Oikos 121: 290–298, 2012
doi: 10.1111/j.1600-0706.2011.19656.x

Subject Editor: Mark Vellend. Accepted 6 May 2011

Analyzing the effects of species gain and loss on ecosystem function using the extended Price equation partition

Jeremy W. Fox and Benjamin Kerr

J. W. Fox (jefox@ucalgary.ca), Dept of Biological Sciences, Univ. of Calgary, 2500 University Dr. NW, Calgary, AB, T2N 1N4, Canada. – B. Kerr, Dept of Biology, Univ. of Washington, Box 351800, Seattle, WA 98195, USA.

In nature species richness and composition, as well as the functioning of individual species, all covary along environmental gradients, making it difficult to tease apart their effects on ecosystem function. Here we use a novel extension of the Price equation to partition the causes of functional variation between any two sites sharing at least one species in common. We use the extension to separate effects of species loss from those of species gain; species gain is analogous to migration in evolution. Previous theoretical and empirical studies of biodiversity and ecosystem function fail to distinguish effects of species gain from those of species loss, and so are conceptually incomplete. Application of this approach to data on total plant biomass along an experimental N enrichment gradient leads to novel empirical insights and reveals subtle effects. For instance, effects of species gain are non-negligible even though enrichment leads to loss of many species and gain of few, and non-random gain of high-biomass species reduces the biomass of the persisting species. We also discuss the interpretation of this new approach, which provides a highly-general partitioning of the factors affecting ecosystem function.

Ecosystems perform many ‘functions’ on which life depends. They produce and decompose biomass, take up nutrients, sequester pollutants, take up CO_2, etc. The rate or level of many ecosystem functions varies with the number and identity of the species performing the function, all else being equal (Balvanera et al. 2006, Cardinale et al. 2009).

However, in nature local environmental conditions strongly affect the functioning of any given species, by affecting both its abundance and its per capita functioning. The same species may function at a high level at one time or place, and a low level at another time or place (Fox 2006, Gonzalez et al. 2009, Gonzalez and Loreau 2010). Species richness, composition, and the functioning of individual species all covary along environmental gradients, making it difficult to separate their effects without manipulative experiments that are not always feasible. It would be useful to have a generally-applicable method to decompose total variation in ecosystem function into components attributable to different effects.

Analogous problems arise in other fields. For instance, in evolutionary biology the properties of an evolving population, such as the mean phenotype, vary over time due to the combined effects of natural selection, migration, mutation, and other evolutionary forces. Addressing many fundamental evolutionary questions requires decomposing the total change in mean phenotype over some period of time into components attributable to different evolutionary forces.

Price (1970, 1972) showed how to decompose the total evolutionary change in mean phenotype between an ancestral and descendent population into components due to natural selection, and to factors that bias transmission. This decomposition is called the Price equation. Fox (2006) extended the Price equation to explain how ecosystem function varies between sites due to variation in species richness, species composition, and in the functioning of species present at both sites.

The original Price equation implicitly assumes zero immigration, and so only describes evolutionary forces acting on members of the ancestral population. Analogously, the approach of Fox (2006) only applies when one site comprises a strictly nested subset of the species in the other, so that the species comprising the less-diverse site can be viewed as having ‘descended’ from the species comprising the more-diverse site. In both cases, this is an important limitation. Few evolving populations are closed to immigration, and few diversity gradients (whether natural or anthropogenic) comprise strictly nested subsets of species (Wright et al. 1998). Recently, Kerr and Godfrey-Smith (2009) showed how to incorporate the evolutionary effects of immigration into the Price equation. Here, we show how to use this extended Price equation to compare ecosystem function between any two sites sharing at least one species in common (Fig. 1). We apply the extended approach to data on total plant biomass along an experimentally-imposed nutrient enrichment gradient (Tilman 1987), and show that it provides novel empirical insights. We also discuss key conceptual issues regarding the interpretation of this extended Price equation partition. In particular, the approach demonstrates that, along non-nested diversity gradients, it is

---

The review and decision to publish this paper has been taken by the above noted SE. The decision by the handling SE is shared by a second SE and the EiC.

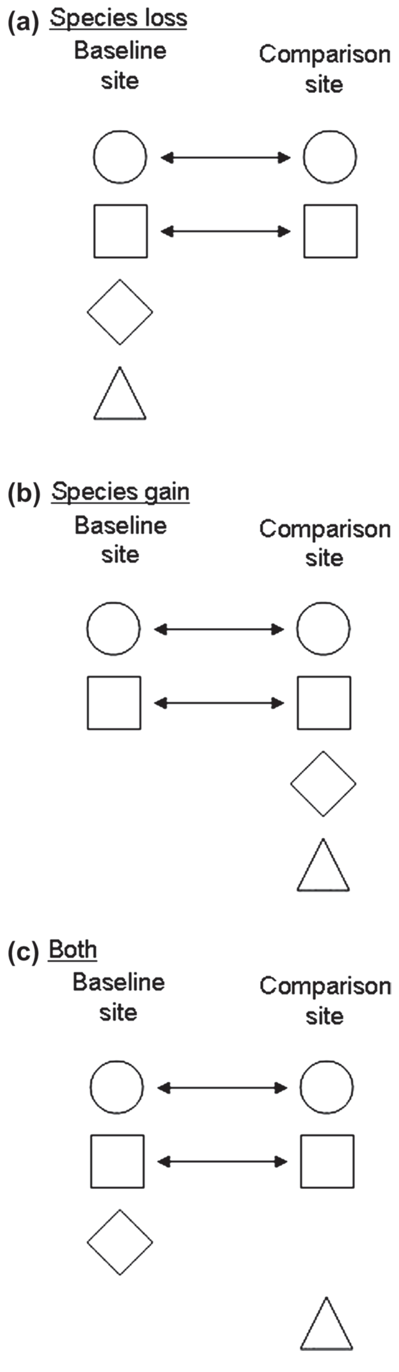


Figure 1. Schematic illustration of species loss and gain. Different shapes represent different species. Species present at both sites are connected by arrows. (a) Species loss. (b) Species gain. (c) Species loss and gain. Note that species gain is the mirror image of species loss (compare (a) and (b)).

useful to recognize more than one effect of species richness or composition on ecosystem function. Specifically, new patterns are revealed when effects arising from species gain ('immigration') are distinguished from those arising from species loss.

