Geometric complexity and the information-theoretic comparison of functional-response models

Mark Novak[1] & Daniel B. Stouffer[2]

[1]Department of Integrative Biology, Oregon State University,
Corvallis, Oregon, 97331 USA
[2]Centre for Integrative Ecology, School of Biological Sciences,
University of Canterbury, Christchurch 8140, New Zealand

RUNNING TITLE: Flexibility of functional-response models

Contents

1 Abstract

2 The assessment of relative model performance using information criteria like AIC and BIC has
3 become routine among functional-response studies, reflecting trends in the broader ecological
4 literature. Such information criteria allow comparison across diverse models because they penal-
5 ize each model's fit by its parametric complexity — in terms of their number of free parameters
6 — which allows simpler models to outperform similarly fitting models of higher parametric
7 complexity. However, criteria like AIC and BIC do not consider an additional form of model
8 complexity, referred to as geometric complexity, which relates specifically to the mathematical
9 form of the model. Models of equivalent parametric complexity can differ in their geometric
10 complexity and thereby in their ability to flexibly fit data. Here we use the Fisher Information
11 Approximation criterion to compare, explain, and contextualize how geometric complexity varies
12 across a large compilation of single-prey functional-response models — including prey-, ratio-,
13 and predator-dependent formulations — reflecting varying levels of phenomenological generality
14 and varying apparent degrees and forms of non-linearity. Because a model's geometric complex-
15 ity varies with the data's underlying experimental design, we also sought to determine which
16 designs are best at leveling the playing field among functional-response models. Our analyses
17 illustrate (1) the large differences in geometric complexity that exist among functional-response
18 models, (2) there is no experimental design that can minimize these differences across all models,
19 and (3) even the qualitative nature by which some models are more or less flexible than others
20 is reversed by changes in experimental design. Failure to appreciate geometric complexity in the
21 empirical evaluation of functional-response models may therefore lead to biased inferences for
22 predator–prey ecology, particularly at low experimental sample sizes where the relative effects
23 of geometric complexity are strongest. We conclude by discussing the statistical and epistemo-
24 logical challenges that geometric complexity poses for the study of functional responses as it
25 relates to the attainment of biological truth and predictive ability.
26 **Keywords:** *consumer-resource interactions, model comparison, model flexibility, structural*
27 *complexity, nonlinearity, experimental design, mutual predator effects, Fisher information, pre-*
28 *diction, statistical inference*

Contribution to Field Statement

The use of criteria like AIC and BIC for selecting among functional-response models is now standard, well-accepted practice, just as it is in the ecological literature as a whole. The generic desire underlying the use of these criteria is to make the comparison of model performance an unbiased and equitable process by penalizing each model's fit to data by its *parametric complexity* (relating to its number of free parameters). Here we introduce the Fisher Information Approximation criterion to the ecological literature and use it to understand how the *geometric complexity* of models — a form of model complexity relating to a model's functional form that is not considered by criteria like AIC and BIC — varies across a large (near exhaustive) compilation of 40 different single-prey functional-response models. Our results add caution against the simplistic use and interpretation of information-theoretic model comparisons for functional-response experiments, showing just how large an effect that geometric complexity can have on inferences of model performance. We therefore use our work to help clarify the challenges that ecologists studying functional responses must face in the attainment of biological truth and predictive ability.

4

1   Introduction

Seek simplicity and distrust it.

Alfred North Whitehead, *The Concept of Nature*, 1919.

The literature contains thousands of functional-response experiments (DeLong & Uiterwaal, 2018), each seeking to determine the relationship between a given predator's feeding rate and its prey's abundance. In parallel, dozens of functional-response models have been proposed (Jeschke *et al.*, 2002; Table 1), each developed to encapsulate aspects of the variation that exists among predator and prey biologies. The desire to sift through these and identify the "best" model on the basis of data is strong given the frequent sensitivity of theoretical population-dynamic predictions to model structure and parameter values (e.g., Aldebert & Stouffer, 2018; Fussmann & Blasius, 2005). Information-theoretic model comparison criteria like the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC) have rapidly become the preeminent tool for satisfying this desire in a principled manner (Okuyama, 2013), mirroring their increasing ubiquity across the ecological literature as a whole (Aho *et al.*, 2014; Ellison, 2004; Johnson & Omland, 2004). Generically, criteria like AIC and BIC make the comparison of model performance an unbiased and equitable process. For standard linear regression models (and most other models), increasing model complexity by including additional free parameters will always result in a better fit to the data. Therefore, by the principle of parsimony or because such increases in fit typically come at the cost of generality beyond the focal dataset, model

63 performance is judged by the balance of fit and complexity (Burnham & Anderson 2002; Höge

64 *et al.* 2018; but see Evans *et al.* 2013; Coelho *et al.* 2019).

65   While differing fundamentally in their underlying philosophies, motivations, and assumptions

66 (Aho *et al.*, 2014; Höge *et al.*, 2018), both AIC and BIC implement the balance of fit and

67 complexity in a formal manner by penalizing a model's likelihood with a cost that depends on

68 its number of free parameters. Specifically, for each model in the considered set of models,

$$\text{AIC} = -2\ln \mathcal{L}(\theta_{mle}|y) + 2k \tag{1}$$

69 and

$$\text{BIC} = -2\ln \mathcal{L}(\theta_{mle}|y) + k\ln(n), \tag{2}$$

70 with the model evidencing the minimum value of one or the other criterion being judged as

71 the best-performing model. For both criteria, the first term is twice the model's negative log-

72 likelihood (evaluated at its maximum likelihood parameter values θ_{mle}) given the data y. This

73 term reflects the model's goodness-of-fit to the data. The second term of each criterion is a

74 function of the model's number of free parameters k. This term reflects a model's parametric

75 complexity. For AIC, a model's complexity is considered to be independent of the data while

76 for BIC it is dependent on the dataset's sample size n; that is, BIC requires each additional

77 parameter to explain proportionally more for datasets with larger sample size. The statistical

78 clarity of the best-performing designation is typically judged by a difference of two information

79 units between the best- and next-best performing models (Burnham & Anderson, 2002; Kass &

80 Raftery, 1995).

6

81   An issue for criteria like AIC and BIC is that a model's ability to fit data is not solely a

82   function of its parametric complexity and mechanistic fidelity to the processes responsible for

83   generating the data. This can be problematic because all models — whether it be due to their

84   deterministic skeleton or their stochastic shell — are, to various degrees, phenomenological, and

85   therefore do not faithfully encode the mechanistic details responsible for generating data (see

86   also Connolly *et al.*, 2017; Hart *et al.*, 2018). Consequently, a given model may fit data better

87   than all other models even when it encodes the mechanisms or processes for generating the data

88   less faithfully.

89   One way in which this can happen is when models differ in their flexibility. A model's

90   flexibility is determined by its mathematical form and can therefore differ among models having

91   the same parametric complexity. For example, although the models $y = \alpha + \beta x$ and $y = \alpha x^{\beta}$

92   have the same number of parameters and can both fit a linear relationship, the second model

93   has a functional form that is more flexible in that it can also accommodate nonlinearities. In

94   fact, the second model may fit some data better than the first even if the first is responsible for

95   generating them. The chance of this happening will vary with the design of the experiment (e.g.,

96   minimizing noise and maximizing the range of x) and clearly decreases as sample size increases.

97   Unfortunately, sample sizes in the functional-response literature are often not large (Novak &

98   Stouffer, 2021), and the degree to which experimental design is important given the variation in

99   mathematical forms that exists among functional-response models has not been addressed.

100   Here our goal is to better understand the contrasting flexibility of functional-response models

7

101   and its impact on their ranking under the information-theoretic model-comparison approach. We

102   quantify model flexibility by geometric complexity (a.k.a. structural complexity) as estimated by

103   the Fisher Information Approximation criterion (Rissanen, 1996). Doing so for an encompassing

104   set of functional-response models across experimental designs varying in prey and predator

105   abundances, we find that geometric complexity regularly differs substantially among models

106   of the same parametric complexity, that differences between some models can be reversed by

107   changes to an experiment's design, and that no experimental design can minimize differences

108   across all models. Although choices among alternative functional-response models should be

109   informed by motivations beyond those encoded by statistical measures of model performance,

110   our results add caution against interpreting information-theoretic functional-response model

111   comparisons merely at face value.

2   Materials and Methods

2.1   Fisher Information Approximation

114   The Fisher Information Approximation (FIA) is an implementation of the Minimum Descrip-

115   tion Length principle (Rissanen, 1978) which Grünwald (2000) introduced as a means for making

116   model comparisons (see also Ly *et al.*, 2017; Myung *et al.*, 2006; Pitt *et al.*, 2002). The Mini-

117   mum Description Length (MDL) principle considers the comparison of model performance as a

118   comparison of how well each model can compress the information that is present in data, with

119   the best-performing model being the one that describes the data with the shortest code length.

120   In the extreme case of random noise, no compression is possible. FIA is asymptotically equiva-

8

121 lent to the normalized maximum likelihood which Rissanen (1996) derived to operationalize the

122 MDL principle, but is easier to implement (Myung *et al.*, 2006). It is computed for each model

123 as

$$\text{FIA} = -\ln \mathcal{L}(\theta_{mle}|y) + \frac{k}{2}\ln\left(\frac{n}{2\pi}\right) + \ln\int_{\mathcal{D}}\sqrt{\det I(\theta)}\,\mathrm{d}\theta\,, \tag{3}$$

124 where the first term is the negative log-likelihood of the model given the data, the second

125 term is a measure of a model's parametric complexity that is dependent on the data via the

126 sample size n (Fig. 1), and the third term is a measure of its geometric complexity (for which

127 we henceforth use the symbol $\mathcal{G}$). As detailed in Box 1, FIA's geometric complexity depends

128 only on the structure of the data (not on n), with structure referring to the range of potential

129 values that the model's parameters can take as well as the structure of the data's predictor

130 variables (i.e. the experimental design). The contribution of geometric complexity to a model's

131 FIA value consequently decreases with increasing sample size relative to the contributions of

132 the likelihood and parametric complexity. This makes the effect of geometric complexity of

133 greatest importance for datasets with low sample sizes. Because both parametric and geometric

134 complexity are independent of the data beyond its sample size and experimental design, the

135 potential importance of model flexibility to the information-theoretic ranking of models may

136 be assessed by comparing their parametric and geometric complexity values and by comparing

137 the geometric complexity values of models having the same parametric complexity. Because

138 FIA converges on half the value of BIC as n becomes large, a one-unit difference in geometric

139 complexity reflects a substantial impact on the relative support that two models of the same
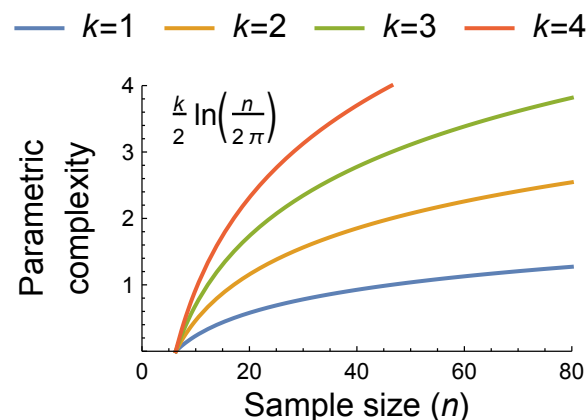
9

Figure 1: The dependence of parametric complexity on data sample size as estimated by the second term of the Fisher Information Approximation (FIA) for models with $k = 1, 2, 3$ and 4 free parameters. The potential importance of model flexibility to the information-theoretic ranking of functional-response models may be assessed by comparing their parametric and geometric complexity values or by comparing the geometric complexity values of models having the same parametric complexity because both measures of complexity are independent of the data beyond its sample size and structure (see main text and Box 1). For context, $n = 80$ was the median sample size of all functional-response datasets collated by Novak & Stouffer (2021).

140 parametric complexity could receive.

