A punny title

A less punny subtitle

Amisha Bhojwani

1st December, 2020

MSc Computational Methods in Ecology and Evolution

Imperial College London 2020-21

Wordcount: 2849

Contents

Abstract

1   Introduction

A functional response is the relationship between consumer predation rate and prey density [13]. Holling (1959) [4] described three functional response models which stem from scenarios that he termed "artificial situations". From these controlled experiments he gathered that there are two behaviours present in all predator-prey situations; time spent searching for prey (a) and the time spent handling prey (h). These insights led to him defining a general functional response equation:

$$y = \frac{ax}{1 + ahx} \tag{1}$$

Where y is consumption rate (resources consumed per time unit and predator) and x is prey density (number of individuals per unit of area or volume). It is important to note that Hollings search rate (a) implies a 100% success rate and is often also referred to as such [4]. The above equation can also be referred to as Holling's disc equation or, what he later defined as, a Type II functional response equation. The curve from this equation presents a negatively accelerated rise to a plateau, unlike the Type I functional response which he describes as a linear rise to a plateau. A Type III functional response is also described, and it accounts for another parameter, according to Holling it is a learning component, that results in an S-shaped curve.

Holling (1959) [4] denotes that the disc equation is general and malleable enough to accommodate interaction specific parameters which may help further explain biological data. Aljetlawi (2004) [1] emphasises this by explaining how the components of an interaction vary between and within species not only on an individual level but also at community level, meaning that the terms a and h in the disc equation may have interaction-specific biological meanings and could encompass more than one mechanism.

The artificial predation systems used to derive Holling's disc equation do not reflect natural ones and the model is therefore considered to be unrealistic by some [6]. To overcome this, we must consider that there are factors which can interfere and influence the biological meaning Holling attributed to his parameters a and h which can be used to explain biological data in more detail [1]. For example, digesting is different from handling but may influence a predator's reason behind searching for new prey. The extensions to the original disc equation made in Holling's 1966 publication [5] are the only ones to treat digestion as a factor of influence to foraging activity due to its relationship with prey density; prey density will determine satiation levels which will, in turn, drive a predator's search rate. Furthermore, in the disc equation, eating time is modelled together with handling time [6], so more detail is required in describing the biological mechanisms that a single parameter represents, as

the mathematical relationship with other parameters may change as a consequence. These considerations point towards the need of fine-tuning the disc equation to the specificity desired in an interaction and incorporating parameters, such as satiation, to further expand the biological meaning the model can provide.

Foraging theory is not the only theoretical framework to be considered among the literature for functional response models; mechanistic models proposed vary in form and nature and tend to branch out from the disc equation. A family tree of functional response models is available in the literature [6], and from it we note the possibility to modulate between models. For example, the disc equation can be modified to include a parameter q that helps modulate from a Type II to a Type III response [8, 9]:

$$y = \frac{ax^{1+q}}{1 + ahx^{1+q}} \tag{2}$$

The value $q = 1$ would mean the equation takes the form of Holling's Type III functional response, whereas at $q = 0$ it equates to a Type II functional response. Jeschke (2002) [6] argues that this equation should be considered phenomenological as it doesn't explain the mechanisms at play in a Type III response, particularly, that the added parameter (q) is not explained mechanistically as needed.

In this study I analyse how well the disc equation (1) can describe data from a multiplicity and variety of species interactions in a large dataset, and compare it to two phenomenological approaches. I use the equation from (1) and do not consider the approach in (2) in order to avoid mixing mechanistic and phenomenological approaches within the same model. The interactions in the studied dataset range from vertebrates, to invertebrates to plants and encompass both terrestrial and aquatic environments, attributing it a wide range of taxa to work on. I do not consider the Holling Type III model in this study because, if the Type II functional response is the most observed in the literature [6], I expect that the mechanistic Holling Type II model will describe a high percentage of our data, moreover, it will do so more appropriately than a phenomenological approach would allow.

I will also investigate if the final confidence set of models for the data would differ depending on the cut-off values used to declare it, as there is discussion that the general rule of $\Delta_{AIC} \leq 2$, where the models in the confidence set have to be less than two units away in AIC from the model with the minimum AIC, is too restrictive and might avoid including the truly most parsimonious model in the confidence set [11, 10].

2   Methods

2.1   Data analysis

The dataset considered was collected from lab and field experiments worldwide and holds the rates of prey consumption by consumers from a variety of species interactions. There's over 50 fields of metadata for this dataset but for the purposes of this report I used fields containing the number of resources consumed per predator per unit time and the resource density, the latter of which varied in units depending on the dimensionality of the interaction studied. The dataset has 308 individually identified interactions (ID's) and each one has at least 4 observations, 142 being the most observations for one interaction. We must note that there might be more than one ID with observations for a particular species interaction, this may be because the data comes from different sources and may represent the interaction in different contexts (laboratory or field settings). Also noteworthy is the existence of replicates in the data, which affect the fitted models as we see in Fig. 1.

