

Trophic niche-space imaging, using resource and consumer traits

Leopold A. J. Nagelkerke^{1,*} and Axel G. Rossberg^{2,3,4}

Published in *Theoretical Ecology* on 2 July 2014

<http://dx.doi.org/10.1007/s12080-014-0229-5>

¹ *Aquaculture and Fisheries Group, Wageningen University, Wageningen Institute of Animal Sciences (WIAS), De Elst 1, 6708 WD Wageningen, The Netherlands, e-mail: Leo.Nagelkerke@wur.nl, Tel: +31 317 483940, fax: +31 317 483937*

² *Lowestoft Laboratory, Centre for Environment, Fisheries and Aquaculture Science (Cefas), Pakefield Road, Lowestoft, Suffolk NR33 0HT, United Kingdom, e-mail: axel.rossberg@cefasc.co.uk, Tel: +44 1502 524575, fax: +44 1502 513865*

³ *School of Biological Sciences, Queen's University Belfast, Belfast BT9 7BL, United Kingdom*

⁴ *School of Environmental Sciences, University of East Anglia, Norwich NR4 7TJ, United Kingdom*

* Author for correspondence

Abstract

The strength of trophic (feeding) links between two species depends on the traits of both the consumer and the resource. But which traits of consumer and resource have to be measured to predict link strengths, and how many? A novel theoretical framework for systematically determining trophic traits from empirical data was recently proposed. Here we demonstrate this approach for a group of 14 consumer fish species (*Labeobarbus* spp., Cyprinidae) and 11 aquatic resource categories coexisting in Lake Tana in northern Ethiopia, analysing large sets of phenotypic consumer and resource traits with known roles in feeding ecology. We systematically reconstruct structure and geometry of trophic niche space, in which link strengths are predicted by the distances between consumers and resources. These distances are then represented graphically resulting in an image of trophic niche space and its occupancy. We find trophic niche to be multi-dimensional. Among the models we analysed, one with two resource and two consumer traits had the highest predictive power for link strength. Results further suggest that trophic niche space has a pseudo-Euclidean geometry, meaning that link strength decays with distance in some dimensions of trophic niche space, while it increases with distance in other dimensions. Our analysis not only informs theory and modelling, but may also be helpful for predicting trophic link strengths for pairs of other, similar species.

Keywords Interaction strength; Trophic links; Food-web model; Trophic traits; Niche space; Cyprinidae; Fishes

Introduction

The understanding of food-web structure and trophic interactions is the subject of continuous attention (Bersier 2007), not only out of scientific interest, but also because it could be instrumental in the effective management of exploited ecosystems (Dejen et al. 2006; Layman et al. 2005; Sibbing et al. 1994).

The question what determines the strength of trophic interactions between species has been addressed at several levels of abstraction by empiricists and theorists, but the fundamental underlying intuition is similar. The trophic link strength a_{ij} between a resource i

and a consumer j is typically modelled as a function of quantitative traits of resource and consumer (Eklöf et al. 2013; Emmerson et al. 2005; Klecka and Boukal 2013; Link 2004; Petchey et al. 2008; Pouilly et al. 2003; Rohr et al. 2010; Rossberg et al. 2006; Rossberg et al. 2008; Rossberg et al. 2010b; Rossberg et al. 2010a; Russo et al. 2008; Sibbing and Nagelkerke 2001; Stouffer et al. 2011; Wainwright and Richard 1995; Williams and Purves 2011; Yoshida 2003). That is, one assumes existence of minimal quantitative characterizations of species in terms of vectors \mathbf{t}_i , \mathbf{t}_j of real numbers (or subsets of real numbers), called the *trophic traits* of species, such that $a_{ij} = a(\mathbf{t}_i, \mathbf{t}_j)$ for any resource-consumer pair (i, j) considered in a model. The function $a(\mathbf{s}, \mathbf{t})$, called the *link-strength function*, is understood to be continuous in its two arguments \mathbf{s} and \mathbf{t} (Rossberg 2013). Because the trophic traits entering the link strength function effectively determine the trophic niche of a species, the subspace of the full traits space that is spanned by trophic traits can be identified with the *trophic niche space*. Defining as *vulnerability traits* those trophic traits that determine the role of a species as a resource, and as *foraging traits* those that determine the role of species as a consumer, one can alternatively interpret trophic niche space as the space spanned, e.g., by the foraging traits alone (Cohen 1977; Rossberg et al. 2010b). We follow this convention here.

An important caveat to be kept in mind in this framework is that the notion of trophic link strength as a quantity determined exclusively by the identities of consumer and resource species is a theoretical idealization (Arditi et al. 2005; Rossberg et al. 2010b). Trophic link strengths can be defined through their roles as parameters in population-dynamical models, typically entering such models *via* formulae for functional responses. The unavoidable simplification of real population dynamics by these formulae directly translates to an inherent inaccuracy of the concept of trophic link strength itself. This, in turn, fundamentally limits the accuracy at which empirical trophic link strengths can be predicted from trophic traits. Parsimony then dictates use of rather simple functional forms for link-strength functions, among which none should be expected to be the single “correct” choice. These considerations support the wide variety of approaches to the problem that can be found in the literature. The question is not so much which particular choices for link-strength function and trophic traits are correct, but rather in which context they can be empirically supported.

It is widely agreed that one of the most important phenotypic traits determining foraging capacity and vulnerability to predation is body size (Berlow et al. 2009; Emmerson et al. 2005; Layman et al. 2005; Stouffer et al. 2011; Vucic-Pestic et al. 2010). In addition, a broad body of natural-history studies suggests that other traits besides size are also relevant for trophic interactions (Barnett et al. 2006; Bhat 2005; Carlson and Wainwright 2010; Naisbit et al. 2012; Rohr et al. 2010; Russo et al. 2008; Schmitz and Price 2011; Spooner and Vaughn 2008). Such traits are typically related to structures involved in feeding, and were mostly studied through correlative approaches (Bhat 2005; Pouilly et al. 2003; Russo et al. 2008), while others focused on the functionality of traits in relation to the mechanics of feeding (Dejen et al. 2006; Link 2004; Sibbing and Nagelkerke 2001; Wainwright and Richard 1995).

In food-web studies, in turn, several models have been developed that reconstruct empirical food-web structure based on hypothetical abstract traits (or “latent traits”) of consumers and resources (Rohr et al. 2010; Stouffer et al. 2011; Williams and Purves 2011), without explicitly linking these to specific phenotypic traits of the species involved. However, this approach faces the question to what extent these abstract traits represent real ecological traits of species, rather than just being fitting parameters of food-web models, contingent on the particular datasets analysed. Recently Rossberg et al. (2010b) proposed a theoretical framework that allows for a systematic deduction of the trophically relevant combinations of

phenotypic traits of consumers and resources from empirical data, and of their relationships to abstract trophic traits. The framework proposes to use generic quadratic polynomials in consumer and resource trait values to model logarithmic link strengths — an approach that was now demonstrated for laboratory data by Klecka and Boukal (2013) — and then a mathematical analysis of these polynomials to extract the relevant trait combinations: a step performed for the first time here. By expressing trophic traits predictive of food-web structure in terms of measured phenotypic traits, the biological plausibility of abstract trophic traits can be assessed and their ontological status as proper ecological traits of species asserted.

Besides the issues of which and how many phenotypic traits predict trophic link strength, there is another question, central to food-web studies, concerning the dimensionality of trophic niche space (Cohen 1977; Eklöf et al. 2013). Trophic niche space would be one-dimensional if a single “niche value” could be computed that would allow good predictions of whether a given consumer feeds on a resource or not. This value could be a single phenotypic trait, such as body size, but also a combination of several phenotypic traits. If more than one such number is needed, trophic niche space can be considered multi-dimensional. The number of independent numerical values determining a species’ role as consumer or as resource (at a given level of accuracy) can be understood as the number of dimensions of trophic niche space. Trophic niche-space dimensionality has implications for food-web structure and the distribution of trophic link strengths (Allesina et al. 2008; Cohen 1977; Rossberg et al. 2010a; Rossberg et al. 2011). Good fits to food-web structure have been demonstrated with models that work with one, or a few dimensions (Rohr et al. 2010), but also for models invoking high-dimensional trophic niche spaces (Rossberg et al. 2006; Rossberg et al. 2008). Considerable ambiguities persist regarding the role of body size and other phenotypic traits in determining food-web structure (Naisbit et al. 2012), and the discussion about trophic niche space dimensionality continues. The present study is a contribution to unravelling these issues.

Our main objective here is to reconstruct trophic niche space by deriving a set of quantitative abstract trophic traits, defined in terms of phenotypic traits, that locate each consumer and resource category in trophic niche space in such a way that the relative positions of consumers and resources predict their link strengths. In the course of this analysis, we will address questions regarding the number and nature of the phenotypic traits needed, and the dimensionality and geometry of trophic niche space.

Our study system is the endemic group of cyprinid fishes (*Labeobarbus* species) from Lake Tana in northern Ethiopia (Nagelkerke et al. 1994; Nagelkerke and Sibbing 2000). We test to what extent the link strengths of 14 of these fish species with the 11 most abundant aquatic food resources (species and food resources in Online Resource 1, Table S1) can be explained based on their phenotypic traits. The *Labeobarbus* group is well-suited for this study since their feeding-related morphological traits and their diets were extensively studied. Moreover, the morphological differences between these species are likely to be the direct result of adaptive radiation, because these fishes most probably originated in the lake from a common ancestor in less than 17,000 years (de Graaf et al. 2010). The selection of phenotypic traits of the fishes and their resources is based on their known relationships to feeding ecology, thereby enabling a direct mechanistic interpretation of the role of these traits in trophic interactions (Sibbing and Nagelkerke 2001; Wainwright and Richard 1995).

The model

The framework developed by Rossberg et al. (2010b) was analysed further and reformulated by Rossberg (2013), who proposed the following link-strength function relating the strength a_{ij} of the trophic link between a resource (indexed by i) and a consumer (j) to their trophic traits:

$$a_{ij} = a_0 \exp \left[V_i^* + F_j^* - \frac{1}{2} \sum_{k=1}^D \sigma_k \left(v_i^{(k)} - f_j^{(k)} \right)^2 \right]. \quad (1)$$

The scale constant a_0 has dimensions of link strength, depending on the particular measure for link strength used. The *baseline vulnerability* V_i^* and the D abstract *vulnerability traits* $v_i^{(k)}$ depend only on phenotypic traits of the resource; and the *baseline foraging strength* F_j^* , and the D abstract *foraging traits* $f_j^{(k)}$ depend only on phenotypic traits of the consumer. The baseline vulnerability and foraging strength (V_i^* and F_j^*) represent overall fitness of resource and consumer with respect to trophic interactions, independent of the interaction partner (Rossberg et al. 2010b; Rossberg 2013). The number D is interpreted as the dimensionality of trophic niche space. The variables $v_i^{(k)}$ and $f_j^{(k)}$, which can be interpreted as coordinates specifying the positions of resources and consumers in trophic niche space, are collectively referred to as *proper trophic traits*. The coefficients σ_k have values of either +1 or -1. They are, next to the dimensionality D , free parameters of the link-strength function and need to be determined empirically. The vector $(\sigma_1, \dots, \sigma_D)$ consisting of D coefficients of values +1 or -1, is called the *signature* of the model. It plays a key role in its geometric interpretation, as we explain in the Discussion. We call our representation of link strengths in terms of trophic traits a *Trophic Trait Model (TTM)*.

By convention, the D dimensions of niche space are indexed in order of decreasing variability of abstract vulnerability and foraging traits (specifically, such that $\text{var } v_i^{(k)} + \text{var } f_j^{(k)}$ decreases with increasing k , based on sample variances over all consumers and resources considered). In principle, the number of dimensions D is large, but if variability of traits decays fast enough with increasing k , terms corresponding to higher dimensions k in Eq. (1) can be approximated by constants and absorbed in the scale constant a_0 . The effective dimensionality D is therefore limited by the desired precision, the precision at which the general form of Eq. (1) holds, and the precision of the available data used to fit the model.

We devised a practical method for fitting the model given by Eq. (1) to empirical data. Our method follows and extends the method proposed by Rossberg et al. (2010b). It assumes additive linear relationships between abstract trophic traits and phenotypic traits for simplicity, but suitable non-linear transformations (e.g. logarithmic or logit transformations) of the phenotypic traits can be performed prior to its application. The method yields the values of D and σ_k ($k = 1, \dots, D$) and results in mapping functions from phenotypic traits of resources i and consumers j onto the abstract traits V_i^* , F_j^* , $v_i^{(k)}$, and $f_j^{(k)}$ ($k = 1, \dots, D$). These are determined such that a_{ij} , given by Eq. (1), approximates the measured link strengths (details of the fitting algorithm are described in Online Resource 2).

