

# Parameter Identification in Partial Integro-Differential Equations for Physiologically Structured Populations

Sylvia Moenickes<sup>\*,1</sup>, Otto Richter<sup>1</sup>, and Klaus Schmalstieg<sup>1</sup>

<sup>1</sup>Technische Universität Braunschweig Carolo Wilhelmina, Institut für Geoökologie, Abt. Umweltsystemanalyse

\*Corresponding author: Langer Kamp 19 c, s.moenickes@tu-bs.de

**Abstract:** Continuous dynamic models, e.g. Comsol based simulations, play – besides statistical or iterative methods – a mayor role in theoretical ecology; in forecasting and management, but also in systems analysis. Ecological issues generally comprise highly interacting agents and/or unknown side effects. We here show how combining direct simulation with Comsol with simple optimization tools in Matlab helps in research in ecology. We analyse the dynamics of a physiologically structured population under changing temperature and food conditions in two ways. At first we conduct virtual experiments aiming at efficient experimental design for parameter identification, afterwards we accomplish simulations for process analysis in real life experiments for the freshwater shrimp *Gammarus pulex*.

**Keywords:** Parameter estimation, physiologically structured populations, experimental design

## 1 Introduction

Climate change surely has an impact on ecological systems and forecasting their reaction is a major goal in nowadays ecological research, [2] and [3]. Dynamic models can be a reliable help in these attempts if they include the governing processes and are correctly parameterized. When it comes to ecological systems comprising highly coupled processes like dynamics of populations in their environment one might choose between compiling a range of well analyzed dependencies with the risk of unjustified omission of others or integrative modelling with the risk of mistaken process attribution. In this paper we show how preceding simulation and subsequent optimization tools can help in de-

signing effective experiments for sound process analysis and parameter identification. We do so for the dynamics of physiologically structured populations of the freshwater shrimp *Gammarus pulex* under changing temperature and food conditions.

## 2 System components and equations

We suppose that the population dynamics of *G. pulex* is mainly affected by its own population density, food supply and temperature. By this reason the following constituting equations are assembled:

Individual growth of an animal follows the von Bertalanffy equation, [6], which assumes that anabolic and catabolic processes scale with the animal's surface and volume respectively:

$$\frac{dw}{dt} = g(w) = \gamma w^{\frac{2}{3}} - \rho w$$

with  $\gamma$  [ $\mu\text{g}^{-\frac{2}{3}} \text{d}^{-1}$ ] and  $\rho$  [ $\mu\text{g}^{-1} \text{d}^{-1}$ ] being rate constants and  $w$  weight [ $\mu\text{g}$ ]. When food  $F$  [1] is limited, the equation modifies to

$$\frac{dw}{dt} = g(F, w) = \gamma \frac{F}{F + F_h} w^{\frac{2}{3}} - \rho w$$

with the half saturation constant  $F_h$  [1]. Besides, when temperature is other than optimal, the equation is complemented by

$$\frac{dw}{dt} = g(F, T, w) = \Phi(T) \left( \gamma \frac{F}{F + F_h} w^{\frac{2}{3}} - \rho w \right)$$

with  $\Phi(T)$  a temperature response function following O'Neill, [5]:

$$\Phi(T) = k \left( \frac{T_{\max} - T}{T_{\max} - T_{\text{opt}}} \right)^p e^{p \frac{T - T_{\text{opt}}}{T_{\max} - T_{\text{opt}}}}$$

where  $p = \frac{1}{400}W^2(1 + \sqrt{1 + \frac{40}{W}})^2$ , with  $W = (q_{10} - 1)(T_{\max} - T_{\text{opt}})$  and shape parameters  $q_{10}$ ,  $T_{\max}$ ,  $T_{\text{opt}}$  and possibly  $T_{\min}$  comprising a number of divergent physiological processes related to temperature.

However, an overall, and weight structured, population is described by a weight density function  $n(w, t)$  and its dynamic follows

$$\frac{\partial n(w, t)}{\partial t} = -\frac{\partial g(F, T, w)n(w, t)}{\partial w} + \lambda \frac{\partial^2 n(w, t)}{\partial w^2} - p_-(F, T, w, \dots) + p_+(F, T, w, \dots)$$

Here  $\lambda \frac{\partial^2 n(w, t)}{\partial w^2}$  describes stochasticity of physiological processes;  $p_-$  and  $p_+$  are birth and death terms.

