Behavioral strategies and the spatial pattern formation of nesting

Authors:

Femke Batsleer¹*(ORCID 0000-0001-5893-5966)

Dirk Maes² (ORCID 0000-0002-7947-3788)

Dries Bonte¹ (ORCID 0000-0002-3320-7505)

¹ Ghent University, Department Biology, Terrestrial Ecology Unit, Karel Lodewijk Ledeganckstraat 35, B-9000 Gent, Belgium

² Research Institute for Nature and Forest (INBO), Herman Teirlinckgebouw, Havenlaan 88 bus 73, B-1000 Brussel, Belgium

*Corresponding author: F. Batsleer, Ghent University, Department of Biology, Terrestrial Ecology Unit. Karel Lodewijk Ledeganckstraat 35, B-9000 Ghent, Belgium. E-mail: femke.batsleer@ugent.be

Short title: Behavior and nest spatial pattern

Key words: approximate Bayesian computation, habitat selection, INLA, spatial self-organization, Crabronidae, Hymenoptera
Abstract

Nesting in dense aggregations is common in central-place foragers, such as group-living birds and insects. Both environmental heterogeneity and behavioral interactions are known to induce clustering of nests, but their relative importance remains unclear. We developed an individual-based model that simulated the spatial organization of nest-building in a gregarious digger wasp Bembix rostrata. This process-based model integrates environmental suitability, as derived from a microhabitat model, and relevant behavioral mechanisms related to local site fidelity and conspecific attraction. The drivers behind the nesting were determined by means of inverse modelling in which the emerging spatial and network patterns from simulations were compared to those observed in the field. Models with individual differences in behavior that include the simultaneous effect of a weak environmental cue and strong behavioral mechanisms yielded the best fit to the field data. The nest pattern formation of a central place foraging insect cannot be considered as the sum of environmental and behavioral mechanisms. We demonstrate the use of inverse modelling to understand complex processes that underlie nest aggregation in nature.
Introduction

The ideal free distribution predicts that organisms optimally distribute themselves across resource patches to minimize resource competition (Kacelnik et al. 1992). This process leads to the sorting of individuals according to their niche (Hutchinson 1957; Holt 2009) and induces spatial patterns of high densities where the environment is most suitable (e.g. in shorebirds; Swift, Rodewald, & Senner, 2017). Animal aggregations are also widely documented in species inhabiting homogenous environments, not only in social species but also in non-social central-place foraging wasps, lizards and birds (Stamps 1988; Tarof and Ratcliffe 2004; Evans and O’Neill 2007). These inherent spatial patterns can emerge from behavioral and internal dynamics, such as the interplay between positive and negative density dependence, and are an example of spatial self-organization (Fortin and Dale 2005; Rietkerk and van de Koppel 2008; Bayard and Elphick 2010; Bradbury and Vehrencamp 2014). The spatial clustering of group-living animals, and more specifically nest clustering, has been explained through several behavioral hypotheses that are intrinsically linked with benefits related to group size (Krause and Ruxton 2002). Groups can provide protection against predation or parasitism (for example via a selfish herd; Hamilton 1971; Larsson 1986; Supplement S1) or against climatic extremes (Gilbert et al. 2008) or simply increase foraging efficiency (Clark and Mangel 1986). Both environmental and behavioral mechanisms can result in the spatial clustering of individuals or their nests. However, their relative importance remains elusive, especially for invertebrates. Moreover, it is unclear at which level these processes can vary: can mechanisms vary among individuals or even during an individual’s lifetime? The relative strengths of environmental and behavioral mechanisms for
spatial clustering are expected to vary among systems (for example, in analogy to bottom-up and top-down regulation of communities; Hunter and Price 1992).

Information use is central to any decision making, and thus also to settlement. Information can be personal when individuals directly use cues from the environment, or inadvertently social when information is generated by the behavior (e.g., foraging, fighting, mating) of other conspecific individuals (Danchin et al. 2004; Dall et al. 2005). Personal information, in addition to self-assessment of the environment, can consist of a female’s experience with previous nesting locations, which results in local site fidelity (Hoi et al. 2012; Asís et al. 2014). Variation in information use among individuals may arise from heterogeneity in these strategies where ‘producers’ rely on personal information and ‘scroungers’ on (inadvertent) social information, or from individuals switching between these sources of information (Barnard and Sibly 1981; Coolen et al. 2007).

The spatial clustering of nests is often regarded as a clearly separated and stepwise process where individuals first collectively select suitable environments at larger spatial scales after which internal dynamics (e.g. competition, attraction, individual and social learning) come into play (Melles et al. 2009; Swift et al. 2017). The prevailing insights are acquired by the analysis of complex spatial patterns and/or from behavioral experiments (e.g. Polidori et al. 2008; Melles et al. 2009; Asís et al. 2014). Environmental heterogeneity and internal dynamics are anticipated to act simultaneously, or even synergistically. Since the emergent patterns of a complex system cannot be predicted from the sum of the underlying components (Bradbury and Vehrencamp 2014), more integrated approaches are needed. Inverse modelling, which can identify the
processes that reproduce a set of observed patterns, has been extremely useful in this perspective (Banks et al. 2014; Curtsdotter et al. 2019; Stouffer 2019).

Here, we apply such an inverse approach to understand the contribution of environmental and behavioral mechanisms in nest aggregations of the ground-nesting digger wasp *Bembix rostrata*. We combine a microhabitat suitability model with an Individual-Based Model (IBM) to investigate the processes underlying the spatial dynamics of nest pattern formation as observed in the field. We include three mechanisms in the IBM: i) environmental suitability, ii) local site fidelity and iii) social cues. The direction and strength of these mechanisms can vary at the population-level (uniform for all), between (individually fixed) and within individuals (individually flexible). The simulated spatial point and network patterns are compared with those recorded in the field using Approximate Bayesian Computation (ABC) to select the most likely combination of environmental and behavioral mechanisms that underlie the observed nesting patterns.

