

HUMBOLDT-UNIVERSITÄT ZU BERLIN



Lebenswissenschaftliche Fakultät  
Institut für Biologie  
Arbeitsgruppe Theoretische Biophysik

## Bachelorarbeit

Zum Erwerb des akademischen Grades  
Bachelor of Science

# Territorial Pattern Formation of Wild Pike

# Modellierungsansatz zur Entstehung von Territorien bei wilden Hechten

vorgelegt von

**Felix Fauer**

Gutachten:

**Edda Klipp**

**Robert Arlinghaus**

Berlin, September 2015



## Content

1	Abstract .....	4
2	Introduction.....	5
2.1	Motivation.....	5
2.2	Territory formation in pike .....	6
2.3	Ecological models of territory formation.....	7
3	Methods .....	10
3.1	Pike tracking data.....	10
3.2	Spatial simulation.....	11
3.3	An approach to model territory formation .....	13
3.3.1	Stigmergy and social spacing.....	13
3.3.2	Outcome of this model.....	16
3.4	Parameter adjustment.....	17
4	Results .....	20
4.1	Adaptation .....	20
4.2	Measured territory size depends on fish size.....	24
4.3	Modelled territory size depends on a set of parameters.....	26
4.4	Most parameters are obtained from measurements.....	29
4.5	Results of parameter adjustment .....	31
5	Discussion.....	35
6	Outlook.....	36
7	Appendix .....	37
7.1	Discrepancy with reimplementa-tion of encounter rate.....	37
7.2	Python Libraries .....	39
8	References.....	40
9	Acknowledgement .....	41
10	Declaration of authorship .....	42

## 1 Abstract

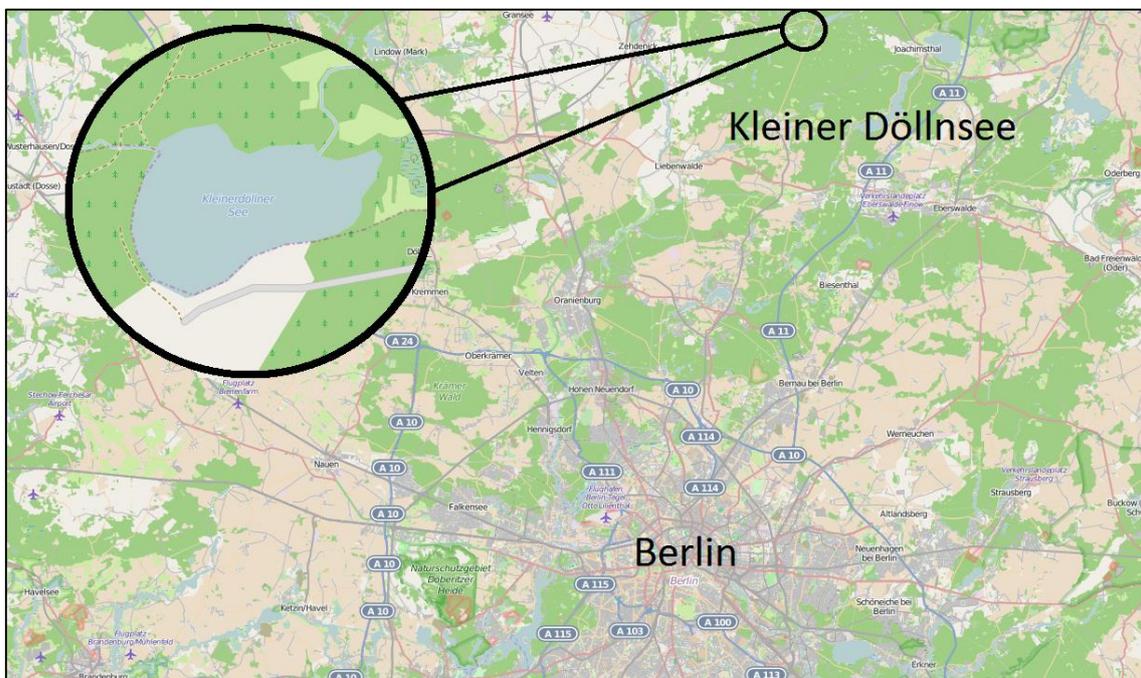
This Bachelor thesis is about the creation of a model which describes the behavior of pike (*Esox lucius*). Pike are cannibalistic loners which form and defend their own territories. The mechanisms that lead to this territory formation are unclear up to now. To model this formation process a scent mediated approach is used where fish respond to each other by encountering scent marks that have been deposited in the past. Pike that meet foreign scent marks return to their own territory with increased probability, depending on their body size ratio. Territory sizes are compared to measured pike tracking data from a small lake near Berlin, Germany to find optimal parameter settings for the model. The resulting model is able to reproduce the measured data and may be used to make predictions about pike behavior in view of different environmental parameters.

Die vorliegende Bachelorarbeit beschreibt ein Modell, welches das Verhalten von Hechten (*Esox lucius*) simulieren soll. Hechte sind kannibalische Einzelgänger, die eigene Territorien bilden und verteidigen. Weitestgehend unbekannt sind jedoch die Mechanismen, die zur Ausbildung dieser Territorien führen. Die Kennzeichnung der Territorien im Modell erfolgt über Duftstoffe, welche über einen gewissen Zeitraum erhalten bleiben und die örtliche Präsenz eines Fisches anzeigen, welche auch zeitlich versetzt noch registriert werden kann. Hechte, die fremde Duftstoffspuren entdecken, ziehen sich in ihr eigenes Territorium zurück, wobei die Stärke dieser Reaktion von der Größe beider beteiligter Fische abhängig ist. Die Größe der Territorien wurde mit gemessenen Daten aus einem kleinen See in der Nähe Berlins verglichen, in dem Hechte mit Sendern versehen und deren Positionen aufgezeichnet wurden. Auf diese Weise wurden Parameter-Einstellungen des Modells optimiert, sodass das Modell in der Lage ist, die gemessenen Daten zu reproduzieren. Zukünftig kann es dafür verwendet werden, Vorhersagen über das Verhalten der Hechte, beispielsweise im Angesicht sich ändernder Umweltbedingungen, zu machen.

## 2 Introduction

### 2.1 Motivation

Lakes are of interest as study object because of their self-contained and natural character. This thesis presents an attempt to analyze whole lake measurement data systematically. Therefore, a reference method [1] of scent-mediated animal interaction has been adapted to the requirements of the given situation in the specific lake so as to find out whether this theoretical model is able to reproduce the data. The obvious question to follow is whether the model furthermore has the power to make predictions about different conditions.



**Figure 1: Position of Lake Döllensee near Berlin, Germany (source: openstreetmap.org)**

Numerous approaches (see chapter 2.3 for a chronological overview of existing models) deal with the problem of modelling animal interactions. I remodeled one existing approach [1] and used an optimization algorithm which helps to adjust the model parameters to the real circumstances. Due to the fact that some parameters cannot be justified by data measurements I looked for an objective function and scanned parameters in order to model the territorial formation of pike which is the main focus of this work.

## 2.2 Territory formation in pike

Pike are one of the most essential carnivore fish in water ecology and play a crucial role in the ecosystem and population rate dynamics [2]. There is already a lot of information about pike, gathered in a review by Raat (1998) [3].

Pike (*Esox lucius*) spawn from April till May. Those time spans can vary for different regions and the named time is a reference for Germany [4]. These oviparous fish start sexual maturity at the age of one year. However not age but rather size is a more important indicator of puberty for pike [5]. Most male pike are at least 19 cm and female 30 cm when reproducing the first time.

It was found that pike tend to be cannibalistic especially when resource availability becomes unsatisfactory. With enough resources pike eat fish of other species but still avoid conspecifics [6]. Size ratio plays a crucial role in behavior of different pikes in encounter situations. Field studies revealed the coexistence of equally sized pike in close areas but almost never pike of different sizes in the same area (about 30-40 m<sup>2</sup>) [7]. It can be concluded that pike show no schooling behavior [3]. Pike that are eaten by conspecifics have averaged one quarter of its predator's size [8] but can also be much bigger up to maximal 70% as recent data from the department of Freshwater Ecology and Inland Fisheries Berlin (group of Robert Arlinghaus) via personal communication showed. Pike that are too small do not provide enough energy and are not in danger to be eaten by bigger pike. Pike diet is dominated by one or two prey species, sometimes more [9]. The presence of pike can be recognized for some time by other fish because of scent marks that are left after eating prey [10]. That is why they often defecate far away from their territory in order not to scare potential new prey [11].

Pike, especially small ones, prefer littoral zones with much vegetation [9] whereas large pike occupy both pelagic and vegetation zones [12]. Large pike also cover wider distances than small ones and are found more often in pelagic zones of lakes [13]. Fish which swim long ways gained weight much faster than resident pike because their chance of prey is bigger [14]. Fish have to take risks and swim out far to gain weight but their risk of being eaten decreases with increasing body length and weight. We have to distinguish between home range and territory although both terms often are

used synonymously. Home ranges cover almost the whole used area where fish feel safe and return to very often and so, home ranges are larger. Burt (1943) defined home ranges as *“that area traversed by the individual in its normal activities of food gathering, mating and caring for young”* [15]. Territories are actively defended by an animal and that is why they have to be smaller than the home range area [15].

Kaukoranta and Lind caught pike, tagged them and released them into the water again. They found that most pike are resident and stay in small limited areas. Seldom pikes are found at positions more than 5 km away from releasing position [16]. The average distance is 500 m. Analysis of activity patterns show that pike stay inactive most of the time [17], especially in littoral zones. Pike of every age prefer definite territories and always return to those areas. This behavior is seen even more often by young pike [18] due to the need to find refuge from predators.

### 2.3 Ecological models of territory formation

A good overview of the recent history of territory formation is given in a review paper by Potts and Lewis [19]. Mainly there are two strands of territorial models: partial differential equations and individual-based approaches.

Ordinary differential equations (ODE) relate functions with their derivative of first or higher order, depending on only one variable. The probably best known ecological ODE is the Lotka-Volterra-equation from 1925 [20]. It describes predator-prey interaction in terms of population size and postulates that there has to be an intrinsic systematic balance. A negative feedback following a change in predator or prey population makes sure that usually neither species is in acute danger of extinction.

In order to add a spatial component, we have to change from ODE to partial differential equations (PDE). PDEs depend on two or more variables. An example is the heat equation which describes the temporal and spatial distribution of heat in a certain region.

One approach of ecological modeling is to investigate the influence of individual behavior on patterns in the population. First who tried to model this microscale-to-macroscale approach were Lewis and Murray [21]. They investigated the occurrence of deer as an answer on the behavior of wolf packs und developed a PDE model which

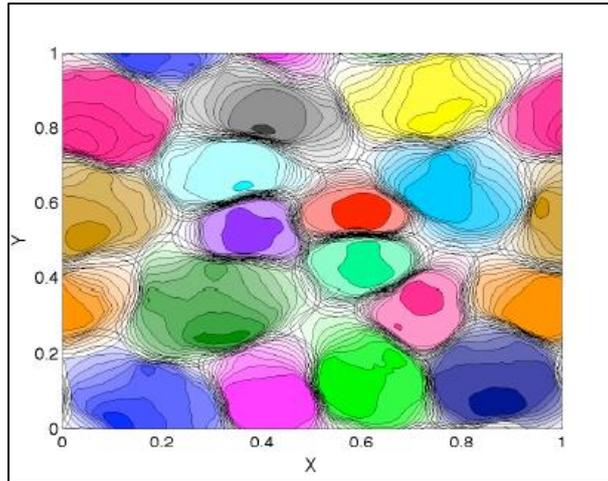
shows how deer as wolves' prey choose to stay most likely in so called buffer regions between different wolf packs. There the occurrence of wolves is less likely and so the buffer zones serve as refuge for deer. Lewis and Murray already use scent-mediated interaction. Their model's assumptions are based on real tracking data but are not compared to it systematically as it is intended in this thesis's work.

A more fine grained way than differential equations are individual-based models (IBM) where each agent is modeled. This approach is advantageous if individuals and their behavior are of interest, so distinct territories can be modelled.

In 2011 Giuggioli, Potts and Harris [22] added timescale observations to this often studied interaction mediated by olfactory signals and containing conspecific avoidance using IBM. They recorded and calculated the mean square displacement of territorial boundaries depending on the time after a territory began to form.

Two years later in 2013 Potts and Giuggioli extended their model from 2011 and presented their findings in a paper which describes animal's stigmergy and social spacing [1]. Marsh and Onof define stigmergy as "*indirect communication mediated by modifications of the environment*" [23]. This is done by scent marks deposited by animals which cause indirect reactions of other individual's. Although this idea is not new, given the fact that Lewis and Murray already have used scent mark mediated reactions, they enriched this scientific field by analyzing and investigating the correlation between encounter rate of animals and individual avoidance behavior [1].

In Potts's model animals are random walkers on a two dimensional lattice with discrete fields. Periodic conditions were implemented so that animals that leave the mesh on one side immediately reappear on the opposite side of the mesh (see figure 2). So no boundary conditions have to be considered. The animals which are supposed to make a step and the duration of a step are drawn randomly by a Gillespie algorithm [1]. Since all animals have the same propensity values it is a simplified version of Gillespie.



**Figure 2: Utilization plot of stigmergy model. This figure is taken from Potts's and Giuggioli's paper [1]. Contour levels show the density of animal occurrence for this area. Animals that are going to make a step outside the whole area, reappear on the corresponding opposite side (periodic conditions).**

Animals leave their scent marks on every field they occupy. The scent mark is active for a certain time  $T_{AS}$ . The whole model is unitless so  $T_{AS}$  represents a number of time steps. In the time between deposition and inactivation the scent mark's strength is decreasing until it fades away completely. The mean value of the last  $T_{AS}$  positions is called centroid, representing the territory area center. The centroid has to be updated with every step.