Mortality might be due to a number of external effects like temperature  $T$  and food  $F$ , again expressed in a way similar to their influence on growth, but it might also be due to internal effects like weight  $w$  and age, e.g. expressed by  $t$ . We come back to that in section 4.

Birth is expressed as the total number of offspring  $B(F, T, t, \dots)$  distributed over small weights  $w$  with a density function  $\Pi(w)$  to  $p_+(w, t, \dots) = B(F, T, t, \dots)\Pi(w)$ . The total number of offspring is the result of an integration of population density over weight, where fertility may depend on weight  $w$ , food  $F$  and temperature  $T$ , e.g.

$$B(F, t) = \int_{w_{\min}}^{w_{\max}} r_{\max} \frac{F}{F + F_h} w^{\frac{2}{3}} n(w, t) dw$$

A number of additional modifications of the population density model itself will be taken into consideration later on. For the time being there are two system compartments left: food and temperature.

The latter can be given explicitly, as a data set of true weather conditions or as a function of time which controlled the temperature during the experiments.

In an experimental setting food  $F$  might also be controlled, for example it may be kept in abundance, however in nature is is coupled to the population as follows:

$$\frac{dF}{dt} = L(t, T) - I_{\max} \Phi(T) \frac{F}{F + F_h} \int_0^{w_{\max}} w^{\frac{2}{3}} n(w, t) dw$$

where  $I_{\max}$  is efficiency of food uptake and  $L(t)$  a possible source term combining litter fall and microbial processing. Note that food is given in an ordinary differential equation whereas population is given in an partial differential equation.

### 3 Parameter identification

One major challenge in ecology is to define system boundaries properly, another is to identify predominant processes correctly. Modern ecological experiments aim at integral studies combined with mathematical tools rather than conducting countless highly specific laboratory studies for single parameters.

We appreciate Comsol multiphysics as a tool with which quickly to do both, exclude, include and modify processes on the one hand and have multiple visualisations of the results on the other hand. However, after creatively defining hypotheses about the ecological system under consideration, their validity should be tested against experimental data by statistical analysis. The facility to combine Comsol multiphysics scripts with Matlab appeared appealing to us as we already used Matlab for data evaluation.

Experimental data on weight structured population dynamics generally consist of frequencies of classified mean weights  $n_i$  for a couple of times  $t_j$ . Optimization tools are confronted with the minimization of the sum of squared residuals

$$L(\theta) = \sum_i \sum_j (n_i(t_j) - \hat{n}_i(t_j, \theta))^2$$

based on simulation results  $\hat{n}_i(t_j, \theta)$  for a given set of parameters  $\theta$ . Goodness of fit was evaluated using model efficiency, [1],

$$\text{ME} = 1 - \frac{\sum_j \sum_i (n_i(t_j) - \hat{n}_i(t_j))^2}{\sum_j \sum_i (n_i(t_j) - \bar{n}(t_j))^2}$$

with  $\bar{n}(t_j)$  being the mean of the mean weights at time  $t_j$ .

### 4 Application in experimental design

Before conducting integral ecological studies it is important to make sure that within

determined boundaries the required processes and related parameters can be identified. We here show how well thought-out experimental design can be achieved based on virtual experiments and can help reduce costs of real life experiments. Synthetic data were acquired via probing direct simulation results on four monitoring times (analogous to a real experimental set up described in the last subsection) and adding a normally distributed error representing measurement errors.

#### 4.1 Reducing experimental effort

The dependence of anabolic and catabolic processes on temperature was expressed as O'Neill function. Earlier experiments determined the related parameters via piles of animals fed and kept at constant but differing temperatures being thoroughly monitored in their growth, [4]. Our first experiment shows that one (possibly subdivided however) population kept at time dependent temperature might suffice to determine all of the parameters appearing in the O'Neill function and the anabolic and catabolic rates  $\gamma$  and  $\rho$ . In this experiment we created a virtual reality that was based on true weather conditions, that is true temperatures measured during the experiment described in the last subsection. In the same vain, we sampled our virtual reality at the same monitoring times as was done in the experiment to be explained.

Figure 1 displays the close match between data and optimization results, also reflected by a model efficiency of 0.9785, compared to 0.9921 for the model based on the given parameters (Remember that the difference to 1 is due to the normally distributed error that was added). However, the table below compares given and found parameters. We find that given parameters are not always within standard deviation of the found parameters and suppose that with temperature data varying in the full range the O'Neill function covers model efficiency would be closer to one.