**Material and Methods**

**Study species**

*Bembix rostrata* (Linnaeus, 1758) (Hymenoptera, Crabronidae, Bembicinae) is a specialized, highly philopatric, gregariously nesting digger wasp found in sandy regions of Europe. They inhabit sun-exposed sand dunes with sparse vegetation (Larsson 1986) and are sensitive to trampling (Bonte 2005). Adults are active from June to August; females construct one nest burrow at a time in which a single larva is progressively provisioned with flies (Nielsen 1945; Field 2005). A female can make a maximum of five nests each with one offspring (Larsson and Tengö
1989). Several kleptoparasitic fly species (Sarcophagidae) lay their larvae (ovi-larviposition; Piwczyński et al., 2017) on the prey provided by females of *Bembix* species (Nielsen 1945; Evans and O’Neill 2007). A selfish herd pattern has been observed in *B. rostrata* with regard to such brood-parasites (Larsson 1986), where the incidence of brood-parasitism per nest decreased with higher nest densities (we found a similar pattern in our field data between *B. rostrata* and *Senotainia albifrons*: Supplement S1).

**Study site and sampling**

Field data were collected in the summer of 2016 in the nature reserve De Westhoek in De Panne (51°04’38”N, 2°33’37”E, Belgium), in a study plot of approximately 40×90m². Surveys took place during 30 days of favorable (sunny and warm) weather conditions for *B. rostrata* (Schöne and Tengö 1991) between 28 June and 15 August. Female wasps were individually tagged with a colored and numbered plastic plate on the thorax (Opalith Zeichenplättchen) and nests were marked with small, handcrafted flags. We recorded visually when an individual started a nest, entered the nest with a prey, and if the prey was ‘infected’ by kleptoparasitic flies (*Senotainia albifrons*, Miltogramminae, Sarcophagidae). The study area was covered several times per day, to sample each nest aggregate as equally as possible. The position of each nest was measured with a Trimble GPS (accuracy of 2cm; Trimble Inc., USA). Remote-sensing imagery was collected using a drone (Rpaswork.com and Didex.be) equipped with a multispectral camera (Red, Green, Blue and Near Infrared bands) at the end of the flight season. These images are processed towards a digital elevation model (DEM; pixel size 2.4×2.4cm²) and the Normalized Difference
Vegetation Index (NDVI; pixel size 1.1×1.1cm²) (Pettorelli 2013). Insolation (as an indicator of the microclimate; pixel size 7.2×7.2cm²) is calculated from the DEM with the ‘solar radiation tool’ and slope (pixel size 7.2×7.2cm²) with the ‘surface toolset’, both extensions of ‘spatial analyst’ in ArcGIS (ESRI 2011).

**Statistical analyses**

The workflow of the analyses is shown in Fig. 1.

*The microhabitat suitability map*

We used Integrated Nested Laplace Approximation (INLA) (Rue et al. 2009; Lindgren and Rue 2011; Martins et al. 2013) to build the microhabitat suitability model. INLA is a Bayesian approach that allows for spatially auto-correlated residuals of the environmental data related to nest location (Zuur et al. 2017) (Supplement S6). We used a generalized linear mixed model with binomial distribution with logit-link for the response variable and a spatial dependency structure modelled with the Matérn covariance function (see online code). NDVI (vegetation) and insolation (microclimate) were used as (normalized) explanatory variables. Every nest was considered as a presence point and absence/zero data points were generated by selecting an equal number of random points that were at least 1m from any nest within the study plot. As the study plot was searched intensely, we considered the generated points as true absences. Models were compared using Watanabe-Akaike information criterion (WAIC: Watanabe 2010; Gelman
et al. 2014), computed within the *inla* function (Rue and Held 2005; Rue et al. 2009). To confirm our a priori choice of a simple linear model, we considered interactions between both variables and an additional covariate, local slope. WAIC differences for these models with a linear model only including *NDVI* and *insolation* were less than 3, so the simple linear model was preferred. The data were split into 70% training and 30% evaluation data. As a cross-validation, the final model was run 10 times using different randomly chosen training and evaluation sets each time. To assess the predictive power of the models, AUC (area under the curve) was calculated with the R-package *ROCR* (Sing et al. 2005). Sensitivity (true positive rate; predictive performance of presences), specificity (true negative rate; predictive performance of absences) and balanced accuracy (overall true rate) were calculated with the R-package *caret* (Kuhn 2008). To calculate the latter three performance measures, predictions were transformed into 0/1 using the prevalence criterion (Manel et al. 2001; Liu et al. 2005): predictions that are larger than the prevalence threshold (proportion of presences/absences in the evaluation datasets: ±0.5 in our case) are classified as 1 and the other predictions as zero. The plotting of the spatial field, the spatial residuals which INLA corrects for, shows if the degree of clustering was a higher (hot spots) or a lower (cold spots) than expected based on the covariates (*NDVI* and *insolation*) in the microhabitat model. Different models that considered different spatial scales were compared (each with cross-validation included): buffers between 0.1m and 10m were drawn around each nest at 9 different radii and the mean of each variable was calculated inside those buffers (QGIS Development Team 2020). The models with buffer scales 0.1, 0.2, 0.5, 1 and 2m had similar WAIC, AUC, sensitivity, specificity, and spatial field plots. We proceeded with the 1m scale, as the suitability predictions of this buffer scale were detailed but also smooth, to balance overfitting.
and poorer estimates (Supplement S7). Within each cross-validation, predictions of the final model were projected back onto the field study plot within a grid of 0.5×0.5m². A detailed habitat suitability map was created with the average of these predictions (with a 0 to 1 scale of probability of nest presence). This map was used as input for the IBM model (Fig. 1 and see section ‘Inverse modelling with IBM and ABC’) to be used as the environmental cue for habitat selection. We did not include the uncertainties of the probabilities in the IBM model, as we deem this extra level of stochasticity negligible.