By relating the abstract trophic traits to phenotypic traits, as we do here, the number of free model parameters does not increase with the number of species, as it otherwise would (Rohr et al. 2010). As a result, the number of parameters becomes much smaller and the fitting procedure more robust to over-fitting. In addition, results are more easily checked for plausibility, and the fitted model can be applied to species for which their relevant phenotypic traits, but not their abstract trophic traits, are known, thus admitting a prediction of their trophic links.

Ideally, a large number of different body plans and behavioural repertoires of consumers and resources would be covered by a single TTM, because this would allow reconstructing complete food webs from the phenotypic traits of their member species. We purposely restrict

ourselves here to the simpler case where all consumers are closely related, while still considering widely varying resources. Since the basic morphological pattern of all consumers analysed is identical, this enhances the functional interpretation of morphological differences between species.

Material and methods

Consumer and resource traits

We used a matrix of 19 morphological traits measured for the 14 *Labeobarbus* species in a previous study (Sibbing and Nagelkerke 2001). These traits relate to all aspects of feeding, from detection to digestion and were selected *a priori* based on their supposed role in the whole foraging and feeding process as currently understood for cyprinid fishes (Sibbing 1991b; Sibbing 1991a; Sibbing and Nagelkerke 2001). The feeding process can be subdivided into search and detection, approach, intake, size selection, taste selection, transport, pharyngeal mastication and digestion. Most of these activities (except for taste selection and mastication) were represented by at least one phenotypic trait. Measurements were performed on more than 1,300 fresh, or freshly frozen and thawed specimens caught in 1992 – 1995. Traits were either of dimension length (and measured in mm) or angular. Measures of dimension length were corrected for overall size by using ratios, or expressing them in units of body length, specifically fork length, the length from snout tip to the fork of the tail fin. Fork length itself was also included in the matrix as a measure of overall fish size. Since all measured specimens attained their adult shape (all were > 15 cm in length) we did not expect complications resulting from allometric growth and ratios were considered as reliable shape descriptors (Sibbing and Nagelkerke 2001). We used mean values per species for all consumer traits and log-transformed them to accommodate variations in the scaling of morphological traits with fork length (Albrecht et al. 1993).

The 11 food resource categories were characterized using 11 traits, which were selected *a priori* based on the supposed challenges they present for fishes feeding on them (Sibbing and Nagelkerke 2001). Traits refer to size, velocity, shape, habitat, digestibility, and mechanical properties of the resources. Resource traits were quantified along ordinal scales with 2 to 5 values and were therefore directly comparable, independent of the actual measuring units (traits of consumers and resources in Online resource 1, tables S2 and S3, and Figure S1).

Trophic link strengths

For simplicity, we assume linear functional responses, and so define link strength as the biomass flux from resource to consumer relative to resource biomass abundance and consumer numerical abundance. This would ideally be calculated from the resource biomass, the consumer's total intake rate and its diet (de Ruiter et al. 1995; Fath et al. 2007). Total intake rates per individual are expected to be related to body size through an allometric scaling law (*i.e.* a power law). In the TTM, Eq. (1), this allometric scaling can be represented by a linear dependence of the baseline foraging traits F_j^* on logarithmic fork length. This means that there will be variation in F_j^* due to size differences between species: the last term of the equation, however, will not be affected and the results for dimensionality D and the proper trophic traits $v_i^{(k)}$ and $f_j^{(k)}$ remain the same. Since absolute intake rates were not available for our study system, we disregarded this allometric variation in F_j^* , taking instead volumetric diet proportions as representative for intake rates. Sufficiently precise direct measurements of resource biomasses were not available either. Instead, we used the mean proportions by volume that resource categories contributed to diets – which were recorded in more than

4,700 specimens in a previous study (Sibbing and Nagelkerke 2001) (Online resource 1, Table S4) – as estimates of “abundance”. Just as for the absolute intake rates, inaccuracies in abundance estimates would not much affect results for niche-space dimensionality D and proper trophic traits $v_i^{(k)}$ and $f_j^{(k)}$, because these could be compensated by adjusting in Eq. (1)

the baseline vulnerability traits V_i^* of resources alone. Thus, we effectively quantified “link strength” here by the proportion by volume that a resource contributes to the diet of a consumer divided by the mean of this proportion over all consumers. With link strength so defined, no resource item can have low link strength with all consumers, which is a reasonable assumption given that resource categories were selected *a priori* based on their relevance for the studied set of consumers. The viability of this simplified quantification of link strength is investigated *a posteriori* through a cross-validation of the TTM and an analysis of how the predictability of link strengths increases with information on traits. If our estimates of link strengths were too inaccurate to be meaningful, increasing information on traits would not help predicting them.

Fitting the trophic trait model (TTM)

Our procedure for estimating the TTM from field data has three main aspects: (i) Estimation of model parameters based on the observed diets, a given set of empirical trait variables, and a prescribed number D of dimensions of trophic niche space. This requires finding the signature $(\sigma_1, \dots, \sigma_D)$ and representations of the abstract trophic traits in Eq. (1) in terms of additive linear functions of phenotypic trait variables, which together reproduce observed link strengths as accurately as possible. In principle, this could be done using any model-fitting procedure. However, standard methods using iterative optimization algorithms tend to be rather slow and sensitive to initial conditions. Another method was therefore used, which we describe in detail in Online resource 2. It exploits the fact that the right-hand side of Eq. (1) is formally similar to a multivariate normal distribution and can, in principle, be interpreted as the probability per unit time for a trophic interaction to occur conditional to both consumer and resource being present. The method is fast and, as we verified using synthetic input data (now shown here), robust to measurement errors.

(ii) Cross-validation of the model to assess its predictive power. Cross-validation asks if a model can make predictions for cases that were not taken into account when fitting the model. This was here done by verifying how well the link strength of any conceivable consumer-resource pair could be predicted from the TTM when all information on this particular consumer and resource was ignored. Predictive power was quantified by the correlation between the so predicted and the measured link strengths. Since link strengths are less accurate for less abundant resources, diet categories were weighted by the squares of their estimated abundances when computing these correlations. This corresponds to 1/variance weighting, which is optimal in the maximum likelihood sense (Piegorisch and Bailer 2005).

(iii) Choice of a set of empirical trait variables appropriate for achieving high predictive power. This was done by an exhaustive search through all combinations of up to two resource traits and two consumer traits, retaining the combinations that scored best in cross-validation. Larger sets of traits were not considered to avoid a combinatorial explosion of the cases considered, which can lead to unnoticed overfitting (Reunanen 2003) and because our goal here is not to find optimal numbers of traits or niche-space dimensions, but only to determine lower bounds. Variable selection requires repeated cross-validation, and cross-validation repeated parameter estimation. This makes this procedure, which we coded in R (R Development Core Team 2010), computation intensive. It was carried out separately for each number of dimensions D ranging from 1 to 4. More than four dimensions were not tested,

Table 1. Correlations between predicted and realized link strengths in models with increasing niche-space dimensionality. The traits of both the resources and the consumers used in the selected models are listed as well as the correlations obtained in leave-one-out cross-validation and with the selected model fitted to the entire dataset.

dimensions	traits of resource used	traits of consumer used	correlation in cross-validation	correlation of selected model
1	diameter; toughness	fork length; eye diameter	0.57	0.14
2	pelagic; macro-reduction	barbel length	0.69	0.66
3	pelagic; elongate	barbel length	0.66	0.65
4	pelagic; macro-reduction	fork length; barbel length	0.74	0.72

because only up to four different phenotypic trait variables were considered. Details on fitting the TTM are provided in Online resource 2.

Evaluating relevance of coordinates

Since the question of the dimensionality of trophic niche space is of high theoretical importance, a specific test was carried out to determine the relevance of the higher niche-space dimensions. Specifically, the statistical significance of the information contributed by higher dimensions was tested against the null-hypothesis that, for a fixed cut-off-dimension D_c , a random assignment of trophic traits $v_i^{(k)}$ and $f_j^{(k)}$ for $k \geq D_c$ could have led to an equally good fit to the data as the assignment determined by our fitting procedure. To test this hypothesis, the values of $v_i^{(k)}$ and $f_j^{(k)}$ were randomly permuted over all consumers and resources, respectively. We did this independently for each $k \geq D_c$. Using 10,000 Monte-Carlo iterations, we computed 2.5% and 97.5% quantiles of the correlation between observed and predicted link strengths, weighted by resource abundance. This analysis was carried out for all TTMs obtained with $D = 1-4$. Results obtained for the TTM with $D = 4$ are shown here, the corresponding results for lower D can be found in online resource 3.

Predicted diet proportions

Based on the predicted link strength and the assumed resource abundances, we computed the expected consumer diet fractions contributed by each resource. The relationship between observed and predicted diet fractions was evaluated using reduced major axis (RMA) regression (Quinn and Keough 2002), because neither of the variables could be considered as ‘independent’.

Results

Optimal combinations of phenotypic trait variables and the predictive power they achieved are listed in Table 1. We note that in Table 1 the correlation between prediction and observation achieved in cross-validation is always higher than the correlation found when the TTM is fitted to the full dataset, a phenomenon that indicates potential overfitting by variable selection as discussed by Reunanen (2003). However, the strength of the effect declines with increasing dimensionality D , and for $D \geq 2$ is less than a few percent. The correlations based on fits to the full dataset are the more reliable measures of predictive power ultimately achieved, because they were not used for variable selection. Containing this effect was one reason for allowing only combinations of up to two resource traits and two consumer traits.

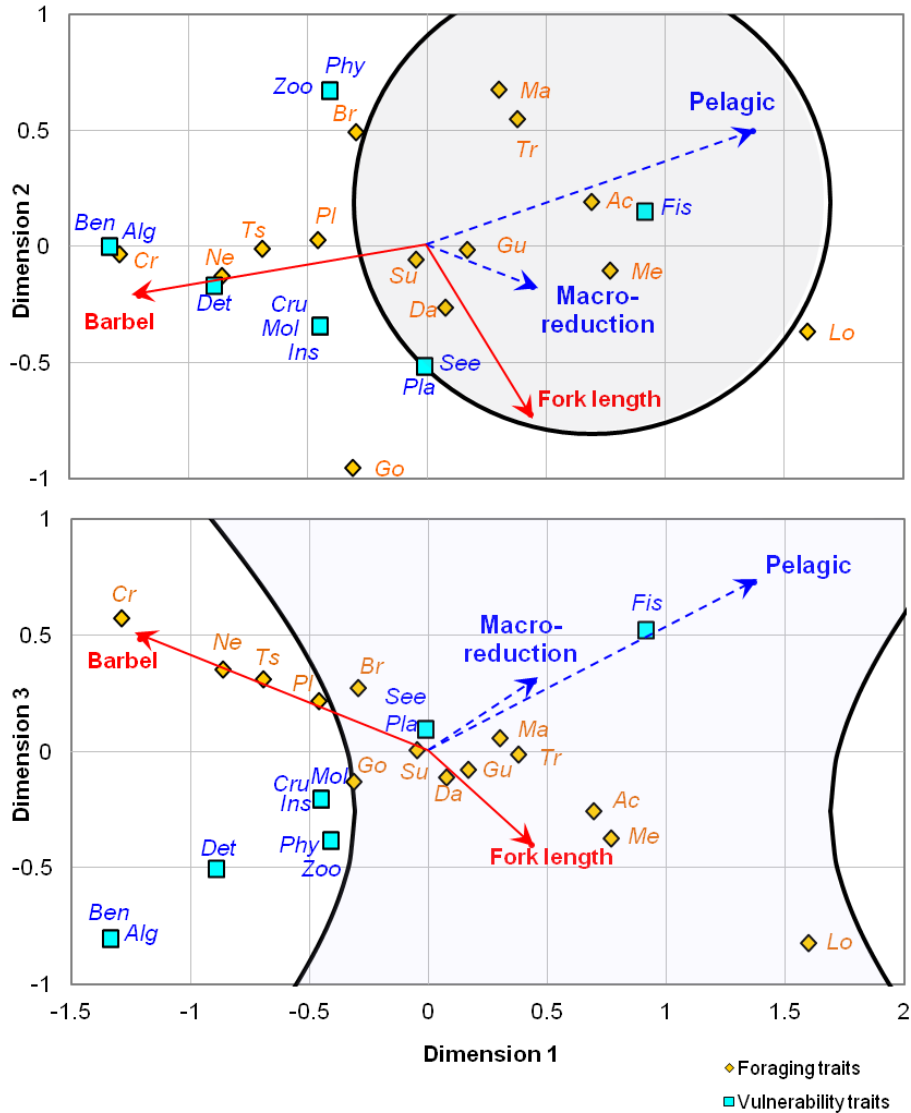


Figure 1. Projection of dimensions 1 and 2 (top) and 1 and 3 (bottom) of the niche space of 14 *Labeobarbus* species (consumers: diamonds) and 11 resource categories (squares). The phenotypic traits which explain the ordination for consumers (solid arrows) and resources (dashed arrows) are also indicated. To outline the trophic niche of one consumer species (*L. acutirostris*, Ac) a circle (top) and a hyperbola (bottom) corresponding to a pseudo-Euclidean distance of 1 have been drawn around this species. Abbreviations as in Online resource 1, Table S1.