	$\gamma$	$\rho$	$T_{opt}$	$q_{10}$
direct	0.16	0.04	17.5	1.70
optimized	0.32	0.06	18.3	2.58
std. dev.	0.02	0.05	0.05	0.10

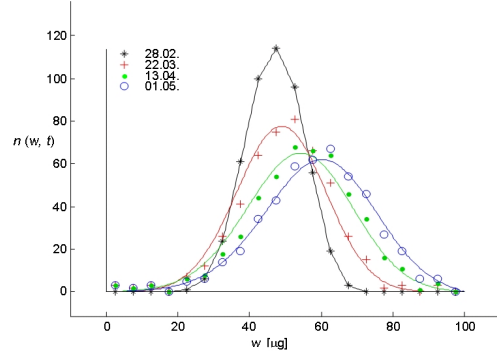


Figure 1: Weight structured population densities taking into account temperature dependent growth, here the optimized parameter set

#### 4.2 Piloting multi-component experiments

Both the traditional type of experiments to find O'Neill parameters as well as the one explained above are based on food kept in abundance. In natural settings food supply varies and is governed by the dynamics of both the population itself and all the actors in the food web. Insufficient food supply results in a decrease of anabolic processes and finally in weight loss of the affected animal. Experiments conducted with a temporarily insufficient food supply allow for a simultaneous estimation of the half saturation constant  $F_h$  which quantifies this effect. However, it is difficult to fully control food supply in an experiment. We here show that it is possible to estimate  $F_h$  even with only rough knowledge of the course of food supply by estimating a food supply function, or to be more exact, its parameters at the same time. On the one hand we can optimize the parameters in an explicit expression, e.g.

$$F(t) = F_0 e^{-(t/t_1)^{p_1}} (1 - e^{-(t/t_2)^{p_2}})$$

or those of the food dynamic given in the second section. The scenario described with the above equation is that food is given only once in the beginning of the experiment, but it becomes accessible after microbial turnover and is depleted after some time.

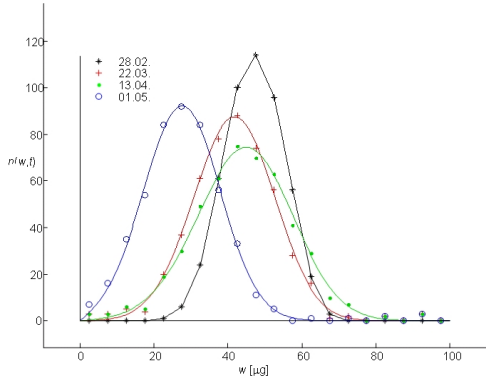


Figure 2: Weight structured population densities taking into account temperature dependent growth and dynamic food supply, here the optimized parameter set

Figure 2 shows the close match between simulation results and virtual data. Note the differences in the underlying dynamic compared to figure 1. The population density shifts to smaller weights in the first monitoring, displaying an overall loss in weight, increasing weight around the second monitoring, just to shift back to smaller weights until the last monitoring. The efficiency of the model parameterized via optimization is equal to that parameterized with given values (0.9947). Given and found parameters vary less than in the previous example.

	$F_h$	$p_1$	$p_2$	$t_1$	$t_2$
direct	4.7	20	25.0	82	98
optimized	4.27	18.4	25.1	84.0	97.9
std. dev.	0.4	0.12	0.83	0.07	0.29

We suppose that the crucial difference between the two optimizations discussed so far is the range of possible values of the environmental component that are adopted during the experiments. The course of the food supply applied here also belongs to the experiment to be described below, however this course comprises ultimate maximum and minimum and is therefore more suitable for optimization.