**Spatial point pattern and network analyses**

A point pattern analysis was carried out with the R package Spatstat (Baddeley et al. 2015). Spatial clustering of nests was investigated using Ripley’s $K$ at scales between 0 and 40m, where a higher $K$ than the calculated expected random distribution at a certain scale or radius is indicative of a clustered pattern within that radius, and a lower $K$ of a regular pattern within the radius (Baddeley et al. 2015). Ripley’s K-values were transformed to represent the relative change compared to complete spatial randomness (CSR) at a scale $r$ with the formula $K_{rel}(r) = (K(r) − CRS(r))/(CRS(r))$.

To assign the nests to different nest aggregates, a k-means cluster analysis was implemented. The optimal number of nest aggregates was 11, considering an elbow-plot (Kassambara and Mundt 2020), visualization of the clusters and topography of the area (Supplement S8).
A network analysis was carried out with the R package *igraph* (Csardi and Nepusz 2006). Nest aggregates from the k-means cluster analysis were considered network nodes in the network analysis and the consecutive nests of individuals as links (or edges) between network nodes. As such, the network nodes are aggregates that are spatially grouped nests and the individuals moving among (and within) the aggregates to a consecutive nest are the links of the network. Five network metrics were calculated for this directed network, defined according to Farine & Whitehead (2015): i) the number of loops, which is the total number of links or subsequent nests of individuals; ii) the number of internal loops, which is the relative number of links that return to the same node, thus an individual that makes consecutive nest in the same aggregate; iii) transitivity (or clustering coefficient), which quantifies how densely nodes are connected: a high transitivity indicates that triads (trios of nodes) have a high degree of being mutually linked; iv) density (or connectance), which is the number of links divided by the total number of possible links between all clusters; v) reciprocity, which is the relative number of reciprocal links between nodes.

*Inverse modelling with IBM and ABC*

We developed an individual-based model (IBM) to simulate and eventually identify the potential drivers behind the species’ nesting dynamics using a pattern-oriented approach (Grimm et al. 2005). Individual wasps, with their different sets of behaviors, are the entities of simulation within a spatially explicit environment. The ODD (Overview, Design concepts, Details) protocol
(Grimm et al. 2006, 2010) is added in Supplement S2, where the detailed explanation, assumptions and parameters are explained. Here we briefly discuss the general set-up.

The three mechanisms—environment, local site fidelity and conspecific attraction—were combined in the model using strengths: the probabilities of the mechanisms being present. Variation in the presence of the mechanisms was possible at three levels as modelled in different submodels or strategies (Fig. 1): 1) the population level: the mechanisms used are uniform across all the individuals in the same population; 2) inter-individual: the mechanisms can vary among individuals in a population but are fixed for an individual; 3) intra-individual: the mechanisms can vary within an individual’s lifetime and are thus flexible. The null model in which random locations were chosen within the study area was used as the fourth submodel. The flow for an individual *Bembix* female when selecting a nesting site was as follows: first, a random position in the area is sampled; then, the average suitability according to the focal mechanisms is assessed after which that position can be stochastically selected according to the calculated suitability (or probability). When the position is not selected, a new one is sampled according to the same procedure (see ODD protocol Supplement S2).

The habitat suitability map serves as a baseline for the environmental cue: the suitability values are used as probabilities for settling. Local site fidelity is implemented as a Gaussian distribution centered around the previous nest, with one parameter \( \sigma_{lsf} \) defining the width of the distribution. As such, positions closer to the previous nest have a higher probability of being chosen. Conspecific attraction is coded in two steps: first, the parameter \( range_{ca} \) defines the radius of the circle in which the number of other nests are counted. Second, settlement probability is implemented with a sigmoid function (Kun and Scheuring 2006; Broly et al. 2016),
with the number of nests counted in the first step as the dependent variable. Two parameters define the sigmoid curve: \( \text{mindens}_{ca} \) is the intercept and \( \sigma_{ca} \) the scale parameter of the function. The Boolean parameter \( \text{beh-excl} \) defines if conspecific attraction and local site fidelity are mutually exclusive: both mechanisms could be strongly present, while not jointly determining an individual’s nest site selection. To optimize convergence time and remain within reasonable ecological boundaries, we applied uniform priors in a valid parameter space (Supplement S2, Table S2.1).

To initialize the model and to define boundary conditions, the following properties derived from the field study were used: total number of individuals sampled in the study site (432); total number of days to run a simulation (30); distribution of the number of nests initiated each day; distribution of the number of nests per individual; and distribution of time between subsequent nests. The latter three are used as probability distributions when initializing the \textit{Bembix} population (Supplement S1, section 3.1; online code).

We verified that the priors were not biasing the analysis towards one of the submodels, by setting them widely for 100,000 simulations (prior predictive check; Supplement S3). Following this analysis, we restricted prior ranges by excluding those ranges where parameters covaried. This step ensures that certain parameter values are not redundant and improves convergence of the actual simulations.

Each one of the four submodels was run 250,000 times with parameters randomly sampled in the prior parameter ranges (Supplement S2, table S2.1). Summary statistics for each
model simulation were calculated as described in the above section ‘Spatial point pattern and network analyses’.

The model was evaluated using Approximate Bayesian Computation (Beaumont 2010; Csilléry et al. 2010; van der Vaart et al. 2016), more specifically rejection-ABC following van der Vaart et al. (2015) with the R package abc (Csilléry et al. 2012). This method is based on the difference between each simulation and the observed field data in the summary statistics of patterns of interest. As this method cannot compare summary statistics that are continuous functions, such as Ripley’s K, values of Ripley’s K at a discrete set of distances were chosen as part of the summary statistics. The complete set of summary statistics were six Ripley’s K values transformed as described in the previous section (at distances between 2 and 40m) and five network metrics (number of total loops, number of internal loops, transitivity, density, and reciprocity), as described in the previous section. The sum of the differences between normalized summary statistics of field and simulation data was calculated. The summation was weighted with the complement of the average Pearson’s correlations (1 − ρ) for each summary statistic, calculated from all simulations. As such, the summary statistics are corrected for their dependence structure. Minimizing this distance, the 1000 simulations (0.1%) closest to the observed field data were retained. We calculated the percentage of accepted simulations for each of the three submodels. Pairwise Bayes factors were calculated for each submodel. A Bayes factor of more than 3 for a model comparison implies that the first model is more substantially supported by the data (Kass and Raftery 1995). Since ABC model selection may be vulnerable to bias (van der Vaart et al. 2015), we scrutinized submodel performance by cross-validation of the model selection (Supplement S4).
The prior and posterior distributions of the summary statistics and parameters were compared as part of the posterior predictive inspection. This allowed us to evaluate how similar the patterns produced by the IBM are to the patterns in the field, estimate parameters and derive which processes are important to reproduce these patterns.