Extending the Price equation partition

We first summarize the extended Price equation partition of Kerr and Godfrey-Smith (2009) and describe its application to ecosystem function (see Kerr and Godfrey-Smith 2009 for mathematical details). The goal of the approach is to compare ecosystem function between two sites, a 'baseline' site comprised of s species and a 'comparison' site comprised of s' species. The species comprising the baseline site are analogous to the individuals comprising an ancestral population, with the species comprising the comparison site being analogous to a descendent population. Let s_c denote the number of species common to both sites, and assume that the two sites share at least one species in common ($s_c \geq 1$). Let the s species at the baseline site be indexed $i = 1,2,...,s$, with the species shared with the comparison site being indexed first (indexing the shared species first is merely a notational convenience). Let the s' species at the comparison site be indexed $j = 1,2,..., s'$, with the shared species being indexed first and in the same order as for the baseline site. We assume that total ecosystem function at each site comprises the sum of the measurable contributions of the species performing the function. Many important ecosystem functions satisfy this assumption (Discussion). Let z_i be the functional contribution of species i at the baseline site, and let z'_j be the functional contribution of species j at the comparison site. Species' functional contributions (z and z' values) are analogous to the phenotypes of individual organisms in evolution. Let $T = \sum_{i=1}^{s} z_i$ and $T' = \sum_{j=1}^{s'} z'_j$ respectively denote total function at the baseline and comparison sites, and let $\bar{z} = \frac{1}{s}\sum_{i=1}^{s} z_i$ and $\bar{z}' = \frac{1}{s'}\sum_{j=1}^{s'} z'_j$ respectively denote mean function per-species at the baseline and comparison sites. Let w_j^i be a variable which indicates whether a species is present at both sites, so that $w_j^i = 1$ if $i = j \leq s_c$ and 0 otherwise. The difference in ecosystem function is then given by

$$\Delta T = T' - T = (s_c - s)\bar{z} + (s' - s_c)\bar{z}' + s_c\left(\bar{z}' - \bar{z}\right) =$$

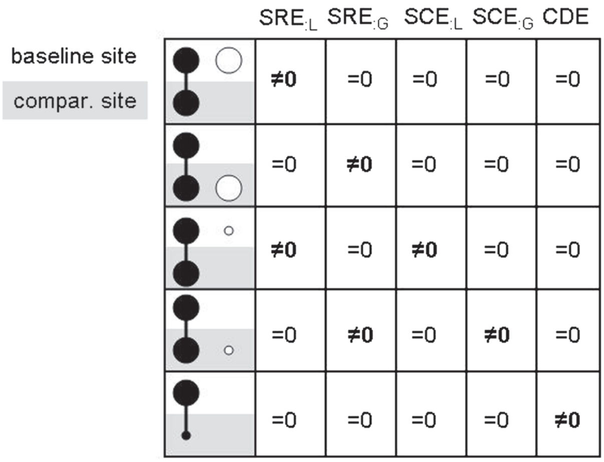
$$(s_c - s)\bar{z} + (s' - s_c)\bar{z}' + \mathrm{Sp}\left(w_{\bullet}^{I}, z\right) + \left[-\mathrm{Sp}\left(w_{J}^{\bullet}, z\right)\right] + \sum_{i=1}^{s_c} \Delta z_i \tag{1}$$

where Sp denotes the sum of products operator, $w_{\bullet}^{I} = \sum_{j=1}^{s'} w_j^i$, $w_{J}^{\bullet} = \sum_{i=1}^{s} w_j^i$, and $\Delta z_i = z'_i - z_i$. We either drop subscripts (e.g. z and z') or use capital scripts (e.g. $w_{\bullet}^{I}$ and $w_{J}^{\bullet}$) to denote general variables as arguments for statistical operators (e.g. Sp). In our approach, $w_{\bullet}^{I} = 1$ if species $i = I$ is common to both sites and 0 otherwise, and $w_{J}^{\bullet} = 1$ if species $j = J$ is common to both sites and 0 otherwise. The quantity $\sum_{i=1}^{s_c} \Delta z_i = \sum_{i=1}^{s}\sum_{j=1}^{s'} w_j^i \Delta z_i$ is the sum of the between-site differences in function for the species present at both sites, since species with $w_j^i = 0$ do not contribute to this sum.

Equation 1 comprises five additive terms. The first term, $(s_c - s)\bar{z}$, captures the effect of loss of species richness, denoted SRE_L. This is the amount by which total baseline function T would be expected to change if species were lost at random from the baseline site (so that $s_c < s$), no species were gained by the comparison site (so that $s_c = s'$), and nothing else changed (Fig. 2). Under these conditions, mean functional contribution per species is the same at both sites ($\bar{z}' = \bar{z}$). This term is always < 0 when species are lost and species make positive functional contributions ($z_i > 0$ for all i). Even when species are lost non-randomly, this term isolates that part of the functional effect of that loss that is not uniquely attributable to the particular species that were lost, and so would have occurred no matter which species were lost (Fox 2006). Mathematically, the SRE_L simply isolates the effect on a sum of removing one or more of the summands, without changing the mean of the summands (Fig. 2; see appendices of Fox 2010 for further discussion). This term equals the species richness effect (SRE) of Fox (2006).

The second term, $(s' - s_c)\bar{z}'$, is the amount by which baseline function T would be expected to change if species were gained at random at the comparison site, no species were lost from the baseline site, and nothing else changed. We refer to this term as the species richness effect due to species gain, denoted SRE_G. The SRE_G is the mirror image of the SRE_L: it is always > 0 when species are gained and species make positive functional contributions ($z_j > 0$ for all j) (Fig. 2).

The sum of the final three terms in Eq. 1 equals $s_c\left(\bar{z}' - \bar{z}\right)$, the between-site difference in mean functional contribution per species, scaled by s_c. The three terms comprising this sum all affect total function by affecting mean functional contribution per species, but they do so in different ways. The third term in Eq. 1, $\text{Sp}\left(w_\bullet^I, z\right)$, equals the species composition effect (SCE) of Fox (2006). This term quantifies the change in ecosystem function attributable to species loss from the baseline site that is non-random with respect to species' baseline functional contributions. Species lost from the baseline site have $w_\bullet^I = 0$, while species common to both sites have $w_\bullet^I = 1$; species present only at the comparison site do not contribute to this term. Non-random loss of high-functioning species reduces ecosystem function more than would random loss of average species, while non-random loss of low-functioning species reduces ecosystem function less than would random loss of average species (Fig. 2). We denote this term SCE_L because it refers to effects of non-random species loss from the baseline site. The SCE_L is formally analogous to the effect of natural selection on an evolving population (Fox 2006). For instance, non-random mortality of large-bodied individuals (a form of selection against large body size) will reduce the mean body size in the next generation. In contrast, mortality that is random with respect to body size generates no selection on size and would not affect mean body size in the next generation.



