Oecologia (2007) 153:973–984
DOI 10.1007/s00442-007-0798-8

COMMUNITY ECOLOGY

Testing the productive-space hypothesis: rational and power

David M. Post

Received: 26 April 2006 / Accepted: 18 June 2007 / Published online: 19 July 2007

Abstract Understanding and explaining the causes of variation in food-chain length is a fundamental challenge for community ecology. The productive-space hypothesis, which suggests food-chain length is determined by the combination of local resource availability and ecosystem size, is central to this challenge. Two different approaches currently exist for testing the productive-space hypothesis: (1) the dual gradient approach that tests for significant relationships between food-chain length and separate gradients of ecosystem size (e.g., lake volume) and per-unit-size resource availability (e.g., g C m^{-1} $year^{-2}$), and (2) the single gradient approach that tests for a significant relationship between food-chain length and the productive space (product of ecosystem size and per-unit-size resource availability). Here I evaluate the efficacy of the two approaches for testing the productive-space hypothesis. Using simulated data sets, I estimate the Type 1 and Type 2 error rates for single and dual gradient models in recovering a known relationship between food-chain length and ecosystem size, resource availability, or the combination of ecosystem size and resource ability, as specified by the productive-space hypothesis. The single gradient model provided high power (low Type 2 error rates) but had a very high Type 1 error rate, often erroneously supporting the productive-space hypothesis. The dual gradient model had a very low Type 1 error rate but suffered from low power to detect an effect of per-unit-size resource availability because the range of variation in resource availability is limited. Finally, I performed a retrospective power analysis for the Post et al. (Nature 405:1047–1049, 2000) data set, which tested and rejected the productive-space hypothesis using the dual gradient approach. I found that Post et al. (Nature 405:1047–1049, 2000) had sufficient power to reject the productive-space hypothesis in north temperate lakes; however, the productive-space hypothesis must be tested in other ecosystems before its generality can be fully addressed.

Keywords Ecosystem size · Food-chain length · Power · Productive-space hypothesis · Resource availability · Type-1 error · Type-2 error

> *The productive space hypothesis implies that maximum food-chain length should be greater, the greater the quantity:area (or volume) occupied by the food web times the productivity of that web*
>
> Schoener (1989)

Introduction

Food-chain length (FCL), a measure of the height of the food web, is a fundamental property of ecological communities (Pimm 1982; Post 2002a; Post and Takimoto 2007) that alters the form and strength of trophic interactions, affects ecosystem and community stability, and modifies nutrient cycling and contaminant bioaccumulation (Carpenter et al. 1987; DeAngelis et al. 1989; Kidd et al. 1995; May 1973; Oksanen et al. 1981; Pimm and Lawton

Communicated by Joel Trexler.

D. M. Post (✉)
Department of Ecology and Evolutionary Biology,
Yale University, New Haven, CT 06520-8105, USA
e-mail: david.post@yale.edu

1977; Schindler et al. 1997; Sterner et al. 1997). In 1989, Schoener (1989) proposed the productive-space hypothesis to address the absence of spatial considerations in the productivity (energy flow) hypothesis (Pimm 1982) for explaining variation in the FCL. The productive-space hypothesis predicts that FCL should increase as the product of ecosystem size (area or volume) and some measure of resource availability per-unit-size (e.g., g C m^{-2} $year^{-1}$). The productive-space hypothesis is a restatement of the energy flow hypothesis outlined by Hutchinson (1959) and Slobodkin (1961) and is based on the second law of thermodynamics. As originally conceived, the energy flow hypothesis proposes that, because a diminishing amount of energy reaches upper trophic levels, FCL should increase as the amount of energy or limiting resources available to top predators increases (Elton 1927; Hutchinson 1959; Lindeman 1942; Slobodkin 1961). In this hypothesis, resource availability would include both autochthonous primary production and allochthonous inputs. The amount of resource reaching the top predators is a function of resource availability at the base of the food web and energetic efficiencies throughout the food web (Pimm 1982; Yodzis 1984). Hutchinson (1959) and Slobodkin (1961) both implicitly incorporated ecosystem size as a determinant of resource availability in their conceptual development of the energy flow hypothesis; however, many modern empirical tests of energetic hypotheses neglected the potential importance of ecosystem size on the total energy available to support top predators (Briand and Cohen 1987; Jenkins et al. 1992; Pimm 1982; Pimm and Kitching 1987; Townsend et al. 1998). The productive-space hypothesis, for the first time, explicitly addressed the spatial component of resource availability.

There have been few explicit tests of the productive-space hypothesis (Post et al. 2000; Schoener1989; Spencer and Warren 1996; Thompson and Townsend 2005; Vander Zanden et al. 1999). Of these studies, Schoener (1989) and Vander Zanden et al. (1999) concluded that their data supported the productive-space hypothesis, while Spencer and Warren (1996) and Post et al. concluded that their data did not support the productive-space hypothesis. The different results may arise because two different methods were used to test the productive-space hypothesis (Post 2002a). The most obvious test of the productive-space hypothesis is a *single gradient approach* (Schoener 1989; Vander Zanden et al. 1999) that tests for a significant relationship between FCL and the productive space (where productive space is the product of ecosystem size and some measure of average per-unit-size resource availability). The single gradient approach is attractive because it makes a simple, direct prediction, but it cannot separate the potentially different effects of ecosystem size and local resource availability.

The second approach is a *dual gradient approach* that tests for significant relationships between FCL and independent gradients of ecosystem size and some measure of per-unit-size-resource availability (Post et al. 2000; Spencer and Warren 1996). Three testable predictions arise from the independent consideration of ecosystem size and per-unit-size resource availability: FCL could be determined by (1) resource availability (productivity hypothesis), (2) ecosystem size alone (ecosystem-size hypothesis), or (3) a combination of resource availability and ecosystem size (productive-space hypothesis) (Post 2002a; Post et al. 2000). The productivity and productive-space hypotheses both derive from the original resource-availability arguments; however, the productivity hypothesis does not explicitly include ecosystem size as a determinant of resource availability. The ecosystem-size hypothesis emerges from a variety of sources, including patterns of community assembly and the effects of ecosystem size on habitat heterogeneity, species richness, the scale of local interactions, and dietary specialization (Cohen and Newman 1992; Holt 1993; Post et al. 2000). The ecosystem-size hypothesis does not imply that energetic considerations are not important (they could be), but rather specifies that they are mediated by ecosystem size and are more complex than the simple application of the second law of thermodynamics as specified by the productivity and productive-space hypotheses. Testing and differentiating among these three hypotheses requires estimates of FCL across independent gradients of ecosystem size and some measure of per-unit-size resource availability (Post 2002a; Post et al. 2000; Spencer and Warren 1996). While this design explicitly addresses the potentially different effects of ecosystem size and per-unit-size resource availability, it may suffer from problems of statistical power because it expands the number of variables analyzed and because there is typically available a smaller range of variation in local resource availability than considered in the ecosystem-size and productive-space hypotheses (Schoener 1989).

