

Electronic Supplementary Material for:
Bayesian characterization of uncertainty in species
interaction strengths

Christopher Wolf^{1*}, Mark Novak², and Alix I. Gitelman¹

¹Department of Statistics, Oregon State University,
Corvallis, Oregon 97331, United States

²Department of Integrative Biology, Oregon State University,
Corvallis, Oregon 97331, United States

*corresponding author (wolfch@science.oregonstate.edu)

S1 *On Novak and Wootton's 'Species x'*

This paper expands on the observational method for estimating attack rates presented by Novak and Wootton (2008):

$$a_i = \frac{F_i A_x}{(F_x - A_x) h_i N_i}, \quad \text{eqn S1}$$

where a_i is the attack rate, h_i is the handling time, and N_i is the abundance, all for the i^{th} prey species. A_i and F_i are the proportions of all predators and feeding predator respectively feeding on the i^{th} prey species. x refers to an arbitrarily chosen prey species that is the same for all a_i . Here we show that this equation can also be written in a more simplified form, showing that the estimates are not dependent on the choice of species x .

Define A_0 to be the observed proportion of predators that are not feeding, so that $A_0 = 1 - \sum_{i=1}^S A_i$. Then, the F_i 's can be obtained by normalizing A_i 's: $F_i = \frac{A_i}{\sum_{j=1}^S A_j} = \frac{A_i}{1 - A_0}$.

Noting that:

$$F_x - A_x = \frac{A_x}{1 - A_0} - A_x = \frac{A_x - A_x(1 - A_0)}{1 - A_0} = \frac{A_x[1 - (1 - A_0)]}{1 - A_0} = \frac{A_x A_0}{1 - A_0}. \quad \text{eqn S2}$$

It follows that

$$\frac{F_i A_x}{F_x - A_x} = \frac{\frac{A_i}{1 - A_0} \cdot A_x}{\frac{A_x A_0}{1 - A_0}} = \frac{A_i A_x}{A_x A_0} = \frac{A_i}{A_0}. \quad \text{eqn S3}$$

This can be further simplified by noting that the A_i 's have a common denominator (total number surveyed). This means that the original attack rate equation can be written as

$$a_i = \frac{A_i}{A_0} \cdot \frac{1}{h_i N_i} = \frac{\# \text{ feeding on } i}{\# \text{ not feeding}} \cdot \frac{1}{h_i N_i}. \quad \text{eqn S4}$$

This shows that the estimate does not involve species x . Moreover, the total number surveyed need not be known to estimate a subset of the attack rates.

S2 *A Bayesian Attack Rate Estimator*

The Bayesian machinery is built around Bayes theorem:

$$f(\boldsymbol{\theta}|data) \propto f(data|\boldsymbol{\theta}) \cdot f(\boldsymbol{\theta}). \quad \text{eqn S5}$$

Here, $f(data|\boldsymbol{\theta})$ is the likelihood: a function specifying the likelihood of the observed data in terms of unknown parameters $\boldsymbol{\theta}$. $f(\boldsymbol{\theta})$ is the prior: a probability density function reflecting prior beliefs or uncertainty about the parameters. Together, these inform $f(\boldsymbol{\theta}|data)$: the posterior distribution of the parameters given the data. Here, we consider only objective (also called non-informative) priors, assuming an absence of prior beliefs or information about the parameters in question (Berger, 2006). In other situations, informative priors constructed from previously obtained knowledge or data may be useful.

A parametric formulation of the attack rate estimator (eqn 2) is

$$\xi_i = \frac{\alpha_i}{\alpha_0} \frac{1}{\nu_i \eta_i}. \quad \text{eqn S6}$$

Here, for the i^{th} prey species, ξ_i is the unknown attack rate, ν_i is the population prey abundance, η_i is the population handling time, α_i is the population proportion of predators feeding, and α_0 is the population proportion of predators that are not feeding on any prey species. In each case, the parameters refer to the broader (statistical) population, rather than sampled data only. By framing the attack rates this way, we are able to estimate them in the context of the broader population about which inference is desired. Note that

eqn S6 is derived from a snapshot estimator of attack rates (eqn 2), and does not imply, for example, that attack rates vary inversely with abundances. Rather, attack rates are parameters in the multi-species Type II functional response (eqn 1) and assumed fixed throughout.

