

# Appendices

## Appendix S1: Relation between parasitism and nest density

To investigate the correlation between the degree of parasitism and nest clustering, the local densities of all nests and parasitized nests were compared, calculated with a kernel estimation from the R-package *Spatstat* (Baddeley et al. 2015). For extra information and clear description for the reader, we added a figure with number of nest per m<sup>2</sup> (Fig. S1.0). The relation between parasitized nest density and nest density is modelled with a linear regression with the basic *lm*-function in R for parasitized nests. Only parasitized nests are used for this analysis because parasitism was overall not easily detected and thus, not all zero-observations of the non-parasitized nests can be considered as true zero-observations. This is related to the observation that assumptions for linear regression were not met for the complete dataset (linearity, homogeneity and independence of residuals, homoscedasticity were all not met). The complete dataset with all nests was used to visualize the variation in rate of parasitism: the rate of parasitism was calculated by dividing the density of parasitized nests by the density of the nests and plotted against density of nests.

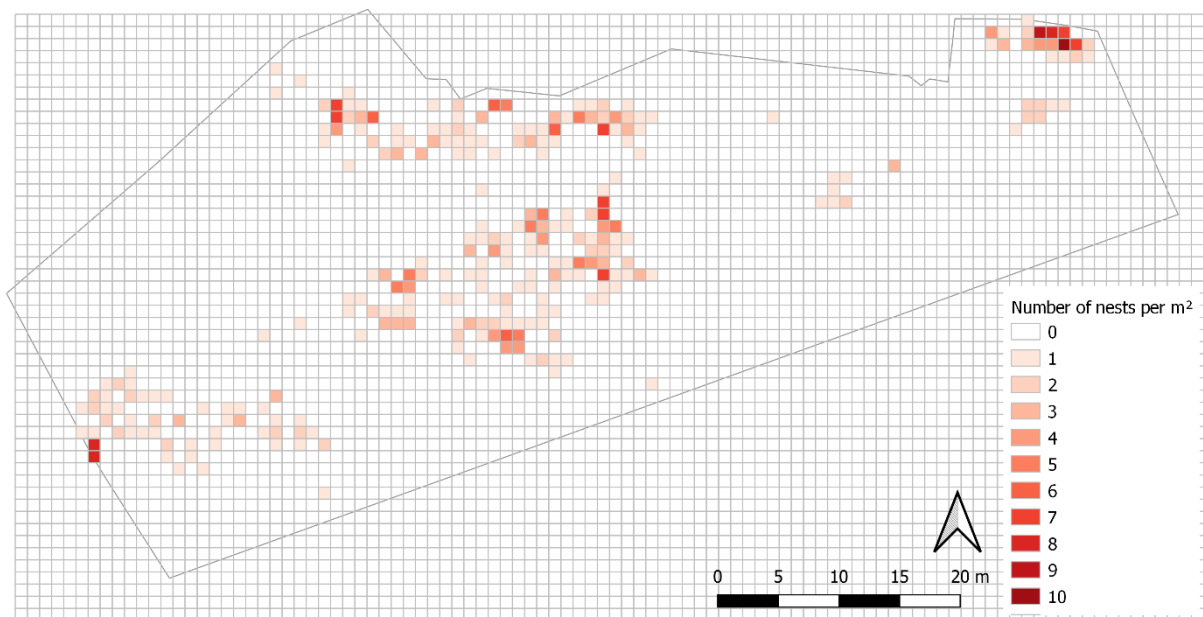


Fig. S1.0: descriptive map of the density of nests in a grid of 1m×1m.

The density of parasitized nests increased with nest density (Fig S1.1;  $\beta = 0.27$ ,  $R^2 = 0.80$ ,  $F(1, 148) = 581.3$ ,  $p < 2 \cdot 10^{-16}$ ). However, as the slope of 0.27 suggests, the number of parasitized nests does not increase at the same rate as the nest density. At higher nest density, the relative number of parasitized nests is therefore lower than at low nest density. Fig S1.2 shows the rate of parasitism in relation to nest density and it is visually clear that there is more variation at lower densities than at larger densities.

This relationship is indicative for a selfish herd pattern, previously suggested to be present in *Bembix rostrata* (Larsson 1986) and another closely related digger wasp, *Crabro cribrellifer* (Wcislo 1984). These studies showed that the incidence of parasitism per nest decreases with nest density or nearest neighbor distance within an aggregation. A similar pattern is found

here in our study, established with a different method. Furthermore, the variation in rate of parasitism decreases with nest density (Fig. S1.2). A higher variation in parasitism incidence at low and intermediate nest densities might favor, for instance, nest choice strategies that include a form of spatial bet hedging (Philippi and Seger 1989). See main text for full discussion.

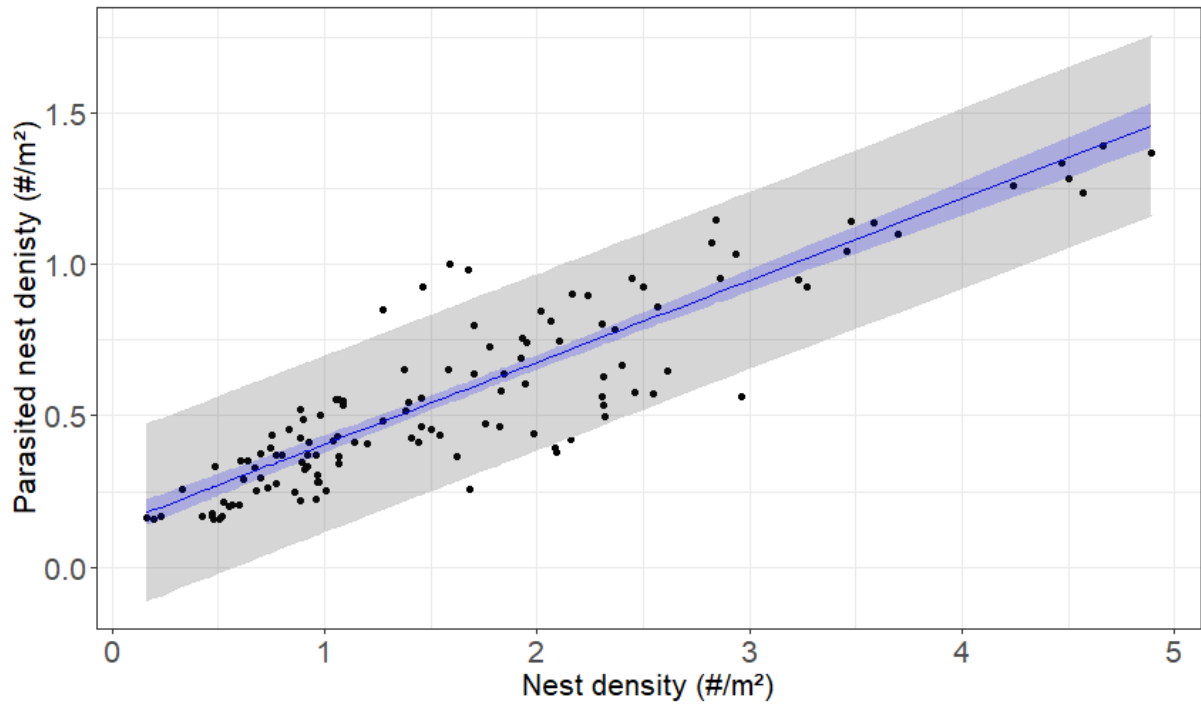


Figure S1.1: estimated parasitized nest density as a function of estimated nest density, for parasitized nests only. Regression line is the dark blue line, confidence interval is light blue shading, prediction interval is grey shading. The slope has an estimated value of 0.27.

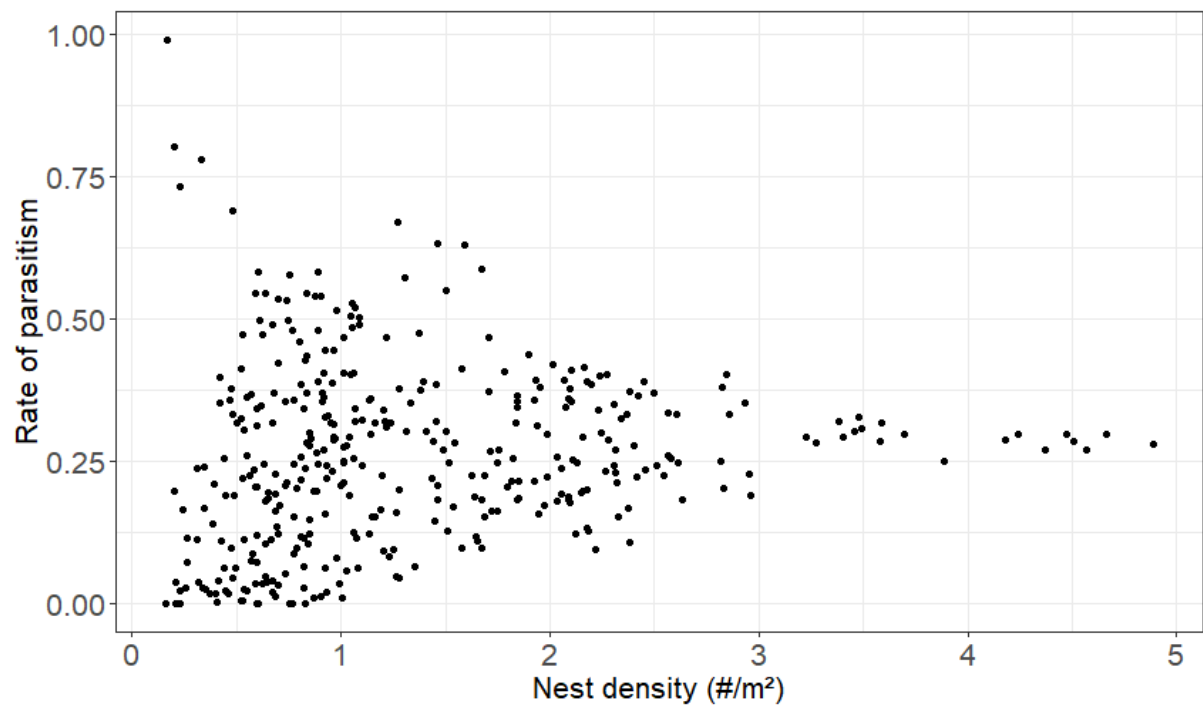


Figure S1.2: Rate of parasitism (estimated density of parasitized nests divided by estimated nest density) as a function of nest density, for all nests (parasitized and non-parasitized).

## Appendix S2: ODD protocol Bembix IBM

### 1. Overview

#### 1.1. Purpose

This IBM is built to understand the mechanisms of spatial pattern formation of nest aggregates in *Bembix rostrata*. More specifically, the IBM is built to understand the relative importance of induced (environmental) and inherent (behavioral) mechanisms, i.e. local site fidelity and conspecific attraction. Combinations of nest choice mechanisms are simulated to understand which behaviors can explain the spatial pattern and social network structure of the field data. The model was implemented with Python (Python Software Foundation 2001).

Model selection and parameterization is done through Approximate Bayesian Computation (ABC)-analysis (Beaumont 2010; Csilléry et al. 2010, 2012; van der Vaart et al. 2016). This method is based on minimizing the difference in summary statistics (spatial pattern, social network) between each simulation and observed field data.

#### 1.2. Entities, state variables and scales (Fig. S2.1)

1. **Population:** consists of a collection of wasps, there is one population per run.

##### State variables:

- Day: the current day-number, which is looped over during the run (integer)
- Wasps: a list of wasp individuals which make up the population
- Strengths of mechanisms: three floats between 0 and 1 that specify the probabilities of having the three mechanisms present.
- Submodel: random, uniform, fixed or flexible (see figure S2 and section 3.3), which specifies at which level the occurrence of mechanisms varies.

2. **Wasps:** these are the core individuals or agents of the model which perform nest site selection.

##### State variables:

- Identity number
- Starting day: the day the wasp will make its first nest (integer)
- Number of nests: how many nests a wasp will make (integer)
- Periods between nests: number of days between consecutive nests (list)
- Mechanisms in use: ENV, LSF or CA; the mechanisms with which it will make its current nest. This can change between nests depending on the submodel of the population.
- Nests: a list of nests made by the wasp

3. **Nests:** these are the nests, made by the wasps

##### State variables:

- Identity number (integer)
- Coordinates x, y: the position of the nest (floats), in meters.
- Day: the day the nest was made (integer)

- Mechanisms used: three Booleans indicating which of the three mechanisms was used during nest choice
4. **Environment:** a grid containing cells with one value, the nest location suitability. The resolution of pixels of the environment is 50x50cm<sup>2</sup>/pixel (the nests have continuous coordinates and are not assigned to a pixel).
  5. **Spatial and temporal scales:** these are chosen according to the field study performed. The spatial extent is 47x97m. The minimal temporal entity is a day and the number of days is 30.