Because the single and dual gradient approaches may provide different answers when testing the productive-space hypothesis, it is critical to understand the advantages and disadvantages of these two approaches. Here I focus on the rationale for and scope of inference provided by each approach, and the effect of each approach on the power to determine the determinants of FCL. In particular, I explore the potential for spurious correlations to emerge when using the single gradient approach (Type 1 error) and the potential lack of power to detect effects of resource availability when using the dual gradient approach (Type 2 error). To address the potential for spurious correlations when using the single gradient approach and lack of power using the dual gradient approach I used simulated data sets to estimate the Type 1 and Type 2 error rates for single and dual gradient models when recovering the known relationship between FCL and ecosystem size, resource availability, or productive space. I am particularly interested in Type 1 errors because they represent false support for a hypothesis that is not correct, which in turn can mislead future research. I also performed a retrospective power analysis to determine if Post et al. (2000), which tested and rejected the productive-space hypothesis using the dual gradient approach, actually had sufficient power to reject the productive-space hypothesis. Finally, I discuss broadly the assumptions and caveats that could hinder future tests of the theory for understanding natural variation in FCL.

Methods

Simulated data sets

To evaluate the Type 1 and Type 2 error rates of the single and dual gradient approaches, I simulated three data sets where FCL was a function of (1) only local resource availability, (2) only ecosystem size, and (3) the productive space. For each data set, I evaluated the relationship between FCL and productive space, ecosystem size, and local resource availability. The models used to simulate the data sets were:

$$\mathrm{FCL} = b_0 + b_1 \times \log(\text{ecosystem size}) + \varepsilon \qquad (1)$$

$$\mathrm{FCL} = b_0 + b_1 \times \log(\text{resource availability}) + \varepsilon \qquad (2)$$

$$\mathrm{FCL} = b_0 + b_1 \times \log(\text{ecosystem size}) + b_2 \times \log(\text{resource availability}) + \varepsilon \qquad (3)$$

where b_0 is the intercept, and b_1 and b_2 are the slope(s) of the relationship between FCL and ecosystem size or resource availability. For all simulations I used an intercept of 3.5. The results are not sensitive to the intercept. I used 0.22 for all slopes (b_1 and b_2) in each of the simulations because (1) it is the slope found by Post et al. (2000) for the relationship between FCL and ecosystem size, and (2) because the productive-space hypothesis suggests that the effects of ecosystem size and resource availability should be similar (i.e., productive space = ecosystem size × local resource availability). Other slopes (often steeper) can be justified from the first principles of ecological energetics (see below), but the shallower slope provides a more conservative estimate of power in analyses of the simulated data sets (power increases with effect size, which is slope in this case). Both ecosystem size and local resource availability were log transformed to linearize their relationship with FCL. Log transformation permitted me to simulate the productive space as the sum rather than the product of ecosystem size and local resource availability [i.e., log(ecosystem size) + log(local resource availability) = log(ecosystem size × local resource availability)]. As a result, the ecosystem size and local resource-availability models (Eqs. 1, 2) are fully nested within the productive space model (Eq. 3), and entire suite of models can be appropriately evaluated using an additive linear model without an interaction term.

For each model, I simulated data for $n = 25$ and $n = 100$ ecosystems in the comparison (e.g., 25 or 100 lakes, islands, streams, etc.). A sample size of $n = 25$ is the same as that used by Post et al. (2000). It is a reasonably large size for the purposes of power, and it is a manageable number of systems to sample for empirical studies. For these n ecosystems, ecosystem size and resource availability were independently simulated from uniform distributions spanning eight orders of magnitude for ecosystem size and three orders of magnitude for resource availability. The simulated gradients were not correlated. Eight and three orders of magnitude variation in ecosystem size and local resource availability, respectively, are reasonable ranges across which these questions could be addressed in natural ecosystems. The error term (ε), included to produce simulated data sets that better represented the data available to test the productive-space hypothesis, was normally distributed with a mean of 0 and a standard deviation (σ_ε) between 0 and 0.5 (analyzed in steps of 0.02).

To evaluate the potential for correlations between variables to reduce power and cause problems with parameter estimation, I simulated a data set using the productive space model outlined in Eq. 3 and $n = 25$ where I held the standard deviation of ε (σ_ε) constant at 0.2 and varied the correlation between ecosystem size and productivity from 0 to 1 (i.e., from no correlation to a perfect correlation). This simulates the empirical example where it may be difficult to work across independent gradients of ecosystem size and resource availability.

Least-squares regression (Model I) assumes that the independent (x) variables are measured with minimal error (Draper and Smith 1981; Sokal and Rohlf 1995). This assumption may be commonly violated where ecosystem size and resource availability are difficult to measure (e.g., Post et al. 2007). When measurement error is present, estimates of the slope(s) can be biased – typically becoming flatter as measurement error increases. Here I simulated three additional data sets to test the effects of measurement error in the independent variables on Type 1 and 2 error rates. The models used to simulate the data sets were:

$$\mathrm{FCL} = b_0 + b_1 \times \log(\text{ecosystem size} + \text{ecosystem size} \times \varepsilon_2) + \varepsilon_1 \quad (4)$$

$$\mathrm{FCL} = b_0 + b_1 \times \log(\text{resource availability} + \text{resource availability} \times \varepsilon_2) + \varepsilon_1 \quad (5)$$

$$\mathrm{FCL} = b_0 + b_1 \times \log(\text{ecosystem size} + \text{ecosystem size} \times \varepsilon_2) + b_2 \times \log(\text{resource availability} + \text{resource availability} \times \varepsilon_2) + \varepsilon_1 \quad (6)$$

Again I used an intercept of 3.5 and slopes of 0.22 for all simulations. In each model, there are now two error terms. The first error term (ε_1) is identical to the measurement error (ε) used in Eqs. 1–3. The second error term (ε_2) simulates errors in the measurement of the independent variables and was simulated as a proportion of the independent variable. Both ε_1 and ε_2 were drawn from normal distributions with a mean of 0. The standard deviation of ε_1 (σ_{ε_1}) was held constant at 0.2, and I varied the standard deviation of ε_2 (σ_{ε_2}) from 0 to 0.5 (analyzed in steps of 0.02). Thus, the true value of each independent variable was altered by a constant percentage of measurement error. At $\sigma_{\varepsilon_2} = 0$ there is no error in the measurement of each independent variable. At $\sigma_{\varepsilon_2} = 0.5$ the difference between the true and simulated independent variables (arithmetic) has an average standard deviation of 0.5.

From each simulated data set, I tested the ability of the single gradient and the dual gradient models to recover the simulated relationship between FCL and the causal environmental variable(s). The single gradient model I used was:

$$\mathrm{FCL} = b_0 + b_{1s} \times \log(\text{productive space}) + \varepsilon \quad (7)$$

where productive space was ecosystem size × resource availability for each ecosystem. The dual gradient model I used was:

$$\mathrm{FCL} = b_0 + b_{1d} \times \log(\text{ecosystem size}) + b_{2d} \times \log(\text{resource availability}) + \varepsilon \quad (8)$$

The fit of these two models to simulated data was evaluated at $\alpha = 0.05$ for 5000 simulations. From the 5000 simulations, I report the mean and the bootstrapped 95% confidence intervals (95% CI) for each regression parameter and the proportion of simulations that resulted in significant parameter estimates from which I estimated Type 1 and Type 2 error rates. There were no differences in the estimated mean values of b_{1s}, b_{1d}, or b_{2d} between simulations using $n = 25$ and $n = 100$. Not surprisingly, $n = 100$ simulations produced smaller 95% confidence intervals. Here I present results for only $n = 25$.