Animals are random walkers only until they discover a field where another animal's scent mark is still active. If they enter such a field they will show retreating behavior. That means that there is a higher chance of walking in direction of its centroid than walking in another direction [1].

Potts and Giuggioli's model suits our requirements well for different reasons. Animals are seen as autonomous individual agents. Territories form just because of their retreating behavior when animals encounter foreign scent marks. Movement is an important part of the model because its dynamic influences the model's results enormously. Furthermore it corresponds well with the data we have, because both model and measured data produce a list of time courses where positions of fish have been recorded and can be analyzed.

## 3 Methods

### 3.1 Pike tracking data

There are several experiments investigating pike behavior. One of the first who investigated fish with modern transmitting technique were Hasler et al. (1969) [24] who tracked white bass (*Roccus chrysops*) with ultrasonic devices. Their main interest lied in the swimming speed and general behavior of fish. They found that white bass often swam directly towards their spawning ground instead of random movement.

Poddubnyi, Malinin and Gaiduk [25] (1970) were the first ones who tracked northern pike. They found territorial patterns among pike although their tracking time was only 50 to 60 hours.

Diana (1977) [26] however doubted the existence of home ranges and argued the tracking time of Poddubnyi and Malinin being too short to make a statement about the existence of home ranges. He tracked pike himself over a time of 47 days and proposed that no home ranges were established by pike.

In 1980 Chapman and Mackay tracked pike and investigated the influence of meteorological factors on pike behavior. As Diana did not determine distribution of habitat types as vegetation and free water, they started a new experiment. There seems to be no correlation between rain and habitat behavior. Furthermore they found that pike stayed close to the shore on sunny days and went to pelagic zones on windy days. Pike's versatility was greater than expected which Chapman and Mackay proposed as an important feature of behavior of top predators [27].

Many years later in 2001 Jepsen et al. radio tagged pike over a time period of 9 months but analyzed data only fragmentary over chosen times of interest. He investigated diel (24-hour-periods) activity patterns, home range sizes and habitat utilizations. He confirmed the existence of pike home ranges. Also he compared behavior of pike between an artificial reservoir and a natural lake and found that in summer pike stayed closer to vegetation, which can be found mainly in the littoral zones, than in winter. Moreover he discovered an equal level of movement during a day period whereas the movement level changes significantly over different times in the year [12].

The group around Robert Arlinghaus started a project in 2010 at a 250 ha lake near Berlin, Germany, called Kleiner Döllnsee, where fish in large numbers and of different species were tagged and their positions recorded every 25 seconds. Characteristics of each fish were documented before releasing them into water again. Those characteristics are length, weight and sex. As part of the experiment fish from other nearby lakes were transferred to Kleiner Döllnsee to investigate the difference in behavior depending on whether fish swim in familiar environment or not. For this thesis all tracked pike were considered and handled equally.

Due to technical restraints the data at some sections is very noisy. Therefore it was preprocessed to get time courses for each fish.

### 3.2 Spatial simulation

As an alternative to PDEs it is possible to model every reaction individually. For this purpose a so called chemical master equation (CME) exists [28]. This equation describes the probability  $b$  of reaction  $R_\mu$  being the next reaction to happen at time point  $t + \tau$  as the product of the probability that nothing else happened before and the probability that the next reaction will be  $R_\mu$ .

Because of its complexity and its character as an ODE there is usually no distinct solution for the CME. Nevertheless, in 1977 Gillespie [28] developed a method in order to choose one possible trajectory of the reaction system according to the possibilities given by the CME. He considered fluctuations and the possibility to integrate rather complex conditions. The crucial difference to the CME is the absence of infinitesimal small time intervals. Instead there is a stochastic choice of reaction and time point.

In Gillespie's description there are two important values: After which time interval  $\tau$  will the next reaction occur and which reaction  $R_\mu$  of all possible reactions ( $R_1, \dots, R_M$ ) will it be? Therefore, we introduce two stochastic values  $r_1$  and  $r_2$ .

The probability  $b$  of reaction  $R_\mu$  occurring in time interval  $(t + \tau, t + \tau + d\tau)$  is  $b = g_\mu * d\tau$ , where  $g_\mu$  is the propensity (=“how likely is a reaction to occur per unit time” [29]) of reaction  $\mu$  and  $g_0$  is the sum of all propensities  $g_\mu$ . Therefore, Gillespie deduced the following formula [28]:

$$b(\tau, \mu) = g_\mu * e^{-g_0\tau} \quad (1)$$

That is the probability density function  $b$ . It describes the probability density that the next reaction will be  $R_\mu$  and will happen at time point  $t + \tau$ . When determining the time step we are not interested in one reaction in particular but all  $M$  possible reactions together so we substitute  $g_\mu$  by  $g_0 = \sum_{\mu=1}^M g_\mu$ . After integration we get the probability distribution that the next reaction occurs in time interval  $(t, t + \tau)$ .

$$B(\tau) = 1 - e^{-g_0\tau}$$

under the condition that for  $\tau \rightarrow \infty$  the value for  $B$  has to become 1. However, we are looking for the probability that the next reaction does *not* occur in  $(t, t + \tau)$  and so we take  $\bar{B} = 1 - B(\tau)$  because that is the time span we want to skip.

$$\bar{B}(\tau) = e^{-g_0\tau}$$

With given probability  $\bar{B} = r_1$  and unknown time step  $\tau$  we obtain:

$$\tau = \frac{1}{g_0} * \ln\left(\frac{1}{r_1}\right) \quad (2)$$

for a random number  $r_1$  ( $0 < r_1 \leq 1$ ). That will be our chosen time step  $\tau$ .

Henceforth, we choose the reaction that will happen. Therefore we have to draw a second random number  $r_2$  ( $0 < r_2 \leq g_0$ ) and the corresponding reaction  $\mu$  such that

$$\sum_{i=1}^{\mu-1} g_i < r_2 < \sum_{i=1}^{\mu} g_i \quad (3)$$

holds true [28]. The whole algorithm follows this scheme:

1. Initialize values for all propensities and parameters.
2. If necessary, update individual propensities.
3. Select time step and reaction and update them.
4. Jump to step 2 until time is over or no molecules are left.

In our model there are no chemical reactions. Instead of choosing a reaction we always choose the fish that's position will be updated next. Propensities of fish are proportional to their measured swimming speed depending on body length (size). It is the product of swimming speed and its activity. Additionally the propensity rises by a fixed factor if the fish is swimming outside of the vegetation region (pelagic zone).

One-dimensional diffusion constants  $D_\mu$  are brought into Gillespie algorithm by the following converting formula as Hepburn et al. showed [30]:

$$g_\mu = \frac{D_\mu}{\Delta x^2} \quad (4)$$

$\Delta x$  is the mean width of one step. For the mean square displacement  $\langle \Delta x \rangle^2 = 2D\Delta t$  holds true because in one dimension there is always the chance to go in one of 2 possible directions with a  $g_\mu$  for each direction. A second dimension would alter this formula to  $\langle \Delta x \rangle^2 + \langle \Delta y \rangle^2 = 4D\Delta t$  because in our lattice grid there are always 4 possible directions to go. That is why in two-dimensional models  $\tau$  is calculated [31] as:

$$\tau = \frac{1}{4 * \frac{D_\mu}{\Delta x^2}} * \ln\left(\frac{1}{r_1}\right) \quad (5)$$

### 3.3 An approach to model territory formation

This chapter explains the model from Potts and Giuggioli, which we call general stigmergy model in the following. It was reimplemented in the Python programming language.

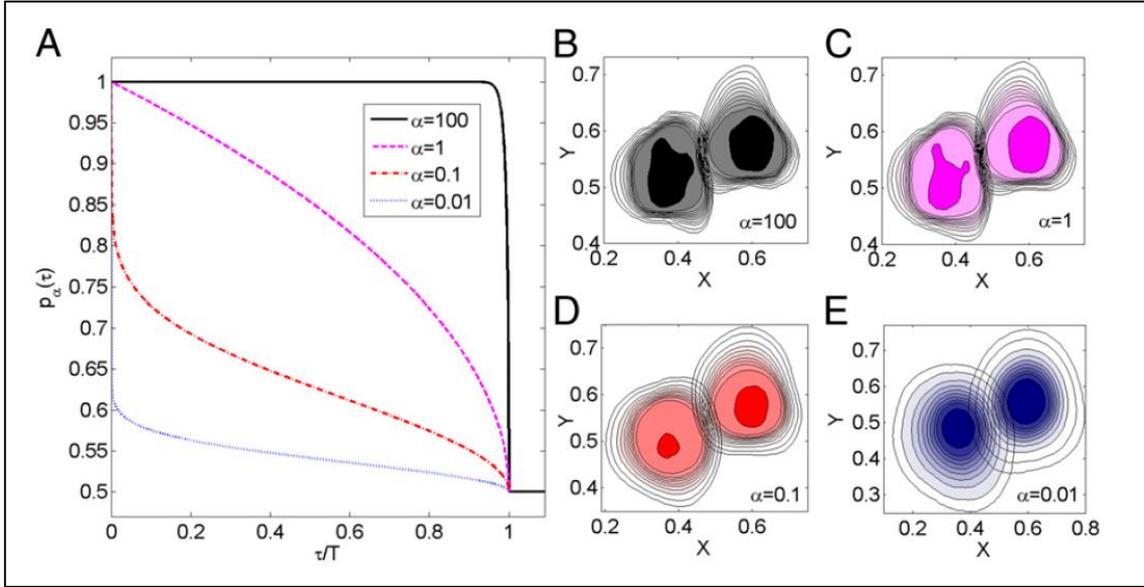
#### 3.3.1 Stigmergy and social spacing

If one of the fish meets a foreign scent mark, it shows retreating behavior towards its centroid. The retreating tendency depends on the current scent mark strength. The decrease in scent mark strength is described by the age  $\tau$  of the mark in relation to the total time the scent mark stays active  $\tau/T_{AS}$ . An additional global parameter  $\alpha$  determines how the age of the scent mark influences the probability  $p$  of retreating behavior.  $\alpha$  is called the degree of stigmergy.

$$p_\alpha(\tau) = \frac{1}{2} + \frac{\sqrt{1 - \left(\frac{\tau}{T_{AS}}\right)^\alpha}}{2} \quad \text{if } 0 \leq \tau \leq T_{AS} \quad (6)$$

$$p_\alpha(\tau) = 0.5 \quad \text{if } \tau > T_{AS} \quad (7)$$

The influence of  $\alpha$  on the probability of retreat  $p$  is shown in figure 3A as well as a contour level plot (B-E) (from now on: utilization plot) which describes the density distribution of animals.



**Figure 3: Influence of  $\alpha$  on probability of retreat  $p_\alpha$ .** This figure is taken from Potts's and Giuggioli's paper [1]. **A:** For high  $\alpha$  the retreating probability stays high even for old scent marks ( $1 - \tau/T$  small). **B-E:** With decreasing  $\alpha$ , boundaries alter from sharp to soft because animal reactions become weaker.

There are different probabilities to make a step in any of the four directions: *up* ( $u$ ), *down* ( $d$ ), *left* ( $l$ ), *right* ( $r$ ). In absence of a foreign scent mark each probability has the same value and a retreat is as likely as no retreat.

$$u = d = l = r = 0.25$$

$$p = 0.5 \quad (8)$$

When encountering a foreign scent mark probabilities for each direction become different and the value for  $p$  increases as shown in figure 4. The following formulas describe how different probabilities are calculated in this case. Following information is required as input: Current position of the animal ( $x, y$ ), position of the centroid ( $x_c, y_c$ ) and the retreating probability  $p$ .

$$u = \frac{1}{4} * \left[ 1 - (2p - 1) * \frac{y - y_c}{\sqrt{(x - x_c)^2 + (y - y_c)^2}} \right]$$

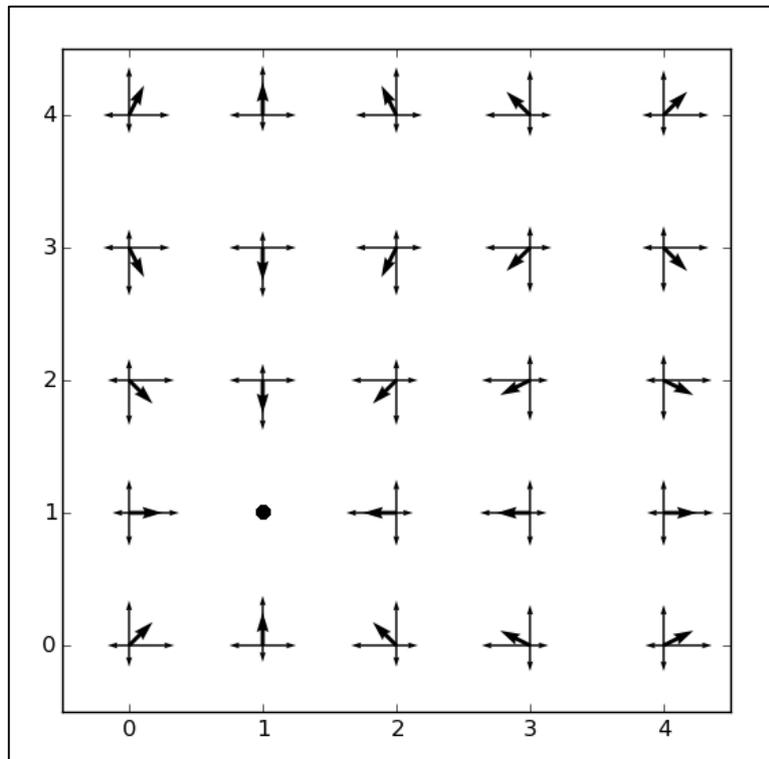
$$d = \frac{1}{4} * \left[ 1 + (2p - 1) * \frac{y - y_c}{\sqrt{(x - x_c)^2 + (y - y_c)^2}} \right]$$

$$l = \frac{1}{4} * \left[ 1 + (2p - 1) * \frac{x - x_c}{\sqrt{(x - x_c)^2 + (y - y_c)^2}} \right]$$

$$r = \frac{1}{4} * \left[ 1 - (2p - 1) * \frac{x - x_c}{\sqrt{(x - x_c)^2 + (y - y_c)^2}} \right] \quad (9)$$

Note that in order to consider periodic conditions, some additional calculations regarding those formulas have to be made. The shortest way to the centroid sometimes crosses the borders of the defined mesh. This is built in by transferring the centroid to the corresponding periodic position outside of the mesh if the distance from current position to centroid is longer than half the mesh's side length.