### 1.3. Process overview and scheduling (Fig. S2.1)

The model is discrete and loops over 30 days. During one day, the wasps in the population are looped to see which individuals will make a new nest that day. The wasp searches a new nesting position according to a combination of maximum three types of possible mechanisms. Random positions in the environment are sampled and evaluated for their suitability. According to the averaged probabilities of the search mechanisms, it is chosen if a new nest is made at the currently evaluated position. If the location was not chosen, new locations are being evaluated. Once a suitable position is chosen, a nest is added to its nest-list.

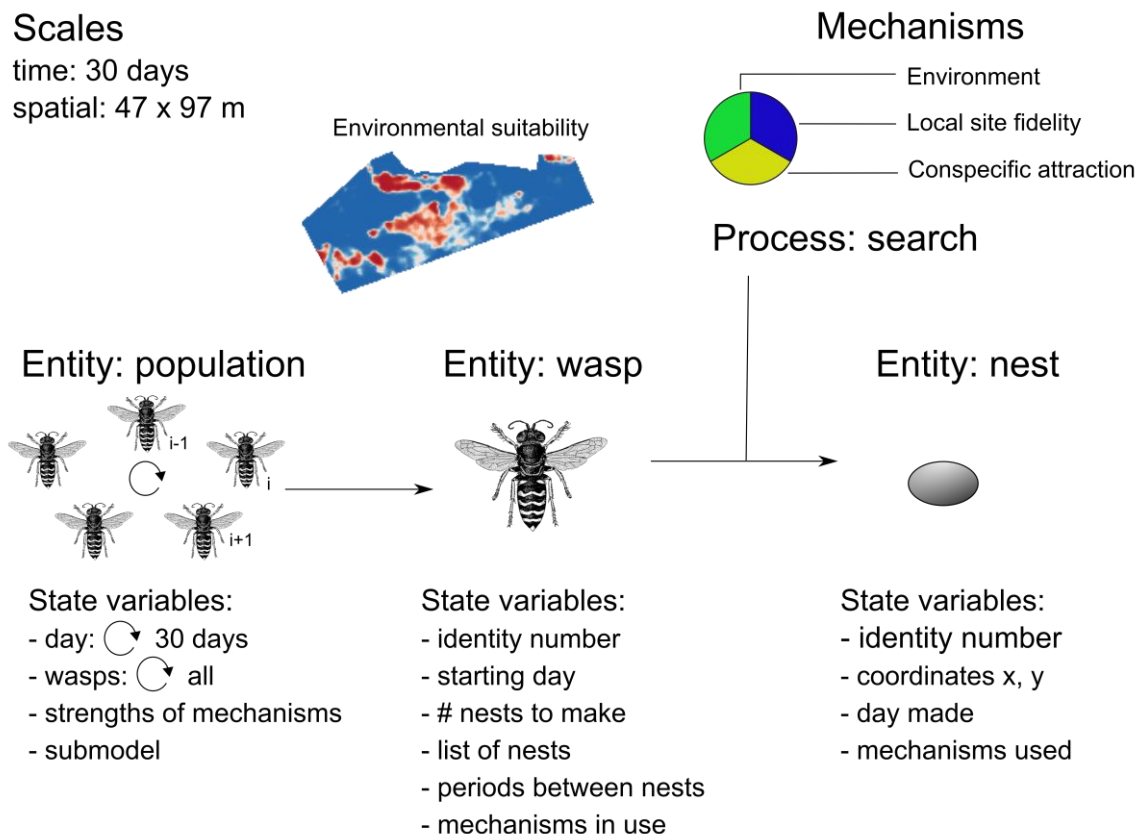


Figure S2.1: schematic overview of entities, their state variables and process scheduling.

## 2. Design concepts

### 2.1. Basic principles (Fig. S2.2)

There are three main mechanisms on which nest selection behavior can be based: environmental suitability (*ENV*), local site fidelity (*LSF*) and conspecific attraction (*CA*). These mechanisms have *strengths*, a value between 0 and 1, the probability each of these mechanisms is present. The selection whether a mechanism is present or absent is made on three different levels (Fig. S2):

- Population level (uniform): the combination of mechanisms for all individuals are identical. The mechanisms are uniformly present or absent across the population.
- Inter-individual level (fixed): the combination of mechanisms can vary between individuals, but is fixed for each.
- Intra-individual level (flexible): the combination of mechanisms can vary when a wasp makes a new nest.

#### Theoretical background of behavioral mechanisms

Environmental and behavioral processes can both result in clustering of individuals or their nests in space, but relative importance of these processes, both within and among individuals, remains elusive. Information used during nest site selection can be personal, when individuals use cues from the environment or from their experience at previous locations, with the latter resulting in local site fidelity (Hoi et al. 2012; Asís et al. 2014). Information can also be inadvertently social when information is generated by the activities of other conspecific individuals (Danchin et al. 2004; Dall et al. 2005).

Both local site fidelity and conspecific attraction have both been put forward as behavioral mechanisms in *Bembix* species and other digger wasps, Crabronidae (Wcislo 1984; Larsson 1986; Larsson and Tengö 1989; Asís et al. 2004; Polidori et al. 2008), but the relative importance of both is unknown.

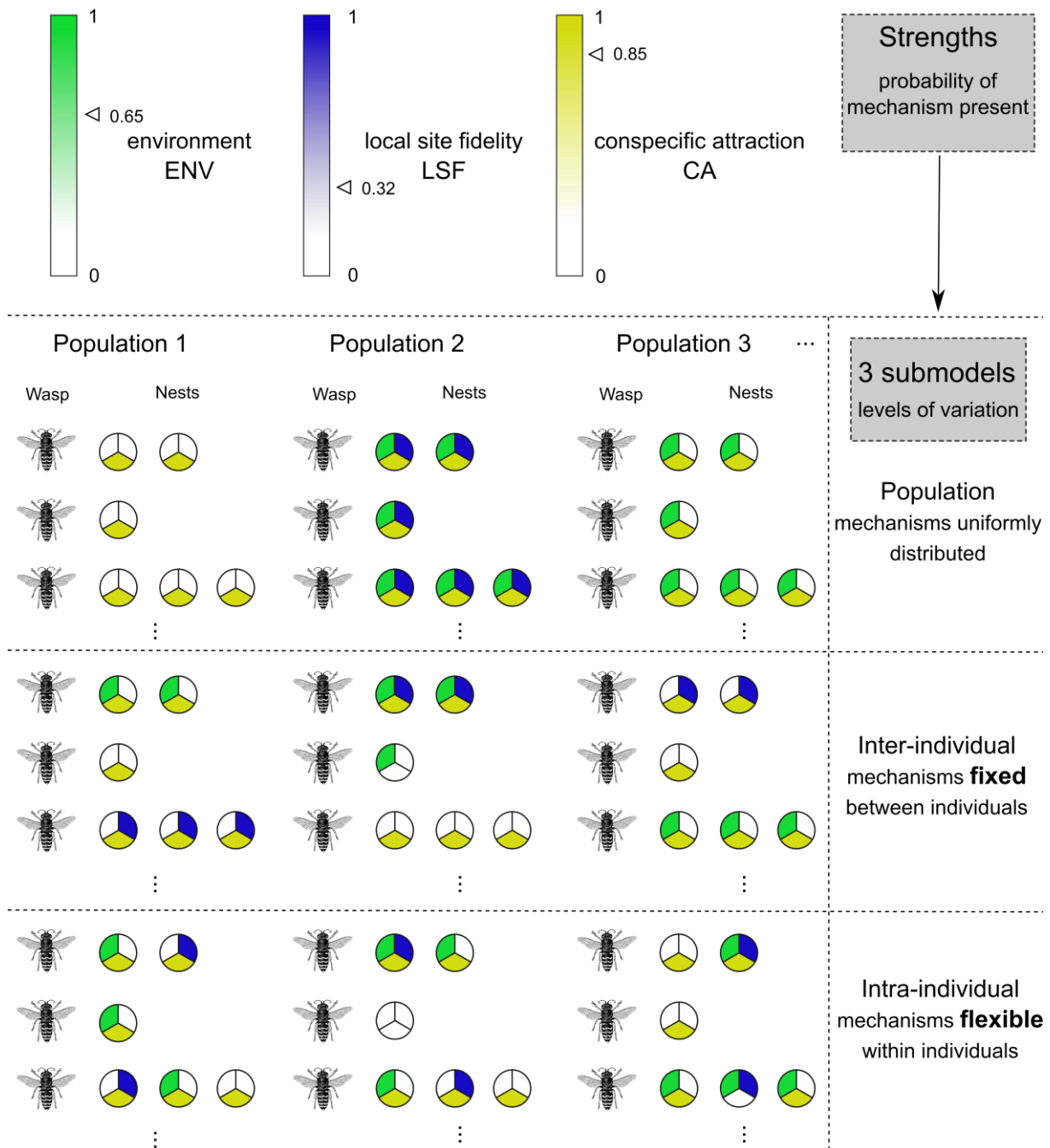


Figure S2.2: Conceptual diagram of *strengths* of the three mechanisms (environment, local site fidelity and conspecific attraction), which defines the probability a mechanism is present or absent. The presence of the mechanisms can be defined on three levels or submodels: population, inter-individual, intra-individual.

## 2.2. Emergence

Different search modes and changes in parameters *strengths* will give rise to different spatial pattern formations of aggregates. This will be compared using Approximate Bayesian Computation with field data to estimate parameters and derive the relative importance of the mechanisms.

## 2.3. Sensing

Wasps can search (1) the environment when trying to find a new nesting place. Wasps can also take into account (2) the position of their previous nest. Wasps can take into account (3) other nests surrounding a certain position. The wasps never have complete knowledge of the environment, but only of the evaluated position.

## 2.4. Interaction

With (1) the environment: when the wasp's search mode includes the environment, they will know the habitat suitability of the pixel under consideration. Their (2) previous nest: when local site fidelity is present, they can take into account the location of their previous nest. The probability of making a nest at a certain distance is according to a Gaussian curve. (3) Other wasps when conspecific attraction is present, the density of nests or wasps present is calculated within a certain range. The probability of making a nest at a considered location is density-dependent related with amount of wasps present, according to sigmoid function, a response function often found and used for density-dependent mechanisms (Kun and Scheuring 2006; Broly et al. 2016).

## 2.5. Stochasticity

Several steps include stochasticity, mostly where a choice is made randomly according to a calculated/given probability.

- Initialisation (see 3.1): values to initialize the state variables of the wasps are chosen randomly according to distributions from the field data.
- Presence of mechanisms: the presence of a mechanism during nest choice (possible on three levels, see Fig. S2.2) is stochastically chosen according to the *strength* of that mechanisms.
- During the nest location search of a wasp: positions that are evaluated for suitability are chosen randomly from the environment. Thus, the wasps never have complete knowledge of the environment or suitability of nesting locations.
- Nest choice: after a position is evaluated for its suitability, whether a nest is made at that location is chosen stochastically according to the calculated probability.

## 2.6. Collectives

Conspecific attraction is one of the mechanisms. It takes into account the number of nests within a radius of a considered position during nest site selection.

## 2.7. Observation



No subsampling of wasps or nests is made to select observed wasps/nests. The main output are the nest locations of each nest, together with its day it was build and which wasp made it. There are three output files for each run:

- The main output: on each row a nest location (xy-coordinates), day, wasp identity and the mechanisms used to make the nest.
- Distances (derived from previous): on each row a distance between two consecutive nests of an individual
- Parameters used in the model and state variables of the population

## **2.8. Design concepts not applicable:**

Adaptation/adaptive traits, fitness/objectives, learning, prediction

### 3. Details

#### 3.1. Initialization

Most initialization parameters are taken or calculated from the field data. The initialization does not include any emergent properties. It consists of boundary conditions/properties of the sampling method, environmental conditions, and size of the population. Thus when comparing summary statistics in the ABC-analysis, the underlying mechanisms that determine spatial patterning can be studied, and not the difference in internal state of the system.

Initialization parameters are:

- number of wasps: 432
- number of days: 30
- frequency distribution of number of nests per individual: to initialize the number of nests for each wasp in the population (see online code and Appendix S9).
- frequency distribution of starting day of first nest per wasp, which are the number of new wasps per day in the population: to initialize the day a wasp will make its first nest (see online code).
- frequency distribution of periods between subsequent nests of individuals: to initialize the period between two consecutive nests (see online code).

Parameters that are to be estimated with the model (see 3.3 submodels) are implemented through priors. The *strengths* of the mechanisms (Fig. S2.2) are chosen from a uniform distribution [0,1] for each run. The parameters of the response functions of the behavioral mechanisms are in a first run chosen from wide distributions (table S2.1) to narrow down possible parameter space for the actual analysis. A sensitivity analysis is performed for these parameters to evaluate the influence on the performance (ABC-analysis).