I subset the data by individual interactions and fit three models to each subset: a quadratic linear model, a cubic linear model and the Holling Type II functional response model [4]. The two first models are phenomenological and I fit them using Ordinary Least-Squares (OLS) methods whilst the last is mechanistic and I used Levenberg–Marquardt Non-Linear Least-Squares (NLLS-LM) methods to fit it.

The Type II model works with two parameters, as seen in (1), and I estimated them using parametrizations before modelling [8, 12]. The parameter h is the inverse of the maximum feeding rate ($h = 1/F_{max}$) and the parameter a is the maximum feeding rate divided by the prey density at which the predator consumes half the amount of its feeding rate ($a = F_{max}/N_{Half}$). The parameter estimations were then sampled from a uniform distribution around the estimated values to further improve their approximation to the data. We bound the parameters at 0 for this to be their lowest possible value, as we know there will never be negative attack rates or handling times.

To compare the three competing models, as well as select the sampled parameters, I used model selection criteria because it considera not just how well the model fits the data but also how complex the model is. Furthermore, this criteria allows us to compare more than two models simultaneously. To compare only the fit of the models, I extracted R^2 values from each model fit [7].

I calculated Akaike's Information Criterion (AIC) derived for small sample sizes (AICc) by Sugiura (1978) and Hurvich and Tsai (1989) [3]. I also calculated Schwartz Criterion for the purposes of model selection. AIC measures the amount of Kullback-Leibler (KL) information loss, accounting for lack of fit, bias correction and model complexity. Schwartz Criterion is also known as the Bayesian Information Criterion (BIC) and is superficially similar to AIC except it includes an expression for sample size dependency and is not based

on KL information theory. I used AICc for our model selection rather than AIC because the number of parameters used in each model exceeded $n/40$, where n is sample size, so AICc allows me to account for the smaller sample sizes of the subsets of data [7]. Furthermore, AICc converges to AIC when sample sizes are large so to use it could avoid errors if one hasn't accounted for sample sizes [3].

According to the literature model selection with model selection criteria such as AIC can work with cut-off values to identify a confidence set of models among which the most parsimonious model will be 95% of the time. Generally, it is considered that if a model has a difference in AIC value less than than 2 ($\Delta \leq 2$) from the model with the smallest AIC value, it falls within the confidence set of models [2]. However, Richards (2011) argues that this consideration is too restrictive and may elude the truly most parsimonious model from the confidence set. Simulations show that for the truly most parsimonious model, according to the data, to be in the confidence set, it should be that $\Delta \leq 6$ [10]. For this study, I considered both these approaches alongside a no-cut-off approach, where the model with the lowest AIC is the most parsimonious in the a priori set of models.

2.2 Computing tools

For the entirety of data analysis part of this project I used R version 4.0.3. R language makes modelling possible in very few lines of code with user friendly and intuitive syntax. The RStudio IDE allows for easy , quick and readily visual data exploration. I point out the visual exploration both because of the integrated graphics device and because packages such as *ggplot2* which make quick plotting very accessible. R also boasts of a rich packages library for statistics, one of which I used for this project: *minpack.lm*. I wanted to keep the dependencies on this project relatively minimal so I havent used any other packages aside from the ones I've mentioned here on R. For writing this report I've used Latex, pdfTeX version 3.14159265-2.6-1.40.20 (TeX Live 2019/Debian), because it's useful to format reports that include figures and data that update in real time. It also has a convenient reference management method, *BibTeX*. Packages I've used to write this report include:

- *graphicx* to include graphics in the report
- *setspace* to indicate text spacing and margins of the document
- *lineno* to enumerate each line
- *amsmath* for writing equations, referencing them and writing in-line maths
- *csvsimple* allows importing of csv files to produce Latex tabulars
- *booktabs* allowed me to format tables

To count the words in the report I used a script in Perl version v5.30.0 called *texcount*. The last computing language I used is Bash, GNU bash version 5.0.17(1), to write a shell

script that compiles the entire project because the syntax is succint.

3   Results

Surprisingly, the phenomenologic cubic model seems to fit the data better than the mechanistic model chosen, as evidenced by the higher proportion of times when it's chosen according to the R^2 values in table 1. BIC values again generally favour the cubic model but do not differ much from the support for the Holling Type II model (table 1).

	Quadratic	Cubic	HollingT2
R-squared	0	0.92	0.08
BIC	0.21	0.41	0.38

Table 1: Proportion of times the model is chosen as: the model with the best fit (highest R^2) or the most parsimonious model (lowest BIC).

