th dilation of the mother wavelet function $\psi_j = \{\psi_j(1), \psi_j(2), \ldots, \psi_j(N)\}$, so that the resulting wavelet coefficients are nominally associated with the 2^j–2^{j+1} unit scale. From Equation (11), it is clear that towards the ends of a data series, the non-zero part of the wavelet filter overlaps the ends of the data. To avoid the problems associated with this, we pad the data by a process of reflection. The wavelet coefficients can be used to calculate the wavelet variance for each scale as:

$$\sigma_j^2 = \frac{1}{n} \sum_{k=f}^{g} \{W_j(k)\}^2, \tag{12}$$

where $n = g - f + 1$. Wavelet coefficients calculated towards the ends of the data series are affected by the wavelet functions overlapping the ends of the data. Typically, f and g are chosen so that only wavelet coefficients that are unaffected or minimally affected are included in the estimate [31]. We can compute scale-specific wavelet covariances for variables u and v in a similar way [32,33], based on the following formula:

$$C_{u,v,j} = \frac{1}{n} \sum_{k=f}^{g} W_{u,j}(k) W_{v,j}(k), \tag{13}$$

and so, wavelet correlations can be computed as:

$$r_{u,v,j} = \frac{C_{u,v,j}}{\sigma_{u,j} \sigma_{v,j}}. \tag{14}$$

The advantages of the wavelet correlation over the cross-variogram are (1) that the wavelet correlation is on a standardized scale that we can more easily interpret and (2) that we can calculate confidence intervals on the wavelet correlations, and so deduce whether they are significantly different from zero. Lark and Webster [32] describe how to compute confidence intervals for these estimates based on effective degrees of freedom for the wavelet correlation. We followed this procedure here.

From the counts in the 0.25 m × 0.25 m quadrats, we formed 41 transects in the direction of the crop rows. Each transect was 164 quadrats long. We allocated each transect to one of four classes: not influenced by either wheel tracks or crop rows, influenced by crop rows, but not a wheel tracks, influenced by wheel tracks, but not crop rows, and influenced by both wheel tracks and crop rows. For each of these, we computed the wavelet coefficients of weeds at scales of 0.5–1 m, 1–2 m, 2–4 m, 4–8 m and 8–16 m. We chose to use Daubechies's extremal phase wavelet [34] with two vanishing moments as our mother wavelet, because it has a compact support and so is well-suited to sets of data on a small grid. We then used the wavelet coefficients to calculate the wavelet correlations between each pair of species separately for all transects pooled. As described above, we included in our calculation only wavelet coefficients that were unaffected or minimally affected by the filters overlapping the ends of the data.

3. Results and Discussion

We recorded more than 82,000 individual seedlings in the plot: 560 of nutsedge, 4869 of johnsongrass, 13,296 of velvetleaf, 53,493 of thornapple, 1148 of nightshade and 8847 of cocklebur. Figure 2 shows their positions.

3.1. Aggregation Analysis

The spatial patterns of the six main species differed substantially from one another. Thornapple was dense across the whole plot with an apparently uniform distribution. Cocklebur infested only the eastern half of the plot heavily. Velvetleaf and johnsongrass infested the plot moderately with variable density. Nutsedge and nightshade were sparse (see Figure 2).

The aggregation analysis (variance:mean ratio) leads us to reject the null hypothesis of randomness in distribution in all six species: all seem aggregated at some scale. This result accords with those of Heijting et al. [35], who studied weeds in maize crops in The Netherlands and with those of Cardina et al. [36] in various crops in USA. San Martín et al. [37] came to a similar conclusion for maize crops in Central Spain, though the spatial resolutions in that investigation lay within only two grid sizes (9 m × 9 m and 1 m × 0.75 m). In the current study, we found an aggregated structure from a very fine scale (Table 2). This scale varied from species to species and was related to the overall weed density in the plot (Figure 2). For example, thornapple was aggregated from the finest scale (3 cm) and nutsedge from the coarsest (18 cm) (see Table 2). It is likely that these differences are caused by biological factors, in particular the weight of seed, to which we have fit the exponential equation:

$$\text{Aggregate size} = 9.04 - 4.89 \left[1 - \exp \left\{ -\frac{\ln(1000\text{-seed-weight})}{1.41} \right\} \right] \tag{15}$$

as shown in Figure 3. Other factors are likely to be management, total abundance and underlying variation in the soil, as well as seed bank, which reflect the history of the field. Indeed, life history traits are related to plant establishment and seed reproduction [38], therefore being important drivers of weed distributions.

Table 2. Smallest size of significant aggregation and associated variance:mean ratio.

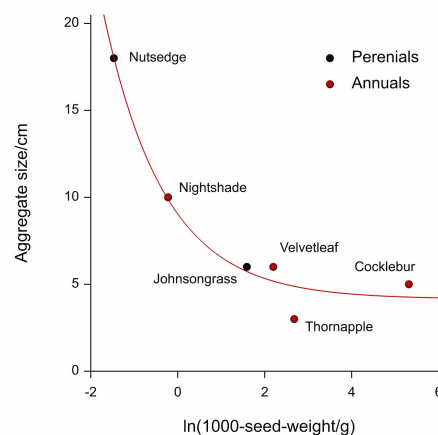| Species	Aggregate Size/cm	Variance:Mean
Nutsedge	18	1.25
Johnsongrass	6	1.25
Velvetleaf	6	1.09
Thornapple	3	1.09
Nightshade	10	1.33
Cocklebur	5	1.20

Figure 3. Relation between seed weight (natural logarithm of the weight of 1000 seeds) and the scale of aggregation (calculated by the Grieg-Smith analysis; values in Table 2). Black discs are for the perennial species, and red ones are for the annual species. The curve is that of the exponential Equation (15) fit by the least-squares algorithm in GenStat.