#### 3.2. Input data

Environment: this is the microhabitat suitability map made for the field study. A microhabitat model was made to explain presence and absence of nests of wasps with the factors *Insolation* (warmth) and *NDVI* (vegetation). This regression model was built with INLA (Lindgren & Rue, 2011; Martins, Simpson, Lindgren, & Rue, 2013; Rue, Martino, & Chopin, 2009), to correct for spatial autocorrelation. Thus, when the model is projected back on the complete study field, these predictions are corrected for spatial autocorrelation. This way, clustering due to other mechanisms than the environment (such as conspecific attraction) are corrected for in the predictions. The environment is a matrix (numpy-array) with values from 0-1 indicating the environmental suitability with one grid cell represent an area of 50×50cm<sup>2</sup>.

#### 3.3. Submodels

The submodels are the different levels at which mechanisms can be present (Fig. S2.2): (1) uniform across the population, (2) variable between individuals (fixed for one individual), (3) flexible within individuals (for different nests) or (4) random. The random model is the null-model. The environment is input data (see 3.2) from a different analysis. Local site fidelity and conspecific attraction are modelled with response functions that are parameterized during the ABC-analysis. Local site fidelity is implemented with a Gaussian response: the probability of nest selection at an evaluated location is related with the distance to the previous nest with a Gaussian function (Table S2.1). The parameters to be estimated is the width (standard

deviation) of the Gaussian function. For conspecific attraction, it is first counted how many nests (density) are present within a radius from the evaluated location, with the radius the first parameter. The probability of nest selection is then related to the density of nests according to a sigmoid function. This sigmoid function is defined by two parameters, which represent the probability where the density is zero and the stretch along the x-axis (standard deviation) of the sigmoid function. Details about the functions, parameters and initial priors can be found in Table S2.1.

### 3.3.1. Parameterization

Parameters of the mechanisms were implemented with uniform priors. To optimize convergence time, we applied uniform priors in an ecologically valid and broad parameter space rather than broad Gaussian priors that would be vague and cover ecologically non-valid values. Even though uniform priors should always be implemented with care (to not exclude valid parameter values), we provide below the rationale and resulting range of response functions with such priors.

Initial range of the uniform priors were first chosen very wide (Table S2.1), taking into account the size of the study area and descriptive values from the literature:

- Local site fidelity: (Larsson and Tengö 1989) found for the distances between consecutive nests a median distance of 1.5–2.5m, mean 3.5–6.5m. In our own study, this was median 4.51 m, mean 11.72m (skewed histogram see S8). We chose the prior of  $\sigma_{lsf}$  very wide such that 95% of the sum of probabilities is within 20cm to 20m from the previous nest.
- Range of conspecific attraction: we chose the prior of  $range_{ca}$  such that in a circle of diameter 40cm to 40m, the nests are counted to which the wasps are attracted to. This maximum distance covers a circle that holds a bit less than the width and half of the length of the study area (which is 47×97m<sup>2</sup>).
- Minimum density of conspecific attraction: we chose the prior of  $mindens_{ca}$  such that the probability at density=0 is between 0.00005 and 0.999
- Stretch of the sigmoid response curve: (Larsson 1986) found a maximum of 18 nests/m<sup>2</sup>, in our study we found a maximum of 12 nests/m<sup>2</sup>. We chose the prior of  $\sigma_{lsf}$  very wide such that the 0.5-probability occurs at 0.2 to 100 counted nests (depending on  $mindens_{ca}$ ).

Table S2.1: Response functions and parameters for the two behavioural mechanisms

Mechanism	Function	Parameters	Values of initials runs for the prior predictive check Priors: uniform	Values of actual runs Priors: uniform
<b>Local site fidelity</b> a subsequent nest is made close to the previous nest	Probability ~ <b>Gaussian function</b> with x the distance to the previous nest ( $\mu=0$ )  $P = \frac{1}{\sigma_{lsf}\sqrt{2\pi}} e^{-\frac{1}{2}\frac{x^2}{\sigma_{lsf}^2}}$	$\sigma_{lsf}$ : defines the width or range of the Gaussian function  quantity: distance (m)	[0.1, 10) ↓ 95% of $\Sigma P$ is within radius of $\pm 20$ cm $\rightarrow \pm 20$ m from the previous nest	[0.1, 10)
	Within the <b>circle</b> with radius $range_{ca}$ , number of nests are counted (#)	<b><math>range_{ca}</math></b> : defines the circle at which nests are taken into account for the conspecific attraction at the position under consideration  quantity: distance (m)	[0.2, 20) ↓ Nests within $\emptyset$ of 40cm $\rightarrow$ 40m are taken into account for conspecific attraction	[0.2, 7) ↓ Nests within $\emptyset$ of 40cm $\rightarrow$ 14m are taken into account for conspecific attraction
<b>Conspecific attraction</b> a location with other nests present attract new nests	Probability ~ <b>sigmoid function</b> for which $+\infty \rightarrow +1$ , where x is the number of nests divided by surface of circle, the density within $range_{ca}$  $P = \frac{1}{1 + e^{-(a+bx)}}$ $= \frac{e^{a+bx}}{1 + e^{a+bx}}$	<b><math>a = mindens_{ca}</math></b> : defines the probability at density=0 for which: $P(0) = \frac{1}{1 + e^{-a}}$  quantity: number of nests (#)	[-10, 10) ↓ $P(0) \in [4.54 \times 10^{-5}, 0.999)$	[-10, -3) ↓ $P(0) \in [4.54 \times 10^{-5}, 0.05)$
		<b><math>b = \sigma_{ca}</math></b> : defines the width or 'stretch' of the sigmoid curve, for which: $P(0.5) = \frac{-a}{b}$  quantity: number of nests (#)	[0.1, 15) ↓ $P(0.5) \in (-100, 100)$ Negative when $P(0) > P(0.5)$	[0.1, 15) ↓ $P(0.5) \in [0.2, 30)$ when $a=-3$ $P(0.5) \in [0.67, 100)$ when $a=-10$
The two behavioral mechanisms are <b>mutually exclusive</b>	Not applicable	<b><math>beh-excl</math></b> : Boolean parameter which sets the possibility that the two behavioral mechanisms are mutually exclusive. Implemented when a consecutive nest of the same wasp is made.	True/False	True/False

## Appendix S3: Prior predictive check

With a prior predictive check, we performed initial simulations to narrow down parameter space, avoid covariation between parameters and check if not too large biases for certain submodels were present. We simulated 25,000 runs of each submodel (including the null model; a total of 100,000 runs). A first ABC-analysis was run, not to perform model selection, but assess the difference between priors and posteriors and the correlation between parameters, especially the *strengths* and the parameters of the mechanisms (Fig. 1 main manuscript, Fig S2.2). This was done for all submodels together, and for the submodels separately (100 best models each). This approach was chosen to take into account the parameterization of all submodels, to not constrain the parameter space that would favour one of the submodels.

Figure S3.1 gives the pairs plots and Pearson's correlation coefficients between the different parameters. Figure S3.2–S3.4 gives detailed scatterplots. We decided to narrow the range of *range<sub>ca</sub>* and *mindens<sub>ca</sub>*, because they covary strongly with *strength<sub>env</sub>* and *strength<sub>lsf</sub>*: very high values of these two parameters are related with very low or very high *strength* values, especially for the fixed and flexible submodels. The narrowed prior parameter ranges can be found in the last column in Table S1, and as dotted lines in the scatterplots.  $\sigma_{ca}$  also covaries with these two *strengths*, but has more accepted values in the higher ranges for the uniform submodel. Thus, we did not narrow this down to not exclude possible parameter values that might be important for the parameterization of the uniform submodel.

$\sigma_{ca}$  is also correlated with *range<sub>ca</sub>* (fig S3.5). First, this is expected, as the larger the range, the more nests are counted in that range, resulting in a wider  $\sigma_{ca}$  for the response function. Second, as *range<sub>ca</sub>* is already narrowed down, narrowing  $\sigma_{ca}$  would be redundant.

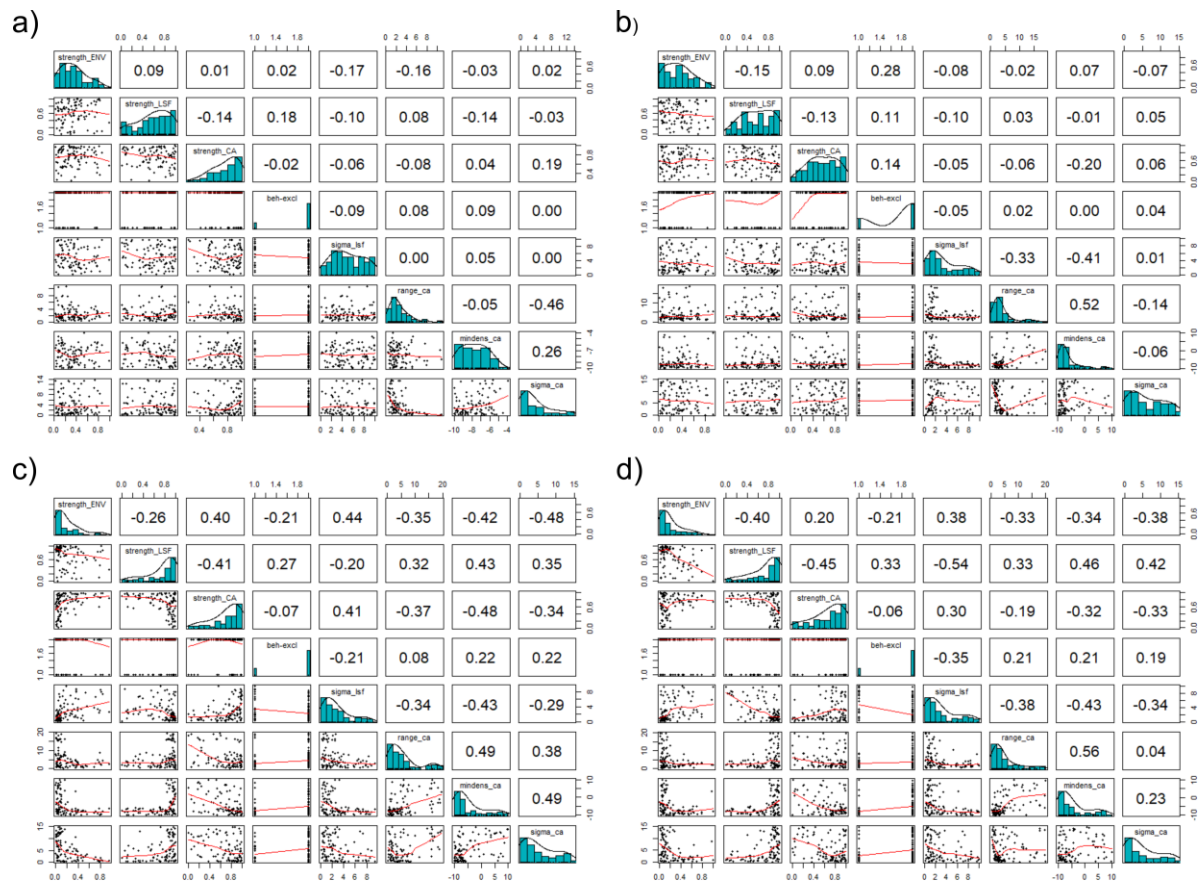
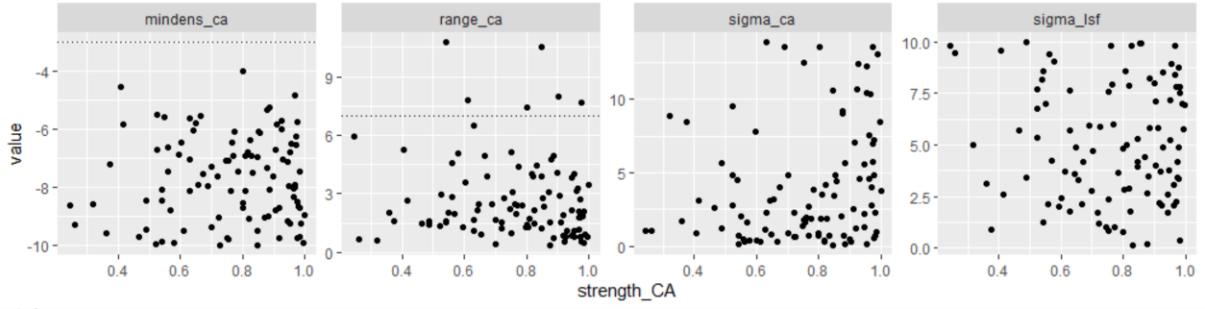
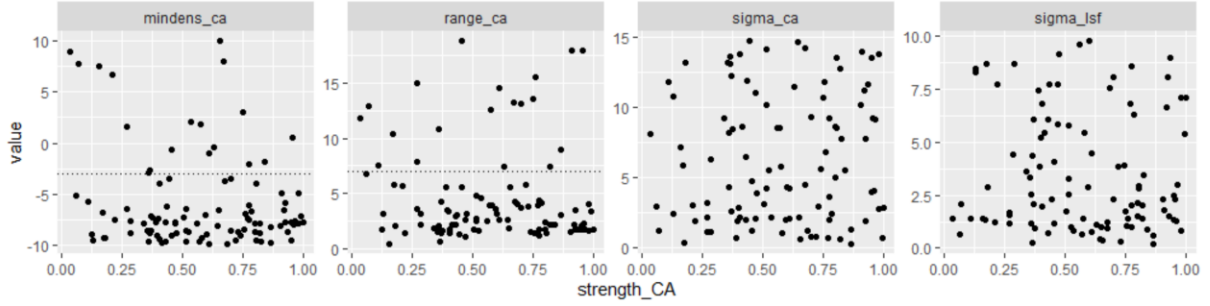


Figure S3.1: pairs plots and Pearson's correlations between the parameters for the prior predictive check.