We derived data from the same dataset at the three different steps in our analysis. Generally, using the same dataset for parameterization and model evaluation risks overfitting and overconfidence in the focal model (termed ‘adaptive overfitting’ in machine learning; Roelofs et al. 2019). Therefore independent datasets should ideally be used as input and evaluation of the model. We argue that the use of complementary or auxiliary data in the different components of the analysis minimizes the risk of overfitting for the following reasons: 1) As the microhabitat model is adjusted for spatial autocorrelation with the use of INLA, the data used here are independent of the clustering of the nests causing the model predictions to fundamentally represent the effect of environment on nest presence. 2) Initialization of the IBM is carried out using auxiliary data or probability distributions from the field data, to have comparable boundary conditions for the simulations. Parameters in the IBM are not derived from the field data and are implemented using wide priors. 3) The ABC compares simulations with field data based on larger-scale emergent patterns of the point pattern: clustering and network metrics.
Results

Field study

A total of 432 individual digger wasps were tagged and 561 nests were marked with flags. Test-digging holes (Nielsen 1945) and tagged individuals that were not observed at a nest were excluded. Of those 432 wasps, 330 had one nest, 78 had two, 21 had three and three wasps had 4 nests (Supplement S9). 150 nests (36.5%) were parasitized by the kleptoparasitic fly *Senotainia albifrons* (Supplement S1). Data are deposited in the Dryad Digital Repository: http://doi.org/10.5061/dryad.g79cnp5q8 (Batsleer et al., 2021).

Microhabitat suitability model

The microhabitat suitability model has a high predictive performance: AUC ± SD of the cross-validated final model was 96.0 ± 1.3%. The sensitivity ± SD (true positive rate) was 73.9 ± 4.1%, the specificity ± SD (true negative rate) 96.7 ± 2.1% and balanced accuracy ± SD (overall true rate) 85.3 ± 1.9%. The model therefore performed better in predicting nest absences than nest presences. Nevertheless, predictive performance was overall high. High NDVI-values decreased and high insolation increased nesting suitability. So, sunny sites with a low vegetation cover have a higher probability of containing nests (Supplement S6). For every run, zero was excluded from the 95% credibility intervals of the effect sizes, indicating that the signs of the effect sizes were clearly determined (this is a Bayesian approach to evaluating statistical significance at a specified level). The predictions (ranging from 0 to 1) for nest suitability in the whole study plot are shown in Fig. 2a.
Fig. 2b shows the spatial random components (spatial field or ‘residuals’ which INLA corrects for in the analysis: Supplement S6) that had clear hot and cold areas, indicating higher and lower clustering respectively than expected based on NDVI and insolation. Such cold and hot spots indicate unmeasured variables that vary in space or other underlying mechanisms that cannot be attributed to the environment, such as behavior.

Spatial point pattern and network analyses

Clustering of nests was present up to 10m (blue dots in Fig. 3). 60% of consecutive nests were made in the same aggregate (internal loops, blue line Fig. 3). The network was not densely connected internally (transitivity, blue line Fig. 3); had a low ratio of possible links present (density, blue line Fig. 3); and had a low level of reciprocal connections (reciprocity, blue line Fig. 3). The distances among consecutive nests had a median of 4.30m, a mean of 11.01m and a maximum of 81.53m (see histogram in Supplement S9).

Inverse modelling with IBM and ABC

The prior predictive check made it possible to restrict two parameters’ prior ranges for the actual simulations (Supplement S3). Data from simulations are deposited in the Dryad Digital Repository: http://doi.org/10.5061/dryad.g79cnp5q8 (Batsleer et al., 2021).

The submodel with fixed (inter-individual) strategies was substantially better supported than the uniform strategy. ABC analysis showed to a lesser extent the superior performance for
the flexible strategy (intra-individual) compared to the uniform one. The fixed and flexible (inter- and intra-individual) submodels perform equally well based on the cross-validation of model selection (Supplement S4) and Bayes’ factors (Table 1). Similar results were obtained with the 10,000 (1%), 500 (0.05%) and 100 (0.01%) best simulations (Supplement S5), indicating that the ABC-analysis converged for the number of simulations run.

Both Ripley’s K values and network metrics calculated from the selected simulations matched the field data well (Fig. 3), despite Ripley’s K having a large range at small distances.

**Strengths** for conspecific attraction and local site fidelity were on average 0.739 (median=0.760, Q₁=0.617, Q₃=0.895) and 0.674 (median=0.716, Q₁=0.535, Q₃=0.853; both skewed distributions to 1; fig. 4a), which means that these behavioral mechanisms are strongly present in the population. The strength for environment was on average 0.209 (median=0.174, Q₁=0.079, Q₃=0.305; distribution skewed toward 0; fig 4a), which means the environmental cue is less strongly present: on average 3.5 and 3.2 times weaker than conspecific attraction and local site fidelity, respectively.