To clarify the meaning of those formulas for the simulation and their constraint there is a quiver plot in figure 4 which shows the different probabilities for each direction and position with a given centroid and a given and constant  $p$ -value.



**Figure 4: Quiver plot of returning direction probabilities. The centroid of this example animal is at position (1, 1). Narrow arrows represent the probabilities for each of the 4 possible directions, according to formulas 9. Thick arrows are the sum of all 4 direction probabilities. They always point exactly towards the centroid and follow the shortest path (periodic conditions).**

Potts and Giuggioli introduced a parameter  $Z$ , called the spatial competition.

$$Z = 4DT_{AS}\rho \quad (10)$$

It contains the diffusion constant  $D$ , the maximum active scent time  $T_{AS}$  and the density  $\rho$ . Since in the stigmergy model the diffusion constant in Pott's model is always

fixed to  $D = 0.25$ ,  $Z$  can be simplified to  $Z = T_{AS} * \rho$ . Increasing the  $Z$ -Parameter is supposed to increase the interaction pressure between animals. With higher  $T_{AS}$  scent marks stay active for a longer time which increases the occurrence of retreating behavior. A higher density  $\rho$  makes the animals move closer due to a smaller area on which the system operates.

The density  $\rho$  is not defined more precisely in Potts and Giuggioli's paper. Because the number of animals  $n$  in all calculations is set to a fixed value, I assume that a variation of  $\rho$  only leads to a change in mesh size. The density  $\rho$  then needs to be negative proportional to the area size of the mesh. That is why area size  $A$  and mesh side length  $s$  are calculated as following:

$$\begin{aligned} A &= n/\rho \\ s &= \sqrt{n/\rho} \quad (11) \end{aligned}$$

In formula 5 (page 13) the time for each jump step is described.  $\Delta x$  equals one lattice site length which is 1. Since  $D = 0.25$  holds true, the time  $\tau$  for each jump is drawn from a random number  $r_1$  via:

$$\tau = \ln\left(\frac{1}{r_1}\right) \quad (12)$$

### 3.3.2 Outcome of this model

An important characteristic of the system is the average encounter rate which tells us, how often animals meet each other.

To make sure that territory sizes do not change with time, the simulation contains a burn-in-time after which the counting process starts. Table 1 shows the parameter set of Potts's model that was used to show the encounter rate with respect to the  $\alpha$ -value for different  $Z$ . For each  $Z$  the values for  $T_{AS}$  and  $\rho$  density are listed separately.

parameter	variable term	value
Time steps	$t_{total}$	110,000
Burn-in-time	$T_{burn}$	10,000
Competition parameter	$Z$	(9, 40, 122, 400)
Active scent time	$T_{AS}$	(2000, 4000, 6000, 10000)
Density	$\rho$	(0.0044, 0.01, 0.0204, 0.04)
Number of animals	$N$	25

**Table 1: Setting for encounter rate calculation [1].**

Figure 5 visualizes the encounter (average of each individual's encounter with other animals) rate with the given parameters. It is normalized to the number of animals, density and number of time steps.

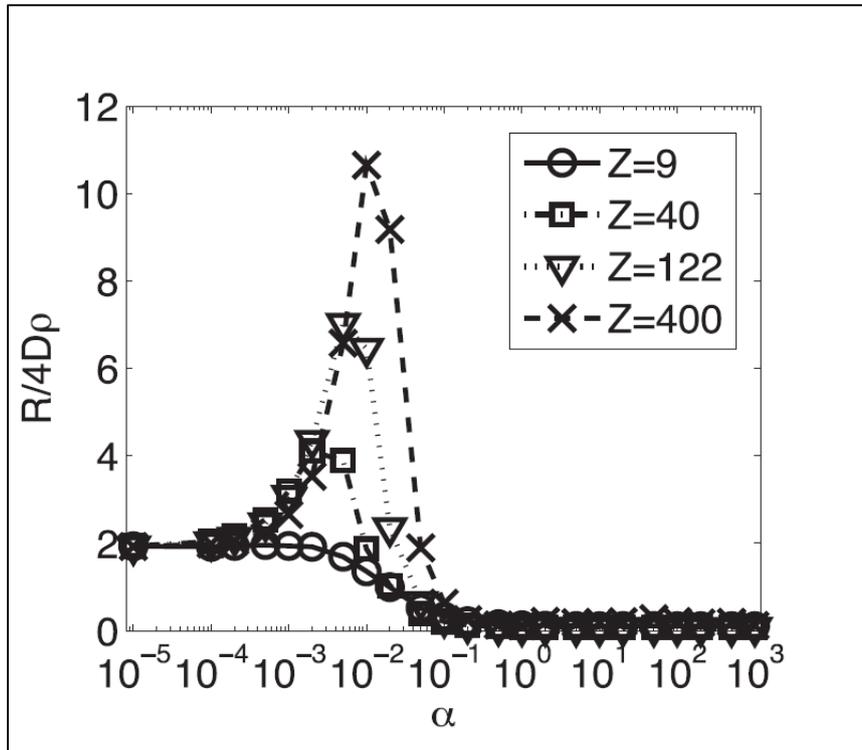


Figure 5: Encounter rate as calculated from Potts and Giuggioli (log-scale on x-axis). This figure is taken from their paper [1]. Encounter rate is normalized to diffusion constant  $D$ , density  $\rho$  and number of steps. For different  $\alpha$ -values the encounter rate changes non-monotonously.  $Z$  describes the competition between animals. For high  $Z$ -values the encounter rate is higher for  $\alpha = 10^{-2}$ . For high  $\alpha$ -values the encounter rate becomes small because animals turn around whenever they meet a foreign scent mark. The chance of encounter at its minimum.

Potts and Giuggioli explain the appearance of the peak around  $\alpha = 10^{-2}$  as follows: With increasing  $\alpha$  the number of animals per marked area (overlap) decreases more steeply than the average territory size does. The ratio of both those quantities, the local population density, increases in this region. The population density scales with the encounter rate, so the encounter rate increases around  $\alpha = 10^{-2}$  [1].

### 3.4 Parameter adjustment

In our model (for information about differences to Potts's model see chapter 4.1) it is unlikely that only one exact parameter setting is appropriate to reproduce original data. The main focus will lie on the  $\alpha$ -parameter and the active scent time  $T_{AS}$ . I expect that low  $\alpha$  values and little active scent times will make animals swim nearly randomly

causing very large territories, larger than those of fish in the reference lake. High  $\alpha$  values and long active scent times however, probably will give rise to sharp boundaries between the territories with almost no overlap which is not coherent with the measured data either.

There are several methods for an automated optimization process, mostly aimed at finding best results. Our project however does not necessarily have one distinct result in this two-dimensional optimization process. For example the modelled territory size could be equivalent to the measured data for small  $\alpha$  in combination with huge  $T_{AS}$  and also for high  $\alpha$  with little  $T_{AS}$  values. The parameters  $\alpha$  and  $T_{AS}$  have the ability to compensate each other when increasing one and decreasing the other. So they are non-identifiable. Because of its stochastic character there are statistic variations which make it difficult for optimization methods to find good results. Work effort would be disproportionate to the saved time because the resolution of model results is not high enough for automated methods. Therefore I forgo an automated optimization and choose a manual parameter scanning procedure instead.

The following parameters can not be obtained from measured data and need to be adjusted:

- number of time steps  $t_{total}$
- degree of stigmergy  $\alpha$
- active scent time  $T_{AS}$

As an objective function we chose the territory size in general for a coarse grained overview and the territory size depending on fish body size for detailed analysis. Territory size is a suitable quantity as objective function because it is easy to calculate with comparatively little computing time and a very crucial property of our model since we want to investigate territory pattern formation.

First I want to find out how many time steps are needed to get steady states in our model. Steady state means in this case, that territory sizes do not change essentially even when modelling time is further increased.

In our model (see chapter 4.1) Gillespie propensities are different for each fish, depending on fish size. This is a crucial difference and it enables the investigation of different behavior of pike according to their size.

Due to the dependency of degree of stigmery  $\alpha$  and active scent time  $T_{AS}$ , I will not optimize both quantities separately. In a two-dimensional function ( $\mathbb{R}^2 \rightarrow \mathbb{R}$ ) I will vary  $\alpha$  and  $T_{AS}$  each in one dimension and determine the value of our objective function given the corresponding two settings for  $\alpha$  and  $T_{AS}$ . For detailed analysis of territory size with respect to fish size I would need a third dimension regarding the body length (size) of fish. Fish are categorized in  $c$  size classes ( $l_1, l_2, \dots, l_c$ ) with each  $l_i$  representing a size range. Size range  $l_i$  always begins where  $l_{i-1}$  ends.

To simplify the optimization process I reduce this third dimension by just calculating the sum of squares  $q$  for each setting  $s$  between original data and modeled territory sizes for every size class  $l_i$ .

$$q_s = \frac{1}{c} * \sqrt{\sum_{i=1}^c (\bar{x}_i - \bar{m}_i)^2} \quad (13)$$

Where  $q_s$  is the sum of squares,  $c$  is the number of size classes,  $\bar{x}_i$  is the modeled average territory size of all fish within size range  $l_i$  and  $\bar{m}_i$  is the measured average territory size of all fish within size range  $l_i$  from original data. Low values of  $q_s$  indicate a good setting where territory sizes are similar to the measured data from the reference lake.

The territory size of the lake data is calculated (Max Flöttmann) by determining the surface area of a minimum convex polygon with a 50% - kernel. That means that the polygon is chosen in a way, such that 50% of all measured data points are inside the polygon. This way we try to remove all data points which represent the fish's position at a certain time but do not belong to its home range where it stays often. Home range size calculation in the model was done in a different way. No kernel has been used there. Just all different points that were occupied by a fish, except from those within burn-in-time, were counted. This is due to the fact that the model is not supposed to simulate the phase which leads to the steady state but just the steady state itself (see chapter 4.1). So, there is no difference in home range size and territory size made in

our model. The model's plotting method however, does not show data points where the fish has been seldom, which is similar to a kernel-method.

The simulation requires a lot of memory and running time so that we were constrained to introduce some simplifications. It is infeasible to simulate the whole lake over a sufficient time period. That is why we reduced the lake dimensions and also the number of fish to maintain an equal density (fish per area). Number of fish is reduced by the reduction factor  $k$ . The area was decreased by the same factor which means that each side length was decreased by  $\sqrt{k}$ . Territory sizes are multiplied by  $k$ , so that they are normalized to the original lake size.

## 4 Results

### 4.1 Adaptation

Our main goal is to find a way to predict the whole system behavior. If there is a chance of varying just microscale reactions affecting the whole system, additionally this model will give us the possibility to make assumptions about behavioral patterns of wild pike.

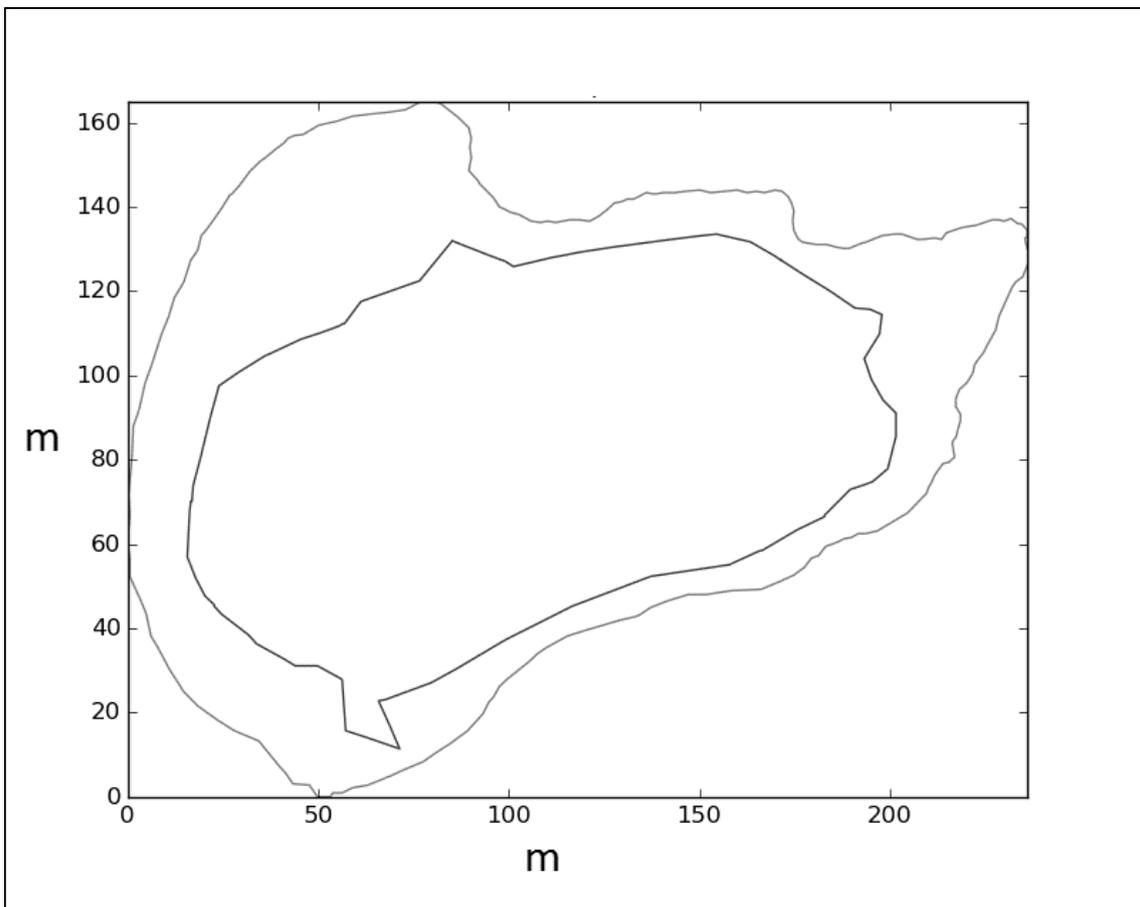
Potts and Giuggioli's model suits our requirements well for different reasons. Animals are seen as autonomous individual agents. Territories form just because of their retreating behavior when animals encounter foreign scent marks. Movement is an important part of the model because its dynamic influences the model's results enormously. Furthermore it corresponds well with the data we have, because both model and measured data produce a list of time courses where positions of fish have been recorded and can be analyzed.