Densities of each species within zones IR, CR and WT show that the influence of soil compaction on the germination of weed seeds varied from species to species. The operation of a maize planter (that was used in our study field) serves as a useful illustration of how a farming operation can affect weeds at a fine scale. During planting, one part of the seeding unit (the opening disc) digs a furrow for the seed and simultaneously exposes buried weed seeds, while the other part (the press wheel) compacts the soil all within a space of tens of centimeters. The disturbance, caused either by the turnover of the soil or compaction, or both, stimulates the weeds to emerge because of the short exposure of seeds to light and changes in the ratios of soil gases and in water content [14,15,39–41]. Compaction can also limit the available oxygen in soil, and competition from crop plants in the crop rows can reduce the density of weeds [42,43]. Thornapple responded positively to compaction under the tractor wheels and, to a lesser degree, by the press wheel of the planter in the crop rows (Figure 4). In contrast, johnsongrass and cocklebur responded in the opposite way, being less dense in the wheel tracks. Although the other three species (nutsedge, velvetleaf and nightshade) had similar behaviors, the effects were less distinct. Johnsongrass is a perennial species, and most of its offspring do not come from seed, but from rhizomes [44]. These subterranean structures are likely to suffer from soil compaction, whereas the small seeds of thornapple do not. However, the seeds of johnsongrass allow the species to create new foci of infestation ('guerrilla' strategy), which finally merge with the main patch in a 'phalanx' strategy through rhizomes [45]. It seems that johnsongrass seeds disperse naturally over short distances (< 5 m) from the parent plants [44]. Although the response of seed germination to soil compaction had already been described [15,16], our analyses provides quantitative evidence of this effect. Soil compaction affects seeds' dormancy and germination in several ways [12–14]. Compaction can encourage seeds to germinate by modifying moisture conditions in the soil [39]. The greater the contact between soil aggregates and seed, the more readily water can move towards seeds [46]. Compaction can also influence the concentration of oxygen in the soil. In this regard, over-compaction might discourage germination or emergence because it creates an anoxic environment [47]. Compaction may also prevent or retard weed emergence.
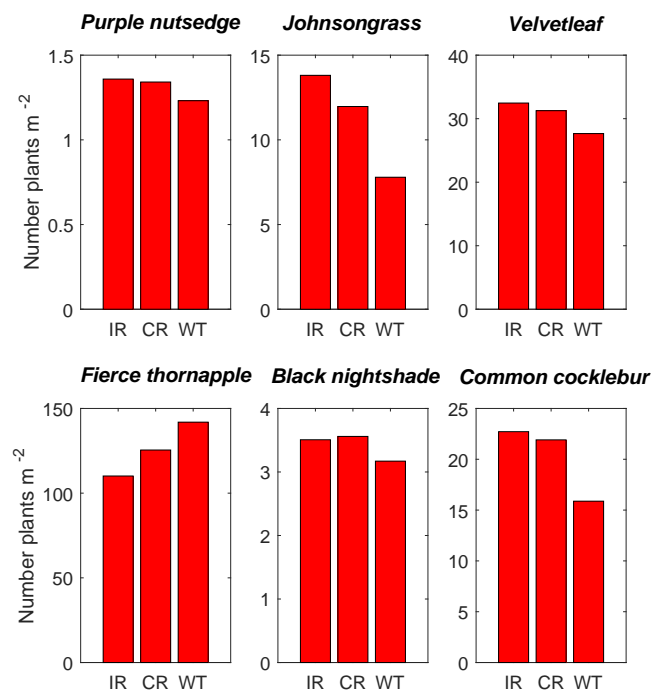
Figure 4. Densities of nutsedge, johnsongrass, velvetleaf, thornapple, nightshade and cocklebur in three zones: (1) inter-rows without wheel tracks (IR), (2) crop rows (CR) and (3) inter-rows with wheel tracks (WT).

3.2. Geostatistical Analysis

The two-dimensional variograms (Figure 5) provided strong evidence of anisotropy, particularly for velvetleaf, thornapple, johnsongrass and cocklebur. For common cocklebur, we removed the trend in data to reveal the short-range variation. As noted above, the variograms of the residuals from the trend showed the largest difference in variation between the direction of the crop rows (i.e., 41.0-m side) and the direction perpendicular to this. The variation appeared to be larger in the direction of the crop rows than in the perpendicular direction. However, differences in the perpendicular direction were greater in thornapple and, to a lesser extent, in cocklebur. To quantify this effect better, we computed the variograms in these two main directions separately.
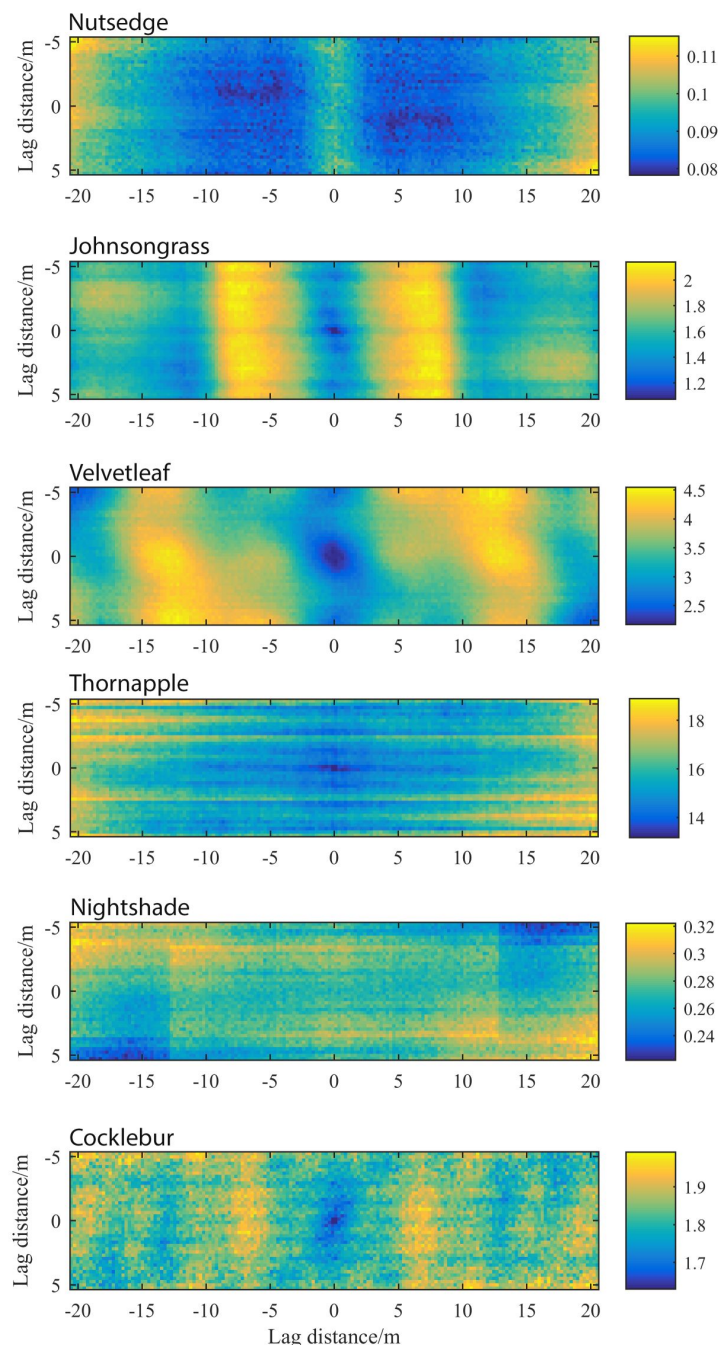
Figure 5. Two-dimensional variograms for the species. The magnitude of the variances are indicated by the scale bars. The trend in cocklebur in the west-east direction has been removed.