The best predictive power was found for a trophic niche space with $D = 4$ dimensions, the highest dimensionality considered (Table 1).

We will, in the following, concentrate on the case $D = 4$. The ecological validity of selected variables and model fit in this case is supported by the observation that the explanatory variables selected for $D = 4$ occur also at lower dimensionalities (Table 1). Moreover, the resulting embedding of resources and consumers in trophic niche space has a clear ecological interpretation. The resource traits selected for $D = 4$ are ‘pelagic’, which characterizes the main habitat, and ‘need for macro-reduction’, which is a categorical trait indicating whether the resource needs to be masticated before swallowing. The selected consumer traits are body size, measured as ‘fork length’, and ‘barbel length’, a trait related to the detection of especially benthic resources (Sibbing and Nagelkerke 2001).

Table 2 specifies the mapping of these phenotypic traits into the four-dimensional trophic niche space by the TTM, the resulting niche-space coordinates of consumers and resources,

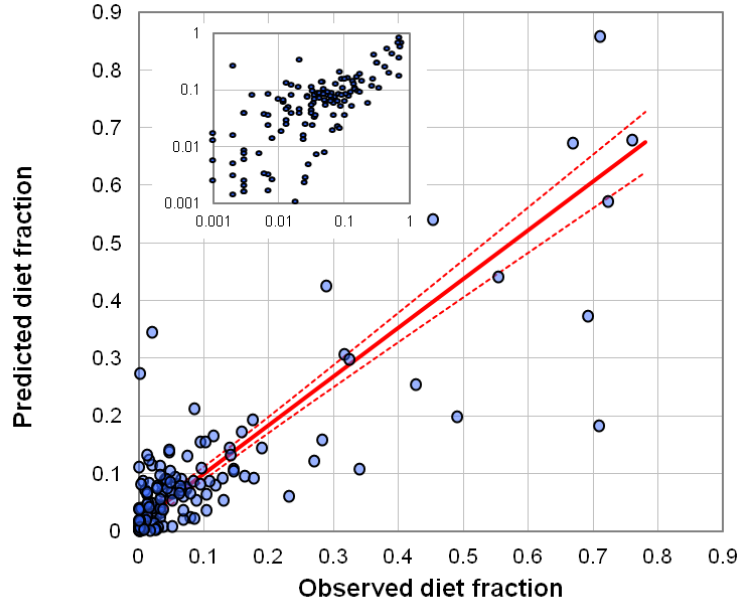


Figure 2. Relationship between observed diet and predicted diets. The solid line is a reduced major axis (RMA) regression line, with the 95% the confidence band indicated by the stippled lines. The inset shows the same data on a double log-scale.

and weight ($\text{var } v_i^{(k)} + \text{var } f_j^{(k)}$) and signature (σ_k) of each dimension k . The σ -values for the leading two dimensions are positive, implying, by Eq. (1), that link strength decays with increasing distance between resource and consumer traits. For dimensions 3 and 4, we find negative σ , meaning that this situation is reversed. Within the leading two dimensions (Fig. 1a), the effects of the consumer trait ‘barbel length’ and the resource trait ‘pelagic’ are collinear, consistent with the understanding that long barbels are of use for detecting resources in or at the sediment, but less important for detecting resources in the water column. Consistent with ecological intuition is also the result that objects requiring much macro-reduction are larger and hence tend to be eaten by larger-bodied consumers (fork length). The TTM “automatically” identifies these relationships.

The relevance of subsequent niche-space coordinates in the four-dimensional trophic niche space is reported in Table 3. When the values of all coordinates are assigned at random to consumers and resources (cut-off dimension $D_c = 1$), the correlation coefficients between measured and predicted link strengths are distributed symmetrically around zero, as expected. With increasing cut-off dimension D_c , correlations increase. When only the values of the fourth coordinate are assigned at random ($D_c = 4$), the confidence interval barely excludes the value obtained without randomization, indicating that the information contained in the fourth coordinate only marginally contributes to improving predictions. This observation supports our decision not to consider niche-space dimensionality beyond $D = 4$.

As a simple model application, we computed the diets predicted by the model. Comparing these with the observed diets (Fig. 2), we found a high and highly significant correlation ($R^2=0.72$, $p<0.0001$, $n = 154$).

Table 2. Description of the four dimensions of niche space, including $\text{var } v_i^{(k)} + \text{var } f_j^{(k)}$ (as a measure of the weight of dimension k), the signature σ of the model, and the projections of all consumers and resources into trophic niche space.

dimension	1	2	3	4	
$\text{var } v_i^{(k)} + \text{var } f_j^{(k)}$	0.953	0.348	0.279	0.218	
σ	+1	+1	-1	-1	
consumers	phenotypic traits				
	fork length	2.154	-3.625	-2.006	2.625
	barbel	-6.063	-1.059	2.423	1.397
	offset	-11.606	3.397	6.191	-1.557
	Projections				
	<i>L. acutirostris</i>	0.691	0.192	-0.256	-0.215
	<i>L. brevicephalus</i>	-0.299	0.493	0.276	-0.357
	<i>L. crassibarbis</i>	-1.292	-0.034	0.571	0.148
	<i>L. dainellii</i>	0.075	-0.265	-0.109	0.200
	<i>L. gorgorensis</i>	-0.314	-0.954	-0.132	0.774
	<i>L. gorguari</i>	0.165	-0.016	-0.079	-0.003
	<i>L. longissimus</i>	1.596	-0.369	-0.823	0.138
	<i>L. macrophtalmus</i>	0.301	0.676	0.058	-0.556
	<i>L. megastoma</i>	0.766	-0.106	-0.375	0.011
	<i>L. nedgia</i>	-0.863	-0.126	0.352	0.179
	<i>L. platydorsus</i>	-0.460	0.029	0.215	0.021
	<i>L. surkis</i>	-0.049	-0.057	0.006	0.049
	<i>L. truttiformis</i>	0.378	0.547	-0.013	-0.462
	<i>L. tsanensis</i>	-0.695	-0.010	0.309	0.073
resources	phenotypic traits				
	pelagic	1.366	0.496	0.726	0.608
	macro-reduction	0.442	-0.173	0.302	-0.053
	offset	-1.777	0.175	-1.109	-0.178
	projections				
	phytoplankton	-0.411	0.670	-0.382	0.430
	sessile algae	-1.336	0.001	-0.807	-0.231
	macrophytes	-0.011	-0.519	0.097	-0.390
	seeds	-0.011	-0.519	0.097	-0.390
	detritus	-0.894	-0.172	-0.506	-0.284
	zooplankton	-0.411	0.670	-0.382	0.430
	macro-crustaceans	-0.452	-0.346	-0.204	-0.337
	benthic invertebrates	-1.336	0.001	-0.807	-0.231
	macro-insects	-0.452	-0.346	-0.204	-0.337
	mollusks	-0.452	-0.346	-0.204	-0.337
	Fish	0.914	0.150	0.522	0.271

Discussion

In this study we explored to what extent link strengths between consumers and resources can be explained by their phenotypic traits through a Trophic Trait Model (TTM). The results provide information on structure and dimensionality of trophic niche space. This study is solely based on previously published data (Sibbing and Nagelkerke 2001), which suggests that the kind of analysis done here is applicable to a wealth of other datasets that were assembled for different purposes. The input for the theoretical model consisted of mean values of 19 phenotypic traits for each consumer species, 11 physical traits for each resource, and a matrix of link strengths based on the presence of the resources in the guts of the consumers. The values for resource traits were semi-quantitative estimates, since no direct measurements were taken. This is likely to have limited the precision of our analysis and, as a result, the resolution of higher niche-space dimensions. Yet, the analysis identified at least two independent trophic traits of resources and two independent trophic traits of consumers that together determine the strengths of trophic links. These two trophic traits of consumers and resources collapse onto a pair of two-dimensional hyper-planes in the four geometric dimensions of trophic niche space. This becomes apparent when projecting niche space onto the plane spanned by dimensions 1 and 3 (Fig. 1b). Then, consumers and resources are concentrated along narrow bands (that is, we are seeing both planes at a flat angle). The question if dimensions 3 and 4 should be included in a measure of the “dimensionality of trophic niche space” therefore depends on the specific context. In the geometrical sense of the model given by Eq. (1) they do, but in a narrower topological sense they may not, because only two-dimensional hyper-planes of trophic niche space are occupied. However, with larger, more precise datasets that admit better model fits and more phenotypic/physical traits to be included, such a complete collapse onto two dimensions would be unlikely to occur.

The concrete identification of the phenotypic traits responsible for foraging and vulnerability to predation adds significant biological meaning to the concept of abstract, or latent, trophic traits (Rohr et al. 2010; Rossberg et al. 2010b; Yoshida 2003), because these traits are now well interpretable in terms of trophic biology. Large barbel length is known to enhance performance in the detection of small benthic resource organisms by cyprinid fishes, but not in the detection of pelagic resources (Kotrschal et al. 1991). On the other hand, large body size of the consumer enhances the capacity of eating large resource organisms, which tend to be pelagic (mostly fish) rather than benthic. The restriction of the analysis to phenotypic traits of known trophic relevance adds to the value of this approach. Recent studies by Eklöf et al. (2013) and Klecka and Boukal (2013) also demonstrated the explanatory power of such trophically relevant traits. The TTM makes no sharp association of specific explanatory variables with specific dimensions of trophic niche space (Fig. 1). Yet, strongly simplifying, one can say that the most important dimension of trophic niche space is – for this dataset – resource habitat and the corresponding foraging strategy, while the second dimension corresponds to consumer body size and the ability of larger consumers to chew larger (and mostly harder) resources. It must be stressed that the question which dimensions are “important”, that is, in which direction of trophic niche space consumers and resources are spread out broadly in comparison with the trophic niche width, depends on the particular set of consumers and resources included in the analysis. The *Labeobarbus* consumer species considered spanned only a factor 2.4 in body size. With a broader variety of consumers, the role of body size is likely to become more prominent. Woodward and Hildrew (2002), for example, presented a detrended correspondence analysis of the diets of freshwater invertebrates in which the effect of body size appears to be similar to that of habitat/behaviour. This analysis differentiated consumer species by ontogenetic stage, thus covering large ranges in body size.

Table 3. The correlation coefficients (ρ) between predicted and observed trophic link strengths with increasing number of predictive dimensions. Results show that all 4 dimensions of fitted trophic niche space contain ecologically relevant information; the fourth dimension, however, only marginally.

D_c	dimensions fixed at best fitting values	dimensions permuted randomly	correlation coefficients (ρ)			
			two-sided 95% confidence interval ^a		actually found in 4D model ^b	
1	-	1, 2, 3, 4	-0.295	- 0.292	0.723	
2	1	2, 3, 4	0.332	- 0.639	0.723	
3	1, 2	3, 4	0.430	- 0.686	0.723	
4	1, 2, 3	4	0.554	- 0.718	0.723	

^a 95% quantile under the null hypothesis that only fixed dimensions are relevant

^b If the value of the actually found correlation falls outside the 95% confidence interval this indicates that the extra dimension significantly add information

Besides this verification of the plausibility of the TTM on ecological grounds, cross-validation was used to ensure reproducibility of results. Fitting models with high-dimensional niche spaces is known to be difficult (the estimates of niche-space dimensionality by Eklöf et al. (2013), for example, suffer from an incorrect application of the Akaike information criterion (AIC) to models which include permutations as “parameters” (see appendix C of Allesina (2011)). In our work, the high correlation between predicted and measured link strengths found in cross-validation of the TTM using a four-dimensional niche space (Table 1, Fig. 2) and the observation that all four dimensions significantly contribute to it (Table 3) demonstrate that our approach has substantial predictive power, despite (a) the rather coarse measure of link strength used, (b) the simple, semi-quantitative classification of resource traits, (c) the simple model structure (Eq. 2) and (d) the fast and robust, but perhaps not optimal model-fitting algorithm. While improvements with regards to points (a) to (d) can be expected to enhance the predictive power of the TTM, the current form of handling these is viable already.

The mechanistic interpretability of the identified trophic traits together with our rigorous cross-validation suggest that the picture of trophic niche space we obtained is a reasonably good representation of reality. As explained, there are various ways by which this “imaging” technique could be improved to yield more accurate representations, but we would not expect these to differ from ours beyond recognition.