Frequentist approaches for combining data from multiple sources to estimate functions of parameters (as in eqn S6) generally rely on bootstrap methods or asymptotics like the multivariate delta-method. Both of these approaches exhibit poor small-sample performance (Efron and Tibshirani, 1994; Kilian, 1998). This is relevant when dealing with predator feeding surveys as the A_i in eqn 2 are often very small for the rare prey species that typify predator diets (Rossberg et al., 2006). Small values of A_i (analogous to having a small sample) can be problematic even when the total number of predators surveyed is large (Agresti and Coull, 1998). Ignoring variation in abundance and handling time estimates to focus on the variation within the feeding surveys may avoid this problem, but will lead to underestimation of the uncertainty in the attack rate estimates. The Bayesian framework circumvents this problem.

If data on prey-specific feeding proportions (\mathbf{F}), abundances (\mathbf{A}), and handling times (\mathbf{H}) are collected independently, the joint likelihood of these distributions may be written as:

$$f(\mathbf{F}, \mathbf{A}, \mathbf{H} | \boldsymbol{\alpha}, \boldsymbol{\nu}, \boldsymbol{\eta}) = f(\mathbf{F} | \boldsymbol{\alpha}) f(\mathbf{A} | \boldsymbol{\nu}) f(\mathbf{H} | \boldsymbol{\eta}). \quad \text{eqn S7}$$

Provided that the corresponding priors are also independent, Bayes theorem implies that

$$f(\boldsymbol{\alpha}|\mathbf{F}) \propto f(\mathbf{F}|\boldsymbol{\alpha})f(\boldsymbol{\alpha}), \quad \text{eqn S8a}$$

$$f(\boldsymbol{\nu}|\mathbf{A}) \propto f(\mathbf{A}|\boldsymbol{\nu})f(\boldsymbol{\nu}), \quad \text{eqn S8b}$$

$$f(\boldsymbol{\eta}|\mathbf{H}) \propto f(\mathbf{H}|\boldsymbol{\eta})f(\boldsymbol{\eta}). \quad \text{eqn S8c}$$

These may therefore be fit with independent models for each component. That is, the posterior distributions of the attack rates in eqn S6 may be estimated by using Markov Chain Monte Carlo (MCMC) to obtain samples from each of the three posterior distributions in eqn S8 and combining these using eqn S6. If the three types of data are not gathered independently, then it is necessary to consider likelihood or prior models that account for this dependence (see Appendix S5).

Model formulation (mathematical details)

Here, we present additional model details for our case study with likelihood and prior distributions given in statistical notation. This supplements the model setup description given in the “Bayesian model formulation” section of the main text.

Modeling the feeding surveys – Letting P be the total number of predators surveyed, X_i the number observed feeding on prey i , and X_0 the number not feeding, we model the combined feeding survey data using a multinomial likelihood with Dirichlet prior:

$$(X_0, X_1, \dots, X_S) \sim \text{Mult}_P(\alpha_0, \alpha_1, \dots, \alpha_S) \quad \text{eqn S9a}$$

$$(\alpha_0, \alpha_1, \dots, \alpha_S) \sim \text{Dirich}(c, c, \dots, c). \quad \text{eqn S9b}$$

The resulting posterior distribution is also Dirichlet:

$$(\alpha_0, \alpha_1, \dots, \alpha_S) | \mathbf{x} \sim \text{Dirich}(c + x_0, c + x_1, \dots, c + x_S). \quad \text{eqn S10}$$

