

# The Spatial Organisation of Nest Aggregates in a Digger Wasp



**Femke Batsleer**

Student number: 01001520

Supervisor & scientific tutor: Prof. Dr. Dries Bonte

Master's dissertation submitted to obtain the degree of Master of Science in Biology

Academic year: 2016 - 2017

**© Faculty of Sciences – research group Terrestrial Ecology**

All rights reserved. This thesis contains confidential information and confidential research results that are property to the UGent. The contents of this master thesis may under no circumstances be made public, nor complete or partial, without the explicit and preceding permission of the UGent representative, i.e. the supervisor. The thesis may under no circumstances be copied or duplicated in any form, unless permission granted in written form. Any violation of the confidential nature of this thesis may impose irreparable damage to the UGent. In case of a dispute that may arise within the context of this declaration, the Judicial Court of Gent only is competent to be notified.

**Cover picture:** Tagged digger wasp individual 0(range)56 at her burrow © Willem Van Echelpoel

# Contents

<b>1</b>	<b>Introduction</b>	<b>5</b>
<b>2</b>	<b>Objectives</b>	<b>8</b>
<b>3</b>	<b>Material &amp; Methods</b>	<b>9</b>
3.1	Data collection . . . . .	9
3.2	Microhabitat model . . . . .	10
3.2.1	Presence-absence model . . . . .	10
3.2.2	Real nests model . . . . .	12
3.2.3	Parasitised nests model . . . . .	12
3.2.4	Temporal microhabitat model . . . . .	12
3.3	Describing the population and spatial pattern . . . . .	12
3.3.1	CMRC: population estimation . . . . .	12
3.3.2	Point pattern analysis . . . . .	13
3.4	The population as a network . . . . .	14
3.4.1	Distance kernel between two subsequent nests . . . . .	14
3.4.2	Cluster analysis . . . . .	14
3.4.3	Social network . . . . .	14
3.5	The causes of clustering . . . . .	15
3.5.1	Microhabitat availability . . . . .	15
3.5.2	Parasite density . . . . .	16
<b>4</b>	<b>Results</b>	<b>17</b>
4.1	Data collection . . . . .	17
4.2	Microhabitat model . . . . .	17
4.2.1	Presence-absence model . . . . .	17
4.2.2	Real nests model . . . . .	18
4.2.3	Parasitised nests model . . . . .	18
4.2.4	Temporal microhabitat model . . . . .	18
4.3	Population estimation and spatial pattern analysis . . . . .	22
4.3.1	CMRC: population estimation . . . . .	22
4.3.2	Point Pattern Analysis . . . . .	22
4.4	The population as a network . . . . .	25
4.4.1	Distance kernel between two subsequent nests . . . . .	25
4.4.2	Cluster analysis . . . . .	26
4.4.3	Social network . . . . .	27
4.5	The causes of clustering . . . . .	33
4.5.1	Microhabitat model . . . . .	33
4.5.2	Parasite density . . . . .	34
<b>5</b>	<b>Discussion</b>	<b>35</b>
5.1	Data collection . . . . .	35
5.1.1	Tagging . . . . .	35
5.1.2	Real nests . . . . .	35
5.1.3	Anti-parasite behaviour . . . . .	35
5.1.4	Prey . . . . .	35
5.2	Microhabitatmodel . . . . .	35
5.2.1	Presence-absence model . . . . .	36
5.2.2	Real nests . . . . .	36
5.2.3	Parasitised nests . . . . .	36
5.2.4	Temporal model . . . . .	37
5.3	Population estimation and spatial pattern analysis . . . . .	37
5.3.1	CMRC: population estimation . . . . .	37
5.3.2	Point Pattern Analysis . . . . .	37
5.4	The population as a network . . . . .	37
5.4.1	Distance kernel between two subsequent nests . . . . .	38
5.4.2	Cluster analysis . . . . .	38

5.4.3	Social network . . . . .	38
5.5	The causes of clustering and implications . . . . .	39
5.5.1	Microhabitat suitability . . . . .	39
5.5.2	Parasite density . . . . .	40
5.5.3	Mechanism of clustering . . . . .	40
5.5.4	The scale of philopatry . . . . .	41
5.6	Implications for conservation . . . . .	41
5.7	Possible future research . . . . .	42
<b>6</b>	<b>Conclusion</b>	<b>43</b>
<b>7</b>	<b>Summary/samenvatting</b>	<b>44</b>
7.1	English summary . . . . .	44
7.2	Nederlandse samenvatting . . . . .	45
<b>8</b>	<b>Acknowledgements</b>	<b>47</b>
<b>9</b>	<b>Bibliography</b>	<b>48</b>
<b>10</b>	<b>Appendices</b>	<b>54</b>

# 1 Introduction

The question of whether **biological systems are regulated bottom-up or top-down**, has for decades been a major topic of discussion in ecology (Hairston *et al.*, 1960; Bowlby & Roff, 1986; Mittelbach *et al.*, 1988; Power, 1992; Hunter & Price, 1992; Banse, 2007). The question of whether the primary control in population, community and ecosystem structure is resource limitation or predation, has been refined and has become more nuanced. The discussion has shifted from a simple dichotomy towards a more complex, integrated view. Some issues are frequently emphasised: (1) **heterogeneity** at any trophic level (Hunter & Price, 1992); (2) the relative importance of top-down and bottom-up effects depending on the biological system or changing environment (Karr *et al.*, 1992; Hunter & Price, 1992; Dunson & Travist, 1991; Burkepile & Hay, 2006). An example of this is Meserve *et al.* (2003) showing that a semi-arid thorn scrub system undergoes shifting bottom-up or top-down control, mainly driven by El Niño events; (3) a spatial-explicit framework within and between systems (Gripenberg & Roslin, 2007); (4) the **simultaneous influence** and relative role of both bottom-up and top-down mechanisms (Hunter *et al.*, 1997; Walker & Jones, 2001; Horswill *et al.*, 2016).

Understanding how **heterogeneous spatial patterns** are formed in nature is important to understand fundamental ecological processes (Rietkerk & van de Koppel, 2008) and the diversity and stability of ecosystems (Solé & Bascompte, 2006; Rietkerk *et al.*, 2004). Local interactions can induce large-scale spatial patterns by **spatial self-organisation**. The mechanism mostly put forward is scale-dependent feedback (Rietkerk & van de Koppel, 2008). Turing (1953) first described this mechanism in a chemical system in which an activator acts locally and an inhibitor (which diffuses faster) acts on a larger scale. This **activator-inhibitor mechanism** is well known in chemical and physical systems, and is also used in ecology to explain regular patterns (Rietkerk & van de Koppel, 2008). Examples abound: vegetation patterns in arid vegetation (Couteron & Lejeune, 2001), see figure 1 formation of coral reefs (Mistr & Bercovici, 2003), sedge tussocks in marshlands (van de Koppel & Crain, 2006)... These are all driven by a facilitating or positive feedback at short distances (water or substrate availability, retention of resources etc.) and a negative feedback on larger scales (water and wind flows, erosion etc.). The negative feedback is the most essential driving force (Rietkerk & van de Koppel, 2008).

The activator-inhibitor mechanism is formulated for growth-based processes and defines one or more interactions that are active on different spatial scales. What it cannot explain is animal aggregates or group living steered by animal behaviour (Liu *et al.*, 2016). Behavioural ecology describes several mechanisms for group living (Davies *et al.*, 2012). For the individuals in a group, predation risk can be reduced by a dilution effect or the selfish herd mechanism, a confusion effect, the collective detection or the collective defence against predators. Group living can also improve foraging because of better food finding or better prey-capture. Understanding the informational flows, through network analysis, is important in understanding the self-organisation of collective behaviour (Sumpter, 2006).

To explain movement driven spatial self-organisation, or group formation driven by behaviour, another mechanism known in physical and chemical systems has only recently been adopted in ecology: the **phase separation principle** (Liu *et al.*, 2013, 2016). The original principle was formulated by Cahn & Hilliard (1958) to explain self-organisation in metal alloys and formation of minerals. Fundamental to the process is density-dependent movement. At low and high densities, species move fast or disperse, while at an intermediate density, they aggregate. The pattern that arises can be similar to the one arising from the activator-inhibitor principle. Examples are mussel clump patterns (van de Koppel *et al.*, 2008; Liu *et al.*, 2013), see figure 1, and aggregations of ants (Theraulaz *et al.*, 2002; Liu *et al.*, 2016). This phase separation principle, based on density-dependent movement, is also able to explain non-stationary patterns, like mobile aggregations of elk (*Cervus elaphus*) (Liu *et al.*, 2016). The activator-inhibitor and phase separation principles are summarised in figure 1, which is adopted from Liu *et al.* (2016).

Spatial patterns of nest aggregates form an interesting subject to investigate self-organisation. These patterns of nestbuilding can be steered by the **simultaneous influence of top-down and bottom-up regulators**. The abiotic nest site preferences are determined by different environmental conditions and resources, and can be seen as different dimensions in an n-dimensional hypervolume in Hutchinson's niche concept (Hutchinson, 1957). These determine a bottom-up template for nest choice. Interactions with other organisms can act as top-down regulators, limiting this abiotic, fundamental niche: biotic processes are important to understand the realised niche of the nestbuilding of a species.

In many cases, clustering of nests can be determined almost solely by **abiotic factors**. The solitary bee *Dieunomia triangulifera*, for instance, has very specific nest site preferences on very small scales and

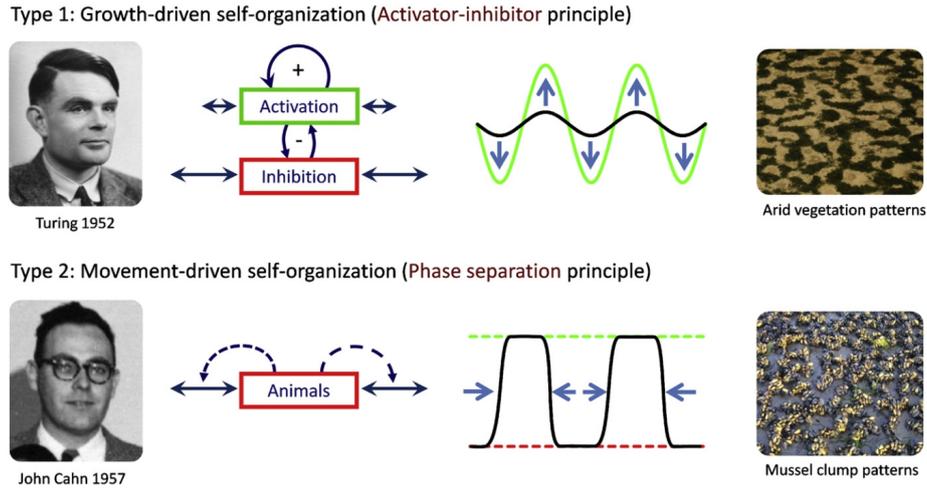


Figure 1: Summary of the activator-inhibitor principle developed by Turing (type 1) and the phase separation principle of Cahn-Hilliard (type 2), both resulting in self-organised spatial pattern formation in biological systems. Adopted from Liu *et al.* (2016)

does not show any preference for places where conspecifics have already made nests (Wuellner, 1999). The clustering pattern in this species is thus mainly determined by the limited availability of the nest site habitat. On larger scales, not only the patch quality, but also the landscape context becomes an important predictor for the presence of species (Mazerolle & Villard, 1999).

The interaction with other organisms, **biotic factors**, can influence spatial patterns of nests as well. American golden plovers (*Pluvialis dominica*) show conspicuous anti-predator behaviour: they chase away predators that come near their nests. The individual predation risk of a nest depends on the density of active plovers and the distance from a nest of an active plover (Giroux *et al.*, 2016). Predation pressure influences nest site selection (Forstmeier & Weiss, 2004; Eggers *et al.*, 2006), which could alter the spatial pattern of these nests. Another example of the importance of a biotic interaction with nest structure is the risk of brood parasitism by cuckoos in a reed warbler which is determined by the spatial characteristics of nests and of the surrounding landscape (Moskat & Honza, 2000).

Clear-cut spatial patterns, which arise from the phase separation principle, can be attenuated by other prevailing, biological mechanisms. For instance, organisms can show behavioural responses and flexibility in uncertain environments (Clark & Mangel, 1984; Inglis *et al.*, 2001; Sih, 2013). The mechanism behind the phase separation principle is based on movement behaviour. Thus, any behavioural response which alters habitat use and movement can attenuate clear spatial patterns arising from a bigger, general process. A strategy to cope with uncertain environments is **spatial bet hedging** (Cohen, 1966), e.g. individuals that choose a variety of nest sites which have different performances or fitness-outcomes under different environmental conditions. It can be likened to betting on different horses in a horse race instead of risking all your money on one supposed champion. This spatial bet hedging could dilute a pattern of social nest aggregates in space and time.

The distribution pattern and establishment of populations of a species are dependent on the habitat selection of that species (Matthiopoulos *et al.*, 2015). There are three important factors that determine the observed use of a habitat: habitat quality, biotic interactions and movement characteristics. **Habitat quality** should be assessed with demographic variables whenever possible (Johnson, 2007), because habitat quality is not always directly linked to the habitat use or density of a species in that habitat (Van Horne, 1983). Densities can be kept high in sink populations or ecological traps. In the latter, maladaptive habitat selection plays an important role (Kristan, 2003). Social cues, and more specifically conspecific attraction, are important information criteria for habitat selection (Hahn & Silverman, 2006; Ward & Schlossberg, 2004; Muller *et al.*, 1997). The mismatch between social cues and habitat quality can contribute to the formation of ecological traps, but this is very difficult to detect (Robertson & Hutto, 2006). Other **biotic interactions**, e.g. the presence of predators, have an influence on the observed habitat choice too (Forstmeier & Weiss, 2004; Eggers *et al.*, 2006). **Movement** characteristics are im-

portant for the functional landscape connectivity (Bélisle, 2005) and can compromise the establishment of a species if this functional connectivity is too low.

Conservation strategies should address all these factors, including conspecific attraction (Campomizzi *et al.*, 2008). But conservation is often unproportionally focused on **connectivity**, and it is precisely this factor that comes with a lot of uncertainties in the estimation of its effect (Hodgson *et al.*, 2009) and that, if thoroughly studied, can be the second factor – after habitat quality – which determines spatial distribution of a species at suitable patches (Bonte *et al.*, 2003). **Increasing habitat area and habitat quality** has more potential for augmenting populations of conservation interest, and will most of the time increase connectivity simultaneously (Hodgson *et al.*, 2009). Therefore it is very important to understand the habitat and niche characteristics of a species of conservation interest.

The gregarious nesting of digger wasps forms an interesting system with which to study the relative effects of top-down and bottom-up regulators of spatial patterns and the relative importance of habitat quality and connectivity. The **digger wasp *Bembix rostrata*** (Linnaeus, 1758; Hymenoptera, Crabronidae, Bembicinae) is a specialised, highly philopatric, gregariously nesting digger wasp found in sandy regions of Europe. This species nests in sand dunes with a low vegetation cover exposed to the sun (Larsson, 1986; Klein & Lefebvre, 2004). The nesting season of *B. rostrata* extends from mid-June to late September (Peeters, 2008). The female constructs one burrow at a time and progressively provisions (Field, 2005) its single larva with flies (Nielsen, 1945), during a nest cycle that takes approximately 12 days in total (Tengö *et al.*, 1996). The adults are nectar-feeders. Several brood-parasitic fly species lay their eggs (or their maggots, in case of the ovoviviparous Sarcophagidae) on the provisioned prey of the female *B. rostrata* (Peeters, 2008; Nielsen, 1945; Larsson, 1986).

*B. rostrata* is **highly endangered** in Belgium, only occurring in its larger sandy regions (dunes at the west coast, inland dunes at Geel-Bel in Antwerp and spoil tips (Dutch: ‘terril’) at Harchies in Wallonia). Records from the Netherlands dating back to 1980 show a severe decline in its distribution (Klein & Lefebvre, 2004). This decline is caused by the loss of suitable sandy habitat and the encroachment of scrubs and grasses observed in most dune systems, which have caused similar declines in other dune arthropod species (van Swaay & Plate, 2009). To counteract the fact that dunes get overgrown, the lack of natural disturbance of aeolic dynamics in dune ecosystems is replaced with the traditional management techniques with large herbivores. These large herbivores have a high negative impact on nest densities on the specialised digger wasp *B. rostrata* (Bonte, 2005). Despite the strong flying capacities of *B. rostrata*, the species is not known to (re)colonize apparently suitable habitat on a large scale. The species is considered highly philopatric: it displays a high site fidelity. Conservation planning which focuses on habitat area and quality instead of connectivity might be especially important in this philopatric species: increasing area and quality are likely the best way to increase and conserve populations of *B. rostrata*.

**This master project focuses on the digger wasp *Bembix rostrata*, aiming to understand the processes involved in the spatial pattern formation of nest aggregates.** Both bottom-up and top-down regulators are described and the mechanism behind the spatial clustering is discussed. Wasps were individually followed up, their nests located and parasite infections observed. The pattern formation will be analysed by means of microhabitat modelling of presence-absence data, point pattern analysis and social network analysis. Clustering in both space and time are investigated. The microhabitat preferences are characterised through remote sensing imagery with a drone. This technique makes it possible to map vegetation with a very high resolution and to measure habitat characteristics that could never be obtained through on-ground field data. It also has advantages for future research: the potential habitat can be mapped on larger areas and could reduce field efforts of inventories. Using drones in ecology and conservation is a hot topic and has the potential to expand ecological knowledge and insights (Koh & Wich, 2012; Marvin *et al.*, 2016).

## 2 Objectives

The main objective of this study is to understand the spatial pattern and formation of nest aggregates of the digger wasp *Bembix rostrata*. The intent is to elucidate the driving factors, bottom-up and/or top-down regulators, of nest clustering. The hypothesis is that microhabitat availability acts as a bottom-up template for clustering. Simultaneously, a top-down regulator has an effect on the clustering: a selfish herd mechanism, driven by parasites, benefits high density nests because they have a lower individual parasite density (Larsson, 1986; Weislo, 1984). The hypothesised mechanism for this clustering is the phase separation principle (see section 1), because both the microhabitat availability and selfish herd mechanism would point to density-dependent movement. Apart from spatial, it is also expected that the clustering is temporal: a non-stationary pattern that can be explained by the phase separation principle.

There are four major sub-objectives to reach this objective. The first is to build a **microhabitat model**, so as to explain the presences of the nests with data derived from the drone and to look for the best scale (area around the nest) to be included in the model. A first hypothesis is that the wasps have specific preferences for the microhabitat to build their nest and that the optimal scale has to include more than the direct surroundings of the nests (cm-m). This predicted habitat suitability, from the microhabitat model, would form the bottom-up template for spatial pattern formation. A second hypothesis is that the microhabitat preferences change with time because later in the season, wasps move on to suboptimal habitat. A third hypothesis is that the microhabitat does not differ between parasitised and non-parasitised nests.

The second major sub-objective is to describe the population. First, by means of a capture-mark-recapture (CMRC) analysis. Populations size, survival probability and migration numbers will be estimated. Second, by means of a **spatial pattern analysis** of the nests. The hypothesis is that the nests are highly clustered in space and time.

The third major sub-objective is to make and analyse **social networks** with the clusters as nodes and with the individuals building a nest in different clusters as links. This is done in order to investigate movement between the clusters as well as the change in time of this movement and the importance of the clusters. It is expected that most of the wasps stay close to their previous nest, because of their high site fidelity, but that the importance of clusters changes with time.

The last major sub-objective is to assert the bottom-up and top-down processes of clustering. A **prediction map** to show suitability of the study area (bottom-up) will be made based on the microhabitat model. The hypothesis is that the suitable area is not completely saturated, because nest site selection is not only influenced by microhabitat availability. The density of the nests is linked to the density of the parasites (top-down). The hypothesis is that a dilution or **selfish herd mechanism** benefits the nests in high densities with a lower individual parasite density.

## 3 Material & Methods

### 3.1 Data collection

Field data were collected from 28-06-2016 until 15-08-2016 in the nature reserve De Westhoek in De Panne (West Flandres, Belgium), see figure 2, in an enclosure of 40 on 90 m. Sampling was done 30 days during a period of 49 days: when it is too cold ( $<20^{\circ}\text{C}$ ), too windy, rainy or cloudy the activity diminishes severely and the wasps stay in their burrows (Schöne & Tengö, 1991; Peeters, 2008). **Female digger wasps were individually tagged** with a plate on the thorax that was coloured and numbered (used for queen bees) with superglue. The fixation and tagging of the wasps was done with a marking tube, see figure 3a. Their **nests were marked with numbered flags**, see figure 3b. If a digger wasp brought a prey to her nest, was parasitised or closed a burrow, it was recorded in the notebook. Re-sights of previously tagged individuals were also recorded.

*Bembix rostrata* females show **test-digging behaviour**: they search and dig in the sand to look for an appropriate nesting place (Schöne & Tengö, 1991; Tengö *et al.*, 1996). Thus, when females were tagged at a place they were digging, it was not certain that it was a real nest (in which they lay an egg and provision the single larva with prey). To make a distinction in the dataset between real nests and all nests (real + apparent test-digging nests), real nests were identified using four criteria: (1) an individual arrives at the nest carrying a prey, (2) the digger wasp is seen at a nest during consecutive days (without prey) and (3) preys are found next to the nest and (4) an individual digger wasp visits or digs at the same nest multiple times during one day (not in the beginning of the season, see Tengö *et al.* 1996). With these criteria, 556 out of 1017 nests were identified as certain real nests.

The **position of the nests** was measured with an accuracy of approximately 2 cm with a **Trimble GPS** (Trimble Inc., USA) from the INBO (Instituut voor Natuur- en Bosonderzoek; Research Institute for Nature and Forest). It took 2 days to cover all the nests. The coordinates of the nests were provided in the Lambert 72 coordinate system.

Remote sensing imagery of the area was executed by Rpaswork and Landron Projects with 2 **drones** on 16-08-2016. A picture of one of the drones is shown in figure 3c. Aerial pictures in the visible and near infrared spectrum were taken. They provided a standard aerial photograph (RGB), a Digital Elevation Model (DEM) and raw data to calculate a NDVI (Normalised Difference Vegetation Index) map.

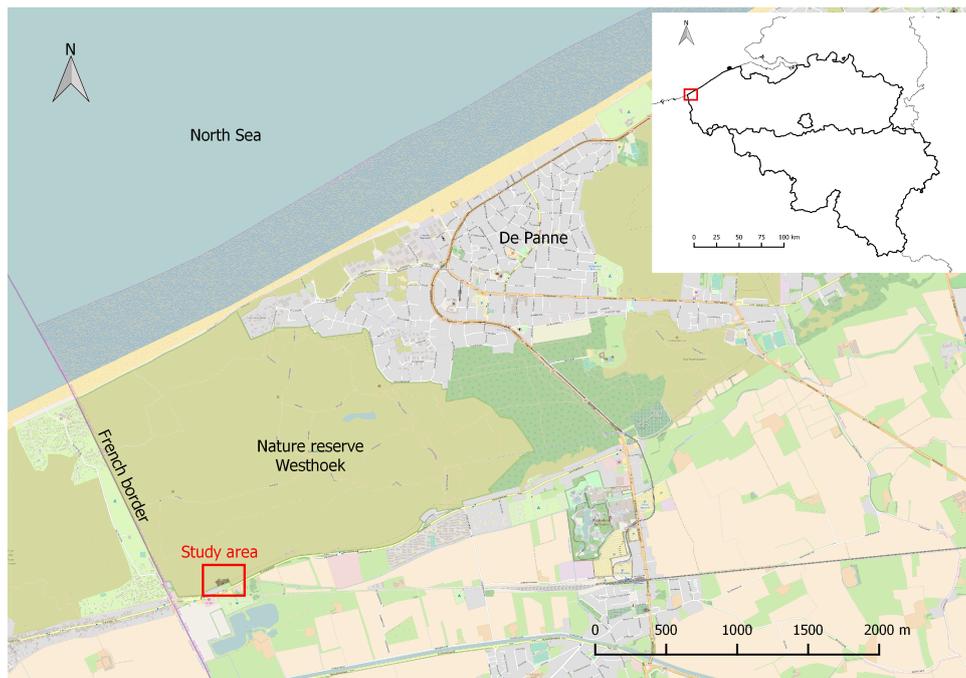


Figure 2: Overview of where the study area is situated

The data were incorporated in a Microsoft Access (2016) **database**, see Appendix 5a. Information of all records (sights), nests and days are present in different tables and linked which each other through relations. Useful data are then quickly obtained by filtering, making queries and pivot tables. Data calculated during analysis are added to the database.