We computed the variograms and fit suitable models for all species except for nutsedge, where we could not fit any authorized model (Table 3). In the direction perpendicular to the rows, effects of the wheel tracks were evident for nightshade, johnsongrass, cocklebur and, more particularly, for thornapple (Figure 6). In all of these cases, a model with a periodic component (an oscillation) was fit. The variance of these models peaked at approximately 75-cm lag, corresponding to half the distance between wheel tracks (1.44 m). This indicates that the largest differences in weed density in this direction can be associated with the alternation of compacted and un-compacted soil. With velvetleaf, it seemed likely that the zig-zag pattern in this direction (see Figure 2) obscures the effect of the wheel

tracks. This zig-zag pattern might have been caused when the harvester caught the fruits containing the seeds and scattered them.

Table 3. Models fit to experimental variograms.

Species	Direction *	Model	c_0	c_1	A	a/m	ω/m	α	β **
Johnsongrass	0	Periodic	1.147		0.601		8.82		
	90 °	Periodic	0.993		0.347		1.29		
Velvetleaf	0	Periodic	2.061		1.592		7.28		
	90 °	Periodic	2.101		0.616		4.32		
Thornapple	0	Exponential	10.49	3.124		3.38			
	90 °	Periodic	11.91		2.451		1.29		
Nightshade	0	Power	0.255					1.75	0.000271
	90 °	Periodic	0.260		0.0164		1.37		
Cocklebur	0	Exponential	1.551	0.938	0.070	7.84	5.24		
	90 °	+Periodic	1.597	0.299	0.025	2.83	1.11		

No models were fit for nutsedge. * Direction 0 is that of the longer side of the plot; direction 90 ° is that of the shorter side. ** The parameters are those defined in Equations (7), (8) and (9)

Along the crop rows, the magnitude of the variation was similar to or somewhat larger than in the perpendicular direction (except for thornapple) (Figure 6). For velvetleaf and johnsongrass, the variance in this direction was larger at lags exceeding 5 m. This indicates the existence of large patches at this scale and that there might be an interaction of the effect of cultivation or harvester dragging seeds with variable soil properties. This pattern is visible in the maps of the raw data as denser patches (Figure 2). Moreover, in velvetleaf, there appears to be a small trend from the left to the right of the plot (see Figure 2). There was a similar trend in this direction for common cocklebur. The peak in the variograms of velvetleaf and johnsongrass at $\approx$ 5–7 m relates to the larger patches that are visible in Figure 2. For nutsedge and nightshade, we found no spatial structure; the variograms are effectively flat.