---

141 **Box 1. Unpacking the third term of the Fisher Information Approximation.**
142 The Fisher Information Approximation estimates the geometric complexity $\mathcal{G}_M$ of a
143 model M as the natural log of the integration (over all parameters θ) of the square root
144 of the determinant of the model's Expected unit Fisher Information matrix $I_M(\theta)$:

$$\mathcal{G}_M = \ln \int_{\mathcal{D}_M} \sqrt{\det I_M(\theta)}\, \mathrm{d}\theta\,. \tag{4}$$

148 The Fisher Information matrix $I_M(\theta)$ is a $k \times k$ matrix comprising the second-order
149 derivatives of the model's negative log-likelihood function with respect to each of
150 its parameters. It therefore reflects the sensitivies of the log-likelihood's gradient
151 with respect to those parameters. The Expected unit Fisher Information matrix is
152 the expected value of this matrix with a sample size of $n = 1$, computed with re-
153 spect to the dataset's independent variables (i.e. its experimental design). For exam-
154 ple, for a functional-response experiment having five prey-abundance treatment levels
155 $N \in \{10, 20, 30, 40, 50\}$ and a single predator-density level, the expectation is taken
156 by associating a $1/5^{th}$ probability to the Fisher Information matrix evaluated at each
157 treatment level.

158 The determinant of a matrix corresponds to its geometric volume. A larger determinant
159 of the Expected unit Fisher Information matrix therefore corresponds to a more flexi-
160 ble model that has higher gradient sensitivities for more of its parameters. Parameters
161 that share all their information — such as parameters that only appear in a model as
162 a product — result in matrix determinants of zero volume. Such non-identifiable mod-
163 els with statistically-redundant parameters require re-parameterization. Models can
164 also be non-identifiable because of experimental design, such as when there is insuffi-
165 cient variation in predictor variables. For example, all predator-dependent functional-
166 response models will be non-identifiable for designs entailing only a single predator
167 abundance level (see Fig. S1).

168 The domain $\mathcal{D}_M$ of the integral reflects the range of values that the model's parameters
169 could potentially exhibit. Three closely related issues are pertinent in this regard:

170 First, a closed-form solution of the indefinite integral in Eq. (4) may not exist, and
171 when it does it is often divergent. This means that numerical integration methods
172 are necessary and that parameter ranges must typically be bounded (i.e. the domain
173 $\mathcal{D}_M$ must be finite). However, how to specify bounds on mathematical grounds is not
174 always obvious. For example, for the ratio- and consumer-dependent models such as the
175 Hassell–Varley (HV) model, the interference strength parameter is not mathematically
176 limited but rather can take on any non-negative value to infinity if the "attack rate"
177 parameter is similarly unconstrained.

11

178 Second, for some experimental designs the range of parameter values may be more
179 empirically restricted than is mathematically permissible. For example, the handling
180 time of the Holling Type II (H2) model (and all other models) is mathematically
181 constrained only to be non-negative, and yet too large a handling time would mean
182 that no prey are ever expected to be eaten, an outcome few experimentalists would
183 consider useful. Similarly, too large an attack rate would prevent an experimentalist
184 from differentiating among models without the use of potentially intractable decreases
185 in an experiment's duration.

186 Third, because a model's geometric complexity reflects the range of parameter val-
187 ues which are considered possible, two models can exhibit different relative geometric
188 complexities for different experimental designs. However, different parameterizations
189 of the same functional form must have the same geometric complexity for a given
190 experimental design when the permissible range of their parameters is limited equiva-
191 lently (see Box 2). This is an issue because recognizing that two models simply reflect
192 alternative parameterizations is not always easy (e.g., contrast the original formula-
193 tion of the Steady State Satiation model by Jeschke *et al.* (2002) in Table S1 to our
194 reformulation in Table 1).

195 In our analyses, we overcome these three issues by imposing parameter constraints in
196 a manner that is indirect and equitable across all models. We do so by imposing the
197 same minimum and maximum constraints on the expected number of prey eaten for all
198 models, rather than on each model's parameters individually (see *Methods: Parameter*
199 *constraints*).

---

200 **Box 2. Imposing equitable integration limits.**

201 Different parameterizations of the same functional form should always have the same
202 geometric complexity for a given experimental design. However, this will only be true
203 when the range of their parameter values over which the integration of Eq. (4) is
204 performed is limited equivalently, which can be challenging. This issue is irrelevant
205 when solutions may be obtained in closed-form, but is not irrelevant when this is not
206 possible, as we suspect is the case for almost all functional-response models.

207 The challenge of determining equitable integration limits is well-demonstrated by a
208 comparison of the Holling and Michaelis–Menten Type II functional-response models
209 (Fig. 2). These are typically written as

$$F_{H2} = \frac{aN}{1 + ahN} \qquad \text{and} \qquad F_{MM} = \frac{\alpha N}{\beta + N}, \tag{5}$$

213 the equivalence of which is demonstrated by substituting $\alpha = 1/h$ (the maximum
214 feeding rate equals the inverse of the handling time) and $\beta = 1/(ah)$ (the abundance
215 at which half-saturation occurs is the inverse of the product of the attack rate and
216 handling time).

12

217 By definition, all four parameters (a, h, α and β) are limited only in that they

218 must be non-negative; they could each, in principle, be infinitely large (i.e. $\mathcal{D}_{H2} =$

219 $\{a \in [0, \infty), h \in [0, \infty)\}$ and $\mathcal{D}_{MM} = \{\alpha \in [0, \infty), \beta \in [0, \infty)\}$). If the integral in

220 Eq. (4) could then be computed analytically for the two models, we would always

221 obtain $\mathcal{G}_{H2} = \mathcal{G}_{MM}$ for any given experimental design.

222 However, because the integrals in Eq. (4) for the two models are divergent, finite

223 limits to $\mathcal{D}_{H2}$ and $\mathcal{D}_{MM}$ must be applied. At first glance, it may seem intuitive

224 to impose these limits on the maximum parameter values. For example, we might

225 consider imposing $a \in [0, a_{max}]$ and $h \in [0, h_{max}]$. Because of their inverse relation-

226 ships, doing so means that the equivalent limits for the Michaelis–Menten model are

227 $\alpha \in [1/h_{max}, \infty]$ and $\beta \in [1/(a_{max}h_{max}), \infty]$, which are not finite and hence cannot

228 solve our problem. Naively, we might therefore instead consider imposing both minima

229 and maxima, $a \in [a_{min}, a_{max}]$ and $h \in [h_{min}, h_{max}]$, so that $\alpha \in [1/h_{max}, 1/h_{min}]$ and

230 $\beta \in [1/(a_{max}h_{max}), 1/(a_{min}h_{min})]$. This, however, does not solve a further problem in

231 that the limits for β depend on the value of α (i.e. $1/h$). That is, we must also impose

232 the additional constraint that $\beta > \alpha/a_{max}$ (Fig. 2), for only then will the computed

233 $\mathcal{G}_M$ of the two models be equal.

234 Problems such as these only compound for models entailing a greater number of pa-

235 rameters. As alluded to in Box 1, our approach to circumventing these model-specific

236 issues is to impose constraints on the expected number of eaten prey (Fig. 2), rather

237 than on the model parameters directly (see *Methods: Parameter constraints*). That is,

238 we require that the minimum expected number of eaten prey is no less than one prey

239 individual in the maximum prey abundance treatment(s) (i.e. $1 \leq \mathbb{E}[F(N_{max}, P, \theta)PT]$

240 for all P in the experimental design) and that the maximum expected number of eaten

241 prey is no greater than N_{max} in any of the treatments (i.e. $\mathbb{E}[F(N, P, \theta)PT] \leq N_{max}$

242 for all $N \times P$ combinations in the experimental design). This imposes natural limits for

243 most (combinations of) parameters (e.g., the handling time or saturation parameters

244 of all models). For other parameters, it does not impose hard limits, but nonethe-

245 less results in their contribution to $\mathcal{G}_M$ tending asymptotically to zero as their value

246 increases (Fig. 2). This is most notably true for the "attack rate" parameter of all

247 models.

248 ## 2.2 Experimental designs

249 We computed the geometric complexity of 40 different functional-response models across a range

250 of experimental designs. We first describe the experimental designs we considered because

251 aspects of these also determined our manner for equitably bounding the permissible parameter
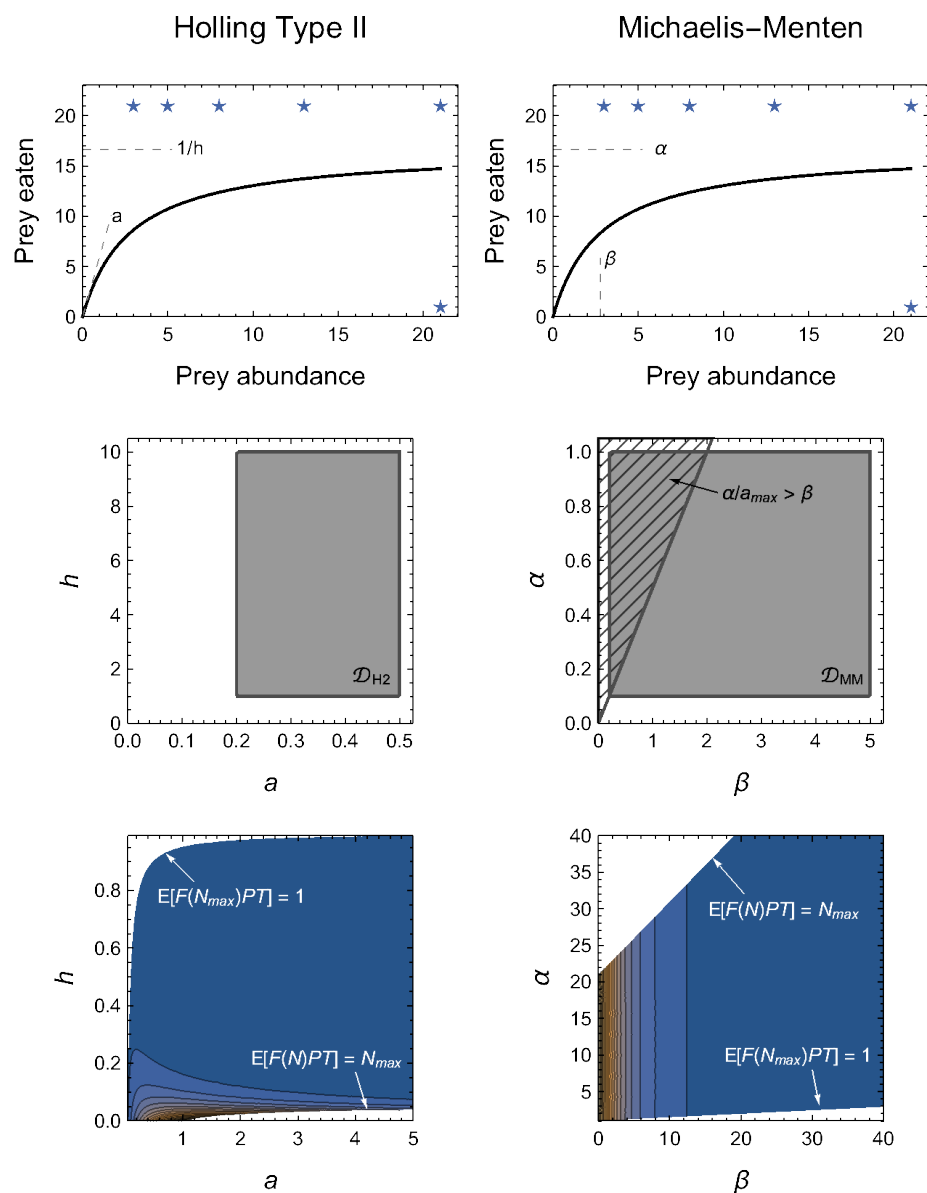
13

Figure 2: Alternative parameterizations of the same functional form should have the same geometric complexity for any given experimental design, but this will only be true in practice when their parameter domains $\mathcal{D}$ are equivalently constrained (see Box 2 for details). *Top row:* Illustration of the functional equivalence and parameter interpretations of the Holling (*left column*) and Michaelis–Menten (*right column*) models. *Middle row:* Direct constraints on $\mathcal{D}_{H2}$ and $\mathcal{D}_{MM}$ necessitate more than potentially arbitrary minimum and/or maximum limits, but must also account for the confounded relationships among parameters. *Bottom row:* We circumvent this challenge by imposing parameter constraints indirectly via the expected number of eaten prey, $\mathbb{E}[F(N, P, \theta)PT]$. (Stars in the top row indicate these limits imposed on the assumed experimental design. Colour-scale in bottom row reflects $\sqrt{\det I_M(\theta)}$ from dark blue (low values) to light orange (high values).