Among conceivable applications of the TTM is its use, similar to the model by Link (2004), to predict links in (changing) food webs. Specifically, predicted link strengths and the coordinates of new resources and consumers in the trophic niche space defined by our $D=4$ dimensional TTM can be computed from their phenotypic traits by the following procedure: (i) To obtain the abstract foraging traits for *Labeobarbus* species (and presumably other cyprinid fishes), measure their fork length FL and their anterior barbel length ABaL (both in cm). ABaL is measured as the maximal natural length of the barbel from its base to its tip when it is extended, but not stretched. (ii) Form the row vector $(\log_{10}(FL), \log_{10}(ABaL/FL), 1)$, and multiply this vector from the right with the matrix formed by the block under “Consumers - Phenotypic traits” in Table 3. (iii) To obtain vulnerability traits for resource objects, let PE=1 if they are pelagic and PE=0 otherwise, and estimate the degree of macro-reduction (MR) they require by a value in the range 0 to 4 on the scale used by Sibbing and Nagelkerke (Sibbing and Nagelkerke 2001). (iv) Form the row vector $(PE, MR, 1)$, and multiply this vector from the right with the matrix formed by the block under “Resources - Phenotypic traits” in Table 3.

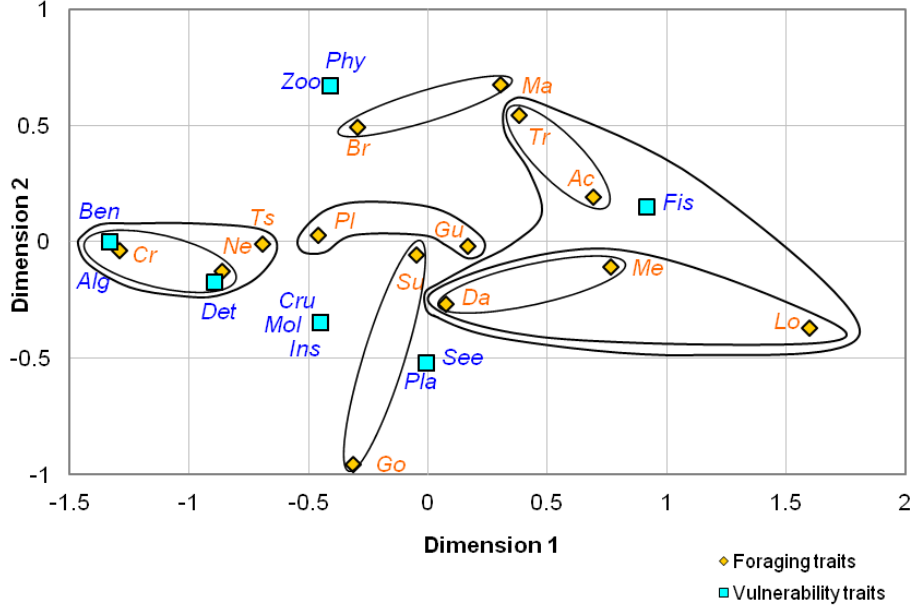


Figure 3. As Figure 1a, with ellipses and curved shapes indicating the hierarchical clusters of diets as found by Nagelkerke and Sibbing (2000) in their Fig. 8b, indicating that our model is consistent with these findings.

(v) The baseline traits V_i^* and F_j^* are always 0 for the $D=4$ model fit. (vi) Predicted link strengths are then given by Eq. (1).

Noteworthy is also that the clustering of consumers with resources found in the TTM is largely consistent with the study of Sibbing and Nagelkerke (2001) (Fig. 3). Similar to that study, the morphologically more ‘extreme’ species with the highest degree of specialization were most consistently associated with their particular resource categories. True piscivores, such as *L. longissimus*, *L. acutirostris*, and *L. megastoma*, clustered consistently with fish in both studies, while *L. nedgia* and *L. tsanensis* consistently clustered with the consumption of benthos and detritus. *L. gorgorensis* and *L. surkis* are identified as consumers of molluscs and plants (Figs 1, 3). The other species have a more generalized morphology or combine trait variables that could fit a wider range of resources, which mostly resulted in more diverse diets and therefore less clear-cut trophic relationships.

A geometric interpretation of the TTM arises from reading the sum $\sum_{k=1}^D \sigma_k (v_i^{(k)} - f_j^{(k)})^2$ in the exponent of Eq. (1) as the squared distance between the points $(v_i^{(1)}, \dots, v_i^{(D)})$ and $(f_j^{(1)}, \dots, f_j^{(D)})$ in trophic niche space. When all signs σ_k are equal to +1 (or all -1), this interpretation is obvious, because the Euclidean distance between $(v_i^{(1)}, \dots, v_i^{(D)})$ and $(f_j^{(1)}, \dots, f_j^{(D)})$ is defined as $\sqrt{\sum_{k=1}^D (v_i^{(k)} - f_j^{(k)})^2}$. Except for the trophic baseline traits V_i^* and

F_j^* , Eq. (1) then has the form of a Gaussian interaction kernel, a form that is widely used in theoretical ecology to model niches (MacArthur and Levins 1967; May and MacArthur 1972). The trophic niche of a consumer can be defined as the (hyper-) sphere, centred in trophic niche space at $(f_i^{(1)}, \dots, f_i^{(D)})$, beyond which trophic link strengths with resources fall below a given threshold. In the present case, this is the situation when only the two leading dimensions of the TTM are considered. To this extent, the historically assumed geometric picture (MacArthur and Levins 1967; May and MacArthur 1972) is here, in essence, empirically

confirmed. It implies that by adapting a consumer's foraging traits its link strength with a given resource species can be maximized.

However, for the full four-dimensional TTM we find a signature (+1, +1, -1, -1) containing both positive and negative signs σ_k . For dimensions with negative σ_k , link strengths increase with increasing separation between vulnerability and foraging traits in niche space. This means that resources can minimize link strength with particular consumer species by adapting corresponding vulnerability traits, i.e. optimize their defence against these species (Rossberg et al. 2010b). The geometric interpretation of the TTM can therefore be upheld only when interpreting the sum in Eq. (1) as a squared distance in a pseudo-Euclidean space (Sokolov 2002). Pseudo-Euclidean geometry differs from Euclidean geometry exactly by employing a distance measure with mixed sign structure.

For our TTM with signature (+1, +1, -1, -1), surfaces of constant link strength form hyperboloids in the 4-dimensional trophic niche space. In Fig. 1, for example, we show sections through a surface of constant link strength with the consumer *L. acutirostris*. The sections are parallel to the 1-2 plane (top panel) and the 1-3 plane (bottom panel) respectively and go through the point corresponding to the foraging traits of *L. acutirostris*. Resources with vulnerability traits located within this hyperboloid have a pseudo-Euclidean distance smaller than 1 from *L. acutirostris*, and therefore, by Eq. (1), a trophic link strength at most $e^{(-1^2/2)}=0.6$ times smaller than a_0 . In a certain sense, this hyperboloid therefore represents the foraging niche of *L. acutirostris*.

The geometric interpretation becomes crucial when comparing or combining our results with other TTMs obtained using different datasets. Different TTMs could have different coordinates that are equivalent in their biological meanings (Rossberg 2013). By appropriate coordinate transformations these TTMs could be directly compared. In a Euclidean niche space such transformations would be solid rotations, translations of the coordinate system, and inversions of the directions of axes. In pseudo-Euclidean geometries, rotations are complemented by other kinds of transformations (French 1968). In general, all those transformations need to be considered that keep link strengths as given by Eq. (1) invariant for all resource-consumer pairs. Correspondingly, characterizations of distributions of consumers and resources in trophic niche space are ecologically relevant only if they are invariant under these coordinate transformations. Rossberg (2013) discusses these questions in detail.

Conclusion

A central finding of this study is that models such as Eq. (1) can indeed be consistently fitted to data in such a way that trophic trait variables are given in terms of phenotypic traits. The model fit lends empirical support to the common preference of modellers for Gaussian interaction kernels with maximal link strength at the centre, at least as a first approximation. However, we showed that, upon closer inspection, the Euclidean geometry of trophic niche space implied by Gaussian kernels gives way to a pseudo-Euclidean geometry as discussed above. Within this pseudo-Euclidian space, the model fit assigns relative positions to the consumers and resources that we studied in such a way that the pseudo-Euclidean distances between them become predictors of (logarithmic) trophic link strength. The relative positions of consumers and resources thus obtained are independent of the specific dataset studied. If sufficient information on other consumers and resources became available, either in terms of their trophic traits or their trophic interaction partners, these could be placed in niche space alongside those studied here. Thus, the relative positions of consumers and resources that we find directly represent aspects of ecological reality that are not contingent on study design (so differentiating this method from other ordination techniques). This is why our analysis can be

understood as providing a direct image of trophic niche space, a concept that has remained largely hypothetical so far. We demonstrated the reality of trophic niche space by taking a picture of it.

While the picture of trophic niche space that we obtained remains diffuse, the positions of 14 cyprinid fishes and their resources are recognizable and consistent with previous analyses (Fig. 3). The picture resolves more than one trophic dimension. This implies not only that several phenotypic traits notably contribute to determining trophic links, but also that these contributions are independent rather than jointly determining a single “niche value”. The fact that with a four-dimensional niche space about $(0.723)^2 \approx 52\%$ of the variance in link strengths could be explained constrains the scope for contributions from even higher niche-space dimensions in this system. This suggests the preliminary conclusion that, whilst trophic niche space is multi-dimensional, a rather small number of dimensions will generally be sufficient to predict trophic links with reasonable accuracy—thus supporting recent analyses of food-web structures based on similar premises (Allesina et al. 2008; Petchey et al. 2008; Rohr et al. 2010; Stouffer et al. 2011; Williams and Purves 2011). Apart from its theoretical significance, this potentially low number of trophic traits encourages efforts to predict trophic links from phenotypic traits also in systems different from the one studied here. It also supports the idea that, in order to preserve functional diversity within a community, it suffices to preserve diversity among certain important traits. Our approach might even be useful for selecting appropriate sets of phenotypic traits to be included in indices of functional diversity for management applications.

Acknowledgements

A.G.R. was supported by the U.K. Department of Environment, Food and Rural Affairs (M1228), the European Commission (agreement no. 308392, DEVOTES) and a Beaufort Marine Research Award carried out under the Irish Sea Change Strategy and the Strategy for Science Technology and Innovation (2006 – 2013).

References

- Albrecht GH, Gelvin BR, Hartman SE (1993) Ratios as a size adjustment in morphometrics. *Am J Phys Anthropol* 91 (4):441-468. doi:10.1002/ajpa.1330910404
- Allesina S (2011) Predicting trophic relations in ecological networks: A test of the Allometric Diet Breadth Model. *Journal of Theoretical Biology* 279 (1):161-168. doi:10.1016/j.jtbi.2010.06.040
- Allesina S, Alonso D, Pascual M (2008) A general model for food web structure. *Science* 320 (5876):658-661. doi:10.1126/science.1156269
- Arditi R, Michalski J, Hirzel AH (2005) Rheagogies: Modelling non-trophic effects in food webs. *Ecological Complexity* 2 (3):249-258. doi:http://dx.doi.org/10.1016/j.ecocom.2005.04.003
- Barnett A, Bellwood DR, Hoey AS (2006) Trophic ecomorphology of cardinalfish. *Marine Ecology Progress Series* 322:249-257. doi:10.3354/meps322249
- Berlow EL, Dunne JA, Martinez ND, Stark PB, Williams RJ, Brose U (2009) Simple prediction of interaction strengths in complex food webs. *Proceedings of the National Academy of Sciences* 106 (1):187-191. doi:10.1073/pnas.0806823106
- Bersier L-F (2007) A history of the study of ecological networks. In: Képès F (ed) *Biological networks*. World Scientific, New Jersey, pp 365-421. doi:10.1016/j.ecocom.2007.06.013
- Bhat A (2005) Ecomorphological correlates in tropical stream fishes of southern India. *Environmental Biology of Fishes* 73 (2):211-225. doi:10.1007/s10641-005-0561-0