The cross-variograms in the direction perpendicular to the crop rows reveal negative correlations between thornapple and other species (Figure 7). Thornapple and johnsongrass appear to exclude one another where the wheels compact the soil (the smallest variances correspond to the 75-cm lags associated with wheel tracks). This is probably the result of allelopathy; sorghum plants inhibit other species by exuding sorgoleone [48]. Although thornapple was negatively correlated with velvetleaf and cocklebur, the structure was not well defined. In the direction of the crop rows, we found only one pair with a clear relation, that was between thornapple and velvetleaf. That correlation was negative and became increasingly negative with increasing lag distance (Figure 8).
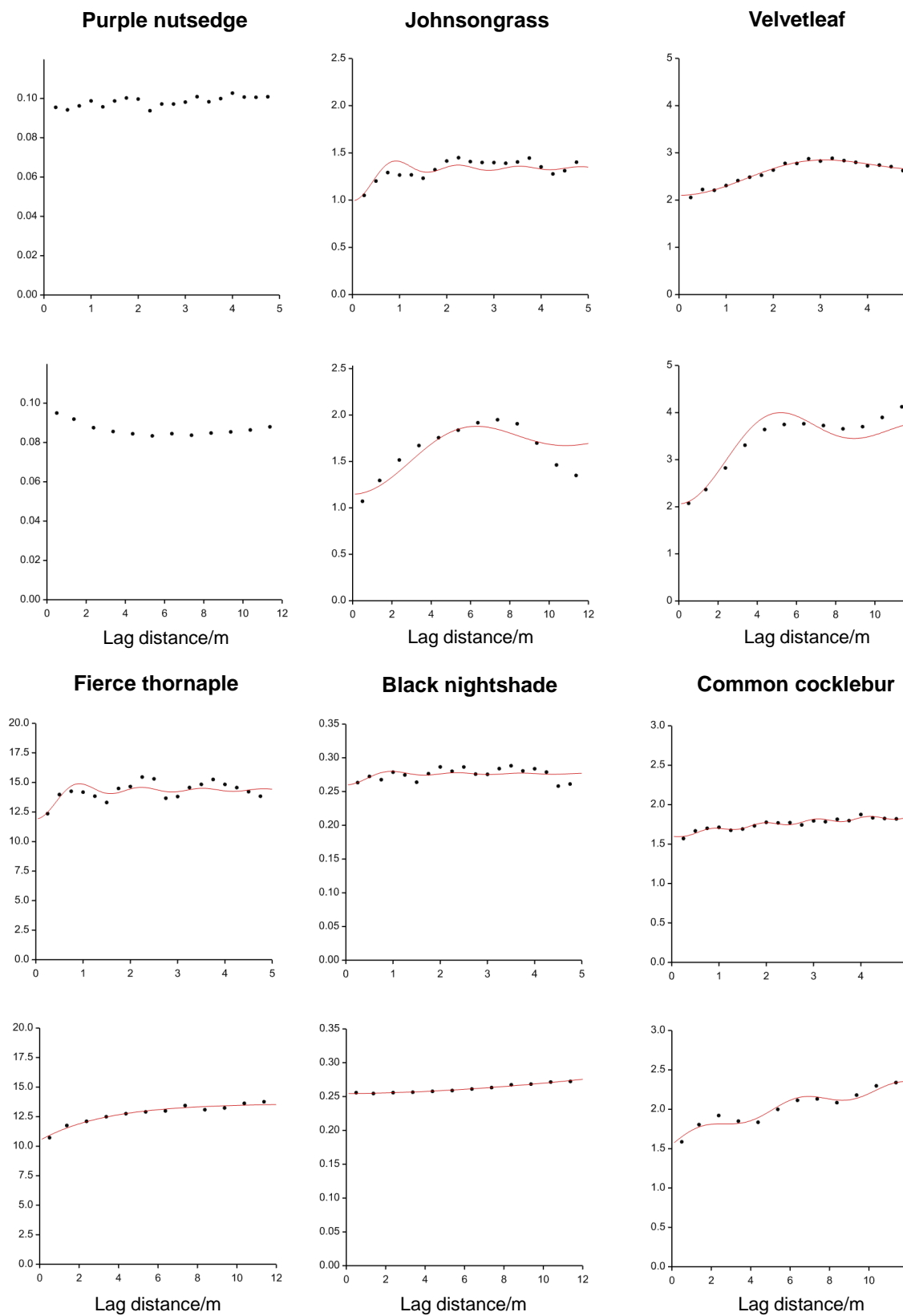
Figure 6. Directional variograms (dots) with fit model (red lines). Up: variograms perpendicular to the crop rows. Bottom: variograms parallel to the crop rows. Models could not be fit to all species and so are not shown in all figures.
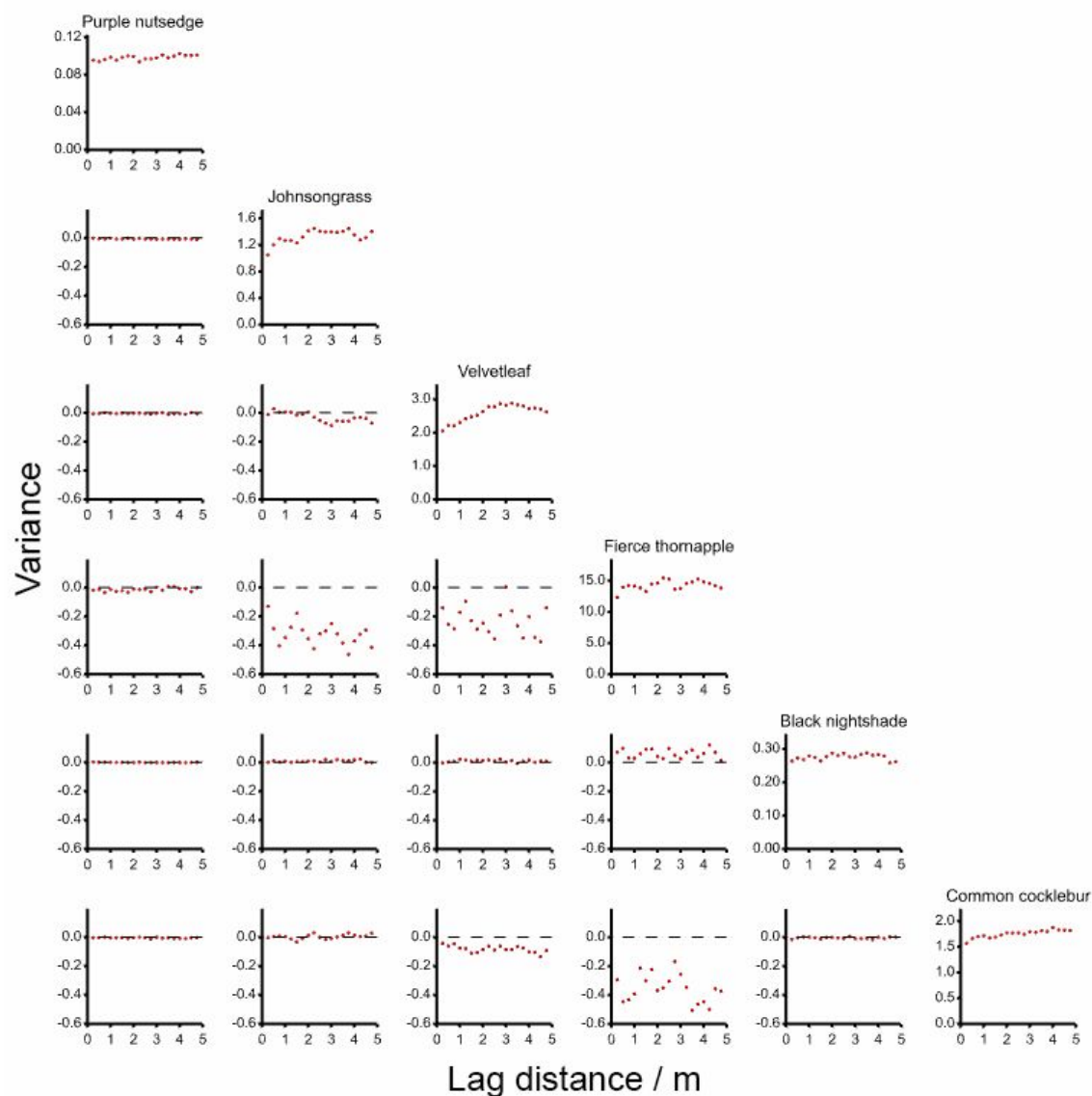
Figure 7. Experimental cross-variograms for pairs of species perpendicular to the crop rows. The dotted line indicates the zero variance line.