Although the R^2 values are high for the quadratic model, for the most part $R^2 \sim 1$, the remaining models manage to have an almost imperceptibly higher fit for most subsets of data. Despite not having the best fit of the three, the quadratic model presents the lowest BIC, deeming it the most parsimonious model of the three. The cubic model's BIC is similar to that of the Holling Type II model, both being chosen as the most parsimonious models at a rate of ~ 0.40.

The model fits are best for the cubic model in subsets with the lowest sample size $(n = 4)$, and the model is chosen as the most parsimonious according to the model selection criteria. In these cases, the model is able to fit the data perfectly $(R^2 = 1)$ and the AIC and BIC values are $-Inf$ (panel two in Fig. 1).

|	No cut-off	Cut-off 2	Cut-off 6
Quadratic	0.14	0.06	0.02
Cubic	0.1	0.08	0.06
HollingT2	0.76	0.65	0.44
Overall	1	0.8	0.51

Table 2: Proportion of times that when said model has the lowest AIC, it is the only one in the confidence set. Conversely, the inverse of the terms represented indicate that there are at least two models in the confidence set. When there is no cut-off, the model that has the lowest AIC is considered the only one in the confidence set. The "Overall" fields are irrespective of the model type, for example overall, when there is no cut-off, there is one most parsimonious model per subset of data all of the time. An example on how to read the table would be as follows: the quadratic model has the lowest AIC value $\sim 15\%$ of the time, if we consider a cut-off at 2 approach $(\Delta \leq 2)$, the quadratic model will be the most parsimonious model, standing alone in the confidence set, $\sim 5\%$ of the time. In contrast, if the cut-off is $\Delta \leq 6$, it will rarely remain alone in the confidence set ($\sim 1\%$ of the time).

In terms of AIC, a general trend we see is that a lower cut-off translates to considering only one model in the confidence set $\sim 80\%$ of time, whereas with a higher cutoff this value is lessened to $\sim 50\%$ of times. If we are considering the former, that cut-off values are $\Delta \leq 2$, the Holling Type II model remains the only model in the confidence set $\sim 85\%$ of the times

that it is the model with the lowest AIC for a subset. Furthermore, when contrasted with the other models, and regardless of the cut-off type, Holling's Type II model is consistently chosen as the most parsimonious model for the data at higher rates than the cubic and quadratic models (table 2).

Visual representations of the different fits can be found in the "Results" directory of this project under a subdirectory named "plots". In this paper we include three examples. Firstly, we see times where the fits of all three models are almost indistinguishable (panel 1 Fig. 1). Secondly, the cases where AIC and BIC values for the cubic model were $-Inf$ happens in subsets like the one in panel 2 of Fig. 1, where the sample size was small ($n = 4$). And thirdly, an example of a subset having replicates for select values of resource densities can be seen in the last panel of Fig. 1.
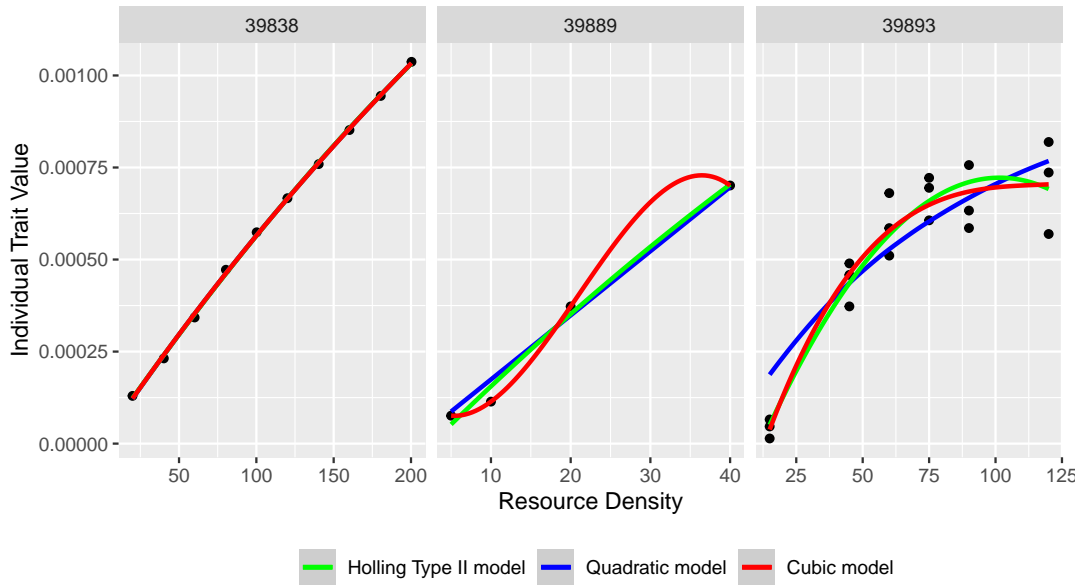
Figure 1: Plots showing the fit of all three models in the a priori model set for three ID numbers. The interaction in ID number 39838 is between *Rhyacophila dorsalis* (consumer) and *Chironomus spp* (resource). The interaction in ID number 39889 is between *Plectrocnemia conspersa* (consumer) and *Nemurella picteti* (resource). The interaction in ID number 39893 is between *Peromyscus maniculatus* (consumer) and *Triticum spp.* (resource).