14

252 space of all functional-response models (Boxes 1 & 2).

253 The experimental designs we considered exhibited treatment variation in prey N and preda-

254 tor P abundances. All designs had at least five prey-abundance levels, a minimum prey-

255 abundance treatment of three prey individuals, and a minimum predator-abundance treatment

256 of one predator individual. The designs varied by their maximum prey and predator abun-

257 dances (N_{max} and P_{max}) which we achieved by correspondingly varying the number of prey

258 and predator treatment levels (L_N and L_P); that is, by including higher abundance levels to

259 smaller experimental designs. We specified the spacing between prey and predator abundance

260 levels to follow logarithmic series. This follows the recommendation of Uszko *et al.* (2020) whose

261 simulations showed that a logarithmic spacing of prey abundance levels performed well for the

262 purpose of parameter estimation. We used the golden ratio ($\phi = 1.618\ldots$) as the logarithmic

263 base and rounded to the nearest integer to generate logistically-feasible abundance series that

264 increase more slowly than typically used bases (e.g., $\log_2$ or $\log_{10}$). We thereby approximated

265 the Fibonacci series $(1, 1, 2, 3, 5, 8, \ldots)$ on which ϕ^n converges for large n. We varied L_N between

266 5 and 10 levels and varied L_P between 1 and 5 levels, thereby affecting N_{max} and P_{max} abun-

267 dances of up to 233 prey and up to 8 predator individuals. All resulting designs are depicted in

268 the *Supplementary Materials.*

269 An important aspect of experimental design which we assumed throughout our analyses

270 was that all eaten prey are continually replaced. The constancy of available prey allowed us

271 to treat observations as Poisson random variates and hence use a Poisson likelihood to express

15

each deterministic functional-response model as a statistical model. This was necessary because computing geometric complexity requires an inherently statistical perspective (see Box 1 and below).

2.3 Functional-response models

The functional-response models we considered ranged from having one to four free parameters (Table 1). We included prey-, ratio-, and predator-dependent models that are commonly assessed in the functional-response literature, as well as many models that have received far less attention, such as those that encapsulate emergent interference, adaptive behavior, or both handling and satiation. We did not consider models that explicitly include more variables than just the abundances of a focal predator-prey pair. Given that our statistical framework was based on experimental designs within which eaten prey are continually replaced, we also did not include any models which explicitly account for prey depletion or reflect the selection of hosts by non-discriminatory parasitoids (e.g., Rogers, 1972). All but two of the considered models are previously published. The exceptions were a three-parameter model (AS) which represents an illustrative generalization of the adaptive behavior A1 model of Abrams (1990), and a four-parameter predator-dependent model (SN2) that extends the Beddington–DeAngelis and Crowley–Martin models and may be interpreted as reflecting predators that cannot interfere when feeding and can partially feed when interfering (see Stouffer & Novak, 2021).

That said, we do not concern ourselves with the biological interpretation of the models as this has been discussed extensively throughout the functional-response literature. Rather, we

16

focus on the models' contrasting mathematical forms. Across the different models, these forms include rational, power, and exponential functions, as well as functions that are linear, sublinear, or superlinear with respect to prey or predator abundances. To highlight their similarities, we reparameterized many models to "Holling form", noting that different parameterizations of the same functional form have the same geometric complexity for a given experimental design (Box 2). This included models that, as originally defined, had statistically-redundant parameters (e.g., the models of Abrams, 1990), were written in "Michaelis-Menten form" (e.g., Sokol & Howell, 1981), or were written with parameters affecting divisions (e.g., we replaced $1/c \rightarrow c$). This also included the Steady State Satiation (SSS) model of Jeschke *et al.* (2002) for which $\mathcal{G}_M$ could not be computed. Fortunately the SSS model can also be derived using the citardauq formula (rather than the quadratic formula) for which $\mathcal{G}_M$ could be computed and which further reveals its similarity to the adaptive behavior A2 model of Abrams (1990) and the predator-dependent model of Ruxton *et al.* (1992). For simplicity and to further clarify similarities among models, we present all model parameters using the symbols a, b, c, and d for non-exponent parameters and u and v for exponent parameters, noting that their biological interpretations frequently differ among models.

17

Table 1: The deterministic functional-response models we considered for describing the per predator rate at which prey are eaten as a function of prey abundance N, predator abundance P, and the parameter(s) θ. From these per predator rates, the total count of prey eaten corresponds to the functional response, $F(N, P, \theta)$, multiplied by the number of predators P and the time period T of the experiment. The number of parameters k refers to the number of free parameters in each model because only these determine the mean and variance of prey eaten under the Poisson likelihood which we assumed. Note that, where appropriate, we use $P - 1$ rather than P for Holling-type predator-dependent models because P represents a count of predators in our synthetic experimental designs (rather than a density) and predator individuals cannot interfere with themselves. Original parameterizations are provided in Table S1.

Name	Abbrev.	Model $F(N, P, \theta)$	Reference
One-parameter models ($k = 1$)			
Holling I	H1	aN	Lotka (1925); Volterra (1926)
Linear ratio-dependent	LR	aN/P	Arditi & Ginzburg (1989)
Barbier–Wojcik–Loreau I	BWL1	$a\sqrt{N}/\sqrt{P}$	Barbier *et al.* (2021)
Two-parameter models ($k = 2$)			
Holling II	H2	$\frac{aN}{1+abN}$	Holling (1959)
Michaelis–Menten	MM	$\frac{aN}{b+N}$	Michaelis & Menten (1913)
Holling III	H3	$\frac{aN^2}{1+abN^2}$	Holling (1965); Real (1977)
Hyperbolic tangent	HT	$\frac{1}{b}\tanh(abN)$	Jassby & Platt (1976)
Gause–Ivlev	GI	$\frac{1}{b}(1-\exp[-aN])$	Gause (1934); Ivlev (1955)
Gause–Ivlev–Aldebert	GIA	$\frac{1}{b}(1-\exp[-abN])$	Aldebert *et al.* (2016a;b)
Gutierrez–Baumgärtner	GB	$\frac{1}{b}(1-\exp[-aN/P])$	Gutierrez & Baumgärtner (1984)
Abrams 0	A0	$\frac{aN}{1+ab\sqrt{N}}$	Abrams (1982)
Abrams I	A1	$\sqrt{\frac{aN}{1+abN}}$	Abrams (1990)
Abrams III	A3	$\frac{a\sqrt{N}}{1+ab\sqrt{N}}$	Abrams (1990)

Name	Abbr.	Formula	Reference
Sokol–Howell	SH	$\frac{aN}{1+abN^2}$	Sokol & Howell (1981)
Arditi–Ginzburg	AG	$\frac{aN/P}{1+abN/P}$	Arditi & Ginzburg (1989); Sutherland (1983)
Cosner–DeAngelis–Ault–Olson	CDAO	$\frac{aN/\sqrt{P}}{1+abN/\sqrt{P}}$	Cosner et al. (1999)
AG–Kratina	AGK	$\frac{a(N/P)^2}{1+ab(N/P)^2}$	Kratina et al. (2009)
Rosenzweig	R	aN^u	Rosenzweig (1971)
Hassell–Varley	HV	aN/P^v	Hassell & Varley (1969)
Three-parameter models ($k = 3$)			
Holling–Real III	H3R	$\frac{aN^u}{1+abN^u}$	Real (1977)
Abrams–Stouffer	AS	$\left(\frac{aN}{1+abN}\right)^u$	*This study*
Hassell–Lawton–Beddington	HLB	$\frac{aN^2}{1+cN+abN^2}$	Hassell et al. (1977)
Monod–Haldane	MH	$\frac{aN}{1+cN+abN^2}$	Andrews (1968)
Tostowaryk	T	$\frac{aN}{1+abN+cN^3}$	Tostowaryk (1972)
Fujii–Holling–Mace	FHM	$\frac{aN\exp[dN]}{1+abN\exp[dN]}$	Fujii et al. (1986)
Abrams II	A2	$\frac{aN}{1+abN+\sqrt{acN(1+abN)}}$	Abrams (1990)
Steady State Satiation	SSS	$\frac{2aN}{1+a(b+c)N+\sqrt{(1+a(b+c)N)(1+a(b+c+4bc)N)}}$	Jeschke et al. (2002)
Ruxton–Gurney–de Roos	RGD	$\frac{2aN}{1+abN+\sqrt{(1+abN)^2+8ac(P-1)}}$	Cosner et al. (1999); Ruxton et al. (1992)
Beddington–DeAngelis	BD	$\frac{aN}{1+abN+c(P-1)}$	Beddington (1975); DeAngelis et al. (1975)
Crowley–Martin	CM	$\frac{aN}{1+abN+c(P-1)+abcN(P-1)}$	Crowley & Martin (1989)
Tyutyunov–Titova–Arditi	TTA	$\frac{aN}{1+abN+cP-(1-\exp[-cP])}$	Tyutyunov et al. (2008)
Barbier–Wojcik–Loreau II	BWL2	$aN^u P^{v-1}$	Barbier et al. (2021)
Arditi–Akçakaya	AA	$\frac{aN/P^v}{1+abN/P^v}$	Arditi & Akçakaya (1990)

Schenk–Bersier–Bacher	SBB	$\frac{a(N/P^v)^2}{1+ab(N/P^v)^2}$	Schenk *et al.* (2005)
Watt	W	$\frac{1}{b}\left(1-\exp[-aN/P^v]\right)$	Watt (1959)
Four-parameter models ($k = 4$)			
BD–Okuyama–Ruyle	BDOR	$\frac{aN^u}{1+abN^u+c(P-1)}$	Okuyama & Ruyle (2011)
CM–Okuyama–Ruyle	CMOR	$\frac{aN^u}{1+abN^u+c(P-1)+abcN^u(P-1)}$	Okuyama & Ruyle (2011)
AA–Okuyama–Ruyle	AAOR	$\frac{aN^u/P^v}{1+abN^u/P^v}$	Okuyama & Ruyle (2011)
Stouffer–Novak I	SN1	$\frac{aN}{1+abN+c(P-1)+abc(1-d)N(P-1)}$	Stouffer & Novak (2021)
Stouffer–Novak II	SN2	$\frac{aN(1+c(1-d)(P-1))}{1+abN+c(P-1)+abc(1-d)N(P-1)}$	*This study*

2.4 Parameter constraints

As mentioned above, we assumed a Poisson statistical model in computing the geometric complexity of each deterministic functional-response model. In a context of fitting models to actual data, the consequent log-likelihood function,

$$\ln \mathcal{L}(\theta|y) = -n\lambda - \prod_{i=1}^{n} \ln\left(y_i!\right) + \ln(\lambda) \sum_{i=1}^{n} y_i \,, \tag{6}$$

expresses the likelihood of a model's parameter values given the observed data, with $\lambda = F(N, P, \theta)PT$; that is, the feeding rate of a predator individual (as per the focal deterministic functional-response model) times the number of predators and the time period of the experiment, which we universally set to $T = 1$. In our context of quantifying $\mathcal{G}_M$, data is not needed because the second term of Eq. (6) drops out when taking derivatives with respect to model parameters and because $I_M(\theta)$ involves the expected value of y_i with respect to an experiment's prey and predator abundance levels (Box 1).

Despite this independence from data, additional information is nonetheless necessary for computing the geometric complexity of models such as those we consider here (Box 1). This information entails the range of values that a model's parameters could potentially exhibit (i.e. their permissible domain $\mathcal{D}_M$ of integration). Encoding this information in an equitable manner that does not bias the inferred geometric complexity of some models over others has several potential issues associated with it (Boxes 1 and 2).