- Carlson RL, Wainwright PC (2010) The ecological morphology of darter fishes (Percidae: Etheostomatinae). *Biological Journal of the Linnean Society* 100 (1):30-45. doi:10.1111/j.1095-8312.2010.01417.x
- Cohen JE (1977) Food webs and the dimensionality of trophic niche space. *Proceedings of the National Academy of Sciences* 74 (10):4533-4536
- de Graaf M, Megens H-J, Samallo J, Sibbing F (2010) Preliminary insight into the age and origin of the Labeobarbus fish species flock from Lake Tana (Ethiopia) using the mtDNA cytochrome b gene. *Molecular Phylogenetics and Evolution* 54 (2):336-343. doi:10.1016/j.ympev.2009.10.029
- de Ruiter PC, Neutel A-M, Moore JC (1995) Energetics, patterns of interaction strengths and stability in real ecosystems. *Science* 269:1257-1260. doi:10.1126/science.269.5228.1257
- Dejen E, Vijverberg J, de Graaf M, Sibbing FA (2006) Predicting and testing resource partitioning in a tropical fish assemblage of zooplanktivorous 'barbs': an ecomorphological approach. *Journal of Fish Biology* 69 (5):1356-1378. doi:10.1111/j.1095-8649.2006.01197.x
- Eklöf A, Jacob U, Kopp J, Bosch J, Castro-Urgal R, Chacoff NP, Dalsgaard B, de Sassi C, Galetti M, Guimarães PR, Lomáscolo SB, Martín González AM, Pizo MA, Rader R, Rodrigo A, Tylianakis JM, Vázquez DP, Allesina S (2013) The dimensionality of ecological networks. *Ecol Lett* 16 (5):577-583. doi:10.1111/ele.12081
- Emmerson MC, Montoya JM, Woodward G (2005) Body size, interaction strength, and food web dynamics. In: de Ruiter PC, Wolters V, Moore JC (eds) *Dynamics food webs, multiple species assemblage, ecosystem development and environmental change*. Academic Press, pp 167-178
- Fath BD, Scharler UM, Ulanowicz RE, Hannon B (2007) Ecological network analysis: network construction. *Ecological Modelling* 208 (1):49-55. doi:10.1016/j.ecolmodel.2007.04.029
- French A (1968) *Special Relativity*. W.W. Norton & Company, New York
- Klecka J, Boukal DS (2013) Foraging and vulnerability traits modify predator-prey body mass allometry: freshwater macroinvertebrates as a case study. *Journal of Animal Ecology* 82 (5):1031-1041. doi:10.1111/1365-2656.12078
- Kotrschal K, Brandstätter R, Gomahr A, Junger H, Palzenberger M, Zaunreiter M (1991) Brain and sensory systems. In: Winfield IJ, Nelson JS (eds) *Cyprinid fishes: systematics, biology and exploitation*. Chapman and Hall, London, pp 284-331
- Layman CA, Langerhans RB, Winemiller KO (2005) Body size, not other morphological traits, characterizes cascading effects in fish assemblage composition following commercial netting. *Canadian Journal of Fisheries and Aquatic Sciences* 62 (12):2802-2810. doi:10.1139/F05-183
- Link JS (2004) A General Model of Selectivity for Fish Feeding: A Rank Proportion Algorithm (*Transactions of the American Fisheries Society*). *Transactions of the American Fisheries Society* 133 (3):655-673. doi:10.1577/t02-142.1
- MacArthur R, Levins R (1967) The limiting similarity, convergence, and divergence of coexisting species. *The American Naturalist* 101 (921):377-385
- May RM, MacArthur RH (1972) Niche overlap as a function of environmental variability. *Proceedings of the National Academy of Sciences* 69 (5):1109-1113
- Nagelkerke LAJ, Sibbing FA (2000) The large barbs (*Barbus* spp., Cyprinidae, teleostei) of Lake Tana (Ethiopia), with a description of a new species, *Barbus osseensis*. *Netherlands Journal of Zoology* 50 (2):179-214. doi:10.1163/156854200505946
- Nagelkerke LAJ, Sibbing FA, van den Boogaart JGM, Lammens EHRR, Osse JWM (1994) The barbs (*Barbus* spp.) of Lake Tana: a forgotten species flock? *Environmental Biology of Fishes* 39 (1):1-22. doi:10.1007/BF00004751

- Naisbit RE, Rohr RP, Rossberg AG, Kehrlı P, Bersier L-F (2012) Phylogeny versus body size as determinants of food web structure. *Proceedings of the Royal Society B: Biological Sciences*. doi:10.1098/rspb.2012.0327
- Petchey OL, Beckerman AP, Riede JO, Warren PH (2008) Size, foraging, and food web structure. *Proceedings of the National Academy of Sciences* 105 (11):4191-4196. doi:10.1073/pnas.0710672105
- Piegorsch WW, Bailer AJ (2005) *Analyzing environmental data*. John Wiley & Sons, Ltd, Chichester, UK
- Pouilly M, Lino F, Bretenoux JG, Rosales C (2003) Dietary-morphological relationships in a fish assemblage of the Bolivian Amazonian floodplain. *Journal of Fish Biology* 62 (5):1137-1158. doi:10.1046/j.1095-8649.2003.00108.x
- Quinn GP, Keough MJ (2002) *Experimental design and data analysis for biologists*. Cambridge University Press, Cambridge, UK
- R Development Core Team (2010) *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria
- Reunanen J (2003) Overfitting in making comparisons between variable selection methods. *J Mach Learn Res* 3:1371-1382
- Rohr RP, Scherer H, Kehrlı P, Mazza C, Bersier L-F (2010) Modeling food webs: exploring unexplained structure using latent traits. *The American Naturalist* 176 (2):170-177. doi:10.1086/653667
- Rossberg AG (2013) *Food webs and biodiversity: foundations, models, data*. 1 edn. John Wiley & Sons, Ltd, Chichester, UK. doi:10.1002/9781118502181
- Rossberg AG, Brännström Å, Dieckmann U (2010a) Food-web structure in low- and high-dimensional trophic niche spaces. *Journal of the Royal Society Interface*. doi:10.1098/rsif.2010.0111
- Rossberg AG, Brännström Å, Dieckmann U (2010b) How trophic interaction strength depends on traits. *Theoretical Ecology* 3 (1):13-24. doi:10.1007/s12080-009-0049-1
- Rossberg AG, Farnsworth KD, Satoh K, Pinnegar JK (2011) Universal power-law diet partitioning by marine fish and squid with surprising stability–diversity implications. *Proceedings of the Royal Society B: Biological Sciences* 278 (1712):1617-1625. doi:10.1098/rspb.2010.1483
- Rossberg AG, Ishii R, Amemiya T, Itoh K (2008) The top-down mechanism for body-mass - abundance scaling (Ecology). *Ecology* 89 (2):567-580. doi:10.1890/07-0124.1
- Rossberg AG, Matsuda H, Amemiya T, Itoh K (2006) Food webs: Experts consuming families of experts. *Journal of Theoretical Biology* 241 (3):552-563. doi:10.1016/j.jtbi.2005.12.021
- Russo T, Pulcini D, O'Leary Á, Cataudella S, Mariani S (2008) Relationship between body shape and trophic niche segregation in two closely related sympatric fishes. *Journal of Fish Biology* 73 (4):809-828. doi:10.1111/j.1095-8649.2008.01964.x
- Schmitz OJ, Price JR (2011) Convergence of trophic interaction strengths in grassland food webs through metabolic scaling of herbivore biomass. *Journal of Animal Ecology* 80 (6):1330-1336. doi:10.1111/j.1365-2656.2011.01882.x
- Sibbing FA (1991a) Food capture and oral processing. In: Winfield IJ, Nelson JS (eds) *Cyprinid fishes: systematics, biology and exploitation*, vol 3. Fish and Fisheries Series. Chapman & Hall, London, pp 377-412
- Sibbing FA (1991b) Food processing by mastication in cyprinid fish. In: Vincent JFV, Lillford PJ (eds) *Feeding and the texture of food*, vol 44. SEB Seminar Series. Cambridge University Press, Cambridge, pp 57-92. doi:10.1017/CBO9780511600555.005

- Sibbing FA, Nagelkerke LAJ (2001) Resource partitioning by Lake Tana barbs predicted from fish morphometrics and prey characteristics. *Reviews in Fish Biology and Fisheries* 10 (4):393-437. doi:10.1023/A:1012270422092
- Sibbing FA, Nagelkerke LAJ, Osse JWM (1994) Ecomorphology as a Tool in Fisheries - Identification and Ecotyping of Lake Tana Barbs (*Barbus-Intermedius* Complex), Ethiopia. *Neth J Agric Sci* 42 (1):77-85
- Sokolov DD (2002) Pseudo-Euclidean space. In: Hazewinkel M (ed) *Encyclopaedia of Mathematics*. Kluwer Academic Publishers, Dordrecht, The Netherlands,
- Spooner D, Vaughn C (2008) A trait-based approach to species' roles in stream ecosystems: climate change, community structure, and material cycling. *Oecologia* 158 (2):307-317. doi:10.1007/s00442-008-1132-9
- Stouffer DB, Rezende EL, Amaral LAN (2011) The role of body mass in diet contiguity and food-web structure. *Journal of Animal Ecology* 80 (3):632-639. doi:10.1111/j.1365-2656.2011.01812.x
- Vucic-Pestic O, Rall BC, Kalinkat G, Brose U (2010) Allometric functional response model: body masses constrain interaction strengths. *Journal of Animal Ecology* 79 (1):249-256. doi:10.1111/j.1365-2656.2009.01622.x
- Wainwright PC, Richard BA (1995) Predicting patterns of prey use from morphology of fishes. *Environmental Biology of Fishes* 44:97-113. doi:10.1007/BF00005909
- Williams R, Purves D (2011) The probabilistic niche model reveals substantial variation in the niche structure of empirical food webs (*Ecology*). *Ecology*:in press. doi:10.1890/11-0200.1
- Woodward G, Hildrew AG (2002) Body-Size Determinants of Niche Overlap and Intraguild Predation within a Complex Food Web. *Journal of Animal Ecology* 71 (6):1063-1074. doi:10.1046/j.1365-2656.2002.00669.x
- Yoshida K (2003) Dynamics of evolutionary patterns of clades in a food web system model. *Ecological Research* 18:625-637. doi:10.1111/j.1440-1703.2003.00585.x

Online resource 1

Electronic supplementary material to:

Journal: Theoretical Ecology

Title: Trophic niche-space imaging, using resource and consumer traits

Authors: Leopold A. J. Nagelkerke ^{1,*} and Axel G. Rossberg ^{2,3}

¹*Aquaculture and Fisheries Group, Wageningen University, Wageningen Institute of Animal Sciences (WIAS), De Elst 1, 6708 WD Wageningen, The Netherlands, e-mail:*

Leo.Nagelkerke@wur.nl, Tel: +31 317 483940, fax: +31 317 483937

²*Lowestoft Laboratory, Centre for Environment, Fisheries and Aquaculture Science (Cefas), Pakefield Road, Lowestoft, Suffolk NR33 0HT, United Kingdom*

³*School of Biological Sciences, Queen's University Belfast, Belfast BT9 7BL, Northern Ireland, United Kingdom, e-mail: axel@rossberg.net, Tel: +44 28 9097 5859, fax: +44 28 9097 5877*

* Author for correspondence

Table S1. Consumer species and resource categories. The 14 consumer species (*Labeobarbus* fish species from Lake Tana, Ethiopia), and the 11 resource categories (Sibbing and Nagelkerke 2001). Abbreviations that are used in figures and text are also given.

consumers		resources	
species name	abbreviation	category name	abbreviation
<i>L. acutirostris</i>	Ac	phytoplankton	Phy
<i>L. brevicephalus</i>	Br	sessile algae	Alg
<i>L. crassibarbis</i>	Cr	macrophytes	Pla
<i>L. dainellii</i>	Da	seeds	See
<i>L. gorgorensis</i>	Go	detritus	Det
<i>L. gorguari</i>	Gu	zooplankton	Zoo
<i>L. longissimus</i>	Lo	macro-crustaceans	Cru
<i>L. macrophthalmus</i>	Ma	benthic invertebrates	Ben
<i>L. megastoma</i>	Me	macro-insects	Ins
<i>L. nedgia</i>	Ne	mollusks	Mol
<i>L. platydorsus</i>	Pl	fish	Fis
<i>L. surkis</i>	Su		
<i>L. truttiformis</i>	Tr		
<i>L. tsanensis</i>	Ts		

Table S2. Consumer traits. Data adapted from Nagelkerke (1997). Traits as defined as in Nagelkerke (1997) and Sibbing and Nagelkerke (2001). All lengths are measured in units of fork length (FL), areas in units of fork length squared, etc. Fork length itself is measured in cm. Reference numbers refer to Figure S1, in some cases expressing how variables were calculated. For instance, body depth / width ratio is a division of measurement 4 and 5 and indicated as (4 / 5).