4   Discussion

From our analyses we can say that the Holling Type II model receives the most support when considering AIC values but the quadratic model receives the most support when considering BIC. Both AIC and BIC are able to incorporate fit and model complexity in their index [7], so there must be a reason for the disagreement between both selection criteria.

Over 70% of the subsets of data have 10 or less observations. The subsets are generally too small for BIC to be reliable for model selection because, for this index, the "true model" is identified with maximum probability as the sample size increases to infitinity (Schwarz

1978 cited in [3]). Another BIC caveat is the notion that the true model is in the a priori set of models under consideration, if this is not the case the criterion is inconsistent [3]. We did not include interaction specific models in the initial set of models, so it may well be that for certain subsets BIC is unreliable because the true model was in fact not considered in the initial set of models.

The mistaken analysis of small sample sizes did not only affect us in considering BIC and appropriate model selection criteria, it also led to overfitting of the cubic model to the data as seen in the second panel of Fig. 1.

These mistakes in model inferencing can be palliated by looking at our AICc values. When we computed AICc values we accounted for small sample sizes. We also know that AIC indexes work with KL Information Theory (IT) and tell us about the amount of real information lost to the model [7]. In this way, the assumption that the true model is in the a priori model set is avoided. Furthermore, AIC looks to maximise fit and minimise information loss, criteria that are attractive for multi-model inferencing.

Now fully trusting our AICc and R^2 values, we can say that while the cubic model presents the best fit $\sim 90\%$ of the time, the Holling Type II model is the most parsimonious among the ones considered at a higher rate than the quadratic and cubic linear models, a finding consistent with our initial hypothesis.

5   Conclusion

References

[1] Albashir A. Aljetlawi, Erik Sparrevik, and Kjell Leonardsson. Prey–predator size-dependent functional response: derivation and rescaling to the real world. *Journal of Animal Ecology*, 73(2):239–252, 2004.

[2] Kenneth P Burnham and David R Anderson. *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*. Springer New York, New York, NY, 2. edition, 2002.

[3] K.P. Burnham, D.R. Anderson, and Kathryn (Kate) Huyvaert. Aic model selection and multimodel inference in behavioral ecology: some background, observations, and comparisons. *Behavioral Ecology and Sociobiology*, 65:23–35, 01 2011.

[4] C. S. Holling. Some characteristics of simple types of predation and parasitism. *The Canadian Entomologist*, 91(7):385–398, 1959.

[5] C. S. Holling. The functional response of invertebrate predators to prey density. *Memoirs of the Entomological Society of Canada*, 98(S48):5–86, 1966.

[6] Jonathan M. Jeschke, Michael Kopp, and Ralph Tollrian. Predator functional responses: Discriminating between handling and digesting prey. *Ecological Monographs*, 72(1):95–112, 2002.

[7] Jerald Johnson and Kristian Omland. Johnson jb, omland ks. model selection in ecology and evolution. trends ecol evol 19: 101-108. *Trends in ecology and evolution*, 19:101–8, 03 2004.

[8] Leslie A. Real. The kinetics of functional response. *The American Naturalist*, 111(978):289–300, 1977.

[9] Leslie A. Real. Ecological determinants of functional response. *Ecology*, 60(3):481–485, 1979.

[10] Shane Richards, Mark Whittingham, and Philip Stephens. Model selection and model averaging in behavioural ecology: The utility of the it-aic framework. *Behav. Ecol. Sociobiol.*, 65:77–89, 08 2011.

[11] Shane A. Richards. Dealing with overdispersed count data in applied ecology. *Journal of Applied Ecology*, 45(1):218–227, 2008.

[12] Benjamin Rosenbaum and Björn C. Rall. Fitting functional responses: Direct parameter estimation by simulating differential equations. *Methods in Ecology and Evolution*, 9(10):2076–2090, 2018.

[13] M. E. Solomon. The natural control of animal populations. *Journal of Animal Ecology*, 18(1):1–35, 1949.