To avoid these issues, we placed no direct constraints on the parameters themselves. Rather,

21

323 we specified infinite domains on the parameters (i.e. $\{a, b, c, u, v\} \in [0, \infty)$ and $d \in (-\infty, \infty)$)

324 and instead placed constraints on them in an indirect manner. Specifically, we imposed the re-

325 quirement that the expected number of eaten prey in all maximum prey abundance treatments

326 was no less than 1 (i.e. $1 \le \mathbb{E}[F(N_{max}, P, \theta)PT]$ across all P treatments) and that the expected

327 number of prey eaten in any treatment was no greater than the number of prey made available

328 in the maximum prey treatment level (i.e. $\mathbb{E}[F(N, P, \theta)PT] \le N_{max}$ for all $N \times P$ treatment

329 combinations). Under the assumed Poisson model, the lower bound corresponds to an expec-

330 tation of observing zero prey being eaten in no greater than 37% of an experiment's maximum

331 prey abundance replicates (since $\mathrm{P}(\mathbb{E}[F(N_{max}, P, \theta)PT] = 1) = 0.37$). The upper bound is

332 similarly arbitrary in a mathematical sense but seems logistically feasible since researchers are

333 unlikely to choose a prey abundance beyond which they could not continually replace consumed

334 individuals. For the SN1 and SN2 models, we imposed the respective additional requirement

335 that $bd \le 1/\max[F(N, P, \theta)PT]$ and $b \le 1/\max[F(N, P, \theta)PT]$ for all treatments to maintain

336 biologically-appropriate (non-negative) predator interference rates (Stouffer & Novak, 2021). We

337 note that our placement of constraints on the expected number of eaten prey is similar to the

338 use of Bayesian prior predictive checks with a joint prior distribution in that we restrict the

339 domain of permissible parameter values based on their conditional inter-dependencies.

340   It is worth noting that some authors defined their models with parameters to be greater than

341 1, rather than 0 as we did. For example, theoreticians often assume $u \ge 1$ for the Hill exponent

342 of the Holling–Real Type III (H3R) model, though Real (1977) did not do so. We consider

22

343 non-negative values less than one to also be biologically and statistically possible (see discussion

344 in Stouffer & Novak, 2021). Indeed, relaxing this constraint and redefining the statistically-

345 redundant parameters of the original A3 model (Abrams, 1990) clarifies, for example, that it is

346 equivalent to H3R with $u = 0.5$.

2.5 Model comparisons

348 Comparisons of geometric complexity can only be made across models of the same paramet-

349 ric complexity; it is in conjunction with its second term that FIA enables comparisons across

350 models in general. Therefore, for each set of models (i.e. for models with $k = 1, 2, 3$ or 4 pa-

351 rameters), we first assessed how an experiment's design determined the geometric complexity of

352 a selected "baseline" model. Because their relationships to each other and most other models

353 are readily apparent, we chose the Holling Type I (H1) model as the baseline for the $k = 1$

354 models, the Holling Type II (H2) model for the $k = 2$ models, the Holling–Real Type III (H3R)

355 and the Beddington–DeAngelis (BD) models for the $k = 3$ models (H3R for the prey-dependent

356 models and BD for the ratio- and predator-dependent models), and the Beddington–DeAngelis–

357 Okuyama–Ruyle (BDOR) model for the $k = 4$ models. We then compared the geometric

358 complexity of the other models within a given set to the set's baseline model(s) by calculat-

359 ing, for each experimental design, the difference between the two model's geometric complexity

360 values (e.g., $\mathcal{G}_{LR} - \mathcal{G}_{H1}$). This difference enables a direct evaluation of the degree to which a

361 model's flexibility influences its information-theoretic ranking because it has the same units of

362 information as the likelihood and parametric complexity terms of the FIA criterion. We refer to

363 models with a greater positive difference to the baseline model (or other comparator) as being

364 "favoured" because commonly-used criteria like AIC and BIC do not consider their more flex-

365 ible (more geometrically complex) functional form. By contrast, models with a more negative

366 difference will be ranked lower, hence "disfavoured", by such criteria.

2.6 Sensitivity to assumptions

368 We evaluated the sensitivity of our inferences to three aspects of experimental design, repeating

369 our analyses for designs that

370   1. varied in the number of prey and predator levels (L_N and L_P) but kept the maximum

371   prey and predator abundances constant at $N_{max} = 233$ and $P_{max} = 5$ (based on results

372   from the main analysis);

373   2. used arithmetically-uniform (rather than logarithmic) series of prey and predator abun-

374   dances; and that

375   3. relaxed the constraint on either the minimum or the maximum expected number of eaten

376   prey by an order of magnitude (i.e. $\mathbb{E}[F(N_{max}, P, \theta)PT] \geq 1/10$ or $\mathbb{E}[F(N, P, \theta)PT] \leq$

377      $10 N_{max}$).

378 All analyses were performed in Mathematica (Wolfram Research Inc., 2020) using the *Local*

379 *Adaptive* integration method and with precision and accuracy goals set to 3 digits.

24

3 Results

3.1 Baseline models & equivalent models

The geometric complexity $\mathcal{G}_M$ of all baseline models (H1, H2, H3R, BD and BDOR) increased with increasing N_{max} and decreasing P_{max} (Figs. 3–6). For these models, $\mathcal{G}_M$ varied more greatly across the considered variation in N_{max} than across the considered variation in P_{max}, with at most a very weak interactive effect occurring between these. The difference in $\mathcal{G}_M$ between the smallest and largest N_{max} for a given P_{max} varied from about 2 information units for the parametrically simplest H1 model to about 5 units for the parametrically most complex BDOR model, with the difference for the other baseline models being intermediate and roughly proportional to their number of free parameters.

As expected (Box 2), alternative parameterizations of the same functional form had the same $\mathcal{G}_M$ for all designs, with numerical estimation errors accounting for deviations from exact equivalence. This was demonstrated by H2 and MM as well as GI and GIA (Fig. 4), which differ only in the biological interpretation of their parameters. Likewise, all ratio-dependent models had the same $\mathcal{G}_M$ as their "corresponding" Holling-type models when there was no variation in predator abundances (e.g., $\mathcal{G}_{LR} \approx \mathcal{G}_{H1}$, $\mathcal{G}_{AG} \approx \mathcal{G}_{H2}$ and $\mathcal{G}_{AGK} \approx \mathcal{G}_{H3}$ when $P_{max} = 1$; Figs. 3–5).

3.2 One-parameter models

For the one-parameter models (Fig. 3), both $\mathcal{G}_{LR}$ and $\mathcal{G}_{BLW1}$ were always greater than $\mathcal{G}_{H1}$ (excepting when $P_{max} = 1$ for LR). The degree to which the linear ratio-dependent (LR) model was favoured relative to the Holling Type I (H1) model decreased with increasing N_{max} and

25

400 decreasing P_{max}. This was also true for the BLW1 model of Barbier *et al.* (2021) when $P_{max} \geq 3$,

401 but for $P_{max} < 3$ its difference to H1 increased with increasing N_{max}. The most equitable designs

402 capable of differentiating among all three models therefore consisted of only two predator levels

403 ($P_{max} = 2$), entailed a $\mathcal{G}_M$ difference among models of about 0.2 information units or more, and

404 slightly favoured LR more for small N_{max} but BWL1 more for large N_{max}. The least equitable

405 design entailed large P_{max} and small N_{max} and favoured LR and BWL1 over H1 by more than

406 1 and 0.8 information units, respectively.

3.3   Two-parameter models

408 There were four categories of two-parameter models qualitatively distinguished by whether they

409 exhibited equivalent, higher, lower or a design-dependent $\mathcal{G}_M$ relative to the H2 baseline model

410 (Fig. 4):

411   (i) MM was equivalent to H2 for all designs (as already mentioned above);

412   (ii) H3, HT, GI, GIA, SH, AG, AGK, GB, CDAO and R were favoured over H2 for all designs

        (had higher $\mathcal{G}_M$, excepting for $P_{max} = 1$ where $\mathcal{G}_{AG} \approx \mathcal{G}_{CDAO} \approx \mathcal{G}_{H2}$);

414   (iii) A0, A1 and A3 were disfavoured relative to H2 for all designs (had lower $\mathcal{G}_M$); and

415   (iv) HV was favoured over H2 by small N_{max} designs and disfavoured by large N_{max}, with large

        and small P_{max} designs respectively favouring and disfavouring it more greatly.

417 H3 was the only model for which the difference from H2 was insensitive to experimental design,

418 always being favoured by about 0.45 information units. For HT, GI, GIA, A0, A1 and A3, the
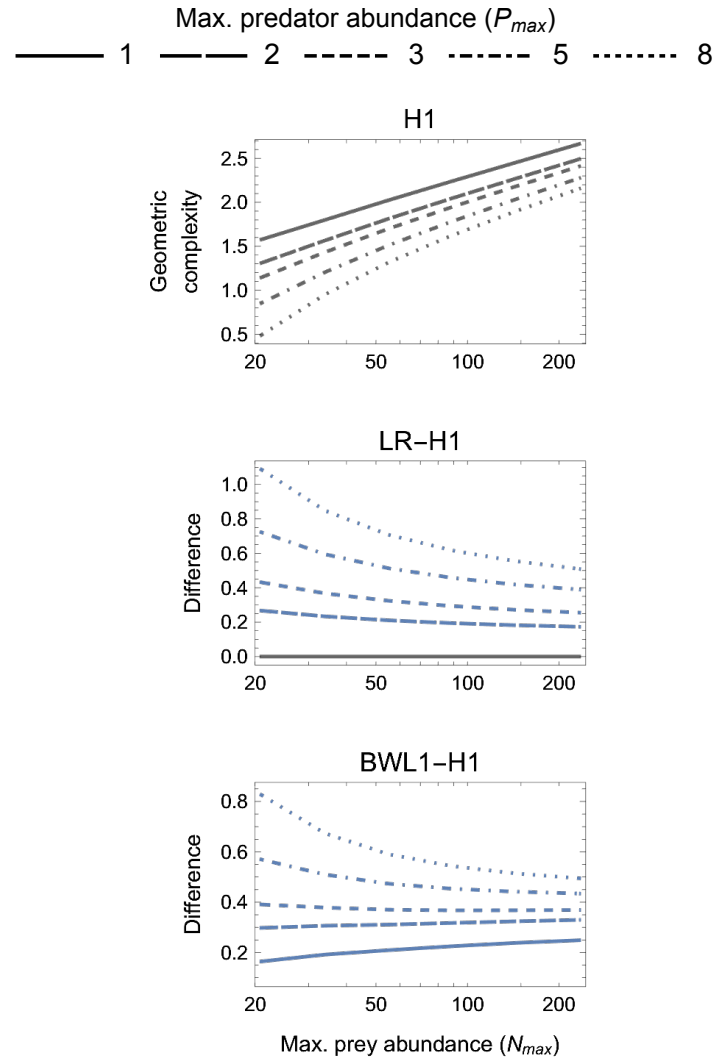
26

Figure 3: *First panel*: The geometric complexity $\mathcal{G}_{H1}$ of the single-parameter ($k = 1$) baseline Holling Type I (H1) model as a function of an experiment's maximum prey and predator abundances (N_{max} and P_{max}). *Other panels*: The difference in $\mathcal{G}_M$ of the linear ratio-dependent (LR) model and the square-root model of Barbier *et al.* (2021; BWL1) relative to the H1 model. Positive differences reflect experimental designs for which a focal model's mathematical flexibility would result in it being favoured by information criteria like AIC and BIC that do not consider this form of model complexity.

difference to H2 was insensitive to P_{max}, but while it increased with increasing N_{max} for HT, GI, GIA and A0 (making small N_{max} designs the most equitable), it decreased with increasing N_{max} for A1 and A3 (making large N_{max} designs the most equitable). The degree to which AG, AGK, GB, CDAO and R were favoured over H2 decreased with increasing N_{max}, but while it increased with increasing P_{max} for AG, AGK, GB and CDAO (making large N_{max}, small P_{max} designs the most equitable), it decreased — albeit weakly — with increasing P_{max} for R. For SH, the difference to H2 first increased from small to intermediate N_{max} then slowly decreased from intermediate to large N_{max}, but was always minimized by large P_{max}. Small N_{max}, large P_{max} designs were therefore the most equitable for SH. Finally, for HV, which was either favoured or disfavoured depending on design, the most equitable designs spanned $N_{max} \approx 30$ for $P_{max} = 2$ to $N_{max} \approx 120$ for $P_{max} = 8$. Overall, A0 and AGK exhibited the greatest potential disparity relative to H2, respectively being disfavoured and favoured by about 1.4 information units by their least equitable design. The greatest potential disparity among all considered two-parameter models was about 2 information units and occurred between HV and A0 for small N_{max}, large P_{max} designs in favour of HV.