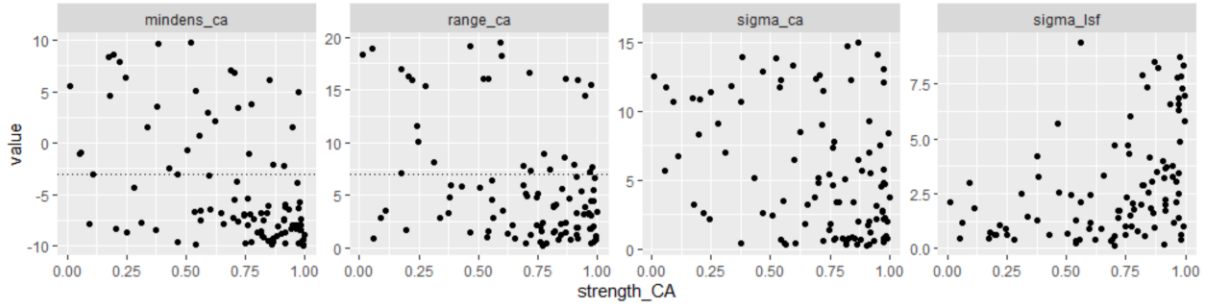
a) All models 100/100 000



b) Uniform 100/25 000



c) Fixed 100/25 000



d) Flexible 100/25 000

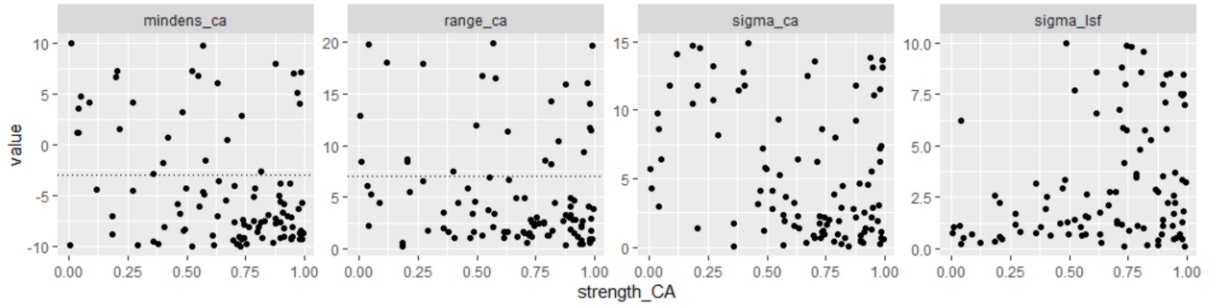
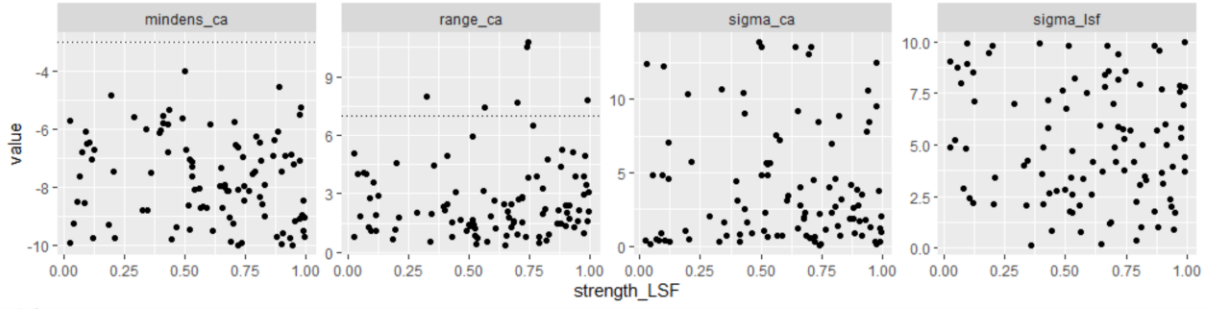
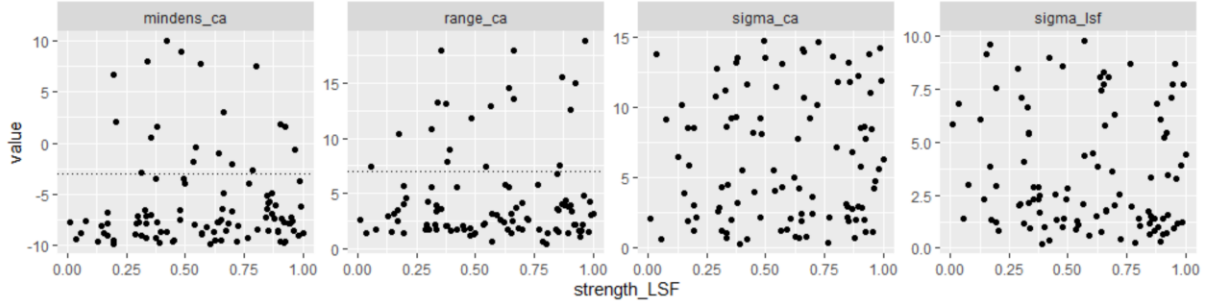


Fig. S3.2: scatterplots between the 4 parameters of the behavioral mechanisms and  $strength_{env}$  of conspecific attraction for the prior predictive check. ABC-analysis done for all submodels together (a), for the three submodels separately (b-d). Dotted lines are upper limit of the narrowed parameters (table S2.1).

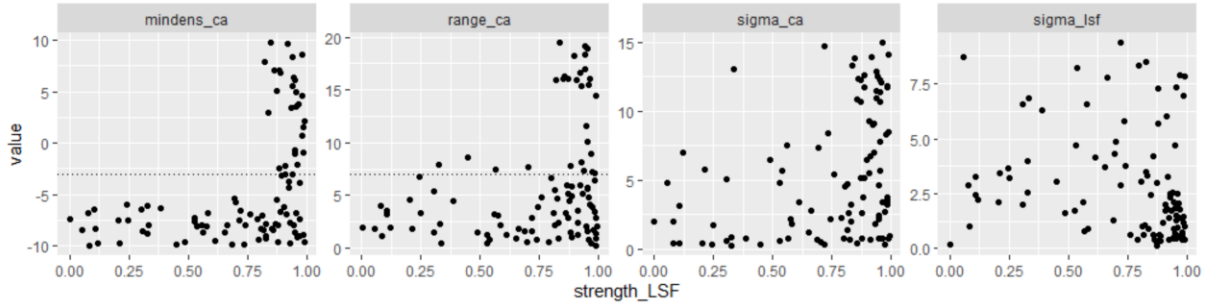
a) All models 100/100 000



b) Uniform 100/25 000



c) Fixed 100/25 000



d) Flexible 100/25 000

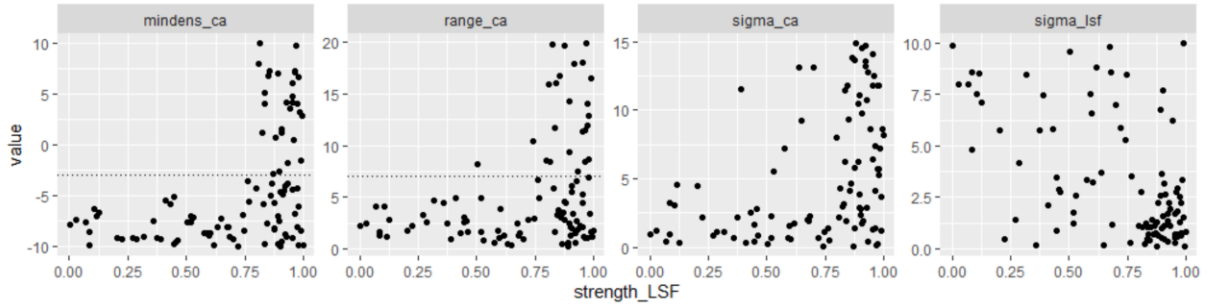
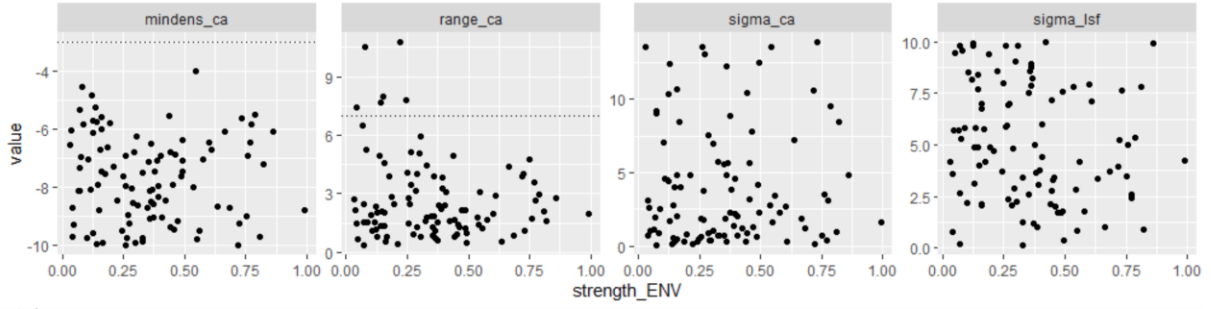


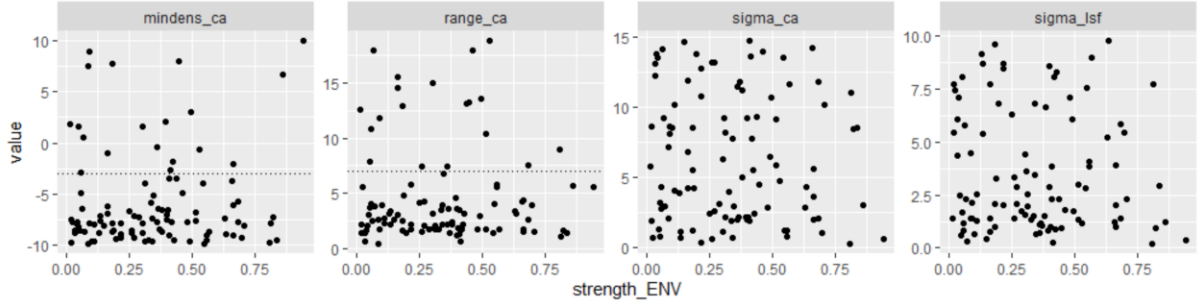
Fig. S3.3: scatterplots between the 4 parameters of the behavioral mechanisms and  $strength_{lsf}$  of conspecific attraction for the prior predictive check. ABC-analysis done for all submodels together (a), for the three submodels separately (b-d). Dotted lines are upper limit of the narrowed parameters (table S2.1).