Trait description	General size	Search and detection		Approach				
	Fork length	Anterior barbel length	Eye diameter	Body depth	Body depth / width ratio	Oral gape / body area ratio	Caudal peduncle depth	Anal fin area
Abbreviation ^a	FL	ABaL	ED	BD	BD/BW	OGAr/Bar	CPD	AfiAr
Reference number in figure S1	–	–	7	4	(4 / 5)	(17 x 18) / (4 x 5)	6	(2 x 3)
<i>Labeobarbus</i> species								
<i>L. acutirostris</i>	21.8	0.028	0.048	0.219	1.99	0.272	0.085	1.07
<i>L. brevicephalus</i>	16.6	0.037	0.049	0.238	2.14	0.111	0.099	1.21
<i>L. crassibarbis</i>	20.3	0.058	0.041	0.243	1.93	0.218	0.103	1.24
<i>L. dainellii</i>	26.6	0.038	0.039	0.215	1.87	0.358	0.088	1.19
<i>L. gorgorensis</i>	38.1	0.050	0.036	0.269	1.99	0.115	0.102	1.32
<i>L. gorguari</i>	23.3	0.035	0.043	0.224	1.71	0.282	0.091	1.19
<i>L. longissimus</i>	32.9	0.023	0.035	0.221	1.82	0.241	0.098	1.12
<i>L. macropthalmus</i>	15.8	0.029	0.059	0.239	2.00	0.209	0.095	1.28
<i>L. megastoma</i>	26.0	0.029	0.041	0.215	1.90	0.285	0.090	1.13
<i>L. nedgia</i>	22.4	0.051	0.042	0.236	1.91	0.112	0.100	1.25
<i>L. platydorsus</i>	21.3	0.043	0.044	0.243	1.88	0.231	0.093	1.24
<i>L. surkis</i>	23.3	0.038	0.045	0.270	2.14	0.083	0.098	1.23
<i>L. truttiformis</i>	17.2	0.029	0.037	0.242	1.83	0.248	0.106	1.30
<i>L. tsanensis</i>	21.3	0.047	0.044	0.253	1.99	0.145	0.102	1.19

Trait description	Intake					Size selection			Digestion		
	Oral gape axis	Protrusion length	Lower jaw length	Pharyngo-opercular volume ^b	Head length	Post-orbital length	Gill arch resistance	Oral gape diameter	Gill raker length	Gill raker profile	Gut length
Abbreviation ^a	OGAx	ProtL	LJL	–	HL	POrL / OpD	GiAR	OGD	GiRL	GiRP	GuL
Reference number in figure S1	16	26	13	–	11	(25 / 19)	–	18	9	10	–
<i>Labeobarbus</i> species											
<i>L. acutirostris</i>											
<i>L. brevicephalus</i>	65.8	0.041	0.109	0.487	0.271	1.657	1.09	0.082	0.0061	1.89	1.698
<i>L. crassibarbis</i>	50.6	0.040	0.077	0.354	0.206	1.149	1.61	0.060	0.0068	4.75	2.054
<i>L. dainellii</i>	46.9	0.066	0.108	0.710	0.250	1.237	1.00	0.076	0.0054	1.20	2.273
<i>L. gorgorensis</i>	46.2	0.051	0.110	0.576	0.274	1.490	0.98	0.090	0.0053	1.50	1.516
<i>L. gorguari</i>	52.7	0.047	0.082	0.605	0.222	1.201	1.29	0.065	0.0060	2.67	3.399
<i>L. longissimus</i>	63.5	0.047	0.107	0.774	0.267	1.449	1.22	0.090	0.0070	1.40	1.655
<i>L. macropthalmus</i>	77.6	0.038	0.109	0.588	0.250	1.431	1.01	0.086	0.0056	1.22	1.708
<i>L. megastoma</i>	68.0	0.048	0.104	0.569	0.251	1.260	1.39	0.084	0.0075	3.25	1.821
<i>L. nedgia</i>	79.6	0.039	0.110	0.488	0.248	1.449	1.06	0.083	0.0056	2.46	1.908
<i>L. platydorsus</i>	47.5	0.051	0.092	0.595	0.247	1.177	1.43	0.073	0.0066	3.10	2.412
<i>L. surkis</i>	65.8	0.040	0.100	0.710	0.253	1.333	1.19	0.086	0.0066	2.30	1.831
<i>L. truttiformis</i>	58.7	0.037	0.075	0.413	0.205	1.114	1.64	0.057	0.0065	3.93	2.898
<i>L. tsanensis</i>	67.1	0.031	0.095	0.631	0.238	1.409	1.13	0.084	0.0071	2.50	1.916
	49.6	0.051	0.090	0.619	0.237	1.230	1.45	0.071	0.0070	3.57	2.277

^a in Sibbing and Nagelkerke (2001): Table 4

^b defined in Nagelkerke (1997)

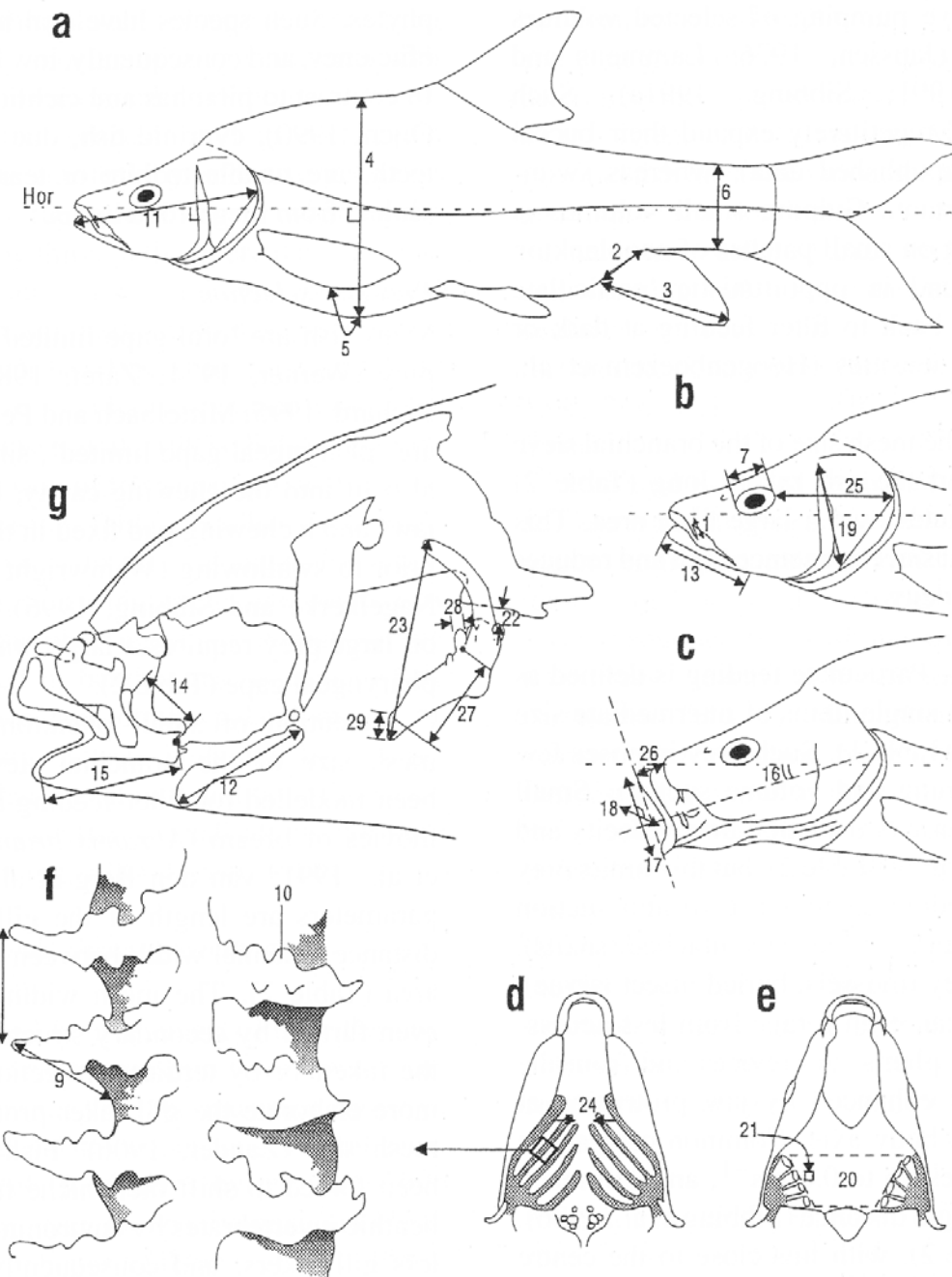


Figure S1. Measurements of phenotypic traits in the Lake Tana *Labeobarbus* species: (a) external traits (Hor: Horizontal line as reference for angular measurements); (b) head with the mouth closed; (c) head with open, protruded mouth; (d) oro-pharyngeal floor; (e) oro-pharyngeal roof; (f) detail of gill arch showing a raker profile (10), coded as value 3 (on a scale from 1 to 5); (g) elements of the head skeleton (not used in this study). Numbers of used traits can be found in Table S2. Figure reproduced with permission from Sibbing and Nagelkerke (2001).

Table S3. Resource traits. Data adapted from Nagelkerke (1997). Data are ordinal and adaptations from the original tables from Sibbing and Nagelkerke (2001).

Resource	Maximum diameter	Escape velocity	Shape (elongation)	Habitat (pelagic)	Chemical composition (protein)	Strength	Compliance	Fibrousness	Toughness	Requiring macro-reduction	Requiring micro-reduction
Phytoplankton	1.0	0.0	0	1	0	0	1.5	0	0	0	3
Sessile algae	2.0	0.0	0	0	0	0	1.5	0	0	1	2
Macrophytes	3.5	0.0	1	0	0	2	2	2	2	4	3
Seeds	3.5	0.0	0	0	0	3	1	1	2	4	1
Detritus	2.0	0.0	0	0	0	0	2	1.5	1	2	1
Zooplankton	1.5	1.5	0	1	1	0	2	1	1	0	1
Macro-crustaceans	4.5	3.5	0	0	1	0	1.5	1	1	3	0
Benthic invertebrates (larvae / worms)	2.0	1.0	1	0	1	0	3	1.5	2	1	1
Macro-insects	3.5	4.0	1	0	1	2	2.5	2	3	3	0
Molluscs	3.5	1.0	0	0	1	3	0	1	2	3	0
Fish	5.0	5.0	1	1	1	2	4	3	5	3	0

Table S4. Consumer diets. Data adapted from Nagelkerke (1997). Volume-% of food categories in the gut. Data per species do not add up to 100% due to a rest category of unidentified particles in the gut.

<i>Labeobarbus</i> species	Phytoplankton	Sessile algae	Macrophytes	Seeds	Detritus	Zooplankton	Macro- crustaceans	Benthic invertebrates (larvae/worms)	Macro-insects	Molluscs	Fish
<i>L. acutirostris</i>	3.2	0.1	2.0	0.0	2.6	0.6	0.0	7.9	3.5	0.7	76.0
<i>L. brevicephalus</i>	7.4	0.2	14.6	0.0	3.9	34.0	0.0	9.4	23.2	5.2	0.2
<i>L. crassibarbis</i>	10.4	0.3	0.7	0.3	28.2	8.8	0.0	28.8	7.3	11.8	2.1
<i>L. dainellii</i>	2.1	0.0	8.5	0.0	1.0	0.0	2.4	5.7	7.3	1.6	69.2
<i>L. gorgorensis</i>	8.5	0.1	31.6	0.1	14.6	0.0	3.3	1.6	0.0	27.0	11.5
<i>L. gorguari</i>	6.9	0.2	15.8	0.7	3.4	1.4	2.4	6.5	1.2	3.6	55.4
<i>L. longissimus</i>	3.8	0.1	12.9	0.7	2.9	0.0	2.6	2.5	0.1	0.8	71.0
<i>L. macrophthalmus</i>	9.6	1.8	8.4	0.3	3.2	6.5	0.5	13.0	10.4	0.3	45.3
<i>L. megastoma</i>	1.3	0.1	13.9	0.0	0.7	1.3	0.3	6.8	6.9	0.3	66.8
<i>L. nedgia</i>	5.0	0.3	4.7	0.3	18.9	3.2	1.1	32.5	10.8	16.3	5.6
<i>L. platydorsus</i>	9.4	0.0	10.3	0.0	3.2	1.2	0.1	17.6	3.6	4.0	49.0
<i>L. surkis</i>	7.9	0.3	70.9	0.6	2.8	5.2	0.2	4.9	4.3	0.4	2.1
<i>L. truttiformis</i>	4.2	0.0	3.3	0.0	3.7	2.8	5.1	1.3	0.0	1.6	72.3
<i>L. tsanensis</i>	6.3	0.2	4.6	0.8	14.2	6.2	0.0	42.6	4.9	17.7	1.3

References

Nagelkerke LAJ (1997) The barbs of Lake Tana, Ethiopia: morphological diversity and its implications for taxonomy, trophic resource partitioning, and fisheries. Wageningen Agricultural University, Experimental Animal Morphology and Cell Biology, Wageningen, The Netherlands

Sibbing FA, Nagelkerke LAJ (2001) Resource partitioning by Lake Tana barbs predicted from fish morphometrics and prey characteristics. *Reviews in Fish Biology and Fisheries* 10 (4):393-437.
doi:10.1023/A:1012270422092

Online resource 2

Electronic supplementary material to:

Journal: Theoretical Ecology

Title: Trophic niche-space imaging, using resource and consumer traits

Authors: Leopold A. J. Nagelkerke ^{1,*} and Axel G. Rossberg ^{2,3}

¹*Aquaculture and Fisheries Group, Wageningen University, Wageningen Institute of Animal Sciences (WIAS), De Elst 1, 6708 WD Wageningen, The Netherlands, e-mail:*

Leo.Nagelkerke@wur.nl, Tel: +31 317 483940, fax: +31 317 483937

²*Lowestoft Laboratory, Centre for Environment, Fisheries and Aquaculture Science (Cefas), Pakefield Road, Lowestoft, Suffolk NR33 0HT, United Kingdom*

³*School of Biological Sciences, Queen's University Belfast, Belfast BT9 7BL, Northern Ireland, United Kingdom, e-mail: axel@rossberg.net, Tel: +44 28 9097 5859, fax: +44 28 9097 5877*

* Author for correspondence

Online resource 2

Electronic supplementary material to:

Journal: Theoretical Ecology

Title: Trophic niche-space imaging, using resource and consumer traits

Authors: Leopold A. J. Nagelkerke and Axel G. Rossberg

A Details of data analysis

Our procedure to obtain the trophic trait vectors of consumers and resources from field data has three main aspects: (1) parametrization of the trophic trait model (TTM) based on observed feeding behaviour for a given set of empirical trait variables, (2) cross-validation of the model to assess its predictive power, and (3) selection an appropriate set of trait variables such as to achieve good predictive power. Variable selection (3) requires repeated cross validation (2), and cross-validation repeated model fitting (1).