3.4 Three-parameter models

Noting that all predator-dependent models are non-identifiable for $P_{max} = 1$ designs (Fig. S1), there were three categories of three-parameter models that were qualitatively distinguished by whether they exhibited higher, lower or a design-dependent $\mathcal{G}_M$ relative to the two baseline models — H3R for prey-dependent models and BD for ratio- and predator-dependent models
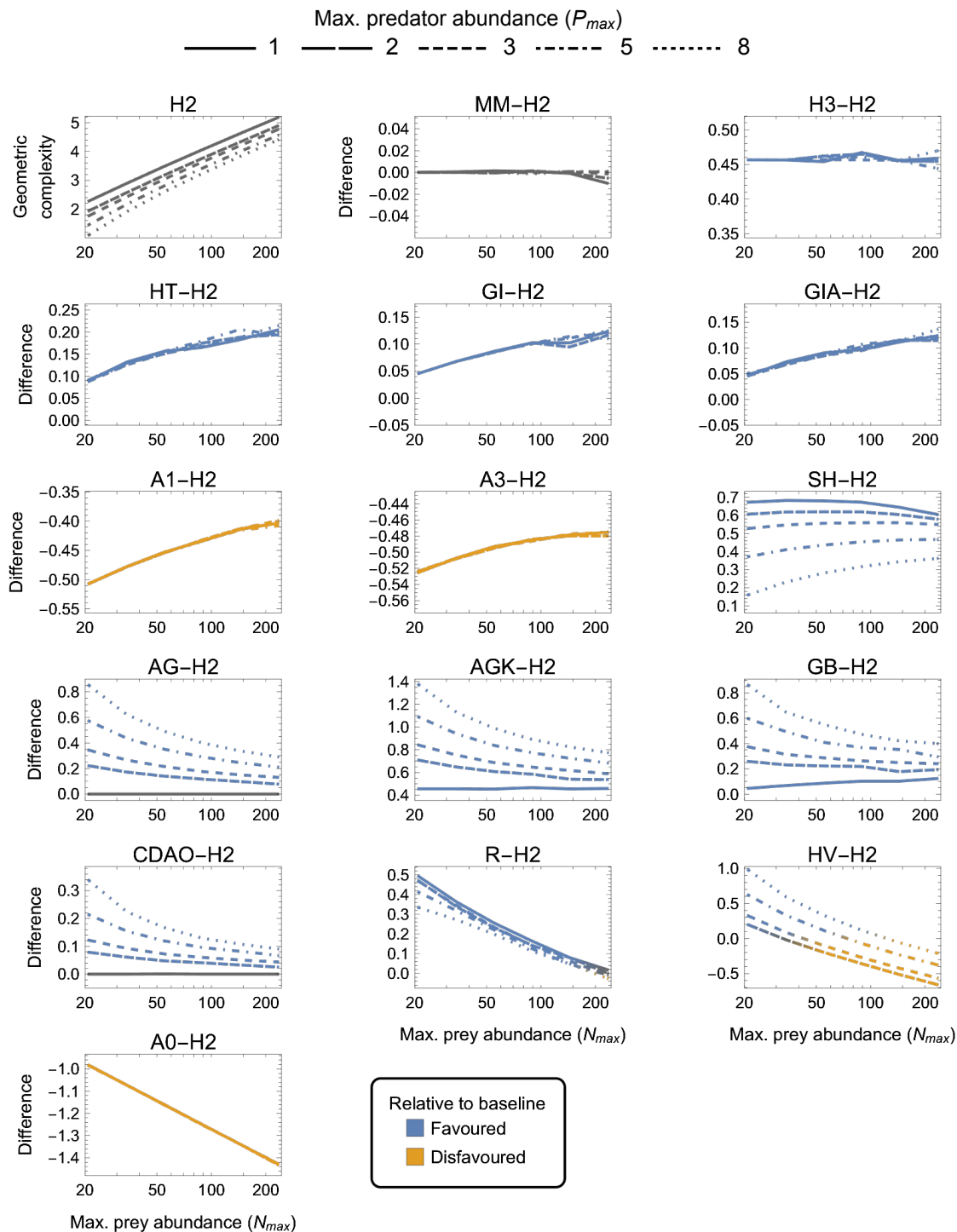
28

Figure 4: As in Fig. 3 but for two-parameter ($k = 2$) functional-response models. *First panel*: The geometric complexity $\mathcal{G}_{H2}$ of the baseline Holling Type II model (H2) as a function of an experiment's maximum prey and predator abundances (N_{max} and P_{max}). *Other panels*: The difference in $\mathcal{G}_M$ of all other two-parameter models relative to the H2 model. As a visual aid, we colour models favoured relative to the baseline in blue and models disfavoured relative to the baseline in yellow.

439 (Fig. 5):

440 (i) FHM and BD were favoured over H3R, and CM, W, SBB and AA were favoured over BD,

441 for all designs (excepting for $P_{max} = 2$ where $\mathcal{G}_{AA} \approx \mathcal{G}_{BD}$);

442 (ii) A2, HLB, MH, AS, SSS and T were disfavoured relative to H3R, and TTA and RGD were

443 disfavoured relative to BD, for all designs; and

444 (iii) BWL2 was favoured over BD by small N_{max}, large P_{max} designs and was disfavoured by

445 large N_{max}, small P_{max} designs.

446 For the ratio- and predator-dependent models, differences to BD were more sensitive to variation

447 in P_{max} than to variation in N_{max}. The degree to which CM, W, SBB and AA where favoured

448 over BD increased with increasing P_{max}, reaching as much as 0.8 information units at $P_{max} = 8$.

449 For these models, the most equitable design therefore entailed small P_{max} regardless of N_{max},

450 but for TTA and RGD, for which the difference to BD decreased with increasing P_{max}, it was

451 designs entailing large P_{max} which disfavoured them the least (by no less than 1.4 and up to

452 2.9 information units). The degree to which the prey-dependent AS, SSS and T models were

453 disfavoured relative to H3R was also more sensitive to variation in P_{max} than in N_{max}, but

454 the degree to which A2, HLB and MH were disfavoured and the degree to which FHM was

455 favoured was relatively insensitive to variation in P_{max}. As N_{max} increased, SSS and T became

456 more disfavoured, A2, HLB, MH and AS became less disfavoured, and FHM became more

457 favoured. For BWL2, which could either be favoured or disfavoured depending on design, the

30

458 most equitable designs spanned those that had the largest considered N_{max} when P_{max} was large

459 to those that had the smallest considered N_{max} when P_{max} was small. Overall, SSS and RGD

460 exhibited the greatest potential disparity relative to their H3R and BD baselines, respectively

461 being disfavoured by about 13 and almost 2.9 information units by the least equitable designs.

462 The greatest potential disparity among all other considered three-parameter models was about

463 11 information units and occurred between SSS and CM for large N_{max}, large P_{max} designs in

464 favour of CM.

465 ## 3.5   Four-parameter models

466 Finally, among the four-parameter models, which exhibited the greatest amounts of numerical

467 estimation noise (Fig. 6):

468   (i) AAOR was favoured over BDRO for all designs (had higher $\mathcal{G}_M$);

469   (ii) SN1 and SN2 were disfavoured relative to BDRO for all designs (had lower $\mathcal{G}_M$); and

470 (iii) CMOR tended to be favoured by large N_{max}, large P_{max} designs and disfavoured by small

471   N_{max}, small P_{max} designs.

472 For CMOR, AAOR and SN1, the difference to BDOR was less sensitive to variation in N_{max}

473 than to variation in P_{max}, but the opposite was true for SN2. Further, while the degree to which

474 AAOR was favoured over BDRO was minimized by $P_{max} = 2$ designs (to about 0.2 information

475 units), the degree to which SN1 was disfavoured relative to BDRO was minimized by $P_{max} = 8$

476 designs (to about 0.5 information units). SN2 was non-identifiable for designs having $P_{max} \leq 3$
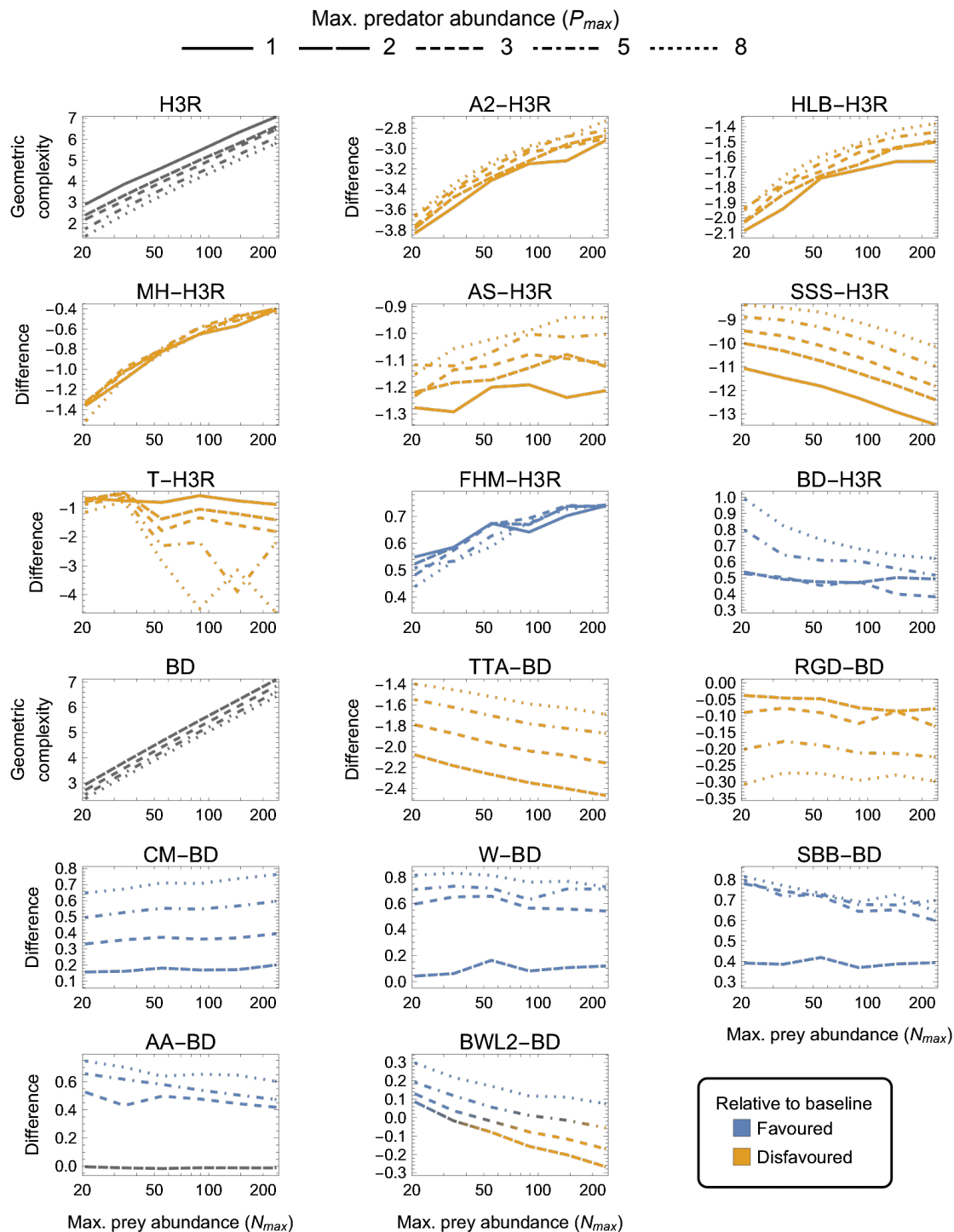
Figure 5: As in Fig. 3 but for three-parameter ($k = 3$) functional-response models. *First and tenth panels*: The geometric complexity $\mathcal{G}_M$ of the baseline Holling–Real Type III (H3R) and Beddington–DeAngelis (BD) models as a function of the experiment's maximum prey and predator abundances (N_{max} and P_{max}). *Other panels*: The difference in $\mathcal{G}_M$ of the other three-parameter prey-dependent (top two rows) and ratio- and predator-dependent (bottom two rows) models relative to the baseline models. As a visual aid, we colour models favoured relative to the baseline in blue and models disfavoured relative to the baseline in yellow.
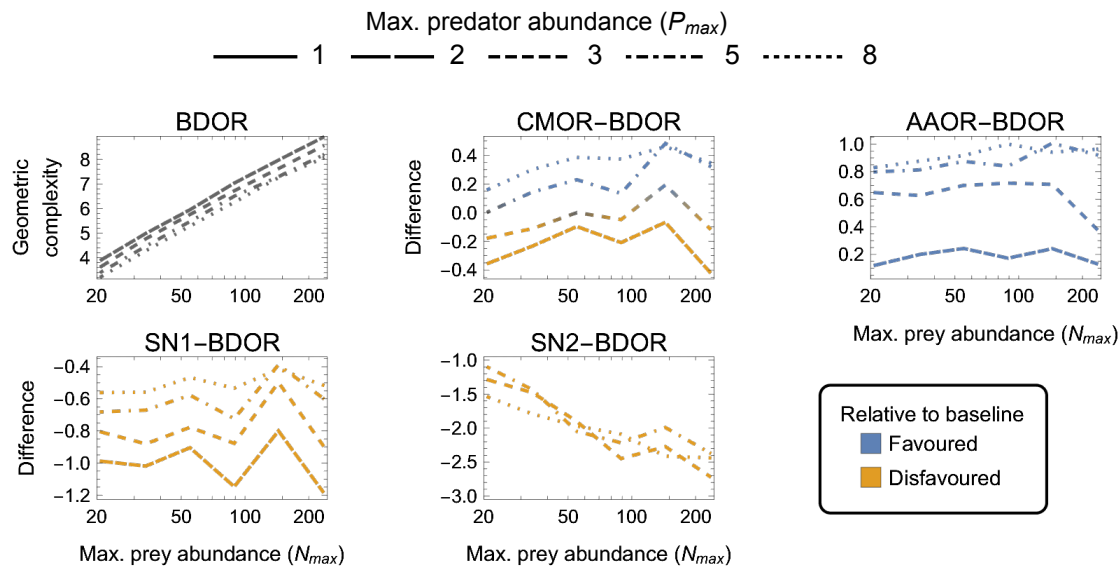
Figure 6: As in Fig. 3 but for four-parameter ($k = 4$) functional-response models. *First panel*: The geometric complexity $\mathcal{G}_{BDOR}$ of the Beddington–DeAngelis–Okuyama–Ruyle (BDOR) model as a function of the experiment's maximum prey and predator abundances (N_{max} and P_{max}). *Other panels*: The difference in $\mathcal{G}_M$ of all other four-parameter models relative to the BDOR model. As a visual aid, we colour models favoured relative to the baseline in blue and models disfavoured relative to the baseline in yellow.