### 4.3 Process analysis

What remained quite unlikelike until now is that no mortality was taken into account. Mortality might depend on the animal's age and on environmental factors as temperature and food. Age is unknown for animals collected from natural environments, and a time-dependence or a constant value is supposed for mortality. As soon as tempera-

ture and food vary with time and as soon as they do so in such a way that optimal conditions for survival are lost, an analysis of experimental data must include assumptions on lethal processes. We here created a virtual experiment with constant, food dependent and temperature dependent mortality and tried to find back their proportions. We supposed that all parameters discussed so far are given and only mortality is to be explained. We supposed an inverse O'Neill-like mortality  $\mu_T(1 - \Phi(T))$  with mortality rate  $\mu_T$  and a parameterization identical to the one included in growth and a food dependency of  $\mu_F \frac{F_\mu}{F_\mu + F}$  with mortality rate  $\mu_F$  and a shape parameter  $F_\mu$ . Together with a constant mortality of  $\mu_0$  we have a survival rate  $\sigma$  of

$$\sigma = (1 - \mu_0)(1 - \mu_T(1 - \Phi(T)))(1 - \mu_F \frac{F_\mu}{F_\mu + F})$$

such that

$$p_-(F, T) = (1 - \sigma)n(w, t).$$

Results differed depending on whether to suppose knowledge of single process shape parameters ( $F_\mu$ ,  $T_{\text{opt}}$  and  $q_{10}$ ), or not, the latter is called 'all' in the table below, the former is called 'rates only'.

	$\mu_0$	$\mu_F$	$\mu_T$
direct	0.0050	0.0010	0.0010
rates only	0.0040	0.0012	0.0041
all	0.0045	0.0000	0.0045

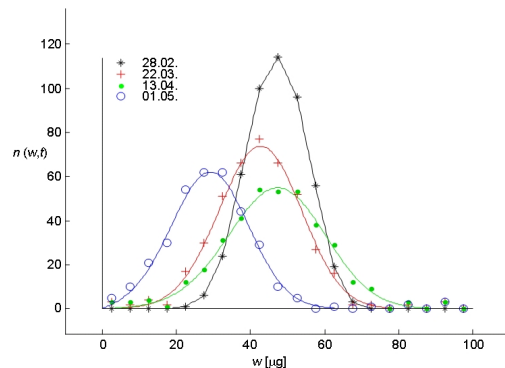


Figure 3: Weight structured population densities taking into account temperature and food effects, especially on mortality, optimized parameter set

It was not possible to find back the true influence of starving related death without previous knowledge of  $F_\mu$ , compare an estimation of  $\mu_F$  of 0.0000 to 0.0012 while direct simulation input was 0.0010. Besides, the

rôle of temperature related death was over-estimated, again we suppose that a lack of temperature amplitude is one possible reason. A model efficiency of 0.99 was found and is illustrated in figure 3

#### 4.4 Coupled systems analysis

Year long experiments in nature comprise monitoring of newborn animals, e.g. unpublished data by Schneider et al. Model equations then comprise  $p_+(F, T, \dots)$ . Other experiments distinguish between female and male animals and focus on differing rates for anabolic and catabolic processes and mortality rates. Any of these modifications results in multi-experiment analysis. Virtual tests performed so far show a fairly acceptable reproduction of the population densities as long as environmental components were excluded or were very pronounced. However, parameter attribution in real life circumstances is difficult as all possible environmental components, like temperature and food supply, will have to be monitored collaterally.

#### 4.5 Application in real life experiments

In data so far unpublished Suhling et al. analyzed the influence of increasing temperature as one effect of climate change on mean weight of *G. pulex*. They included four temperature regimes, one being true weather conditions, the others differed to this by two, four and six degree Celsius respectively. A single food reservoir was furnished and mean weight was monitored during little more than sixty days.

We supposed that all of the processes described in the previous subsections are relevant. Differences in dynamic behaviour between the four regimes would then be fully explained by temperature dependency of both anabolic processes and death. As food stock will be depleted depending on the rate of anabolic processes we suppose that food dynamic shape parameters may vary between the four temperature regimes.

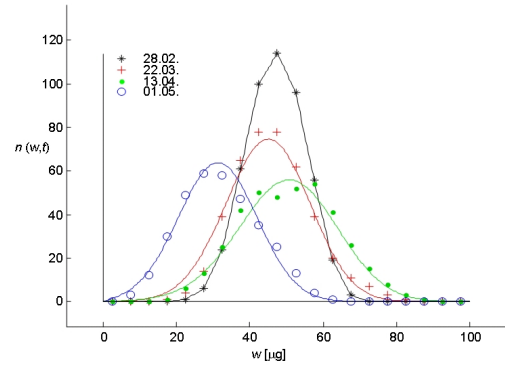


Figure 4: Optimized parameter set for experimental data with outside temperature

Including all of the processes given in the previous subsection we found impressive model efficiencies of 0.98 when optimizing all parameters for the outside temperature regime, see figure 4. The overall dynamic of the measured and modelled population density is identical and close to what has been modelled during virtual experiments in section 4.3. However, the bigger the temperature difference to the latter, the smaller the model efficiency, see figure 5 and the table below. The series of shrinking, growth and shrinking is amplified with rising temperature, and death is rarer than modelled.