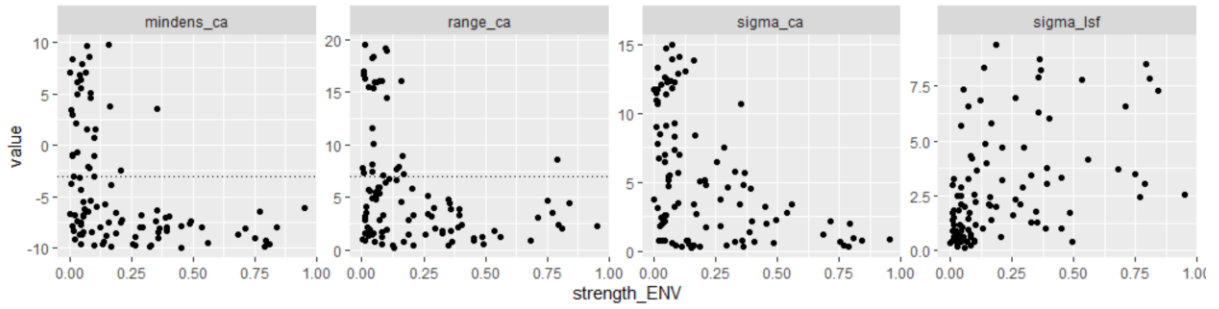
a) All models 100/100 000



b) Uniform 100/25 000



c) Fixed 100/25 000



d) Flexible 100/25 000

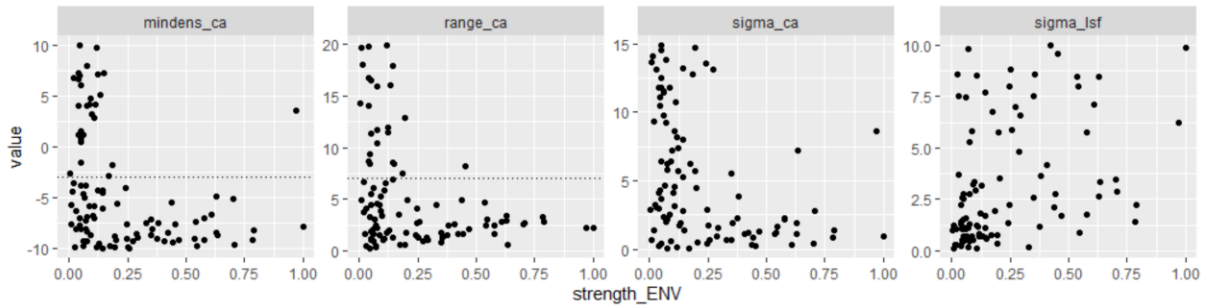


Fig. S3.4: scatterplots between the 4 parameters of the behavioral mechanisms and  $strength_{ca}$  of conspecific attraction for the prior predictive check. ABC-analysis done for all submodels together (a), for the three submodels separately (b-d). Dotted lines are upper limit of the narrowed parameters (table S2.1).

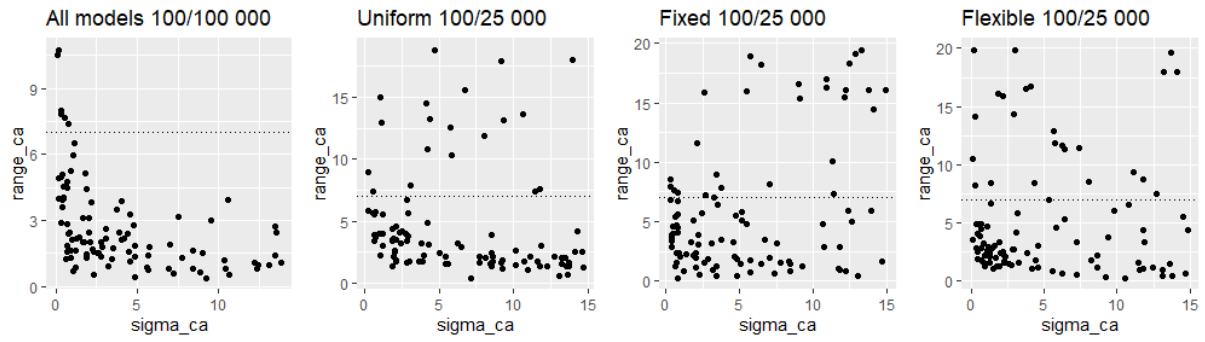


Fig. S3.5: scatterplots between  $range_{ca}$  and  $\sigma_{ca}$  for the prior predictive check. ABC-analysis done for all submodels together and for the three submodels separately. Dotted lines are upper limit of the narrowed parameters (table S2.1).

## Appendix S4: Cross validation of the ABC model selection

Cross validation was performed with the `cv4postpr`-function of the `abc` package in R (Csilléry et al. 2012). This functions performs a leave-one-out cross validation for model selection with ABC. We ran it for 50 simulations for each model ( $n_{val} = 50$ ). Details for the different submodels (ODD-protocol) can be found in Appendix S2.

Table S4.1 shows the mean model posterior probabilities of the cross validation of the 1000 best models. The rows are the true models and the columns how these models are classified by the ABC-analysis. Fig. S4.1 is a figure representing table S4.1.

All models are for the largest part classified as themselves by the cross-validation, apart from the inter-individual (fixed) and intra-individual (flexible) submodels. This is also reflected in the model selection with ABC (see main manuscript and appendix S5), where the two submodels are equally likely to produce the spatial patterning and network metrics.

Table S4.1: Mean model posterior probabilities of the cross validation. The submodels represent at which level the mechanisms can vary: population, inter-individual (fixed), intra-individual (flexible).

True model	Classification			
	Random	Population	Inter-individual	Intra-individual
Random	0.8407	0.1191	0.0193	0.0209
Population	0.0943	0.6047	0.1494	0.1516
Inter-individual	0.0057	0.1533	0.4159	0.4251
Intra-individual	0.0335	0.1198	0.4227	0.4239

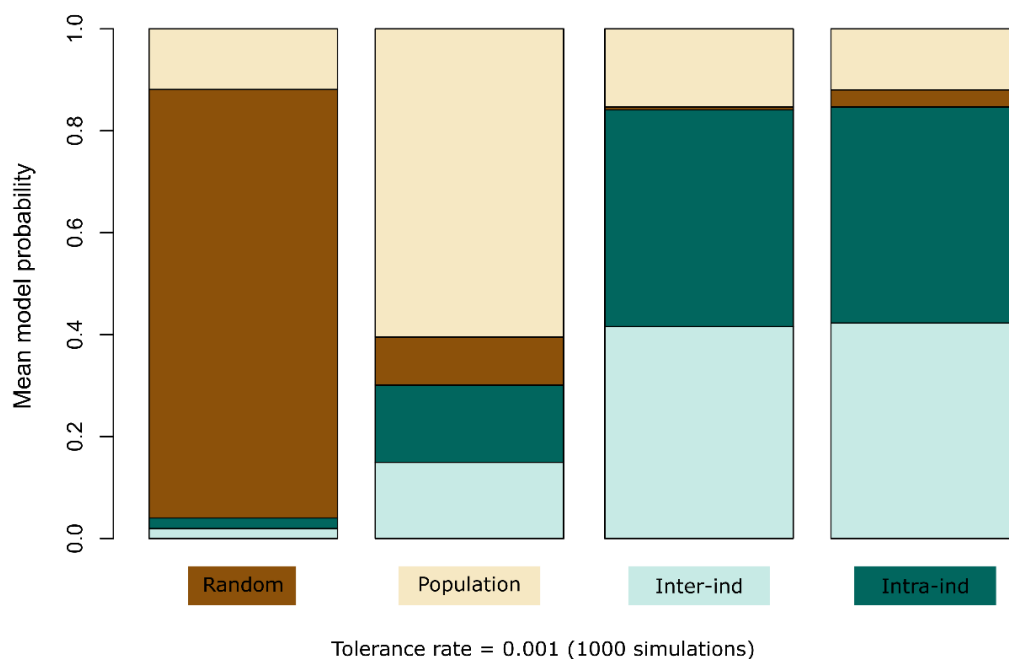


Figure S4.1: Visualization of confusion matrix: mean model posterior probabilities of the cross validation.

## Appendix S5: Convergence check of the IBM and ABC-analysis

The convergence of the IBM model was checked by repeating the ABC analysis for different acceptance rates. In the main manuscript 1000 simulations out of 1,000,000 simulations are used (rate of 0.1%). Here we give model-selection results with Bayes-factor tables for 10,000 (1%) in table S5.1, 500 (0.05%) in table S5.2, 100 (0.01%) in table S5.3 simulations accepted. Bayes' factors (BF) are the ratios of the posterior probabilities of two models, indicating the *strength* of evidence for model  $M_1$  (rows) relative to model  $M_0$  (columns), given the data. Evidence categories according to (Kass and Raftery 1995) are:  $BF < 1$  more evidence for  $M_0$  than  $M_1$ ;  $1 < BF < 3$  weak evidence for  $M_1$  compared to  $M_0$ ;  $3 < BF < 10$  substantial evidence for  $M_1$  compared to  $M_0$ .

Table S5.1: Bayes' factors and percent of accepted models for model selection with ABC-analysis retaining the 10,000 best simulations of 1,000,000 (1%). The submodels represent at which level the mechanisms can vary: population, inter-individual (fixed), intra-individual (flexible). Bayes' factors (BF) are the ratios of the posterior probabilities of two models, indicating the *strength* of evidence for model  $M_1$  (rows) relative to model  $M_0$  (columns), given the data.

	Random	Population	Inter-individual	Intra-individual
Random	-	0.00	0.00	0.00
Population	$\infty$	1.00	0.31	0.32
Inter-individual	$\infty$	3.25	1.00	1.05
Intra-individual	$\infty$	3.09	0.95	1.00
% accepted simulations	0	13.63%	44.26%	42.11%

Table S5.2: Bayes' factors and percent of accepted models for model selection with ABC-analysis retaining the 500 best simulations of 1,000,000 (0.05%). The submodels represent at which level the mechanisms can vary: population, inter-individual (fixed), intra-individual (flexible). Bayes' factors (BF) are the ratios of the posterior probabilities of two models, indicating the *strength* of evidence for model  $M_1$  (rows) relative to model  $M_0$  (columns), given the data.

	Random	Population	Inter-individual	Intra-individual
Random	-	0.00	0.00	0.00
Population	$\infty$	1.00	0.31	0.36
Inter-individual	$\infty$	3.27	1.00	1.18
Intra-individual	$\infty$	2.77	0.85	1.00
% accepted simulations	0	14.2%	46.4%	39.4%

Table S5.3: Bayes' factors and percent of accepted models for model selection with ABC-analysis retaining the 100 best simulations of 1,000,000 (0.01%). The submodels represent at which level the mechanisms can vary: population, inter-individual (fixed), intra-individual (flexible). Bayes' factors (BF) are the ratios of the posterior probabilities of two models, indicating the *strength* of evidence for model  $M_1$  (rows) relative to model  $M_0$  (columns), given the data.

	Random	Population	Inter-individual	Intra-individual
Random	-	0.00	0.00	0.00
Population	$\infty$	1.00	0.41	0.54
Inter-individual	$\infty$	2.42	1.00	1.31
Intra-individual	$\infty$	1.84	0.76	1.00
% accepted simulations	0	19%	46%	35%

## Appendix S6: Microhabitat suitability model: posterior distributions and spatial autocorrelation

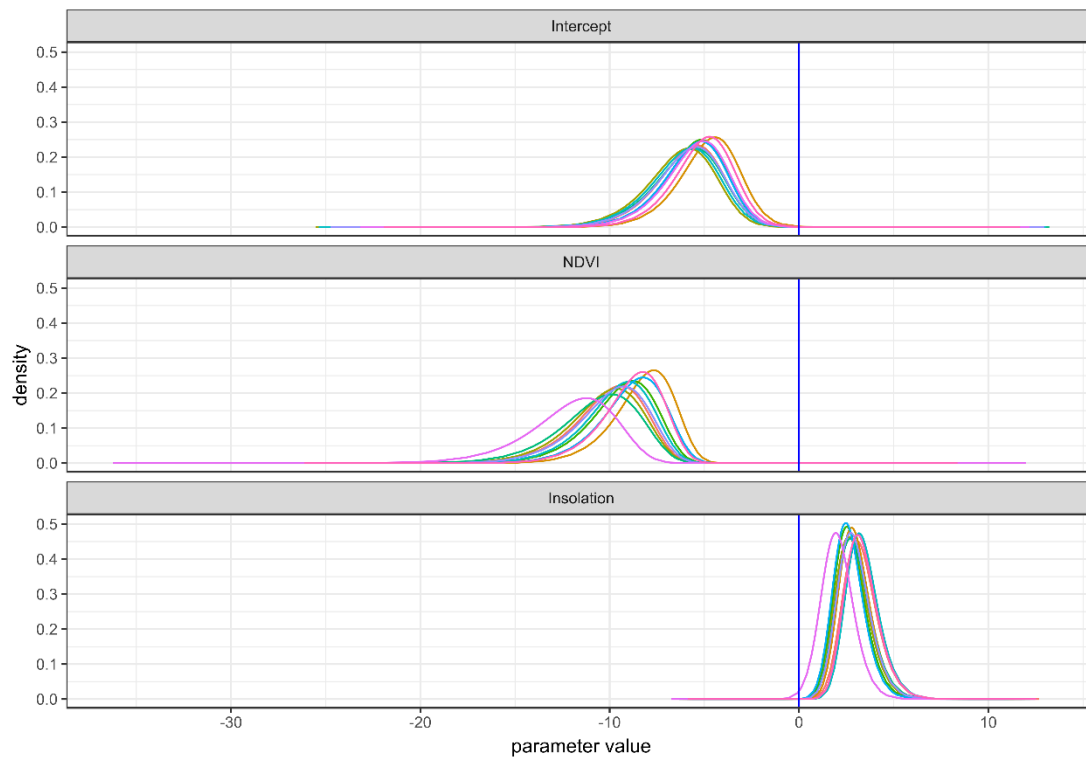


Figure S6.1: posterior distributions of the covariates (*intercept*, *NDVI*, and *Insolation*) for the 10 cross-validation runs of the model.