Power analysis

I use data from Post et al. (2000) to perform retrospective power analyses (e.g., Thomas 1997) to evaluate the potential lack of power using the dual gradient approach. The central issue here is the assertion that resource availability did not determine FCL in the Post et al. (2000) data set as tested by the dual gradient approach. Power is $1 - \beta$, where β is the probability of either making a Type 2 error or of accepting the null hypothesis of no effect when there is actually an effect. Although there is no convention for what constitutes sufficient or high power, various authors suggest power in the range of 80–95% is appropriate for most purposes (Cohen 1988; Peterman 1990); Toft and Shea (1983) recommend using the same stringent standard for rejecting a null hypothesis as for accepting a positive results (i.e., $\alpha = \beta = 0.05$, power = 0.95). Power is a function of sample size, sampling variance, α-level, and effect size (see Cohen 1988; Toft and Shea 1983; Thomas 1997, for a full discussion of retrospective power analyses). For my analyses, sample size was fixed ($n = 25$ lakes), and sample variance was estimated from the data in Post et al. (2000). I calculated power given different assumptions about both α (ranging from 0.05 to 0.2) and the postulated effect size. Because power and α are positively related, too stringent an α (e.g., 0.05 or 0.001) can lead to low power. Thus, I provide analyses that relax α to increase power in an effort to be conservative about making a Type 2 error (accepting the null hypothesis that there is no influence of resource availability on FCL when there really is an effect). I performed analyses to establish power within both a simple linear regression framework, which tests the productivity hypothesis, and a multiple regression framework, which directly tests the productive-space hypothesis.

Post et al. (2000) used total phosphorous (TP) as their index of resource availability because north temperate lakes are typically phosphorus-limited and TP is a strong predictor of primary production in lakes (Schindler 1978; Vollenweider 1979). Primary production is a good measure of resource availability in lakes; however, the relationship between TP and primary productivity is non-linear, with the rate of increase in primary productivity decreasing as TP increases (Vollenweider 1979). Although only an approximation, a one order of magnitude increase in TP produces about a 0.62 order of magnitude of increase in per-unit-size productivity (g C m^{-2} $year^{-1}$; for the range of TP used in this study, 2.6–230 μg P l^{-1}). Thus, Post et al. (2000) sampled across approximately 1.24 orders of magnitude of variation in per-unit-size primary productivity (g C m^{-2} $year^{-1}$). Post et al. (2000) used lake volume as their index of ecosystem size because it captured both vertical and horizontal components of habitat in lakes and sampled across 6.6 orders of magnitude of variation in ecosystem size.

Using linear regression to test for a relationship between resource availability and FCL does not consider other variables, thereby making it a test of the productivity hypothesis, which is the simplest form of the energetic hypothesis. The power analysis for linear regression used slope as a measure of effect size, and used slope, the standard deviation of X (per-unit-size productivity; 0.36 for the 25 lakes spread over 1.24 orders of magnitude of resource availability) and Y (maximum trophic position; 0.4) to determine sample variance (standard deviation of the residuals for this analysis). I used the linear regression power analysis to calculate the power to detect an effect of per-unit-size resource availability over the range of variation in per-unit-size resource availability used by Post et al. (2000). Effect size for these models depends upon ecological efficiency, which averages around 10% (Pauly and Christensen 1995; Slobodkin 1961) but the latter is variable and likely ranges between 3 and 55% (Pauly and Christensen 1995; Pimm 1982). Low ecological efficiencies provided conservative estimates of effect size (shallow slope). At an ecological efficiency of 3%, the productive-space hypothesis predicts that there should be a one trophic level increase in FCL for each 1.52 order of magnitude increase in per-unit-size resource availability. This translates into a slope of 0.66 for the relationship between FCL and resource availability. An ecological efficiency of 50% predicts a one trophic level increase in FCL for each 0.5 order of magnitude increase in resource availability, or a slope of 2 for the relationship between resource availability and FCL. Thus, a postulated effect size (slopes) for resource availability in the range of 0.65 to 2 is realistic. This would translate into a slope of 0.4–1.8 for a relationship between FCL and TP, as used to measure resource availability by Post et al. (2000)

The power analysis for multiple regression used R^2 as the measure of standardized effect size (incorporating effect size and sample variance into one variable; Cohen 1988). I used this framework to calculate the power to detect an effect of per-unit-size resource availability on FCL over and above that already explained by ecosystem size ($R^2 = 0.8$; Post et al. 2000), if there was indeed an effect of per-unit-size resource availability on FCL. Because the productive-space hypothesis postulates that per-unit-size resource availability and ecosystem size should have equal effects on FCL (Schoener 1989), effect size should be proportional to the range of variation in ecosystem size and per-unit-size productivity (both log-transformed). There were 6.6 orders of magnitude variation in ecosystem size and 1.24 orders of magnitude of variation in per-unit-size resource availability. This suggests an effect size for per-unit-size resource availability of around 1.24/7.84 of the total R^2, or $R^2 = 0.15$, given that ecosystem size alone had an $R^2 = 0.8$. Because a total R^2 of nearly 1.0 may not be reasonable, it is more realistic and conservative to hypothesize an effect size of $R^2 > 0.1$.

Results

Simulated datasets

The first question I address is whether the dual gradient model can reliably recover the original model used to simulate each data set. There are two parts to this evaluation: (1) did the model recover the original slope and (2) how often did the model find the slope significant (an issue of power). The dual gradient model reliably recovered a slope of 0.22 for the parameters used to simulate each data set (Fig. 1d, h, f, i) and a slope of 0.0 for parameters not used to simulate the data sets (b_{2d} in the ecosystem size model and b_{1d} in the resource-availability model; Fig. 1e, g). The 95% confidence intervals (bootstrapped) around the parameters increased with σ_ε (as would be expected), although in all cases the increase in uncertainty was greater for b_{2d} (resource availability) than b_{1d} (ecosystem size).

The dual gradient model had high power (low Type 2 error rate) and reliably recovered ecosystem size as the only determinant of FCL in the ecosystem-size simulations (Eq. 1) and as one of the determinants of FCL in the productive-space simulations (Eq. 3) in simulations across all values of σ_ε (Fig. 2a, c). The dual gradient model had high power (low Type 2 error rate) to recognize resource availability as the determinant of FCL in the resource-availability simulations (Eq. 2) and as a determinant of FCL in the productive-space simulations (Eq. 3) for values of $\sigma_\varepsilon < 0.25$ (Fig. 2b, c); however, as σ_ε increased above 0.25, power fell (Type 2 error rate increased), and the dual gradient model often failed to find a significant relationship between resource availability and FCL even when that relationship was used to simulate the data (Fig. 2b, c). The Type 1 error rate (rate of finding a significant relationship between ecosystem size or resource availability and FCL when they were not used to determine FCL), was around $1/2\alpha$ for b_{1d} in the resource-availability simulations and b_{2d} in the ecosystem size simulations where $\sigma_\varepsilon > 0$ (Fig. 2a, b). At $\sigma_\varepsilon = 0$, the Type 1 error rate was approximately α (Fig. 2a, b).

