S-1. Prey assemblages and size distributions

Prey-size distributions often follow a log-normal function (Armstrong 2003, Quintana et al. 2008, Schartau et al. 2010). More complicated distribution forms can be represented by a superposition of $N$ log-normal functions. Let the coefficients $p_n$ obeying $\sum_i p_i = 1$, denote the relative contribution of each sub-distribution which again is characterized by specific values for mean (log) size $\bar{l}_i$ and variance $\sigma^2_i$.

$$P_i = P_{\text{tot}} \sum_{j = 1}^{\mathcal{N}} \frac{p_i}{\sqrt{2\pi\sigma^2_{ij}}} e^{-(\bar{l}_i - \bar{l}_j)^2/2\sigma^2_{ij}}$$  \(\text{(S1)}\)

This representation also emulates a multi-species (model) assemblage so that $\sigma^2_{ij}$ describes the size variance of a single population. The superposition notion enables collecting individual species $i'$ of a large assemblage into groups $i$. Within multi-species food-web models, different partitioning of total food biomass among many prey species leads to different outcomes and, thus, to an inconsistent dependency of total grazing on the partitioning resolution $N$ (Tilman 1982, Visser & Fiksen 2013). A species assembling algorithm based on Eq. (S1) avoids this anomaly. Using the group index $i$ the algorithm aggregates the species resolution by calculating group-based grazing rates $G_i$ that also allow meaningful interpretation of feeding half-saturation and affinity (see also Appendix A-IV in the main article). Still, the group-based rates $G_i$ have to be linked to the species-specific grazing rate $G_{i'}$ on prey $i'$. Each species contributes to 2 adjacent groups $i$ and $i+1$ where the group index $i$ derives from the condition $l_i \leq l_{i+1}$. The group characteristics $p_n, \bar{l}_i$, and $\sigma^2_i$ are recursively calculated from the contributing species, weighted according to the respective distance $l_{i+1} - l_i$. The procedure is repeated until the mean group sizes $l_i$ converge, which is generally the case after a few iterations. If the initial spacing or bin width, $\Delta l = l^0_{i+1} - l^0_i$, is set sufficiently large (e.g. $\Delta l^0 = 1/4$), total grazing rate $G = \sum_i G_i$ becomes largely independent from species resolution; the overall food availability in a broader size class enables a consistent application of functional grazing responses ($G_i(x)$, see below). $G_i$ is down-scaled to the individual species $i'$ using the selection kernel and Eq. (4): $G_{i'} = G_i \cdot P_i e^{-x/(l_{i+1} - l_i)} P_i e^{-x/(l_{i+1} - l_i)}$ with $s' = s/(1 + 2s \sigma^2_i)$, and subsequent renormalization by $\sum_i G_i$, to ensure mass conservation of the algorithm.

Eq. (S1) can also be used to represent multimodal distribution functions that can arise from intensive grazing within a certain sub-range (Hahn & Hofle 1999, Havlicek & Carpenter 2001, Schartau et al. 2010). In correspondence to grazing holes in the food spectrum, multi-modality of $P_i (N \geq 2)$ gains importance when assessing a model with size-selective grazing. Multi-modal or complicated prey size-distributions are to be studied in subsequent works, which can be realized via complementing the formulas derived here by the summation $\sum_{i=1}^{\mathcal{N}}$. Still, for many experiments with restricted food repertoire, a single elementary distribution ($N = 1$) will suffice. Even in situ samples for distinct functional groups (e.g. phytoplankton, zooplankton) are often fitted by one log-normal function if a reduced accuracy can be tolerated (Quintana et al. 2008). Therefore, the number of peaks in the food particle distribution is in this study shrunk to the case $N = 1$.

S-II. Functional grazing response and optimal affinity

The mechanistic grazing rate $G$ (Eq. A11 in the main article) derived in Appendix A-IV in the main article and Wirtz (2012) contains as major argument the ratio between external ($A_e$, $P$) and internal ($I_{\text{max}}$) food processing rates. At very large $x$, i.e. at high encounter capture rates, we have $G = I_{\text{max}} \cdot x \cdot (-x^n)/(-x^{n+1}) = I_{\text{max}}$; at very small $x$, grazing is affinity...
controlled \((G = \bar{I}_{\text{max}} \cdot x = A_x \cdot P)\). For the more typical intermediate values of \(x\), the grazing function \(G(n, x)\) covers a wide range of situations, from those with high intermittency in feeding sub-stages \((\text{low } n)\) to those with high synchrony \((\text{large } n)\); corresponding to these situations, \(G(n, x)\) in Eq. (A11) changes from a Holling-type-II \((n = 1)\), or an Ivlev \((n = 2)\), to a rectangular Holling-type-I \((n \to \infty)\) functional representation. \(A_x\) in Eq. (A11) represents the food affinity that in turn is directly related to variable feeding activity of the consumer. Within the size-based approach proposed here, it is hardly conceivable how this activity can be distinguished with respect to single prey size classes. Activity can either be absent or present, no matter what the specific characteristics of the next encountered particle are. Activity is a prerequisite of selection \(f_{\text{sel}}\) and cannot be formulated as a function of prey size. A predator may reject a certain prey size class \(\text{(described by } f_{\text{sol}}\text{)}, or can decide to become inactive under generally unfavorable feeding conditions, a situation that may reflect the prey size distribution as a whole, but the pursue state \(A_x\), as such, cannot be written in meaningful terms of \(f\). As a consequence, \(G\) also has to be formulated independently from specific prey size classes, but using bulk properties of the prey field.