The results from our reproduction of the encounter rate with the given data from the paper look different because we cannot reproduce the characteristic peak (Figure 5, see Appendix in chapter 7 for further information). But still, we decided to use the model because we think that the main idea was implemented correctly.

To adapt the model to our needs we had to modify it substantially. Periodic conditions are not useful in a real lake and have been substituted by natural blocking boundary conditions. Now if fish are supposed to jump to a position outside the lake, their

randomly chosen step will be calculated again until it is inside. Boundaries are determined by the original shape of Lake Döllnsee. The lake's dimensions have a direct impact on the systems behavior because the available space scales with lake area size in square. The chosen resolution is  $r \approx 1 \text{ m}^2$  per lattice site.

There is vegetation at the littoral zone of our lake. I needed to implement it in our model because fish show different behavior in the pelagic zones as mentioned in the introduction. The littoral zone is laid over the whole lake as another layer (shown in figure 6).



**Figure 6: Schematic picture of Lake Döllnsee and its pelagic zone (inside the inner polygon).**

Especially small but also bigger fish tend to form territories in the lake's pelagic zone with the vegetation region. We consider this by introducing different weighting factors for the calculation of the fish's centroids. All occupied fields inside the vegetation region are weighted higher by a factor  $w_{veg} > 1$ , when calculating the centroid. All fields in the open water where no vegetation can be found are weighted normally ( $w_{veg} = 1$ ). A fish that has been in the littoral zone at least once and is forced to

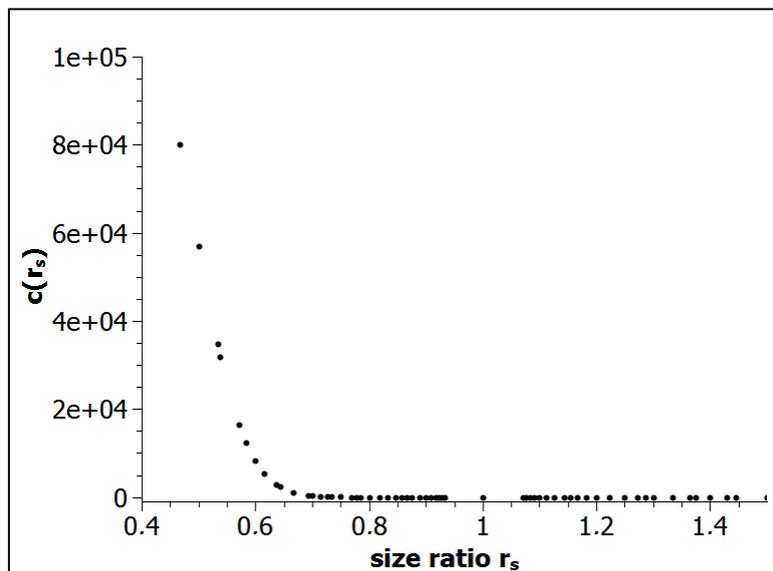
retreat to its centroid – which only happens if it meets another foreign scent mark – will be slightly shifted to the littoral zone.

As mentioned in the introduction pike show different behavior according to their body size. Since pike are cannibalistic, small fish usually keep a distance to bigger fish. Size dependence is considered in our model by a change in the  $\alpha$ -value. We assume that fish can recognize the size of their interaction partner which deposited the encountered scent mark. A function  $c$  determines by which factor the standard value  $a_0$  is raised if one fish meets a scent mark of a larger fish with respect to the size ratio  $r_s$ :

$$a_{r_s} = a_0 * c(r_s) \quad (14)$$

$$c(r_s) = A * 2^{-\left(\frac{r_s - \mu}{\sigma}\right)^2} + 1 \quad (15)$$

Parameter  $A$  indicates the maximum increase of  $\alpha$  whereas  $\mu > 0$  shifts the local extremum to the right.  $\mu$  should be equal to the size ratio at which the likelihood to be eaten is at its peak.  $\sigma$  influences the function's vertical dilation or compression. The new value of  $\alpha$ , which is  $\alpha_{r_s}$ , will be applied once to this encounter and set back to  $\alpha_0$  immediately afterwards. Figure 7 shows how  $\alpha$  changes for different size ratios  $r_s$ .



**Figure 7: Raising factor  $c$ .  $\alpha$  will be multiplied by  $c$ . Fish react on other fish depending on their own size compared to the other fish's size. Since fish size ranges from 35 to 70, there is no size ratio smaller than 0.5. For size ratios bigger than 0.75 there is no raise in  $\alpha$  (raising factor = 1) because pike show no predatory behavior. Used parameters are:  $A = 10^5$  (maximum  $\alpha$  raising factor),  $\mu = 0.41$  (size ratio for maximum  $\alpha$  raising factor),  $\sigma = 0.1$  (vertical dilation of this function)**

Parameters are chosen such that fish can be eaten from size ratio of  $r_s = 0.3$  up to  $r_s = 0.7$  [32].

Bigger fish should swim faster than smaller fish (for speed data see Table 3, chapter 4.4). In addition to that, fish are supposed to have smaller swimming speed in the pelagic zone. That's why I need to modify our Gillespie simulation. There are several methods to realize differences in swimming speed:

Currently if a fish is supposed to make a step, it chooses one of its neighboring lattice sites to enter. A fish with higher swimming speed could choose a field to enter with a certain distance bigger than one. I tried this possibility and found that it is very disadvantageous because the model's idea is based on solid borders around a territory. When animals can jump from one lattice site  $s_1$  to another one  $s_2$  without entering the lattice sites between  $s_1$  and  $s_2$  gaps in the border line arise. There is the possibility that fish ignore each other even when they technically cross routes. That means other animals can pass these gaps without showing retreating behavior because no scent mark is deposited on these fields.

A more accurate way is to give fish different propensities. Fish with higher propensity are drawn more often by the Gillespie algorithm which is equivalent to a higher swimming speed. The propensity  $g_f$  of a fish  $f$  with size  $s_f$  and position  $pos_f$  is calculated as follows:

$$g_f = g_{size}(s_f) * g_{veg}(pos_f) \quad (16)$$

The standard propensity for fish  $g_f$  is multiplied by a factor  $g_{size}$  according to the fish size and by a second factor  $g_{veg}$  according to the position in the lake depending on whether it is in the littoral or pelagic zone.

The original stigmergy model [1] had a built-in diffusion constant for each animal. Diffusion constants describe the average mobility of molecules and are deduced from the mean square covered distance per time. Diffusion constants do not play an essential role in this implementation because, the movement speed is based on propensities (see formula 16).

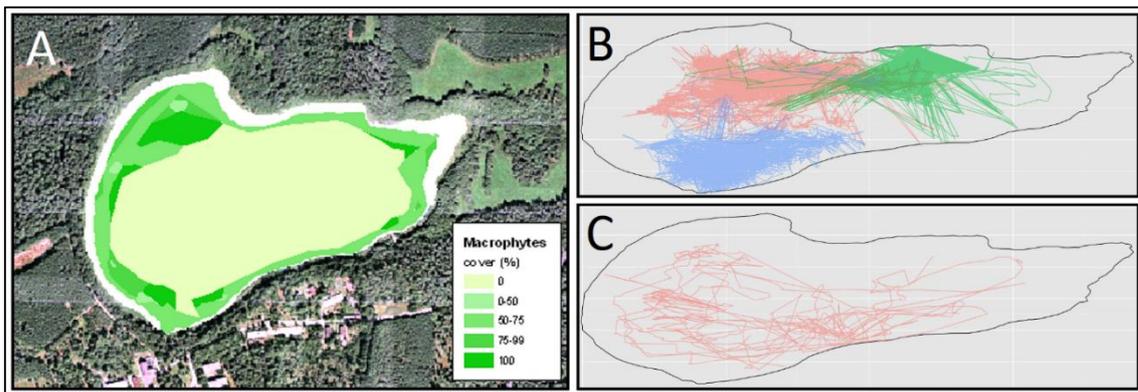
Time steps  $t_{total}$  in our model are scaled so that one step accords with one second, by adjusting propensities correspondingly. However, this unit has to be handled with care

because the model does not necessarily simulate the forming process of territories but rather the finite state which shows complete home ranges and territories.

The burn-in-time in our model is fixed to 10% of the whole simulation time. Moreover I do not distinguish between male and female pike and do not consider the age of fish.

## 4.2 Measured territory size depends on fish size

There is much information about Kleiner Döllnsee provided by the group of Robert Arlinghaus. The lake has a size of 250,000 m<sup>2</sup>, a mean depth of 4.1 m and a maximum depth of 7.8 m [33]. Its shape is available as polygon data points (Klefoth and Kobler, 2007). We have very detailed knowledge about the macrophyte coverage (figure 8A) but simplified it for our model into Boolean values (True: coverage > 0%) and got a second polygon as result (Figure 6).



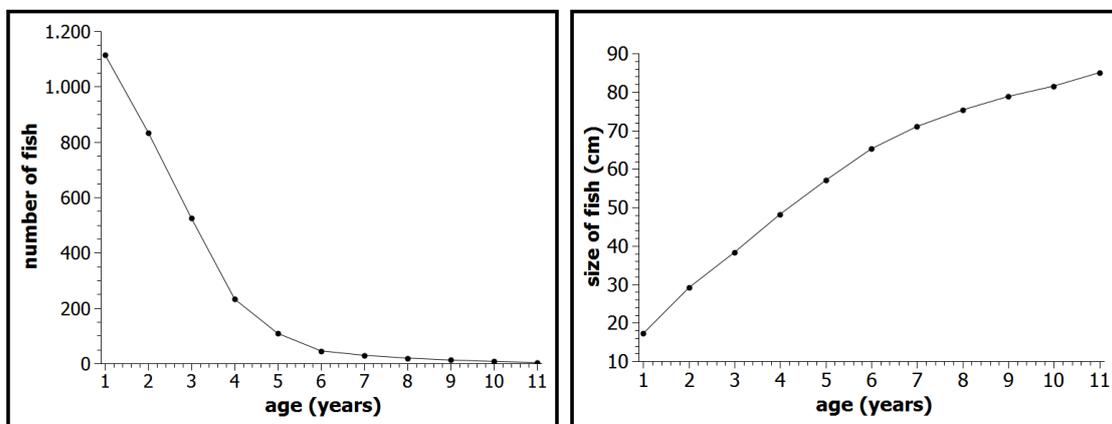
**Figure 8:** Excerpt from data provided by the group of Robert Arlinghaus. **A:** Macrophyte coverage of Lake Döllnsee (source: presentation by David March-Morlà). In our model I used only the one polygon which represents the 0% - threshold of coverage. **B:** Trace plot of three different fish, domiciled in Lake Döllnsee. **C:** Trace plot of a fish that was transferred from Lake Wuckersee, near Berlin, Germany. It seems that foreign fish need to explore the whole available area first before forming territories.

It is found that foreign transferred fish need more time to find territories than fish that have spawned in Lake Döllnsee (figures 8B and 8C). However, those fish domiciled to the lake still often swim far distances before finding areas in which they stay for most of the time, which we call home ranges. Of all 146 tracked fish, 43 have spawned in Lake Döllnsee. Table 2 shows the abundance of pike in the lake. Furthermore it is visualized by figure 9.

Age	Mean length	Number
1	17.2 ± 3.6	1115
2	29.1 ± 5.3	834

Age	Mean length	Number
3	38.3 ± 6.2	524
4	48.2 ± 6.9	232
5	57.0 ± 6.9	108
6	65.2 ± 6.3	45
7	71.0 ± 4.7	29
8	75.3 ± 3.3	19
9	78.8 ± 2.8	12
10	81.5 ± 3.1	7
11	85.0 ± 1.7	3

**Table 2: Abundance of pike population and mean body length for different age classes (data provided by the group of Robert Arlinghaus). Of all pike only those of size 35 – 70 cm are modelled because there is no tracking data available from smaller or bigger pike.**



**Figure 9: visualization of abundance (Table 2), number of pike in Lake Döllnsee, their age and size. A: There are many young pike and very few old pike. B: size of pike correlates with age.**

Territory size  $h$  is correlated with body size of pike. Figure 10A shows the size of territories depending on fish body size. Figure 10B shows the same data but body sizes of fish have been binned to nearest lower multiple of 5 cm, so that the correlation can be recognized more clearly.