(a) Marking tube with a tagged wasp, used for marking queen bees by beekeepers.



(b) Flags positioned next to the nests at an aggregate.



(c) One of the drones, ready for take-off.

Figure 3: Pictures of tagging, flags and a drone.

## 3.2 Microhabitat model

### 3.2.1 Presence-absence model

To investigate the nest site preferences of the digger wasps, a microhabitat model was build. **3 parameters** were considered, which were available through remote sensing imagery with the drones: NDVI (Normalised Difference Vegetation Index), slope and insolation.

Vegetation indices are a powerful tool to differentiate grey dunes, blond dunes and grasslands in monitoring coastal dunes (Provoost *et al.*, 2004). The NDVI is calculated for each pixel with formula (1), where NIR is the Near Infrared band (band 1 in the raw data) and Red the Red band (band 3 in the raw data). This was calculated in QGIS (QGIS Development Team, 2009), using the **Raster calculator**.

$$NDVI = \frac{NIR - Red}{NIR + Red} \quad (1)$$

The insolation (warmth) and slope were calculated from the Digital Elevation Model (DEM), by Sam Provoost at the INBO with ArcGIS (ESRI, 2011b):

- The **insolation**, warmth or Area Solar Radiation was calculated with the `Solar Radiation tool` (ESRI, 2011a), a `Spatial Analyst` extension. This takes into account the DEM and a time frame in which the insolation is summed (thus care should be taken when comparing different maps). The time frame was taken from 1-06-2016 to 31-08-2016. The unit of the calculated value at each pixel is watt hours per square meter (WH/m<sup>2</sup>). Details on the calculations of the insolation can be found on ESRI (2011a).
- The **slope** is calculated with the `Surface toolset`, also a `Spatial Analyst` extension. The slope is expressed in percent (%).

To have absences in the model, **random points** were generated with the function `runifpoint` in `Spatstat` in R (R Core Team, 2016) within the raster window of the point pattern. These random points are chosen to be at least 1 m away from any nest, an arbitrary cut-off level.

Before a statistical model was implemented, a **data exploration** was carried out. Correlations were calculated and pair plots were made for the explanatory variables, to see if there is any sign of collinearity. The mean and standard deviation SD of the variables for the nests is calculated, to have an idea of the range of these variables. The spatial autocorrelation was plotted in R with `spline.correlog` in the package `ncf` (Bjornstad, 2016). This is comparable to what Ripley's K (see formula 12a in section 3.3.2) represents.

To build the **statistical model**, the `glmmPQL`-function from the package `MASS` (Venables & Ripley, 2002) was chosen. With this function, a non-linear error distribution (binomial) of the response variable can be fitted together with the spatial autocorrelation (Dormann *et al.*, 2007). Because this function performs mixed models, at least one random factor had to be passed to `glmmPQL`: site was generated as an extra term in the data set, which had only 1 group, which is similar to what (Dormann *et al.*, 2007) did. The downside of this function is that no AIC's or Log likelihoods can be calculated. Thus, model selection was done with an alternative approach: the data was split up in training data and the (complementary) test data. The training data is used to build a model, predictions for the test data are calculated from this model. From this, the **AUC (Area Under Curve)** of the ROC-curve (Receiver Operating Characteristics) is calculated with the R-package `ROCR` (Sing *et al.*, 2005). The ROC-curve is the True Positive Rate plotted against the False Positive Rate of the predictions from the model of the test data. The function shows the amount of true positives among all positive predictions plotted against the amount of false positives (type I error) among all negative predictions at a certain threshold of prediction-value. The AUC is the area under the curve (integral) of this function. It is a measure used for binary data in which an AUC = 1 indicates perfect prediction. Thus, the closer the AUC to 1, the better the model predicts the test data. More information how ROC and AUC are exactly calculated and their interpretation can be found in Brown & Davis (2006).

Before a conclusive model was built, preliminary models/tests had to be performed to make choices about which autocorrelation function to implement, which explanatory terms to include and which spatial scale is most optimal to explain nest presences. To choose an appropriate **autocorrelation function**, models were run with 10 different subsets of 200 random points of the dataset. The mean AUC of the different autocorrelation functions are compared. To select **which explanatory terms** should be included in the model, also considering quadratic terms, a subset of 1000 random points was considered. The full model is  $Presence \sim NDVI + insolation + NDVI^2 + insolation^2$ . An interaction term is also tested in the model  $Presence \sim NDVI + insolation + NDVI : insolation$ . For different models, different terms of the full model were combined or an interaction term considered, the AUC is calculated and the level of significance of the terms is considered too. It is investigated at which **scale** nest site preferences is most optimally predicted. From the maps of the 3 parameters, values were extracted at different radii from the points (nests and random points) in QGIS, see figure 4. For this, **buffers** of 0.1 m, 0.2 m, 0.5 m, 0.8 m, 1 m, 2 m, 3 m, 5 m and 10 m were drawn around each point with the Geoprocessing tool **Fixed distance buffer**. In these buffers, the mean (and the amount of pixels (count), median and standard deviation) of the values were calculated using the plugin `Zonal statistics`. To get the pixel values of the points, the plugin `Point sampling tool` was used. All these data were joined, exported and added to the Access database, to have one data file with presence-absence points, their coordinates and explanatory parameters. 10 different subsets of 200 random data points in the presence-absence data set are used to run models for all the different buffers. The predicted values are each time calculated for the complement of the 200 random points. Here too, the mean AUC is used as selection criterium. The

selection of the terms that should be included is again run at the chosen scale.

The **final conclusive model** is run with all nests and random points, the chosen correlation function, linear terms of NDVI and insolation and scale. The training data was 70% randomly chosen data points from the full data set and the test data was the other 30% complementary data points.

### 3.2.2 Real nests model

To test if the microhabitat characteristics differ between real nests (presences) and other nests (absences), the full model from previous section was ran with the nest points. Presence was defined by a real nest, absence by another nest.

### 3.2.3 Parasitised nests model

To test if parasitised nests have other microhabitat characteristics, a full model was ran again. Presence is defined as parasitised nests and absence as non-parasitised nests.

### 3.2.4 Temporal microhabitat model

To know if the microhabitat preferences change with time, the predicted values of the nests are plotted against day the nest was marked. A Generalised Linear Mixed Model with spatial autocorrelation is fitted to this relation, with the `glmmPQL`-function of the package `MASS` (Venables & Ripley, 2002). If there would be a difference in microhabitat suitability with time, the factor day will be added to the full model. When interactions with the other factors (NDVI, slope or insolation) would be included, it could be deduced which factor is responsible for the change in time of the habitat suitability.

## 3.3 Describing the population and spatial pattern

### 3.3.1 CMRC: population estimation

To estimate the amount of female digger wasps in the focal area, a population estimation was carried out based on the capture mark recapture data (CMRC; re-sights of previously tagged individuals each day). A Jolly-Seber model was applied, which has the assumption of an open population: there was a constant influx of untagged wasps, and tagged wasps vanished from the focal area. Considering the biology of *B. rostrata* (hatching only takes place at the beginning and mortality at the end of the season), the wasps most likely immigrated from and emigrated to adjacent areas. Mortality is considered as a minor cause for the vanishing, because during the whole period of data collection in the focal area, only a few dead males (which were not tagged) and only one dead female (in mid-August after several cold days) were found. Calculating the survival probability  $\phi$  can test this hypothesis of minor mortality during the season. Formulas and a spreadsheet to calculate the estimations for the CMRC-data were found on Larsen (2014). Estimations of the population size  $N_i$ , probability of survival  $\phi_i$  and joining number  $B_i$  are calculated for subsequent sample periods (days). The mean of those 3 parameters gives an estimation of that parameter for the complete study period. The survival probability  $\phi_i$  is between 0 and 1, the joining number  $B_i$  can have positive or negative values which indicate immigration or emigration numbers, respectively.

The following equations (2) are used for the estimations, with  $i$  the number of the day:

$$N_i = \frac{n_i \cdot M_i}{m_i} \quad \phi_i = \frac{M_{i+1}}{M_i + (n_i - m_i)} \quad B_i = N_{i+1} - N_i \cdot \phi_i \quad (2)$$

with  $M_i = \frac{n_i \cdot z_i}{r_i} + m_i$  the amount of tagged individuals before day  $i$ ;  $m_i$  amount of tagged individuals caught at day  $i$ ;  $n_i$  the total amount of individuals caught at day  $i$ ;  $r_i$  the amount of individuals caught and released at day  $i$  that are caught in the future and  $z_i$  the amount of individuals tagged before day  $i$ , not caught at day  $i$  but caught at a later day.

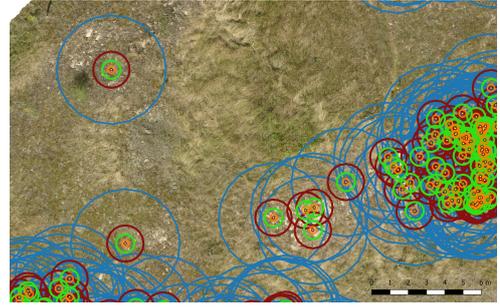


Figure 4: Buffers around the nests (yellow): 0.2 m (orange), 0.5 m (green), 1 m (dark red) and 3 m (blue).

This CMRC analysis was repeated in Program MARK (White & Burnham, 1999), which uses numerical maximum likelihood techniques to estimate the parameters. The POPAN parametrisation of the Jolly Seber model was used, which gives one estimation of the population size  $N$ . The probability of survival was calculated for the different intervals, but the probability of entry (pent) and the recapture probability (p) were kept constant.

### 3.3.2 Point pattern analysis

To examine the spatial pattern of the nests, a point pattern analysis was carried out in R (R Core Team, 2016) with the R-package `Spatstat` (Baddeley *et al.*, 2015), according to the guidelines of the manual by Baddeley (2010).

The coordinates of the nests are used to make a **ppp-object** (planar point pattern) in `Spatstat`, and used to easily plot the data and make calculations. Information of the points can be included in a ppp-object as ‘marks’: if it is a real nest (Real Nest; 0-1), if it is parasitised (Par; 0-1), the day of marking the nest (Day; 1-49) and probability of presence (ranging from 0 to 1) based on the microhabitat model prediction of each nest .

In `Spatstat`, **densities** can be estimated for a ppp-object. This density-function computes an intensity function from a point pattern with an isotropic smoothing kernel (with sigma the standard deviation for this kernel). It estimates for every point in a pixel image the expected density based on the values of the kernels of surrounding points at that pixel.

Spatial 2D-statistics can be based on 3 different measures of distance in point patterns: pairwise distances, empty space distance and nearest neighbour distance. For each of these distances, corresponding cumulative distribution functions exist. These functions can be used to examine if the point pattern is clustered, random or regular. The null model for a cumulative distribution function is complete spatial randomness, modelled with a Poisson distribution. In the main text, pairwise distances and the corresponding **Ripley’s K** and **pair correlation function** will be used. The definition and application on the focal data of empty space distance, nearest neighbour distance and their corresponding functions, can be found in Appendix 2c. Pairwise distances are represented as a distance matrix: for each point, the distances to the other points is measured. The Ripley’s K function calculates for each point the relative amount of points that are present within the distance radius, this for several distances. The calculation of Ripley’s K for distance  $d_s$  is shown in formula (3).

$$K(d_s) = \frac{1}{\lambda} \sum_{i=1}^N \sum_{j=1}^N \frac{\delta_{ij}(|x_i - x_j| \leq d_s)}{N} \quad (3)$$

With  $\lambda$  the density of points in the whole area,  $N$  the total amount of points,  $d_s$  the focal distance for which Ripley’s K is calculated and  $\delta_{ij}$  the dirac delta function.

Ripley’s K, for a selected distance range, is represented in graphs with the values of K on the vertical axis and the distance on the horizontal axis. Edge-corrections can be calculated, which take into account that a point pattern can extend the window (or sampling) considered. When K is larger than expected for the Poisson-K, the points are clustered. Simulation envelopes are made around the Poisson-K to be able to see the possible range of complete spatial randomness. These envelopes are based on 99 Monte Carlo simulations.

Related to Ripley’s K, the pair correlation function gives the probability to observe a pair of points separated by a distance  $r$ .

To take into account the types of the points (or marks in the ppp-object) in the analysis of a point pattern, **mark equality functions** are calculated. It gives the probability that two points have the same type (or mark) when separated by a distance  $r$ . This is calculated for the real nests, the parasitised nests, day the nest was marked and predicted probability of presence from the microhabitat model. Simulation envelopes, based on 99 Monte Carlo simulations, are plotted around the random mark correlation (1.0-line) for each of these marks. This to see if the mark equality function is different from complete spatial randomness.

## 3.4 The population as a network

### 3.4.1 Distance kernel between two subsequent nests

The distances between two subsequent nests from the same female were quantified. A distance matrix for the nests was calculated and exported from ArcGIS (ESRI, 2011b). The information of which real nest belongs to which individual was extracted with a query from the database. Those two files were combined in a Python script (Python Software Foundation, 2001) to filter the distances of subsequent nest of the same individual. A histogram of the distances grouped per 2 meter is made to give the distance kernel for subsequent nests.

### 3.4.2 Cluster analysis

When a point pattern is clustered, a cluster analysis can be performed to identify the clusters present and assign each point to such a cluster. In R, **k-means clustering** is performed. This method uses the Euclidian distance between points and assigns points to a cluster when the centre of the cluster is the closest for that point. The number of clusters is an input parameter. The analysis will be carried out for several numbers of clusters, for all and for real nests. Visualisation of the clusters, consideration of density plots and the radius of correlation between nests found in the point pattern analysis, helps to decide the most parsimonious amount of clusters to use in further analysis.

### 3.4.3 Social network

Exchange between clusters of subsequent, real nests of individuals is further analysed by means of network analysis. The R package **igraph** (Csardi & Nepusz, 2006) is used. The nodes (or vertices) are defined as the clusters from the cluster analysis and the links (or edges) as an individual that makes it subsequent nest in another cluster.

The total amount of real nests, relative parasite density and microhabitat suitability are added to the information of the nodes as **node properties**. The first two measures are filtered from the database using a Python script (Python Software Foundation, 2001), with the relative parasite density the amount of parasitised nests divided by the amount of real nests in a cluster. To calculate the latter measure, the microhabitat suitability, buffers around each nest in each cluster are made and merged for each cluster. Mean NDVI, slope and insolation are calculated for the merged buffers to predict the suitability of each cluster with the full model from section 3.2.1.

The correlation between the node properties is calculated, to assess which independent properties can to be used in further analysis.

Both **directed and non-directed networks** are considered, because links could define a basic connection between two clusters (undirected) or the direction of the links might matter (directed). For the focal system it is not straightforward which to choose in advance. Non-aggregated networks (in which each individual link between the nodes is represented) are aggregated, in which each link gets an attribute 'weight': the count of links between the two clusters connected.

Following **node and network metrics** are calculated to describe the network, these are based on Farine & Whitehead (2015) and Ognyanova (2016):

Network metrics:

1. The amount of **internal loops**: the relative amount of links that return to the same node. This happens when a digger wasp builds its next nest in the same cluster.
2. **Density or Connectance**: the number of (weighted) links in a network divided by the total possible links.
3. **Reciprocity** (for directed networks): gives the relative amount of connections that are in both directions from one node to another.
4. **Assortativity**: gives the correlation in properties between connected nodes. Negative correlation implies that nodes with opposite value of the property are linked, positive correlation implies linkage between nodes with similar values.

5. **Transitivity or Clustering Coefficient:** calculates the proportion of triads (trios of nodes) that have three links divided by the amount of triads that have two links. It quantifies the clustering of the ‘global’ network: if node x is connected to node y, which is connected to node z, x is also connected to z.

Node metrics:

1. **Degree** (binary network) or **Strength** (aggregated network): the count (or weighted count for Strength) of number of edges connected to the node. For a directed network these can be partitioned into in-degree and out-degree.
2. **Centrality measures:** 3 centrality measures are used, but more exist (see reference manual **igraph** (Csardi & Nepusz, 2013)). They all quantify how ‘central’ a node is in the network with different methods and different definitions for centrality. The **Eigen-vector centrality** captures the influence of a certain node, this by taking into account its own degree but also the degree of the nodes it is connected to. **Betweenness centrality** is the count of shortest paths that flow through the node. This measures how important a node is to connect disparate parts of a network. **Closeness centrality** is calculated as the sum of all the shortest paths from the node to all the other nodes in the network. Other centrality measures are often a variation on these three (especially on the eigen-vector centrality).
3. **Hubs & Authorities** (for directed networks): a high hub-score represents the tendency to have many outgoing links, a high authority-score represents many ingoing links.
4. **Transitivity or Clustering Coefficient:** this can also be calculated for each ‘local’ node. How many triads the node forms with other nodes.

The relationship between those metrics and (1) the size of the nodes, (2) the relative amount of parasites present and (3) microhabitat suitability, is looked at through General Linear Models in R (the **lm**-function) or Generalised Linear Models (the **glm**-function). The results are put in a table for all the metrics to summarise and be able to make overall conclusions.

To see if the matrix of the amount of links between the different clusters is related to the matrix of Euclidian distance between the clusters, a **Mantel test** is carried out. This to investigate if exchange between clusters is related to the distance between clusters.

The network is split up in 3 time frames (a time-aggregated network), a **temporal network**. The breaks are chosen at every 10 days of sampling, to investigate the temporal shift in the node and network properties. This is done with the R-packages **timeordered** (Blonder, 2015) and **igraph** (Csardi & Nepusz, 2006).

Network plots are made for the different timeframes and these are qualitatively compared. Boxplots are made to explore the mean size, amount of parasites, relative parasite density, degree, eigenvector centrality and transitivity of the clusters for the 3 timeframes.

The variability ( $CV = sd/mean$ ) of the different network metrics for each cluster are related with habitat suitability to explore if the variability during the 3 timeframes is related with the microhabitat model predictions.

## 3.5 The causes of clustering

### 3.5.1 Microhabitat availability

A map is made to project the predictions of the full microhabitat model on the study area. In QGIS regular spaced points are generated with the Research tool **Regular points**. The spacing between the points is chosen as 0.5 m. The buffer of the appropriate scale of the microhabitat model (see 3.2.1) is drawn and mean values of NDVI, slope en insolation are extracted for all the regular points. Predictions of the full microhabitat model are made in R. These are plotted as a contour map with all the nests added. The covering of the nests is compared with the predicted suitability of the microhabitat.

### 3.5.2 Parasite density

**Density plots** were made for both nest and parasite presence with the density function of the ppp-class (see 3.3.2). Same sigma (= 1) is taken for nests and parasites, to have comparable density-estimations. Densities at each pixel are extracted in a data frame to make scatterplots and infer relations between density of parasites and density of nests. The individual chance of parasitism, or parasite density per individual, was calculated for each pixel as shown in formula (4).

$$\text{Chance of infection per nest} = \frac{\text{Density parasites}}{\text{Density nests}} \quad (4)$$

Only real nests were taking into account because it is only relevant for those to be infected by brood parasites.

## 4 Results

The database, raw data and the maps used in GIS can be found in Appendix 5.

### 4.1 Data collection

A total of **732 individual digger wasps were tagged** and **1017 nests were marked with a flag**. In total, **1807 records** were noted down. While tagging the wasps, glue was spilled over their wings sometimes, especially in the beginning of the field work, when the handling still had to be learned. 16 digger wasps were lost this way, which is 2.1 % of the tagged individuals. Another casualty was that wasps sometimes lost their tags, which could have an influence on the uncertainty of CMRC-estimations and the detection of consecutive nests of individuals. 7 tags were found in the sand and 5 wasps were caught which had traces of glue on their thorax, which is at least 7 to 12 tags that were lost (0.9-1.6%).

#### Prey

Several times, wasps were seen dragging prey (pieces and complete individuals) from their burrows. The prey were collected and in 5 (out of 17) samples, small Diptera larvae emerged. Also, at previously parasitised nests, prey laying next to a burrow were found. So when prey were found laying next to a nest, these were also collected, and in 9 (out of 36) samples Diptera larvae developed.

The hoverflies that were collected were identified as being from the genera *Eristalis*, *Helophilus*, *Eupeodes* and *Episyrphus*. Other small unidentified flies (Diptera) were also seen or collected as prey. It was observed once that a digger wasp was flying around with a specimen of *Bombus terrestris* (Apidae, Hymenoptera).



Figure 5: Tagged wasp with a prey (red arrow) and a parasite on the prey (dark grey arrow). Picture by Tom Van Assche

### 4.2 Microhabitat model

The R-code and complete analysis can be found in Appendix 1. The raw data and maps to use in GIS can be found in Appendix 5.

#### 4.2.1 Presence-absence model

Figure 6 gives maps of the 3 parameters NDVI, insolation and slope for the microhabitat model. The random points are shown in blue in figure 7a. The correlogram indicates spatial correlation, see figure 7b, or clustering up to a radius of 15-18 m. This correlation is negative between 18 m and 35 m, which indicates a more regular pattern at these scales, probably caused by the space between the large nest clusters. This is comparable to Ripley's K (see section 4.3.2), but here also the random points are considered in the analysis, which makes this correlogram not as reliable as Ripley's K to determine scale for clustering. This implies that for the microhabitat model, all datapoints must be corrected for spatial autocorrelation.

Figure 7c shows the **collinearity** between the 3 parameters at a scale of 2 m. The correlation of these parameters is calculated:  $\rho_{NDVI,ins} = -0.54$ ,  $\rho_{NDVI,slope} = 0.24$  and  $\rho_{ins,slope} = -0.59$ . Considering figure 7c and these correlations, slope was not included in the model, because it is largely correlated with insolation. Although insolation is to a lesser extent collinear with NDVI, figure 7c shows there is still a lot of variation between insolation and NDVI. The hockey-stick shape between insolation and slope indicates a more important and stronger (non-linear) correlation. Insolation is also biologically a more relevant measure to interpret for nest choice in *Bembix rostrata*: ground temperature might be a more direct cue for this xerothermic digger wasp for e.g. the development of or correct microclimate for its larva. More indirect arguments exist for slope, as e.g. the stability of the nest entrance. After the complete model was made, the model was ran with slope instead of insolation, and it performed a little less good. Insolation was used for the final model, keeping in mind the biological interpretation of both.

The mean insolation of the nests is  $448\,194 \text{ WH/m}^2 \pm 14\,533 \text{ WH/m}^2 \text{ SD}$  and the mean NDVI of the nests is  $0.1204 \pm 0.0347 \text{ SD}$ .

From the several **autocorrelation functions** tested for the model, only two converged. The model was a linear combination of NDVI and insolation at a scale of 2 m. The test with 10 different subsets of training data of 200 points for the two remaining functions yielded an  $\text{AUC} = 0.913 \pm 0.003 \text{ SD}$  for the exponential correlation function and  $\text{AUC} = 0.899 \pm 0.03 \text{ SD}$  for the Gaussian correlation function. The two functions perform equally at scale = 2 m. At lower scales, the Gaussian correlation function performed slightly worse, thus the **exponential correlation function** was used in the following analyses.