	$\text{SRE}_{:\text{L}}$	$\text{SRE}_{:\text{G}}$	$\text{SCE}_{:\text{L}}$	$\text{SCE}_{:\text{G}}$	CDE
	≠0	=0	=0	=0	=0
	=0	≠0	=0	=0	=0
	≠0	=0	≠0	=0	=0
	=0	≠0	=0	≠0	=0
	=0	=0	=0	=0	≠0

Figure 2. Illustration of maximally-simple limiting cases of the application of Eq. 1, in which as many terms as possible equal zero. These limiting cases help clarify the interpretation of the terms in Eq. 1. Within each row, the first column illustrates a hypothetical baseline site (open area) and comparison site (grey area). Filled and open circles represent different species, the size of the circle represents the value of the species' functional contribution, and shared species are connected by lines. The remaining columns indicate which terms in Eq. 1 are non-zero. More complex cases in which all terms are non-zero are illustrated in Fig. 4.

Note that the SCE_L necessarily is accompanied by, but is entirely distinct from, the SRE_L. Species loss necessarily generates a SRE_L (unless $\bar{z} = 0$), but may or may not generate a SCE_L, depending on whether or not species loss is non-random (Fig. 2). Schmid et al. (2002) argued that the effects of species richness and species composition are inherently confounded and so not completely separable, but this is only true when these effects are defined in certain ways (Fox 2006). Equation 1 defines all its effects in such a way that they are never confounded. The values of s, s', z_i, z_j' and w_j^i uniquely determine the values of every term in Eq. 1, with no between-site variation that cannot be uniquely attributed to one of those terms.

The fourth term in (1), $-\text{Sp}\left(w_J^\bullet, z'\right)$, is the mirror image of the SCE_L. This term captures the effect of non-random species gain in the comparison site, and so we denote this term SCE_G. Gained species are only present in the comparison site and have $w_J^\bullet = 0$, while species common to both sites have $w_J^\bullet = 1$; species present only at the baseline site do not contribute to this term. This term has a minus sign in front of it, so that non-random gain of high-functioning species increases total ecosystem function at the comparison site more than would random gain of average species. Conversely, non-random gain of low-functioning species increases total ecosystem function at the comparison site less than would random gain of average species (Fig. 2). The SCE_G is formally analogous to the effect of non-random immigration into an evolving population (Kerr and Godfrey-Smith 2009). For instance, if immigrants are larger on average than the resident population, the mean body size of the population will increase. The SCE_G necessarily is accompanied by, but is entirely distinct from, the SRE_G (Fig. 2).

The final term in Eq. 1, $\sum_{i=1}^{s_c} \Delta z_i$ is equivalent to the context dependence effect (CDE) of Fox (2006). This is the sum, over the species common to both sites, of the between-site differences in their functional contributions (Fig. 2). Species present only at one of the two sites do not contribute to this sum. Many underlying biological mechanisms can contribute to this sum, including between-site differences in both environmental conditions and species interactions. For example, if environmental conditions at the comparison site are poor compared to conditions at the baseline site, the species common to both sites might function worse at the comparison site, so that the CDE is < 0. As a second example, if the species at the baseline site are competitors, loss of some species might release the remaining species from competition, thereby allowing them to make increased functional contributions at the comparison site, so that the CDE is > 0. We denote this term as the context dependence effect because it captures the extent to which species' functional contributions vary between ecological 'contexts' (sites). The CDE is formally analogous to biased transmission in evolution. Biased transmission occurs when phenotypes of offspring deviate systematically from those of their parents (Frank 1997). For instance, an environmental change that causes lower growth in offspring than their parents will tend to reduce mean body size in the offspring generation compared to the parental generation, independent of selection or other evolutionary forces.

One way to clarify the interpretation of the final three terms in Eq. 1, and highlight that they are the only three factors which affect mean function per species, is to rewrite them as follows:

$$s_c\left(\bar{z}'-\bar{z}\right)=s_c\left(\bar{z}_c-\bar{z}\right)+\left[-s_c\left(\bar{z}'_c-\bar{z}'\right)\right]+s_c\left(\bar{z}'_c-\bar{z}_c\right) \quad (2)$$

where $\bar{z}_c = \frac{1}{s_c}\sum_{i=1}^{s_c} z_i$ and $\bar{z}'_c = \frac{1}{s_c}\sum_{i=1}^{s_c} z'_i$ respectively give the average pre-loss and post-loss functional contribution of the s_c species common to both sites. The term $s_c\left(\bar{z}_c-\bar{z}\right)$ equals the SCE_L, and captures whether the species common to both sites differ in their average functioning at the baseline site from all species at that site. If they do, this implies that the lost species comprise a non-random subset of all baseline site species. Similarly, the term $-s_c\left(\bar{z}'_c-\bar{z}'\right)$ equals the SCE_G. This term captures whether the species common to both sites differ in their average functioning at the comparison site from all species at that site. If they do, this implies that the gained species comprise a non-random subset of all comparison site species. Finally, the term $s_c\left(\bar{z}'_c-\bar{z}_c\right)$ equals the CDE. This term captures between-site variation in the average functioning of the species common to both sites.

Illustrative application

Next we apply the extended Price equation partition to data from the ongoing long-term N enrichment experiment at Cedar Creek, MN, USA (see Tilman 1987 for methodological details). As an illustration, we apply the extended Price equation partition to data from Field C in 1998, the 16th year of the experiment (an arbitrary choice made purely for illustrative purposes; these data are representative of the long-term outcome of the experiment in other fields). The data are available from <www.cedarcreek.umn.edu/research/data/>.

The ecosystem function of interest is total aboveground plant biomass. N enrichment leads to a curvilinear, saturating increase in total plant biomass, and a curvilinear decline in species richness (Fig. 3a). However, plots of the same enrichment level vary substantially in both species richness and total biomass (Fig. 3a). Plots of the same richness also vary in species composition (Fig. 3b). Although enriched plots have fewer species on average than unenriched plots, the diversity gradient is not strictly nested. Many enriched plots have species that are absent from some or all of the unenriched plots. Compositional similarity to the control plots declines in curvilinear fashion with increasing enrichment, indicating compositional turnover along the enrichment gradient (Fig. 3b).