The single gradient model estimated a significant relationship between FCL and productive space (b_{1s}, ecosystem size × resource availability) in all model simulations. For data sets simulated with ecosystem size as the only determinant of FCL (Eq. 1), the single gradient model found a significant relationship between productive space and FCL in simulations across all values of σ_ε (a high Type 1 error rate; Fig. 2a) and estimated an average slope (b_{1s}) of 0.053 (Fig. 1a). For the data set simulated with resource

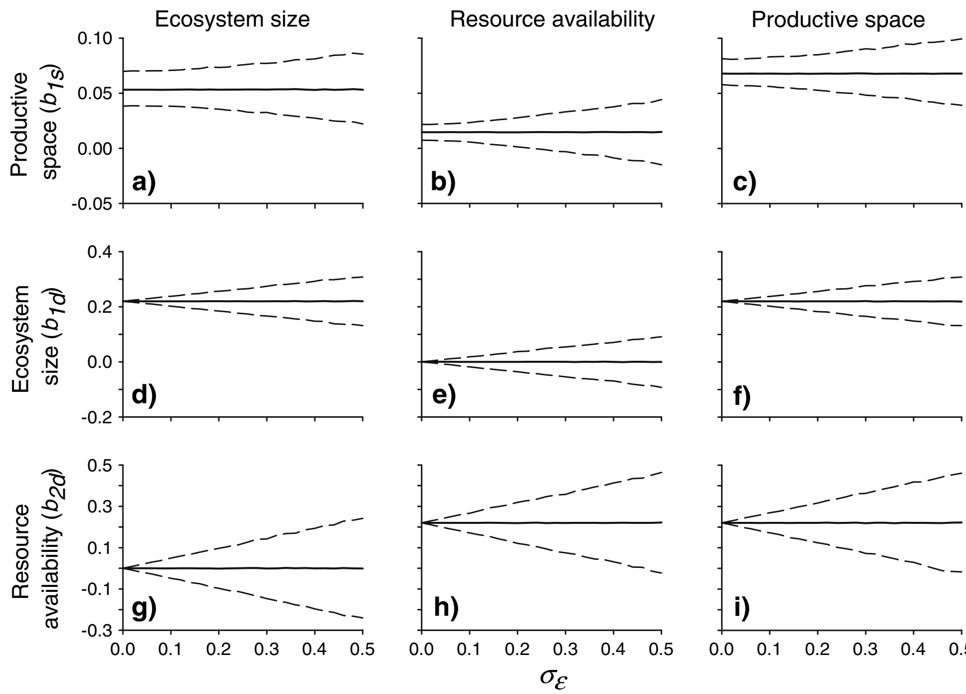

Fig. 1 Parameter estimates (*solid lines*) and 95% confidence intervals (*broken lines*) across a range of parameter independent measurement error (σ_ε) in simulations where ecosystem size (**a**, **d**, **g**), resource availability (**b**, **e**, **h**), or both ecosystem size and resource availability (the productive space; **c**, **f**, **I**) were the underlying determinant(s) of food-chain length (FCL; Eqs. 1–3). Confidence intervals are bootstrapped estimates from 5000 simulations.

availability as the only determinant of FCL (Eq. 2), the single gradient model again returned significant relationships between productive space and FCL across all values of σ_ε (Fig. 2b), although the Type 1 error rate fell as σ_ε increased, and estimated an average slope of 0.015 (Fig. 1b). Overall, the single gradient model had higher Type 1 error rates when ecosystem size, rather than resource availability, was the only determinant of FCL. Finally, the single gradient model recovered significant relationships between FCL and productive space across all values of σ_ε (i.e., it had high power; Fig. 2c) when both ecosystem size and resource availability were determinants of FCL (Eq. 3). The estimated average slope, 0.068, was higher than that estimated by the single gradient model for the other simulated datasets (Fig. 1).

To provide a graphic illustration of model fit, I selected a single-example data set simulated using ecosystem size as the only determinant of FCL (Eq. 1 at $\sigma_\varepsilon = 0.2$; Fig. 3). In this simulation, the dual gradient model accurately recovered the original model by finding a significant relationship between FCL and ecosystem size but not resource availability (Fig. 3a, b). The single gradient model found a significant relationship between FCL and the productive space (Fig. 3c). A major difference between the single and dual gradient models was the amount of variation explained. The dual gradient model explained much more of the variation in FCL ($R^2 = 0.867$; the average R^2 for all 5000 simulations was 0.873) than the single gradient model ($R^2 = 0.565$; average R^2 for all 5,000 simulations was 0.522) because the productive-space variable included resource availability, which was not correlated with FCL and therefore acted as a second error term.

Degree of correlation

The strength of the correlation between ecosystem size and resource availability (the independent variables) in the data set had some impact on power to detect significant relationships between FCL and ecosystem size (b_{1d}) or resource availability (b_{2d}) using the dual gradient model (Fig. 4). As was seen across variation in σ_ε, increasing the correlation most strongly affected the power to detect an effect of resource availability on FCL (Fig. 4). Correlations between the independent variables had little effect on parameter estimates for the dual gradient model until the correlation exceeded 0.98, at which point confidence intervals became extremely large. There was no effect of correlation on parameter estimates (around 0.07 for b_{1s}) or power when using the single gradient model (Fig. 4).

Error in independent variables

Measurement error had only modest effects on the estimates of slope, power, and Type 1 error rates. In all cases, measurement error biased slope estimates downward, although the bias was minimal in all cases (Appendix A). The effects of increasing $\sigma_{\varepsilon 2}$ were similar to the effects of increasing σ_ε on power and Type 1 error rates (Figs. 5, 6). At all values of σ_{ε_2}, the dual gradient model had sufficient power to recover the relationship between FCL and

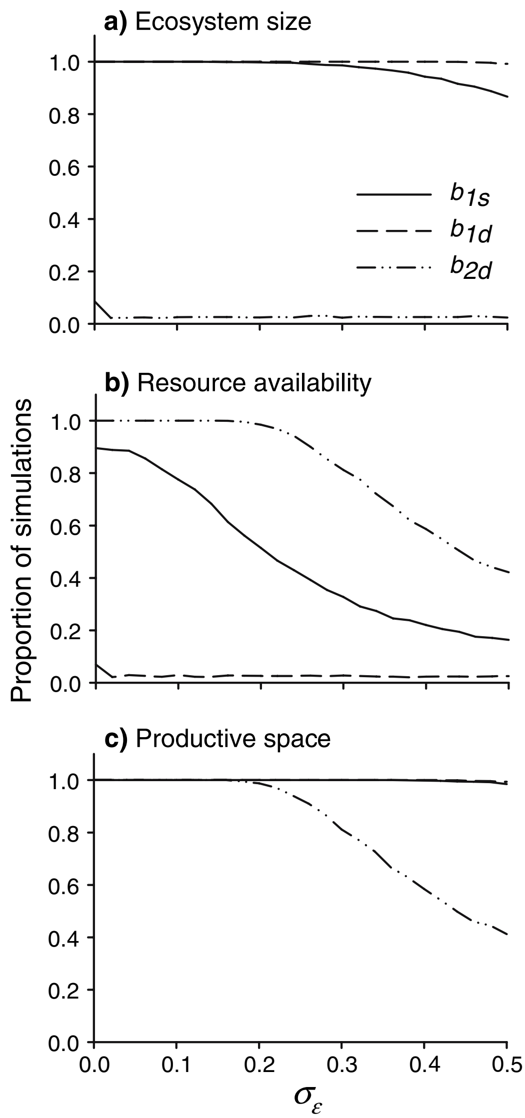

Fig. 2 The proportion of 5000 simulations that resulted in significant parameter estimates for the parameters relating FCL to ecosystem size (b_{1d}) and resource availability (b_{2d}) in the dual gradient model and to productive space (b_{1s}) in the single gradient model. Data sets were simulated (n = 25) across a range of parameter independent measurement error (σ_ε) added to the data sets with ecosystem size as the only determent (Eq. 1) (**a**), resource availability as the only determent (Eq. 2) (**b**) or both ecosystem size and resource availability as determents (**c**) of FCL (e.g., productive space; Eq. 3)