Models were compared which included different combinations of **linear and quadratic explanatory terms** (executed with a subset of 1000 points training data, the complementary, 1018 points used as test data). The AUC of the different models and p-values of each term in a model are given in table 1. The AUC-values are all very close to each other and considered equal. Different combinations give equally performing models (considering p- and AUC-values). The linear model is chosen to proceed with in further analysis: it performs as good as any quadratic model, is the most simple (and thus most parsimonious) and is more straightforward to interpret biologically.

The AUC is plotted against **scale r (m)**, see figure 7d. The model used is a linear combination of NDVI and insolation. AUC at each scale is calculated for 10 iterations of the model with a subset of 200 training data points. Models with scales between 2 and 3 m perform the best. 2 m was chosen to use in further analysis. The presence of nests is influenced by insolation and NDVI up to a radius of 2 m. The performance declines after the 5 m scale, which corresponds to the radius of clustering (see section 4.3.2). This scale measure indicates that the surrounding vegetation and insolation up to a radius of 2 m is important for the habitat selection of a nest for *B. rostrata*.

A **full model** was run with exponential autocorrelation function, scale = 2 m and linear terms NDVI and insolation. 70% of the points was used as training data, 30% as test data. For this full model, an AUC of 0.9104 was calculated. For both the insolation-term and NDVI-term  $p < 0.00001$ . Formula 5 is the complete predictive formula for the full model at scale = 2 m. The nests have higher insolation and less vegetation (lower NDVI) than random points in the study area.

$$\ln\left(\frac{\text{Presence}}{1 - \text{Presence}}\right) = 0.000045 \cdot \text{insolation} - 27.28477 \cdot \text{NDVI} \quad (5)$$

#### 4.2.2 Real nests model

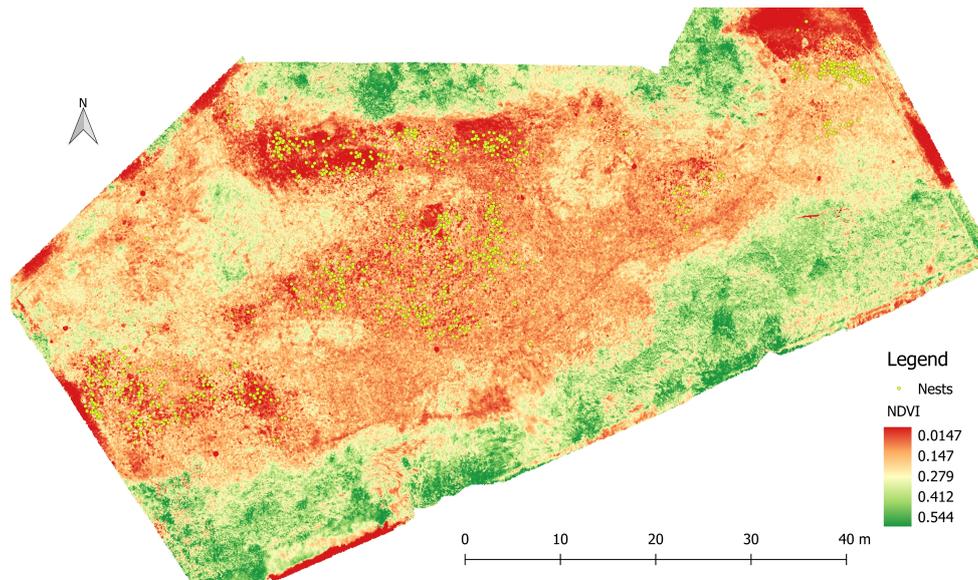
To see if real nests have other abiotic characteristics, a full model was run for the real nests (presence) and other nests (absence). This resulted in an  $\text{AUC} = 0.515$  and p-values for insolation and NDVI of 0.493 and 0.199, respectively. These terms cannot detect any difference in microhabitat conditions between real and the other nests. An AUC of around 0.5 indicates a random real positive or false positive prediction.

#### 4.2.3 Parasitised nests model

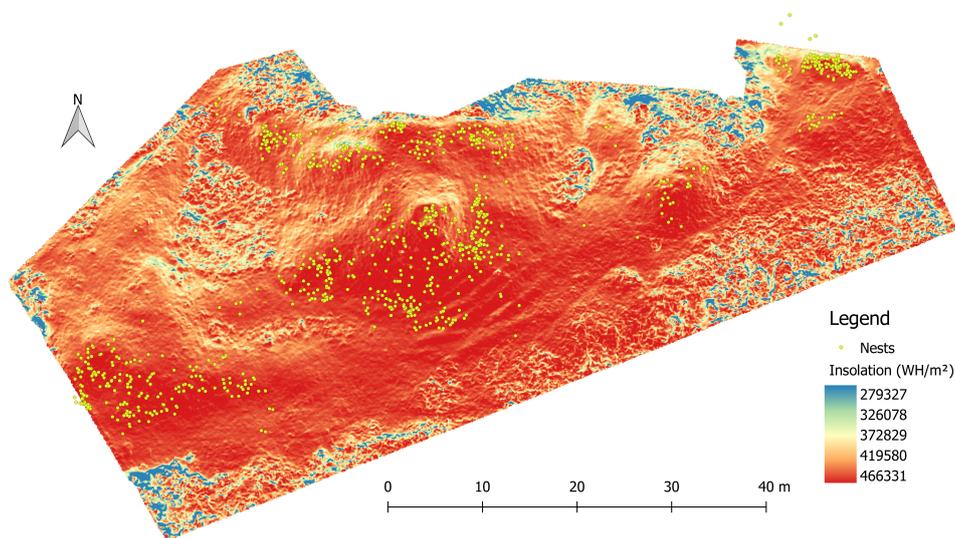
A similar model was run for parasitised nests (presence) and non-parasitised nests (absence). This resulted in an  $\text{AUC} = 0.625$  and p-values for insolation and NDVI of 0.569 and 0.754, respectively. These terms cannot detect any difference in microhabitat conditions between parasitised and non-parasitised nests.

#### 4.2.4 Temporal microhabitat model

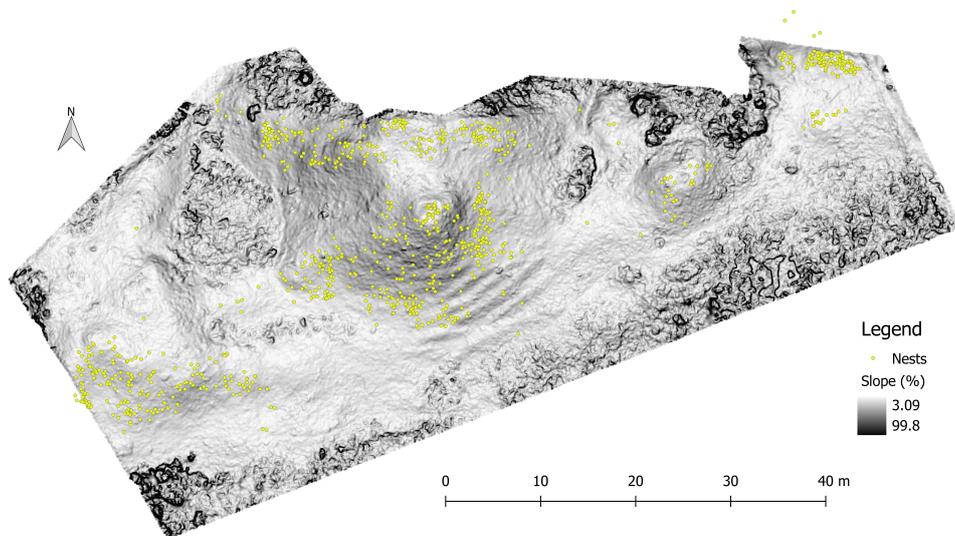
The predicted values were plotted against day number, see figure 8, the colours of the dots indicate the cluster number. The GLMM with spatial autocorrelation structure returned a model with p-value = 0.96 for the term day. There is no difference in habitat suitability with time. Figure 8 indicates that the poorly predicted points (model prediction between 0.10 and 0.25; false negatives, type II errors) are mainly nests from the northeast corner on the map.



(a) False colour map of Normalised Difference Vegetation Index (NDVI). Resolution 1 cm.

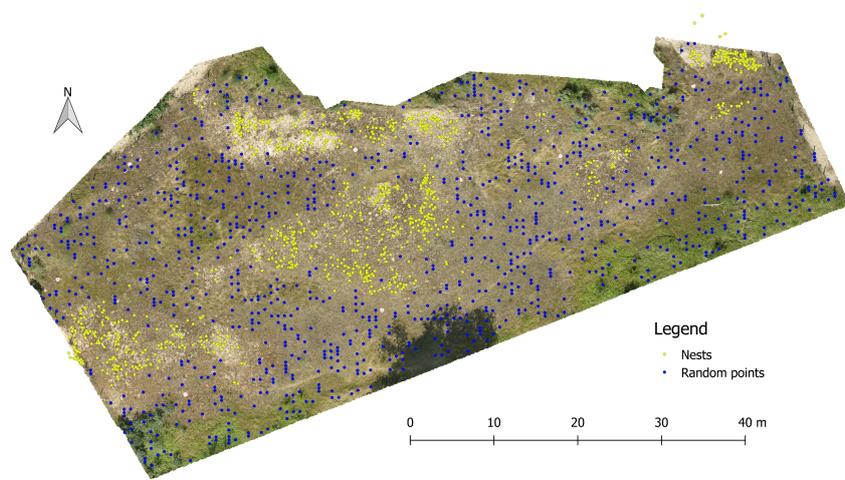


(b) False colour map of insolation. Resolution 7 cm.

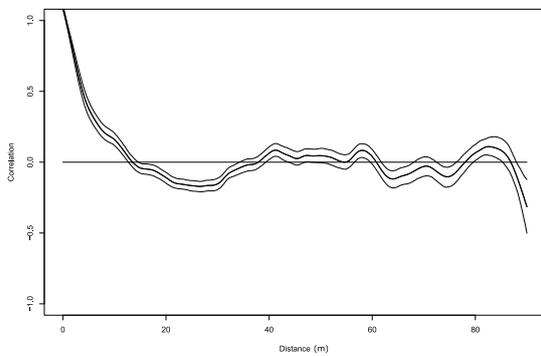


(c) False colour map of the slope. Resolution 7 cm.

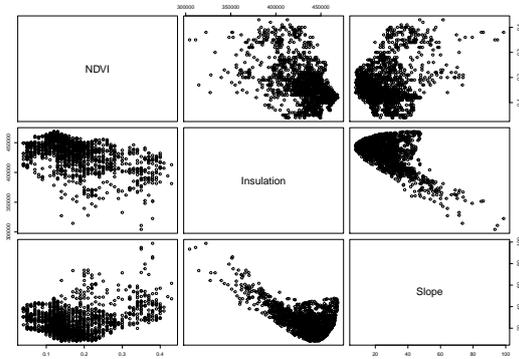
Figure 6: False colour maps of the 3 parameters in the microhabitatmodel. Nests are plotted as yellow dots.



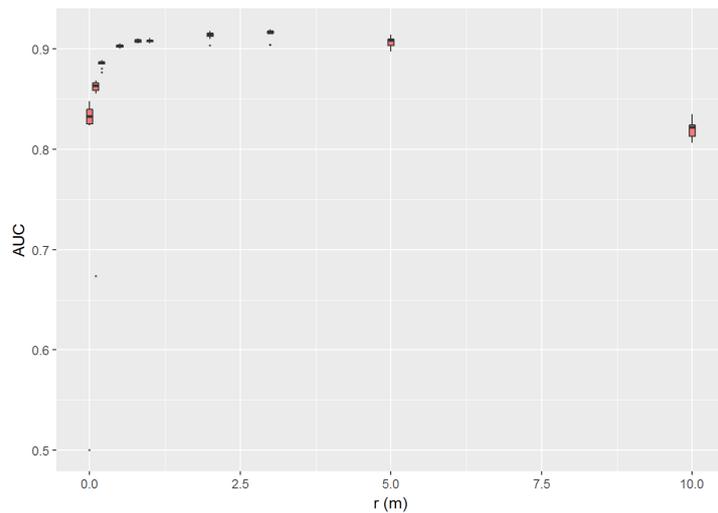
(a) Random points and nests shown on the aerial photograph.



(b) The correlogram of the spatial autocorrelation of the nests.



(c) Pair plots of the 3 parameters at scale = 2 m to explore collinearity in the data.



(d) AUC at different spatial scales, with boxplots for each scale-value (middle line is median, upper Q3 and lower Q1 quartiles are limits of the box, whiskers maximum values excluding outliers, small dots are outliers ( $>Q3 + 1.5 \cdot \text{boxlength}$  or  $<Q1 - 1.5 \cdot \text{boxlength}$ )).

Figure 7: Random points plotted on the aerial map, spatial autocorrelation, collinearity of the terms and the spatial scale plot.

Table 1: Model comparison of models which include different linear and quadratic terms. Significant terms are in bold, chosen combination is highlighted in red.

<i>Terms included in the model</i>	<i>p-values</i>	<i>AUC</i>
NDVI <sup>2</sup> insolation <sup>2</sup> NDVI insolation	0.7788 <b>0.0162</b> 0.1422 <b>0.0121</b>	0.914
insolation <sup>2</sup> NDVI insolation	<b>0.0162</b> < <b>0.00001</b> <b>0.0121</b>	0.914
<b>NDVI</b> <b>slope</b>	< <b>0.00001</b> < <b>0.00001</b>	<b>0.915</b>
NVDI insolation NVDI:insolation	0.6148 <b>0.0072</b> 0.3579	0.916
NDVI insolation <sup>2</sup>	< <b>0.00001</b> < <b>0.00001</b>	0.915
NDVI <sup>2</sup> insolation <sup>2</sup>	< <b>0.00001</b> < <b>0.00001</b>	0.916
NDVI <sup>2</sup> insolation	< <b>0.00001</b> < <b>0.00001</b>	0.916

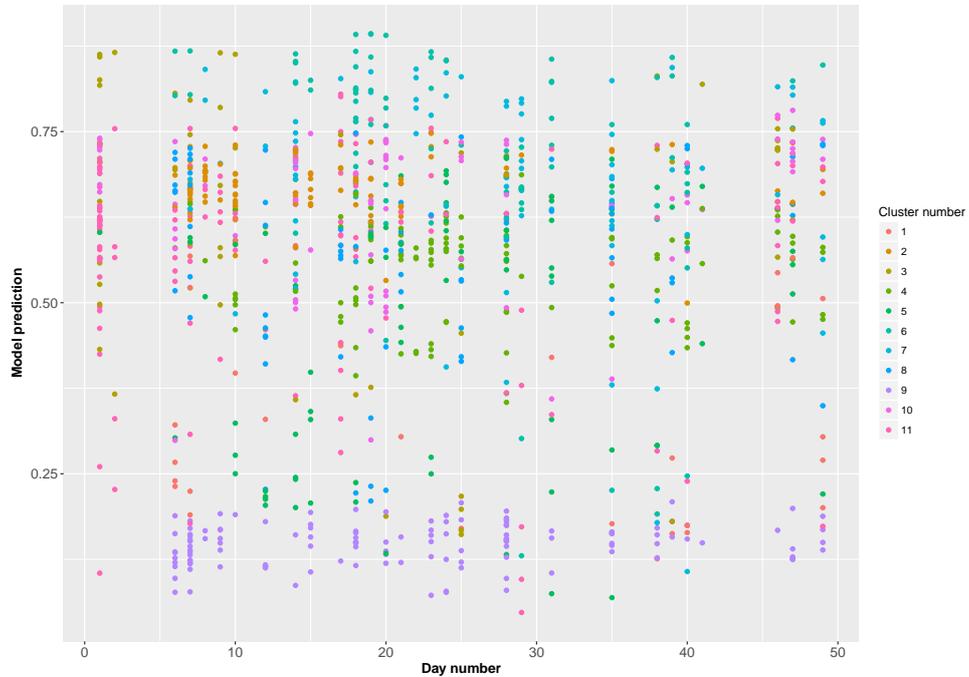


Figure 8: Predicted value of the model plotted against the day number.

## 4.3 Population estimation and spatial pattern analysis

### 4.3.1 CMRC: population estimation

The Jolly Seber population estimation yielded a population size of on average  $\bar{N} = 1125 \pm 119$  SE female individuals. The survival probability was on average  $\bar{\phi} = 1.15 \pm 0.14$  SE, which supports the hypothesis that there is no significant mortality during the sampling period. The average joining number  $\bar{B} = -16.9 \pm 61$  SE indicates no net immigration or emigration, but strongly fluctuating positive and negative numbers during the complete period (see Appendix 2a and 2b) indicate that migration to and from adjacent areas might be of considerable importance. Considering that the mortality is almost non-existing during the field period, this importance of migration is confirmed by the fact that 318 out of 732 were tagged and never seen again (43%). Birth is considered not to influence the joining numbers  $B_i$ , as in this system all individuals emerge in the beginning of the season (end of June).

Tables of all estimations of subsequent days can be found in Appendix 2a and 2b.

The estimation in Program MARK provided a population size of  $N = 1256 \pm 55$  SE, and a survival probability of on average  $\bar{\phi} = 0.95$ . These are consistent with the manually calculated estimations.

### 4.3.2 Point Pattern Analysis

All code, complete analyses and extra analyses can be found in Appendix 2c.

#### Visualisation

The nests are plotted on the aerial photograph taken by the drone (resolution of 1 cm), see figure 9.

The four different marks are plotted with *Spatstat*, see figure 10. The configuration of the nests themselves are plotted in figure 10a. The plot for day or ‘temporal gradient’ (10b) suggests that there is clustering in time: the colours are often grouped, but there is also a lot of variation. The plot for real nests (10c) does not indicate any visual pattern. The plot for parasitised nests (10d) also does not show any visual pattern. The plot for the microhabitat model prediction (10e) suggests clustering in space of the microhabitat and that the edges of the aggregates are less well predicted. It also indicates the bad predictions for the nests in the northeast corner.

Density plots are made, see figure 11. This is done for all the nests and the real nests only (sigma = 2 was taken constant, see 3.3.2). These indicate clustering of the nests.

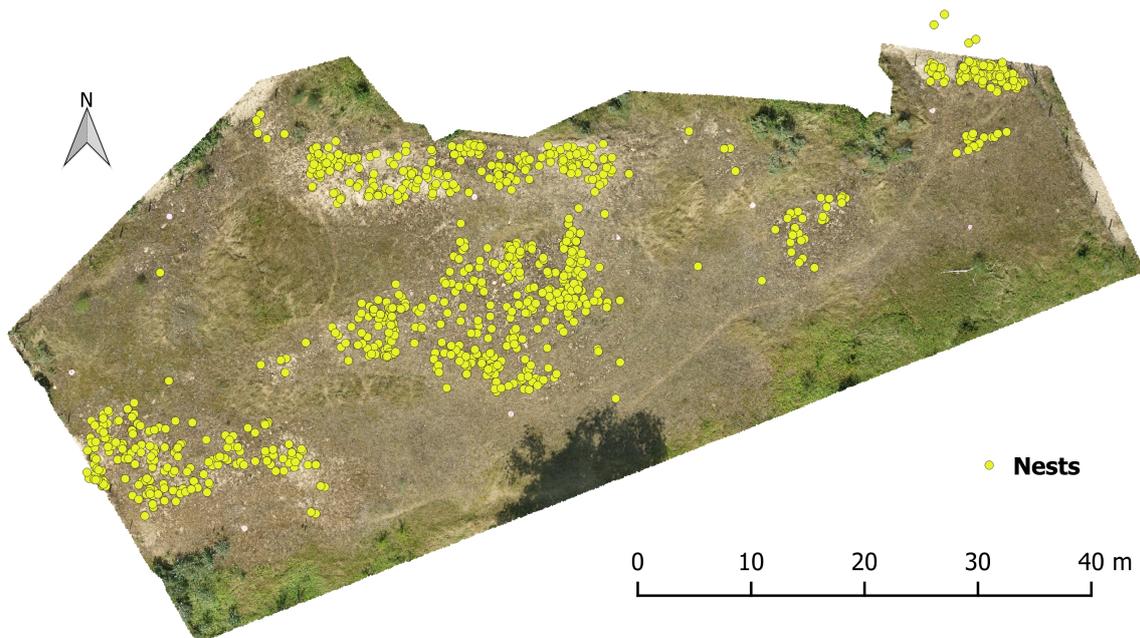


Figure 9: Map of the study area. Nests are indicated with yellow dots. Pixel resolution is 1 cm.

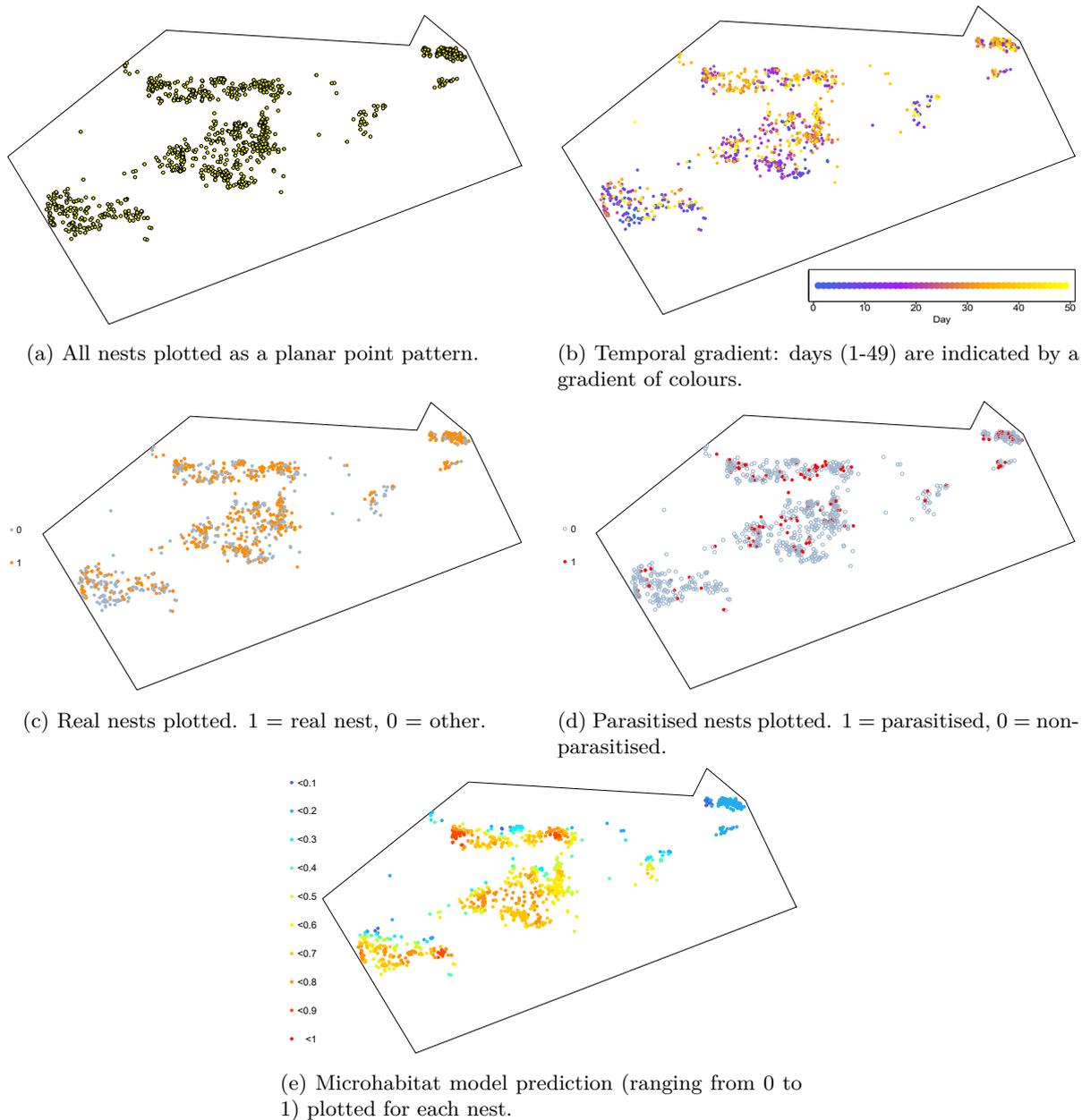


Figure 10: Nests plotted as a ppp-object in **Spatstat** and plots with the 4 marks: day, real nests, parasitised nests and model prediction.