Modeling the abundance surveys – Letting Y_1, \dots, Y_n correspond to the n prey abundance measurements, and by conditioning on whether or not a zero occurs, we can write the likelihood density of the zero-inflated gamma (ZIG) distribution as

$$g(y; \alpha, \beta, \rho) = \rho^{I[y=0]} [(1 - \rho) f(y, \alpha, \beta)]^{I[y>0]}, y \geq 0, \quad \text{eqn S11}$$

where ρ is the probability of a zero, $f(y; \alpha, \beta)$ is the usual gamma density with shape α , rate β , and mean $\frac{\alpha}{\beta}$, and $I[\cdot]$ is the indicator function that equals 1 when its argument is true and 0 otherwise (Ospina and Ferrari, 2012). The ZIG density is separable in ρ and (α, β) – that is, it can be expressed as the product of a function of ρ and a function of (α, β) . So, we can model the zero-inflation parameter separately, provided that a separable prior is used. Thus, for each prey species, we model the number of observed zeros using a binomial distribution with a uniform prior on ρ and we take the gamma distribution parameter priors to be $\log(\alpha) \sim \text{Unif}(-100, 100)$ and $\log(\beta) \sim \text{Unif}(-100, 100)$ to approximate the independent scale-invariant non-informative prior

$f(\alpha, \beta) = f(\alpha)f(\beta) \propto \frac{1}{\alpha} \frac{1}{\beta}$, which is equivalent to an (improper) uniform prior on the logarithmic scale (Syversveen, 1998).

Modeling the handling time experiments – We consider the i^{th} handling time observation for a given prey species to be associated with a covariates vector \mathbf{X}_i consisting of 1 followed by temperature, predator size, and prey size (all log transformed). We then model

the likelihood of the i^{th} handling time with a modified-normal likelihood written as

$$H_i \sim N_{l_i}(e^{\mathbf{X}_i^T \boldsymbol{\beta}}, \sigma^2) \quad \text{eqn S12}$$

where the subscript l_i refers to the censoring “window” length and indicates that we added a $Unif(-\frac{l_i}{2}, \frac{l_i}{2})$ error to the normal distribution (corresponding to the interval censoring with which handling times were observed). As noted in the main text, the exponential link of eqn S12 avoids negative mean handling time estimates.

Treating the field covariates (predator size, prey size, and temperature) as random to account for sampling variability, we model the distributions of the (log-transformed) covariate observations $\mathbf{X}_1, \dots, \mathbf{X}_N$, where N is the total number of field observations, as being independent, identically distributed, and drawn from a multivariate normal distribution with mean vector $\boldsymbol{\mu}$ and covariance matrix Σ' . We use non-informative multivariate normal and inverse Wishart priors for $\boldsymbol{\mu}$ and Σ' respectively (Fink, 1997). Letting \mathbf{X}^* follow the posterior predictive distribution (our estimate of the distribution of the covariates), we may write the mean handling time as

$$E(H) = E[E(H|\mathbf{X}^*)] = E(e^{\boldsymbol{\beta}^T \mathbf{X}^*}). \quad \text{eqn S13}$$

As described in the main text, we can estimate this expectation by sampling from the regression parameters’ posterior distribution, sampling new covariates from their posterior predictive distribution, computing $e^{\boldsymbol{\beta}^T \mathbf{X}^*}$ for each sample, and averaging across all samples.

Model implementation: Putting the pieces together to estimate per capita attack rates

Using the likelihoods and priors of the feeding surveys, abundances and handling times, we draw samples from the parameters' posterior distributions using Markov Chain Monte Carlo (MCMC). We use JAGS with the R package 'rjags' for MCMC sampling (Plummer and Stukalov, 2014). We then combine parameter samples to produce samples from the attack rate posterior distribution on each prey species (see eqn S6). This treats handling times, \mathbf{H} , as being independent of the predator feeding surveys, \mathbf{F} , even though we use covariate observations of predator size, prey size and temperature from the feeding surveys informing \mathbf{F} to inform \mathbf{H} by combining them with the laboratory-based handling time regression coefficients associated with these covariates. We establish the validity of this assumption by examining the relationship between feeding proportions and covariate averages between the individual surveys (Appendix S5).