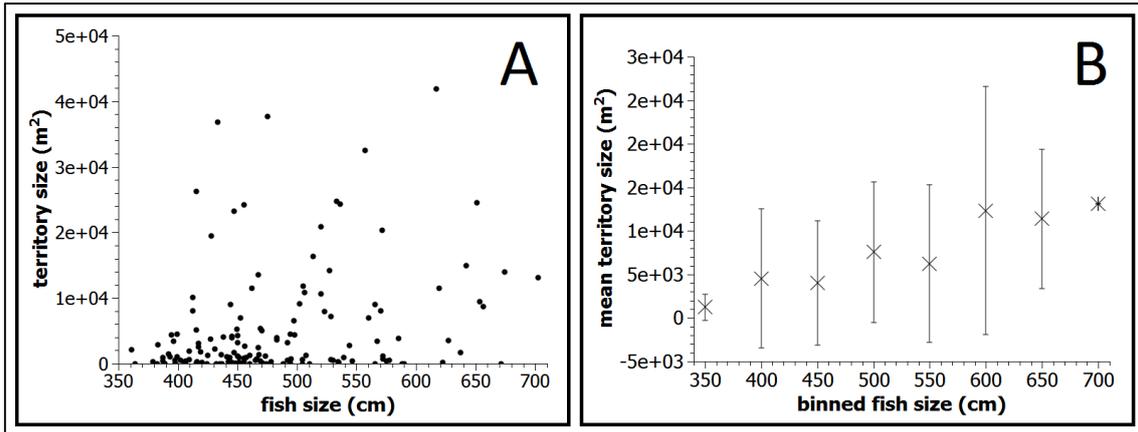


Figure 10: Correlation of fish's body size ( $cm$ ) and measured territory size ( $m^2$ ). A: Raw data of territory size depending on fish size. B: All body sizes were rounded to the next lower multiple of 5. Then the mean of all territory sizes for each body size class has been calculated. Vertical lines represent the mean standard displacement. This illustration provides a better demonstration of the correlation between both quantities.

### 4.3 Modelled territory size depends on a set of parameters

Due to memory restrictions I had to reduce the amount of fish as well as the lake's dimensions by the reduction factor  $k$ . The influence of this measure is shown by confronting both ways in figures 11A and B. Rough extends of territory pattern remain constant. I assume that territory size is not influenced significantly as long as the lake's boundaries do not have a high impact on limiting territory dimensions on one territory flank. Since territories should be much smaller than the lake's dimensions this measure can be implemented without corrupting the results too much. The density remains constant. The calculated territory size is normalized by multiplying it by  $k$ .

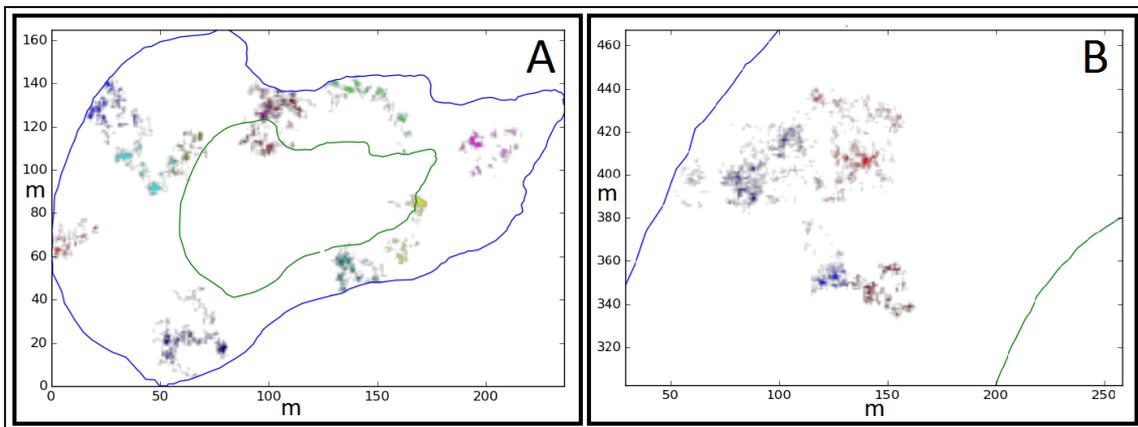


Figure 11: Demonstration of the reduction factor's influence. A: reduction factor  $k = 10$ . The lake's area is 10 times smaller than the original lake. B:  $k = 1$  (no reduction). There is a 10-fold zoom applied to the plot to make both pictures (A+B) comparable. No change in area

patterns in comparison of A and B is visible. Territory Size and extend shapes seem not to be influenced significantly. Density remains constant.

Territory patterns are mainly defined by two values: The degree of stigmergy  $\alpha$  and the active scent time  $T_{AS}$ . High  $\alpha$  and  $T_{AS}$  cause sharp boundaries between territories and lead to almost no overlap so that the whole available space is split to all fish in different extends depending on size and initial conditions. Low values for  $\alpha$  and  $T_{AS}$  on the other hand, cause large overlap of territories because animals almost do not react on each other. Figure 12 shows the influence of both parameters on the utilization plot.

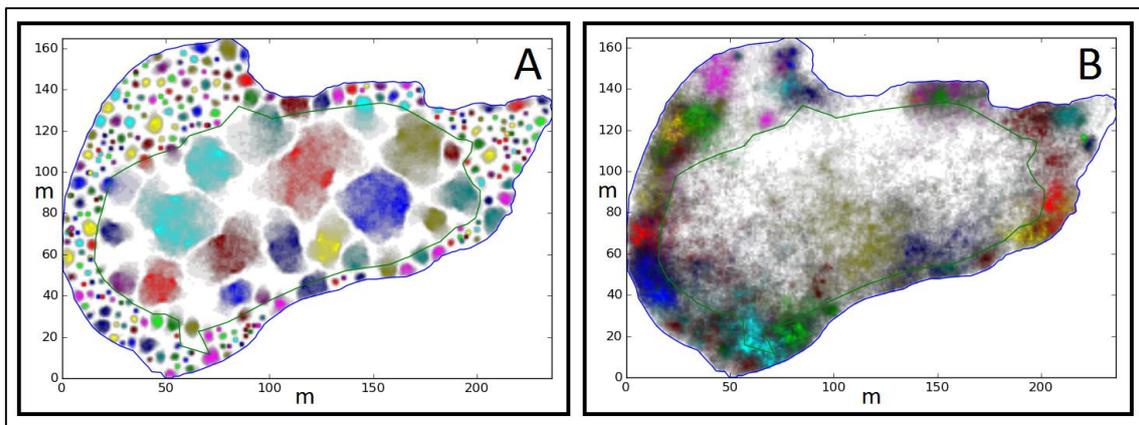


Figure 12: Influence of degree of stigmergy  $\alpha$  and active scent time  $T_{AS}$  on territory boundaries. A:  $\alpha = 10^{-0.5}$  and  $T_{AS} = 10^8$ . Territories are well delimited and there is no overlap between the territories of different fish. B:  $\alpha = 10^{-2.5}$  and  $T_{AS} = 10^{4.5}$ . Many territory sizes are only limited by the simulation time but not by the interaction with other fish. There is huge overlap and little reaction on each other. The chosen plotting method removes all position points, where fish do not have been often.

Territory size is also defined by different parameters, the most direct being the number of simulation steps  $t_{total}$ . The longer a simulation is running the larger territories become. In search of a steady state I look for a  $t_{total}$  value where territory size  $h$  does not change significantly any more when further increasing  $t_{total}$ .

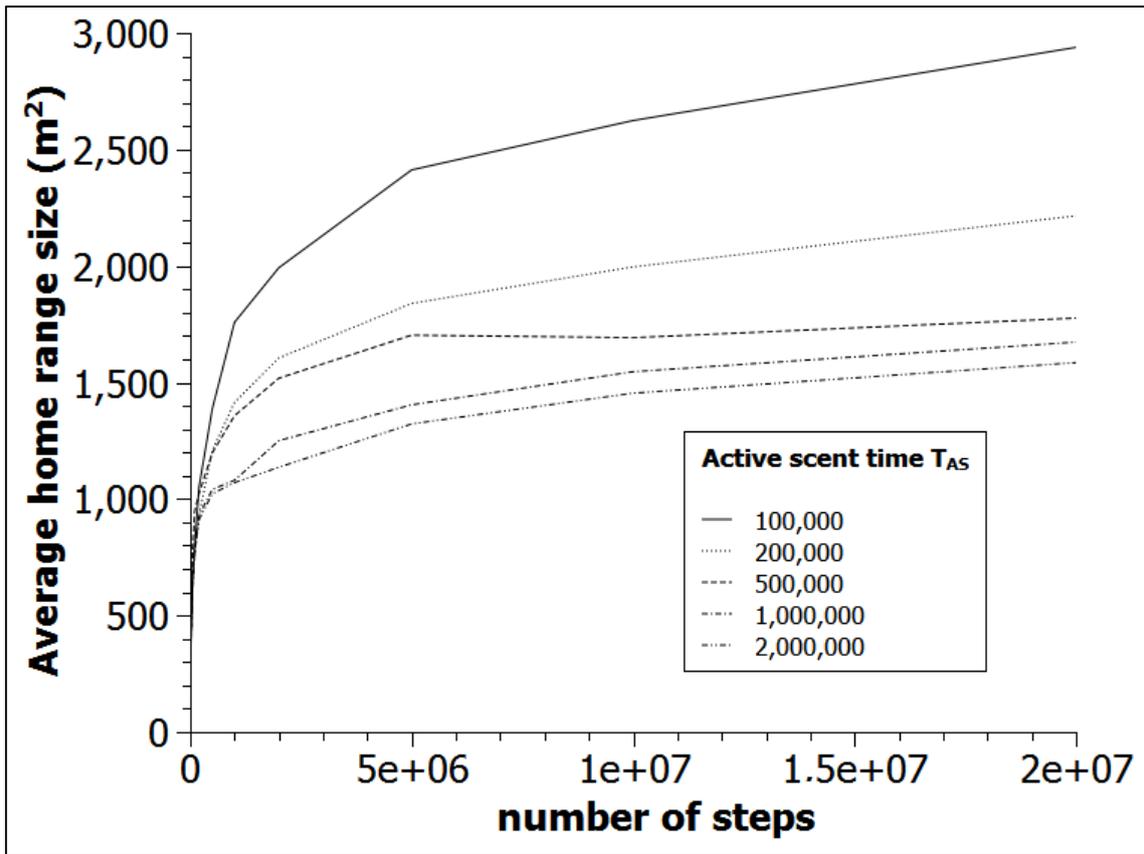
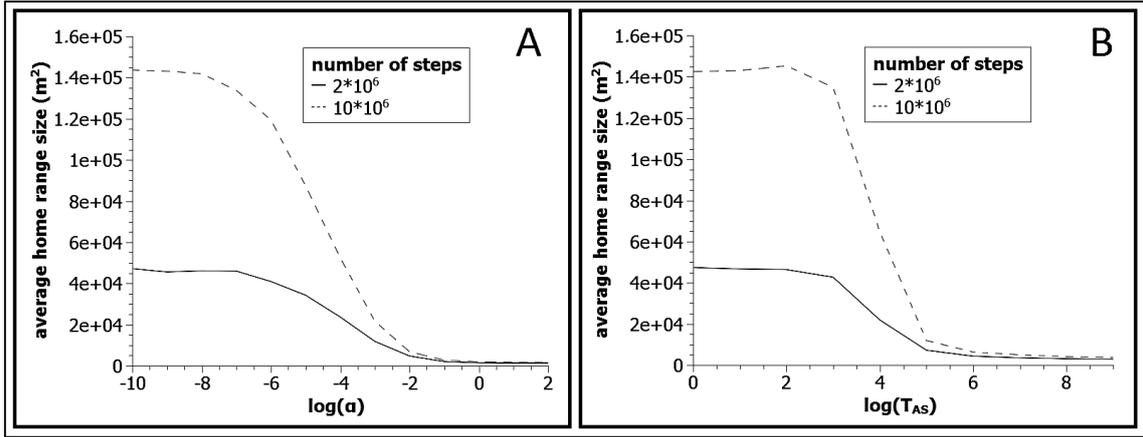


Figure 13: Average modelled home range size depending on number of steps for different  $T_{AS}$ . For certain thresholds on the x-axis the average territory size should reach saturation. Obviously the number of steps is not sufficient for some saturations with certain  $T_{AS}$ . This is because of the limited computer memory and prevention of long running times.

As can be seen in figure 13, there is still no  $t_{total}$  visible for which the territory size derivative  $\frac{dh}{dt}$  is zero. Nevertheless I chose  $t_{total} = 5 * 10^6$  (57.8 days) because much higher values extravagate the main memory and increase simulation time. An increase in reduction factor  $k$ , to compensate the higher  $t_{total}$ , would lead to more corrupted results.

Besides  $t_{total}$ , the values for  $\alpha$  and  $T_{AS}$  play a crucial role in territory size. I first investigated the influence of  $\alpha$  on territory size  $h$  with constant  $T_{AS}$  and then the influence of  $T_{AS}$  on  $h$  with constant  $\alpha$  (see Figure 14).

Both  $\alpha$  and  $T_{AS}$  show critical values for the upper/lower limit (figure 14) which marks the point where a further increase/decrease does not change territory size significantly. These will be the chosen limits for the optimization process (see chapter 4.5).