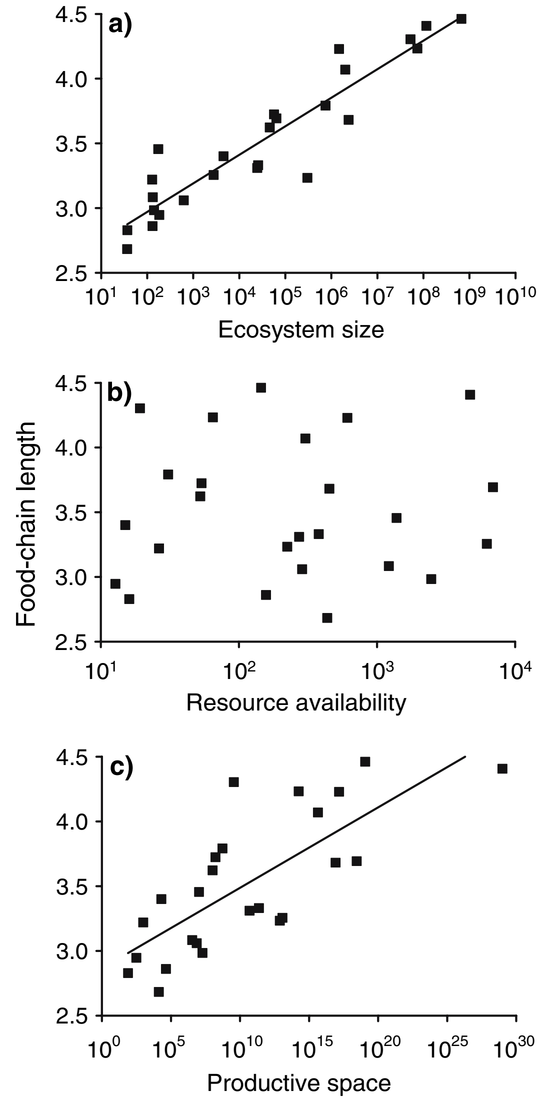

Fig. 3 Results from a randomly selected single simulation with ecosystem size (Eq. 1) as the determinant of FCL and at $\sigma_\varepsilon = 0.2$. The recovered relationship: for **a**, ecosystem size was FCL = 2.53 + 0.22 × log(ecosystem size), $p < 0.001$, $R^2 = 0.867$ (the average R^2 for all simulations at $\sigma_\varepsilon = 0.2$ was 0.873); for **b**, resource availability was not significant, $p = 0.87$; for **c**, the productive space was FCL = 2.87 + 0.06 × log(productive space), $p < 0.001$, $R^2 = 0.565$ (the average R^2 for all simulations at $\sigma_\varepsilon = 0.2$ was 0.522)

ecosystem size or resource availability (Fig. 5a, b), and the single gradient model had sufficient power to recover the relationship between FCL and productive space (Fig. 5c). High σ_{ε_2} reduced the power of the dual gradient model to detect an effect of resource availability when both resource availability and ecosystem size determined FCL (Fig. 5c). Measurement error had little effect on the high Type 1 error rate of the single gradient model when ecosystem size or resource availability were the only determinants of FCL (Fig. 5a, b).

Power analysis for the Post et al. (2000) data set

In a linear regression framework with 1.24 orders of magnitude of variance in resource availability (two orders of variation in TP), there was sufficient power to detect an effect of resource availability on FCL if the slope of the relationship between resource availability and FCL was greater than about 0.45 (at $\alpha = \beta = 0.2$; Fig. 5a). Under the very conservative conditions of $\alpha = \beta = 0.05$ (95% power), the minimum detectable slope would be 0.67. Relaxing α to

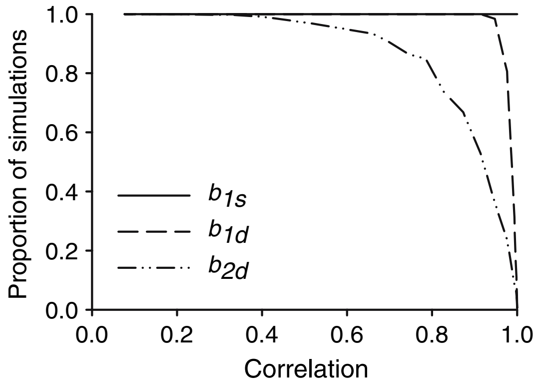

Fig. 4 The proportion of 500 simulations that resulted in significant parameter estimates for the parameters relating FCL to ecosystem size (b_{1d}) and resource availability (b_{2d}) in the dual gradient model and to productive space (b_{1s}) in the single gradient model. The data set was simulated with varying degrees of correlation between ecosystem size and resource availability and assuming that both ecosystem size and resource availability are determents of FCL (e.g., productive space; Eq. 3)

0.2 allowed a minimum detectable slope of 0.55 with 95% power. There was very little power to detect an effect of resource availability on FCL if the actual slope was less than 0.3 (Fig. 5a).

In the multiple regression framework, where lake size explained 80% of the variation in FCL, there was sufficient power to detect an effect of resource availability on FCL if resource availability actually explained just an additional 4 or 5% of the variation in FCL (power > 0.8; Fig. 5b). Again, using the conservative conditions of $\alpha = \beta = 0.05$ (95% power), the minimum detectable R^2 would be 0.073 (7.3% of additional variance explained). Relaxing α to 0.2 allowed a minimum detectable R^2 of 0.053 with 95% power. There was very little power to detect an effect of resource availability on FCL if the actual R^2 was less than 0.03 (Fig. 5b).