We verify Markov chain convergence using trace plots and the Gelman and Rubin convergence diagnostic (Gelman and Rubin, 1992), remove samples obtained before the chains had converged (i.e. burn-in time), and thin each chain to ensure independence among the remaining samples. We compute scale reduction factors – a convergence diagnostic that compares 'within' versus 'between' chain variability – using 250 independent chains with random initial values. We select burn-in times and thinning values separately for feeding survey, prey abundance, and handling time models based on trace plots and autocorrelation function plots. We base inferences on 1,000 samples after confirming that independent sets of 1,000 samples led to the same conclusions.

S3 *F-distribution median*

In general, the median of the F-distribution does not have a closed form. However, we can derive an approximation by relating the F-distribution to the beta-distribution.

Let $X \sim F_n^m$. We can express X as the ratio of scaled, independent Chi-squared distributions $C_m \sim \chi_m^2$ and $C_n \sim \chi_n^2$:

$$X = \frac{C_m/m}{C_n/n} \tag{eqn S14}$$

It follows that we can express X as the ratio of scaled independent gamma distributions

$G_m \sim \text{gamma}(\frac{m}{2}, 2)$ and $G_n \sim \text{gamma}(\frac{n}{2}, 2)$:

$$X = \frac{G_m/m}{G_n/n} \tag{eqn S15}$$

$$= \frac{n}{m} \frac{G_m}{G_n} \tag{eqn S16}$$

We can then normalize the gamma distributions:

$$X = \frac{n}{m} \frac{\frac{G_m}{G_m+G_n}}{\frac{G_n}{G_m+G_n}} \tag{eqn S17}$$

Letting $(D_1, D_2) \sim \text{Dirich}(\frac{m}{2}, \frac{n}{2})$ and using the relationship between Dirichlet and gamma distributions,

$$X = \frac{n}{m} \frac{D_1}{D_2} \tag{eqn S18}$$

Using the marginal distribution for Dirichlet components result and the fact that

$D_1 + D_2 = 1$, we have that

$$X = \frac{n}{m} \frac{B}{1 - B} \quad \text{eqn S19}$$

where $B \sim \text{Beta}(\frac{m}{2}, \frac{n}{2})$. Note that this is a monotone transformation of B , so it preserves the median. When $m > 2$ and $n > 2$, the median of B is approximately $\frac{\frac{m}{2} - \frac{1}{3}}{\frac{m}{2} + \frac{n}{2} - \frac{2}{3}}$ (Kerman, 2011). Substituting this result, we have that

$$\text{med}(X) = \frac{n}{m} \frac{\text{med}(B)}{1 - \text{med}(B)} \quad \text{eqn S20}$$

$$= \frac{n}{m} \frac{\frac{\frac{m}{2} - \frac{1}{3}}{\frac{m}{2} + \frac{n}{2} - \frac{2}{3}}}{\frac{\frac{m}{2} - \frac{1}{3}}{\frac{m}{2} + \frac{n}{2} - \frac{2}{3}}} \quad \text{eqn S21}$$

$$= \frac{n}{m} \frac{\frac{m}{2} - \frac{1}{3}}{\frac{n}{2} - \frac{1}{3}} \quad \text{eqn S22}$$

$$= \frac{n}{m} \frac{3m - 2}{3n - 2} \quad \text{eqn S23}$$

$$= \frac{n}{2n - 2} \frac{3m - 2}{m} \quad \text{eqn S24}$$

S4 *A Hierarchical Model for Abundances*

We also consider a hierarchical model as in Cressie et al. (2009) to account for the spatial and temporal structure in the prey abundance data. Although we lack sufficient data to estimate the parameters in this model (i.e. MCMC chains fail to converge with non-informative priors), we present the details here to show how our Bayesian approach can be extended to account for dependence due to spatial, temporal, or other structure.