## 2D-statistics

Extra spatial statistics are explained and applied to the data in Appendix 2c.

**Ripley's K** function, see figure 12 (12a includes edge-corrections, see 3.3.2; figure 12b shows simulation envelopes around the Poisson Ripley's K), confirms the conclusion of the density plots, that the nests are indeed clustered. The Ripley's K line (black) is larger than the Poisson-K line (blue in 12a and red in 12b) up to a scale of 30 m. This indicates that nests are clustered up to a radius of about 15 m. The density plots (fig. 11) and aerial map (fig. 9) suggest that there is clustering at even smaller scales, within radii of 3-5 m.

The **pair correlation function** is shown in figure 13. It shows the probability to observe a pair of points separated by a distance  $r$ . This graph confirms that there is clustering at smaller scales of up to 5 m.

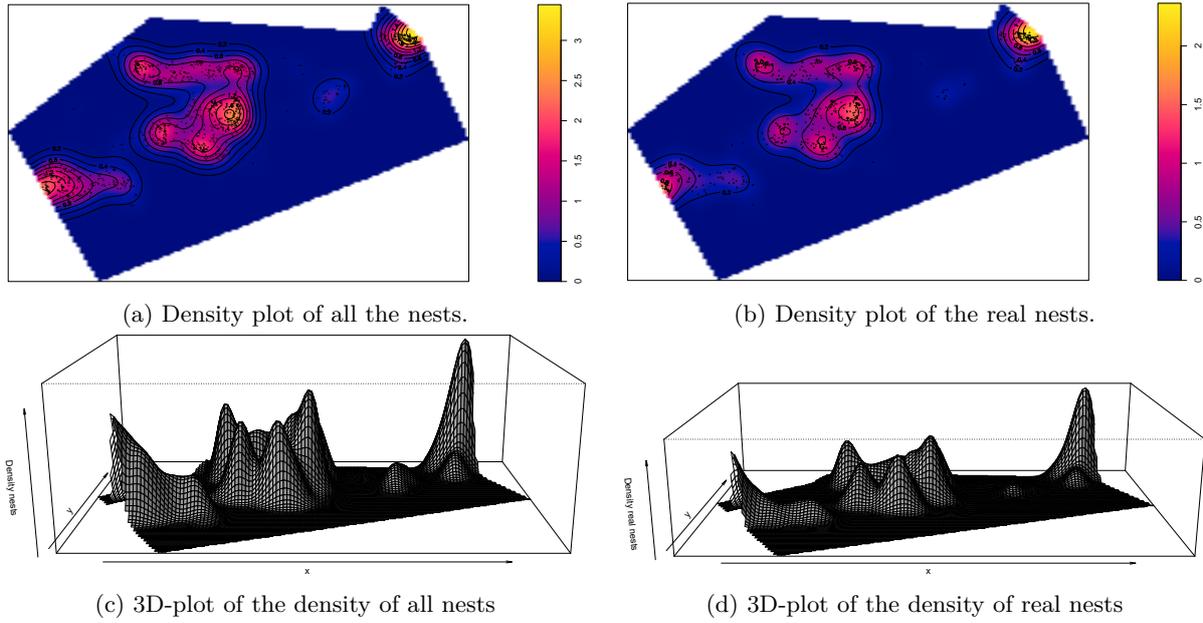
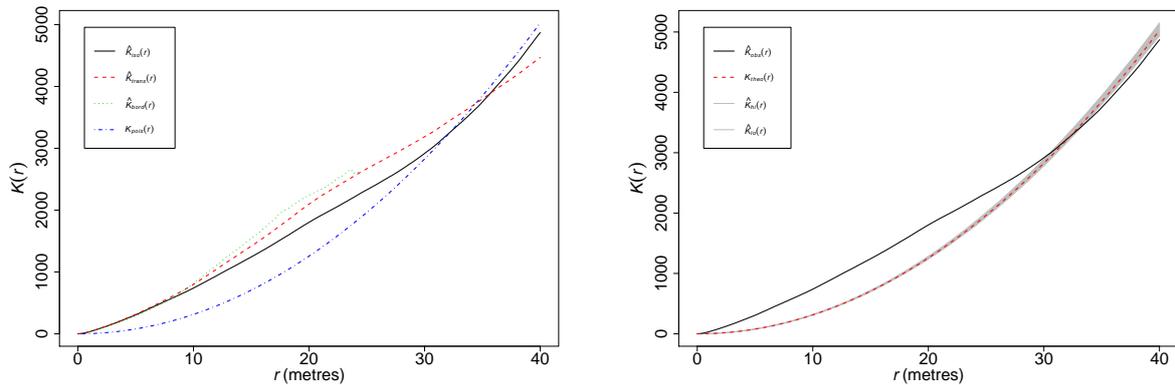


Figure 11: Density-plots, all calculated with  $\sigma = 2$ .

### Mark equality function

The mark equality plots, see figure 14 give the probability that two points have the same type (or mark) when separated by a distance  $r$ . The marks are *real nest*, *day*, *parasitised* and *prediction*. Monte Carlo simulation envelopes are plotted around the random correlation line, to see if the mark equality function deviates from complete spatial randomness.

These plots confirm the spatial plots of the marks (fig. 10). The mark equality plot of the real nests (fig. 14a) are equally or randomly distributed over all the nests: they are a spatial subset of all the nests. The mark equality plot of day (fig. 14b) confirms that there is temporal clustering of the nests, up to a spatial scale of 5 m, the same radius as the clusters. The parasitised nests (fig. 14c) are randomly distributed over all nests. The microhabitat predictions (fig. 14d) are spatially clustered, this suggest that nests near each other (up to 10 m) tend to have the same microhabitat suitability. This suggest that the habitat suitability will differ between spatial clusters. Clusters spatially close to each other will not differ as much as further away clusters.



(a) Ripley's K ( $K_{iso}$ ).  $K_{trans}$  and  $K_{bord}$  are edge-corrections.  $K_{pois}$  is the random, Poisson Ripley's K.

(b) Ripley's K with simulation envelopes for  $K_{pois}$ , these are based on 99 Monte Carlo simulations, with the assumption of complete spatial randomness.

Figure 12: Ripley's K function

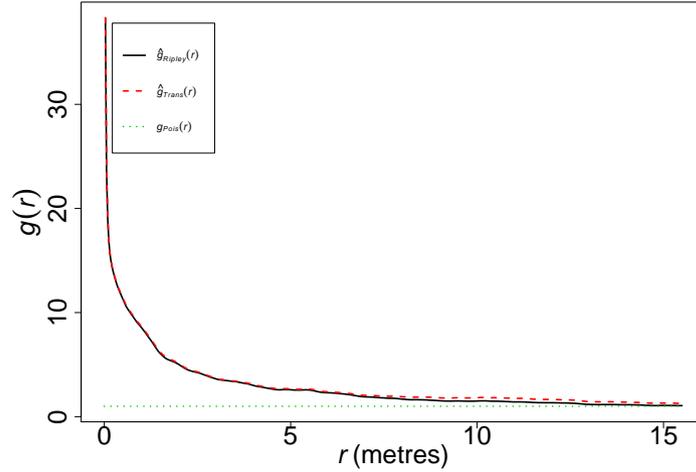
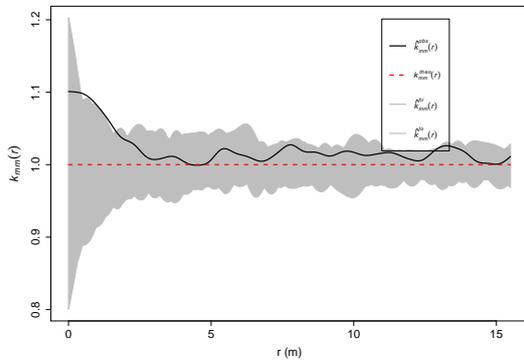
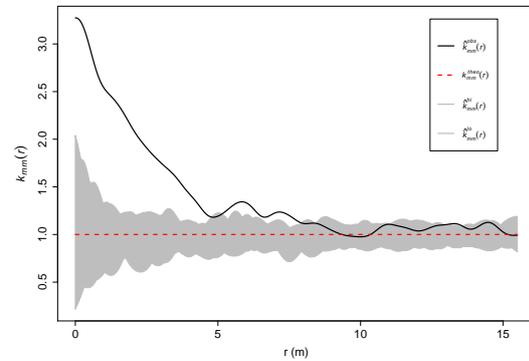


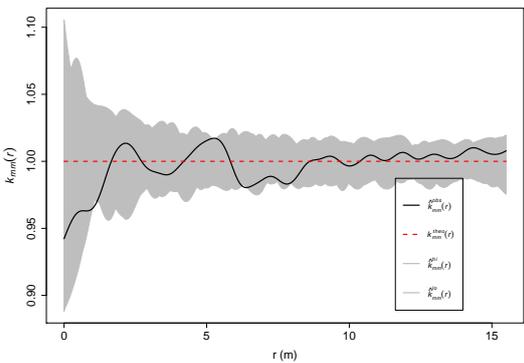
Figure 13: The pair correlation function.  $g_{\text{trans}}$  is an edge-correction.



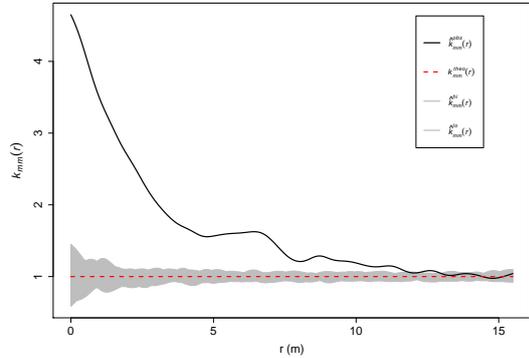
(a) Mark equality/correlation of real nests.



(b) Mark equality/correlation of day of the nests.



(c) Mark equality/correlation of parasitised nests.



(d) Mark equality/correlation of microhabitat model prediction of the nests.

Figure 14: Mark equality plots. Simulation envelopes are based on 99 Monte Carlo simulations with the assumption of complete spatial randomness.

## 4.4 The population as a network

All code and R-scripts can be found in Appendix 3.

### 4.4.1 Distance kernel between two subsequent nests

The Python script to filter the distances between subsequent nests of each individual out of the distance matrix, can be found in Appendix 3a. The logic here was to first assign all the nests to an individual and

to secondly couple a distance to subsequent nests (chronologically the next nest) from one individual.

145 distances for real nests were obtained with the Python script. The histogram of distances of subsequent real nests, see figure 15, has a lot of small distances, but also a fat tail. The distances are grouped by 2 m. The **mean distance is 11.72 m**, the **median is 4.51 m**. This is indicative of the skewness of a kernel with a fat tail.

Taking all (not only real nests) into account, gave 377 distances. It gave qualitatively the same kind of histogram (a lot of low distances but a fat tail, not shown).

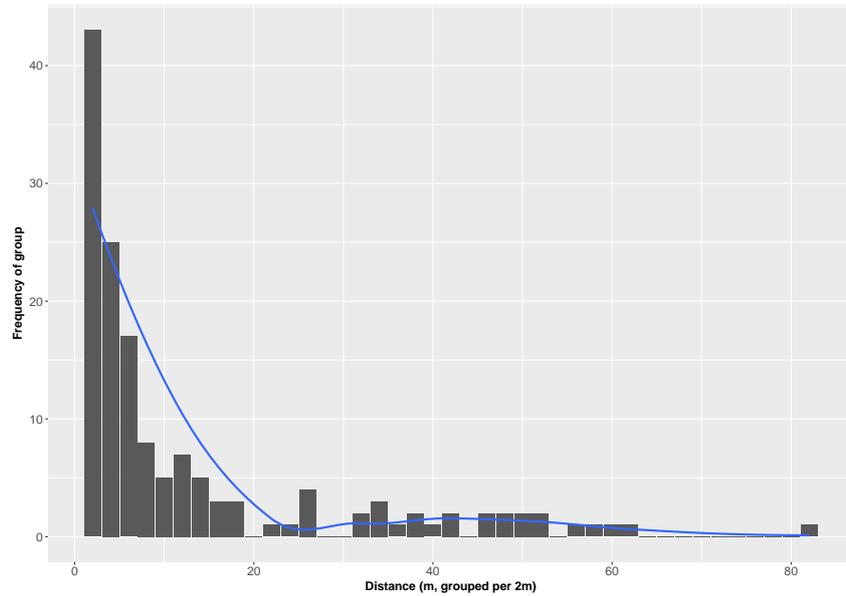


Figure 15: Histogram of distances of subsequent nests of each individual.

#### 4.4.2 Cluster analysis

The result of the cluster analysis is shown in figure 16, where the nests that belong to a cluster have the same colour. The analysis was carried out separately for all and only real nests, they gave roughly the same results apart from a few nests that were assigned to an adjacent cluster. Based on the results of the spatial pattern analysis, see section 4.3.2, the radius of the clusters is considered between 3-5 m. Considering this, together with the visualisation of the clusters (figure 16) and density plots (figure 11), the amount of clusters was chosen as 11.

The analysis assigned a number to each cluster, which were coupled to the nests in the Microsoft Access database, to use the cluster numbers in further analysis. The R-code and complete analysis can be found in Appendix 3b.

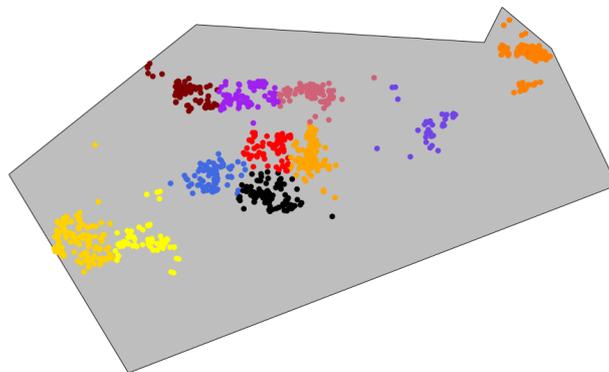


Figure 16: Result of cluster analysis. All nests belonging to a cluster have the same colour. All nests are included.

### 4.4.3 Social network

#### Full network

How the network data were constructed can be found in Appendix 3c. Complete code and detailed output of the network analysis can be found in Appendix 3e. The metadata of the nodes are amount of real nests, amount of parasites, relative parasite density (parasites/nests) and microhabitat suitability. These were filtered from the data with a Python code, see Appendix 3d, the predictions were added to the data file during the microhabitat model analysis.

Figure 18 shows the total network (undirected and directed) in the correct spatial configuration. The size of the nodes is proportional to the size of the cluster (amount of real nests) and the redness indicates relative parasite density. The thickness of the links is proportional to the amount of individual links that were present in a non-aggregated network.

Figure 19 shows the undirected network; the size of the nodes indicates the amount of real nests and the colour indicates microhabitat model prediction (blue low probability, yellow high probability). Cluster 1 and 9 have many incoming links, high relative parasite density and low microhabitat suitability.

The correlations between the four properties of the nodes (amount of real nests, amount of parasitised nests, relative parasite density and microhabitat suitability) were calculated (see Appendix 3e) and indicated that only size and amount of parasites are correlated: thus size, relative density of parasites and microhabitat suitability are used for further analysis. Microhabitat suitability was slightly ( $\rho = -0.57$ ) correlated with relative parasite density. This might indicate that clusters with suboptimal habitat also have higher relative parasite density and can be considered as ‘marginal clusters’ (see cluster 1, 5 and 9 in figures 18, 19)

**Descriptive network metrics** are given in table 2. They show that the amount of internal loops is 56% of all loops (significantly higher than 50%; exact one-sided binomial test, with the null-hypothesis that the internal loops are half of the links, returned a  $p = 0.092$ ), which is consistent with the distance kernel of subsequent nests. The network is quite dense: 38% (for the directed) and 62% (for the undirected network) of all the possible links are present and the clustering coefficient is 64%. 38% of the links are reciprocal: the links between clusters seem directional. The assortativity shows that clusters slightly tend to connect to clusters with different size, relative parasite density and prediction of habitat suitability, which might be a mere depiction of the high variation in properties of the clusters, seen in the network representations of size, relative parasite density and habitat suitability prediction (figures 18 and 19).

Table 3 gives an overview of the relationships between the **node metrics** and node properties (size, relative parasite density and microhabitat suitability).

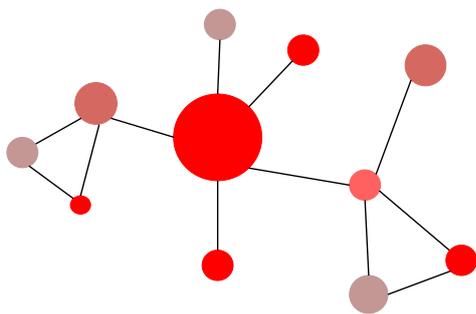


Figure 17: Schematic representation of the interpretation of the positive relation between degree and centrality with size and the negative relation between clustering coefficient with size. Big nodes have many links, but the ones they are linked to are not that much linked to each other. Smaller nodes do not have many links, but they do form clustering triads.

The three important trends are coloured in red, purple and yellow. The red colour highlights the positive trend of degree, strength & centrality with the size of the nodes. The amount of links coming to a node (degree and strength), and the centrality of a node in the network are positively correlated with the size of the node: bigger nodes have more links and are more central in the network. On the other hand, the blue highlights the negative correlation between clustering coefficient and the size of the nodes. Bigger nodes tend to be less clustered (form triads). Those two trends seem contradictory, but it can be understood when bigger clusters are considered *star-shaped* in the network, see scheme in figure 17. A lot of nodes are linked to them, but those nodes are not that much linked to each other. Smaller nodes do not have many links, but they do form clustering triads.

The yellow highlights the positive relation between authorities (which have many incoming links) and relative parasite density. It seems contradictory that individuals would go more to nodes with higher relative parasite density, but the causality is not known; it indicates a correlation. It could be the other way around: parasites could follow the many incoming individuals in a node. A non-significant trend

between size and hubs is also present: bigger clusters have more incoming links. Here also the causation is not known (are the nodes attractive because they are so big or are they becoming big because of the many incoming links?).

These relationships are visualised in figure 20. 20a shows the eigenvector centrality: the size of the nodes is proportional to the centrality-measure and the bright greenness indicates size of the clusters. Figure 20b gives a visualisation of the authorities: the size is proportional to the authority-measure (authorities have many incoming links) and the redness indicates relative parasite density. Figure 20c show the hubs: the size is proportional to the hubs measure (hubs have many outgoing links). Nodes that have similar sizes in the hubs and authority plot have as many incoming as outgoing links, but if the size is larger in one of the two figures, the node is predominantly an authority (many incoming links) or a hub (many outgoing links). Cluster 1, 6 and 9 are authorities; cluster 2 and 11 are hubs.

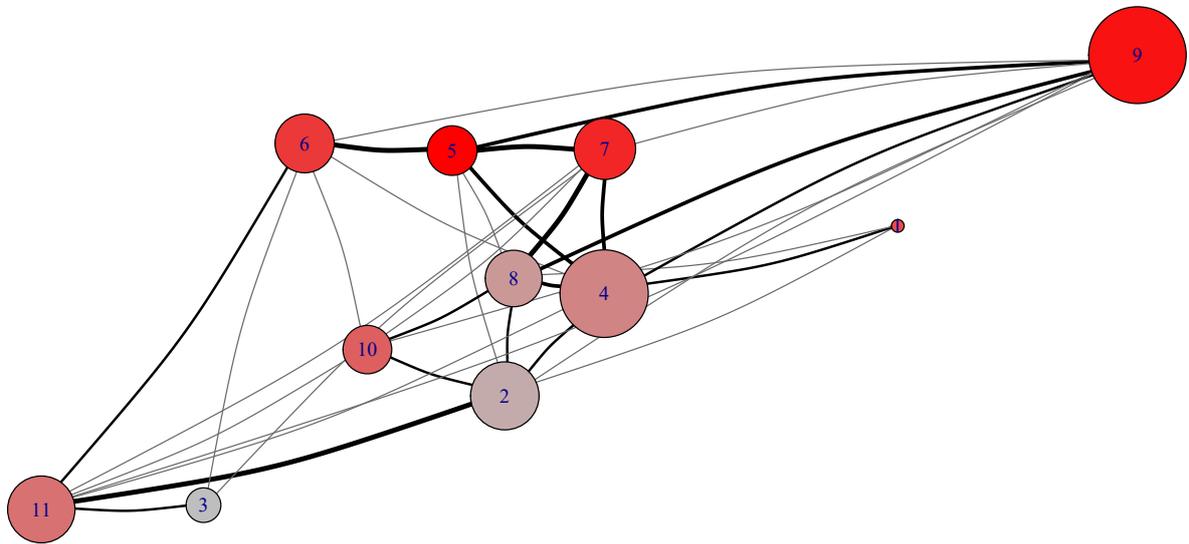
The Mantel test between the weight matrix of the amount of links between the different clusters and the Euclidian distance matrix, gave a correlation of  $-0.27$  with  $p = 0.956$ . Thus, the distance matrix and amount of links matrix are not correlated. This means that the amount of links is not correlated with the distance between nodes. This is not contradictory to the distance kernel (figure 15), because the tendency to build a new nest in the same cluster as the previous nest is not confined by mobility, it is steered by other processes. Thus, if the wasps go outside a cluster to build a next nest, the distance to another cluster (at the scale considered) is not important.

Table 2: Network metrics.

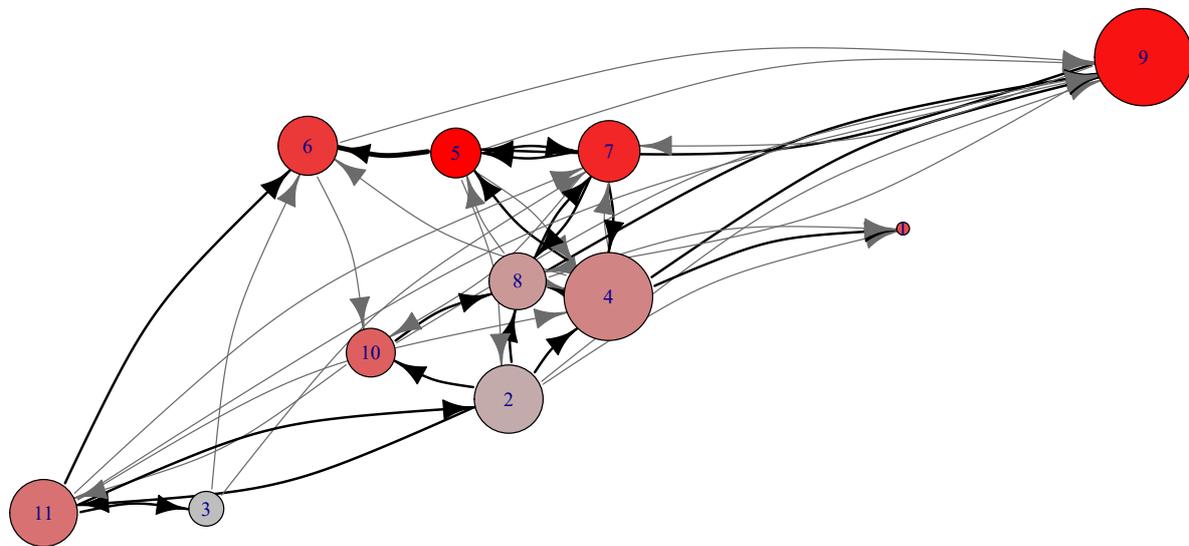
<i>Network metric</i>	<i>Undirected</i>	<i>Directed</i>
Internal loops	55.9%	
Density/Connectance	61.8%	38.2%
Reciprocity	NA	38.1%
Assortativity	Size $-0.14$	Size $-0.07$
	Par $-0.12$	Par $-0.19$
	Pred $-0.22$	Pred $-0.20$
Transitivity or Clustering Coefficient	64.4%	

Table 3: Relationships between node metrics and size, relative parasite density (par) and microhabitat model prediction (pred) of the nodes.  $\oplus$  indicates positive and  $\ominus$  indicates negative significant relation ( $p < 0.05$ ),  $\backslash$  indicates no relation and - or + indicates negative or positive trend ( $0.05 < p < 0.10$ ).