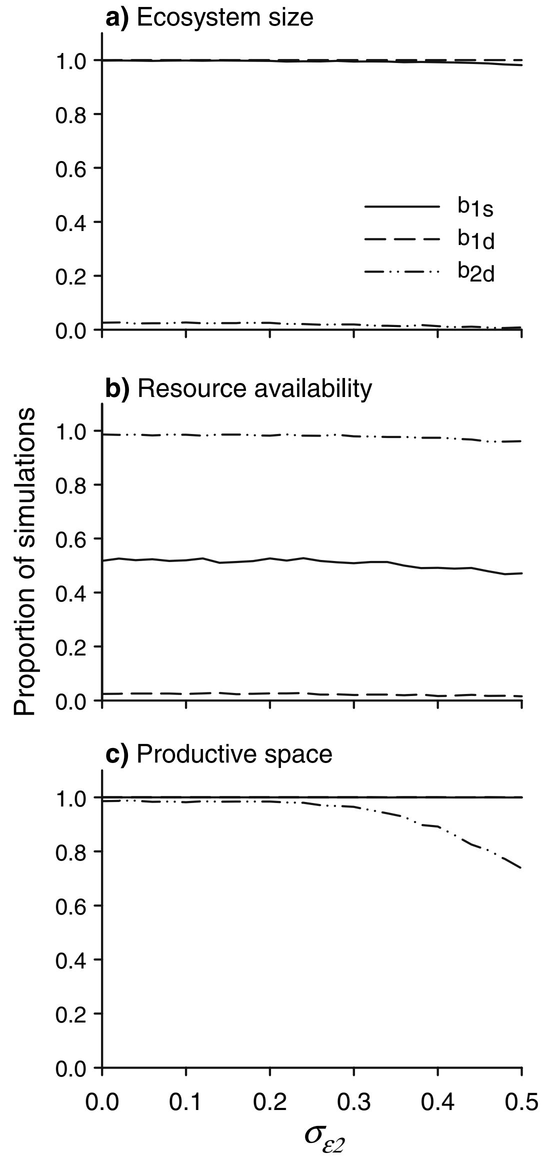

Fig. 5 The proportion of 5000 simulations that resulted in significant parameter estimates for the parameters relating FCL to ecosystem size (b_{1d}) and resource availability (b_{2d}) in the dual gradient model and to productive space (b_{1s}) in the single gradient model. Data sets were simulated ($n = 25$) across a range of error in the independent variables (σ_{ε_2}) for data sets with ecosystem size as the only determent (Eq. 4) (**a**), resource availability as the only determent (Eq. 5) (**b**) or both ecosystem size and resource availability (**c**) as determents of FCL (e.g., productive space; Eq. 5). In all cases, variance in the independent error was held constant at $\sigma_{\varepsilon_2} = 0.2$

Discussion

Addressing the efficacy of the single and dual gradient approaches requires placing the statistical performance of each approach (i.e., Type 1 and Type 2 error rates) into the context of the necessary and sufficient evidence for testing the productive-space hypothesis. This issue is important here because we are using patterns of natural variation in FCL to identify mechanisms that likely regulate FCL. An increase in FCL with either increasing ecosystem size or increasing resource availability supports the pattern outlined by the productive-space hypothesis but does not support the mechanistic foundation of this hypothesis (inefficiency in the transfer of energy up the food web ultimately limits FCL), which predicts FCL should increase with both ecosystem size and resource availability. If ecosystem size influences FCL only through its effect on total resource availability, then an increase in either ecosystem size or resource availability would be sufficient to support the productive-space hypothesis. Ecosystem size, however, can affect FCL in ways that are separate from its

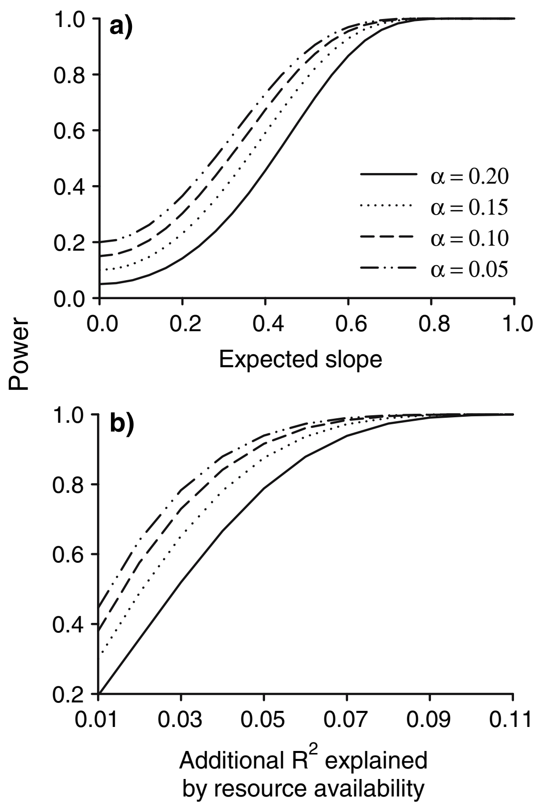

Fig. 6 Power to detect the influence of productivity on FCL versus effect size using a simple linear regression framework (**a**) and a multiple regression framework (**b**). Effect size for the multiple regression analysis is the postulated additional R^2 explained by productivity, given that log(ecosystem size) had an $R^2 = 0.80$. Effect size for the simple linear regression is the slope of postulated relationship between log(productivity) and FCL. The *separate lines* present power given different decisions about alpha (α; the probability of making a Type 1 error)

effect on total resource availability (e.g., habitat heterogeneity and species richness; Post et al. 2000, Post 2002a), and a positive relationship between FCL and both ecosystem size and resource availability is necessary to support the productive-space hypothesis and the mechanism upon which it was based.

Simulated data

Analyses of the three simulated data sets indicate that the single gradient approach does not provide the evidence necessary to support the productive-space hypothesis. The type 1 error rates were very high for the single gradient model because a relationship between FCL and productive space can be caused by the influence of ecosystem size alone (Fig. 2a), productivity alone (Fig. 2b at $\sigma_\varepsilon < 0.3$), or the combination of size and productivity, as specified by the productive-space hypothesis (Fig. 2c). The high Type 1 error rates indicates that the single gradient model may find support for the productive-space hypothesis because of spurious correlations that obscure the underlying mechanistic relationship between FCL and ecosystem size and resource availability (e.g., Fig. 3). Because an increase in FCL with both increasing ecosystem size and resource availability is necessary to support the productive-space hypothesis and the energetic mechanisms it is based on, the single gradient approach provides necessary but not sufficient evidence to support the productive-space hypothesis.

The dual gradient approach, on the other hand, provides the necessary and sufficient evidence to support the productive space hypothesis but has problems with power (Type 2 error), particularly in detecting potential effects of resource availability on FCL. The problem of power is not surprising because the dual gradient approach requires an additional variable for the analysis, and the range of resource availability in natural ecosystem is limited (Lawton 1989; Schoener 1989). Across all ecosystem types, the range of naturally observed primary productivity is roughly 1–3,000 g C m^{-2} $year^{-1}$ (Pimm 1982; Schoener 1989; Whittaker 1975), although there is some evidence that this widely cited range both over-estimates the lower end of the range and under-estimates the upper end of the range (e.g., Oksanen et al. 1981). Assuming that the range of resource availability, which can include allochthonous inputs, roughly matches the range of primary productivity in natural ecosystems, there are only a few orders of magnitude of variation across which we can evaluate variation in FCL. This range may be further constrained when working within a single ecosystem type (e.g., north temperate lakes likely span less than two orders of magnitude of variation in resource availability), although it is certainly true that the inclusion of a very low productivity ecosystem, when available, can greatly extend this range. In contrast to resource availability, there are more than12 orders of magnitude of variation in ecosystem size across which FCL can be evaluated, although realistically there are only eight to nine orders of magnitude of variation within any given ecosystem type, such as north temperate lakes and Caribbean islands. The single gradient approach has extremely high power across all values of σ_ε because there are 20–40 orders of magnitude of variation in productive space across which FCL can be measured (given the maximum range of variation in ecosystem size and resource availability and a range of correlation between the two variables).