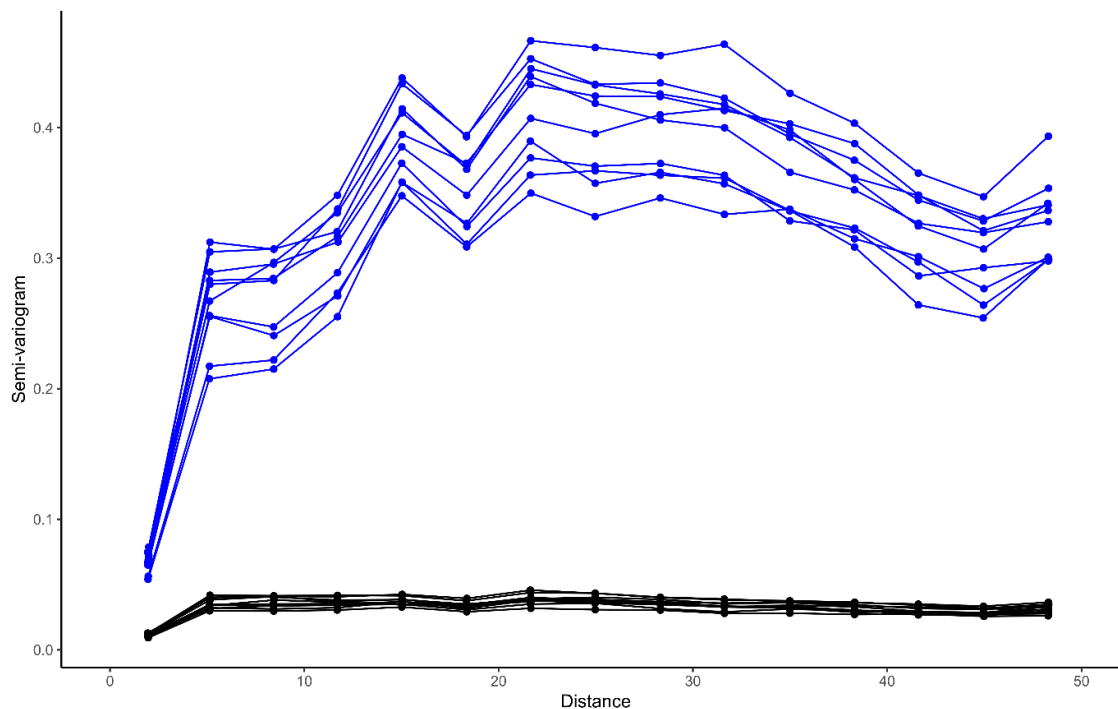


Figure S6.2: Variogram based on the Pearson's residuals for the 10 cross-validation runs of the model in INLA with (black dots and lines) and without (blue dots and lines) a spatially auto-correlated term included.

## Appendix S7: Deciding scale for microhabitat model

AUC, sensitivity, specificity and accuracy are given in table S6.1 for the scales that performed the best (> 2 m decreased in AUC). For each scale, cross-validation as described in the main manuscript was performed. Sensitivity and accuracy increase with increasing scale. All AUC's and specificities are similar. Figures S7.1–S7.5 show the predictions of the microhabitat suitability map for the 5 scales. Very small scales are quite granular, indicative of a certain amount of overfitting. The scale of 2 m gives the smoothest predictions. We chose the scale of 1m to proceed with in the main analysis, as this scale is more smooth than the three smallest scales, but shows more detail than the scale of 2 m.

Table S7.1: results for INLA-models for 5 scales (higher scales > 2m gave poorer results), based on 10 random cross-validations. AUC, Sensitivity, Specificity and Accuracy, including their standard deviations, are given (the latter 3 calculated for predictions transformed into 0/1 using the prevalence criterion;  $\pm 0.5$  in our case). In bold the scale that was chosen.

Scale	0.1 m		0.2 m		0.5 m		<b>1 m</b>		2 m	
	mean	sd	mean	sd	mean	sd	<b>mean</b>	<b>sd</b>	mean	sd
Mean auc	0.9501	0.0093	0.9461	0.0088	0.9459	0.0074	<b>0.9599</b>	<b>0.0128</b>	0.9546	0.0092
Sensitivity	0.6247	0.0421	0.6238	0.0598	0.6642	0.0654	<b>0.7394</b>	<b>0.0411</b>	0.7729	0.0901
Specificity	0.9749	0.0148	0.9703	0.0210	0.9575	0.0187	<b>0.9673</b>	<b>0.0213</b>	0.9633	0.0169
Accuracy	0.7998	0.0190	0.7971	0.0211	0.8109	0.0253	<b>0.8534</b>	<b>0.0191</b>	0.8681	0.0400

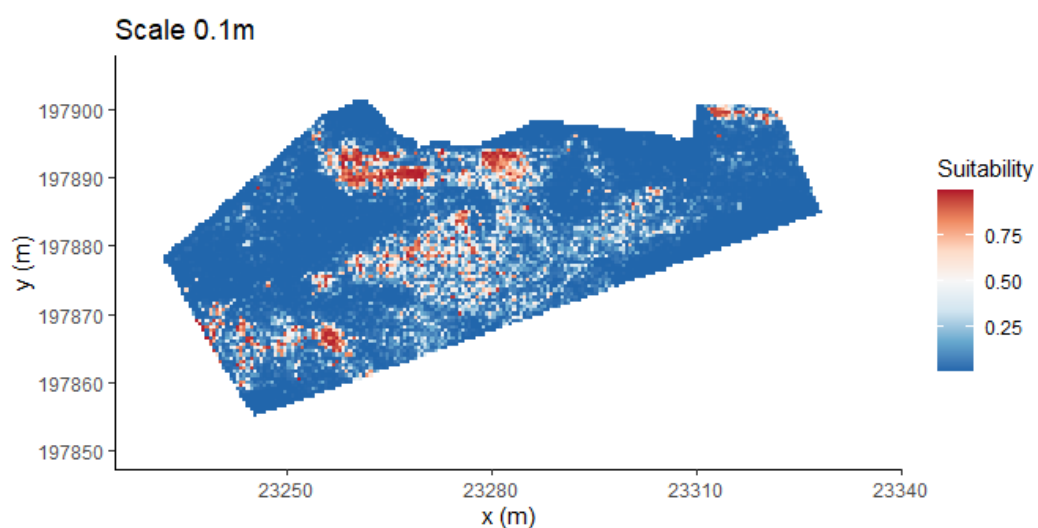


Figure S7.1: Suitability map for scale (buffers around the nests) of 0.1 m.

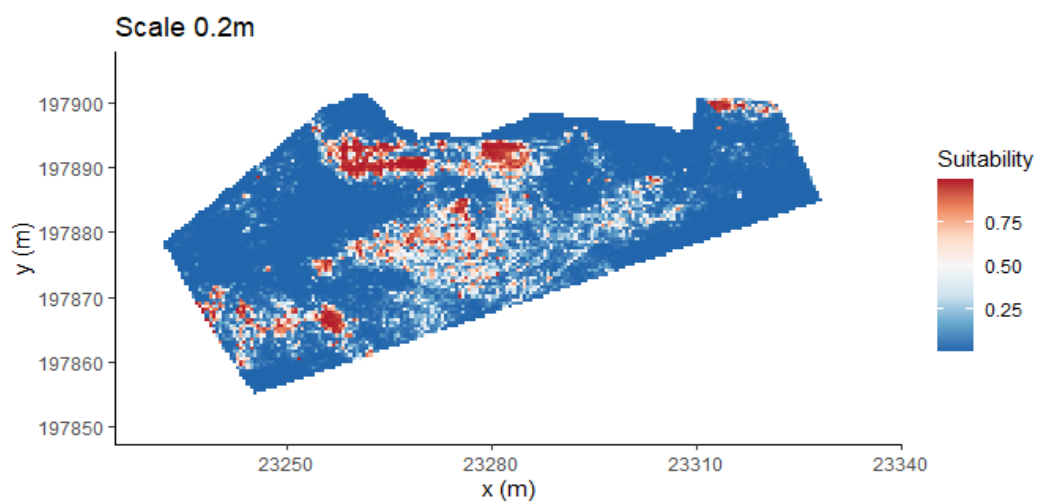


Figure S7.2: Suitability map for scale (buffers around the nests) of 0.2 m.

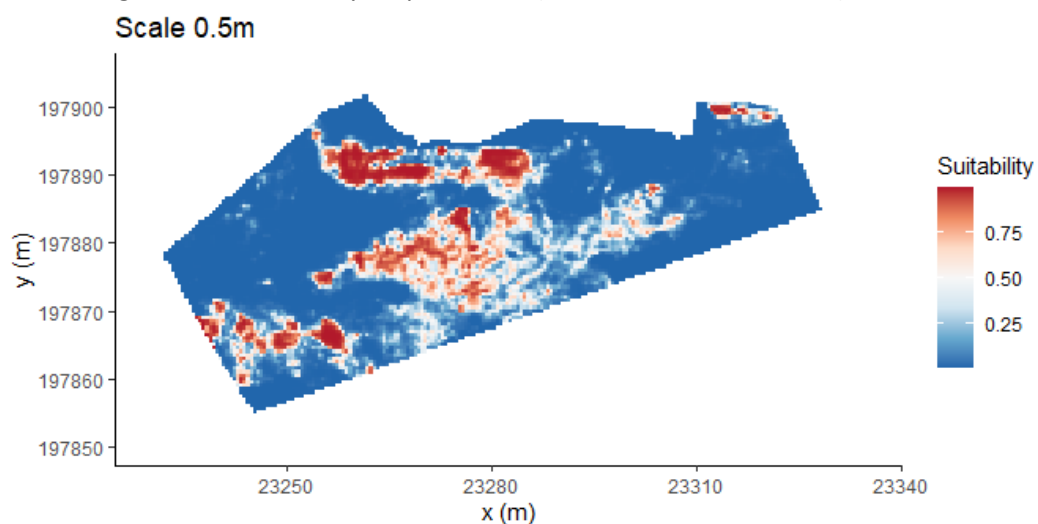


Figure S7.3: Suitability map for scale (buffers around the nests) of 0.5 m.



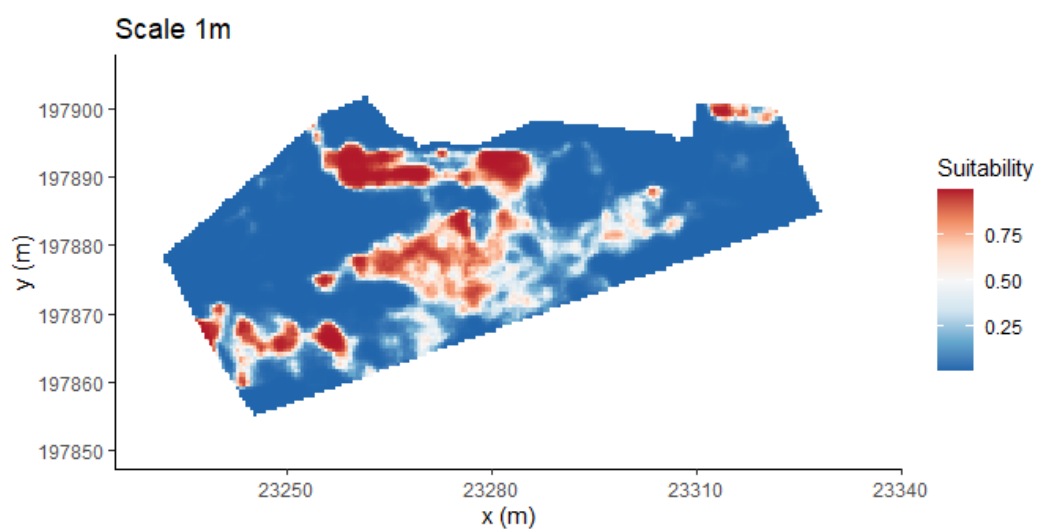


Figure S7.4: Suitability map for scale (buffers around the nests) of 1 m.