477 (Fig. S1), but for $P_{max} > 3$ designs it was disfavoured by at least 1 information unit. The most

478 equitable designs for CMOR and BDOR entailed intermediate predator abundances ($P_{max} =$

479 3–5). Overall, the greatest potential disparity to the BDOR baseline model occurred for the

480 SN2 model (about 2.5 information units) at the largest N_{max}. The greatest potential disparity

481 among all considered four-parameter models occurred for the SN2 and AAOR models (about

482 3.5 information units) for the largest N_{max}, largest P_{max} design in favour of AAOR.

33

3.6 Sensitivity analyses

Fixing $N_{max} = 233$ and $P_{max} = 5$ and varying the number of prey and predator treatment levels (L_N and L_P) to below the numbers used in our primary analysis showed that $\mathcal{G}_M$ was relatively insensitive to variation in L_N for most models (Figs. S2–S5). In contrast, the degree to which models were favoured or disfavoured relative to their baseline was far more sensitive to variation in L_P. For most of the L_P-sensitive models, decreasing L_P increased their difference to the baseline model, but for an almost equal number the difference decreased. The largest effects of L_P most often occurred when reducing from two predator levels ($P \in \{1, 2\}$) to only a single-predator level (or the corresponding reduction of three to two levels for the four-parameter models). Setting aside these last-mentioned and in some ways trivial changes to L_P, the greatest effects of changing L_N and L_P were to change the degree to which models were favoured or disfavoured relative to their baseline models by up to about 0.6 information units (excepting T and SSS for which they affected a change of up to 2.5 units).

The use of designs with arithmetic rather than logarithmic spacings of prey and predator abundances also had little to no effect on the degree to which models were favoured or disfavoured relative to their baselines (Figs. S6–S9). The notable exceptions included the manner in which (i) HV was favoured relative to H2 (arithmetic spacings making HV invariably favoured rather than favoured or disfavoured depending on N_{max} and P_{max}), (ii) BD was favoured relative to H3R (arithmetic spacings making it more favoured by large rather than small N_{max}), and (iii) CMOR was favoured relative to BDOR (arithmetic spacings making CMOR invariably

34

503 disfavoured rather than favoured or disfavoured depending on N_{max} and P_{max}).

504 Finally, relaxing the indirect constraints we imposed on parameters by changing the minimum

505 or the maximum expected number of eaten prey by an order of magnitude had similarly little

506 effect (Figs. S10–S17). The notable consequences were that increasing the maximum expected

507 number of eaten prey across all treatments from N_{max} to 10 N_{max} caused (i) CDAO to be

508 disfavoured rather than favoured relative to H2, (ii) T and W to be favoured or disfavoured

509 relative to H3R and BD in a design-dependent rather than design-independent manner; (iii)

510 CMOR to become favoured over BDOR for a greater range of designs, and (iv) $\mathcal{G}_{SN1}$ and $\mathcal{G}_{SN2}$

511 to no longer be estimable, even after a month of computation on a high-performance computing

512 cluster (all other models each requiring only seconds to minutes on a 4-core computer).

4 Discussion

514 The functional-response literature is replete with models, even among those that only consider

515 variation in the abundances of a single predator-prey pair (Table 1, Jeschke *et al.*, 2002). Each

516 of these many models was proposed to encapsulate a different aspect of predator-prey biology,

517 though frequently even very different biological processes lead to very similar or even the same

518 model form (Table 1). Information-theoretic criteria, which balance model fit and complexity,

519 represent the principal, most general, and most accessible means for comparing the statistical

520 performance of these models when they are confronted with data (Okuyama, 2013). The primary

521 contribution of our analyses is to show that existing models, independent of the biology they

522 are meant to reflect, frequently also differ in their flexibility to fit data, even among models

35

having the same parametric complexity. Differences in model flexibility as assessed by the

geometric complexity term $\mathcal{G}$ of the FIA criterion were frequently greater than 0.5 information

units, spanned values up to 13 information units, and for several models were never below 1

information unit even for the most equitable of considered experimental designs. Secondarily,

our analyses demonstrate just how dependent a model's flexibility can be on the experimental

design of the data (i.e. what the range and combinations of prey and predator abundances are).

In some instances this design dependency was great enough to cause models that were less

flexible than other models for some experimental designs to become more flexible than the same

models for different designs.

Our use of the FIA criterion allows us to contextualize the importance of this variation in

flexibility in two rigorous and quantitative ways: First, we can compare $\mathcal{G}$ among models of

the same parametric complexity for a given experimental design assuming their goodness-of-fit

to a hypothetical dataset to be the same. In this scenario, the potential significance of model

flexibility to the information-theoretic comparison of functional-response models is evidenced in

a general manner by the fact that a 2-unit difference in AIC or BIC among competing models

— equivalent to a 1-unit difference in FIA — represents "substantial" support (a weight-of-

evidence of 2.7 to 1) for one model over another (Burnham & Anderson, 2002). (Such a difference

reflects a probability of 0.73 that the first of only two competing models is "better" than the

other.) Second, we can compare $\mathcal{G}$ to a model's parametric complexity for hypothetical datasets

of differing sample size assuming its goodness-of-fit to these data remains the same. In this

543 scenario, the potential significance of model flexibility to the inferences of functional-response

544 studies performed in the past is evidenced by the fact that our estimated differences in $\mathcal{G}$ are

545 comparable to the values of parametric complexity that are associated with the median and

546 even maximum sample sizes seen in the large collection of datasets recently compiled by Novak

547 & Stouffer (2021) (Table 2). That is, as feared by Novak & Stouffer (2021), sample sizes

548 among existing empirical datasets are often sufficiently small that the likelihood and parametric

549 complexity differences of many models is unlikely to have sufficiently out-weighed the influence

550 of their functional flexibility in determining their information-theoretic rankings.

Table 2: The value of FIA's parametric complexity term (the second term of Eq. (3) depicted in Fig. 1) for models of $k = 1$, 2, 3 and 4 parameters evaluated at the sample sizes of the smallest ($n = 10$), median ($n = 80$), and largest ($n = 528$) sized datasets in the set of 77 functional-response datasets having variation in both prey and predator abundances compiled by Novak & Stouffer (2021). These values serve as reference for gauging the magnitude differences in geometric complexity between models reported here and, thereby, for judging the likely influence of model flexibility on prior inferences of relative model performance using AIC and BIC.

	Sample size (n)		
k	10	80	528
1	0.2	1.3	2.2
2	0.5	2.5	4.4
3	0.7	3.8	6.6
4	0.9	5.1	8.9

4.1 What makes models (in)flexible?

552 Given that the influence of model flexibility on information-theoretic model comparisons of the

553 past is likely substantial, that its influence will likely not change dramatically in the future

554 given the logistical challenges of standard experimental approaches, and because there is no

555 experimental design that can make the comparison of functional-response models universally

37

556   equitable with respect to their flexibility, an important question is: What aspects of their

557   mathematical formulation make models more or less flexible for certain experimental designs?

558   For the one-parameter models the answer is relatively accessible given the specifics of our

559   analyses. The linear ratio-dependent (LR) model is more flexible than the Holling Type I (H1)

560   model because the division of prey abundances by a range of predator abundances allows a

561   greater range of parameter a ("attack rate") values to satisfy the condition that the resulting

562   expected numbers of eaten prey will lie within our specified minimum and maximum bounds

563   (i.e. satisfying both $\mathbb{E}[F(N_{max}, P, \theta)PT] \geq 1/10$ and $\mathbb{E}[F(N, P, \theta)PT] \leq 10N_{max}$). Relative to

564   H1 for which high N_{max} and low P_{max} maximize the potential range of attack rates that an

565   individual predator could express in an experiment, having many predators "interfering" in a

566   ratio-dependent manner enables each individual predator to express an even greater attack rate

567   without all predators in total consuming too many prey. The effects on the maximum versus

568   the minimum prey eaten are asymmetric in magnitude (i.e. the maximum potential value of a

569   increases more than the minimum potential value of a) because division by P in LR has an

570   asymmetric effect on the per predator number of prey eaten (relative to the multiplication by

571   P that is common to all models); it is symmetric only on a logarithmic scale. The magnitude of

572   this effect is dampened in the BWL1 model of Barbier *et al.* (2021) because it entails a ratio of

573   the square roots of (is sublinear with respect to) prey and predator abundances, making BWL1

574   more flexible than H1 but less flexible than LR.

575   The same rationale applies to all other models and explains the varied (in)sensitivities that

38

576 their model flexibility has with respect to experimental design. That said, the situation is

577 often more complicated for models with multiple parameters because of (i) the interdependent

578 influences that parameters have on the number of prey that are eaten, and (ii) the fact that,

579 for some models, the minimum and the maximum boundaries on the expected number of eaten

580 prey come into play at different points in parameter- and species-abundance space.