**Figure 14: Average territory size depending on  $\alpha$  (A) or  $T_{AS}$  (B) for different number of steps (log-scale on x-axis). This calculation helps to find a range for both parameter within which they should be optimized because there is a territory size limit for both very high and very low values of  $\alpha$  and  $T_{AS}$ . (Used parameters: A:  $T_{AS} = 1/5 * \text{number of steps}$ . B:  $\alpha = 0.01$ )**

#### 4.4 Most parameters are obtained from measurements

For many parameter values we can resort to measurement data. So we know the number of fish out of the abundance and their size distribution (see table 2). The data gives us insight into swimming speed and activity (Shinnosuke Nakayama). The total velocity has been determined by multiplying the activity (percentage of time a fish swims actively) and the median velocity. The results are shown in Table 3.

size class (cm)	median velocity ( $\frac{m}{s}$ )	activity (%)	total speed ( $\frac{m}{s}$ )
25 – 35	0.2089141	0.01813534	0.003788728
35 – 45	0.2100969	0.03219665	0.006764416
45 – 55	0.190974	0.03960992	0.007564465
55 – 65	0.1742125	0.07384041	0.012863922

**Table 3: swimming speed for different size classes (Shinnosuke Nakayama, columns 1-3). Total speed is calculated as the product of median velocity and activity.**

Larger fish are faster than smaller fish. Larger pike show much more activity (percentage of whole measurement time) than small ones but they are not necessarily faster. Not for all fish size classes that I want to optimize speed data is available. That is why I have to extrapolate. We chose a sigmoidal curve of the form:

$$g_{size} = d_1 * \frac{d_2 * \overline{l_{size}} - d_3}{\sqrt{1 + (d_2 * \overline{l_{size}} - d_3)^2}} + d_4 \quad (17)$$

where  $g_{size}$  is the part of the propensity which depends on fish size. Additionally,  $g_{size}$  is a measure for swimming speed.  $\overline{l_{size}}$  is the mean size in size class  $l_{size}$  (see chapter

3.4).  $d_1, \dots, d_4$  are free parameters. Those parameters were adjusted manually, such that the given data matches the whole distribution. The result is shown in figure 15.

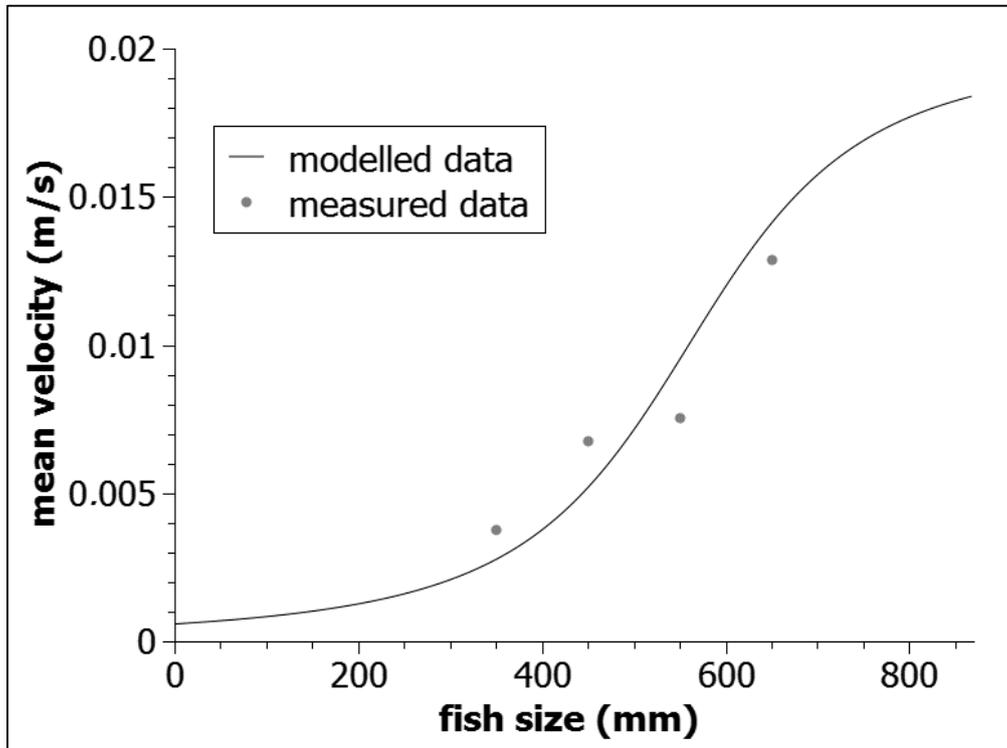


Figure 15: approximation of velocity depending on fish body size. Dots represent measured data whereas the lined curve is modelled sigmoidal in order to find a continuous function that converts fish body size into velocity. Under a linear fit, small fish would have been assigned negative velocities. That is why we chose a sigmoidal form. Chosen values:  $d_1 = 0.01$ ,  $d_2 = 0.005$ ,  $d_3 = 2.8$ ,  $d_4 = 1$ .

Some parameters cannot be obtained from measurement data. The part of the propensity which depends on the fish's position,  $g_{veg}$ , is set to 3, which means, that fish in the pelagic zone will swim 3 times faster than those in the littoral zone. Because fish stay inactive in littoral zones most of the time (see chapter 2.2) the swimming inhibition, induced by  $g_{veg}$ , could be much higher, but the model is not construed for fish that swim slow and seldom. The fish's movement is a crucial action that leads to territory formation, that is why  $g_{veg}$  was set to a moderate value.

We chose to set the centroid weighting factor  $w_{veg}$  to 10, because especially small fish are not supposed to form territories in pelagic zones, justified by the observed behavior. Since large fish will react on other fish less strong they are not influenced that much by  $w_{veg}$ .

Parameter  $A$ , introduced in formula 15, tells us how much the degree of stigmergy  $\alpha$  is increased, when small fish meet large fish. It was set arbitrarily to  $A = 10^5$ , because there is no relevant data which helps to optimize  $\alpha$ , which is directly connected to  $A$  and small pike need to react very strong on big fish.

#### 4.5 Results of parameter adjustment

The main parameters that need to be scanned are  $T_{AS}$  and  $\alpha$ . They present the core of the whole modelling approach. The range of these parameters has been determined as shown in chapter 4.3. The following parameter set in table 4 has been implemented.

parameter	variable term	value
Number of fish	$N$	2928
Time steps	$t_{total}$	$5 * 10^6$
Burn-in-time	$T_{burn}$	$5 * 10^5$
Vegetation centroid weighting factor	$w_{veg}$	10
Vegetation velocity factor	$p_{pos}$	3
Size classes	$(l_1, l_2, \dots, l_c)$	(35, 40, 45, 50, 55, 60, 65, 70)
Abundance (% of $t_{total}$ )		(13, 29, 25, 15, 10, 4, 3, 1)
Total velocity ( $10^{-4} \frac{m}{s}$ )		(27, 37, 51, 71, 93, 118, 138, 154)
Reduction factor	$k$	10

**Table 4: Setting for optimization of  $\alpha$  and  $T_{AS}$**

With variation of  $T_{AS}$  and  $\alpha$  we get a heat map for the territory size in a two-dimensional matrix which is shown in figure 16.

Red color indicates large territories whereas blue color indicates small ones. As expected, territory size increases with small  $\alpha$ -values because the strength of interaction is small which leads to almost random walk, inducing large territories. The same applies to the active scent time  $T_{AS}$ . When scent marks fade away fast (small  $T_{AS}$ ), there is no basis for interaction, resulting in territory increase.

The sum of squares serves as quality criterion of the given setting. Figure 17A shows the sum of squares for different  $T_{AS}$  and  $\alpha$  over the whole range that has been determined as mentioned in chapter 4.3. To achieve better perceptibility, in figure 17B the range for both parameters  $T_{AS}$  and  $\alpha$  has been zoomed and for better definition the intervals have been reduced. Furthermore the reciprocal is plotted because only very small values are of interest to us.

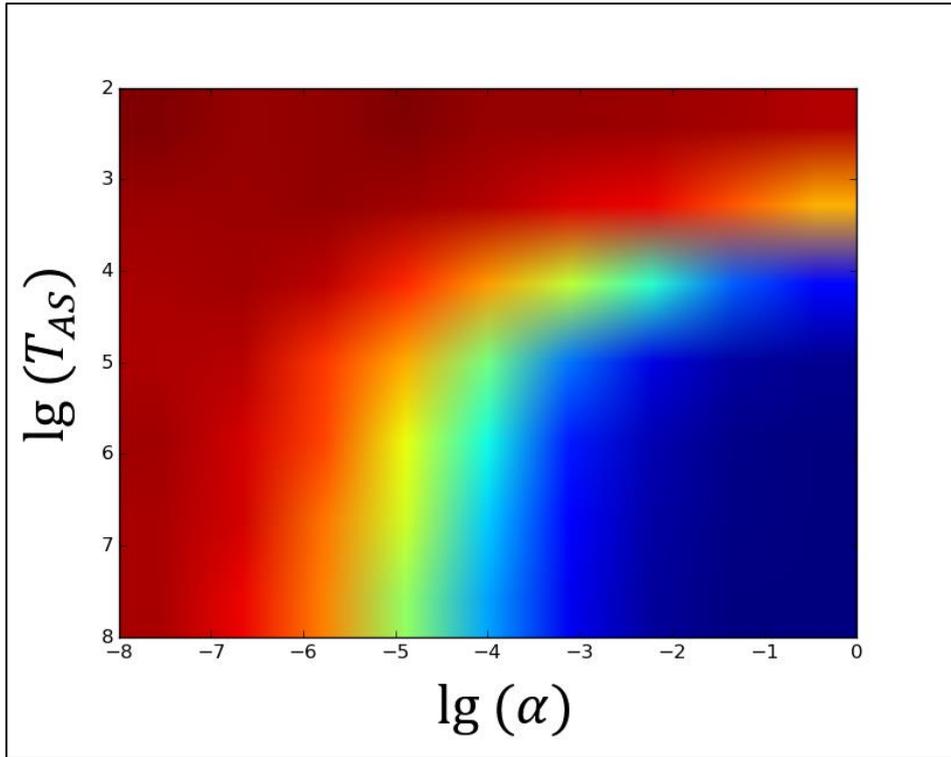


Figure 16: heat map of modelled home range sizes depending on  $\alpha$  and  $T_{AS}$  (inverse scale on y-axis). Blue color stands for small, yellow for medium and red for large home ranges. With increasing  $\alpha$  and  $T_{AS}$  home ranges become smaller because of the higher competition between fish.

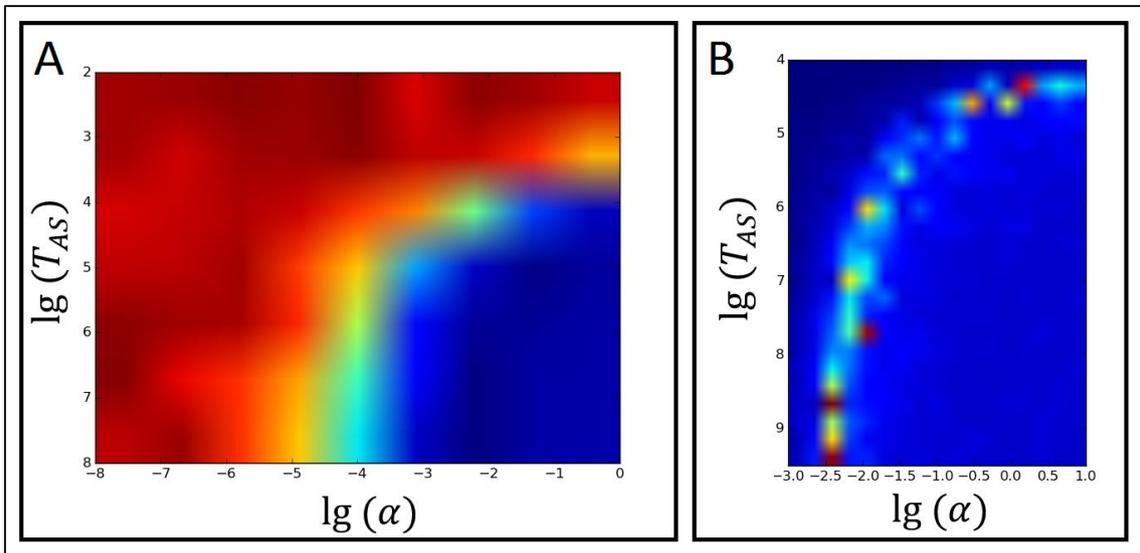


Figure 17: Sum of squares as quality criterion of parameter setting. Both modeled and measured territory sizes are compared and the sum of squares was calculated. A: Sum of Squares. There is a faint dark blue band starting from  $\lg(\alpha) = -2.5$ . B: Reciprocal ( $1/S$ ) of sum of squares with reduced range for both  $\alpha$  and  $T_{AS}$ . So the setting with good values (dark blue band in A, light blue and yellow/red color in B) are better visible.  $\alpha$  and  $T_{AS}$  compensate each other (non-identifiable character). In the outstanding area (colors different from dark blue) modeled territory sizes are very similar to measured values.

