Modeling How Community Assembly Alters the Functioning of Ecosystems

Authors and Affiliations

Koffel, Thomas, W.K. Kellogg Biological Station, Michigan State University, 3700 E Gull Lake Dr., Hickory Corners, MI 49060. U.S.A. koffelth@msu.edu

Kremer, Colin T, W.K. Kellogg Biological Station, Michigan State University, 3700 E Gull Lake Dr., Hickory Corners, MI 49060. U.S.A. kremerco@msu.edu

Bannar-Martin, K. kbannarm@gmail.com

Ernest, S.K. Morgan, Department of Wildlife Ecology and Conservation, University of Florida, Gainesville, FL 32611, USA.

Eisenhauer, Nico, [1] German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Deutscher Platz 5e, 04103 Leipzig, Germany; [2] Institute of Biology, Leipzig University, Deutscher Platz 5e, 04103 Leipzig, Germany. nico.eisenhauer@idiv.de

Roscher, Christiane, [1] Helmholtz Centre for Environmental Research (UFZ), Department Physiological Diversity, Permoserstrasse 15, 04318 Leipzig, Germany; [2] German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Deutscher Platz 5e, 04103 Leipzig, Germany. christiane.roscher@ufz.de

Cabral, Juliano Sarmento, Ecosystem Modeling, Center for Computational and Theoretical Biology (CCTB), University of Würzburg, Emil-Fischer-Str. 32, 97074 Würzburg, Germany. juliano.sarmento_cabral@uni-wuerzburg.de

Leibold, Mathew A., Department of Biology, University of Florida, Gainesville, FL 32611, USA. mleibold@ufl.edu

Abstract:

Although the effects of species richness on ecosystem functioning have been extensively studied, there is increased interest in understanding how community assembly in a more general sense can alter the functioning of ecosystems. We identify two complementary approaches for evaluating how community assembly influences biomass production. The first quantifies the relative importance of complementarity and selection by contrasting monocultures with polycultures. The second identifies the effects of species replacements (losses and/or gains of species relative to the baseline polyculture) as well as the indirect effects on other species' productivity. We compare and contrast these two approaches using simulated communities where species compete for different resource types and experience different levels of environmental heterogeneity and find that these metrics can jointly diagnose the mechanisms of competition driving productivity. We apply our methods to data from a long-term biodiversity-ecosystem experiment (the Jena Experiment) and find that the data do not correspond to any of the competition scenarios we modeled. We address two additional possible complications: facilitation by nitrogen fixing plants, and non-equilibrium behavior during community assembly, and find that facilitation/competition by nitrogen fixing plants is the more likely explanation for the results obtained at Jena.

Introduction:

Community assembly drives shifts in community structure (i.e., species richness and composition) which can have important consequences for the functioning of ecosystems. While the impact of changes in species richness on ecosystem function (commonly called biodiversity-ecosystem function, or BEF, effects; reviewed by Cardinale et al. 2012, Hooper et al. 2012, Tilman et al. 2012) are the best-studied, the importance of all of community assembly's impacts on community structure (referred to here as community assembly-functioning of ecosystem, or CAFE, effects) is increasingly recognized (e.g. Ptacnik et al. 2010, Thompson and Gonzalez 2016, Hodapp et al. 2016, Leibold et al. 2017, Bannar-Martin et al. 2018). Drivers of community assembly include species interactions, disturbance, dispersal, environmental heterogeneity, and landscape effects that can lead to changes in community composition, without necessarily affecting species richness (Brown et al. 2001). Community assembly is thus complex and depends on multiple factors and processes that may also consequently affect ecosystem function (e.g. Thompson and Gonzalez 2016, Hodapp et al. 2016, Leibold et al. 2017). Therefore, it is community assembly processes as a whole that ultimately link community composition, including species richness, to ecosystem attributes.