3.3. Wavelet Analysis

The positions of the transects in relation to crop rows and wheel tracks did not significantly affect the correlations between the pairs of weeds. Therefore, we present only the results for the whole population (Figure 9). We observed the five following basic patterns of correlation.

(1)   There were significant positive correlations between the pair nutsedge and thornapple and the pair velvetleaf and nightshade at the coarsest scale (16–32 m), suggesting that these pairs of weeds grow in similar regions of the plot (Figure 9a,b).

(2)   There were significant negative correlations between nightshade and cocklebur and between johnsongrass and cocklebur at the coarser scales (Figure 9c,d); it seems that the weeds in each pair exclude one another spatially, being located in different regions in the crop.

(3)   A weak, but significant negative correlation between thornapple and cocklebur appears at the finest scales (0.5–2 m); this suggests that there is local competition between these species (Figure 9e).

(4)   For the pairs velvetleaf and johnsongrass, velvetleaf and cocklebur, and velvetleaf and thornapple, we observed weak, but significant negative correlations at the mid-range scales (4–16 m) (for example, Figure 9f).

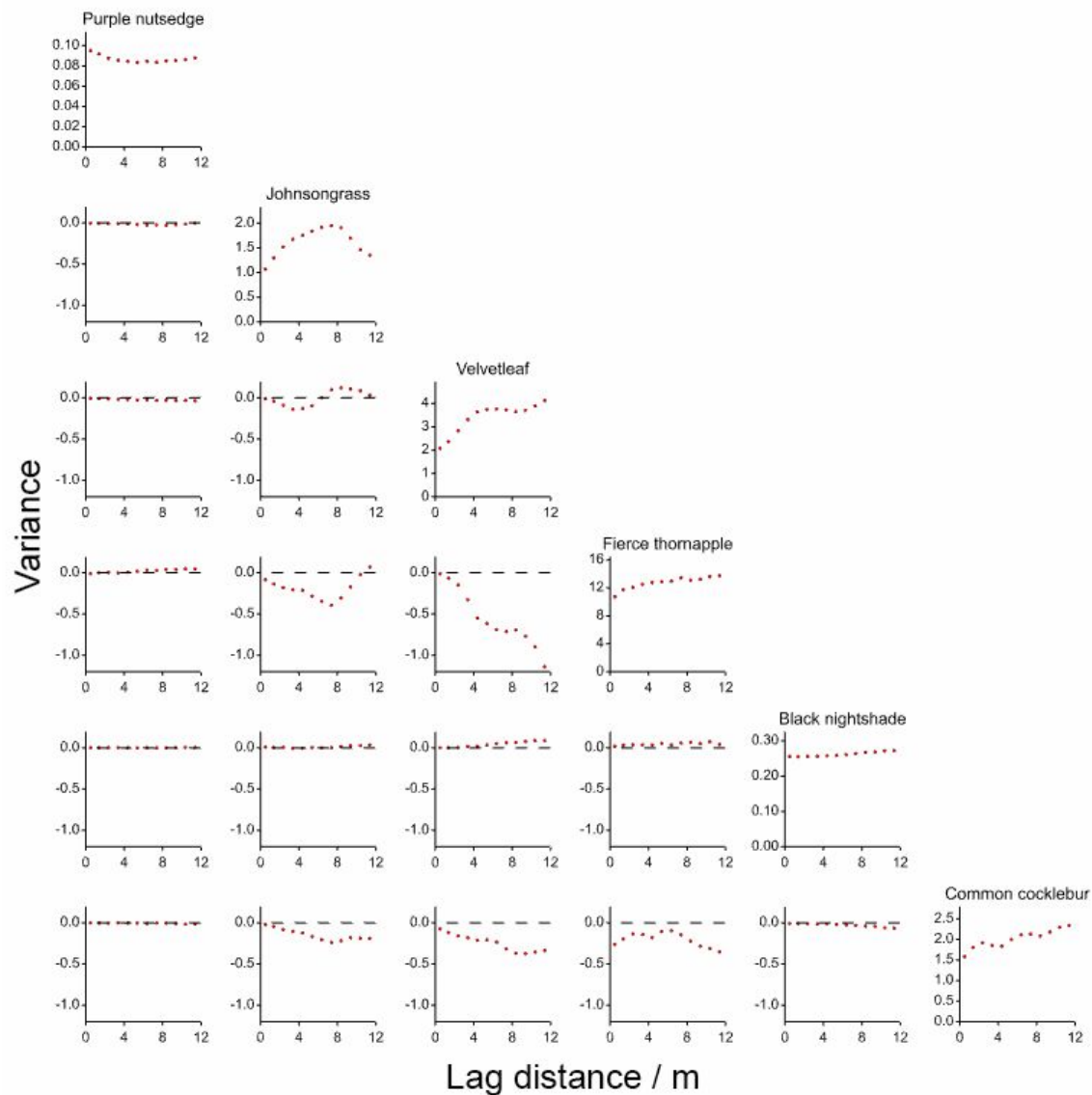(5)   All other pairs had weak correlations across the scales that were not significantly different from zero.