The estimated range of the conspecific attraction was on average 2.29m (median=1.98m, Q₁=1.29m, Q₃=3.05m; fig. 4a). The two parameters for the sigmoid response function for conspecific attraction were both low (mindens<sub>ca</sub>: mean=-8.13, median=-8.29, Q₁=-9.25, Q₃=-7.27; σ<sub>ca</sub>: mean=3.16, median=1.74, Q₁=0.69, Q₃=4.28; see Supplement 2 for detailed information about parameter ranges), indicating that the response function had an intercept very close to zero and a steep slope (fig. 4b): the probability of nest site selection becomes large at low densities of conspecific nests. The parameter σ<sub>lsf</sub> for local site fidelity tended towards a narrow
Gaussian distribution, with a scale up to 10m (fig. 4b). 76.5% of the accepted simulations included the parameter \textit{beh-excl}, indicating that the two behavioral mechanisms are mutually exclusive.

**Discussion**

We used an inverse modelling approach to study the processes underlying the aggregative nest pattern formation in the digger wasp \textit{Bembix rostrata}. The observed patterns in nature were best predicted by simultaneously considering effects of the environment, conspecific attraction and local site fidelity. We found that nest pattern formation cannot be decomposed into a stepwise process of environmental filtering and behavioral effects. Rather, it represents a complex system with varying nest choice strategies that rely on the simultaneous integration of environmental and behavioral mechanisms with differing strengths. The spatial patterns of nesting are primarily explained by models with individual differences in behavior, including that an individual uses either personal information or inadvertent social information.

Conspecific attraction is widespread in digger wasps (Evans and O’Neill 2007). Individuals can be attracted to conspecifics as their presence provides an honest cue for habitat suitability. In such situations, the use of social information may have adaptive payoffs by reducing the investment of time and energy in sampling of the environment (Dall et al. 2005). Conspecific attraction strongly affects nest site selection in the studied population (Buxton et al. 2020), while environmental cues appear to have a weaker effect. While the use of social information is best known in vertebrates and social insects, non-social insects also possess individual and even social learning abilities that eventually contribute to higher fitness (Coolen et al. 2005). \textit{Bembix rostrata}
is known to perform test-digging behavior, in which individuals seem to sample the environment by digging shallow burrows in the sand across the nesting area before starting to dig an actual nest (Nielsen 1945). The use of social cues for habitat suitability is therefore likely adaptive as it reduces the time and energy spent on this behavior. The large contribution of social attraction in predicting nest patterning likely explains the high levels of philopatry in *B. rostrata* (Nielsen 1945; Larsson 1986; Blösch 2000), i.e. their tendency to remain in the same nesting area for several consecutive generations.

The selfish herd hypothesis states that individuals within a population attempt to reduce their predation or parasitism rate by putting other conspecifics between themselves and predators or parasites (Mooring and Hart 1992). This theory has been invoked to explain the aggregation of *B. rostrata* (Larsson 1986) and another closely related digger wasp, *Crabro cribrellifer* (Wcislo 1984). These studies, along with our data (Supplement S1), show that the incidence of parasitism per nest decreases with nest density. In our system, the presence of *Senotainia albifrons* brood parasites is anticipated to convey information to the wasps and to induce conspecific attraction. However, how much the actual presence of these brood parasites contributes to the overall conspecific attraction is uncertain. Since we lack more information on the dynamics of these parasites, such mechanisms were not directly incorporated in the model, but were instead included as part of the primary process of attraction. Several nests nevertheless occurred at quite low densities, where the parasitism rate, and especially its variation, is higher (Supplement S1 Fig. S1.2). Potential reasons for this more risky behavior could include, non-exhaustively, spatial bet hedging (Philippi and Seger 1989), imperfect information of parasitism risk (Koops and Abrahams 1998) or avoidance of perceived intraspecific competition (Polidori et
al. 2008). Regarding the latter mechanism, conspecific kleptoparasitism (wasps that steal prey from neighboring individuals) has been observed in five other Bembix-species, but not (yet) in B. rostrata (Evans and O’Neill 2007).

The most likely nest selection strategy identified by our model is one with individually consistent but mutually exclusive behaviors (parameter beh-excl): when local site fidelity is used, conspecific attraction is not used simultaneously for nest site selection. Consistent individual variation in movement behavior, with individuals relying on either personal or social information, can be responsible for the emergence of ecological patterns at larger spatial scales (Spiegel et al. 2017). Such heterogeneity in behavior due to individual specialization, may be especially relevant in populations experiencing high levels of intraspecific competition (Araújo et al. 2011). The second most probable, but slightly less supported model, considered individual behavioral flexibility during an individual’s lifetime. Shifts in individual behavior have been found across taxa in, for example, foraging in heterogeneous environments (Newlands et al. 2004; Webber et al. 2020), mating (Perrill et al. 1982), migration (Eggeman et al. 2016) and seasonal aggregation (Bonar et al. 2020). These shifts arise from plasticity in response to environmental and demographic changes (e.g. density, competition, predation). We modeled behavioral shifts as a stochastic process since any information on the potential conditionality of such shifts was lacking. Explicitly considering thresholds that underlie movement changes is nevertheless important to explain larger-scale patterns (Morales and Ellner 2002; Newlands et al. 2004; Goossens et al. 2020). In our study, a flexible strategy is not clearly distinguishable from a consistent one in explaining the spatial nest pattern. Therefore, our results show there are clear behavioral differences between individuals, but it is not conclusive if these behaviors vary over time.
Moreover, in some species (e.g. caribou, *Rangifer tarandus*), both consistent and flexible strategies can be present in a population depending on the socio-spatial position of individuals (Bonar et al. 2020). The relative importance of different strategies can change within and among populations, seasons or with different levels of parasitic pressure (Spiegel et al. 2017). More fine-scaled, individual studies or experiments are therefore required to explore the importance of consistent and flexible strategies.