581 For example, for the Holling Type II (H2) model, requiring that at least one prey on average

582 be eaten in the highest prey abundance treatments causes high handling times to impose a lower

583 limit on each individual's attack rates only if and when prey abundances are sufficiently high to

584 affect saturation. The Holling Type III (H3) model experiences this same effect as well, hence

585 its *relative* flexibility is insensitive to variation in maximum prey abundances. H3 is nonetheless

586 more flexible than H2 because it is superlinear with respect to prey abundance (when handling

587 times or prey abundances are low) and can therefore satisfy the minimum of one-prey-eaten-per-

588 predator constraint for smaller attack rate values than can H2. Similarly, the exponential form

589 of the Gause–Ivlev models (GI and GIA) makes them more flexible than H2 because they are

590 superlinear with respect to prey abundance, while the A1 and A3 models of Abrams (1990) are

591 less flexible than H2 because they are sublinear with respect to prey abundance. The insensitivity

592 of the relative flexiblity of all these models to variation in predator abundances occurs because

593 the total prey eaten they effect is determined by predator abundance in the same proportional

594 manner as for H2. That is, just like most other two-parameter prey-dependent models, the

595 relative flexibility of H2 and these models is similarly uninfluenced by the ratio of prey and

39

596 predator abundances, in contrast to the way that all ratio- and predator-dependent models are

597 affected (as per the contrast of H1, LR and BWL1 discussed above).

598   The prey-dependent Type IV model of Sokol & Howell (1981) (SH) represents an informa-

599 tive exception to all other two-parameter prey-dependent models in that its relative flexibility is

600 sensitive to predator abundance. Whereas all monotonically increasing prey-dependent models

601 only ever come up against the maximum prey abundance constraint as predator abundances in-

602 crease, increasing predator abundances additionally alleviate the constraint that SH experiences

603 uniquely due to the eventual decline of its feeding rate at high prey abundance; high preda-

604 tor abundances permit the total number of prey eaten to stay above the minimum-of-one-prey

605 constraint for greater maximum prey abundances than is possible for low predator abundances

606 given the parameter values.

607   The dependence of model flexibility on predator abundance emerges among the prey-dependent

608 three-parameter models for similar reasons. For example, although the feeding rates of neither

609 the HLB model of Hassell *et al.* (1977) nor the A2 model of Abrams (1990) decline with respect

610 to prey abundance, increasing their c parameter does make their denominators more sensitive to

611 maximum prey abundances where the minimum of one-prey-eaten-per-predator constraint comes

612 in to play. Therefore, just as for SH, increasing predator abundances increase the number of

613 prey eaten to allow for larger values of c to satisfy the minimum-of-one-prey constraint. That is,

614 although increasing predator abundance would limit the range of c due to the minimum-of-one-

615 prey constraint if all else were to be held constant, all else is not constant. Rather, high predator

40

616 abundance enables a greater range of a values for a given value of c before the maximum-prey-

617 eaten constraint is violated. This is also the reason why all predator-dependent models exhibit

618 increasing relative flexibility as predator abundance increases even as the absolute flexibility of

619 their respective baseline models decreases.

4.2 Additional aspects of experimental design

621 Our sensitivity analyses on the role of experimental design reinforce the inferences of our main

622 analysis. They also speak to the likely generality of our results to addtional aspects of ex-

623 perimental design which we did not specifically address. For the two-parameter models whose

624 relative flexibility was insensitive to the ratio of prey and predator abundances, using arith-

625 metic rather than logarithmic designs had little or no qualitative influence because arithmetic

626 spacings did not alter maximum prey abundances where the constraints on the number of prey

627 eaten are incurred. By contrast, models for which changes to spacings or the prey-eaten con-

628 straints did alter their relative flexibility were either ratio- or predator-dependent models, or

629 were prey-dependent models whose additional (third) parameter made their flexibility sensitive

630 to predator abundance. We conclude from this that the precise spacings of prey and predator

631 abundances are less important from a model flexibility perspective than are their maxima and

632 combinatorial range, but that these aspects of design become more important as the parametric

633 complexity of the considered models increases.

634 Nonetheless, searching for equitable experimental designs as we did is different from searching

635 for optimal designs for model-specific parameter uncertainty, bias, or identifiability (e.g., Moffat

41

636 *et al.*, 2020; Sarnelle & Wilson, 2008; Uszko *et al.*, 2020; Zhang *et al.*, 2018). A precedence of

637 other motivations for an experiment, such as maximizing the precision of parameter estimates,

638 may therefore lead to different and likely model-specific conclusions about which design aspects

639 are important. Fortunately, given our results, some aspects of experimental design may be of

640 little consequence. For example, independent of the maximum prey abundance used, the general

641 utility of a logarithmic spacing of prey makes intuitive sense given that, for many models, most

642 of the action that differentiates model form occurs at low prey abundances (i.e. their derivatives

643 with respect to N are greatest at low values of N). Intuition likewise suggests that designs

644 should preclude total prey consumption being overwhelmed by the overall effect of interference

645 among predators and hence that predator abundances shouldn't be high. In this regard our

646 results indicate that just a little variation across a range of low predator abundances is often —

647 though far from universally — best from a relative model flexibility standpoint, just as it would

648 be expected to be best for parameter estimation.

649   Our analyses did not consider questions regarding the treatment-specific distribution of ex-

650 perimental replicates, important though these often are given logistical constraints. All of our

651 analyses assumed uniformly-balanced designs, the effect of which future analyses could easily

652 assess by changing the assumed probability of each experimental treatment when computing

653 the Expected unit Fisher Information matrix underlying $\mathcal{G}$ (see Box 1). We anticipate, however,

654 that shifting replicates from lower prey and predator abundances to higher abundances will have

655 a similar effect to that seen in the comparison of logarithmic to arithmetic spacings. Therefore,

42

656   from a model flexibility standpoint alone, we expect such a shift to have a greater effect for

657   models of high parametric complexity.

658   A final important aspect of experimental design that our analyses did not address was the

659   assumed likelihood function connecting each deterministic functional-response model to an ex-

660   periment's design (i.e. the structure of the data). We assumed a Poisson likelihood and therefore

661   that eaten prey are continually replaced, that the mean and variance of prey eaten are equal for

662   a given combination of predator and prey abundances, and that all feeding events are indepen-

663   dent. Model flexibility as assessed by geometric complexity may be different under alternative

664   likelihoods such as the binomial likelihood (which would be appropriate for depletion designs)

665   or the negative binomial likelihood (which allows for under- or over-dispersion). Indeed, for the

666   binomial likelihood even the linear Holling Type I deterministic function response results in a

667   non-linear statistical model (Novak & Stouffer, 2021), hence its geometric complexity may be

668   quite different. That said, the maximum likelihood parameter estimators under Gaussian and

669   log-Normal likelihoods are the same as under a Poisson likelihood for many — and possibly all

670   — of the models we considered (Novak & Stouffer, 2021), so it is likely that our inferences would

671   be little changed under these commonly assumed alternatives.

4.3   Model flexibility as problem and desirable property

673   There are many perspectives on the purpose of models and why we fit models to data. Shmueli

674   (2010) articulates two primary axes of motivation that align well to the functional-response

675   literature: *explanation* (where the primary motivation is to infer biologically- and statistically-

43

significant causal associations the nature of which models are meant to characterize) and *predic-*

tion (where the primary motivation is to best describe as yet unseen out-of-sample data).[1] The

ability to satisfy both motivations converges as the amount of data and the range of conditions

the data reflect increase, thereby mirroring the inferential convergence of information criteria

as sample sizes increase and cause differences in goodness-of-fit to dominate measures of model

complexity. Model flexibility, and with it our analyses, would thus be irrelevant if the sample

sizes of functional-response experiments were sufficiently large. Instead, sample sizes for many

studies are such that model flexibility — as well as other forms of statistical and non-statistical

bias (Novak & Stouffer, 2021) — preclude the conclusion that models deemed to perform best

on the basis of their information-theoretic ranking are also closest to biological truth.

Empiricists fitting functional-response models to data must therefore make the explicit choice

between explanation, for which criteria such as BIC and FIA are intended, and prediction,

for which AIC(c), cross-validation, model-averaging, and most forms of machine learning are

intended (Aho *et al.*, 2014; Höge *et al.*, 2018; Shmueli, 2010). If data is limited and explanation

is the goal, then design-dependent differences in model flexibility represent a critical problem

for commonly-used criteria like BIC because more flexible models will be conflated for the

truth. In such contexts, it would be wise to identify the most equitable design for a specifically

chosen subset of hypothesis-driven models (see also Burnham & Anderson, 2002), or, in lieu

of a better reasoned solution, to use a design or multiple designs that stack the deck against

---

[1]A third axis, *description*, remains common in the functional-response literature and typically takes the form of fitting "non-mechanistic" polynomial models to evaluate the statistical significance of the individual parameters.

leading hypotheses associated with the most flexible models. On the other hand, if data is limited and out-of-sample prediction is the goal, then model flexibility could be considered an advantage if it causes more-complex-than-true models to be selected because they are deemed to perform better, especially when the true model may not even be among those being compared (Höge *et al.*, 2018). More generally, there are clearly contexts in which ecologists wish to have generic, flexible functional-response models that merely approximate aspects of the truth in a coarse manner, be it in more descriptive statistical contexts or in theoretical contexts where the potential role of these aspects in determining qualitatively different regimes of population dynamics is of interest (e.g., AlAdwani & Saavedra, 2020; Arditi & Ginzburg, 2012; Barbier *et al.*, 2021). In these contexts, and since all models are phenomenological and hence agnostic with respect to precise mechanistic detail (as Table 1 underscores; see also Connolly *et al.*, 2017; Hart *et al.*, 2018), we consider the results of our analyses to be useful for making *a priori* choices among models given that more flexible models likely capture and exhibit a greater amount of biologically insightful variation in a more analytically tractable manner.

4.4 Conclusions

Several syntheses evidence that there is no single model for characterizing predator functional responses in general (Novak & Stouffer, 2021; Skalski & Gilliam, 2001; Stouffer & Novak, 2021). This is consistent with the fact that, to a large degree, the statistical models of the functional-response literature characterize aspects of predator-prey biology for which there is evidence in data, not whether specific mechanisms do or do not occur in nature (see also Connolly

45

715 *et al.*, 2017). In light of the fact that functional-response data are hard to come by, our study

716 demonstrates that a model's functional flexibility should be considered when interpreting its

717 performance. That said, we are not advocating for FIA as an alternative to more commonly-used

718 information criteria; more fundamental issues exist that pertain to the explicit consideration of

719 study motivation. Indeed, we submit that questions of motivation are ones that the functional-

720 response literature as a whole needs to grapple with more directly. Even in the specific context

721 of prediction, for example, functional-response studies rarely address explicitly what their study

722 and their data are intending to help better predict (e.g., feeding rates or population dynamics).

723 Valuable effort would therefore be expended in future work to consider the relationship of model

724 flexibility to the parametric- and structural sensitivities of models when it comes to drawing

725 inferences for population dynamics (e.g., Adamson & Morozov, 2020; Aldebert & Stouffer, 2018).

726 Likewise, it would also be useful to clarify the relevance of model flexibility to the rapidly

727 developing methods of scientific machine learning, including the use of symbolic regression,

728 neural ordinary differential equations, and universal differential equations for model discovery

729 (e.g., Bonnaffé *et al.*, 2021; Guimerà *et al.*, 2020; Martin *et al.*, 2018; Rackauckas *et al.*, 2020).

Acknowledgments

735 and Lucile Packard Foundation grant to the Partnership for Interdisciplinary Studies of Coastal

736 Oceans.

Conflict of Interest Statement

738 The authors declare no conflicts of interest.

Author contributions

740 MN led the study and wrote the first draft. Both authors contributed to the analyses and

741 revisions.

Code and data availability

743 All code has been archived at FigShare:XXXXX◇and is available at `https://github.com/`   MN: To be uploaded upon acceptance

744 `marknovak/GeometricComplexity`.

References

746 Abrams, P. A. (1982). Functional responses of optimal foragers. *The American Naturalist*, 120,

747   382–390.

748 Abrams, P. A. (1990). The effects of adaptive-behavior on the type-2 functional-response. *Ecol-*

749   *ogy*, 71, 877–885.

750 Adamson, M. W. & Morozov, A. Y. (2020). Identifying the sources of structural sensitivity in

751   partially specified biological models. *Scientific Reports*, 10, 16926.

47

752 Aho, K., Derryberry, D. & Peterson, T. (2014). Model selection for ecologists: the worldviews

753   of AIC and BIC. *Ecology*, 95, 631–636.