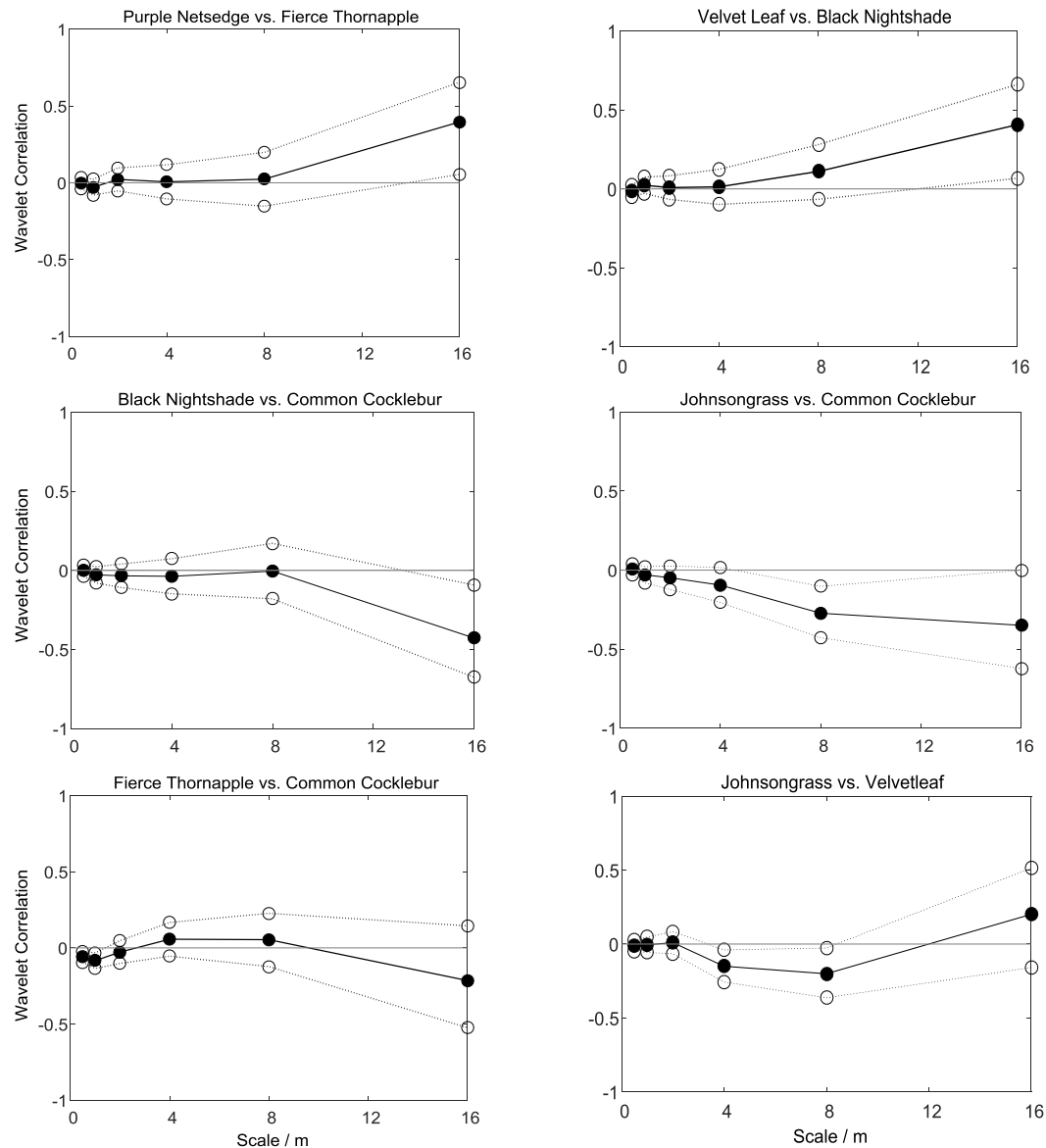**Figure 8.** Experimental cross-variograms for pairs of species parallel to the crop rows.

Figure 9. Examples of wavelet correlations for pairs of species. Black discs are the estimates; open circles show the 95% confidence limits.

4. Implications

The presence of a species in a given place is determined not only by germination. To flourish, a plant needs both favorable conditions for germination and for its establishment [49], as well as to complete its whole life cycle [50]. Further, the continuity of a species in a place is strongly related to its capacity to produce offspring there. When plants are very dense, as in the field we studied, intra-specific and inter-specific competition is important [51]. We found that thornapple excludes johnsongrass where wheel tracks have compacted the soil (or, alternatively, johnsongrass excludes thornapple in non-compacted areas). Our results do not allow us to distinguish definitively the relative importance of abiotic (compaction) and biotic (competitive exclusion) factors in causing the spatial distribution of the six species. They do, however, give a novel insight into the scales at which these factors operate and enable us to form hypotheses about the plant traits that explain the propensity of the plant species to form patches and the role of the local environment in facilitating or interrupting spread that we can test experimentally.

Our spatial analysis reveals the scales at which the weed densities vary and the effects of crop rows and tractor wheels on the spatial distribution of weeds. It suggests that finer-scale patterns result from a combination of biological, environmental and management factors. While the relationship between seed size and patch size is correlative and confounded by the spread of two species by rhizome fragments, it supports the hypothesis that traits of weed dispersal can enable us to predict the propensity of different species to form patches at different scales. Weed species with contrasting traits will, therefore, vary in their spatial dynamics at the fine scale, resulting in different scales of aggregation and stability of patches. For example, the scale of aggregation revealed by the Grieg-Smith analysis broadly increased with decreasing seed size (Tables 1 and 2). There was evidence that management and its effect on the local environment influenced the density of several species (Figures 4–8). This idea could be supported by more focused experimentation and mechanistic spatially-explicit models of weed populations that predict the spread of contrasting weeds in a heterogeneous environment. Simulation models could be used in the future to investigate the effect of other biotic and abiotic factors on the spatial distribution of weeds.