<i>Node metric</i>	<i>Undirected</i>			<i>Directed</i>		
	<i>size</i>	<i>par</i>	<i>pred</i>	<i>size</i>	<i>par</i>	<i>pred</i>
Strength (weighted)	$\oplus$	$\backslash$	$\backslash$	$\oplus$	$\backslash$	$\backslash$
Degree (binary)	$\oplus$	$\backslash$	$\backslash$	$\oplus$	$\backslash$	$\backslash$
Eigenvector centrality	$\oplus$	$\backslash$	$\backslash$	$\backslash$	$\oplus$	$\backslash$
Betweenness centrality	$\oplus$	$\backslash$	$\backslash$	$\backslash$	$\backslash$	$\backslash$
Closeness centrality	$\oplus$	$\backslash$	$\backslash$	$\oplus$	$\backslash$	$\backslash$
Transitivity or Clustering Coefficient	$-$	$\backslash$	$\backslash$	$\ominus$	$\backslash$	$\backslash$
Hubs	NA	NA	NA	+	$\backslash$	$\backslash$
Authorities	NA	NA	NA	$\backslash$	$\oplus$	$\backslash$



(a) Undirected network



(b) Directed network

Figure 18: Visualisation of the network, undirected and directed. Grey links are links that appeared only once; thickness of links is proportional to the amount of summed individual links; size of the nodes is proportional to the amount of real nests; colour of nodes is proportional to relative parasite density (red = high relative density). The number of each cluster is written inside the circle of the node.



## Temporal network

The R-code and complete analysis can be found in Appendix 3f.

Figure 21 shows the networks for the three time frames. The size and relative parasite density of the different nodes is quite variable between the timeframes. Which links are present between which nodes, also changes with time. Figure 23 shows the change in degree and eigenvector centrality for the different nodes for the three timeframes. These are also variable with timeframe.

Figure 22 gives the node properties and metrics averaged over all nodes for the three timeframes. This indicates the differences in metrics between the three timeframes. The size (a) and amount of parasites (b) are higher for timeframe 2 (they are correlated with each other), which make the relative parasite density (c) staying quite constant for the three timeframes. The degree (d) and centrality (e) of nodes become more divergent with time: nodes have more variable and extreme degree or centrality with time. Nodes have a higher clustering coefficient in timeframe 2. This is again indicative of the variability in time: there is not a single continuous process that gives rise to the resulting, full network and its properties.

The variability (CV) of the node properties and metrics for the clusters (mean size, amount of parasites, relative parasite density, degree, eigenvector centrality and transitivity of the clusters) was not related with the habitat suitability. The variability between the timeframes cannot be linked to the node properties and metrics.

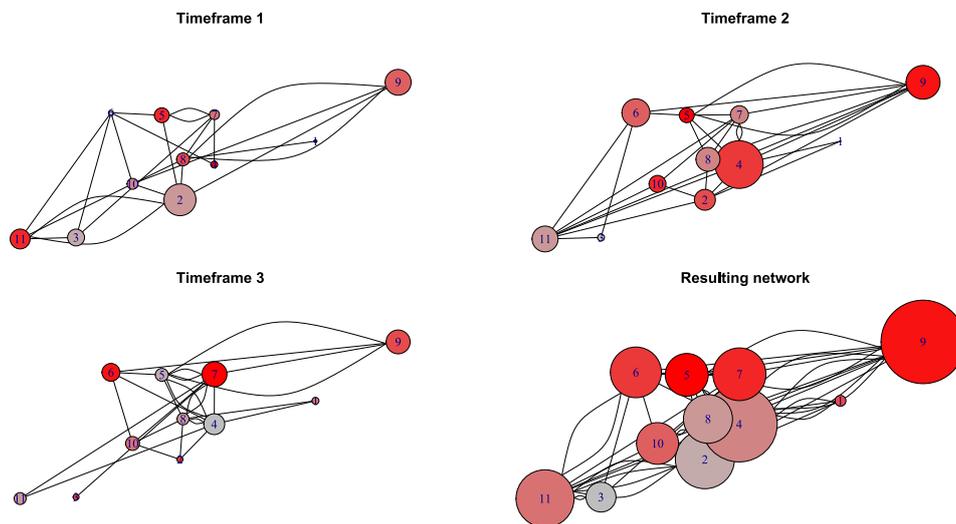


Figure 21: Three time-aggregated networks: timeframe 1 is sampling days 1-10, timeframe 2 is sampling days 11-20 and timeframe 3 is sampling days 21-30. Resulting network with same scale is given too. Size is proportional to the amount of real nests in the cluster and redness indicates relative parasite density of the nodes.

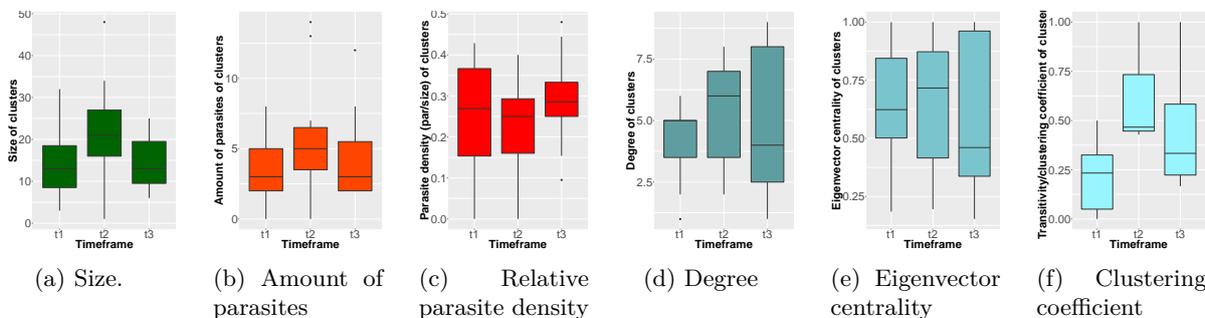
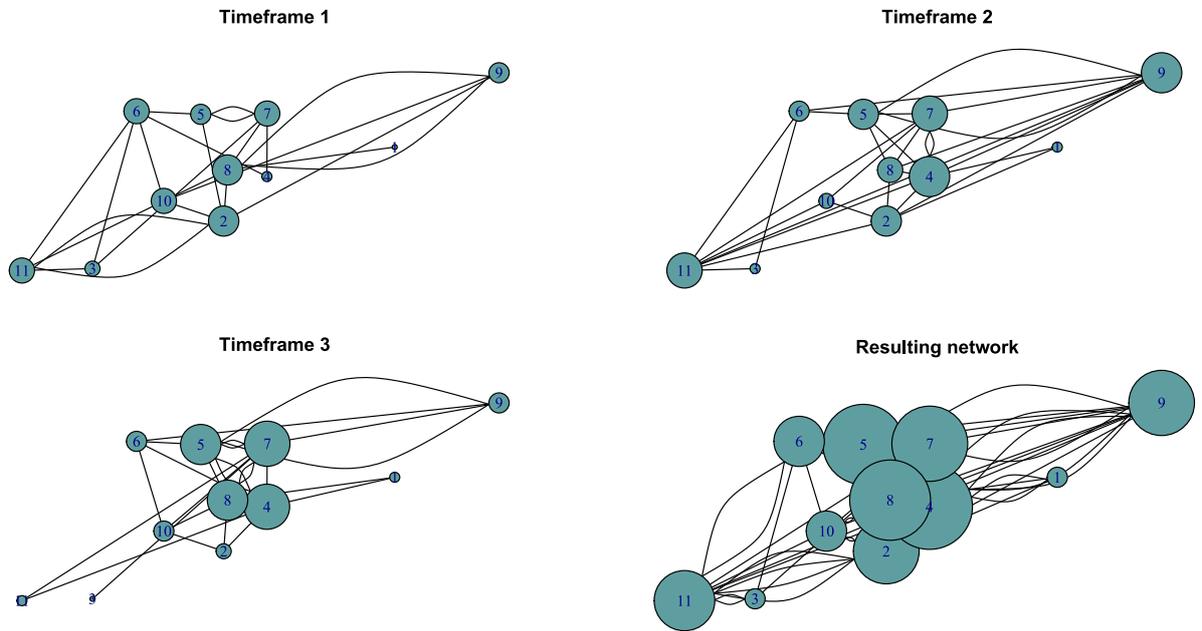
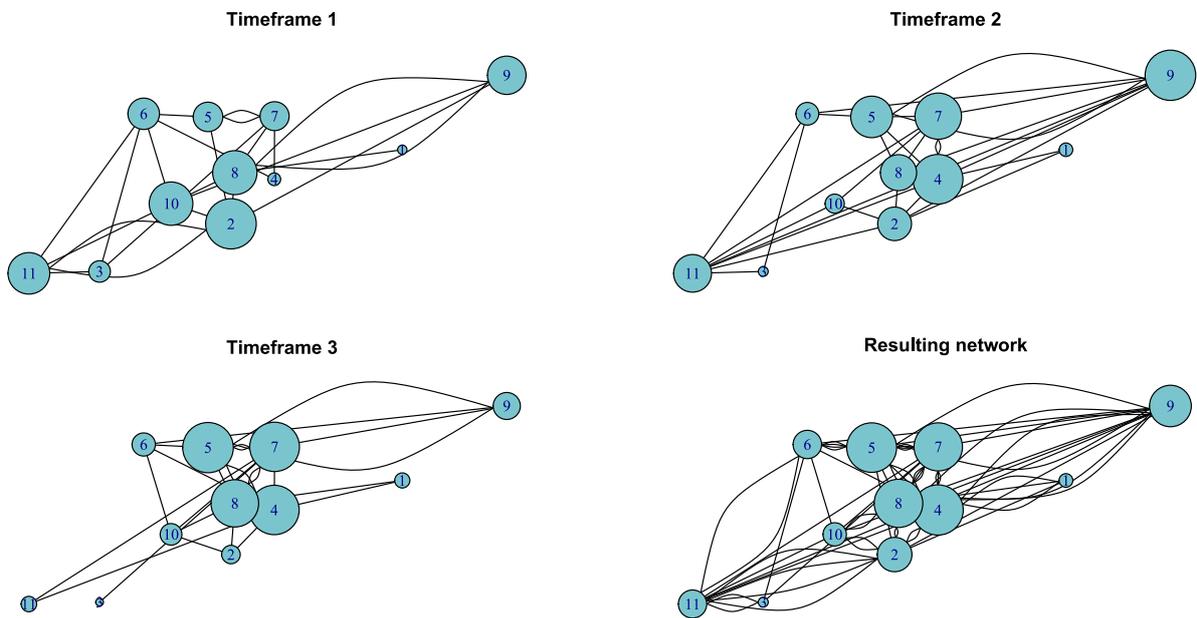


Figure 22: Boxplots of node properties and metrics of the clusters for the three timeframes (middle line is median, upper Q3 and lower Q1 quartiles are limits of the box, whiskers are maximum values excluding outliers, small dots are outliers ( $>Q3 + 1.5 \cdot \text{boxlength}$  or  $<Q1 - 1.5 \cdot \text{boxlength}$ )).



(a) Visualisation of the degree for the different timeframes. Size is proportional to degree.



(b) Visualisation of the eigenvector centrality for the different timeframes. Size is proportional to eigenvector centrality.

Figure 23: Visualisation of degree and eigenvector centrality for the three timeframes.

## 4.5 The causes of clustering

The R-code and complete analysis can be found in Appendix 4.

### 4.5.1 Microhabitat model

Figure 24 shows the map of the predicted microhabitat suitability for the complete study area. The nests are plotted as yellow dots. Figure 25 shows a boxplot of the predicted microhabitat suitability of the nests of the different clusters. A one-way ANOVA demonstrates that the microhabitat suitability differs between the cluster ( $p = 3.75 \cdot 10^{-7}$ ). The most optimal habitat is filled up with nests but other suitable habitat is not saturated. The northeast corner (cluster 9) is situated in suboptimal habitat.

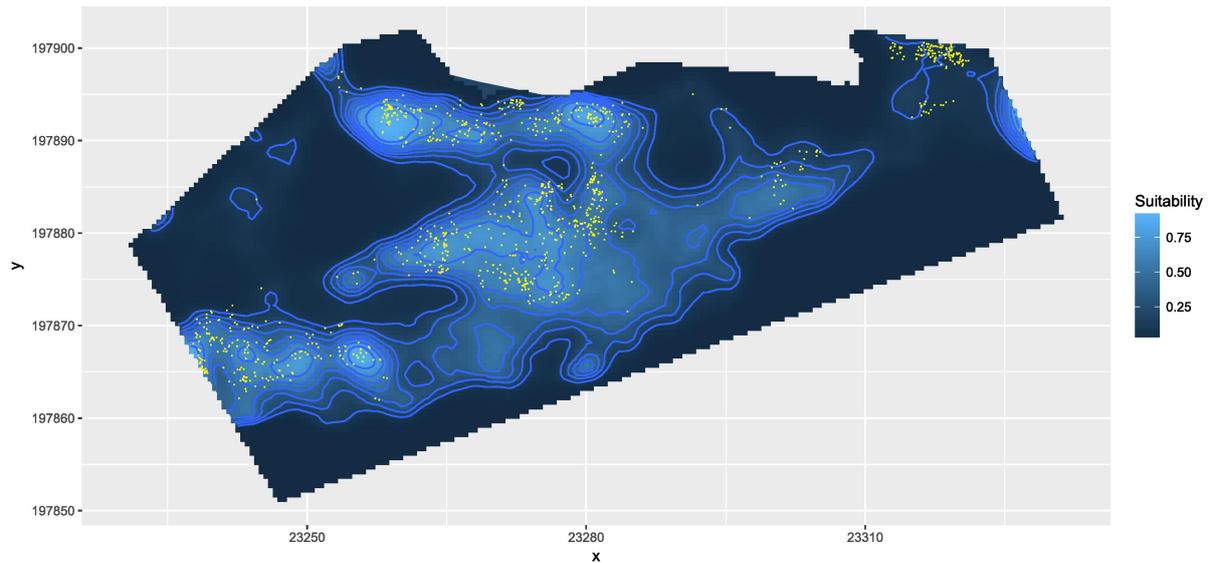


Figure 24: Microhabitat model suitability predictions for the complete study area. Nests are plotted as yellow dots. The lighter the blue, the higher the probability of presence because the more suitable the microhabitat is for nests of *B. rostrata*.

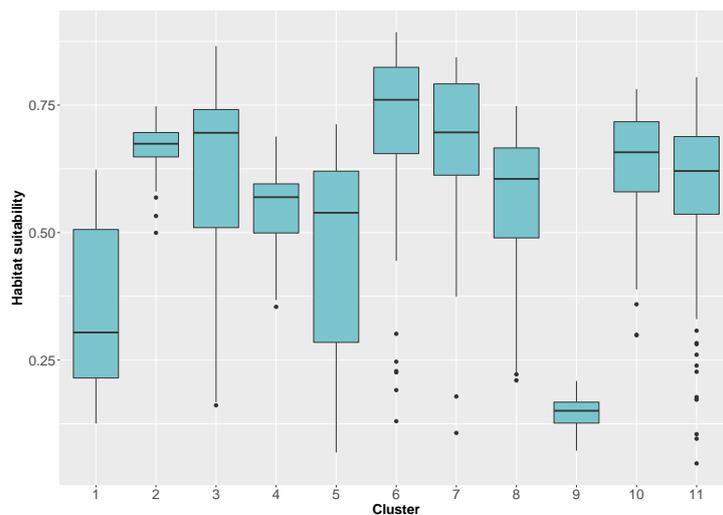
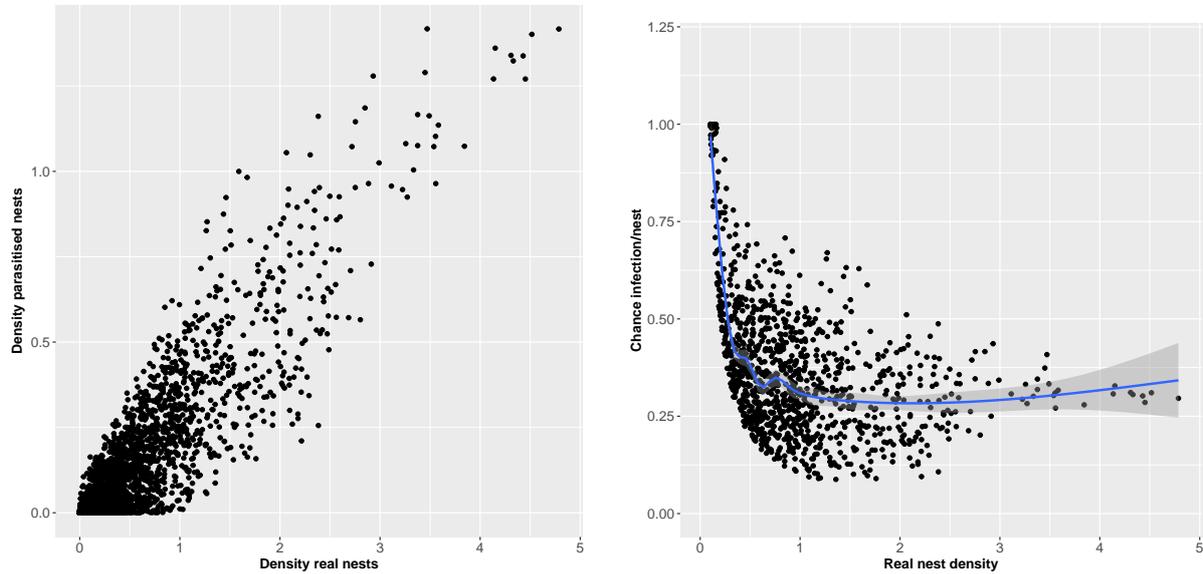


Figure 25: Boxplots of microhabitat suitability of the nests in the different clusters (middle line is median, upper Q3 and lower Q1 quartiles are limits of the box, whiskers are maximum values excluding outliers, small dots are outliers ( $>Q3 + 1.5 \cdot \text{boxlength}$  or  $<Q1 - 1.5 \cdot \text{boxlength}$ )).

### 4.5.2 Parasite density

Figure 26a shows the correlation between the density of parasitised nests and density of (real) nests. In figure 26b, the chance of infection per nest (formula 4) is plotted against nest density. The line is not a fitted function, but a smoothed line, provided in `ggplot` (Wickham, 2009) in R. Although parasite density increases with nest density, the individual parasite density decreases with increasing nest density.



(a) Correlation between density of parasitised nests and density of nests.

(b) Chance of infection per nest in function of the nest density. The line is a smoothed line, not a fitted function.

Figure 26: Correlation plot and chance of infection.

## 5 Discussion

Both a bottom-up and top-down regulator have been identified for the clustering of nests of *Bembix rostrata* in the results (section 4): the microhabitat suitability predictions and the selfish herd mechanism. The findings will be elaborately discussed in this section. First, remarks on the data collection will be explained. Second, all the analyses will be discussed: the microhabitat model, the CMRC analysis and spatial pattern analysis, the network analysis, including the distance kernel between subsequent nests, ending with the causes of clustering and the implications. Finally, implications for conservation and possible future research are considered.

### 5.1 Data collection

#### 5.1.1 Tagging

2.1% of the tagged individuals were lost due to glue spilled on their wings. This is considered as a tolerated amount of casualties.

At least 0.9-1.6% of tagged animals lost their tag. A relative amount that is considered having a minor influence on CMRC-estimations or sampling of consecutive nests of individuals, even if the real number of lost tags would be much higher than the detected losses (this detection is considered relatively high).

#### 5.1.2 Real nests

The criteria to label a nest as a real nest are listed in section 3.1. This was done because *B. rostrata* shows test-digging behaviour before it makes an actual nest. The nests that are not labelled as real nests are not by definition test-digging nests, they are ‘uncertain nests’. The real nests are a subsampling of all the nests which are certainly real nests and should be considered this way. The relevance of this regarding analysis will be addressed in the related sections.

#### 5.1.3 Anti-parasite behaviour

Several times, *B. rostrata* individuals were seen dragging prey out of their burrows. Cell cleaning has been reported for other Bembicini-species (Evans & O’Neill, 2007), but the behaviour observed for these *B. rostrata* individuals does not match the description for cell cleaning. As prey that are seen being dragged from a burrow cannot be contaminated or parasitised when laying next to the nest, they were infected beforehand. I considered this prey-dragging behaviour as an anti-parasite behaviour. Thus, the 54 samples of prey were taken as an extra indication for parasitised nests, apart from 98 regular observed parasite infections at nests. More detailed research would be needed to determine if this behaviour is an exclusive anti-parasite behaviour or if it can have other causes.

#### 5.1.4 Prey

The prey caught by *B. rostrata* are known to be flies (Diptera), mainly from the families Syrphidae, Muscidae, Calliphoridae, Sarcophagidae and Tabanidae (Evans & O’Neill, 2007; Peeters, 2008). However, this master project **reports for the first time that *B. rostrata* catches a species of the order Hymenoptera**, namely *Bombus terrestris*. Other *Bembix*-species are known to exploit a wide variety of prey-species, including Hymenoptera, Odonata and Lepidoptera (Evans & O’Neill, 2007). There are several evolutionary shifts in prey use within the genus but also several *Bembix*-species show opportunistic, plastic behaviour when novel prey species appear in the habitat. Taking this into account, it is probably not that surprising to see *B. rostrata* which carries a bumblebee, it confirms the genus *Bembix* its flexible behaviour to catch suitable, abundant flying insects, present in the individual its territory.

### 5.2 Microhabitatmodel

Nest site locations and preferences (and the influence on the breeding performance) have been elaborately investigated in birds, taking the specific nest site (Slagsvold, 1987; Belles-Isles & Picman, 1986), microclimate (Lloyd & Martin, 2004), habitat (Storch, 1991; Mahon *et al.*, 2007; Löhmus, 2003) and/or

landscape context (Martínez *et al.*, 2003; Sánchez-Zapata & Calvo, 1999) into account. In Hymenoptera, there has also been a lot of attention to the characteristics of preferred nest sites (Seeley & Morse, 1978; Wuellner, 1999), taking the landscape context (Frankie *et al.*, 1988; Svensson *et al.*, 2000; Kells & Goulson, 2003) and scale (Westphal *et al.*, 2006) into account. Digger wasp (Sphecidae and Crabronidae) nesting behaviour has been studied in detail (Tinbergen, 1951; Brockmann & Dawkins, 1979). Their nest site characteristics have been related to abiotic (soil temperature, hardness, moisture and slope) and vegetation structure (Wcislo, 1984; Vanommeslaeghe, 2015). The characterisation was done with point or quadrant measurements. This master project however, characterised abiotic factors (slope and insolation) and vegetation (NDVI) at a very high resolution, which could not be accomplished with other field measurements than remote sensing. The microlandscape around the nest could be taken into account and simultaneously the habitat characteristics of the complete study area were considered. This kind of integrative approaches have the potential to link micro-scale landscape with larger scale habitat properties and link ecological processes important at both scales.