754 AlAdwani, M. & Saavedra, S. (2020). Ecological models: higher complexity in, higher feasibility

755   out. *Journal of The Royal Society Interface*, 17, 20200607.

756 Aldebert, C., Nerini, D., Gauduchon, M. & Poggiale, J. C. (2016a). Does structural sensitivity

757   alter complexity–stability relationships? *Ecological Complexity*, 28, 104–112.

758 Aldebert, C., Nerini, D., Gauduchon, M. & Poggiale, J. C. (2016b). Structural sensitivity and

759   resilience in a predator–prey model with density-dependent mortality. *Ecological Complexity*,

760   28, 163–173.

761 Aldebert, C. & Stouffer, D. B. (2018). Community dynamics and sensitivity to model structure:

762   towards a probabilistic view of process-based model predictions. *Journal of The Royal Society

763   Interface*, 15, 20180741.

764 Andrews, J. F. (1968). A mathematical model for the continuous culture of microorganisms

765   utilizing inhibitory substrates. *Biotechnology and Bioengineering*, 10, 707–723.

766 Arditi, R. & Akçakaya, H. R. (1990). Underestimation of mutual interference of predators.

767   *Oecologia*, 83, 358–361.

768 Arditi, R. & Ginzburg, L. R. (1989). Coupling in predator-prey dynamics: Ratio-dependence.

769   *Journal of Theoretical Biology*, 139, 311–326.

770 Arditi, R. & Ginzburg, L. R. (2012). *How species interact: altering the standard view on trophic*

771   *ecology*. Oxford University Press.

772 Barbier, M., Wojcik, L. & Loreau, M. (2021). A macro-ecological approach to predation density-

773   dependence. *Oikos*, 130, 553–570.

774 Beddington, J. R. (1975). Mutual interference between parasites or predators and its effect on

775   searching efficiency. *The Journal of Animal Ecology*, 44, 331–340.

776 Bonnaffé, W., Sheldon, B. C. & Coulson, T. (2021). Neural ordinary differential equations

777   for ecological and evolutionary time-series analysis. *Methods in Ecology and Evolution*, 12,

778   1301–1315.

779 Burnham, K. P. & Anderson, D. R. (2002). *Model selection and multimodel inference: A practical*

780   *information-theoretic approach*. 2nd edn. Springer, New York.

781 Coelho, M. T. P., Diniz-Filho, J. & Rangel, T. F. (2019). A parsimonious view of the parsimony

782   principle in ecology and evolution. *Ecography*, 42, 968–976.

783 Connolly, S. R., Keith, S. A., Colwell, R. K. & Rahbek, C. (2017). Process, mechanism, and

784   modeling in macroecology. *Trends in Ecology & Evolution*, 32, 835–844.

785 Cosner, C., DeAngelis, D. L., Ault, J. S. & Olson, D. B. (1999). Effects of spatial grouping on

786   the functional response of predators. *Theoretical Population Biology*, 56, 65–75.

787 Crowley, P. H. & Martin, E. K. (1989). Functional responses and interference within and between

788 year classes of a dragonfly population. *Journal of the North American Benthological Society*,

789 8, 211–221

790 DeAngelis, D. L., Goldstein, R. A. & O'Neill, R. V. (1975). A model for trophic interaction.

791 *Ecology*, 56, 881–892.

792 DeLong, J. P. & Uiterwaal, S. F. (2018). The FoRAGE (Functional Responses from Around

793 the Globe in all Ecosystems) database: a compilation of functional responses for consumers

794 and parasitoids. *Knowledge Network for Biocomplexity*, doi:10.5063/F17H1GTQ.

795 Ellison, A. M. (2004). Bayesian inference in ecology. *Ecology Letters*, 7, 509–520.

796 Evans, M. R., Grimm, V., Johst, K., Knuuttila, T., de Langhe, R., Lessells, C. M., Merz, M.,

797 O'Malley, M. A., Orzack, S. H., Weisberg, M., Wilkinson, D. J., Wolkenhauer, O. & Benton,

798 T. G. (2013). Do simple models lead to generality in ecology? *Trends in Ecology & Evolution*,

799 28, 578–583.

800 Fujii, K., Holling, C. & Mace, P. (1986). A simple generalized model of attack by predators and

801 parasites. *Ecological Research*, 1, 141–156.

802 Fussmann, G. F. & Blasius, B. (2005). Community response to enrichment is highly sensitive

803 to model structure. *Biology Letters*, 1, 9–12.

804 Gause, G. F. (1934). *The Struggle for Existence*. Williams and Wilkins, Baltimore.

805 Grünwald, P. (2000). Model selection based on minimum description length. *Journal of Math-*

806   *ematical Psychology*, 44, 133–152.

807 Guimerà, R., Reichardt, I., Aguilar-Mogas, A., Massucci, F. A., Miranda, M., Pallarès, J. &

808   Sales-Pardo, M. (2020). A bayesian machine scientist to aid in the solution of challenging

809   scientific problems. *Science Advances*, 6, eaav6971.

810 Gutierrez, A. & Baumgärtner, J. (1984). Multitrophic models of predator-prey energetics I:

811   Age-specific energetics models – Pea aphid *Acrythosiphon pisum* (Homoptera: Aphidae) as

812   an example. *The Canadian Entomologist*, 116, 923–932.

813 Hart, S. P., Freckleton, R. P. & Levine, J. M. (2018). How to quantify competitive ability.

814   *Journal of Ecology*, 106, 1902–1909.

815 Hassell, M. P., Lawton, J. H. & Beddington, J. R. (1977). Sigmoid functional responses by

816   invertebrate predators and parasitoids. *The Journal of Animal Ecology*, 46, 249–262.

817 Hassell, M. P. & Varley, G. C. (1969). New inductive population model for insect parasites and

818   its bearing on biological control. *Nature*, 223, 1133–1137.

819 Höge, M., Wöhling, T. & Nowak, W. (2018). A primer for model selection: The decisive role of

820   model complexity. *Water Resources Research*, 54, 1688–1715.

821 Holling, C. S. (1959). Some characteristics of simple types of predation and parasitism. *The*

822   *Canadian Entomologist*, 91, 385–398.

51

823 Holling, C. S. (1965). The functional response of predators to prey density and its role in mimicry

824 and population regulation. *Memoirs of the Entomological Society of Canada*, 45, 3–60.

825 Ivlev, V. S. (1955). *Experimental ecology of the feeding of fishes*. Yale University Press.

826 Jassby, A. D. & Platt, T. (1976). Mathematical formulation of the relationship between photo-

827 synthesis and light for phytoplankton. *Limnology and Oceanography*, 21, 540–547.

828 Jeschke, J. M., Kopp, M. & Tollrian, R. (2002). Predator functional responses: Discriminating

829 between handling and digesting prey. *Ecological Monographs*, 72, 95–112.

830 Johnson, J. B. & Omland, K. S. (2004). Model selection in ecology and evolution. *Trends in*

831 *Ecology & Evolution*, 19, 101–108.

832 Kass, R. E. & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Associ-*

833 *ation*, 90, 773–795.

834 Kratina, P., Vos, M., Bateman, A. & Anholt, B. R. (2009). Functional responses modified by

835 predator density. *Oecologia*, 159, 425–433.

836 Lotka, A. J. (1925). *Elements of physical biology*. Williams & Wilkins.

837 Ly, A., Marsman, M., Verhagen, J., Grasman, R. P. & Wagenmakers, E.-J. (2017). A tutorial

838 on Fisher information. *Journal of Mathematical Psychology*, 80, 40 – 55.

839 Martin, B., Munch, S. & Hein, A. (2018). Reverse-engineering ecological theory from data.

840 *Proceedings of the Royal Society B: Biological Sciences*, 285, 20180422.

Michaelis, L. & Menten, M. L. (1913). Die Kinetik der Invertinwirkung. *Biochem. Z.*, 49, 333–369.

Moffat, H., Hainy, M., Papanikolaou, N. E. & Drovandi, C. (2020). Sequential experimental design for predator–prey functional response experiments. *Journal of The Royal Society Interface*, 17, 20200156.

Myung, J. I., Navarro, D. J. & Pitt, M. A. (2006). Model selection by normalized maximum likelihood. *Journal of Mathematical Psychology*, 50, 167–179.

Novak, M. & Stouffer, D. B. (2021). Systematic bias in studies of consumer functional responses. *Ecology Letters*, 24, 580–593.

Okuyama, T. (2013). On selection of functional response models: Holling's models and more. *BioControl*, 58, 293–298.

Okuyama, T. & Ruyle, R. L. (2011). Solutions for functional response experiments. *Acta Oecologica*, 37, 512–516.

Pitt, M. A., Myung, I. J. & Zhang, S. (2002). Toward a method of selecting among computational models of cognition. *Psychological review*, 109, 472

Rackauckas, C., Ma, Y., Martensen, J., Warner, C., Zubov, K., Supekar, R., Skinner, D., Ramadhan, A. & Edelman, A. (2020). Universal differential equations for scientific machine learning. *arXiv preprint arXiv:2001.04385*.

859 Real, L. A. (1977). The kinetics of functional response. *The American Naturalist*, 111, 289–300.

860 Rissanen, J. (1978). Modeling by shortest data description. *Automatica*, 14, 465–471.

861 Rissanen, J. J. (1996). Fisher information and stochastic complexity. *IEEE Transactions on*
862   *Information Theory*, 42, 40–47.

863 Rogers, D. (1972). Random search and insect population models. *Journal of Animal Ecology*,
864   41, 369–383.

865 Rosenzweig, M. L. (1971). Paradox of enrichment: Destabilization of exploitation ecosystems in
866   ecological time. *Science*, 171, 385–387.

867 Ruxton, G., Gurney, W. & De Roos, A. (1992). Interference and generation cycles. *Theoretical*
868   *Population Biology*, 42, 235–253.

869 Sarnelle, O. & Wilson, A. E. (2008). Type III functional response in Daphnia. *Ecology*, 89,
870   1723–1732

871 Schenk, D., Bersier, L.-F. & Bacher, S. (2005). An experimental test of the nature of predation:
872   neither prey- nor ratio-dependent. *Journal of Animal Ecology*, 74, 86–91.

873 Shmueli, G. (2010). To explain or to predict? *Statististical Science*, 25, 289–310.

874 Skalski, G. T. & Gilliam, J. F. (2001). Functional responses with predator interference: Viable
875   alternatives to the Holling Type II model. *Ecology*, 82, 3083–3092.

54

Sokol, W. & Howell, J. A. (1981). Kinetics of phenol oxidation by washed cells. *Biotechnology and Bioengineering*, 23, 2039–2049.

Stouffer, D. B. & Novak, M. (2021). Hidden layers of density dependence in consumer feeding rates. *Ecology Letters*, 24, 520–532.

Sutherland, W. J. (1983). Aggregation and the 'ideal free' distribution. *Journal of Animal Ecology*, 52, 821–828.

Tostowaryk, W. (1972). The effect of prey defense on the functional response of *Podisus modestus* (Hemiptera: Pentatomidae) to densities of the sawflies *Neodiprion swainei* and *N. pratti banksianae* (Hymenoptera: Neodiprionidae). *The Canadian Entomologist*, 104, 61–69.

Tyutyunov, Y., Titova, L. & Arditi, R. (2008). Predator interference emerging from trophotaxis in predator–prey systems: An individual-based approach. *Ecological Complexity*, 5, 48–58.

Uszko, W., Diehl, S. & Wickman, J. (2020). Fitting functional response surfaces to data: a best practice guide. *Ecosphere*, 11, e03051.

Volterra, V. (1926). Fluctuations in the abundance of a species considered mathematically. *Nature*, 118, 558–560.

Watt, K. (1959). A mathematical model for the effect of densities of attacked and attacking species on the number attacked. *The Canadian Entomologist*, 91, 129–144.

Whitehead, A. N. (1919). The concept of nature. *Convergence*, 62, 79.

894 Wolfram Research Inc. (2020). *Mathematica, v.12.1.* Champaign, IL. URL https://www.

895 wolfram.com/mathematica.

896 Zhang, J. F., E., P. N., Theodore, K. & C., D. C. (2018). Optimal experimental design for

897 predator–prey functional response experiments. *Journal of The Royal Society Interface*, 15,

898 20180186.