The joint contributions of the environment and internal dynamics to predict spatial nest patterns in our system suggest synergism among multiple processes underlying spatial pattern formation. These mechanisms are often studied in isolation, as behavioral and landscape ecologists traditionally work on very different scales and units of research (Lima and Zollner 1996). These research fields have accordingly developed their specific analytical methods that can be regarded as separate approaches to explain spatial pattern formation of nests in a stepwise manner (Melles et al. 2009). Such combined approaches have been applied in insect (Polidori et al. 2008; Asís et al. 2014) and bird-oriented research (Brown and Brown 2000; Perry and Andersen 2003; Melles et al. 2009), most often from a conservation perspective (Etterson 2003; Ward and Schlossberg 2004; Bayard and Elphick 2010; Swift et al. 2017). Spatial pattern analysis of point data from homogeneous landscapes allows inference of putative feedbacks that eventually lead to spatial self-organization (Rietkerk and van de Koppel 2008). The assumption of environmental homogeneity is not always valid. We likewise first decoupled environment from behavior by building the microhabitat model with INLA, but integrated both types of mechanisms again in the IBM. IBMs are ideal for bridging and integrating these seemingly separate processes at different scales and enable quantifying their relative importance and synergism. This inverse
approach has much to offer for understanding behavioral mechanisms underlying spatial organization, including identifying its own methodological limits. Spatial organization processes are inherently stochastic and replications are hard to obtain as the strength or shape of processes are likely to change with space and time (Wagner and Fortin 2005). The general approach we illustrate here could thus be applicable in teasing apart coexisting, context-dependent processes, which are pervasive in ecological systems.

Acknowledgements

FB is supported by Research Foundation – Flanders (FWO). DB is supported by Research Foundation – Flanders (FWO) project INVADED G018017N.

The computational resources (Stevin Supercomputer Infrastructure) and services used in this work were provided by the VSC (Flemish Supercomputer Center), funded by Ghent University, FWO and the Flemish Government – department EWI.

We thank Richard Sibly and his research group at the University of Reading, for advice and workshop on the ABC analysis, and Johan Lamaire (ANB) for help and permission to access the field study area in nature reserve ‘De Westhoek’. We are grateful for logistical help during fieldwork from Wouter Van Gompel, Daan Dekeukeleire, Steven Goossens, Nadine De Schrijver and Marc Batsleer. We thank Sam Provoost for feedback on this study in an early stage. We thank Martijn L. Vandegehuchte for comments on an earlier version of the manuscript and Frederik Mortier for feedback on Python code.
We thank two anonymous reviewers and associate editor Dr. Benjamin M. Bolker for the detailed and in-depth comments which improved the manuscript greatly.

**Statement of Authorship**

All three authors conceived the ideas and designed methodology; FB collected and analyzed the data, developed the model and led the writing of the manuscript. All authors contributed critically to the drafts and gave final approval for publication.

**Data and Code Accessibility**

Data deposited in the Dryad Digital Repository: [http://doi.org/10.5061/dryad.g79cnp5q8](http://doi.org/10.5061/dryad.g79cnp5q8) and code deposited in Zenodo: [http://doi.org/10.5281/zenodo.5212680](http://doi.org/10.5281/zenodo.5212680)

**Literature Cited**


References Cited Only in the Online Enhancements


Table 1: Bayes’ factors and proportions of accepted models for model selection with ABC-analysis.

<table>
<thead>
<tr>
<th></th>
<th>Random</th>
<th>Population</th>
<th>Inter-individual</th>
<th>Intra-individual</th>
<th>% accepted simulation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random</td>
<td>-</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0</td>
</tr>
<tr>
<td>Population</td>
<td>∞</td>
<td>1.00</td>
<td>0.31</td>
<td>0.36</td>
<td>14.3%</td>
</tr>
<tr>
<td>Inter-individual</td>
<td>∞</td>
<td>3.20</td>
<td>1.00</td>
<td>1.14</td>
<td>45.7%</td>
</tr>
<tr>
<td>Intra-individual</td>
<td>∞</td>
<td>2.80</td>
<td>0.88</td>
<td>1.00</td>
<td>40.0%</td>
</tr>
</tbody>
</table>

The ABC-analysis retained the 1000 best simulations of 1,000,000 (0.1%). The submodels represent at which level the mechanisms can vary: population, inter-individual, intra-individual. 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 $M_0$. 
Figure Legends

Figure 1: Conceptual figure of the workflow of the analyses. (1) The relation between the nest positions and the environment is investigated by building a microhabitat suitability model with INLA. Because the clustering is much higher than expected based solely on the microhabitat model, (2) the environment and behavioral mechanisms (i.e. local site fidelity and conspecific attraction) are simultaneously modelled with an Individual-Based Model (IBM). The simulations differ in strength of the mechanisms. This is implemented through strength parameters that represent the probability of the mechanisms being present. The presence of the mechanisms can vary on three different levels: the population, inter-individual and intra-individual. (3) The simulations from the IBM are compared with the spatial pattern of the field data using Approximate Bayesian Computation (ABC) to infer which submodels best approach the field data.

Figure 2: a) Predictions for microhabitat suitability (0-1) on the field study plot based on vegetation (NDVI) and sun irradiance (insolation), mean of all 10 iterations. b) Posterior mean values of the spatial field, from one of the iterations, others were similar. The spatial field shows where the spatial autocorrelation, corrected for with INLA, deviated from zero indicating higher clustering (hot spots; dark orange) or lower clustering (cold spots; dark purple) than expected based on NDVI and insolation. This indicates that other mechanisms must be involved in the clustering of the nests. Black dots are nest locations. Pixel size is 50×50cm² on ground. Cartesian coordinate reference system used is Belgian Lambert 72, epsg:31370.