A.1 Parametrization of the TTM

Input trait variables The procedure for parametrizing the TTM starts from two given sets of phenotypic traits, one for resources and one for consumers. The values of the resource trait variables are known for each of R kinds of resource, here diet categories, and the values of the consumer trait variables are known for each of Q kinds of consumers, here labeobarb species. For each resource category, the values of n trait variables are given, and for each labeobarb species the values of m trait variables. The trait values for resource i are given by an n -component vector of real values $\tilde{\mathbf{t}}_i$, that is $\tilde{\mathbf{t}}_i$ and trait values for consumer j by an m -component vector $\tilde{\mathbf{s}}_j$. The trait values of labeobarbs were log-transformed. By taking logarithms, the problem of how to best account for dependencies of trait variables on body size is mitigated, because different normalizations in terms of body size correspond to different linear combinations of logarithmic traits. The best-fitting abstract foraging traits, which are linear combinations of logarithmic phenotypic traits, will automatically combine these traits in such a way as to achieve the most appropriate normalization.

Trait standardization The input trait values were now standardized, by transforming them into equivalent sets of trait values that were statistically uncorrelated and had mean zero and variance one. This was achieved as follows: from the vectors of phenotypic traits, we first computed the means over the sets of consumers and resources considered, i.e. $\bar{\mathbf{t}} = R^{-1} \sum_{i=1}^R \tilde{\mathbf{t}}_i$ and $\bar{\mathbf{s}} = Q^{-1} \sum_{j=1}^Q \tilde{\mathbf{s}}_j$, and then formed the raw resource trait matrix defined by $\tilde{\mathbf{T}} = (\tilde{\mathbf{t}}_1 - \bar{\mathbf{t}}, \dots, \tilde{\mathbf{t}}_R - \bar{\mathbf{t}})$ and raw consumer trait matrix defined by $\tilde{\mathbf{S}} = (\tilde{\mathbf{s}}_1 - \bar{\mathbf{s}}, \dots, \tilde{\mathbf{s}}_Q - \bar{\mathbf{s}})$. Then we sought matrices \mathbf{A} and \mathbf{B} such that $(R - 1)^{-1} \mathbf{A} \tilde{\mathbf{T}} \tilde{\mathbf{T}}^T \mathbf{A}^T = \mathbf{I}$, and $(Q - 1)^{-1} \mathbf{B} \tilde{\mathbf{S}} \tilde{\mathbf{S}}^T \mathbf{B}^T = \mathbf{I}$, and defined new, scaled trait matrices as $\mathbf{T} = \mathbf{A} \tilde{\mathbf{T}}$, $\mathbf{S} = \mathbf{B} \tilde{\mathbf{S}}$.¹ The columns of $\mathbf{T} = (\mathbf{t}_1, \dots, \mathbf{t}_R)$ and $\mathbf{S} = (\mathbf{s}_1, \dots, \mathbf{s}_Q)$ are equivalent representations of the original phenotypic traits. The original trait values can be recovered by making use of \mathbf{A} , \mathbf{B} , $\bar{\mathbf{s}}$, and $\bar{\mathbf{t}}$.

¹To find an appropriate transformations \mathbf{A} , we computed the eigendecomposition of the covariance matrix $(R - 1)^{-1} \tilde{\mathbf{T}} \tilde{\mathbf{T}}^T = \mathbf{V} \mathbf{D} \mathbf{V}^T$, with the diagonal matrix \mathbf{D} containing the eigenvalues and the columns of \mathbf{V} the eigenvectors, and set $\mathbf{A} = \mathbf{D}^{-1/2} \mathbf{V}^T$. We proceeded analogously for consumer traits to obtain \mathbf{B} .

One might be concerned that this standardization procedure enhances measurement errors for trait variables that show little variability over the data set considered or for combinations of trait variables that are strongly correlated. However, selections of trait variables that include such cases will be identified by the cross-validation procedure that we apply to each candidate combination, and the variable-selection procedure will reject them.

Relation between link-strength function stomach-content data There are various ways in which trophic link strength can be defined. Here we use a particularly simple definition: We interpret $a(\mathbf{t}, \mathbf{s})\Delta t$ as the probability that a consumer individual with traits \mathbf{s} , upon encountering a resource item with traits \mathbf{t} , will feed on it within a small time interval of length Δt . Mathematically, this is the conditional probability $a(\mathbf{t}, \mathbf{s})\Delta t = P[\mathbf{s} \text{ eats } \mathbf{t} \text{ within } \Delta t | \mathbf{s} \text{ meets } \mathbf{t}]$. By basic probability theory and the assumption that encounters occur at random, a feeding interaction will then occur with probability

$$\begin{aligned} P[\mathbf{s} \text{ meets and eats } \mathbf{t} \text{ within } \Delta t] &= a(\mathbf{t}, \mathbf{s})P[\mathbf{s} \text{ meets } \mathbf{t}]\Delta t \\ &\propto a(\mathbf{t}, \mathbf{s}) \times (\text{density of } \mathbf{s}) \times (\text{density of } \mathbf{t}). \end{aligned} \tag{1}$$

The precise meaning of Eq. (1) depends on the statistical ensemble considered. Here, we choose the population of one labeobarb species at random with uniform probability and randomly pick one of its individuals. Within this ensemble, the density of labeobarbs with traits \mathbf{s} is a fixed constant when a species with traits \mathbf{s} exists, and zero otherwise.² Resources we consider sampled at random from the volume of the lake. The probability of encounter with prey that has traits \mathbf{t} is proportional to the volumetric abundance of prey with traits \mathbf{t} in the lake if such prey exists, and zero otherwise. If, upon encounter, a labeobarb individual eats the prey item, this ends up in its gut. The probability on the left hand side of Eq.(1) is therefore approximately proportional to the volumetric proportion in the gut content: $P[\mathbf{t} \text{ meets and eats } \mathbf{s} \text{ within } \Delta t] \propto (\text{proportion of } \mathbf{s} \text{ in gut of } \mathbf{t})$ if labeobarbs with traits \mathbf{t} exists, and zero otherwise.

Estimation of the link-strength function To estimate $a(\mathbf{t}, \mathbf{s})$ by Eq. (1), we approximated the three densities entering Eq. (1) by normal distributions. To determine the parameters of these normal distributions, we computed the mean $\boldsymbol{\mu}_{\mathbf{t}}$ and covariance matrix $\mathbf{C}_{\mathbf{t}}$ of occurrences of resource traits \mathbf{t} in the lake, the mean $\boldsymbol{\mu}_{\mathbf{s}}$ and covariance matrix $\mathbf{C}_{\mathbf{s}}$ of consumer traits \mathbf{s} over all labeobarb species with equal weights, and the mean $\boldsymbol{\mu}_{\mathbf{f}}$ and covariance matrix $\mathbf{C}_{\mathbf{f}}$ of pairs (\mathbf{t}, \mathbf{s}) in feeding interactions. Since, unfortunately, reliable direct measurements of the densities of resource items are not available, we use the mean proportion of prey items in labeobarb gut content as a rough surrogate, as discussed in the main text. With f_{rc} denoting the proportion of item r in the gut of labeobarb species c (*diet proportion*) this gives estimated densities $\rho_r = Q^{-1} \sum_{c=1}^Q f_{rc}$. Abbreviating further $N = \sum_{rc} f_{rc}$, we obtain

²Mathematically inclined readers will notice that we gloss over a few Dirac-delta functionals and questions regarding proportionality constants here. Discussing these subtleties would not yield any additional insights.

$$\begin{aligned}
\boldsymbol{\mu}_t &= \left(\sum_{r=1}^R \rho_r \right)^{-1} \sum_{r=1}^R \mathbf{t}_r \rho_r = \frac{1}{N} \sum_{c=1}^Q \sum_{r=1}^R \mathbf{t}_r f_{rc} \\
\boldsymbol{\mu}_s &= \frac{1}{Q} \sum_{c=1}^Q \mathbf{s}_c, \\
\boldsymbol{\mu}_f &= \frac{1}{N} \sum_{c=1}^Q \sum_{r=1}^R \begin{pmatrix} \mathbf{t}_r \\ \mathbf{s}_c \end{pmatrix} f_{rc},
\end{aligned} \tag{2}$$

and

$$\begin{aligned}
\mathbf{C}_t &= \frac{1}{N} \sum_{c=1}^Q \sum_{r=1}^R (\mathbf{t}_r - \boldsymbol{\mu}_t) (\mathbf{t}_r - \boldsymbol{\mu}_t)^\top f_{rc}, \\
\mathbf{C}_s &= \frac{1}{Q} \sum_{c=1}^Q (\mathbf{s}_c - \boldsymbol{\mu}_s) (\mathbf{s}_c - \boldsymbol{\mu}_s)^\top, \\
\mathbf{C}_f &= \frac{1}{N} \sum_{c=1}^Q \sum_{r=1}^R \begin{bmatrix} \mathbf{t}_r \\ \mathbf{s}_c \end{bmatrix} - \boldsymbol{\mu}_f \begin{bmatrix} \mathbf{t}_r \\ \mathbf{s}_c \end{bmatrix} - \boldsymbol{\mu}_f \begin{bmatrix} \mathbf{t}_r \\ \mathbf{s}_c \end{bmatrix}^\top f_{rc}.
\end{aligned} \tag{3}$$

Three short technical notes help understanding details of calculations in Eqs. (2) and (3). Note 1: The diet proportions of a consumer do not always add up to 1, e.g., when resource items are omitted from the statistics or when portions of diets remain unclassified. The normalization constant N takes this into account. Note 2: Because we computed weighted co-variances, no attempt of a Bessel correction (a denominator of “ $n - 1$ ” rather than “ n ”) was made. Note 3: Due to the initial data standardization $\boldsymbol{\mu}_s = 0$, and \mathbf{C}_s is proportional to an identity matrix.

Putting the estimated normal distributions into Eq. (1) gives

$$\begin{aligned}
&\exp \left\{ -\frac{1}{2} \begin{bmatrix} \mathbf{t} \\ \mathbf{s} \end{bmatrix} - \boldsymbol{\mu}_f \right\}^\top \mathbf{C}_f^{-1} \begin{bmatrix} \mathbf{t} \\ \mathbf{s} \end{bmatrix} - \boldsymbol{\mu}_f \left\} \propto \\
&a(\mathbf{t}, \mathbf{s}) \exp \left[-\frac{1}{2} (\mathbf{t} - \boldsymbol{\mu}_t)^\top \mathbf{C}_t^{-1} (\mathbf{t} - \boldsymbol{\mu}_t) \right] \exp \left[-\frac{1}{2} (\mathbf{s} - \boldsymbol{\mu}_s)^\top \mathbf{C}_s^{-1} (\mathbf{s} - \boldsymbol{\mu}_s) \right].
\end{aligned} \tag{4}$$

Solving this for $a(\mathbf{t}, \mathbf{s})$ yields

$$a(\mathbf{t}, \mathbf{s}) = a_1 \exp \left[\mathbf{b}^\top \begin{pmatrix} \mathbf{t} \\ \mathbf{s} \end{pmatrix} + \frac{1}{2} \begin{pmatrix} \mathbf{t} \\ \mathbf{s} \end{pmatrix}^\top \mathbf{C} \begin{pmatrix} \mathbf{t} \\ \mathbf{s} \end{pmatrix} \right] \tag{5}$$

with

$$\mathbf{b} = \mathbf{C}_f^{-1} \boldsymbol{\mu}_f - \begin{pmatrix} \mathbf{C}_t^{-1} \boldsymbol{\mu}_t \\ \mathbf{C}_s^{-1} \boldsymbol{\mu}_s \end{pmatrix} \tag{6}$$

and

$$\mathbf{C} = -\mathbf{C}_f^{-1} + \begin{pmatrix} \mathbf{C}_t^{-1} & 0 \\ 0 & \mathbf{C}_s^{-1} \end{pmatrix}. \tag{7}$$

We measure time in units of stomach turnover time, assuming this does not vary much between consumer species. The constant of proportionality a_1 can then be chosen such that the mean of Eq. (5) over all consumer-resource pairs, weighted by resource abundance ρ_r , approximates the actual weighted mean link strength $(Q \sum_r \rho_r)^{-1} \sum_{rc} f_{rc} \rho_r$.