Differences in the range of ecosystem size and resource availability used had an impact on the Type 1 error rate of the single gradient approach. The Type 1 error rate was the highest for the single gradient model in simulations when ecosystem size was the only determinant of FCL (Fig. 2a) because of the large leverage provided by the larger range of variation in ecosystem size. In simulations where resource availability was the determinant, the Type 1 error rate was lower, although it was always higher than expected by chance alone (Fig. 2b).

Both Post (2002a) and Spencer and Warren (1996) have argued that gradients of ecosystem size and resource availability should be independent for testing the productive-space hypothesis using the dual gradient approach. This appears to be partially true. At an σ_ε of 0.2, I find that strong correlations between ecosystem size and resource availability do reduce the power to detect the relationship between these independent variables and FCL using the dual gradient approach (Fig. 4). Strong correlations also caused uncertainty in the parameter estimates. Correlations had a greater effect on power for detecting the effects of resource availability than ecosystem size, again because of the limited range of variation available for resource availability. Correlations below 0.5 had no impact on power (again at an σ_ε of 0.2). Parameter correlations had little impact on power using the dual gradient model. It is important to note that the impact of parameter correlations on power will increase with increasing σ_ε (and measurement error – $\sigma_{\varepsilon 2}$); where σ_ε is large, correlations between the independent variables will cause great difficulties in estimating parameters and detecting the underlying effect of the separate variables (in particular resource availability).

Power analysis of the Post et al. (2000) results

A retrospective power analysis of the Post et al. (2000) data is useful in this context because they used the dual gradient approach and a relatively narrow range of variation in per-unit-size resource availability (1.24 orders of magnitude) to test the productive-space hypothesis. These are the conditions that can lead to insufficient power to detect effects of per-unit-size resource availability on FCL, and it raises the question: did Post et al. (2000) have sufficient power to detect an effect of resource availability on FCL – if indeed there was an effect? A critical component of retrospective power analyses is determining an appropriate effect size (Thomas 1997). Because there was no significant effect of per-unit-size resource availability on FCL in the Post et al. (2000) data, I must suppose a reasonable effect size to perform the retrospective power analysis using either the simple linear regression or multiple regression frameworks.

Using a simple linear regression framework, I ask if there is an effect of resource availability without taking into account ecosystem size (e.g., the productivity hypothesis). Effect size for linear regression is the slope of the hypothesized relationship between FCL and per-unit-size resource availability. Because the productivity hypotheses predict that FCL is ultimately determined by the inefficiency of energy transfer between steps in a food web (Hutchinson 1959; Pimm 1982; Schoener 1989; Slobodkin 1961), effect size for a simple linear regression emerges from assumptions about ecological efficiency. With an ecological efficiency of between 3 and 55% (Pauly and Christensen 1995; Pimm 1982), a postulated effect size (slopes) for resource availability in the range of 0.65–2 is realistic. Even with only 1.24 orders of magnitude of variation in per-unit-size primary productivity (g C m^{-2} $year^{-1}$), the Post et al. (2000) data set has very high power to detect the minimum postulated effect size of 0.65 (Fig. 6).

The multiple regression framework explicitly tests the productive-space hypothesis using the dual gradient approach. Because the productive-space hypothesis postulates that per-unit-size resource availability and ecosystem size have equal effects on FCL (Schoener 1989), effect size should be proportional to the range of variation in ecosystem size and per-unit-size resource availability (both log-transformed). Post et al. (2000) worked across 6.6 orders of magnitude variation in ecosystem size, and 1.24 orders of magnitude of variation in per-unit-size primary productivity (g C m^{-2} y^{-1}), suggesting an effect size for per-unit-size productivity of around 1.24/7.84 of the total R^2, or $R^2 = 0.15$ given that ecosystem size alone had an $R^2 = 0.8$. The retrospective power analysis shows that Post et al. (2000) had sufficient power to detect an R^2 as low as 0.05 – well below what we expect for an effect size – and very high power to detect R^2 between 0.1 and 0.15 (Fig. 6).

Both power analyses suggest that it is very unlikely that there is actually an undetected direct effect of per-unit-size productivity on FCL in the lakes studied by Post et al. (2000). If there is an effect, it is small and much less than expected given current versions of the productive space and productivity hypotheses. It is noteworthy that the slope (effect size) I estimated from a simple application of the second law of thermodynamics under the energetic hypothesis is considerably higher (between 0.65 and 2) than that estimated assuming ecosystem size and resource availability have similar effects on FCL based on the Post et al. (2000) observations (0.22). In contrast, the slopes recovered for the relationship between FCL and productive space using the single gradient model were also quite low (although how this might relate to a "true" relationship is not clear).

Additional considerations

Power and Type 1 error rates are only two of many the difficulties that must be addressed when testing hypotheses for natural variation in FCL. Important additional considerations include reliable and robust measures of FCL and the environmental gradients being tested. For example, ecosystem size can be hard to estimate in a spatially open ecosystem where physical boundaries are absent or where there is little congruence between physical boundaries, community membership, and ecosystem function, such as resource supply (Post et al. 2007). Likewise, resource availability can be spatially and temporally variable and should be scaled to match the spatial scale of the food web in which FCL is being measured. These problems suggest that measurement error may be a common problem when testing the productive-space hypothesis and other hypotheses for variation in FCL. The results of my simulations suggest that measurement error can have some effect on power and Type 1 error rates but, perhaps most importantly, it may bias estimates of the slope of the relationship between FCL and independent variables (Appendix A). This is particularly problematic when using the single gradient approach.

Difficulties in estimating FCL are of less statistical concern, but they do create fundamental problems for testing the productive-space and other hypotheses. Central issues include the currency used to estimate FCL (e.g., energy flow vs. interaction strength; Post 2002a), the method used to estimate FCL (e.g., connectance webs vs. stable isotopes; Post 2002b), and the set of top predators considered when estimating FCL (Post and Takimoto 2007). Mobile top predators, such as birds and anadromous fish, can cause considerable problems when estimating FCL. For example, birds that feed across several lakes or habitat patches may have strong top–down effects on local food webs (Power et al. 1989; Steinmetz et al. 2003; Wootton 1995) and, therefore, must be considered in estimates of functional FCL (Post 2002a). For a test of hypotheses related to realized FCL, such as the productive-space hypothesis, ecosystem size and resource availability should be scaled to the resource shed over which they are foraging (Post et al. 2007; Power and Rainey 2000). Ecosystem size for an osprey feeding on a small lake is not the area or volume of the lake, but rather the area of volume of all the lakes over which the osprey forages. Thus, estimates of resource availability, delineation of ecosystem size, and definitions of community membership are all interrelated when testing the productive-space and other hypotheses for variation in FCL.

Conclusion

The productive-space hypothesis is the focus of contemporary research on the determinants of variation in FCL. It is important, therefore, to evaluate the efficacy of the two models used to test this hypothesis – the single and dual gradient models. The single gradient model is attractive because it provides a simple framework and high power to detect a significant effect of the productive space on FCL because of the large available range of variation in productive space. It cannot, however, explicitly differentiate among the potentially different effects of ecosystem size, local resource availability, and total ecosystem resource availability. The resulting high Type 1 error rate means the single gradient model often found support for the productive-space hypothesis when it should have found no support for the hypothesis. In the end, the single gradient model fails because the effects of local resource availability and ecosystem size, which can affect FCL through multiple mechanisms, are nested within the productive space variable allowing for spurious correlation.