### 5.2.1 Presence-absence model

The performance of the model is looked at through the Area Under Curve (AUC), see section 3.2.1 for definition. This measure is sometimes highly criticised (Lobo *et al.*, 2008). But these critics are mainly for predictive distribution models of presence-only data. AUC calculations from ROC plots are considered effective measures for model performance of presence-absence data (Fielding & Bell, 1997; Allouche *et al.*, 2006). It is recommended to use the AUC mainly to compare models and care should be taking when using it as a performance measure or measure for predictive accuracy of a model (Manel *et al.*, 2001). This seems only of importance when the model would be projected onto other areas – where a priori presences are not known – outside the study area. But in the case of this microhabitat model, the AUC can and is used as a predictive accuracy of the nests in the study area, because the model was build on the same data: about 9% of the nests is not well predicted by the model.

Models with scales between 2 and 3 m performed equally good (see figure 7d). The **2 m scale** is chosen to be included in the full model. The influence of insolation and vegetation up to this radius around the nest is of importance for nest presence or habitat selection for nestbuilding in *B. rostrata*. This ‘model scale’ can be interpreted as the ‘selection scale’ of nests in *B. rostrata*.

The presence of the nests, based on the linear full model with an exponential autocorrelation function at a scale of 2 m, is predicted by formula (6) with both insolation and NDVI highly significant factors and an AUC = 0.91:

$$\text{Probability of presence} = \frac{e^{0.000045 \cdot \text{insolation} - 27.28477 \cdot \text{NDVI}}}{1 + e^{0.000045 \cdot \text{insolation} - 27.28477 \cdot \text{NDVI}}} \quad (6)$$

The nests have higher insolation and less vegetation (lower NDVI) than random points in the study area.

### 5.2.2 Real nests

Microhabitat predictions of real nests is not different from uncertain nests. This is as expected, because they are considered a random subset of all the nests. Although in other analysis, like the network analysis, only real nests are considered because differences could be apparent then. Including uncertain nests in such analysis are considered as extra noise.

Following up in detail a few quadrants during the season would help defining actual real nests and actual test-digging nests. This method, for instance, would be better to test the hypothesis that real nests have different microhabitat values than test-digging nests. The method used in this master project is less appropriate: the real nests are ‘certain nests’ and the other nests ‘uncertain nests’. There is too much variation on the uncertain nests. The real nests are the ‘certain subset’ of the nests for other analyses.

### 5.2.3 Parasitised nests

Microhabitat predictions of parasitised nests is not different from non-parasitised nests. This is as expected, because the microhabitat is not expected to influence parasite-presence. They are expected to be

randomly distributed over the nests.

The three models given above (presence-absence, real nests and parasitised nests) could be made stronger by iterations of different random subsets of 70% training data and 30% test data. But the runtime of one iteration for the presence-absence model was already  $\pm 2$  hours, thus this was not considered a priority in this master project. Certainly because for the presence/absence model, the p-values are very low and the AUC very high. It would be the most interesting to do this for the real and parasitised model. For the real model, when different training data sets were run manually, the p-value indicated once or twice a possible trend (of little less warm places for real nests). Looking at the frequency of these ‘possible trends’ would make the conclusion of the model for real and parasitised nests stronger.

#### 5.2.4 Temporal model

The prediction values are not dependent of day (see figure 8): the hypothesis that the microhabitat suitability changes with time is not correct. There is no evidence that the wasps have to settle in more suboptimal habitat later in the season.

What also can be seen in figure 8 is that mainly **cluster 9** is poorly predicted. This is further discussed in section 5.5.1

### 5.3 Population estimation and spatial pattern analysis

#### 5.3.1 CMRC: population estimation

Both the manually performed Jolly Seber estimation and the one performed in Program MARK, gave values in agreement with each other.

A population size of around 1000 female individuals of *Bembix rostrata* is quite a big amount, compared to an estimation made by Vanommeslaeghe (2015), 2 years earlier in a population 2.6 km away from the focal study area, of around 70 female individuals.

The survival probability  $\phi$  is around 1. This confirms the hypothesis that females do not die massively before the end of the season.

The highly variable and fluctuating joining numbers  $B_i$  and the 43% of the tagged females that were never seen again, suggest that a lot of the wasps probably migrated from and to adjacent areas.

A Jolly-Seber population estimation was carried out because the assumption was made of an open population. Because of the life-history of *B. rostrata*, and because  $\phi = 1$ , little or no females were born or have died during the study period.

#### 5.3.2 Point Pattern Analysis

The 2D-statistic Ripley’s K (figure 12) indicates **spatial clustering** up to 30 m, thus clustering up to a radius of 15 m. But the plots of the nests (figure 9 and 10a), density plots (figure 11) and the pair correlation (figure 13) suggest a smaller clustering within that 15 m radius of about 3-5 m.

The non-random clustering of nests made around the same day (figures 10b and 14b) suggest **temporal clustering**. The clustering of the microhabitat prediction (figures 10e and 14d) suggest that different clusters or groups of nests have different microhabitat suitability.

### 5.4 The population as a network

Integrating social network analysis from sociology into biological research has started about 15 years ago, which mainly focused on primates (Sih *et al.*, 2009). Network analysis has an enormous potential in studying non-human animal systems (Krause *et al.*, 2009). It can help exploring informational flows (Aplin *et al.*, 2015), parasite transmission (Poulin, 2010; Godfrey *et al.*, 2010), social group dynamics (Kerth *et al.*, 2006) and patterns of sociality (Chaverri, 2010). Social network analysis on insects have been mainly focused on eusocial colony behaviour (Naug, 2008, 2009). Understanding informational flow in eusocial insects is important in trying to understand how collective behaviour works and has evolved (Sumpter, 2006; Fewell, 2003). Applying network analysis in a gregarious nesting, non-eusocial insect is has not been done before. Social networks can be defined by any kind of ‘node’ (Sih *et al.*, 2009) (individuals, group of individuals, territory etc.). In this master project, the clusters or aggregates of individuals of *Bembix rostrata* are defined as nodes, as these are the entities of interest to investigate dynamics and structure of the system. Looking at one-by-one individual interactions would not be feasible

or informative in this system. Another reason to look at the aggregates as nodes is that they are the informative, social factors regarding the mechanisms behind clustering and conspecific attraction (see discussion section 5.5). As the evolution of eusociality is still not completely understood, it is interesting and helpful to look at social systems, behavioural and ecological processes of mating and (gregarious) nesting in all clades of Hymenoptera (Lin & Michener, 1972; Brockmann & Dawkins, 1979; Hughes *et al.*, 2008).

#### 5.4.1 Distance kernel between two subsequent nests

The histogram (figure 15) shows that most of the distances between subsequent nests is between 0 and 15 m. Those nests are most likely build in the same cluster. But the histogram also has a fat tail: several nests are build up to 60 m from the previous nest, one even 82 m. Larsson & Tengö (1989) reported a maximum of 40 m, but also reported ‘long-distance’ dispersals of at least 650 m. The tail of the histogram is probably prolonged, but these rare long-distance dispersals are very hard to detect and the method of this master project was not appropriate for studying these. The method used by Vanommeslaeghe (2015) is more suitable, because she had very low sample sizes, which also makes it harder to detect any long-distance dispersals. Because of the CMRC-data and the long-tail of the distance kernel, it would not be surprising if long-distance migration or exchange would be more ubiquitous than previously thought. A population genetics approach would even be better to investigate the amount of exchange between populations. It could also be that males are the dispersive sex in the beginning of the season, which is not known for *B. rostrata* and a likely hypothesis in this mating system with promiscuity (Perrin & Mazalov, 2000), in which where males show high competition for females (Schöne & Tengö, 1981).

#### 5.4.2 Cluster analysis

The amount of clusters was chosen as 11, taking into account the radius of 3-5 m of a cluster and the visualisation of these clusters on a map (figure 16). The clusters seem to overlap each other, but the 3D-density plots (figures 11a and 11b) show that the clusters are defined by a high density of nests in their centres.

#### 5.4.3 Social network

##### Full network

The internal loops make up the biggest amount of connections (significantly higher than 50%). Most of the next nests are made in the same cluster. This corresponds with the distance histogram (figure 15), showing that most of the nests are build close to the previous nest (<10 m). The network is quite a dense and clustered network because of the high connectance and transitivity. Only 38% of the links are reciprocal in the directed network, which might indicate that the links are indeed better defined as directed.

The relations between the node properties and node metrics indicate two trends. The first trend (red and purple in table 3) is the ‘star-shaped bigger nodes’ interpretation. Smaller nodes have more exchange with other smaller nodes and bigger nodes are the most central and form the connectance between the smaller nodes. The second trend (yellow in table 3) is that authorities (nodes with many incoming links) have higher relative parasite density. This is a correlation, but it is the most probable that parasites follow the incoming individuals. The other possibility, many incoming individuals when the relative parasite density or individual parasite density is high, would not be consistent with the selfish herd mechanism (see discussion 5.5). Thus, most likely, parasites follow incoming individuals.

Two striking nodes in the analysis are node 9 and node 1. These have low habitat suitability prediction (figure 19), high authority scores and high relative parasite density (figure 20). Together with cluster 5, they steer the correlation between low habitat probability and relative parasite density. Cluster 1 and 9 are very important in the relation between relative parasite density and authorities. These clusters can be interpreted as ‘marginal clusters’.

The Mantel test showed that there is no relation between the amount of nodes and links and the Euclidian distance. This is not contradictory with the distance histogram/kernel, it indicates that if a nest is not made in the same cluster (internal links), distance does not matter. Distance, at this spatial scale, is not a constraint when making a nest in another cluster.

## Temporal network

The time-aggregated networks showed that there is high variability in size, relative parasite density, links between clusters, degree and centrality with time. The average and range of these properties over all nodes for the three timeframes are also quite variable, indicating that there is no single continuous process resulting in the full network. The full network is the outcome or summation of a very variable network in time, with shifts in importance of nodes and their dynamics.

There was no difference in change of network metrics with habitat suitability: higher microhabitat predicted clusters are not less or more variable in node metrics with time.

Thus, there is indication of temporal clustering, but the pattern is attenuated. The temporal clustering is not a completely clear-cut pattern. The temporal network is not build out of discrete entities; it does not look very modular or forms the puzzle pieces of the complete network. This is also seen in the point pattern figure 10b. The mechanism behind the clustering (see discussion section 5.5.3) is partially attenuated. A hypothesised mechanism is spatial bet hedging (Cohen, 1966): digger wasps make their different nests in a variety of nest sites, which have different performances in an uncertain environment (uncertain environmental conditions, but also chance of e.g. trampling of the nests when they are all close to each other). This attenuates the dominant mechanism of density-dependent temporal (and spatial) clustering. Cluster 1 and 9 seem visually a bit more stable in the temporal networks (figures 21 and 23) and have a lot of incoming links (authorities), this might indicate their importance in the bet hedging strategy. As mentioned before, cluster 1 and 9 can be considered as ‘marginal clusters’, because they have low predicted habitat suitability and high parasite density. But these marginal clusters might be indicative and/or important for this bet hedging strategy. The other nodes are more regulated by the temporal clustering.

## 5.5 The causes of clustering and implications

### 5.5.1 Microhabitat suitability

The map showing the predicted habitat suitability for the complete study area, see figure 24, indicates that the best available microhabitat is taken, but a little less preferable microhabitat is not yet saturated.

The nests of cluster 9 (northeast corner) are poorly predicted, see also figures 8 and 19. This cluster has 130 nests (on average the clusters have 92 nests) which makes up 6.4% of the poorly predicted data (AUC = 0.89, thus 11% of the nests and random points is poorly predicted: more than half of it consists of cluster 9 nests). When looking at the data, cluster 9 has a much higher NDVI than the other clusters (0.17, while the other clusters all have an NDVI between 0.10 and 0.15; insolation was not irregular). Thus, cluster 9 has more vegetation ground cover and/or more vascular plants than mosses. This anomaly could be (1) due to edge effects (although on the left edge, those nests are well predicted), or (2) because another explanatory variable is missing (e.g. wind: it was noticed during the fieldwork that this site was a very sheltered part of the study area where less wind penetrated, which could compensate for the higher vegetation ground cover) or (3) the area is suboptimal but a mismatching cue generates high clustering, indicative of an ecological trap (Robertson & Hutto, 2006) (e.g. *B. rostrata* is known to make nests in places that were used in the past by conspecifics, but this site could have only recently become non-suitable) or (4) as explained in the discussion of the social network (5.4.3), it could be a ‘marginal’ cluster used by many individuals for spatial bet hedging.

The mean predictions of the nests per cluster are shown in figure 25. Cluster 9 is poorly predicted, but cluster 1 also seems poorly predicted. This is the cluster just southwest of the northeast corner. This cluster did not have much nests (33, the lowest of all clusters which had a mean of 92 nests), thus it is not strange it would be suboptimal microhabitat.

These predictions of microhabitat suitability are based on the correlation between nest presence and the microhabitat characteristics. As mentioned in the introduction and stated by Johnson (2007), abundance is an ambiguous measure for habitat quality. Following up the clusters subsequent years and assessing reproductive performance (by capturing the emerging individuals at the beginning of the season), could help assess a fitness measure to the cluster and its microhabitat. Investigating fitness and preferences are essential when one wants to detect ecological traps (Robertson & Hutto, 2006), resulting from a mismatch between social cue and habitat quality, but can be very difficult and time consuming to assess. Determining emergence rates and body conditions the next year could be feasible for *B. rostrata*, but true preference seems much harder to correctly investigate (through e.g. controlled choice experiments).

On the east edge of the map there is an apparently highly suitable site present. When compared to the aerial picture (figure 9), this was outside the grid. But high densities were not noticed there. That part outside the grid is grazed by cattle and horses, although very extensively, but trampling traces were conspicuous. This is known to be very detrimental for *B. rostrata* (Bonte, 2005). Something similar can be seen for 2 or 3 spots on the northwest side of the map, but there, a few digger wasps were noticed outside the grid: those sandy patches were more surrounded by dense shrubs that keep away grazers.

The microhabitat suitability map forms the bottom-up template for clustering: suitable microhabitat is constrained to specific patches in the study area.

### 5.5.2 Parasite density

Larsson (1986) described a selfish herd mechanism responsible for the clustering in *B. rostrata*. The parasite density increases with increasing nest density, but parasite density per individual (chance of infection per nest) decreases with increasing nest density. Larsson (1986) explained this as a selfish herd mechanism, or dilution effect: for an individual it is better to build a nest at a place with high nest density because it makes the relative/individual parasite density lower. This is highly likely to promote high density aggregates in *B. rostrata*, because it provides a selective advantage for females in dense aggregates. Thus, this supports the selfish herd hypothesis by Hamilton (1971).

Larsson (1986) used a different method than this master project: he used standard timeframes to make censuses of activity of both *B. rostrata* and the parasites present in quadrants. The method in this master project uses the resulting density pattern of all cumulative observations. The latter method is more indicative of the actual selective advantage of higher nest densities having an individual lower infection chance, while the first method is more indicative of the direct behavioural responses of *B. rostrata* on the presence of broodparasitic flies. *B. rostrata* reacts strongly on the presence of flies near her nests when the wasps try to enter with a prey (personal observations): the wasp darts away when a fly mounts her prey or keeps making zigzagging flight patterns near her nest, chased by several flies which try to keep track of her.

Larsson (1986) already pointed out that the kleptoparasite/broodparasite, predator and parasitoid he observed at a colony of *B. rostrata* could have an influence on the spatial pattern of the nest aggregates. This is also seen in the results of this master project (section 4.5.2), that the parasites have an influence on the nest densities: this makes up a top-down regulator for the spatial pattern of nestbuilding in *B. rostrata*.

It is currently unclear if the selfish herd mechanism is directly or indirectly caused by the presence of parasites. It could either be that the presence of parasites in a population triggers the wasps to nest in dense clusters or that the wasps display an a priori protective behaviour, which would be indicative of an evolutionary adaptation. Comparing densities of nests at parasite-free clusters (in the field or in the lab), with the microhabitat as covariate, could help disentangle this direct or indirect effect. There were no significant movement patterns found in the full and temporal cluster analysis, related to parasite density. Although they react heavily on parasites near their nests, they show no direct response to parasite density with spatial distribution in different clusters. Maybe the parasites are too ubiquitous in the whole study area. Then a selfish herd mechanism is the best strategy to ‘escape’ from the parasites. The indirect, evolutionary adaptation is the most likely mechanism, as it is consistently found in digger wasps (Larsson, 1986; Weislo, 1984).

### 5.5.3 Mechanism of clustering

Relative importance of both bottom-up and top-down controls have received a lot of attention, at community (Power, 1992; Burkepile & Hay, 2006; Dyer & Letourneau, 2003) and population level (Suryan *et al.*, 2006; Horswill *et al.*, 2016; Hunter *et al.*, 1997). Including heterogeneity, scale and spatial ecology in examining the relative importance of both regulators, has been a priority too (Hunter & Price, 1992; Gripenberg & Roslin, 2007; Seitz *et al.*, 2017). Investigating the influence of top-down or bottom-up controls on spatial patterning is looking at the reversed process (regulators have an effect on heterogeneity instead of the other way around). Such studies often only take one of the two effects into account, e.g. predator influences on spatial distribution of preys (Livingston *et al.*, 2017; Hammond *et al.*, 2007; Orrock *et al.*, 2008) or plant-soil interactions (Ettema & Wardle, 2002; Reinhart *et al.*, 2003). Both regulators have been taken into account on very large scales (Bailey *et al.*, 1996) and in some marine

systems (Ji *et al.*, 2013). But no solid literature was found about the simultaneous influence of top-down and bottom-up effects on spatial patterning at comparable scales.

Both bottom-up template and a top-down regulator of the spatial patterning of *B. rostrata* nests are presented in the two sections above: microhabitat suitability and a selfish herd mechanism driven by broodparasite density.

The relative strength of both regulators remains unclear. Individual based modelling, or other mechanistic or process-oriented modelling, might help to elucidate the relative importance of these effects.

It might seem in the study that the bottom-up regulator is the most important, but that is not necessarily the case. Wcislo (1984) tested the influence of 3 abiotic factors (soil temperature, hardness and moisture) on clustering of the digger wasp *Crabro cribrellifer* (Crabronidae). These factors did not have any significant influence on the non-random distribution of the aggregates (or it could not be detected by their sampling method). Thus, there was no evidence for a bottom-up effect. But he found a correlation between activity of the digger wasp and the broodparasite *Metopia campestris* (Sarcophagidae). He excavated nests and found that the probability of parasitism decreased with increasing nest density, which points to a selective advantage for females in dense clusters, and thus supports the selfish herd mechanism of Hamilton (1971), the main process that steers clustering in that system. The findings of Wcislo (1984) and of the conducted study suggest that this top-down regulator might be of considerable importance in the formation of nest-aggregates in digger wasps (Crabronidae).

Both the point pattern and social network analysis suggest clustering in space and time of the nest aggregates of *B. rostrata*. The top-down regulator of parasites and the high preference for nest sites with conspecifics (Peeters, 2008) point towards density-dependent movement between clusters. Those two arguments (non-stationary spatial pattern and density-dependent movement) promote the **phase separation principle** for the self-organisation of nests of *B. rostrata*. This could be tested with a pattern-oriented modelling framework: to investigate a phase separation mechanism as explanatory ecological process for nest aggregates in *B. rostrata*, without taking into account the mechanisms behind the density-dependent movement (microhabitat and selfish herd mechanism).

The density-dependent movement, social network analysis, as well as the CMRC-analysis, point out that big and dense aggregates have a crucial role in the dynamics of the population.

#### 5.5.4 The scale of philopatry

*Bembix rostrata* is known to be philopatric (Peeters, 2008), the tendency to return or stay in its home area. A detailed study performed on another digger wasp *Cerceris arenaria* (Crabronidae) by Polidori *et al.* (2006) showed its highly philopatric behaviour: they re-use the nests they emerged from, choose consecutive nests very close to their first nest and rarely dig new nests. They show philopatry at a short scale. However, this master project showed that *B. rostrata* also tends to stay in the neighbourhood of its first nest (see distance histogram, figure 15); the amount of internal links in the network analysis (table 2). But as the network analysis (section 5.4.3) and CMRC-data (section 5.3.1) indicate, *B. rostrata* is much more mobile when it makes subsequent nests. Together with the mechanism found here, density-dependent selfish herd, this indicates that individuals do migrate to and from adjacent areas, but they are bound to the presence of nests of conspecifics. This could be seen as philopatry at a larger, landscape scale. Both scales give the same pattern and behaviour from a conservationist's perspective (the species rarely (re)colonise new available habitat), but the biological processes are different.

## 5.6 Implications for conservation

It could be that **connectivity** is as big of a constraint to *B. rostrata* as is the habitat quality. Long-distance dispersal might be not as limiting as previously thought (see discussion section 5.4.1), but the biggest constraint is probably the habitat quality and more importantly, the **presence of conspecifics**. As the density-dependent clustering is steered by parasites (either directly by their presence or indirectly by a priori protective behaviour), growth of low density populations are inhibited by a possible higher parasitic load. This is indicative of an Allee effect (Stephens *et al.*, 1999), which inhibits growth of a population at low density.

More research, on a larger spatial scale, would be needed to investigate what the constraints are at which scale for the establishment of new populations of *B. rostrata* (functional connectivity or presence of conspecifics). This highlights the uncertainties of connectivity in conservation planning (Hodgson *et al.*,

2009), which is in the case of *B. rostrata* a prevailing issue. Enlarging areas and their quality has less uncertainties, and is more straightforward to implement, when a conservationist wants to ameliorate a population of conservation interest.

When aeolic dynamics are hard or almost impossible to restore at a landscape level, livestock **grazing** is used as ‘the tool’ to open up landscapes. The trampling by large grazers (and vacationers) is very detrimental for the nesting of *B. rostrata*. As grazing is a system-approach which focuses on creating semi-natural disturbance dynamics (and applied in the major part of the nature reserve De Westhoek), some conservationists might be reluctant to revert to species-specific micro-scale conservation for *B. rostrata*. But as this conservation approach with grazers also has detrimental effects on other threatened specialised dune invertebrates (Maes & Bonte, 2006), specific measures are needed to conserve the community of typical natural dune species.

Based on those three factors, I provide some recommendations for the conservation of *B. rostrata* in the nature reserve De Westhoek. (1) Large populations of *B. rostrata* should not be grazed and should be fenced to serve as source populations. (2) Suitable adjacent areas should be cleared from thickets and grazers could help to reduce the thickness of the grass layer, to recover early stages of grey dunes (xerotherm moss dominated vegetation), as these are not any more maintained by natural, aeolic dune processes in Belgium and other European countries (Provoost *et al.*, 2004). Once these adjacent areas seem suitable, grazers are removed. Nearby high densities of *B. rostrata* of the source populations makes the chance of establishment high through a spill-over effect (cfr. distance kernel section 5.4.1 and temporal shift sections 5.4.3 and 5.3.2). (3) If large populations are becoming to densely vegetated and grazing should be performed, it is recommended to do this in autumn/winter and protect vulnerable large clusters by fencing them (micro-scale conservation). Another option would be to implement grazing in phases, e.g. 20% of a large population each year. The large, dense colonies have a crucial role in the local dynamics, survival and establishment in adjacent areas.

The conservation of *B. rostrata* on larger scales, national or European level, is constrained by the same factors (connectivity, presence of conspecifics, trampling). And of course, the connectivity constraint will be of higher importance at larger scales. When *B. rostrata* seems not to be able to (re)colonize on large scale apparently suitable habitat (Peeters, 2008), and the presence of conspecifics proves to be the major constraint, translocation, although highly controversial (Thomas, 2011b; Vilà & Hulme, 2011; Thomas, 2011a; Webber *et al.*, 2011), could be considered. Or while ameliorating coastal and inland dunes, we could also wait for the rare incidents of *B. rostrata* popping up at isolated sites, as observed by Gertjan Bisschop in 2016 in Wetteren (Waarnemingen.be, 2016). However, to know if a population is able to establish at that location, it has to be monitored the following years.