For a single prey species, the data have the form Y_{ijkl} where i is the year (2005 or 2006), j is the season (Summer or Winter), k is the transect (1 or 2), and l is the quadrat (1 to 5). There are no data for summer 2005. Transects are different for each year and

season. To account for the structure of the data, the following hierarchical model can be used (distributions independent except for μ^* s):

$$Y_{ijkl} \sim ZIG(\alpha = \frac{\mu_{ijk}^{*2}}{\tau}, \beta_{ij} = \frac{\mu_{ijk}^*}{\tau}, \rho)$$

$$\rho \sim Unif(0, 1)$$

$$f(\tau) \propto \frac{1}{\tau}$$

$$\log(\mu_{ijk}^*) = \mu + \theta_i + \psi_j + \phi_{k(ij)}$$

$$\mu \sim N(0, 1000^2)$$

$$\theta_i \sim N(0, \sigma_\theta^2), \quad \psi_j \sim N(0, \sigma_\psi^2), \quad \phi_{k(ij)} \sim N(0, \sigma_\phi^2)$$

$$f(\sigma_\theta^2) \propto \frac{1}{\sigma_\theta^2}, \quad f(\sigma_\psi^2) \propto \frac{1}{\sigma_\psi^2}, \quad f(\sigma_\phi^2) \propto \frac{1}{\sigma_\phi^2}$$

In this model, the responses are ZIG with non-informative prior on the gamma variances and probability of a zero. The mean responses are related to covariates using a log link. The overall mean of the logarithms has non-informative prior $\mu \sim N(0, 1000^2)$. The year, season, and transects are independent and normally distributed with non-informative priors on the normal distribution variances. The parameter of interest – the overall mean – is $(1 - \rho)e^\mu$.

Hierarchical modeling of predator feeding survey and handling time data (field measurements of predator/prey size and temperature) can be done with a similar approach when sufficient data are available. In particular, more data at broader scales (e.g. year, transect) is needed to estimate the variability at these scales.

S5 *Accounting for dependence among information sources*

In our dataset, predator feeding surveys included covariate information (predator size, prey size, and temperature) that was used to estimate field handling times on the basis of regression models for handling times parameterized using laboratory data. In estimating attack rates we treat the field covariates as part of the handling times data \mathbf{H} and assume they are independent of the feeding proportions data \mathbf{F} . We assess the validity of this assumption by plotting the regression covariates versus the observed feeding proportions, as shown in Fig. S1. In this figure, every point represents a single feeding survey. The x-axes are the averages of the (log-transformed) covariate and the y-axes are the proportions of predators feeding. Only two species had sufficient data to be plotted and showed little evidence of a dependence.

If a lack of independence were evident it would need to be accounted for in the covariates distribution model. That is, although our model for the covariates was a multivariate normal, feeding survey level information (specifically proportions of predators feeding on each prey species) could be added to the model to affect its multivariate mean. This way, the mean covariate vector would be a function of the proportion of predators feeding on that prey type. Posterior distribution sampling could then be done by first sampling from the feeding proportions posterior distributions and then using the sampled feeding proportions to obtain samples from the handling times.