	+0°C	+2°C	+4°C	+6°C
ME	0.98	0.79	0.687	0.535

Not shown here but even worse are model efficiencies when only food dynamic parameters are left to be optimized to the temperature regimes of plus two, four and six degree Celsius.

We therefore conclude that another temperature dependent process is hidden in the experiment. Our final clue to this assumption is that even direct simulation of the experimental set up with the parameters found for basis temperatures did not show the same dynamic as revealed by the experiments.

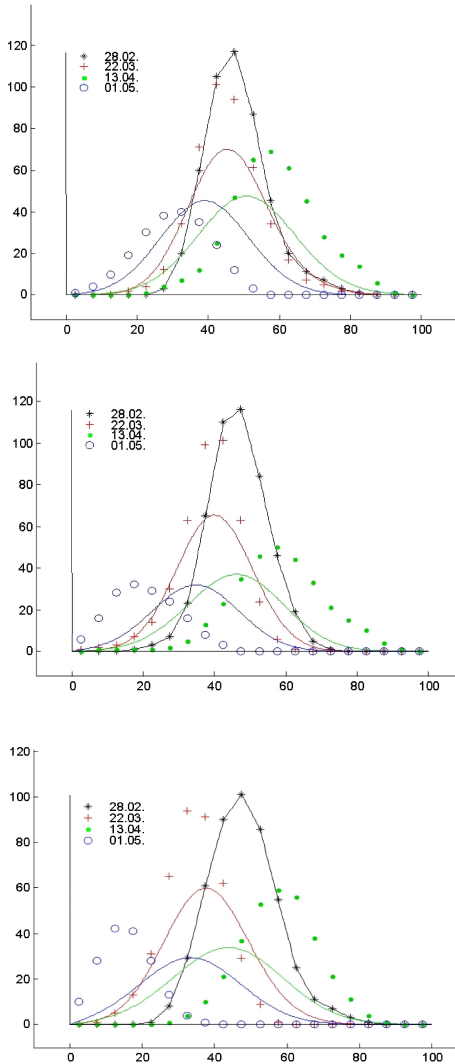


Figure 5: Optimized parameter set for experimental data with  $+2^{\circ}\text{C}$ ,  $+4^{\circ}\text{C}$ , and  $+6^{\circ}\text{C}$  respectively

## 5 Conclusion

Combining direct simulation with Cmsol with simple optimization tools in Matlab helps in research in ecology in two ways. On the one hand it is an efficient way to analyse coupled processes found in integral experi-

ments, on the other hand it can guide researchers to design efficient experiments for single processes in varying contexts.

## References

- [1] J. E. Nash and J. V. Sutcliffe, *River flow forecasting through conceptual models part i a discussion of principles*, Journal of Hydrology **10** (1970), 282–290.
- [2] Camille Parmesan, *Ecological and evolutionary responses to recent climate change*, Annual Review of Ecology, Evolution, and Systematics **37** (2006), 637–669.
- [3] Manfred Pöckl, Bruce Webb, and David W. Sutcliffe, *Life history and reproductive capacity of gammarus fossarum and g. roeseli under naturally fluctuating water temperatures: a simulation study*, Freshwater biology **48** (2003), 53–66.
- [4] D. W. Sutcliffe, T. R. Carrick, and L.G. Willoughby, *Effects of diet, body size, age and temperature on growth rates in the amphipod gammarus pulex*, Freshwater biology **11** (1981), 183–214.
- [5] Shripad Tuljapurkar and Hal Caswell, *Structured-population models in marine, terrestrial, and freshwater systems*, vol. 21, Chapman & Hall, 1997.
- [6] L. von Bertalanffy, *Quantitative laws in metabolism and growth*, Quarterly Review of Biology **32** (1957), 217–231.

## Acknowledgements

The authors would like to thank PD Dr. Frank Suhling for freely sharing with us both the data he acquired by the experiments described above and his knowledge in ecology.