Figure 3: Violin plots of the distribution of prior summary statistics of all 1,000,000 simulations (grey), the distribution of the posterior summary statistics of the 1000 (0.1%) best models (green), and field data (blue dots or lines). Yellow distributions are from the null model (submodel random). Summary statistics are of two types: spatial clustering (Ripley’s K, relative change compared to complete spatial randomness CSR) and network metrics (all and internal loops, transitivity, density and reciprocity).
Figure 4: a) priors (transparent red) and posteriors (dark grey) of the strengths (of the mechanisms: ENV environment; LSF local site fidelity; CA conspecific attraction) and parameters of the behavioral mechanisms. Parameter beh-excl defines if conspecific attraction and local site fidelity are mutually exclusive (1) or not (0). b) Effect of the posteriors on the response functions, defined by the bottom 4 parameters, plotted for median (50% quantile), 20% and 80% quantiles of the corresponding parameter. Other parameters than the focal are held constant at the median of the posterior distributions. Local site fidelity (first graph in b) is implemented with a Gaussian curve with the center at the previous nest. Conspecific attraction is implemented with a sigmoid curve (3 graphs on the right of b), with the dependent variable the density of nests (number of nests within range_{ca}). See main text and Supplement S2 for further details on parameter definitions and ranges.
1) Microhabitat model

with NDVI and Insolation

Environment

Local site fidelity

Conspecific attraction

2) Individual-Based Model

Strengths

Probability of mechanism present

Submodels

Level of variability of mechanisms

3) ABC4IBM

Spatial clustering & network between aggregates

Field data
Figure 3

This is the author's accepted manuscript without copyediting, formatting, or final corrections. It will be published in its final form in an upcoming issue of The American Naturalist, published by The University of Chicago Press. Include the DOI when citing or quoting: https://doi.org/10.1086/717226
Copyright 2021 The University of Chicago Press.
Supplementary material - Behavioral strategies and the spatial pattern formation of nesting

Femke Batsleer¹*(ORCID 0000-0001-5893-5966)
Dirk Maes² (ORCID 0000-0002-7947-3788)
Dries Bonte¹ (ORCID 0000-0002-3320-7505)

¹ Ghent University, Department Biology, Terrestrial Ecology Unit, Karel Lodewijk Ledeganckstraat 35, B-9000 Gent, Belgium
² Research Institute for Nature and Forest (INBO), Herman Teirlinckgebouw, Havenlaan 88 bus 73, B-1000 Brussel, Belgium

*Corresponding author: F. Batsleer, Ghent University, Department of Biology, Terrestrial Ecology Unit. Karel Lodewijk Ledeganckstraat 35, B-9000 Gent, Belgium. E-mail: femke.batsleer@ugent.be

The American Naturalist
Supplement 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). 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.

The density of parasitized nests increased with nest density (Fig S1.1; β = 0.27, R² = 0.80, F(1, 148) = 581.3, p = < 2·10⁻¹⁶). 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).
Supplement 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)
Behavior and nest spatial pattern

- 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 50x50cm²/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 Supplement 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 50×50cm².

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 47×97m²).
- 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/m², in our study we found a maximum of 12 nests/m². We chose the prior of $\alpha_{lsf}$ very wide such that the 0.5-probability occurs at 0.2 to 100 counted nests (depending on $mindens_{ca}$).
<table>
<thead>
<tr>
<th>Mechanism</th>
<th>Function</th>
<th>Parameters</th>
<th>Values of initials runs for the prior predictive check</th>
<th>Values of actual runs</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Local site fidelity</strong></td>
<td>Probability ~ Gaussian function with x the distance to the previous nest (μ=0)</td>
<td>$a_{lsf}$: defines the width or range of the Gaussian function</td>
<td>[0.1, 10]</td>
<td>[0.1, 10]</td>
</tr>
<tr>
<td></td>
<td>$P = \frac{1}{\sigma_{lsf} \sqrt{2\pi}} e^{-\frac{x^2}{2\sigma_{lsf}^2}}$</td>
<td>quantity: distance (m)</td>
<td>↓</td>
<td>↓</td>
</tr>
<tr>
<td></td>
<td>$\sigma_{lsf}$: defines the width or range of the Gaussian function</td>
<td></td>
<td>95% of $P$ is within radius of ±20 cm → ±20 m from the previous nest</td>
<td>95% of $P$ is within radius of ±20 cm → ±20 m from the previous nest</td>
</tr>
<tr>
<td></td>
<td>$r_{range}$: defines the circle at which nests are taken into account for the conspecific attraction at the position under consideration</td>
<td>quantity: distance (m)</td>
<td>[0.2, 20]</td>
<td>[0.2, 7]</td>
</tr>
<tr>
<td></td>
<td>$P = \frac{1}{1 + e^{-a}}$</td>
<td></td>
<td>Nests within Ø of 40 cm → 40m are taken into account for conspecific attraction</td>
<td>Nests within Ø of 40 cm → 14m are taken into account for conspecific attraction</td>
</tr>
<tr>
<td></td>
<td>$a = mindens_{ca}$: defines the probability at density=0 for which:</td>
<td>$[-10, 10]$</td>
<td>$P(0) \in [4.54 \times 10^{-5}, 0.999]$</td>
<td>$P(0) \in [4.54 \times 10^{-5}, 0.05]$</td>
</tr>
<tr>
<td></td>
<td>$P(0) = \frac{1}{1 + e^{-a}}$</td>
<td></td>
<td>↓</td>
<td>↓</td>
</tr>
<tr>
<td></td>
<td>$P(0.5) \in (-100, 100)$</td>
<td></td>
<td>Negative when $P(0) &gt; P(0.5)$</td>
<td>Negative when $P(0) &gt; P(0.5)$</td>
</tr>
<tr>
<td></td>
<td>$b = a_{ca}$: defines the width or ‘stretch’ of the sigmoid curve, for which:</td>
<td>$[0.1, 15]$</td>
<td>$P(0.5) \in (-100, 100)$ when $a=3$</td>
<td>$P(0.5) \in [0.67, 100]$ when $a=10$</td>
</tr>
<tr>
<td></td>
<td>$P(0.5) = \frac{-a}{b}$</td>
<td></td>
<td>↓</td>
<td>↓</td>
</tr>
<tr>
<td><strong>Conspecific attraction</strong></td>
<td>a location with other nests present attract new nests</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Probability ~ sigmoid function for which $+\infty \rightarrow +1$, where x is the number of nests divided by surface of circle, the density within range$_{ca}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$P = \frac{1}{1 + e^{-(a+bx)}}$</td>
<td>quantity: number of nests (#)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$a_{ca}$: defines the probability at density=0 for which:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$P(0) = \frac{1}{1 + e^{-a}}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$P(0) \in [4.54 \times 10^{-5}, 0.999]$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$P(0.5) \in (-100, 100)$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The two behavioral mechanisms are mutually exclusive</td>
<td>Not applicable</td>
<td>beh-excl: 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.</td>
<td>True/False</td>
<td>True/False</td>
</tr>
</tbody>
</table>

The two behavioral mechanisms are mutually exclusive
Supplement 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$_{ca}$ and mindens$_{ca}$, because they covary strongly with strength$_{env}$ and strength$_{lsf}$. 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. σ$_{ca}$ 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.