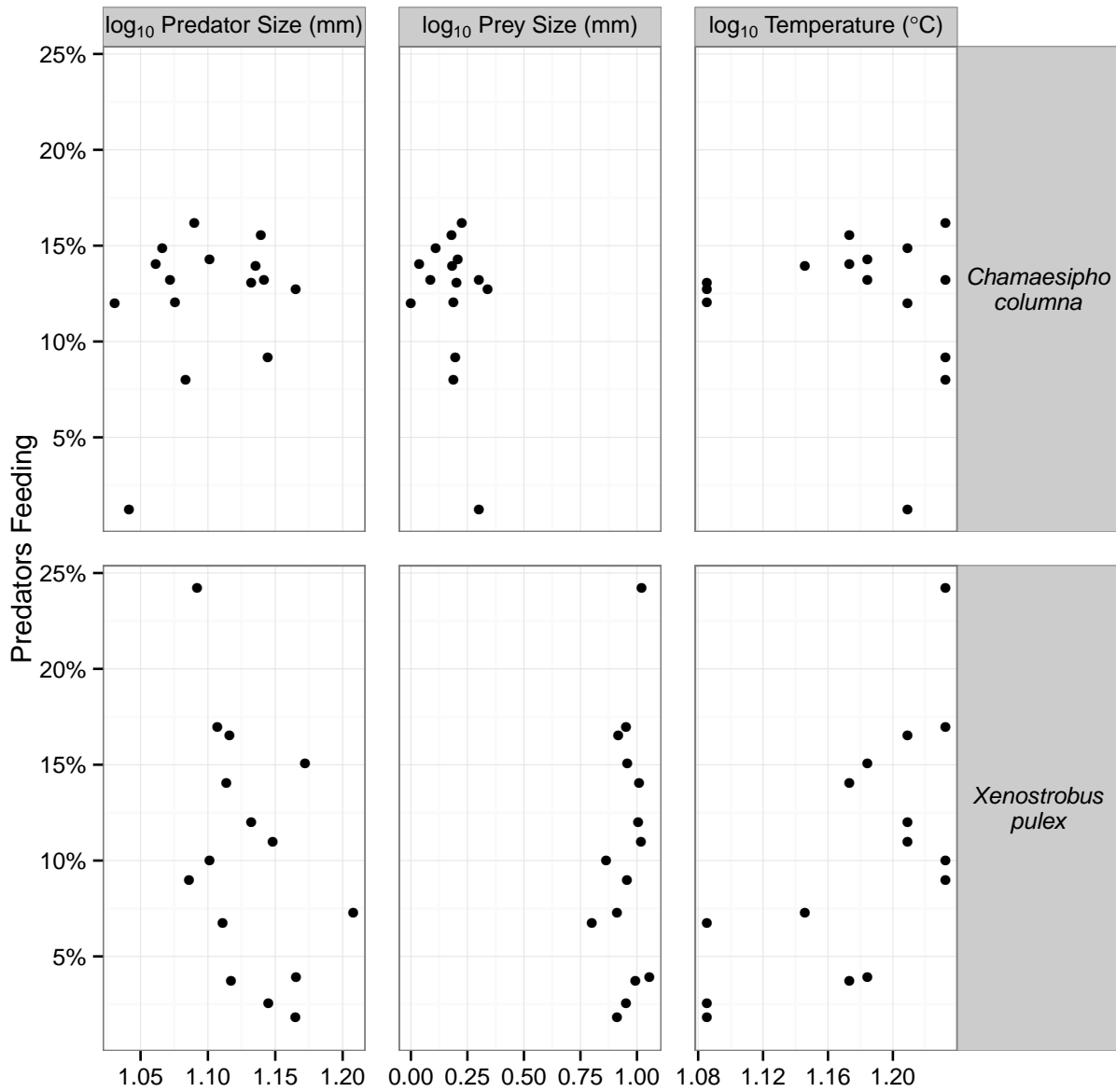


Fig. S1. Average field covariates versus feeding proportions. Each point corresponds to a single feeding survey. Only species that appeared in more than three separate feeding surveys are shown. Of the eight species and three covariates, only *Xenostrobis pulex* showed any evidence of a relationship between feeding proportions and feeding covariates (i.e., between F and H in eqn S7)