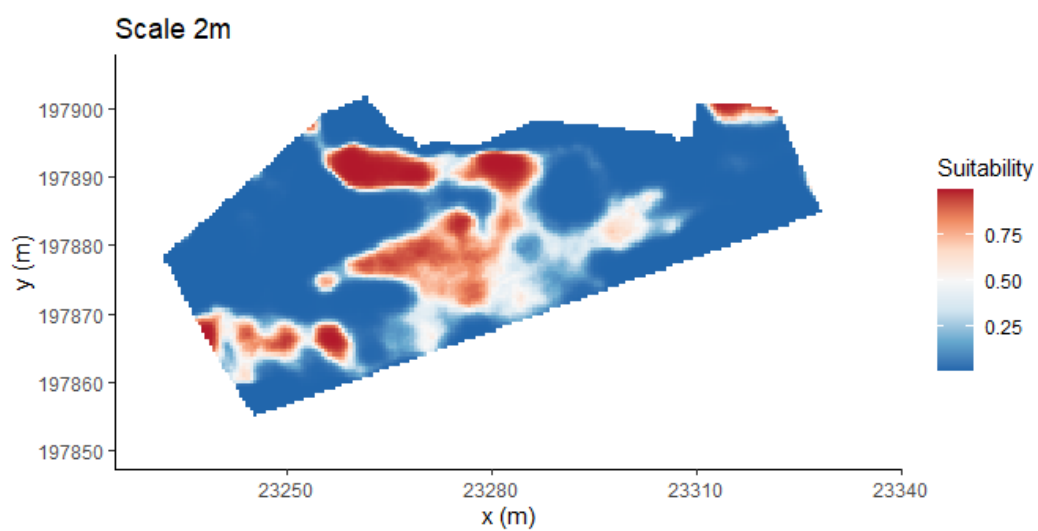


Figure S7.5: Suitability map for scale (buffers around the nests) of 2 m.

## Appendix S8: Deciding optimal number of aggregates with k-means cluster analysis

To divide the spatial pattern into aggregates, we performed a k-means cluster analysis. To decide on the number of clusters, we made an elbow-plot (Kassambara and Mundt 2020), visualized the clusters, looked at the stability of the cluster-assignment of nests and took into account the topography of the study area.

- 1) The elbow-plot (Fig. S8.1) suggests an optimal number of 8 or 9 clusters, as the curve starts to flatten there, taking into account the abnormal 'bump' in the curve at 7 clusters.
- 2) As an elbow-plot never gives full disclosure on how many clusters should be taken into account, we performed the k-means cluster analysis for cluster numbers between 5 and 12 and plotted these (Fig. S8.2).
- 3) We noticed that depending on the number of random sets (*nstart* in the *kmeans*-function), the spatial position of the clusters what not stable for cluster number up to 10 (example given in Fig. S8.3). For cluster number 11 this was stable.
- 4) We also took into account the topography of the study area (digital elevation model; Fig. S8.4), as this is something that is not taken into account during the k-means spatial cluster analysis (which works 2D, xy-coordinates). We compared this with the results of the different cluster-analyses to make sure that clusters on top of a hill (see center or top of map) and separate slopes were considered separate aggregates, as these were physically separated due to the topography.

Taking these factors into account (elbow-plot, spatial plotting, stability of cluster-assignment and topography), we decided on 11 to be the optimal number of clusters (Fig. S8.5).

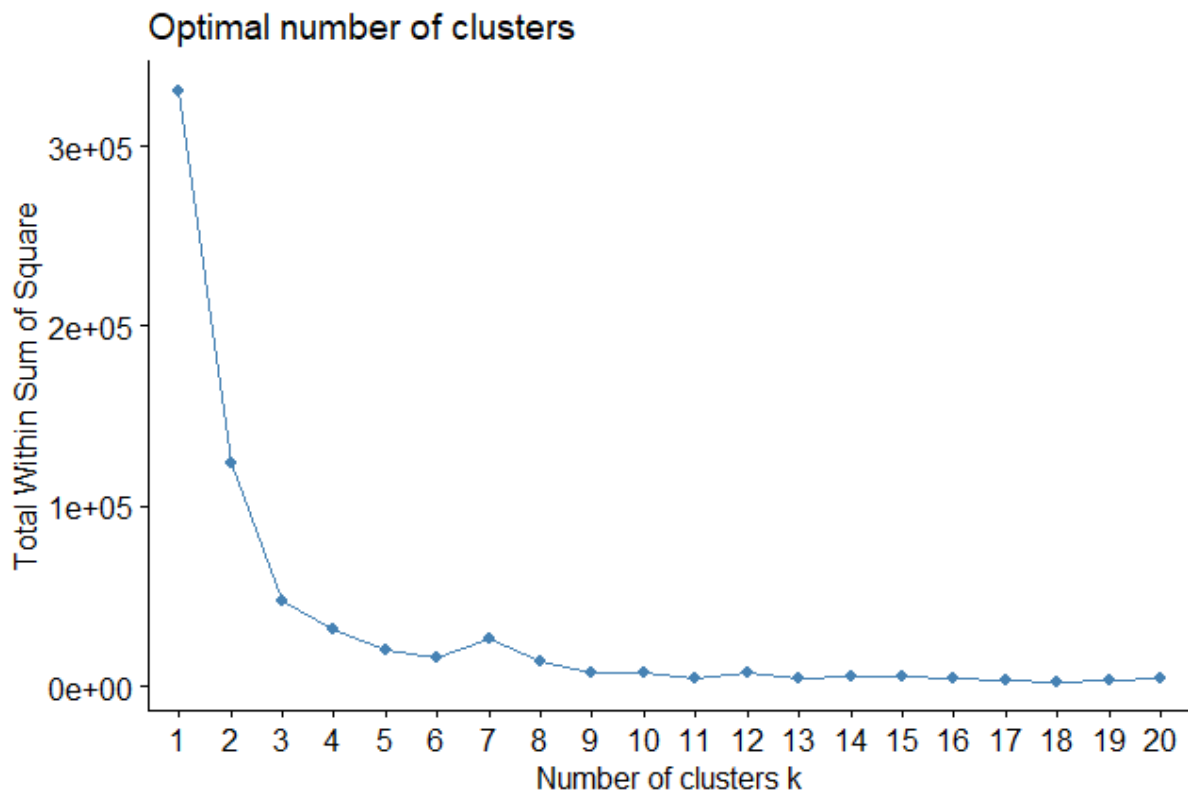


Figure S8.1: elbow-plot to define the optimal number of clusters for the k-means clustering.

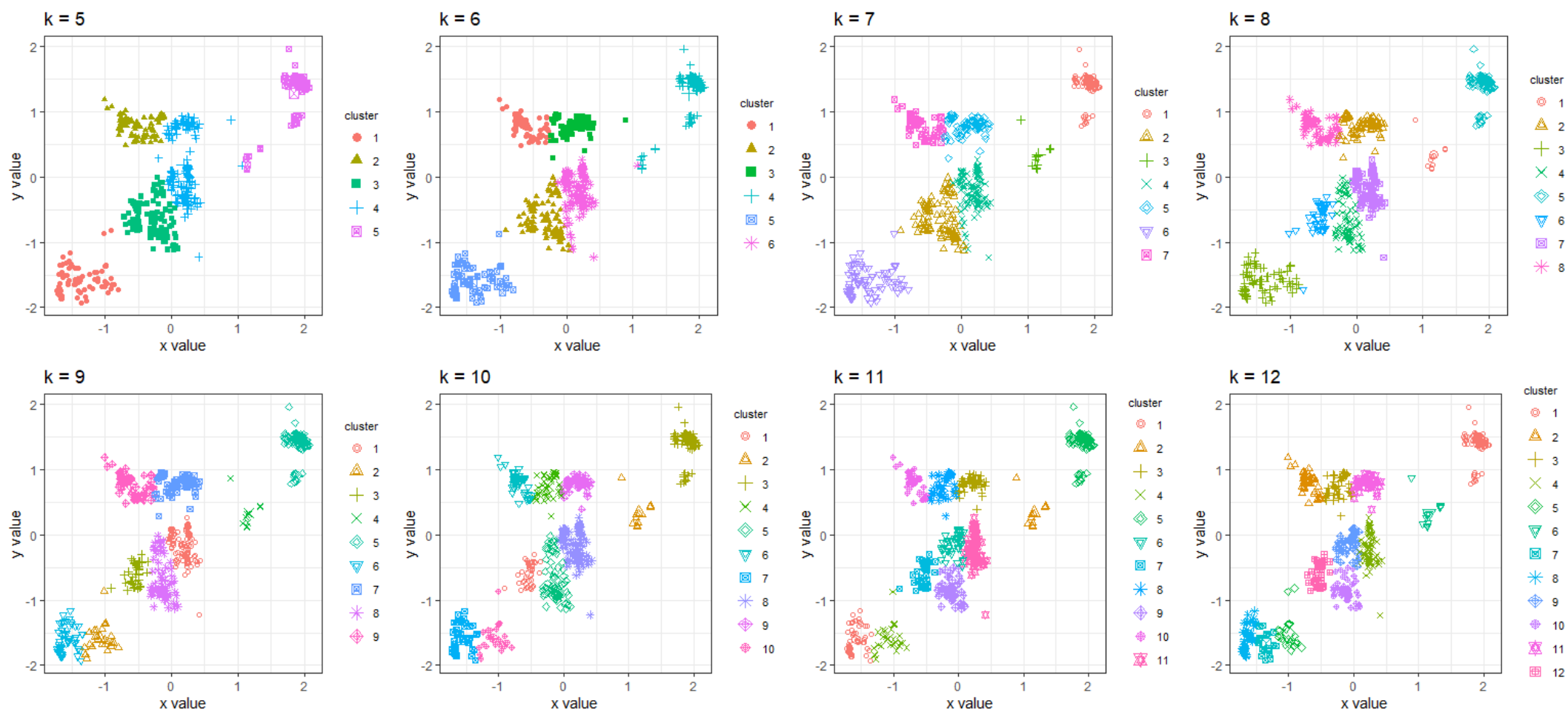


Figure S8.2: plotting of k-means cluster analysis for number of clusters between 5 and 12 ( $nstart = 20$ ).

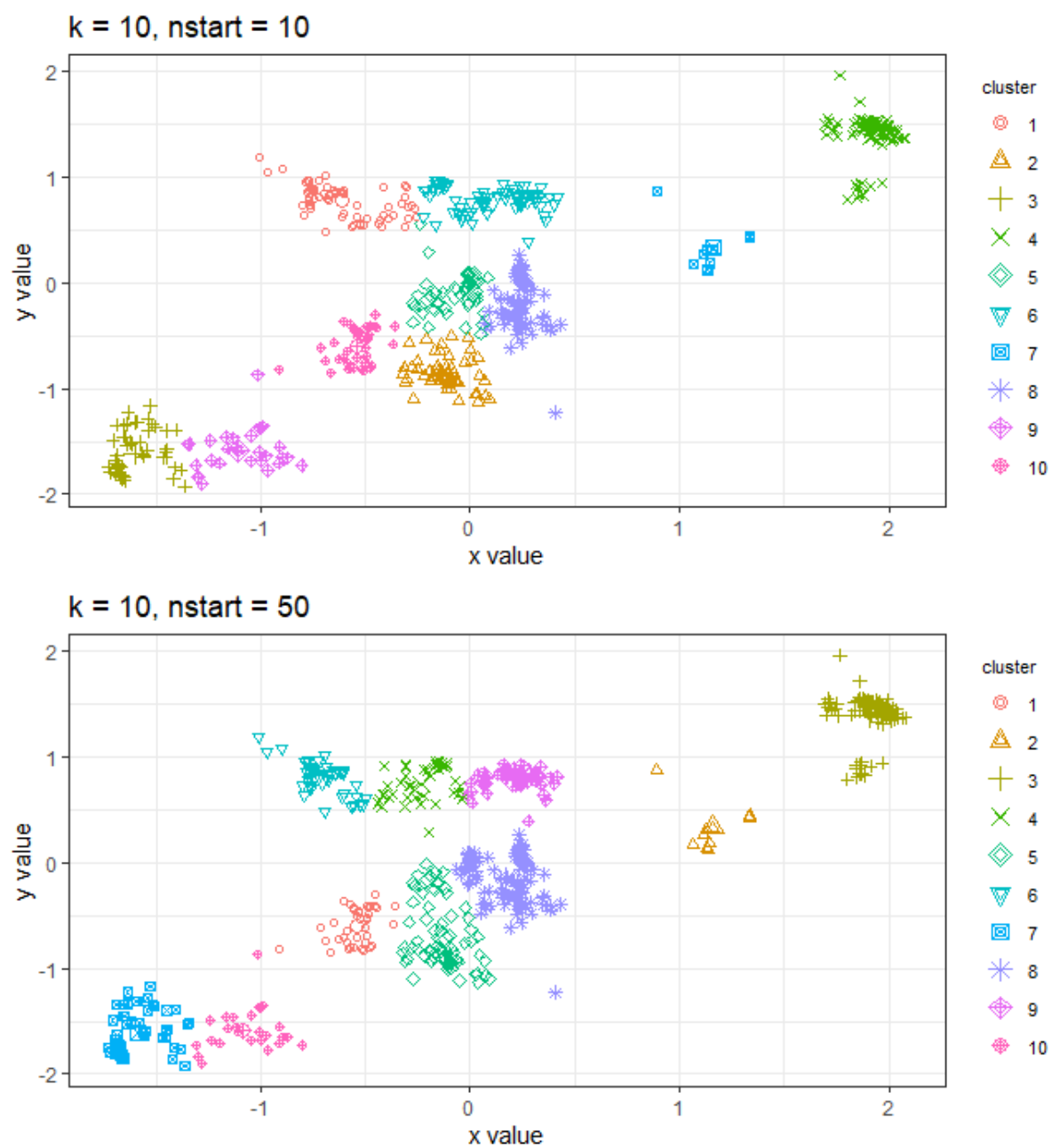


Figure S8.3: the effect of the number of random sets (*nstart*) on the cluster assignment of the nests.