## 5.7 Possible future research

The study of this thesis could be complemented with modelling of the mechanisms with e.g. an individual based model (IBM), to see if the pattern (and network parameters) can be replicated with a bottom-up/mechanistic model. Structural equation modelling could assist the parameterisation of such an IBM. A top-down/pattern-oriented model approach could be used to test the phase separation principle on this example. These two models can be complementary and inverse modelling can link parameters of the two models with each other.

A projection of the microhabitat suitability model on the complete nature reserve or adjacent dune areas could be performed. This could help in directed searches for populations, to complement regular inventarisations. Probably, some technical difficulties will have to be tackled, such as the calibration of the drone-maps and current model to large scale maps covering the complete nature reserve.

Investigating the impact of the parasites across generations could help understanding the strength of the selective advantage of dense aggregates. Complementary to this, behaviours against these predators, hasty flying when parasites are near or dragging infected preys out of the burrow, could be assessed at in more detail.

Upscaling this research, and thus a large project regarding the mobility at several spatial and temporal scales of *Bembix rostrata*, would be highly interesting, both for fundamental ecology and conservation. It could integrate the other research options mentioned above, study in detail the different processes acting at individual, local and landscape scale, elucidate the constraints of its movements, investigate its space use and structure as a metapopulation *sensu lato* (Hanski *et al.*, 1995; Fronhofer *et al.*, 2012). A molecular approach could be included, to investigate genetic exchange on longer temporal scales and test the hypothesis of sex-biased dispersal.

## 6 Conclusion

The research conducted for the master project has demonstrated the **simultaneous influence of both bottom-up and top-down regulators** on the spatial patterning of *Bembix rostrata*. The microhabitat model determined the bottom-up regulator: both insolation (warmth) and the vegetation index (NDVI) are significant factors in predicting the presence or absence of nests in the study area. The selfish herd mechanism is the top-down regulator: the individual parasite density decreases with increasing nest density, which provides a selective advantage for females in dense aggregates.

The microhabitat model predicts the probability of nest presence with formula (6). Sites with **nests have a higher insolation and a lower vegetation index** than do random points in the field. The optimal scale for prediction was 2 m: the influence of the surrounding insolation and slope is of importance for the presence of the nests up to 2 m. This confirms the hypothesis that wasps have specific preferences for the microhabitat and that the characteristics up to a radius of a few metres around the nest has to be taken into account. Contrary to the expectations, there was no change in microhabitat preferences with time. The hypothesis that there was no microhabitat difference between parasitised and non-parasitised nests was confirmed.

CMRC analysis estimated a high **population size around 1000 female** individuals, which is high. The **survival probability was around 1** throughout the sampling period; this confirms that the mortality of females mainly occurs at the end of the season. The joining numbers indicate a highly fluctuating number of **emigration and immigration from and to adjacent areas**.

The spatial pattern of the nests was clustered up to a radius of 15 m. Within that radius, smaller clusters were made of a **radius between 3-5 m**. The spatial pattern analysis also indicated that the nests were temporally clustered and that the prediction value of the microhabitat model was similar for nearby nests. Parasitised nests were randomly distributed across the nests.

The wasps mainly made nests close to their previous nest, but a considerable number also made their nests in another cluster, resulting in quite a dense network.

The importance and centrality of the clusters change with time, which confirms the **temporal clustering** of the nests. This pattern was not completely clear-cut; **spatial bet hedging** is hypothesised as a possible mechanism for the attenuation of the temporal clustering.

The **prediction map** indicated that the most optimal habitat was saturated, but still some good sites were vacant. Cluster 1 and 9 were poorly predicted. Multiple explanations were given to account for this rarity (edge effects, missing variable in the model, ecological trap, marginal cluster for spatial bet hedging). This site selection map is the bottom-up template for nest clustering.

The **selfish herd mechanism** is suggested as being the mechanism behind the observation of decreasing individual parasite density with increasing nest density. This is the top-down regulator of nest clustering, a mechanism probably of considerable importance in the formation of spatial and temporal clustering.

This selfish herd mechanism and the temporal clustering are indicative of a system with density-dependent movement, which promotes the phase separation principle for self-organisation of nests of *B. rostrata*. This could be tested with pattern-oriented modelling.

*B. rostrata* is probably a **less philopatric** species than previously thought: the biggest issue for displaying such high site fidelity is most likely the need for the **presence of conspecifics**. This shows the importance of conservation which focuses as hard on habitat quality and quantity as it does on connectivity. Specific measures to maintain suitable microhabitat are important for *B. rostrata*, as well as for other dune specialist arthropods, when grazing is the dominant disturbance dynamic.

## 7 Summary/samenvatting

### 7.1 English summary

One of the major discussions in ecology has been about the **bottom-up or top-down regulation** of biological systems. Heterogeneity and the simultaneous influence of both regulators have been considered in this discussion, shifting from a simple dichotomy towards a more complex, integrated view. The formation of **heterogeneous spatial patterns** has been explained through spatial self-organisation, which concerns an activator-inhibitor mechanism or phase separation principle. The first one is based on a negative feedback on large scales and a positive feedback on small scales, the latter one is based on density-dependent movement. Both mechanisms result in large scale spatial patterns. The **simultaneous influence** of bottom-up and top-down regulators are relevant when looking at spatial patterns of, for instance, nestbuilding: both abiotic factors and biotic interactions are essential. Of importance in habitat selection of nest sites is habitat quality, biotic interactions and movement (in relation to connectivity). Increasing habitat quality and area has the potential to augment populations with less uncertainties than does the connectivity which conservation often focuses on.

The **focal species of this master project is *Bembix rostrata***, a specialised, highly philopatric, gregariously nesting digger wasp (Crabronidae) found in sandy regions in Europe. The females show brood care: they bring flies to provision their larvae in their burrows built in the sand. Several brood-parasitic fly species lay their eggs or maggots on the provisioned prey of the females. It is a highly endangered species in Belgium.

This master project investigated **processes involved in the spatial pattern formation of nest aggregates of *B. rostrata***, looking at both bottom-up and top-down regulators. Female wasps were individually tagged, their nests marked with a flag and parasitic infections observed. This was done in a study area of 50 on 90 m in the nature reserve De Westhoek. Remote sensing imagery was made by means of a drone, in order to characterise the microhabitat of the wasps (vegetation, slope and insolation). A microhabitat model was made to identify the nest site preferences. A Capture-Mark-Recapture (CMRC) analysis was carried out to investigate population size, survival probability and migration numbers. The spatial and temporal clustering were studied in detail with spatial pattern analysis and network analysis. It was investigated if a selfish herd mechanism is present in the system, in which high nest densities have lower individual chance of parasite infection. The hypothesis is that nest clustering is regulated by (1) the bottom-up template defined by the nest site suitability which is predicted by the microhabitat model and by (2) the top-down regulator which is defined by the selfish herd mechanism.

The **microhabitat model** indicated that nests have a higher insolation and lower vegetation index than do random points in the study area. The model performed best at a scale of 2 m. Thus, the insolation and vegetation of the surroundings up to that scale have an influence on nest habitat selection by *B. rostrata*. The microhabitat preferences did not change with time, and thus the wasps do not have to settle in suboptimal habitat later in the season.

The **CMRC-analysis** gave an estimate of around 1000 female individuals. The survival probability estimate was around 1, which indicates that mortality largely occurs at the end of the season. Highly fluctuating (positive and negative) joining number estimates, indicate that migration to and from adjacent areas is of considerable importance.

**Point pattern analysis** showed that there is a clear clustering up to a radius of 15 m, with smaller clusters forming within that radius of 3-5 m. This analysis also indicated that there is temporal clustering.

The **network analysis** showed that most of the nests are made very close to their previous nest, but that less than half of the nests are made in another cluster. When a nest is built in another cluster, the distance between the two clusters is of no importance. With time, the network is very variable and the importance of clusters changes. The pattern is not completely clear-cut; spatial bet hedging is hypothesised as a possible mechanism which attenuates the temporal clustering.

The **prediction map**, based on the microhabitat model, showed that not all preferred microhabitat is saturated. It is striking that cluster 9 was very poorly predicted (northeast corner). Some explanations are hypothesised: edge effects, missing explanatory information, mismatching cues resulting in an ecological trap or a marginal cluster used for spatial bet hedging.

The **selfish herd mechanism** provides a selective advantage for females in dense aggregates, as the density of parasites per individual decreases with increasing nest density. The prediction map forms the

bottom-up template and the selfish herd mechanism forms the top-down regulator of nest clustering in *B. rostrata*.

The **phase separation principle** is hypothesised as being the mechanism of self-organisation of the spatial pattern, because both spatial and temporal clustering are present as well as density-dependent movement, as indicated by the selfish herd mechanism. Modelling these dynamics could confirm this mechanism and could help to pinpoint the relative importance of bottom-up and top-down regulators.

Because this master project indicates that *B. rostrata* is less philopatric than previously thought, which has important **conservation implications**. Colonisation of *B. rostrata* seems the most constraint by the presence of conspecifics. Livestock grazing, used as the tool in dunes to create disturbance dynamics (because aeolic dynamics are no longer present), is detrimental to the nesting of *B. rostrata*. Aggregates of *B. rostrata* should be protected with micro-conservation, using fences, when grey dunes become too overgrown and grazing is recommended. The dense aggregates play a central part in the dynamics of the population.

## 7.2 Nederlandse samenvatting

De vraag of biologische systemen dan wel **bottom-up of top-down gereguleerd** zijn, is één van de grootste discussies in de ecologie. Heterogeniteit en de gelijktijdige invloed van beide regulatoren werden ook bij de discussie betrokken, waardoor de inzichten verschoven van een dichotome naar een meer complexe, geïntegreerde visie. **Heterogene ruimtelijke patronen** worden verklaard aan de hand van zelf-organisatie, met behulp van een activator-inhibitor of een fasescheiding mechanisme. Het eerste mechanisme is gebaseerd op een negatieve feedback op grote schaal en een positieve feedback op kleine schaal, het tweede is gebaseerd op dichtheidsafhankelijke bewegingen. Beide mechanismen zorgen voor ruimtelijke patronen op grote schaal. De **gelijktijdige invloed van bottom-up en top-down regulatoren** zijn belangrijk als ruimtelijke patronen van o.a. nestbouw worden onderzocht: zowel abiotische factoren als biotische interacties zijn van belang. Bij het kiezen van geschikte nestplaatsen zijn habitatkwaliteit, biotische interacties en beweging (in het opzicht van connectiviteit) van belang. Wanneer populaties worden beschermd, biedt het verhogen van habitatkwaliteit en -oppervlakte meer zekerheid dan het verbeteren van connectiviteit, alhoewel natuurbeheer vaak focust op het laatste.

**De soort waarop in dit masterproject werd gefocust is de harkwesp *Bembix rostrata***, een gespecialiseerde, plaatsgetrouwe graafwesp (Crabronidae) die nesten bouwt in aggregaten. De harkwesp komt voor in zanderige regio's in Europa. De vrouwtjes doen aan broedzorg door prooien naar hun nest, gevangen in het zand, te brengen om hun larve te voeden. Verschillende soorten broedparasitische vliegen leggen hun eitjes of larven op de prooien die de vrouwtjes aanbrengen. De harkwesp is sterk bedreigd in België.

**Dit masterproject onderzoekt de processen betrokken bij de vorming van de ruimtelijk patronen van nesten in nestaggregaten bij de harkwesp**, waarbij zowel bottom-up als top-down regulatoren worden onderzocht. Vrouwelijke harkwespen werden individueel getagged, hun nesten met een vlag aangeduid en infecties door parasieten genoteerd. Dit in een studiegebied van 50 op 90 m, in het natuureservaat de Westhoek. Teledetectiebeelden werden gemaakt met een drone, om het microhabitat van de harkwesp te karakteriseren (vegetatie, helling en zonne-instraling). Een microhabitat model werd opgesteld om de nestplaatsvoorkeuren te bepalen. Een vangst-hervangst analyse werd uitgevoerd om de populatiegrootte, overlevingskans en migratie-aantallen te onderzoeken. De clustering in ruimte en tijd werd in detail bestudeerd met ruimtelijke patroon analyses en netwerk analyses. Er werd onderzocht of hoge nestdichtheden zorgen voor een lagere individuele kans op parasietinfectie, het 'selfish herd' mechanisme. De hypothese is dat nestclustering wordt gereguleerd door (1) de bottom-up template gevormd door de geschiktheid van de nestplaats, voorspeld aan de hand van het microhabitat model en (2) de top-down regulator gevormd door het 'selfish herd' mechanisme.

Het **microhabitat model** toonde aan dat nesten een hogere insolatie en lagere vegetatie index hebben dan random geselecteerde punten in het onderzoeksgebied. Het model presteerde het best op een schaal van 2 m, wat erop neerkomt dat zonne-instraling en vegetatie van de omgeving rond een nest tot op een straal van 2 m van belang zijn voor de habitat selectie van de harkwesp. De microhabitat voorkeuren veranderden niet met tijd, dus de harkwespen nemen geen suboptimaal habitat in later in het seizoen.

De **vangst-hervangst analyse** gaf een schatting van rond de 1000 vrouwelijke individuen. De

geschatte overlevingskans was ongeveer 1, wat aangeeft dat de mortaliteit grotendeels plaatsvindt op het einde van het seizoen. Hard fluctuerende (positieve en negatieve) migratie-aantallen duiden erop dat de migratie van en naar nabijgelegen gebieden van belang is.

De **punt-patroon analyse** toonde dat er duidelijk clustering optreedt, tot op een straal van 15 m, met kleinere clusters gevormd binnen die 15 m, met een straal van 3-5 m. Deze analyse duidde ook op clustering in de tijd.

De **netwerkanalyse** toonde dat de meeste nesten dichtbij het vorige nest van hetzelfde individu werden gemaakt. Iets minder dan de helft van de nesten werden in aan andere cluster gegraven. Als een nest in een andere cluster wordt gemaakt, speelt de afstand tussen de clusters geen rol. Het netwerk is erg variabel en het belang van elke cluster verandert met de tijd. Het patroon is niet volledig afgelijnd; ruimtelijke 'bet hedging' werd naar voren geschoven als mogelijk mechanisme voor het verzwakte patroon.

De **voorspellingskaart**, gebaseerd op de voorspellingen van het microhabitat model, toonde dat niet al het voorkeursmicrohabitat gesatureerd is. Het was opvallend dat cluster 9 (noordoostelijke hoek) zo slecht werd voorspeld. Enkele mogelijke verklaringen zijn randeffecten, ontbrekende verklarende variabelen, mismatchende aanwijzingen die resulteren in een ecologische val of een marginale cluster gebruikt voor ruimtelijke 'bet hedging'. Het '**selfish herd**' mechanisme zorgt voor een selectief voordeel voor vrouwtjes in dichte aggregaten, want de dichtheid van parasieten per individu neemt af met het stijgen van de nestdichtheid. De voorspellingskaart, die de geschiktheid van het microhabitat weergeeft, vormt de bottom-up template en het selfish herd mechanisme is de top-down controle van het clusteren van nesten bij de harkwesp.

Het **fasescheidingsprincipe** werd naar voor gebracht als het mechanisme voor zelf-organisatie van het ruimtelijk patroon van de nesten, want er is zowel clustering in ruimte en tijd als dichtheidsafhankelijke beweging, aangegeven door het 'selfish herd' mechanisme. Modelleren van deze dynamieken zou dit principe kunnen bevestigen en zou kunnen helpen om het relatief belang van bottom-up en top-down controles op te helderen.

Dit masterproject wijst er op dat de harkwesp in mindere mate plaatsgetrouw (philopatrisch) is dan eerder gedacht, wat belangrijke gevolgen heeft voor beheer. De kolonisatie van de harkwesp lijkt het meest beperkt door het al dan niet aanwezig zijn van soortgenoten. Begrazing door grote grazers, gebruikt als de maatregel om verstoringdynamieken te creëren in duinen (omdat de natuurlijke winddynamieken zijn weggevalen), is erg nefast voor de harkwesp. Nestaggregaten van de harkwesp zouden beschermd moeten worden door microbeheer, met behulp van rasters, als de mosduinen te hard dichtgroeien en begrazing aangewezen is. Dichte nestaggregaten hebben een centrale rol in de dynamiek van een populatie harkwespen.

## 8 Acknowledgements

I would like to thank the following people for their support and help, making this dissertation possible:

- My supervisor and tutor Prof. Dr. Dries Bonte. For proposing me this intriguing subject (in a bar in Helsinki), guiding me in my fieldwork and analyses with great consideration for my own autonomy. His door was always open and there was always ample time for discussion. I am very grateful for this opportunity.
- Rpaswork and Landron Projects for the drone-flights. Especially Yves Landron for his quick responses and help with questions.
- Wouter Van Gompel, who I spent two days with geolocating more than 1000 nests in the study area with a Trimble GPS.
- Johan Lamaire for letting me camp in his garden and giving me advice regarding the species in the nature reserve.
- Sam Provoost for making the insolation and slope maps in GIS and for being a reading commissioner.
- Dirk Maes for discussing possibilities of microhabitat modelling and for being a reading commissioner.
- People from the TEREK-lab: Hans Mattheve for introducing me to ArcGIS and Steven Goossens for help in the field and help with tagging.
- The professors, lecturers and assistants involved in my training as a biologist.
- Thijs Dekeukeleire for his valuable language advice and Filip Mangeleer for diving into the text.
- My friends and fellow students for the pleasant years in biology.
- Willem Van Echelpoel for taking such great pictures.
- My sister Joke for her support and Tom Van Assche for his vital guidance with GIS.
- My parents for their unconditional support, help in the complete process of this dissertation and the possibility of continuing my studies in biology. Their stimulation of my curiosity and interests since my childhood has been the foundation of my development as a science student.
- Daan, the love of my life, for his endless support and elaborate discussions about the subject.

## 9 Bibliography

### References

- Allouche, O., Tsoar, A. & Kadmon, R. (2006). Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *Journal of Applied Ecology*, 43, 1223–1232.
- Aplin, L.M., Farine, D.R., Morand-Ferron, J., Cockburn, A., Thornton, A. & Sheldon, B.C. (2015). Experimentally induced innovations lead to persistent culture via conformity in wild birds. *Nature*, 518, 538–541.
- Baddeley, A. (2010). *Analysing spatial point patterns in R*. 4th edn. CSIRO. pp. 232.
- Baddeley, A., Rubak, E. & Turner, R. (2015). *Spatial Point Patterns: Methodology and Applications with R*. London: Chapman and Hall/CRC. pp. 810. <http://www.spatstat.org/>.
- Bailey, D.W., Gross, J.E., Laca, E.A., Rittenhouse, L.R., Coughenour, M.B., Swift, D.M. & Sims, P.L. (1996). Mechanisms That Result in Large Herbivore Grazing Distribution Patterns. *Journal of Range Management*, 49, 386–400.
- Banse, K. (2007). Do we live in a largely top-down regulated world? *Journal of Biosciences*, 32, 791–796.
- Bélisle, M. (2005). Measuring Landscape Connectivity: the Challenge of Behavioural Landscape Ecology. *Ecology*, 86, 1988–1995.
- Belles-Isles, J.C. & Picman, J. (1986). Nesting Losses and Nest Site Preferences in House Wrens. *The Condor*, 88, 483–486.
- Bjornstad, O.N. (2016). ncf: Spatial Nonparametric Covariance Functions. <https://cran.r-project.org/package=ncf>.
- Blonder, B. (2015). timeordered: Time-ordered and time-aggregated network analyses. <https://cran.r-project.org/package=timeordered>.
- Bonte, D. (2005). Anthropogenic induced changes in nesting densities of the dune-specialised digger wasp *Bembix rostrata* (Hymenoptera: Sphecidae). *Eur. J. Entomol.*, 102, 809–812.
- Bonte, D., Lens, L., Maelfait, J.P., Hoffmann, M. & Kuijken, E. (2003). Patch quality and connectivity influence spatial dynamics in a dune wolfspider. *Oecologia*, 135, 227–233.
- Bowlby, J.N. & Roff, J.C. (1986). Trophic Structure in Southern Ontario Streams. *Ecology*, 67, 1670–1679.
- Brockmann, H.J. & Dawkins, R. (1979). Joint Nesting in a Digger Wasp as an Evolutionarily Stable Preadaptation to Social Life. *Behaviour*, 71, 203–245.
- Brown, C.D. & Davis, H.T. (2006). Receiver operating characteristics curves and related decision measures: A tutorial. *Chemometrics and Intelligent Laboratory Systems*, 80, 24–38.
- Burkepile, D.E. & Hay, M.E. (2006). Herbivore vs. nutrient control of marine primary producers: Context-dependent effects. *Ecology*, 87, 3128–3139.
- Cahn, J.W. & Hilliard, J.E. (1958). Free Energy of a Nonuniform System. 1. Interfacial Free Energy. *The Journal of Chemical Physics*, 28, 258–267.
- Campomizzi, A.J., Butcher, J.A., Farrell, S.L., Snelgrove, A.G., Collier, B.A., Gutzwiller, K.J., Morrison, M.L. & Wilkins, R.N. (2008). Conspecific attraction is a missing component in wildlife habitat modeling. *Journal of Wildlife Management*, 72, 331–336.
- Chaverri, G. (2010). Comparative social network analysis in a leaf-roosting bat. *Behavioral Ecology and Sociobiology*, 64, 1619–1630.
- Clark, C.W. & Mangel, M. (1984). Foraging and Flocking Strategies: Information in an Uncertain Environment. *The American Naturalist*, 123, 626–641.
- Cohen, D. (1966). Optimizing reproduction in a randomly varying environment. *J. Theoret. Biol.*, 12, 119–129.
- Couteron, P. & Lejeune, O. (2001). Periodic spotted patterns in semi-arid vegetation explained by a propagation-inhibition model. *Journal of Ecology*, 89, 616–628.
- Csardi, G. & Nepusz, T. (2006). The igraph software package for complex network research. *InterJournal, Complex Sy*, pp. 1695. <http://igraph.org>.
- Csardi, G. & Nepusz, T. (2013). *igraph Reference Manual*. pp. 550. <http://igraph.sourceforge.net/doc/igraph-docs>.
- Davies, N.B., Krebs, J.R. & West, A.S. (2012). *An Introduction to Behavioural Ecology*. 4th edn. Wiley-Blackwell. pp. 520.