σ$_{ca}$ is also correlated with range$_{ca}$ (fig S3.5). First, this is expected, as the larger the range, the more nests are counted in that range, resulting in a wider σ$_{ca}$ for the response function. Second, as range$_{ca}$ is already narrowed down, narrowing σ$_{ca}$ would be redundant.
Figure S3.1: pairs plots and Pearson’s correlations between the parameters for the prior predictive check.
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).
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).
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 $\text{range}_{\text{ca}}$ and $\sigma_{\text{ca}}$ 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).
Supplement 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 function performs a leave-one-out cross validation for model selection with ABC. We ran it for 50 simulations for each model (nval = 50). Details for the different submodels (ODD-protocol) can be found in Supplement 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 Supplement 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).

<table>
<thead>
<tr>
<th>True model</th>
<th>Classification</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Random</td>
</tr>
<tr>
<td>Random</td>
<td>0.8407</td>
</tr>
<tr>
<td>Population</td>
<td>0.0943</td>
</tr>
<tr>
<td>Inter-individual</td>
<td>0.0057</td>
</tr>
<tr>
<td>Intra-individual</td>
<td>0.0335</td>
</tr>
</tbody>
</table>

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

Tolerance rate = 0.001 (1000 simulations)
Supplement 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.

<table>
<thead>
<tr>
<th></th>
<th>Random</th>
<th>Population</th>
<th>Inter-individual</th>
<th>Intra-individual</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random</td>
<td>-</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Population</td>
<td>-</td>
<td>1.00</td>
<td>0.31</td>
<td>0.32</td>
</tr>
<tr>
<td>Inter-individual</td>
<td>-</td>
<td>3.25</td>
<td>1.00</td>
<td>1.05</td>
</tr>
<tr>
<td>Intra-individual</td>
<td>-</td>
<td>3.09</td>
<td>0.95</td>
<td>1.00</td>
</tr>
<tr>
<td>% accepted simulations</td>
<td>0</td>
<td>13.63%</td>
<td>44.26%</td>
<td>42.11%</td>
</tr>
</tbody>
</table>

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.

<table>
<thead>
<tr>
<th></th>
<th>Random</th>
<th>Population</th>
<th>Inter-individual</th>
<th>Intra-individual</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random</td>
<td>-</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Population</td>
<td>-</td>
<td>1.00</td>
<td>0.31</td>
<td>0.36</td>
</tr>
<tr>
<td>Inter-individual</td>
<td>-</td>
<td>3.27</td>
<td>1.00</td>
<td>1.18</td>
</tr>
<tr>
<td>Intra-individual</td>
<td>-</td>
<td>2.77</td>
<td>0.85</td>
<td>1.00</td>
</tr>
<tr>
<td>% accepted simulations</td>
<td>0</td>
<td>14.2%</td>
<td>46.4%</td>
<td>39.4%</td>
</tr>
</tbody>
</table>

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.
## Table

<table>
<thead>
<tr>
<th></th>
<th>Random</th>
<th>Population</th>
<th>Inter-individual</th>
<th>Intra-individual</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random</td>
<td>-</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Population</td>
<td>∞</td>
<td>1.00</td>
<td>0.41</td>
<td>0.54</td>
</tr>
<tr>
<td>Inter-individual</td>
<td>∞</td>
<td>2.42</td>
<td>1.00</td>
<td>1.31</td>
</tr>
<tr>
<td>Intra-individual</td>
<td>∞</td>
<td>1.84</td>
<td>0.76</td>
<td>1.00</td>
</tr>
<tr>
<td>% accepted</td>
<td>0</td>
<td>19%</td>
<td>46%</td>
<td>35%</td>
</tr>
<tr>
<td>simulations</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Supplement 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.
Supplement 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; ± 0.5 in our case). In bold the scale that was chosen.

<table>
<thead>
<tr>
<th>Scale</th>
<th>0.1 m</th>
<th>0.2 m</th>
<th>0.5 m</th>
<th>1 m</th>
<th>2 m</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>mean</td>
<td>sd</td>
<td>mean</td>
<td>sd</td>
<td>mean</td>
</tr>
<tr>
<td>Mean auc</td>
<td>0.9501</td>
<td>0.0093</td>
<td>0.9461</td>
<td>0.0088</td>
<td>0.9459</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>0.6247</td>
<td>0.0421</td>
<td>0.6238</td>
<td>0.0598</td>
<td>0.6642</td>
</tr>
<tr>
<td>Specificity</td>
<td>0.9749</td>
<td>0.0148</td>
<td>0.9703</td>
<td>0.0210</td>
<td>0.9575</td>
</tr>
<tr>
<td>Accuracy</td>
<td>0.7998</td>
<td>0.0190</td>
<td>0.7971</td>
<td>0.0211</td>
<td>0.8109</td>
</tr>
</tbody>
</table>
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.
Supplement 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 was not stable for a cluster number up to 10 (example given in Fig. S8.3). For a cluster number of 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).

![Optimal number of clusters](https://example.com/figure.png)

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 $k = 11$ ($n_{start} = 20$).
Supplement S9: Information about consecutive nests in the field study

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

<table>
<thead>
<tr>
<th>Number of nests</th>
<th>frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>330</td>
</tr>
<tr>
<td>2</td>
<td>78</td>
</tr>
<tr>
<td>3</td>
<td>21</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
</tr>
</tbody>
</table>

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 supplementary material