How can we study these community assembly effects on ecosystem function? One approach would be to comprehensively study the detailed interactions among the species involved and theoretically analyze these interactions to fully disentangle how community assembly affects ecosystem attributes. This approach, however, is logistically demanding and likely to give answers that are system specific. An alternative is to find general metrics that parse the effects of various components of community assembly on ecosystem attributes. Several methods are available (e.g. Kirwan et al. 2009, Bell et al. 2009, Hector et al. 2009, Bannar-Martin et al. 2018); in this paper we focus on two of these. The first, which we call the Complementarity-Selection Effects method (originally called the 'additive partitioning method'), is based on comparing the contributions of species to ecosystem function when they

are in monocultures to those that emerge due to species interactions in polycultures (Loreau and Hector 2001). Such comparisons allow total ecosystem functioning in a community to be partitioned into the function that results from selection effects (when species that are most productive in monocultures come to dominate multispecies assemblages and their functioning) from what is due to complementarity effects (when productivity in multispecies communities is higher than expected from proportional monoculture contributions due to more complex species interactions such as niche partitioning or facilitation). Both of these effects can be positive (e.g. when selection favors the more productive monocultures or when complementarity results in enhanced production relative to that expected or when there is facilitation among species) or negative (when the reverse is true). It is generally thought that selection effects occur due to environmental filtering which selects for more productive species that are better suited to the local environment, whereas complementarity is related to niche partitioning that results in enhanced coexistence and increased resource-use efficiency. This approach is carefully tailored to the way most BEF experiments are actually conducted because BEF experiments require data on monocultures for all the component species and they use null models to compare final outcomes of polycultures to these monocultures instead of comparing them to other multispecies communities that may alternatively occur.

Another approach for linking community assembly effects with ecosystem function is based on pairwise comparisons of communities and quantifies how differences in total function relate to changes in composition and individual species' functions by calculating 'Price components' (Fox 2006, Fox and Kerr 2012, Bannar-Martin et al. 2018). Although this approach is constrained to ecosystem functions that can be calculated as the sum of contributions of their constituent species (Fox 2006), it has the advantage that it can be more generally used to look how <u>any</u> change in community structure (i.e., species richness, taxonomic turnover, changes in resident abundances or function) alters ecosystem functioning. Calculating Price components can be used flexibly to make inferences based on complex pairwise

comparisons (Bannar-Martin et al. 2018) or on temporal effects related to environmental change (Kremer et al. in prep). Here we will focus on comparisons between the communities with the most diverse species pool and those with lower levels of richness. There are many different ways the Price components can be configured (Bannar-Martin et al. 2018), but we employ a three-part arrangement that separately evaluate the functional effects of species gains ('SG'), species losses ('SL') and the differences in density/function among species present in both communities being compared (the Context Dependent Effect or 'CDE'). In Bannar-Martin et al. (2018) these are called the Community Assembly components; we use this terminology in the rest of this paper.

The Complementarity-Selection Effects (Loreau and Hector 2001) and Community Assembly components methods (Fox and Kerr 2012, Bannar-Martin et al. 2018) provide complementary information about the impact of community assembly on ecosystem function. The first focuses on mechanisms related to environmental filtering and niche relations that affect dominance and coexistence, whereas the second focuses on critical aspects of community assembly involving the loss and gain of species as well as the consequent effects on species that remain in the community. We study how the two approaches might be combined and compared to allow us to understand underlying mechanisms by applying these methods to a community model we constructed based on resource-competition theory (Tilman 1982, Tilman et al. 1997, Schreiber and Tobiason 2003, Gross and Cardinale 2007, Hodapp et al. 2016, Koffel et al. 2016). To model the community assembly process, we viewed a BEF experiment as having the following sequence of events: a) there is a filter on the species pool (taken from the total species pool of the experiment), that is imposed by the researchers via selective seeding and/or weeding, b) interactions among the remaining species lead to some of the species going extinct, and c) there are density-dependent responses among the remaining species that then determine ecosystem attributes. We then used this mathematical model to solve the assembly of communities from all possible subsets of a given total species pool, imitating a standard BEF experiment. We applied the model to various assembly scenarios to