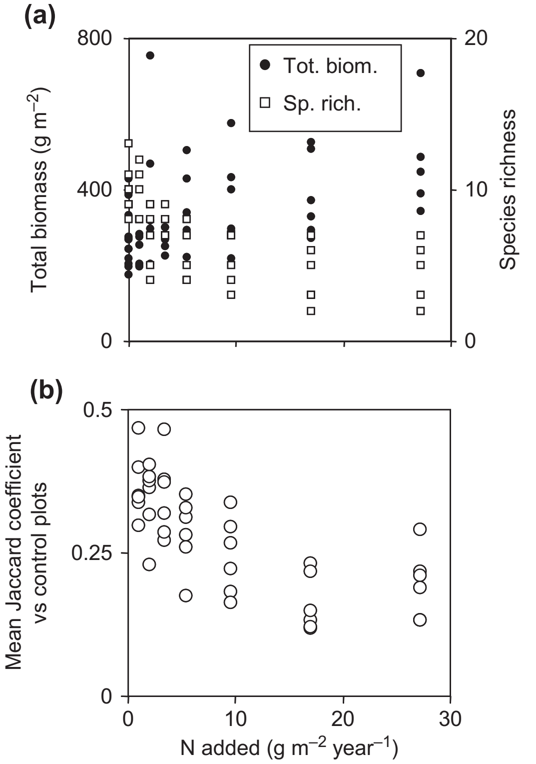


Figure 3. (a) Total aboveground plant biomass and plant species richness as a function of the annual rate of N addition. (b) Mean Jaccard coefficient between each N-enriched plot and the n = 12 unenriched plots, as a function of the annual rate of N addition. The Jaccard coefficient is a measure of similarity of species composition.

Previous analyses of this and other N enrichment experiments focus on statistically describing and explaining the effects of enrichment on plant species richness and composition (Tilman 1987, Clark et al. 2007). Here we seek to understand how loss and gain of species, together with between-plot variation in the biomass of persisting species, generates variation in total plant biomass along the enrichment gradient. The answer to this question is not at all obvious, due to the substantial variation in species richness, composition, and total biomass within and among enrichment levels. Indeed, because species richness and composition change in a correlated fashion along the enrichment gradient (compare Fig. 3a and 3b), their separate effects on total plant biomass might be thought impossible to tease apart.

We designated the unenriched control plots as the baseline sites, to which all of the enriched plots were compared. We compared each enriched plot to each of the n = 12 unenriched plots. To provide a compact summary of the results, we averaged the results for each enriched plot across the unenriched plots, as in Fox (2006).

The SRE_G is always positive and the SRE_L is always negative, as they must be given that species' biomasses cannot be negative (Fig. 4a). Interestingly, the SRE_G generally is equal or larger in absolute magnitude than the SRE_L (Fig. 4a), even though many more species are lost than gained with increasing

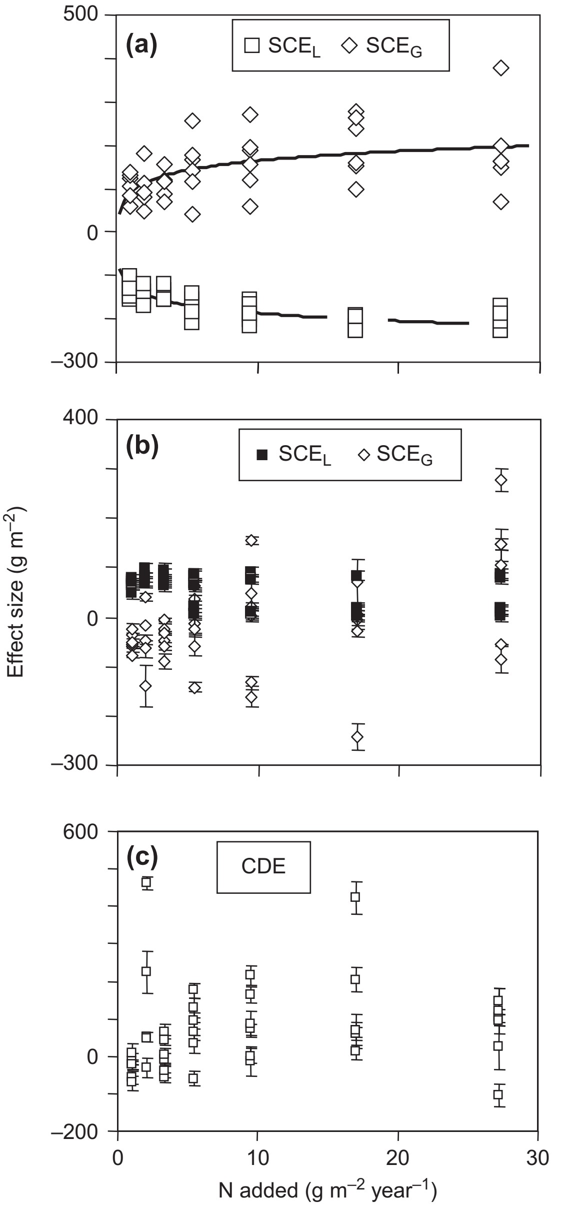


Figure 4. Effect sizes in the extended Price equation partition as a function of the annual rate of N addition. Each point gives the mean ± SE for one N addition plot. Error bars, which are sometimes too small to be visible, reflect variation in species presence/absence and biomasses among the control (baseline) plots to which the N addition (comparison) plots are compared. (a) The SRE$_L$ and SRE$_G$. Regression lines of the form $y = a + b\ln(x)$ are shown and are significant at the $p < 0.05$ level. (b) The SCE$_L$ and SCE$_G$. (c) The CDE.

enrichment (Fig. 3a). The SRE$_G$ is larger in absolute magnitude when the mean biomass per gained species in enriched plots exceeds the mean biomass per lost species in the unenriched plots. The SRE$_G$ increases asymptotically with increasing enrichment, while the SRE$_L$ decreases asymptotically (Fig. 4a). These curvilinear trends in the SRE$_G$ and SRE$_L$ reflect the curvilinear trend in species richness along the enrichment gradient. That the trends in the SRE$_G$ and the SRE$_L$ mirror each other does not mean they are confounded (all the terms in Eq. 1 necessarily are orthogonal). Rather, the ecology of this particular system causes the SRE$_G$ and SRE$_L$ to vary in correlated fashion along the enrichment gradient.