- Dormann, C.F., McPherson, J.M., Araújo, M.B., Bivand, R., Bolliger, J., Carl, G., Davies, R.G., Hirzel, A., Jetz, W., Kissling, W.D., Kühn, I., Ohlemüller, R., Peres-Neto, P.R., Reineking, B., Schröder, B., Schurr, F.M. & Wilson, R. (2007). Methods to account for spatial autocorrelation in the analysis of species distributional data: A review. *Ecography*, 30, 609–628.
- Dunson, W.A. & Travis, J. (1991). The Role of Abiotic Factors in Community Organization. *The American Naturalist*, 138, 1067–1091.
- Dyer, L.A. & Letourneau, D. (2003). Top-down and bottom-up diversity cascades in detrital vs. living food webs. *Ecology Letters*, 6, 60–68.
- Eggers, S., Griesser, M., Nystrand, M. & Ekman, J. (2006). Predation risk induces changes in nest-site selection and clutch size in the Siberian jay. *Proceedings of the Royal Society B*, 273, 701–706.
- ESRI (2011a). ArcGIS Desktop: Release 10. ArcMap Tools, Solar Radiation toolset. Environmental Systems Research Institute.
- ESRI (2011b). ArcGIS Desktop: Release 10. Environmental Systems Research Institute. Redlands, CA, USA.
- Ettema, C.H. & Wardle, D.A. (2002). Spatial soil ecology. *Trends in Ecology and Evolution*, 17, 177–183.
- Evans, H.E. & O’Neill, K.M. (2007). *The Sand Wasps: Natural History and Behavior*. Harvard University Press, Cambridge, Massachusetts, USA. pp. 340.
- Farine, D.R. & Whitehead, H. (2015). Constructing, conducting and interpreting animal social network analysis. *Journal of Animal Ecology*, 84, 1144–1163.
- Fewell, J.H. (2003). Social insect networks. *Science*, 301, 1867–1870.
- Field, J. (2005). The evolution of progressive provisioning. *Behavioral Ecology*, 16, 770–778.
- Fielding, A.H. & Bell, J.F. (1997). A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation*, 24, 38–49.
- Forstmeier, W. & Weiss, I. (2004). Adaptive plasticity in nest-site selection in response to changing predation risk. *Oikos*, 104, 487–499.
- Frankie, G.W., B, V.S., Newstrom, L.E. & Barthell, J.F. (1988). Nest Site and Habitat Preferences of Centris Bees in the Costa Rican Dry Forest. *Biotropica*, 20, 301–310.
- Fronhofer, E.A., Kubisch, A., Hilker, F.M., Hovestadt, T. & Poethke, H.J. (2012). Why are metapopulations so rare? *Ecology*, 93, 1967–1978.
- Giroux, M.A., Trottier-paquet, M., Bêty, J., Lamarre, V. & Lecomte, N. (2016). Is it safe to nest near bold neighbours? Spatial patterns in predation risk associated with the density of American Golden-Plover nest. *PeerJ*, 4, 1–13.
- Godfrey, S.S., Moore, J.A., Nelson, N.J. & Bull, C.M. (2010). Social network structure and parasite infection patterns in a territorial reptile, the tuatara (*Sphenodon punctatus*). *International Journal for Parasitology*, 40, 1575–1585.
- Gripenberg, S. & Roslin, T. (2007). Up or down in Space? Uniting the Bottom-up versus Top-down Paradigm and Spatial Ecology. *Oikos*, 116, 181–188.
- Hahn, B.A. & Silverman, E.D. (2006). Social cues facilitate habitat selection: American redstarts establish breeding territories in response to song. *Biology Letters*, 2, 337–40.
- Hairston, N.G., Smith, F.E. & Slobodkin, L.B. (1960). Community Structure, Population and Competition Control. *The American Naturalist*, XCIV, 421–425.
- Hamilton, W.D. (1971). Geometry for the selfish herd. *Journal of Theoretical Biology*, 31, 295–311.
- Hammond, J.I., Luttbeg, B. & Sih, A. (2007). Predator and prey space use: Dragonflies and tadpoles in an interactive game. *Ecology*, 88, 1525–1535.
- Hanski, I., Pakkala, T., Kuussaari, M. & Lei, G. (1995). Metapopulation Persistence of an Endangered Butterfly in a Fragmented Landscape. *Oikos*, 72, 21–28.
- Hodgson, J.A., Thomas, C.D., Wintle, B.A. & Moilanen, A. (2009). Climate change, connectivity and conservation decision making: Back to basics. *J. Appl. Ecol.*, 46, 964–969.
- Horswill, C., Ratcliffe, N., Green, J.A., Phillips, R.A., Trathan, P.N. & Matthiopoulos, J. (2016). Unravelling the relative roles of top-down and bottom-up forces driving population change in an oceanic predator. *Ecology*, 97, 1919–1928.
- Hughes, W.O.H., Oldroyd, B.P., Beekman, M. & Ratnieks, F.L.W. (2008). Ancestral Monogamy Shows Kin Selection is Key to the Evolution of Eusociality. *Science*, 320, 1213–1216.

- Hunter, M.D. & Price, P.W. (1992). Playing Chutes and Ladders: Heterogeneity and the Relative Roles of Bottom-Up and Top-Down Forces in Natural Communities. *Ecology*, 73, 724–732.
- Hunter, M.D., Varley, G.C. & Gradwell, G.R. (1997). Estimating the relative roles of top-down and bottom-up forces on insect herbivore populations: a classic study revisited. *Proceedings of the National Academy of Sciences of the United States of America*, 94, 9176–81.
- Hutchinson, G.E. (1957). Concluding remarks. *Cold Spring Harbor Symposia on Quantitative Biology*, 22, 415–427.
- Inglis, I.R., Langton, S., Forkman, B. & Lazarus, J. (2001). An information primacy model of exploratory and foraging behaviour. *Animal Behaviour*, 62, 543–557.
- Ji, R., Stegert, C. & Davis, C.S. (2013). Sensitivity of copepod populations to bottom-up and top-down forcing: A modeling study in the Gulf of Maine region. *Journal of Plankton Research*, 35, 66–79.
- Johnson, M.D. (2007). Measuring Habitat Quality: a Review. *The Condor*, 109, 489–504.
- Karr, J.R., Dionne, M. & Schlosser, I.J. (1992). Bottom-up versus top-down regulation of vertebrate populations: lessons from birds and fish. In: *Effects of resource distribution on animal-plant interactions* (eds. Hunter, M.D., Ohgushi, T. & Price, P.W.). Academic Press, San Diego, California, USA, pp. 243–286.
- Kells, A.R. & Goulson, D. (2003). Preferred nesting sites of bumblebee queens (Hymenoptera: Apidae) in agroecosystems in the UK. *Biological Conservation*, 109, 165–174.
- Kerth, G., Ebert, C. & Schmidtke, C. (2006). Group decision making in fission-fusion societies: evidence from two-field experiments in Bechstein’s bats. *Proceedings of the Royal Society B*, 273, 2785–90.
- Klein, W.F. & Lefeber, V. (2004). Crabronidae graafwespen. In: *De wespen en mieren van Nederland* (eds. Peeters, T. & Al., E.). Naturalis en KNNV, pp. 356–430.
- Koh, L.P. & Wich, S.A. (2012). Dawn of drone ecology: low-cost autonomous aerial vehicles for conservation. *Tropical Conservation Science*, 5, 121–132.
- van de Koppel, J. & Crain, C.M. (2006). Scale-dependent inhibition drives regular tussock spacing in a freshwater marsh. *The American Naturalist*, 168, E136–E147.
- van de Koppel, J., Gascoigne, J.C., Theraulaz, G., Rietkerk, M., Mooij, W.M. & Herman, P.M.J. (2008). Experimental evidence for spatial self-organization and its emergent effects in mussel bed ecosystems. *Science*, 322, 739–42.
- Krause, J., Lusseau, D. & James, R. (2009). Animal social networks: An introduction. *Behavioral Ecology and Sociobiology*, 63, 967–973.
- Kristan, W.B. (2003). The role of habitat selection behavior in population dynamics: source-sink systems and ecological traps. *Oikos*, 103, 457–468.
- Larsen, D.R. (2014). *Mark and Recapture Methods: Jolly-Seber Method*. Natural Resources Biometrics. Retrieved 15-11-2016 at <http://oak.snr.missouri.edu/nr3110/topics/jolly.php>.
- Larsson, F. (1986). Increased Nest Density of the Digger Wasp *Bembix rostrata* as a Response to Parasites and Predators. *Entomol. Gener.*, 12, 71–75.
- Larsson, F. & Tengö, J. (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.
- Lin, N. & Michener, C.D. (1972). Evolution of Sociality in Insects. *The Quarterly Review of Biology*, 47, 131–159.
- Liu, Q.X., Doelman, A., Rottschäfer, V., de Jager, M., Herman, P.M., Rietkerk, M. & van de Koppel, J. (2013). Phase separation explains a new class of self-organized spatial patterns in ecological systems. *Proceedings of the National Academy of Sciences of the United States of America*, 110, 11905–11910.
- Liu, Q.X., Rietkerk, M., Herman, P.M.J., Piersma, T., Fryxell, J.M. & van de Koppel, J. (2016). Phase separation driven by density-dependent movement: A novel mechanism for ecological patterns. *Physics of Life Reviews*, 19, 107–121.
- Livingston, G., Fukumori, K., Provete, D.B., Kawachi, M., Takamura, N. & Leibold, M.A. (2017). Predators regulate prey species sorting and spatial distribution in microbial landscapes. *Journal of Animal Ecology*, 86, 501–510.
- Lloyd, J.D. & Martin, T.E. (2004). Nest-site preference and maternal effects on offspring growth. *Behavioral Ecology*, 15, 816–823.

- Lobo, J.M., Jiménez-valverde, A. & Real, R. (2008). AUC: A misleading measure of the performance of predictive distribution models. *Global Ecology and Biogeography*, 17, 145–151.
- Lõhmus, A. (2003). Do Ural owls (*Strix uralensis*) suffer from the lack of nest sites in managed forests? *Biological Conservation*, 110, 1–9.
- Maes, D. & Bonte, D. (2006). Using distribution patterns of five threatened invertebrates in a highly fragmented dune landscape to develop a multispecies conservation approach. *Biol. Conserv.*, 133, 490–499.
- Mahon, C.L., Martin, K. & Steventon, J. (2007). Habitat attributes and chestnut-backed chickadee nest site selection in uncut and partial-cut forests. *Canadian Journal of Forest Research*, 37, 1272–1285.
- Manel, S., Williams, H.C. & Ormerod, S.J. (2001). Evaluating presence absence models in ecology; the need to count for prevalence. *Journal of Applied ecology*, 38, 921–931.
- Martínez, J.A., Serrano, D. & Zuberogoitia, I. (2003). Predictive models of habitat preferences for the Eurasian eagle owl *Bubo bubo*: a multiscale approach. *Ecography*, 26, 21–28.
- Marvin, D.C., Koh, L.P., Lynam, A.J., Wich, S., Davies, A.B., Krishnamurthy, R., Stokes, E., Starkey, R. & Asner, G.P. (2016). Integrating technologies for scalable ecology and conservation. *Global Ecology and Conservation*, 7, 262–275.
- Matthiopoulos, J., Fieberg, J., Aarts, G., Beyer, H.L., Morales, J.M. & Haydon, D.T. (2015). Establishing the link between habitat-selection and animal population dynamics. *Ecological Monographs*, 85, 150119094816002.
- Mazerolle, M.J. & Villard, M.A. (1999). Patch characteristics and landscape context as predictors of species presence and abundance: A review. *Écoscience*, 6, 117–124.
- Meserve, P.L., Kelt, D.A., Milstead, W.B. & Gutiérrez, J.R. (2003). Thirteen Years of Shifting Top-Down and Bottom-Up Control. *BioScience*, 53, 633–646.
- Mistr, S. & Bercovici, D. (2003). A theoretical model of pattern formation in coral reefs. *Ecosystems*, 6, 61–74.
- Mittelbach, G.G., Osenberg, C.W. & Leibold, M.A. (1988). Trophic Relations and Ontogenetic Niche Shifts in Aquatic Ecosystems. In: *Size-Structured Populations* (eds. Ebenman, B. & Persson, L.). Springer, Berlin, Heidelberg, pp. 219–235.
- Moskat, C. & Honza, M. (2000). Effect of nest and nest site characteristics on the risk of cuckoo *Cuculus canorus* parasitism in the great reed warbler *Acrocephalus arundinaceus*. *Ecography*, 23, 335–341.
- Muller, K.L., Stamps, J.A., Krishnan, V.V. & Willits, N.H. (1997). The Effects of Conspecific Attraction and Habitat Quality on Habitat Selection in Territorial Birds (*Troglodytes aedon*). *The American Naturalist*, 150, 650–661.
- Naug, D. (2008). Structure of the social network and its influence on transmission dynamics in a honeybee colony. *Behavioral Ecology and Sociobiology*, 62, 1719–1725.
- Naug, D. (2009). Structure and resilience of the social network in an insect colony as a function of colony size. *Behavioral Ecology and Sociobiology*, 63, 1023–1028.
- Nielsen, E. (1945). *Moeurs des Bembex*. 7th edn. Spoolia Zool Mus Haun, København. pp. 174.
- Ognyanova, K. (2016). Network Analysis and Visualisation with R and igraph. Tech. rep., NetSciX 2016 School of Code Workshop, Wrocław, Poland. pp. 64.
- Orrock, J.L., Grabowski, J.H., Pantel, J.H., Peacor, S.D., Barbara, L., Sih, A. & Werner, E.E. (2008). Consumptive and Nonconsumptive Effects of Predators on Metacommunities of Competing Prey. *Ecology*, 89, 2426–2435.
- Peeters, T. (2008). De harkwesp *Bembix rostrata* - een literatuurstudie. *Nieuwsbrief NEV*, pp. 40–44.
- Perrin, N. & Mazalov, V. (2000). Local Competition, Inbreeding, and the Evolution of Sex-Biased Dispersal. *Am. Nat.*, 155, 116–127.
- Polidori, C., Casiraghi, M., Di Lorenzo, M., Valarani, B. & Andrietti, F. (2006). Philopatry, nest choice, and aggregation temporal-spatial change in the digger wasp *Cerceris arenaria* (Hymenoptera: Crabronidae). *Journal of Ethology*, 24, 155–163.
- Poulin, R. (2010). Network analysis shining light on parasite ecology and diversity. *Trends in Parasitology*, 26, 492–498.
- Power, M.E. (1992). Top-Down and Bottom-Up Forces in Food Webs: Do Plants Have Primacy? *Ecology*, 73, 733–746.

- Provoost, S., Ampe, C., Bonte, D., Cosyns, E. & Hoffmann, M. (2004). Ecology, management and monitoring of grey dunes in Flanders. *J. Coast. Conserv.*, 10, 33–42.
- Python Software Foundation (2001). Python Language Reference, version 3.4. <http://www.python.org>.
- QGIS Development Team (2009). QGIS Geographic Information System. Open Source Geospatial Foundation. <http://qgis.osgeo.org>.
- R Core Team (2016). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. <https://www.r-project.org/>.
- Reinhart, K.O., Packer, A., Van Der Putten, W.H. & Clay, K. (2003). Plant-soil biota interactions and spatial distribution of black cherry in its native and invasive ranges. *Ecology Letters*, 6, 1046–1050.
- Rietkerk, M., Dekker, S.C., de Ruiter, P.C. & van de Koppel, J. (2004). Self-organized patchiness and catastrophic shifts in ecosystems. *Science*, 305, 1926–1929.
- Rietkerk, M. & van de Koppel, J. (2008). Regular pattern formation in real ecosystems. *Trends in Ecology and Evolution*, 23, 169–175.
- Robertson, B.A. & Hutto, R.L. (2006). A framework for understanding ecological traps and an evaluation of existing evidence. *Ecology*, 87, 1075–1085.
- Sánchez-Zapata, J.A. & Calvo, J.F. (1999). Raptor distribution in relation to landscape composition in semi-arid Mediterranean habitats. *Journal of Applied Ecology*, 36, 254–262.
- Schöne, H. & Tengö, J. (1981). Competition of Males, Courtship Behaviour and Chemical Communication in the Digger Wasp *Bembix rostrata*. *Behaviour*, 77, 44–66.
- Schöne, H. & Tengö, J. (1991). Homing in the Digger Wasp *Bembix rostrata* - Release Direction and Weather Conditions. *Ethology*, 87, 160–164.
- Seeley, T.D. & Morse, R.A. (1978). Nest site selection by the honey bee, *Apis Mellifera*. *Insectes Sociaux*, 25, 323–337.
- Seitz, R., Lipcius, R. & Hines, A. (2017). Consumer versus resource control and the importance of habitat heterogeneity for marine bivalves. *Oikos*, 126, 121–135.
- Sih, A. (2013). Understanding variation in behavioural responses to human-induced rapid environmental change: A conceptual overview. *Animal Behaviour*, 85, 1077–1088.
- Sih, A., Hanser, S.F. & McHugh, K.A. (2009). Social network theory: new insights and issues for behavioral ecologists. *Behavioral Ecology and Sociobiology*, 63, 975–988.
- Sing, T., Sander, O., Beerenwinkel, N. & Lengnauer, T. (2005). ROCRC: visualizing classifier performance in R. *Bioinformatics*, 21, 7881.
- Slagsvold, T. (1987). Nest Site Preference and Clutch Size in the Pied Flycatcher *Ficedula hypoleuca*. *Ornis Scandinavica (Scandinavian Journal of Ornithology)*, 13, 165–175.
- Solé, R.V. & Bascompte, J. (2006). *Complexity in Ecological Systems*. Princeton University Press. pp. 284.
- Stephens, P.A., Sutherland, W.J. & Freckleton, R.P. (1999). What Is the Allee Effect? *Oikos*, 87, 185–190.
- Storch, I. (1991). Habitat Fragmentation, Nest Site Selection, and Nest Predation Risk in Capercaillie. *Ornis Scandinavica (Scandinavian Journal of Ornithology)*, 22, 213–217.
- Sumpter, D.J.T. (2006). The principles of collective animal behaviour. *Philos. Trans. R. Soc. B*, 361, 5–22.
- Suryan, R.M., Irons, D.B., Brown, E.D., Jodice, P.G.R. & Roby, D.D. (2006). Site-specific effects on productivity of an upper trophic-level marine predator: Bottom-up, top-down, and mismatch effects on reproduction in a colonial seabird. *Progress in Oceanography*, 68, 303–328.
- Svensson, B., Lagerlöf, J. & G. Svensson, B. (2000). Habitat preferences of nest-seeking bumble bees (Hymenoptera: Apidae) in an agricultural landscape. *Agriculture, Ecosystems & Environment*, 77, 247–255.
- van Swaay, C. & Plate, C. (2009). Grootste klappen in de soortenrijke duinen. *Vlinders (meetnet vlinders)*, pp. 14–15.
- Tengö, J., Schöne, H., Kühme, W., Schöne, H. & Kühme, L. (1996). Nesting cycle and homing in the digger wasp *Bembix rostrata*. *Ethol. Ecol. Evol.*, 8, 207–211.
- Theraulaz, G., Bonabeau, E., Nicolis, S.C., Solé, R.V., Fourcassié, V., Blanco, S., Fournier, R., Joly, J.L., Fernández, P., Grimal, A., Dalle, P. & Deneubourg, J.L. (2002). Spatial patterns in ant colonies. *PNAS*, 99, 9645–9649.

- Thomas, C.D. (2011a). Anthropocene Park? No alternative. *Trends in Ecology and Evolution*, 26, 497–498.
- Thomas, C.D. (2011b). Translocation of species, climate change, and the end of trying to recreate past ecological communities. *Trends in Ecology and Evolution*, 26, 216–221.
- Tinbergen, N. (1951). *The Study of Instinct*. Oxford Clarendon Press.
- Turing, A.M. (1953). The Chemical Basis of Morphogenesis. *Philos. Trans. R. Soc. B*, 237, 37–72.
- Van Horne, B. (1983). Density as a Misleading Indicator of Habitat Quality. *The Journal of Wildlife Management*, 47, 893–901.
- Vanommeslaeghe, A. (2015). *Movement patterns, nesting preferences and population estimates of the digger wasp *Bembix rostrata**. Msc thesis, [MSc thesis]. Unpubl. manuscript, UGent, Depart. Biology.
- Venables, W. & Ripley, B. (2002). *Modern Applied Statistics with S*. 4th edn. Springer, Yew York. pp. 501.
- Vilà, M. & Hulme, P.E. (2011). Jurassic Park? No thanks. *Trends in Ecology and Evolution*, 26, 493–494.
- Waarnemingen.be (2016). Harkwesp - *Bembix rostrata*; 25-06-2016; Wetteren - Speelbos Warandeduinen [OV]; Gertjan Bisschop; waarnemingen.be.
- Walker, M. & Jones, T.H. (2001). Relative roles of top-down and bottom-up forces in terrestrial tritrophic plant-insect herbivore-natural enemy systems. *Oikos*, 93, 177–187.
- Ward, M.P. & Schlossberg, S. (2004). Conspecific Attraction and the Conservation of Territorial Songbirds. *Conservation Biology*, 18, 519–525.
- Wcislo, W.T. (1984). Gregarious nesting of a digger wasp as a "selfish herd" response to a parasitic fly (Hymenoptera: Sphecidae; Diptera: Sacrophagidae). *Behavioral Ecology and Sociobiology*, 15, 157–160.
- Webber, B.L., Scott, J.K. & Didham, R.K. (2011). Translocation or bust! A new acclimatization agenda for the 21st century? *Trends in Ecology and Evolution*, 26, 493–494.
- Westphal, C., Steffan-Dewenter, I. & Tscharrntke, T. (2006). Bumblebees experience landscapes at different spatial scales: possible implications for coexistence. *Oecologia*, 149, 289–300.
- White, G.C. & Burnham, K.P. (1999). Program MARK: survival estimation from populations of marked animals. *Bird Study*, 46, S120–S139.
- Wickham, H. (2009). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. pp. 213.
- Wuellner, C.T. (1999). Nest site preference and success in a gregarious, ground-nesting bee *Dieunomia triangulifera*. *Ecological Entomology*, 24, 471–479.

## 10 Appendices

The url-link to the datasets and the output of the data can be found on this dropbox-link:

<https://www.dropbox.com/sh/skghyd8r1k8r8nm/AAC0tc1Ckh775uvMBkQhZny2a?dl=0>

At that link there are 5 maps containing R-outputs (html-files made with R-Markdown), data-files used in the code (text-files) that were derived from the database or filtered with Excel or a quick Python script (not added), maps and raw data. Below is a list with the file names of the appendices. In the main text, references were made to this list.

### 1. Microhabitat model

- (a) MH\_Random\_points.html with Nest\_parasiet\_coor\_date.txt as data file
- (b) MH\_Data\_exploration.html with MHmodel.txt as data file
- (c) MH\_Spatial\_autocorrelation\_structure.html with MHmodel.txt as data file
- (d) MH\_Terms.html with MHmodel.txt as data file
- (e) MH\_Scale.html with MHmodel.txt as data file
- (f) MH\_Full\_models.html with MHmodel.txt, RealnestsModel.txt, TemporalModel.txt, Nests\_Cluster\_predictions.txt, Coordinates\_Middle\_clusters.txt and RegularPointsData.txt as data files

### 2. Population estimation and spatial pattern analysis

- (a) CMRC\_estimation.xls (calculations)
- (b) CMRC\_estimation.doc (results)
- (c) PPA\_code.html with Nests\_Cluster\_predictions.txt as data file

### 3. The population as a network

- (a) NETW\_Code\_distances\_freq.py with Ind\_nest.txt and Ind\_nest\_all.txt as data files
- (b) NETW\_Clusteranalysis.html with Nest\_parasiet\_coor\_date.txt as data file
- (c) NETW\_customize\_data.html with afstanden\_opeenvolgend.txt, Nest-Cluster.txt, afstanden\_opeenvolgend\_all.txt and Hist\_distances.txt as data files
- (d) Netw\_code\_cluster\_properties.py with NestParCluster.txt as data file. With this same code, properties of the different time frames were filtered, using other input data (per data frame exported from the data base).
- (e) NETW\_Networkanalysis.html with cl\_cl.txt and Cluster\_predictions.txt as data files
- (f) NETW\_Tempralnetwork with afstanden\_opeenvolgend.txt, Nest-Cluster-Day.txt, Cluster\_predictions.txt and Coordinates\_Middle\_clusters\_time\*.txt as data files

### 4. The causes of clustering

- (a) Prediction\_map.html with RegPoints.txt and Nests\_Cluster\_predictions.txt as data files
- (b) Selfish\_herd.html with Nest\_parasiet\_coor\_data.txt as data file

### 5. Raw data and maps for GIS

- (a) DataBase\_Bembix\_rostrata.accdb
- (b) Aerial\_map.tif
- (c) DEM\_RawData.tif
- (d) Insulation6cm.tif
- (e) Slope6cm.tif
- (f) NDVI\_calculated.tif calculated in QGIS with RGB\_NIR\_Raw.tif