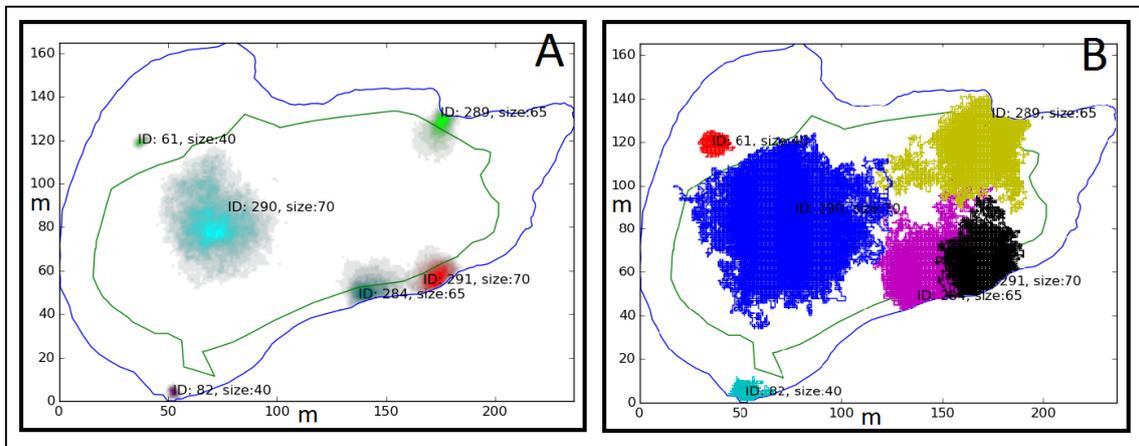
As expected, not only one single setting leads to a small sum of squares, but a series of non-identifiable settings starting from small  $\alpha$ -values (bottom left corner in figure 17B) and long active scent times  $T_{AS}$  turning to small  $T_{AS}$  and high  $\alpha$  (top right corner).

On the matrix's bottom right side territories are smaller than the target values and on the upper left side they are larger (compare to Figure 16). Due to the limited resolution ( $\Delta \lg(\alpha) = \Delta \lg(T_{AS}) = 0.25$ ) there is much noise in the results of the data, but the main idea of their non-identifiable character (Figure 17B) is made clear. Some values have been extremely small, such that the color palette shifted to dark blue for almost all regions except the extreme values. Those values have been removed from the figure.

Figure 18A shows a utilization plot with the parameter setting, printed in table 5.

parameter	variable term	value
Degree of stigmergy	$\alpha$	$10^{-0.75}$
Active scent time	$T_{AS}$	$10^{4.5}$

**Table 5: Parameter setting for figure 18. All other parameters remain unchanged. The parameter adjustment results state these parameters as good values for the simulation.**



**Figure 18: Example plot for different fish. A: Utilization plot where territories for selected fish are shown. B: Trace plots for same fish which represent all positions where fish have been over the entire simulation time. Complete traces (B) are much bigger than territories (A) and show great overlap whereas territories are delimited.**

Territories seem to have sharp boundaries in figure 18A, but a look on the complete trace plot in figure 18B shows that the utilization goes much farther than just the plotted territory. The heat map in figure 18A is instructed to show densities of utilization by varying the transparency of its color. So those regions with high densities of measurements are printed full colored, others have a higher level of transparency.

In this way, only territories are plotted and not all fields of its home range that have been visited by a fish. This is an adequate visualization method since we want to investigate territory formation. The whole trace of fish reaches far into the neighboring territories, while territories do not overlap. In this model the trace plot area is similar to its home range.

Figure 19 compares modelled and measured data. Shown is the mean of 10 simulations in the first (lined) curve and the mean of all territory sizes for each fish size in the second (dotted) curve. The average sum of squares of those 10 simulations is 3594, which is a comparatively good value. The parameter setting for this simulation is the one shown in Table 5. The vertical lines show the mean square displacement of each of all 10 simulations. Especially for big fish, the mean square displacement is very high, because of the little number of fish with this size.

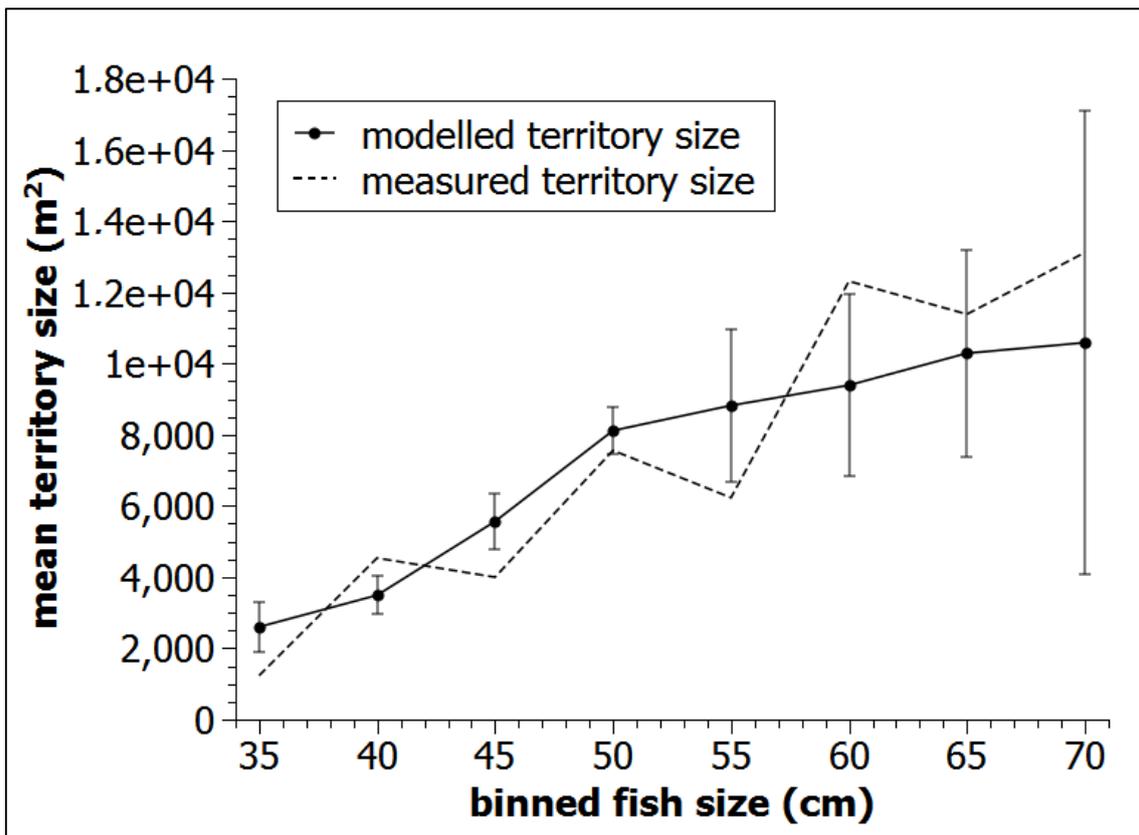


Figure 19: Comparison of modelled with measured territory size. Measured data is averaged over all fish of the corresponding size. Modelled size is averaged over 10 simulations. Vertical lines represent mean square displacement of modelled territory sizes. The parameter setting is shown in Table 5.

## 5 Discussion

The main achievement of this work is the connection between active scent times  $T_{AS}$  and degree of stigmergy  $\alpha$ . They stand representatively for the duration of chemical cues indicating recent presence and the strength of interaction between fish. Those findings provide insight in behavior of pike and help to understand the formation of territories.

First I tried to reverse engineer the stigmergy model by Potts and Giuggioli and estimated its quality by comparing their encounter rate curve with mine. There are still differences between the curves. After personal communication with the authors we believe that those discrepancies probably arise from different calculation methods regarding the centroid determination. We adapted the stigmergy model to our needs and used parameters deduced by the measurements of the group of Robert Arlinghaus from the department of fisheries Berlin. In the last step we optimized all parameters that cannot be obtained from empiric results by varying them and comparing them with original data of territory size as a quality criterion of our model. We found settings that match both data and model which we will now discuss in order to find their advantages and deducible statements.

A frequently asked question is: Are scent marks a realistic assumption when investigating pike interactions? It is proven that at least pike prey (European minnows, *Phoxinus phoxinus*) leave pheromones which serve as Schreckstoff for other fish [10]. So pike can be recognized via scent marks. There is no data available about the time the scent marks stay active. Another approach is that small  $T_{AS}$  stand representatively for sight range of fish instead of chemical cues which stay for a longer time. Therefore we act on the assumption of very small active scent times. The parameter optimization of  $\alpha$  and  $T_{AS}$  leave open, which values should be chosen to get a realistic model. The optimization just reveals a dependency of both values. Small active scent times  $T_{AS}$  require high  $\alpha$ -values as shown in Figure 17B. This is consistent with the fact that fish, if they meet each other, will react very strongly, especially when small pike meet large ones [7]. This leads to the conclusion that of all possible  $\alpha - T_{AS}$  - combinations those are in favor where  $T_{AS}$  is low and  $\alpha$  is high. The results for  $T_{AS}$  that we get from the

optimization range down to about 8.7 hours, which is about 0.6 % of the whole simulation time. For lower  $T_{AS}$  territory sizes become too small, independent of  $\alpha$ .

There are huge differences in home range and territory sizes even for equally sized pike (see error bars in Figure 19). A reason for this is among other things the differences in initial conditions. Fish are initialized randomly somewhere in the littoral zone. Fish, for example, that are initialized in the lake's corners tend to form larger territories because they have to defend them only to one side. Moreover, fish that were initialized very close form smaller territories than those that have been assigned large almost exclusive regions by coincidence. Because of a missing temporal forming stage, in which fish would show different behavior und just try to find own territories, territory size depends very much on initial conditions and only averages over many territory sizes give hints about the dependencies of territory sizes on other parameters.

For reasons of time some parameter influences have not been tested entirely. So I am not sure how much burn-in-time is necessary to get saturation in territories sizes. But in our model an accurate setting of burn-in-time is less essential than for total time steps  $t_{total}$ , because territory sizes do not depend on interim but final values. The number of time steps should be higher than the chosen number to make sure that model output is not influenced by a too-short modelling time. Because of the available main memory I had to limit modelling time.

## 6 Outlook

For time reasons this thesis is limited to the actual state of working process which is a good conclusion because the model is able to show pike's territory formation. Also I found a series of parameter settings which can reproduce measured data. Continuing investigation could improve the results of this model.

The value for  $A$  (factor that raises  $\alpha$ -value for special encounter size ratios, formula 15) was set arbitrarily to  $A = 10^5$ . Further testing could help to understand the influence of this parameter. Same is true for  $w_{veg}$  (higher weighting factor of centroid in littoral zone) and  $g_{veg}$  (increase of swimming speed in pelagic zone). Lake measurement data shows that fish stay inactive for most of the time. This part was ignored in the model

where only the average velocity was implemented. An implementation of active and inactive times depending on fish size could increase the model's output accuracy. Small fish just avoid large fish in our model. An interesting feature would be the cannibalistic predatory behavior, so that fish are eaten by conspecifics. The number of small fish in the parameter setting then had to be much higher and it could be investigated which behavior of small fish leads to success in growth and rise in food chain.

Some of the concepts used in this model could be adapted to other fields. So the Gillespie part could be taken over in spatial molecule simulations.

The current state of the model allows us to roughly reproduce measurements. The next step should be the deduction of predictions about pike behavior. The influence of climate change for example forces animals to reactions, often even before they can adapt their behavior accordingly [19]. Influences of temperature, sunshine duration and wind velocity can be added to the environmental parameters, such that changes of those values could be modelled. Refined data analysis of the recorded fish positions, like better thresholds and filters for territory sizes would improve the model. A more precise look on the classifications, like sex or original spawning lake (transferred fish from other lakes) would give hints, what time foreign fish require to form territories and adapt their behavior.

## 7 Appendix

### 7.1 Discrepancy with reimplementing of encounter rate

When trying to reproduce the encounter rate from Potts and Giuggioli [1], I first made a mistake in centroid calculation. The wrong calculation lies in the fact that the centroid is always determined by taking the mean of the last  $n$  positions occupied by the animal ( $n$  is the number of steps which an animal makes on average during time  $T_{AS}$ ). Because each new position is periodically corrected first and taken into account for the mean value afterwards, there is a strong bias for the centroids to wander to the center of the mesh (The mean value of random numbers between 0 and  $x$  will be  $\frac{x}{2}$ ). If the centroids are close, the animals will meet very often (see figure 20 for better understanding).

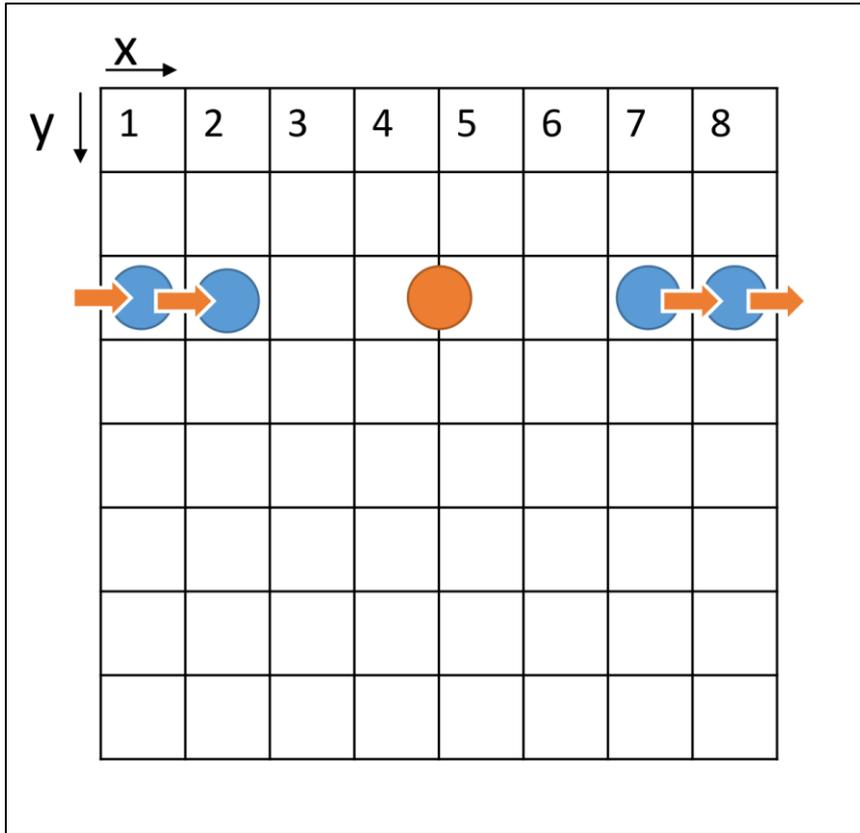


Figure 20: demonstration of wrong centroid calculation. Fish that successively occupy the fields 7 – 8 – 9 – 1 – 2 get a value of  $c = 4.5$  as their centroid which is wrong, because the centroid needs to be at  $c = 0.0$  on the outer edge of the operating area (8.0 is excluded and jumps to 0.0 immediately).