A fundamental understanding of the interactions between weed biology and the local environment at the fine scale will be important if precision weed management is to be widely adopted. Farmers will need to be sure that they have correctly identified the circumstances for successful patch spraying and to prevent weeds escaping from their control. The combination of GPS navigation, digital imaging from an on-board camera and pre-programmed statistical analysis of the data should enable farmers to achieve that goal. It should allow farmers (1) to predict where infestation by weeds is likely and so be prepared to spray herbicide there and (2) help them to manage soil conditions as part of an integrated approach to weed control.

Author Contributions: All authors have contributed to the development of the research described. Cesar Fernández-Quintanilla, Dionisio Andújar and José Dorado conceived of and designed the survey and recorded the data. Carolina San Martín processed the imagery and prepared the files of data for analysis, which was done by Carolina San Martín, Alice E. Milne, Richard Webster and Jonathan Storkey. Carolina San Martín, Alice E. Milne and Richard Webster prepared the script, and all of us responded to the reviews.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Maxwell, B.D.; Luschei, E.C. Justification for site-specific weed management based on ecology and economics. *Weed Sci.* **2005**, *53*, 221–227.
2. San Martín, C.; Andújar, D.; Barroso, J.; Fernández-Quintanilla, C.; Dorado, J. Weed decision threshold as a key factor for herbicide reductions in site-specific weed management. *Weed Technol.* **2016**, *30*, 888–897.
3. Stoate, C.; Boatman, N.D.; Borralho, R.J.; Rio Carvalho, C.; de Snoo, G.R.; Eden, P. Ecological impacts of arable intensification in Europe. *J. Environ. Manag.* **2001**, *63*, 337–365.
4. Andújar, D.; Ribeiro, Á.; Fernández-Quintanilla, C.; Dorado, J. Accuracy and feasibility of optoelectronic sensors for weed mapping in wide row crops. *Sensors* **2011**, *11*, 2304–2318.
5. Burgos-Artizzu, X.P.; Ribeiro, A.; Guijarro, M.; Pajares, G. Real-time image processing for crop/weed discrimination in maize fields. *Comput. Electron. Agric.* **2011**, *75*, 337–346.
6. Christensen, S.; Søgaard, H.T.; Kudsk, P.; Nørremark, M.; Lund, I.; Nadimi, E.S.; Jørgensen, R. Site-specific weed control technologies. *Weed Res.* **2009**, *49*, 233–241.

7. Herrera, P.J.; Dorado, J.; Ribeiro, Á. A novel approach for weed type classification based on shape descriptors and a fuzzy decision-making method. *Sensors* **2014**, *14*, 15304–15324.
8. Dieleman, J.A.; Mortensen, D.A.; Buhler, D.D.; Cambardella, C.A.; Moorman, T.B. Identifying associations among site properties and weed species abundance. I. Multivariate analysis. *Weed Sci.* **2000**, *48*, 567–575.
9. Dieleman, J.A.; Mortensen, D.A.; Buhler, D.D.; Ferguson, R.B. Identifying associations among site properties and weed species abundance. II. Hypothesis generation. *Weed Sci.* **2000**, *48*, 576–587.
10. Rew, L.J.; Cousens, R.D. Spatial distribution of weeds in arable crops: Are current sampling and analytical methods appropriate? *Weed Res.* **2001**, *41*, 1–18.
11. Metcalfe, H.; Milne, A.E.; Webster, R.; Lark, R.M.; Murdoch, A.J.; Storkey, J. Designing a sampling scheme to reveal correlations between weeds and soil properties at multiple spatial scales. *Weed Res.* **2016**, *56*, 1–13.
12. Benech-Arnold, R.L.; Sanchez, R.A.; Forcella, F.; Kruk, B.C.; Ghersa, C.M. Environmental control of dormancy in weed seed banks in soil. *Weed Sci.* **2000**, *67*, 105–122.
13. Chauhan, B.S.; Singh, R.G.; Mahajan, G. Ecology and management of weeds under conservation agriculture: A review. *Crop Prot.* **2012**, *38*, 57–65.
14. Gallagher, R.; Cardina, J. The effect of light environment during tillage on the recruitment of the various summer annuals. *Weed Sci.* **1998**, *46*, 214–216.
15. Tardif-Paradis, C.; Simard, M.-J.; Leroux, G.D.; Panneton, B.; Nurse, R.E.; Vanasse, A. Effect of planter and tractor wheels on row and inter-row weed populations. *Crop Prot.* **2015**, *71*, 66–71.
16. Longchamps, L.; Panneton, B.; Simard, M.-J.; Leroux, G.D. Could weed sensing in corn interrows result in efficient weed control? *Weed Technol.* **2012**, *26*, 649–656.
17. Andújar, D.; Ruiz, D.; Ribeiro, Á.; Fernández-Quintanilla, C.; Dorado, J. Spatial distribution patterns of Johnsongrass (*Sorghum halepense*) in corn fields in Spain. *Weed Sci.* **2011**, *59*, 82–89.
18. Ruiz, D.; Escribano, C.; Fernández-Quintanilla, C. Assessing the opportunity for site-specific management of *Avena sterilis* in winter barley fields in Spain. *Weed Res.* **2006**, *46*, 379–387.
19. Howard, C.L.; Mortimer, A.M.; Gould, P.; Putwain, P.D.; Cousens, R.; Cussans, G.W. The dispersal of weeds: Seed movement in arable agriculture. In *Proceedings of Brighton Crop Protection Conference–Weeds*; British Crop Protection Council: Brighton, UK, 1991; pp. 821–828.
20. Paice, M.E.R.; Day, W.; Rew, L.J.; Howard, A. A stochastic simulation model for evaluating the concept of patch spraying. *Weed Res.* **1998**, *38*, 373–388.
21. Boyd, N.; Van Acker, R. Seed and microsite limitations to emergence of four annual weed species. *Weed Sci.* **2004**, *52*, 571–577.
22. Royal Botanic Gardens Kew. Seed Information Database (SID). Version 7.1. Available online: http://data.kew.org/sid/ (accessed on 8 February 2018).
23. San Martín, C.; Andújar, D.; Fernández-Quintanilla, C.; Dorado, J. Spatio-temporal dynamics of *Sorghum halepense* in poplar short-rotation coppice under several vegetation management systems. *Forest Ecol. Manag.* **2016**, *379*, 37–49.
24. Greig-Smith, P. The use of random and contiguous quadrats in the study of the structure of plant communities. *Ann. Bot.* **1952**, *16*, 293–316.
25. Food and Agriculture Organization (FAO). World Reference Base for Soil Resources. In *A Framework for International Classification, Correlation and Communication*; World Soil Resources Reports 103; Food and Agriculture Organization (FAO): Rome, Italy, 2006.
26. Environmental Systems Research Institute (ESRI). *ArcGIS Desktop: Release 10.1*; Environmental Systems Research Institute: Redlands, CA, USA, 2013.
27. Matérn, B. Spatial variation: stochastic models and their application to problems in forest surveys and other sampling investigations. *Meddelanden från Statens Skogforskinstitut* **1960**, *49*, 1–144.
28. Webster, R.; Oliver, M.A. *Geostatistics for Environmental Scientists*, 2nd ed.; John Wiley & Sons: Chichester, UK, 2007.
29. Payne, R.W. (Ed.) *The Guide to GenStat Release 18—Part 2: Statistics*; VSN International: Hemel Hempstead, UK, 2016.
30. Percival, D.B.; Walden, A.T. *Wavelet Methods for Time Series Analysis*; Cambridge University Press: Cambridge, UK, 2000.
31. Milne, A.E.; Webster, R.; Lark, R.M. Spectral and wavelet analysis of gilgai patterns from air photography. *Aus. J. Soil Res.* **2010**, *48*, 309–325.