Transformation to trophic traits The symmetric matrix \mathbf{C} defined by Eq. (7) can be represented in terms of a complete orthonormal set of $n + m$ eigenvectors \mathbf{e}_k and the corresponding real eigenvalues λ_k ,

$$\mathbf{C} = \sum_{k=1}^{n+m} \mathbf{e}_k \lambda_k \mathbf{e}_k^T. \quad (8)$$

Without loss of generality, we assume that the eigenvalues are ordered such that $|\lambda_1| \geq |\lambda_2| \geq \dots \geq |\lambda_{n+m}|$.

Now define, for each $k = 1, \dots, n + m$, partial eigenvectors \mathbf{e}'_k , \mathbf{e}''_k consisting of the first n and remaining m components of \mathbf{e}_k , i.e. the components corresponding to resource and consumer (Rossberg et al., 2010) respectively. With this definition:

$$\mathbf{e}_k = \begin{pmatrix} \mathbf{e}'_k \\ \mathbf{e}''_k \end{pmatrix}. \quad (9)$$

Since the eigenvectors \mathbf{e}_k form a complete orthonormal basis, we can decompose any $(n + m)$ -component vector \mathbf{u} as

$$\mathbf{u} = \sum_{k=1}^{n+m} \mathbf{e}_k (\mathbf{e}_k^T \mathbf{u}). \quad (10)$$

If \mathbf{t} represents the phenotypic traits of some resource item, and \mathbf{s} the phenotypic traits of a labeobarb species, we therefore get, using Eqs. (9) and (10),

$$\begin{pmatrix} \mathbf{t} \\ \mathbf{s} \end{pmatrix} = \sum_{k=1}^{n+m} \mathbf{e}_k (\mathbf{e}'_k{}^T \mathbf{t} + \mathbf{e}''_k{}^T \mathbf{s}) = \sum_{k=1}^{n+m} \mathbf{e}_k |\lambda_k|^{-1/2} (w^{(k)} - f^{(k)}) \quad (11)$$

where

$$w^{(k)} = |\lambda_k|^{1/2} \mathbf{e}'_k{}^T \mathbf{t}, \quad f^{(k)} = -|\lambda_k|^{1/2} \mathbf{e}''_k{}^T \mathbf{s}. \quad (12)$$

Putting Eq. (11) into the quadratic term in expression (5) and taking Eq. (8) and the orthogonality of the vectors \mathbf{e}_k into account yields

$$a(\mathbf{t}, \mathbf{s}) = a_1 \exp \left[\sum_{k=1}^{n+m} \mathbf{b}^T \mathbf{e}_k \mathbf{e}_k^T \begin{pmatrix} \mathbf{t} \\ \mathbf{s} \end{pmatrix} - \frac{1}{2} \sum_{k=1}^{n+m} \sigma_k (w^{(k)} - f^{(k)})^2 \right], \quad (13)$$

where

$$\sigma_k = -\text{sign } \lambda_k. \quad (14)$$

The variable-selection procedure will provide a dimensionality $D \geq 1$ of the trophic niche space to be used or tested. We therefore keep only the first D terms in the sum over $(w^{(k)} - f^{(k)})^2$. This is legitimate, because, as a result of the initial data standardization, the vectors \mathbf{t} and \mathbf{s} have the same variance in all directions. A large eigenvalue of \mathbf{C} therefore immediately implies that trait variations in the direction of the corresponding eigenvector are more important than those directions corresponding to smaller eigenvalues. An eigenvalue $\lambda_k = 0$ corresponds to a direction in trait space that does not matter for the quadratic term at all. It can therefore always be dropped. The case that $\lambda_k = 0$ and $\sigma_k = 0$ is therefore excluded in the following.

To simplify the first sum in Eq. (13), we define the trophic baseline traits as

$$V^* = \sum_{k=D+1}^{n+m} \mathbf{b}^T \mathbf{e}_k \mathbf{e}_k^T \mathbf{t} \quad \text{and} \quad F^* = \sum_{k=D+1}^{n+m} \mathbf{b}^T \mathbf{e}_k \mathbf{e}_k^T \mathbf{s}, \quad (15)$$

accounting for the last $n + m - D$ terms of this sum, and the constants

$$d^{(k)} = \sigma_k |\lambda_k|^{-1/2} \mathbf{b}^T \mathbf{e}_k \quad (k = 1, \dots, D), \quad (16)$$

encapsulating the information from the first D terms. The constants $d^{(k)}$ are then combined with the raw vulnerability traits to obtain the (final) vulnerability traits as

$$v^{(k)} = w^{(k)} - d^{(k)} \quad (k = 1, \dots, D). \quad (17)$$

It can be verified using Eqs. (12) to (17) that Eq. (5) is equivalent to

$$a(\mathbf{t}, \mathbf{s}) = a_0 \exp \left[V^* + F^* - \frac{1}{2} \sum_{k=1}^D \sigma_k \left(v^{(k)} - f^{(k)} \right)^2 + X \right], \quad (18)$$

with X denoting the remaining $n + m - D$ terms in the sum over k , and

$$a_0 = a_1 \exp \left[- \sum_{k=1}^D \frac{(\mathbf{b}^T \mathbf{e}_k)^2}{2\lambda_k} \right]. \quad (19)$$

The approximation that niche space is D -dimensional is equivalent to setting $X = 0$. With $a_{ij} = a(\mathbf{t}_i, \mathbf{s}_j)$, this yields Eq. (2) of the main text.

Formulae (12), (15), and (14) define a transformation from standardized physical/phenotypic traits to trophic traits. For any potential resource-consumer pair characterized by raw physical traits $\tilde{\mathbf{t}}$ and $\tilde{\mathbf{s}}$, the vectors \mathbf{t} and \mathbf{s} entering these equations are obtained as

$$\mathbf{t} = \mathbf{A} (\tilde{\mathbf{t}} - \bar{\mathbf{t}}), \quad \mathbf{s} = \mathbf{B} (\tilde{\mathbf{s}} - \bar{\mathbf{s}}), \quad (20)$$

where \mathbf{A} , \mathbf{B} , $\bar{\mathbf{t}}$, and $\bar{\mathbf{s}}$ are the coefficients of the initial standardization transformation of the data used for determining \mathbf{b} and \mathbf{C} . By combining this transformations with the transformation from standardized trait vectors to trophic traits, Eqs. (12), (15), (17), the final result in Eq. (18) can be used to predict trophic link strengths from phenotypic traits.

A.2 Cross-validation of the TTM

To assess, for a given set of physical/phenotypic trait variables, the predictive power of the TTM as obtained above, we performed a two-fold leave-one-out cross validation over all resource categories and labeobarb species. That is, running over all possible resource–labeobarb pairings (p, q) , the TTM was parametrized with the given set of phenotypic trait variables, and the predicted link strength of the omitted resource category p with the omitted labeobarb species q using this parametrization was correlated with the actual link strength.

The procedure is complicated by the fact that link strength is more difficult to measure for rare diet items than for abundant ones, leading to larger measurement errors. In computing the correlation coefficient, link strengths were therefore weighted with the square of our proxy for resource abundance, that is, its mean contribution to consumer diets: $\bar{f}_r = Q^{-1} \sum_{c=1}^Q f_{rc}$. Hence, the correlation was computed as

$$\rho = \frac{\sum_{p,q} (\hat{a}_{pq} - c_1)(a_{pq} - c_2) \bar{f}_p^2}{\left[\sum_{p,q} (\hat{a}_{pq} - c_1)^2 \bar{f}_p^2 \right]^{1/2} \left[\sum_{p,q} (a_{pq} - c_2)^2 \bar{f}_p^2 \right]^{1/2}} \quad (21)$$

with averaged link strengths

$$c_1 = \frac{\sum_{p,q} \hat{a}_{pq} \bar{f}_p^2}{\sum_{p,q} \bar{f}_p^2}, \quad c_2 = \frac{\sum_{p,q} a_{pq} \bar{f}_p^2}{\sum_{p,q} \bar{f}_p^2}. \quad (22)$$

This is very similar to directly computing the correlation between predicted and observed diet proportions, but, because of the quadratic weights entering Eq. (22), not the same.

The value of ρ was used as a measure for the predictive power of the TTM in the cross-validation procedure.

A.3 Selection of phenotypic trait variables

The estimation of the trophic link strength function $a(\mathbf{t}, \mathbf{s})$ requires estimation of $1 + (n + m)/2 + (n + m)^2/2$ real numbers entering the coefficients a_1 , \mathbf{b} , and \mathbf{C} in Eq. (5) from $R \times Q$ measured link strengths. As a rule of thumb, the numbers of physical resource traits n and phenotypic consumer traits m for which $a(\mathbf{t}, \mathbf{s})$ is estimated should therefore be small compared to both the number of resource categories R and the number of consumer species Q . In our case, many more trait variables have been measured. A systematic method for choosing appropriate sets of trait variables is therefore required. The number of trait variables can be reduced at two levels: immediately at the level of the raw trait variables, or at the level of trophic traits by choosing the number D of dimensions of trophic trait space to be retained. Here, both approaches are combined. We seek combinations of trait variables and a value of D that yield good predictive power according to the cross-validation procedure, i.e., combinations that give a large value of ρ defined by Eq. (21) above.

We conducted an exhaustive search through all combinations of up to $n = 2$ and $m = 2$ trait variables and values of D between 1 and 4. Reasons for limiting the search space to this range were given in the main text: It reduces the risk of unnoticed overfitting (Reunanen, 2003) and we find that dimension 4 makes only a small contribution to improving the fit of our model. There may be procedures that lead to good model fits also for larger values of n , m , or D . To support the conclusions of the present study, however, our conservative choices were sufficient.

References

- Reunanen, J. 2003. Overfitting in making comparisons between variable selection methods. *The Journal of Machine Learning Research* 3:1371–1382.
- Rossberg, A. G., Å. Brännström, and U. Dieckmann. 2010. How trophic interaction strength depends on traits — A conceptual framework for representing multidimensional trophic niche spaces. *Theoretical Ecology* 3:13–24.

Online resource 3

Electronic supplementary material to:

Journal: Theoretical Ecology

Title: Trophic niche-space imaging, using resource and consumer traits

Authors: Leopold A. J. Nagelkerke ^{1,*} and Axel G. Rossberg ^{2,3}

¹*Aquaculture and Fisheries Group, Wageningen University, Wageningen Institute of Animal Sciences (WIAS), De Elst 1, 6708 WD Wageningen, The Netherlands, e-mail:*

Leo.Nagelkerke@wur.nl, Tel: +31 317 483940, fax: +31 317 483937

²*Lowestoft Laboratory, Centre for Environment, Fisheries and Aquaculture Science (Cefas), Pakefield Road, Lowestoft, Suffolk NR33 0HT, United Kingdom*

³*School of Biological Sciences, Queen's University Belfast, Belfast BT9 7BL, Northern Ireland, United Kingdom, e-mail: axel@rossberg.net, Tel: +44 28 9097 5859, fax: +44 28 9097 5877*

* Author for correspondence

Results obtained for the TTMs with D = 1, 2, or 3.

The correlation coefficients (ρ) between predicted and observed trophic interaction strengths with increasing number of predictive dimensions.

D=3

D_c	dimensions fixed at best fitting values	dimensions permuted randomly	correlation coefficients (ρ)			
			two-sided 95% confidence interval ^a		actually found in 3D model ^b	
1	-	1, 2, 3	-0.299	- 0.283	0.655	
2	1	2, 3	0.356	- 0.614	0.655	
3	1, 2	3	0.655	- 0.655	0.655	

D=2

D_c	dimensions fixed at best fitting values	dimensions permuted randomly	correlation coefficients (ρ)			
			two-sided 95% confidence interval ^a		actually found in 2D model ^b	
1	-	1, 2	-0.290	- 0.286	0.656	
2	1	2	-0.045	- 0.493	0.656	

D=1

D_c	dimensions fixed at best fitting values	dimensions permuted randomly	correlation coefficients (ρ)	
			two-sided 95% confidence interval ^a	actually found in 1D model ^b
1	-	1	-0.283 - 0.291	0.140