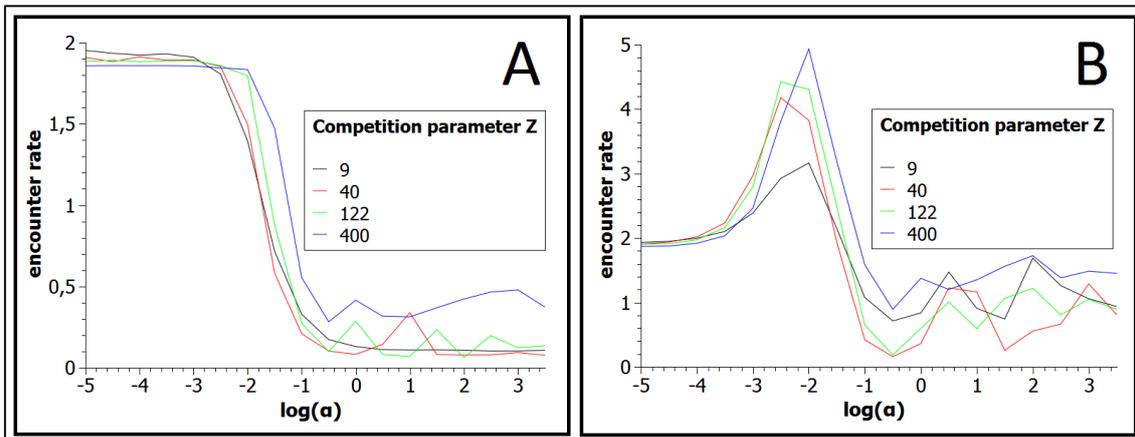
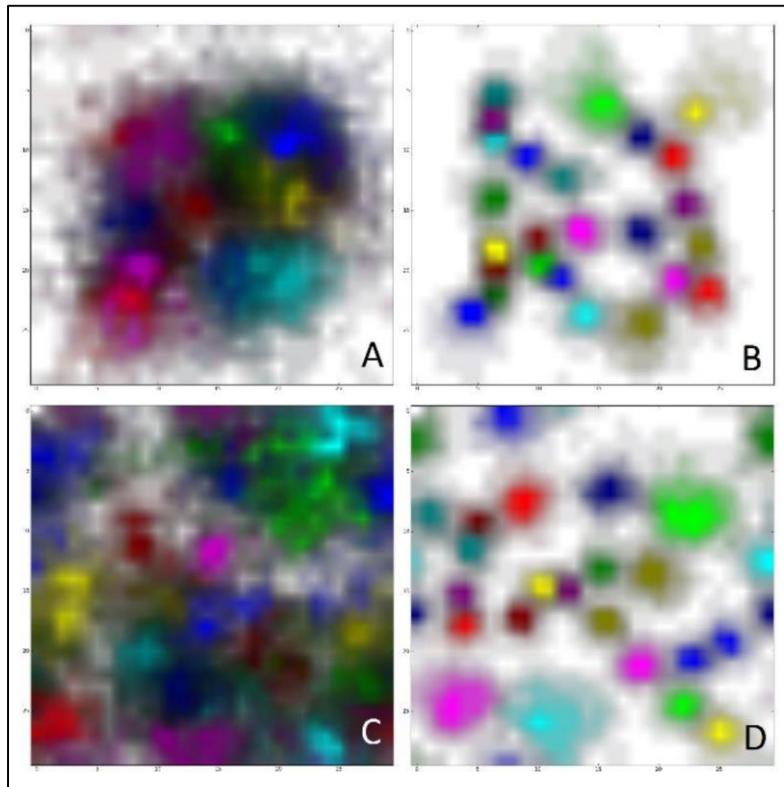


Figure 21: Comparison of encounter rate with correct (A) and incorrect (B) calculation. Encounter rate is normalized to diffusion constant, density, number of steps. The peak at  $\alpha = 10^{-2}$  vanishes after fixing the error. The limit for small  $\alpha$  values remains in both cases at just under 2.0.

After correcting this mistake the encounter rate curve looks different (figure 21A). Figure 21B looks more similar to the one from Potts and Giuggioli (see chapter 3.3.2,

figure 5). For both ways to calculate centroids, in figure 22 a utilization plot for different  $\alpha$ -values is shown.



**Figure 22:** Figure with utilization plot for different alpha values and correct and incorrect centroid calculation in comparison. A: Wrong centroid calculation,  $\alpha = 10^{-2}$ . The wrong calculation has huge influence on the high encounter rate because animals are close. B: Wrong centroid calculation,  $\alpha = 10^{-1}$ . Because of the higher  $\alpha$  value animal reaction on each other is stronger and encounter rate is influenced to a lesser extent. C: Proper centroid calculation,  $\alpha = 10^{-2}$ . Animals are equally distributed and there is high overlap. No peak is to be expected. D: Proper centroid calculation,  $\alpha = 10^{-1}$ . There is less overlap than in (C) because of the higher  $\alpha$ -value but equal distribution of animals as well.

In figures 22A and B centroids are calculated in a wrong way. That is why animals gather in the middle of the whole area. For  $\alpha = 10^{-2}$  (figure 22A) the effect has a higher influence because territories are larger and overlap as well as encounter probability are higher. Potts and Giuggioli justified this peak with the density being higher around this value (see chapter 3.3.2). We are in contact with the authors to find the reason for this difference.

## 7.2 Python Libraries

Programming language:	Python (version 2.7)
Plots:	matplotlib.pyplot (version 1.4.3)
Saving and Loading:	cPickle (version 1.7.1)

Calculations:                    numpy (version 1.8.2)  
                                      bisect  
Random numbers:                random

## 8 References

1. Giuggioli, L., et al., *Stigmergy, collective actions, and animal social spacing*. Proc Natl Acad Sci U S A, 2013. 110(42): p. 16904-9.
2. Prejs, A., et al., *Food web manipulation in a small, eutrophic Lake Wirbel, Poland: effect of stocking with juvenile pike on planktivorous fish*. Hydrobiologia, 1994. 275(1): p. 65-70.
3. Raat, A.J.P., *Synopsis of biological data on the Northern Pike, Esox lucius Linnaeus*, . FAO Fish.Synop., 1998. (30)Rev.2:178 p.
4. Hegemann, M., *Beitrage zur Biologie des Hechtes im Brackwassergebiet des Greifswalder Boddens*. Z. Fisch, 1958. 7: p. 459-477.
5. Frost, W.E. and C. Kipling, *A study of reproduction, early life, weight-length relationship and growth of pike, Esox lucius L., in Windermere*. The Journal of Animal Ecology, 1967: p. 651-693.
6. Frost, W.E., *The Food of Pike, Esox lucius L., in Windermere*. Journal of Animal Ecology, 1954. 23(2): p. 339-360.
7. Larsen, K., *Studies on the biology of danish stream fishes: III. On seasonal fluctuations in the stock density of yellow eel in shallow stream biotopes, and their causes*. MEDD DAN FISK HAVUNDERS, VOL 7, NO 2, P 23-46, 1972. ILLUS., 1972.
8. Nursall, J.R., *Some Behavioral Interactions of Spottail Shiners (Notropis hudsonius), Yellow Perch (Perca flavescens), and Northern Pike (Esox lucius)*. Journal of the Fisheries Research Board of Canada, 1973. 30(8): p. 1161-1178.
9. Chapman, C.A. and W.C. Mackay, *Versatility in habitat use by a top aquatic predator, Esox lucius L.* Journal of Fish Biology, 1984. 25(1): p. 109-115.
10. Magurran, A.E., *Acquired recognition of predator odour in the European minnow (Phoxinus phoxinus)*. Ethology, 1989. 82(3): p. 216-223.
11. Brown, G.E., D.P. Chivers, and R.J.F. Smith, *Localized defecation by pike: a response to labelling by cyprinid alarm pheromone?* Behavioral Ecology and Sociobiology, 1995. 36(2): p. 105-110.
12. Jepsen, N., et al., *Behavior of pike(Esox lucius L.)> 50 cm in a turbid reservoir and in a clearwater lake*. Ecology of Freshwater Fish, 2001. 10(1): p. 26-34.
13. Poddubnyi, A., *Ecological topography of fish populations in reservoirs*. 1976.
14. Vostradovsky, J., *Horizontal distribution on individually tagged fish in the Lipno reservoir*. EIFAC Technical Papers (FAO). no. 23 (suppl. 1, v. 2). 1975.
15. Burt, W.H., *Territoriality and home range concepts as applied to mammals*. Journal of mammalogy, 1943. 24(3): p. 346-352.
16. Kaukoranta, E. and E. Lind, *The pike, Esox lucius L., in the estuary of the Oulujoki River. 1. Ecology*. Ichthyologia Fennica Borealis, 1975: p. 1-2.
17. Diana, J.S., *Diel activity pattern and swimming speeds of northern pike (Esox lucius) in Lac Ste. Anne, Alberta*. Canadian Journal of Fisheries and Aquatic Sciences, 1980. 37(9): p. 1454-1458.
18. Carbine, W.F., *Growth potential of the northern pike (Esox lucius)*. Pap. Mich. Acad. Sci, 1945. 30: p. 205-220.
19. Potts, J.R. and M.A. Lewis, *How do animal territories form and change? Lessons from 20 years of mechanistic modelling*. Proceedings of the Royal Society B-Biological Sciences, 2014. 281(1784).
20. Lotka, A.J., *Elements of Physical Biology*. Williams and Wilkins Company, 1925.

21. Lewis, M.A. and J.D. Murray, *Modelling territoriality and wolf-deer interactions*. Letters to Nature, 1993.
22. Giuggioli, L., J.R. Potts, and S. Harris, *Animal interactions and the emergence of territoriality*. PLoS Comput Biol, 2011. 7(3): p. e1002008.
23. Marsh, L. and C. Onof, *Stigmergic epistemology, stigmergic cognition*. Cognitive Systems Research, 2008. 9(1-2): p. 136-149.
24. Hasler, A.D., et al., *Open-water orientation of white bass, *Roccus chrysops*, as determined by ultrasonic tracking methods*. Journal of the Fisheries Board of Canada, 1969. 26(8): p. 2173-2192.
25. Poddubnyi, A., L. Malinin, and V. Gaiduk, *Experiment in telemetric observations under ice of the behaviour of wintering fish*. Biol. Vnutr. Vod Inf. Byull, 1970. 6: p. 65-70.
26. Diana, J.S., W. Mackay, and M. Ehrman, *Movements and habitat preference of northern pike (*Esox lucius*) in Lac Ste. Anne, Alberta*. Transactions of the American Fisheries Society, 1977. 106(6): p. 560-565.
27. Chapman, C. and W. Mackay, *Versatility in habitat use by a top aquatic predator, *Esox lucius* L.* Journal of Fish Biology, 1984. 25(1): p. 109-115.
28. Gillespie, D.T., *Exact Stochastic Simulation of Coupled Chemical-Reactions*. Journal of Physical Chemistry, 1977. 81(25): p. 2340-2361.
29. Roussel, M.R. *Modelling Biochemical Reaction Networks*. Available from: <http://people.uleth.ca/~roussel/C4000biochemnet/slides/14CME.pdf>.
30. Hepburn, I., et al., *STEPS: efficient simulation of stochastic reaction-diffusion models in realistic morphologies*. BMC systems biology, 2012. 6(1): p. 36.
31. Okubo, A. and S.A. Levin, *Diffusion and ecological problems: modern perspectives*. Vol. 14. 2013: Springer Science & Business Media.
32. Persson, L., A.M. De Roos, and A. Bertolo, *Predicting shifts in dynamics of cannibalistic field populations using individual-based models*. Proceedings of the Royal Society of London B: Biological Sciences, 2004. 271(1556): p. 2489-2493.
33. Pagel, T., *Determinants of individual reproductive success in a natural pike (*Esox lucius* L.) population: a DNA-based parentage assignment approach*. 2009.

## 9 Acknowledgement

I would like to say thank you to all the people who supported me during the creation of this Bachelor thesis.

First I want to express my gratefulness to my supervisors Max Flöttmann, Wolfgang Giese and Jannis Uhlendorf. Thank you for your patience, the contribution of your knowledge and your kind advices in all forms.

Thank you to Edda Klipp for letting me write my thesis in her group and to Robert Arlinghaus and his group, notably Daniel Hühn, David March-Morlà, Shinnosuke Nakayama and Thilo Pagel for providing me with the data for my thesis.

Last but not the least, I would like to thank my friends, especially Lea Heinze and Dennis Nguyen for proofreading and also my family for mental assistance.

## 10 Declaration of authorship

I hereby declare that I wrote this Bachelor thesis independently and no other than indicated sources have been used.

Berlin, 22. September 2015

Signature: \_\_\_\_\_