The SCE$_L$ is positive on average (Fig. 4b), indicating that enrichment leads to loss of species that produce below-average biomass in unenriched plots. Non-random species loss with increasing enrichment therefore reduces total biomass less than would random species loss. Interestingly, the SCE$_L$ is independent of enrichment (Fig. 4b), despite the fact that increasing enrichment leads to increasing shifts in species composition (Fig. 3b). That the SCE$_L$ is relatively independent of enrichment indicates that increasing enrichment does not lead to increasingly non-random species loss.

The SCE$_G$ varies greatly, even among equally-enriched comparison plots (Fig. 4b). The SCE$_G$ increases with increasing enrichment, taking on negative values at low enrichment and largely positive values at high enrichment (Fig. 4b). This indicates that gained species often attain below-average biomass at low levels of enrichment compared to other species, but tend to attain higher-than-average biomass at high levels of enrichment. In particular, *Agropyron repens* attains low biomass at low enrichment but can become a dominant species at high enrichment.

The CDE is positive on average, but generally takes on negative values at the lowest enrichment level (Fig. 4c). This pattern of variation in the CDE is surprising, because it indicates that species present in both unenriched and enriched plots exhibit a U-shaped response to enrichment: they generally exhibit reduced biomass at low enrichment compared to unenriched plots, but increased biomass at high enrichment. The CDE also varies greatly among comparison plots at higher enrichment levels, and negative values can occur even at very high enrichment (Fig. 4c).

The CDE may vary among post-loss plots because of variation in the identity of the lost and gained species. For instance, it might be expected that gain of highly-competitive, high-biomass species would reduce the biomass of the persisting species, while loss of highly-competitive, high-biomass species would have the opposite effect (Fox and Harpole 2008). The results partially support this hypothesis: the CDE is significantly negatively correlated with the SCE$_G$ (Fig. 5a), as expected if gain of high-biomass species reduces the biomass of the persisting species, while gain of low-biomass species permits increased biomass of the persisting species in response to increased enrichment. The CDE is not significantly correlated with the SCE$_L$, and the trend is in the opposite of the expected direction (Fig. 5b), indicating that losing high-biomass species does not increase the biomass of the persisting species. Lack of a significant relationship between the SCE$_L$ and the CDE contrasts with the results of the Biodiversity II experiment at Cedar Creek, which simulated species loss while holding other factors constant (Fox and Harpole 2008). Possibly, the effects of species gain and enrichment swamp the effects of species loss on the persisting species. Note that SCE$_G$ and SCE$_L$ are not themselves significantly correlated ($r = 0.12$).

Biodiversity and abiotic environmental conditions sometimes have been posed as distinct drivers of ecosystem function, and questions asked about their relative importance. For instance, Loreau et al. (2001) suggests that biodiversity strongly affects ecosystem function at small spatial scales, while abiotic environmental conditions drive variation in function at large spatial scales. But for many ecosystem

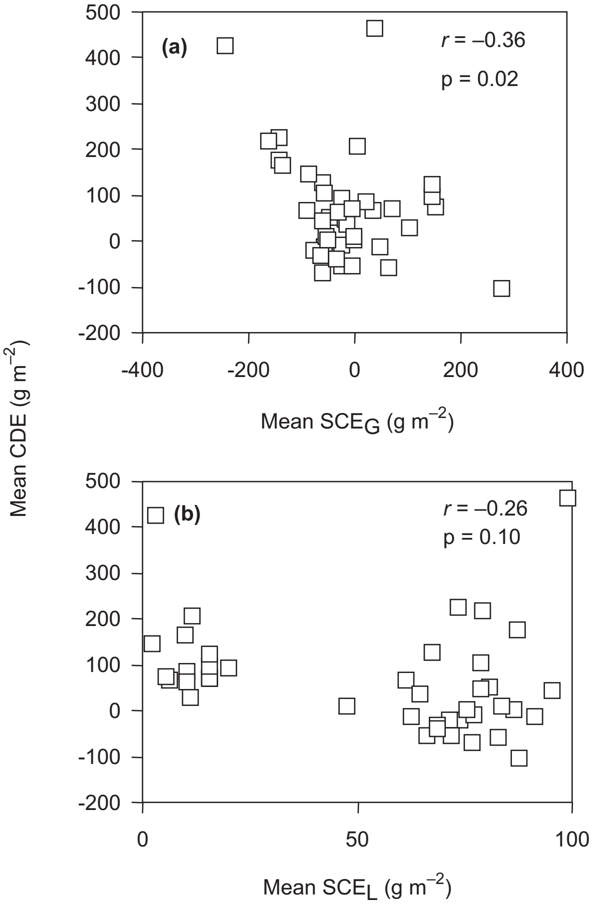


Figure 5. The CDE as a function of (a) the SCE_G and (b) the SCE_L. Each point gives means for one N addition plot. Pearson's correlation coefficient (*r*) is shown. The CDE correlated significantly with the SCE_G, but not with the SCE_L.

functions, including total biomass, changes in abiotic conditions only affect ecosystem function via their direct and indirect effects on the species performing the function. Here, we have used the extended Price equation partition to separate the distinct 'routes' by which enrichment affects total biomass. While all five terms in the extended Price equation partition are non-negligible in absolute magnitude in the case examined here, the strongest effects of enrichment on total biomass often come about via context dependence. This is a surprising result, given the dramatic effects of enrichment on species richness and composition in this study (Tilman 1987). Nor does enrichment necessarily lead to positive context dependence: negative CDEs occur. Negative CDEs apparently arise at least in part from gains of highly-competitive species, which suppress the persisting species. The extended Price equation partition reveals these patterns because it makes full use of information about the performance of individual species at specific sites, rather than averaging away this information in order to obtain a more compact but less informative summary of the data.

It would have been difficult to recover these results from a conventional statistical analysis. Species richness, species composition, and the performance of individual species all covary along the enrichment gradient, so that their separate effects necessarily are at least partially confounded from the perspective of conventional statistical approaches such as general linear models (Tilman and Downing 1994). Further, the results of a conventional statistical analysis would be sensitive to the choice of predictor variables and statistical model, a choice that necessarily has an arbitrary element (Schmid et al. 2002, Bell et al. 2009). In contrast, the extended Price equation partition provides a complete, exact description of all possible sources of variation in ecosystem function, ensuring that no effects are omitted and no redundant terms included.