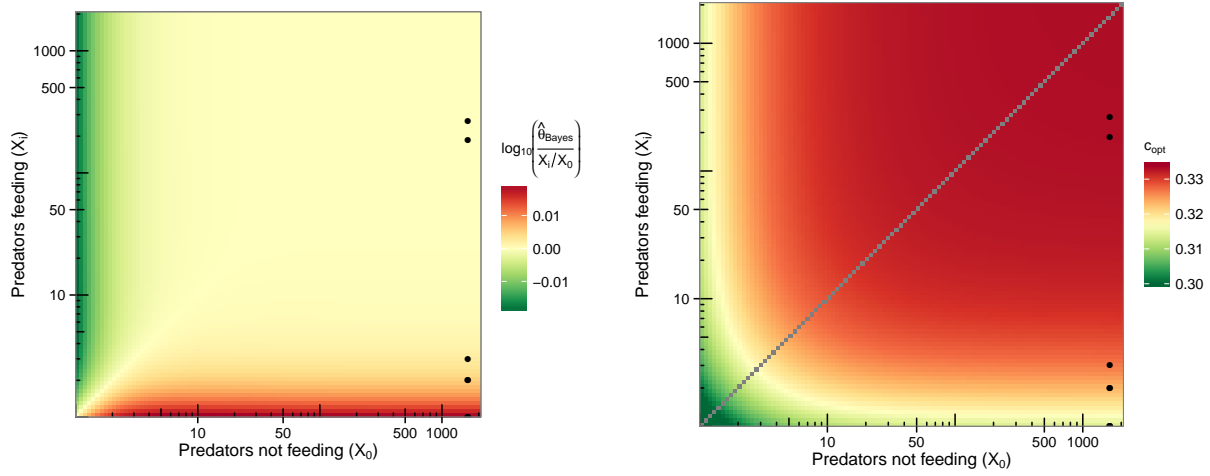


Fig. S2. Given the skewed nature of prey-specific per capita attack rate posterior probability distributions, the distribution median serves as a more appropriate point estimate than the mean. Fig. 1 illustrates the difference between the posterior median and maximum likelihood estimate of the ratio of feeding and non-feeding predators as a function of the number of feeding individuals, showing how the neutral ($c = \frac{1}{3}$) prior minimizes this difference. As a generalization of Fig. 1, in the left panel, we illustrate this difference as a function of both the number of predators observed feeding and the number observed not feeding. The right panel shows that the “optimal” value of c that minimizes this difference (a function of both feeding and non-feeding individuals) is typically around $\frac{1}{3}$. In both cases, the survey data from our example are shown as black dots

Table S1. Summary of notation used in this manuscript. We use capital letters for random variables (e.g. X_i) and lower-case letters for realizations of the random variables (e.g. x_i). Minor notation in the appendix that is not used elsewhere is generally not shown here

General type II functional response	
F_i	functional response on the i^{th} prey species
a_i	attack rate on the i^{th} prey species
h_i	handling time for the i^{th} prey species
N_i	abundance of the i^{th} prey species
S	number of prey species
Observed data	
F, A, H	data for feeding surveys, abundances, and handling times respectively
X_i, x_i	number of predators observed feeding on the i^{th} prey species
X_0, x_0	number of predators observed not feeding
N	total number of feeding predator field observations
\mathbf{X}_i	covariates vector for the i^{th} feeding predator
l_i	length of censoring window for the i^{th} handling time experiment
Probability distributions	
$f(x)$	generic probability distribution function
$Bin(n, p)$	binomial distribution with size n and probability p
$Mult_P(\alpha_0, \dots, \alpha_S)$	multinomial distribution with size P and probabilities $\alpha_0, \dots, \alpha_S$
$Dirich(c, \dots, c)$	Dirichlet distribution with concentration parameters c, \dots, c