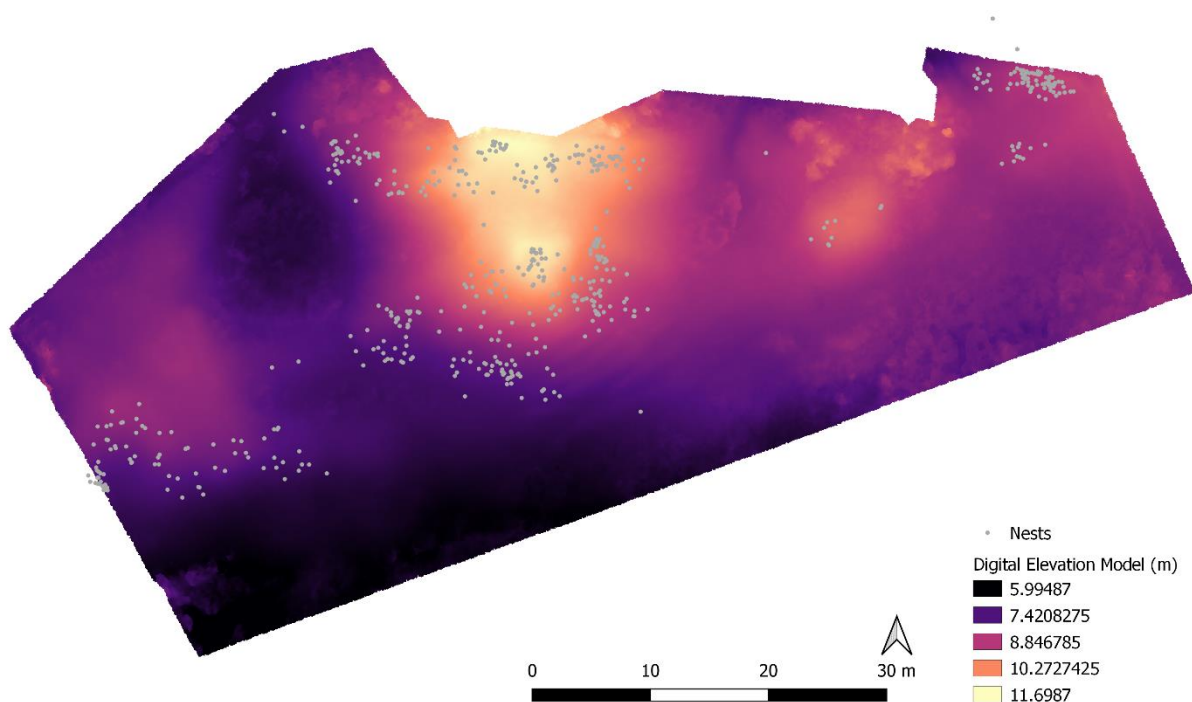


Figure S8.4: Digital elevation model (DEM) for the study area, with the positions of the nests marked with a grey dot.

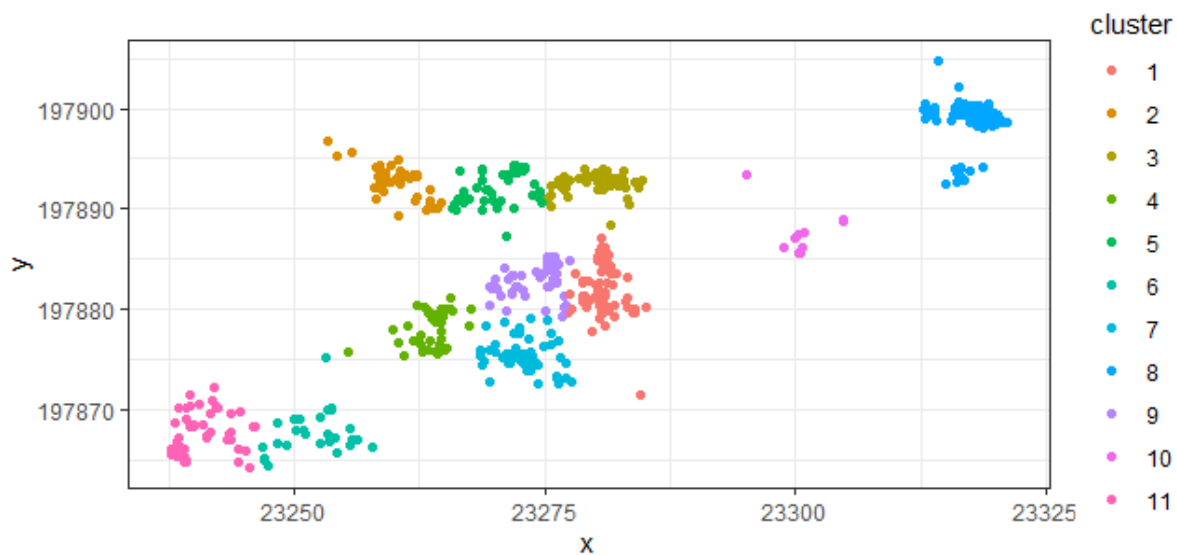


Figure S8.5: The spatial visualisation of the chosen clustering of the nests with  $k = 11$  ( $nstart = 20$ ).

Appendix S9: Information about consecutive nests in the field study

Table S9.1: Frequency of number of individuals having 1 to 4 labelled nests.

Number of nests	frequency
1	330
2	78
3	21
4	3

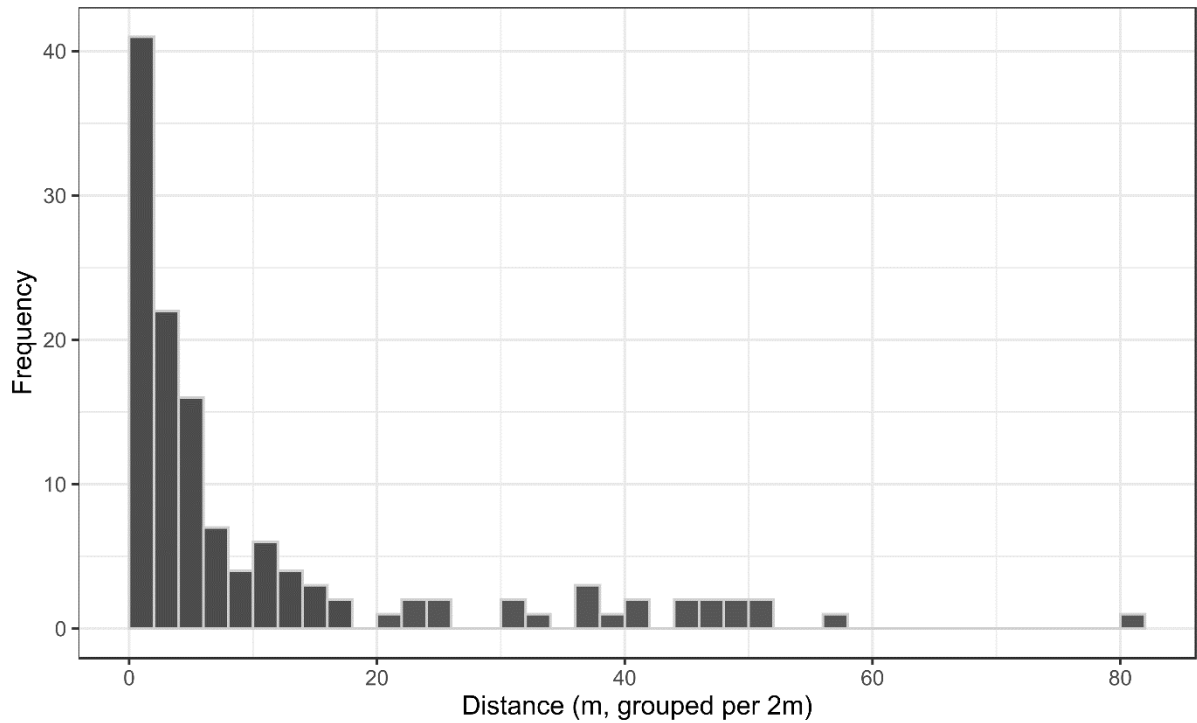


Figure S9.1: histogram of distances between consecutive nests, for a total of 129 links between previous and consecutive nests of a wasp.

## References used in the appendices

- Asís, J. D., Y. Ballesteros, J. Tormos, L. Baños-Picón, and C. Polidori. 2014. Spatial nest-settlement decisions in digger wasps: Conspecifics matter more than heterospecifics and previous experience. *Ethology* 120:340–353.
- Asís, J. D., J. Tormos, and S. F. Gayubo. 2004. Nesting behaviour and provisioning in *Bembix merceti* and *Bembix zonata* (Hymenoptera: Crabronidae). *Journal of Natural History* 38:1799–1809.
- Baddeley, A., E. Rubak, and R. Turner. 2015. *Spatial point patterns: methodology and applications* with R. Chapman & Hall/CRC.
- Beaumont, M. A. 2010. Approximate Bayesian Computation in evolution and ecology. *Annual Review of Ecology, Evolution, and Systematics* 41:379–406.
- Broly, P., Q. Ectors, G. Decuyper, S. C. Nicolis, and J.-L. Deneubourg. 2016. Sensitivity of density-dependent threshold to species composition in arthropod aggregates. *Scientific Reports* 6:1–13.
- Csilléry, K., M. G. B. Blum, O. E. Gaggiotti, and O. François. 2010. Approximate Bayesian Computation (ABC) in practice. *Trends in Ecology and Evolution* 25:410–418.
- Csilléry, K., O. François, and M. G. B. Blum. 2012. abc: an R package for approximate Bayesian computation (ABC). *Methods in Ecology and Evolution* 3:475–479.
- Dall, S. R. X., L. A. Giraldeau, O. Olsson, J. M. McNamara, and D. W. Stephens. 2005. Information and its use by animals in evolutionary ecology. *Trends in Ecology and Evolution* 20:187–193.
- Danchin, E., L. A. Giraldeau, T. J. Valone, and R. H. Wagner. 2004. Public information: from nosy neighbours to cultural evolution. *Science* 305:487–491.
- Hoi, H., A. Kristín, F. Valera, and C. Hoi. 2012. Traditional versus non-traditional nest-site choice: alternative decision strategies for nest-site selection. *Oecologia* 169:117–124.
- Kassambara, A., and F. Mundt. 2020. *factoextra: Extract and Visualize the Results of Multivariate Data Analyses*.
- Kun, Á., and I. Scheuring. 2006. The evolution of density-dependent dispersal in a noisy spatial population model. *Oikos* 115:308–320.
- Larsson, F. K. 1986. Increased nest density of the digger wasp *Bembix rostrata* as a response to parasites and predators. *Entomologia generalis* 12:71–75.
- Larsson, F. K., and J. Tengö. 1989. It is not always good to be large; some female fitness components in a temperate digger wasp, *Bembix rostrata* (Hymenoptera: Sphecidae). *Journal of the Kansas Entomological Society* 62:490–495.
- Philippi, T., and J. Seger. 1989. Hedging one's evolutionary bets, revisited. *Trends in Ecology and Evolution* 4:41–44.
- Polidori, C., P. Mendiola, J. D. Asís, J. Tormos, J. Selfa, and F. Andrietti. 2008. Female – female attraction influences nest establishment in the digger wasp *Stizus continuus* (Hymenoptera: Crabronidae). *Animal Behaviour* 75:1651–1661.
- Python Software Foundation. 2001. *Python Language Reference*, version 3.4. <http://www.python.org>.
- van der Vaart, E., A. S. A. Johnston, and R. M. Sibly. 2016. Predicting how many animals will be where: How to build, calibrate and evaluate individual-based models. *Ecological Modelling* 326:113–123.

Wcislo, W. T. 1984. Gregarious nesting of a digger wasp as a "selfish herd" response to a parasitic fly (Hymenoptera: Sphecidae; Diptera: Sarcophagidae). *Behavioral Ecology and Sociobiology* 15:157–160.