It would be interesting to know if these results generalize to other ecosystem functions and other sources of variation in species richness and composition. The effects of different factors on species richness and composition, and on ecosystem function, often are studied separately. The extended Price equation partition provides a way to unify the study of different special cases within a more general framework.

Discussion

Species loss and gain have distinct effects

Motivated by concerns about the effects of species loss, theoretical and empirical studies of biodiversity and ecosystem function have focused on nested species compositions (Spehn et al. 2005, Fox 2006, Balvanera et al. 2006, Cardinale et al. 2009). A key goal of these studies has been to separate the effect of species richness from that of species composition. But in nature biodiversity gradients rarely are nested; species gain often accompanies species loss. The key insight of the present work is that the conceptual framework of previous studies of biodiversity and ecosystem function is seriously incomplete. Species loss and gain are distinct, 'mirror image' processes (Fig. 1, and Kerr and Godfrey-Smith 2009). For this reason, when species are both lost and gained, there is no single effect of species richness on ecosystem function, and no single effect of species composition. Rather, there are two of each. This perspective highlights the value of the formal theoretical framework provided by Eq. 1. Previous discussions of the effects of species richness vs. composition assume that each is (or can be summarized usefully as) a single effect, even in the context of non-nested biodiversity gradients (Tilman and Downing 1994, Schmid et al. 2002). Effects of species gain and loss must be distinguished from one another because their randomness or non-randomness must be judged against different reference standards. Gained and lost species must be compared to others growing at the same site, otherwise differences in functioning among species (the SRE and SCE) could not be separated from differences in functioning between sites (the CDE). While the SRE_L and SRE_G could be summed to obtain a single 'overall' effect of species richness, and the SCE_L and SCE_G could be summed to obtain a single 'overall' effect of species composition, no ecological insight would be gained by doing so (in fact, insight would be lost).

Our approach takes advantage of the fact that species gain and loss are "mirror images" by using the indexing variable w_j^i to define a 1:1 correspondence between the species present in or absent from one site, and the species present in or absent from the other site (Price 1995, Kerr and Godfrey-Smith 2009). This 1:1 correspondence is what allows a distinction to be drawn between random loss or gain of species (SRE_L, SRE_G), and non-random loss or gain of species (SCE_L, SCE_G). The 1:1 correspondence between the species comprising the two sites is somewhat like the relationship between two paired samples in statistics. In statistics, analyzing paired samples as if they were unpaired discards useful information and results in a less-powerful test. Here, retaining the information about which species were lost or gained, rather than simply treating the two sites as two different 'unpaired' species compositions, allows the SRE, SCE, and CDE to be distinguished (Fox 2006, Hector et al. 2009).

Fox and Harpole (2008, their Appendix B) attempted to compare both gained and lost species to the species at the baseline site. Using the baseline site as a 'reference standard' for both gained and lost species treats species gain and loss asymmetrically, and makes it impossible to determine the randomness of both species gain and loss. Fox and Harpole (2008) incorrectly concluded that the Price equation partition of Fox (2006) could not be extended to cover simultaneous species gain and loss.

Note that, in describing species gain and loss as mirror-image processes, we make no claims about the underlying biological mechanisms causing species gain and loss. The underlying biological mechanisms causing species loss (e.g. competitive exclusion, disease epidemics, disturbances, demographic stochasticity) are not necessarily mirror images of the mechanisms causing species gain (e.g. invasion, experimental introduction). The Price equation is not a model of these underlying mechanisms. Rather, it is a higher-level summary of the ecosystem-level effects these underlying mechanisms necessarily must have. Similarly, in evolutionary biology one term of the Price equation summarizes the effects of all underlying biological mechanisms that result in covariation between organism phenotype and organism fitness. This covariation is natural selection, no matter what underlying biological mechanisms give rise to it. And it is only by generating natural selection, and/or one of the other forces identified by the Price equation, that any underlying biological mechanism can affect the direction of evolution.

On the interpretation of the Price equation partition

Our extended Price equation partition is simply a mathematical identity. At first glance, it resembles a conventional mathematical model, or a statistical model such as a general linear model. Unfortunately, both of these resemblances are only superficial, and are liable to be misleading. Here we discuss some key issues related to the interpretation of our approach. Some of these issues, and others, are discussed further elsewhere (Fox 2006, 2010, Fox and Harpole 2008, Rice and Papadopoulos 2009).

Equation 1 comprises the sum of five terms, but it was not derived by assuming that it would comprise additive terms (Kerr and Godfrey-Smith 2009). Rather, the additivity of the partitioned effects emerges naturally from the definition of the problem (Rice and Papadopoulos 2009). The approach developed here is consistent with the operation of complex, non-linear, non-additive mechanisms affecting species richness, composition, and species' functional contributions, including both natural and artificial mechanisms (experimental manipulations). Our approach simply summarizes the functional effects of all these complex mechanisms in a useful way.

Equation 1 is exact. There is no possibility that any factors are omitted, not even residual error. This allows Eq. 1 to solve the conceptual problem of explaining between-site variation in ecosystem function in a more complete and convincing way than would be possible by attempting to generalize from conventional models based on simplifying assumptions. The Price equation is exact because it is retrospective: it considers only the known, observed outcomes of past events. Rice (2008) derived a prospective equivalent of the Price equation which incorporates uncertainty arising from the fact that future data have not yet been observed and so must be characterized by probability distributions.

In reality, species' z and z' values rarely are known without error even in retrospect, because of sampling error. However, sampling error is a feature of the data input into the Price equation, not of the Price equation itself (Gardner and Grafen 2009). Whether or not the data input into the Price equation are viewed as complete census data, or as random sample data, is a separate question. If the data are viewed as sample data, then sampling error can be reduced in the usual way: by taking replicate samples and averaging the results. For instance, we calculated means and standard errors by averaging across control plots, which we treated as replicate baseline sites (Fig. 4).