$(\alpha_0, \dots, \alpha_S) \mathbf{x}$	distribution of $(\alpha_0, \dots, \alpha_S)$ conditional on \mathbf{x}
$Unif(-100, 100)$	uniform distribution with minimum -100 and maximum 100
$N_{l_i}(\mu, \sigma^2)$	normal dist. (mean μ , var. σ^2) plus $Unif(-\frac{l_i}{2}, \frac{l_i}{2})$ censoring error
Parameters	
α_i	population proportion of predators feeding on the i^{th} prey species
α_0	population proportion of predators not feeding
γ_i	ratio of multinomial probabilities $\frac{\alpha_i}{\alpha_0}$
c	Dirichlet distribution concentration parameters
α	gamma distribution shape
β	gamma distribution rate
ρ	probability of a zero for the zero-inflated gamma distribution
σ^2	handling time model variance
$\boldsymbol{\mu}$	mean vector for (log-transformed) field covariates
Σ'	covariance matrix for field covariates
ξ_i	population attack rate on the i^{th} prey species
ν_i	population abundance of the i^{th} prey species
η_i	population handling time for the i^{th} prey species
Other	
\mathbf{X}^*	random variable following the posterior dist. of the field covariates

Table S2. Predator feeding survey results grouped by predator size class. Predators were split into eight groups based on their size in millimeters (shown in top row). For the most frequently observed prey species (*Chamaesipho columna* and *Xenostrobus pulex*), we applied our Bayesian method using feeding survey results from each size class separately to assess how attack rates varied with predator size

Prey Species	6-10	11	12	13	14	15	16	17-28
<i>Austrolittorina antipodum</i>	0	1	0	0	1	1	0	0
<i>Austrolittorina cincta</i>	1	1	0	0	0	0	0	0
<i>Chamaesipho columna</i>	37	29	37	49	46	31	24	12
<i>Epopella plicata</i>	0	0	0	0	0	2	0	0
<i>Mytilus galloprovincialis</i>	0	1	0	0	0	0	0	0
<i>Notoacmea Radial</i>	0	0	0	0	0	1	0	0
<i>Risellopsis varia</i>	1	0	0	0	0	0	0	0
<i>Xenostrobus pulex</i>	19	13	22	37	31	27	20	16
Not Feeding	321	181	234	241	254	200	102	95
Total Surveyed	379	226	293	327	332	262	146	123

References

- Agresti A, Coull BA (1998) Approximate is better than “exact” for interval estimation of binomial proportions. *The American Statistician* 52:119–126
- Berger J (2006) The case for objective bayesian analysis. *Bayesian Analysis* 1:385–402
- Cressie N, Calder CA, Clark JS, Hoef JMV, Wikle CK (2009) Accounting for uncertainty in ecological analysis: the strengths and limitations of hierarchical statistical modeling. *Ecological Applications* 19:553–570
- Efron B, Tibshirani RJ (1994) *An Introduction to the Bootstrap*. CRC Press
- Fink D (1997) A compendium of conjugate priors. Technical Report
- Gelman A, Rubin DB (1992) Inference from iterative simulation using multiple sequences. *Statistical science* :457–472
- Kerman J (2011) A closed-form approximation for the median of the beta distribution. arXiv preprint arXiv:1111.0433
- Kilian L (1998) Small-sample confidence intervals for impulse response functions. *The Review of Economics and Statistics* 80:218–230
- Novak M, Wootton JT (2008) Estimating nonlinear interaction strengths: an observation-based method for species-rich food webs. *Ecology* 89:2083–2089
- Ospina R, Ferrari SL (2012) A general class of zero-or-one inflated beta regression models. *Computational Statistics & Data Analysis* 56:1609–1623
- Plummer M, Stukalov A (2014) *rjags: Bayesian graphical models using MCMC*
- Rosberg AG, Yanagi K, Amemiya T, Itoh K (2006) Estimating trophic link density from

quantitative but incomplete diet data. *Journal of Theoretical Biology* 243:261–272

Syversveen AR (1998) Noninformative bayesian priors. interpretation and problems with construction and applications. *Preprint statistics* 3