32. Lark, R.M.; Webster, R. Changes in variance and correlation of soil properties with scale and location: Analysis using an adapted maximal overlap discrete wavelet transform. *Eur. J. Soil Sci.* **2001**, *52*, 547–562.

33. Whitcher, B.; Guttorp, P.; Percival, D.B. Wavelet analysis of covariance with application to atmospheric time series. *J. Geophys. Res. Atmos.* **2000**, *105*, 14941–14962.

34. Daubechies, I. Orthonormal bases of compactly supported wavelets. *Comm. Pure Appl. Math.* **1988**, *41*, 909–996.

35. Heijting, S.; Van der Werf, W.; Kruijer, W.; Stein, A. Testing the spatial significance of weed patterns in arable land using Mead's test. *Weed Res.* **2007**, *47*, 396–405.

36. Cardina, J.; Johnson, G.A.; Sparrow, D.H. The nature and consequence of weed spatial distribution. *Weed Sci.* **1997**, *45*, 364–373.

37. San Martín, C.; Andújar, D.; Fernández-Quintanilla, C.; Dorado, J. Spatial distribution patterns of weed communities in corn fields of Central Spain. *Weed Sci.* **2015**, *63*, 936–945.

38. Colbach, N.; Roger-Estrade, J.; Chauvel, B.; Caneill, J. Modeling vertical and lateral seed bank movements during mouldboard ploughing. *Eur. J. Agron.* **2000**, *13*, 111–124.

39. Brown, A.D.; Dexter, A.R.; Chamen, W.C.T.; Spoor, G. Effect of soil macroporosity and aggregate size on seed-soil contact. *Soil Tillage Res.* **1996**, *38*, 203–216.

40. Jurik, T.M.; Zhang, S. Tractor wheel traffic effects on weed emergence in central Iowa. *Weed Technol.* **1999**, *13*, 741–746.

41. Theisen, G.; Bastiaans, L. Low disturbance seeding suppresses weeds in no-tillage soyabean. *Weed Res.* **2015**, *55*, 598–608.

42. Ballaré, C.; Scopel, A.; Sánchez, R. Foraging for light: Photosensory ecology and agricultural implications. *Plant Cell Environ.* **1997**, *20*, 820–825.

43. Norsworthy, J.K. Soybean canopy formation effects on pitted morningglory (*Ipomoea lacunosa*), common cocklebur (*Xanthium strumarium*), and sicklepod (*Senna obtusifolia*) emergence. *Weed Sci.* **2004**, *52*, 954–960.

44. Barroso, J.; Andújar, D.; San Martín, C.; Fernández-Quintanilla, C.; Dorado, J. Johnsongrass (*Sorghum halepense*) seed dispersal in corn crops under Mediterranean conditions. *Weed Sci.* **2012**, *60*, 34–41.

45. Andújar, D.; Barroso, J.; Fernández-Quintanilla, C.; Dorado, J. Spatial and temporal dynamics of *Sorghum halepense* patches in maize crops. *Weed Res.* **2012**, *52*, 411–1420.

46. Benech-Arnold, R.L.; Sanchez. *Handbook of Seed Physiology. Application to Agriculture*; Food Products Press: Binghamton, NY, USA, 2004; 480p.

47. Dexter, A. Advances in characterization of soil structure. *Soil Tillage Res.* **1988**, *11*, 199–238.

48. Głąb, L.; Sowiński, J.; Bough, R.; Dayan, F.E. Allelopathic potential of sorghum (*Sorghum bicolor* (L.) Moench) in weed control: a comprehensive review. *Adv. Agron.* **2017**, *145*, 44–95.

49. Calviño-Cancela, M. Seed and microsite limitations of recruitment and the impacts of post-dispersal seed predation at the within population level. *Plant Ecol.* **2007**, *192*, 35–44.

50. Scott, A.J.; Morgan, J.W. Dispersal and microsite limitation in Australian old fields. *Oecologia* **2012**, *170*, 221–232.

51. Borgy, B.; Perronne, R.; Kohler, C.; Grison, A.-L.; Amiaud, B.; Gaba, S. Changes in functional diversity and intraspecific trait variability of weeds. *Weed Res.* **2016**, *56*, 102–113.