Each effect in Eq. 1 can be interpreted as identifying what the total difference in ecosystem function between two sites would be, 'all else being equal', so that other effects equal zero. In partitioning a total effect in this way, our approach appears analogous to familiar statistical approaches such as ANOVA. However, it is important to recognize that the extended Price equation partition relies on a stricter interpretation of the condition 'all else being equal' than is typical in conventional statistics. For instance, consider a hypothetical experiment in which two identical baseline sites (i.e. two baseline sites with the same species, where each species has exactly the same z_i value at both sites) are subjected to different treatments. At one site, a species is added, while the other site is left as an unmanipulated control. After some period of time, the functioning of each species is remeasured (this is an unreplicated experiment, but replication is not relevant to the point at hand). From an experimental perspective, this experiment isolates the effect of species gain on ecosystem function, in the sense that the only difference between the two experimentally-imposed treatments is that one of the treatments includes a gained species. It might be thought that if these two treatments were each analyzed with the extended Price equation partition, they would differ only in the SRE_G and SCE_G, as these are the two terms in the extended Price equation partition that isolate the effects of species gain. But the CDE also could differ between treatments because species gain might alter the functioning of the remaining species (see 'Illustrative application' for an example). The difference between the two experimental treatments therefore does not reveal what would happen to ecosystem function if a species were gained and 'all else was equal'. Our approach uses a strict interpretation of the phrase 'all else being equal' to distinguish effects propagated through different causal pathways. In this hypothetical experiment, species gain affects ecosystem function directly via the SRE_G and SCE_G, and indirectly via its effect on the performance of the persisting species. Equation 1 separates these direct and indirect effects rather than lumping them together.

The importance of distinguishing 'summed' ecosystem functions from other types

Like the Price equation partition of Fox (2006), the extended approach developed here only applies to ecosystem functions comprising the summed contributions of the species performing the function. This limitation arises because the approach treats the functional contribution of a species as formally analogous to the phenotype of an individual organism (Fox 2006). Many ecosystem functions, not just total biomass or primary productivity, meet this assumption. For instance, total CO_2 flux is sum of the CO_2 fluxes of all the species that produce and/or take up CO_2 (McGrady-Steed et al. 1997). The total amount of dung buried by dung beetles over some period of time is the sum of the amounts of dung buried by each dung beetle species (Larsen et al. 2005). Many other examples could be given. In all these cases, the functioning of each species is measurable (at least in principle), so that total function could be calculated by measuring the functioning of each species and then summing across all species. It is necessary to define the functional contribution of an individual species in order to isolate the effect of losing or gaining that species, all else being equal. This is not to say that 'non-summed' ecosystem functions are rare or unimportant, or that they cannot be affected by species loss or gain, merely that they differ in key respects from summed functions. For instance, if the species of interest do not perform the function of interest, but only affect function indirectly via their interactions with the species performing the function of interest, our approach does not apply. See Appendix A of Fox and Harpole (2008) and Appendix 5 of Fox (2010) for further discussion of the distinction between 'summed' and 'non-summed' functions.

Even with 'non-summed' ecosystem functions, it is still possible to measure a statistical association between total functioning, and species richness or the presence/absence of particular species. Various statistical methods have been proposed to estimate such associations (Tilman and Downing 1994, McGrady-Steed et al. 1997, Hille Ris Lambers et al. 2004, Bell et al. 2009, Hector et al. 2009). Such statistical associations can provide valuable clues to the ecological mechanisms by which species richness and composition affect ecosystem function. However, it is important to recognize that similar statistical patterns can arise for distinct mechanistic reasons. For instance, an increasing saturating relationship between species richness and various ecosystem functions arises in studies to which our approach applies (e.g. effect of plant diversity on total plant biomass in Tilman et al. 2001 and Spehn et al. 2005), and in studies to which it does not apply (e.g. effects of protist diversity on decomposition rate and invasibility in McGrady-Steed et al. 1997). Different effects besides those identified by Eq. 1 can give rise to commonly-observed statistical relationships between biodiversity and ecosystem function. When different effects are at work in different studies, any statistical similarity in the biodiversity–function relationship across different studies becomes difficult to interpret.

Other authors emphasize different distinctions between ecosystem functions. For instance, some authors suggest that standing stocks will respond differently to changes in diversity than will rates of production (Petchey 2003, Balvanera et al. 2006). The appropriate classification of ecosystem functions depends on the purpose of the classification. Distinguishing between stocks and rates may be appropriate for purposes of empirical prediction. Distinguishing summed functions (which include both stocks and rates) from non-summed functions is appropriate for purposes for developing a general theory of biodiversity and ecosystem function.

Conclusion: the value of general theoretical frameworks

The Price equation is not a conventional statistical or mathematical model, and so was little-understood for many years. Price (1970) was initially rejected by Nature (Harman 2010). The Price equation has subsequently proven its value in evolutionary biology (Hamilton 1970, Frank 1997, Rice 2004, Okasha 2006, Gardner 2008, Kerr and Godfrey-Smith 2009, Rice and Papadopoulos 2009), and has been applied to problems in fields ranging from economics (Andersen 2004) to linguistics (Jäger 2008) to computer science (Poli et al. 2010) to archaeology (Brantingham and Perrault 2010) to population biology (Schauber et al. 2007, Rebke et al. 2010). That it has been found so useful in so many contexts is a strong prima facie argument for its utility in analyzing ecosystem function. The application presented here, along with other recent applications (Fox 2006, 2010, Fox and Harpole 2008, Collins and Gardner 2009), have generated novel insights and show that this prima facie argument is borne out in practice.

Conceptual and empirical progress in many areas of science often has hinged on development of general theoretical frameworks, which organize and subsume special cases that would otherwise be loosely related at best (Chesson 2000). The theory of evolution by natural selection is a familiar example – indeed, so familiar that it is easy to forget its organizing role. Without it, there would be no apparent connection between, say, changes over time in the prevalence of bacterial antibiotic resistance, and in the armor plating of sticklebacks (Baquero and Blázquez 1997, Bell et al. 2004). Nor would there be any apparent distinction between what we now call evolutionary change and what we now call developmental change. After all, both involve 'change over time' and in the Victorian era they were often regarded as analogous (Bowler 1983). It remains to be seen whether the extended Price equation partition can be subsumed by an even more-general theory of ecosystem function.

Acknowledgements – This work was supported by an NSERC Discovery Grant to JWF. Thanks to David Tilman for permission to use the data.

References

Andersen, E. S. 2004. Population thinking, Price's equation, and the analysis of economic evolution. – Evol. Inst. Econ. Rev. 1: 127–148.

Balvanera, P. et al. 2006. Quantifying the evidence for biodiversity effects on ecosystem functioning and services. – Ecol. Lett. 9: 1146–1156.