Behavioral down-regulation of clearance activity \(A_x\) has been proposed by Wirtz (2013) to follow optimal foraging theory. Optimal clearance activity in \(x\) \((A_x)\) fulfills
\[
\frac{\partial G}{\partial x} = \bar{I}_{\text{max}} \cdot x'  \tag{S2}
\]
with \(x' = \frac{P_{\text{crit}}}{P}\) and
\[
\frac{\partial G}{\partial x} = \bar{I}_{\text{max}} \frac{1 + (nx - n - 1)x^n}{(1 - x^n)^2}  \tag{S3}
\]
\(P_{\text{crit}}\) defines the food concentration required to compensate activity costs and risks. These consumer losses comprise energy expenditures for filtering activity and swimming (Lehman 1976), or the related risk of being detected by a predator (Visser et al. 2009). Eq. (S2) can be solved using the numerical approximation: \(A_x = \bar{I}_{\text{max}} / P e^{-y}\) with \(y = 10x'/(n + 1) - 4/(n + 2)\).

Using such estimates of \(A_x\), we could further integrate over the functional response dependent on \(I_{\text{max}}(t)\) and \(P\), times the feeding kernel \(f_{\text{sol}}\). In theory, the average grazing rate \(\int \bar{G}(A_x, f_{\text{sol}}(t), P, I_{\text{max}}(t)) dt\) is a more precise quantity than the grazing function computed at the average characteristics \(\bar{I}_{\text{max}}\) and \(P\). If a second-order moment approximation is used, however, this integral again becomes a function of \(I_{\text{max}}\) and \(P\) and a term containing the prey size variance \(\sigma^2\) (e.g. Bolker & Pacala 1997, Wirtz 2000). Hence the 2 average feeding characteristics \((\bar{I}_{\text{max}}\) and \(P\)) calculated here already facilitate incorporation of size-selective grazing into ecosystem models. The same holds for multi-species prey representations as shown in Appendix A-II in the main article.

### S-III. Optimality in feeding selectivity

In this work, optimal foraging theory is also used to estimate the degree of selectivity that maximizes grazing. As shown in Fig. 6 in the main article, \(G\) is a uni-modal function of \(s\) because higher selectivity increases the efficacy of ingestion \((\bar{I}_{\text{max}})\) but lowers the availability of food \((P)\). The mathematical condition for grazing optimization in terms of selectivity,
\[
\frac{\partial G}{\partial s} = \frac{\partial G}{\partial \bar{I}_{\text{max}}} \frac{\partial \bar{I}_{\text{max}}}{\partial s} + \frac{\partial G}{\partial P} \frac{\partial P}{\partial s} = 0  \tag{S4}
\]
contains the 2 partial derivatives that follow from Eq. (A11)
\[
\frac{\partial G}{\partial \bar{I}_{\text{max}}} = \frac{\partial G}{\partial \bar{I}_{\text{max}}} \frac{x}{I_{\text{max}}} \frac{\partial I_{\text{max}}}{\partial x} \frac{\partial x}{\partial \bar{I}_{\text{max}}} \quad \text{and} \quad \frac{\partial G}{\partial P} = \frac{\partial G}{\partial P} \frac{x}{\partial \bar{I}_{\text{max}}} \tag{S5}
\]
Thus, the selectivity effect on maximal ingestion rate \(\frac{\partial \bar{I}_{\text{max}}}{\partial s} = 3\bar{I}_{\text{max}}/(4s^2 + 6s)\) can be calculated from Eq. (5), and the effect on food availability \(\frac{\partial P}{\partial s} = -P\sigma^2/(1 + 2s\sigma^2)\) from Eq. (4). Although exact solutions of Eq. (S3) need to be found numerically, model implementation may be facilitated by approximate, closed expressions, similar to the solution for \(A_x\). At low food \((P \approx 5P_{\text{crit}})\), we can rewrite the grazing term \(G = \bar{I}_{\text{max}} \cdot x \cdot n/(n + 1)\), the last factor coming from a linearization of the terms in Eq. (A11) \((e.g. x^n \approx 1 - n\log x)\). At low to intermediate values of \(s\), we can also simplify \((1 + 2s\sigma^2)/(4s^2 + 6s)\) to \((1 + 3\sigma^2)/12\). Then, by using the activity optimization Eqs. (S2) & (S3), we obtain an estimate for the optimal selectivity,
\[
s = \frac{1 + 3\sigma^2}{12} \frac{n}{P_{\text{crit}}} \frac{n + 1}{P_{\text{crit}}}  \tag{S6}
\]
where the additional correction in the denominator \((+1/12)\) stabilizes the expression at low \(\sigma^2\). At highly abundant food levels \((P \approx 5P_{\text{crit}})\), optimal selectivity according to Eq. (S3) is in the order of \(\bar{I}_{\text{max}}/(\partial G/\partial x)^{-1}\) which can be very large dependent on the functional response at high \(P\), here quantified by the synchrony parameter \(n\). At large \(n\) and \(P\), selectivity optimization Eq. (S3) at \(P \gg P_{\text{crit}}\) may result in \(s \approx 10\) up to 100. The latter case translates to a size feeding range of \(1-2 \mu m\) for a consumer specialized on \(10 \mu m\) items.
LITERATURE CITED