In contrast, the dual gradient model can distinguish between effects of local resource availability, total ecosystem resource availability, and ecosystem size. It has a very low Type 1 error rate (no more than expected by chance) but suffers from a lower power than the single gradient model. This is a particular problem for detecting effects of per-unit-size resource availability because of the small available range of variation in resource availability. Despite their use of the dual gradient approach, retrospective power analysis demonstrates that Post et al. (2000) had sufficient power to reject the productive-space hypothesis. Whether the productive-space hypothesis will receive support in other ecosystems remains to be seen, but where possible, future tests should use the dual gradient (or multiple gradients) approach. In applying the dual gradient approach, power to detect an effect of per-unit-size resource availability on FCL should receive extra attention.

Acknowledgments I thank C. A. Layman, L. Oksanen, L.M. Puth, G. Takimoto, J. Trexler, and an anonymous reviewer for comments and suggestions that improved this paper. This research was supported by the National Science Foundation (DEB no. 0316679).

References

Briand F, Cohen JE (1987) Environmental correlates of food chain length. Science 238:956–960
Carpenter SR et al. (1987) Regulation of lake primary productivity by food web structure. Ecology 68:1863–1876
Cohen J (1988) Statistical power analysis for the behavioral sciences, 2nd edn. Lawrence Erlbaum, Upper Saddle River
Cohen JE, Newman CM (1992) Community area and food-chain length: theoretical predictions. Am Nat 138:1542–1554
DeAngelis DL, Bartell SM, Brenkert AL (1989) Effects of nutrient recycling and food-chain length on resilience. Am Nat 134:778–805
Draper NR, Smith H (1981) Applied regression analysis, 2nd edn. Wiley, New York
Elton C (1927) Animal ecology. Sidgwick and Jackson, London
Holt RD (1993) Ecology at the mesoscale: the influence of regional processes on local communities. In: Ricklefs RE, Schluter D (eds) Species diversity in ecological communities. University of Chicago Press, Chicago, pp 77–88
Hutchinson GE (1959) Homage to Santa Rosalia; or, why are there so many kinds of animals? Am Nat 93:145–159

Jenkins B, Kitching RL, Pimm SL (1992) Productivity, disturbance and food web structure at a local spatial scale in experimental container habitats. Oikos 65:249–255

Kidd KA, Hesslein RH, Fudge RJP, Hallard KA (1995) The influence of trophic level as measured by delta –15N on mercury concentrations in freshwater organisms. Water Air Soil Pollut 80:1011–1015

Lawton JH (1989) Food webs. In: Cherrett JM (eds) Ecological concepts. Blackwell, Oxford, pp 43–78

Lindeman RL (1942) The trophic-dynamics aspect of ecology. Ecology 23:399–418

May RM (1973) Stability and complexity in model ecosystems. Princeton University Press, Princeton

Oksanen L, Fretwell SD, Arruda J, Niemelä P (1981) Exploitation ecosystems in gradients of primary productivity. Am Nat 118:240–261

Pauly D, Christensen V (1995) Primary production required to sustain global fisheries. Nature 374:255–257

Peterman RM (1990) Statistical power analysis can improve fisheries research and management. Can J Fish Aquat Sci 47:2–15

Pimm SL (1982) Food webs. Chapman and Hall, London

Pimm SL, Lawton JH (1977) The number of trophic levels in ecological communities. Nature 275:542–544

Pimm SL, Kitching RL (1987) The determinants of food chain lengths. Oikos 50:302–307

Post DM (2002a) The long and short of food-chain length. Trends Ecol Evol 17:269–277

Post DM (2002b) Using stable isotopes to estimate trophic position: models, methods, and assumptions. Ecology 83:703–718

Post DM, Takimoto G (2007) Proximate structural mechanisms for variation in food-chain length. Oikos 116:775–782

Post DM, Pace ML, Hairston NG (2000) Ecosystem size determines food-chain length in lakes. Nature 405:1047–1049

Post DM, Doyle MW, Sabo JL, Finlay JC (2007) The problem of boundaries in defining ecosystems: a potential landmine for uniting geomorphology and ecology. Geomorphology 89:111–126

Power ME, Rainey WE (2000) Food webs and resource sheds: towards spatially delimiting trophic interactions. In: Hutchings MJ, John EA, Stewart AJA (eds) Ecological consequences of habitat heterogeneity. Blackwell, Oxford, pp 291–314

Power ME, Dudley TL, Cooper SD (1989) Grazing catfish, fishing birds, and attached algae in a panamanian stream. Environ Biol Fish 26:285–294

Schindler DW (1978) Factors regulating phytoplankton production and standing crop in the world's lakes. Limnol Oceanogr 23:478–486

Schindler DE, Carpenter SR, Cole JJ, Kitchell JF, Pace ML (1997) Influence of food web structure on carbon exchange between lakes and the atmosphere. Science 277:248–251

Schoener TW (1989) Food webs from the small to the large. Ecology 70:1559–1589

Slobodkin LB (1961) Growth and regulation of animal populations. Holt, Rinehart and Wilson, New York

Sokal RR, Rohlf FJ (1995) Biometry, 3rd edn. W.H. Freeman, New York

Spencer M, Warren PH (1996) The effects of habitat size and productivity on food web structure in small aquatic microcosms. Oikos 75:419–430

Steinmetz J, Kohler SL, Soluk DA (2003) Birds are overlooked top predators in aquatic food webs. Ecology 84:1324–1328

Sterner RW, Bajpai A, Adams T (1997) The enigma of food chain length: absence of theoretical evidence for dynamic constraints. Ecology 78:2258–2262

Thomas L (1997) Retrospective power analysis. Conserv Biol 11:276–280

Thompson RM, Townsend CR (2005) Energy availability, spatial heterogeneity and ecosystem size predict food-web structure in streams. Oikos 108:137–148

Toft CA, Shea PJ (1983) Detecting community-wide patterns: estimating power strengthens statistical inference. Am Nat 122:618–625

Townsend CR, Thompson RM, McIntosh AR, Kilroy C, Edwards E, Scarsbrook MR (1998) Disturbance, resource supply, and food-web architecture in streams. Ecol Lett 1:200–209

Vander Zanden MJ, Shuter BJ, Lester N, Rasmussen JB (1999) Patterns of food chain length in lakes: a stable isotope study. Am Nat 154:406–416

Vollenweider VRA (1979) Das Nährstoffbelastungskonzept als grundlage für den externen eingriff in den eutrophierungsprozeß stehender gewässer und talsperren. Z Wasser-u Abwasser-Furschung 12:46–56

Whittaker RH (1975) Communities and ecosystems, 2nd edn. Macmillan, New York

Wootton JT (1995) Effects of birds on sea urchins and algae: a lower-intertidal trophic cascade. Ecoscience 2:321–328

Yodzis P (1984) Energy flow and the vertical structure of real ecosystems. Oecologia 65:86–88