Baquero, F. and Blázquez, J. 1997. Evolution of antibiotic resistance. – Trends Ecol. Evol. 12: 482–487.

Bell, M. A. et al. 2004. Twelve years of contemporary armor evolution in a threespine stickleback population. – Evolution 58: 814–824.

Bell, T. B. et al. 2009. A linear model method for biodiversity-ecosystem functioning experiments. – Am. Nat. 174: 836–849.

Bowler, P. J. 1983. Evolution: the history of an idea. – Univ. of California Press.

Brantingham, P. J. and Perreault, C. 2010. Detecting the effects of selection and stochastic forces in archaeological assemblages. – J. Archaeol. Sci. 37: 3211–3225.

Cardinale, B. J. et al. 2009. Effects of biodiversity on the functioning of ecosystems: a summary of 164 experimental manipulations of species richness. – Ecology 90: 854.

Chesson, P. 2000. Mechanisms of maintenance of species diversity. – Annu. Rev. Ecol. Syst. 31: 343–366.

Clark, C. M. et al. 2007. Environmental and plant community determinants of species loss following nitrogen enrichment. – Ecol. Lett. 10: 596–607.

Collins, S. and Gardner, A. 2009. Integrating physiological, ecological, and evolutionary change: a Price equation approach. – Ecol. Lett. 12: 744–757.

Fox, J. W. 2006. Using the Price equation to partition the effects of biodiversity loss on ecosystem function. – Ecology 87: 2687–2696.

Fox, J. W. 2010. Partitioning the effects of species loss on community variability using multi-level selection theory. – Oikos 119: 1823–1833.

Fox, J. W. and Harpole, W. S. 2008. Revealing how species loss affects ecosystem function: the trait-based Price equation partition. – Ecology 89: 269–279.

Frank, S. A. 1997. The Price equation, Fisher's fundamental theorem, kin selection, and causal analysis. – Evolution 51: 1712–1729.

Gardner, A. 2008. The Price equation. – Curr. Biol. 18: R198–R202.

Gardner, A. and Grafen, A. 2009. Capturing the superorganism: a formal theory of group adaptation. – J. Evol. Biol. 22: 659–671.

Gonzalez, A. G. and Loreau, M. 2010. The causes and consequences of compensatory dynamics in ecological communities. – Annu. Rev. Ecol. Evol. Syst. 40: 393–414.

Gonzalez, A. G. et al. 2009. Biodiversity as spatial insurance: the effects of habitat fragmentation and dispersal on ecosystem functioning. – In: Naeem, S. et al. (eds), Biodiversity, ecosystem functioning, and human wellbeing. Oxford Univ. Press, pp. 134–146.

Hamilton, W. D. 1970. Selfish and spiteful behaviour in an evolutionary model. – Nature 228: 1218–1220.

Harman, O. 2010. The price of altruism: George Price and the search for the origins of kindness. – W. W. Norton, New York.

Hector, A. et al. 2009. The analysis of biodiversity experiments: from pattern toward mechanism. – In: Naeem, S. et al. (eds), Biodiversity, ecosystem functioning, and human wellbeing. Oxford Univ. Press, pp. 94–104.

Hille Ris Lambers, J. et al. 2004. Mechanisms responsible for the positive diversity-productivity relationship in Minnesota grasslands. – Ecol. Lett. 7: 661–668.

Jäger, G. 2008. Language evolution and George Price's "General theory of selection". – In: Cooper, R. and Kempson, R. (eds), Language in flux: dialogue coordination, language variation, change and evolution. King's College, London, pp. 53–80.

Kerr, B. and Godfrey-Smith, P. 2009. Generalization of the Price equation for evolutionary change. – Evolution 63: 531–536.

Larsen, T. H. et al. 2005. Extinction order and altered community structure rapidly disrupt ecosystem functioning. – Ecol. Lett. 8: 538–547.

Loreau, M. et al. 2001. Biodiversity and ecosystem functioning: current knowledge and future challenges. – Science 294: 804–808.

McGrady-Steed, J. et al. 1997. Biodiversity regulates ecosystem predictability. – Nature 390: 162–165.

Okasha, S. 2006. Evolution and the levels of selection. – Oxford Univ. Press.

Petchey, O. L. 2003. Integrating methods that investigate how complementarity influences ecosystem functioning. – Oikos 101: 323–330.

Poli, R. et al. 2010. Theoretical results in genetic programming: the next ten years? – Genetic Programming and Evolvable Machines 11: 285–320.

Price, G. R. 1970. Selection and covariance. – Nature 227: 520–521.

Price, G. R. 1972. Extension of covariance selection mathematics. – Ann. Hum. Genet. 35: 485-489.

Price, G. R. 1995. The nature of selection. – J. Theor. Biol. 175: 389–396.

Rebke, M. et al. 2010. Reproductive improvement and senescence in a long-lived bird. – Proc. Natl Acad. Sci. USA 107: 7841–7846.

Rice, S. H. 2004. Evolutionary theory: mathematical and conceptual foundations. – Sinauer.

Rice, S. H. 2008. A stochastic version of the Price equation reveals the interplay of deterministic and stochastic processes in evolution. – BMC Evol. Biol. 8: 262.

Rice, S. H. and Papadopoulos, A. 2009. Evolution with stochastic fitness and stochastic migration. – PLoS One 4: e7130.

Schauber, E. M. et al. 2007. Spatial selection and inheritance: applying evolutionary concepts to population dynamics in heterogeneous space. – Ecology 88: 1112–1118.

Schmid, B. et al. 2002. The design and analysis of biodiversity experiments. – In: Loreau, M. et al. (eds), Biodiversity and ecosystem functioning: synthesis and perspectives. Oxford Univ. Press, pp. 61–78.

Spehn, E. M. et al. 2005. Ecosystem effects of biodiversity manipulations in European grasslands. – Ecol. Monogr. 75: 37–63.

Tilman, D. 1987. Secondary succession and the pattern of plant dominance along experimental nitrogen gradients. – Ecol. Monogr. 57: 189–214.

Tilman, D. and Downing, J. A. 1994. Biodiversity and stability in grasslands. – Nature 367: 363–365.

Tilman, D. et al. 2001. Diversity and productivity in a long-term grassland experiment. – Science 294: 843–845.

Wright, D. H. et al. 1998. A comparative analysis of nested subset patterns of species composition. – Oecologia 113: 1–20.