consider a wide range of possible ways resource competition may operate to affect community assembly in a BEF experiment - including competition for complementary vs. antagonistic resources with homogeneous or heterogeneous resource supply at the local scale. We apply the Complementarity-Selection Effects method (Loreau and Hector 2001) and the Community Assembly components method (Bannar-Martin et al. 2018) to the results of the models to assess if these methods can be used to distinguish among the various assembly scenarios. Finally, we compare our theoretical findings to analyses of data from the long term BEF experiment conducted in grasslands at Jena (Roscher et al. 2004, 2005; Weigelt et al. 2010) and find that the data do not conform to any of the scenarios we initially modelled. We consequently modelled two other scenarios that we hypothesized might be important, inadequate time to achieve steady-state competitive outcomes and facilitation by nitrogen fixing plants to illustrate how we can improve our understanding of how community assembly affects the functioning of ecosystems and find that the data conform to results expected if facilitation is important but not to lack of steady-state competition outcomes.

Methods:

To model community assembly in BEF experiments, we used basic resource competition models for two resources as initially developed by MacArthur (1972) and further elaborated by Tilman (1982):

$$\frac{dB_j}{dt} = [g_j(R_1, R_2) - m_j]B_j$$

$$\frac{dR_i}{dt} = l_i(S_i - R_i) - \sum_{j=1}^{n} C_{ij}(R_1, R_2)B_j \qquad (1)$$

This model describes the population dynamics of n species j (whose biomasses are denoted B_j) that compete for 2 resources, R_1 and R_2. Here $g_j(R_1, R_2)$ and m_j respectively determine species j's per capita growth and mortality rates, S_i is the

supply of resource R_i, which is diluted at rate l_i, and $C_{ij}(R_1, R_2)$ is the per capita consumptive impact of B_j on R_i. Species j's interactions with the two resources can be divided in two parts: (1) its requirements (through the per capita growth rate g_j) and; (2) its impacts on the resources (through the per capita consumptive impacts C_{ij}).

Competitive Scenarios: We considered a range of competitive scenarios, classified by the types of resources involved, including resources that are 'strongly complementary' (virtually 'essential', *sensu* Tilman 1982) vs. 'weakly complementary' and 'weakly antagonistic' (these two cases correspond closely to 'substitutable' in Tilman's terminology, the perfectly substitutable model has undesirable mathematical properties for our purpose) vs. 'strongly antagonistic'. To implement this variation in resource types we used a flexible form of the $g_j(R_1, R_2)$ function developed by Schreiber and Tobiason (2003):

$$g_j(R_1, R_2) = [(x_{1j}R_1)^\alpha + (x_{2j}R_2)^\alpha]^{\frac{1}{\alpha}} \qquad (2)$$

Here α is a tunable shape parameter that determines the type of resource (i.e. complementary or antagonistic). The traits x_{1j} and x_{2j} are species j's acquisition rates for R_1 and R_2, and control how efficient species j is at acquiring resources. Thus we modelled the four different resource scenarios by picking four different values for the shape parameter α : strongly complementary ($\alpha \ll 1$), strongly antagonistic ($\alpha \gg 1$), weakly complementary ($\alpha \lesssim 1$) and weakly antagonistic ($\alpha \gtrsim 1$).

To fully specify the model we also had to tune species' consumptive impacts on their resources, $\boldsymbol{C_{ij}(R_1, R_2)}$. To do so we followed Koffel et al. (2016) and model per capita consumptive impacts separately for each resource scenario to better match the biology of resource uptake.

For strongly complementary resources ($\alpha \ll 1$), we assume that both resources are necessary to grow, which means that growth translates into proportional uptakes of both resources through fixed stoichiometric coefficients q_{ij}:

$$C_{ij}(R_1, R_2) = q_{ij}\, g_j(R_1, R_2) \qquad (3)$$

This is consistent with the way essential resources are usually modeled (Tilman 1982). We further assume that $q_{1j} = x_{2j}$ and $q_{2j} = x_{1j}$, which ensures that every species has a greater impact on the resource that most limits its growth, a standard necessary condition for coexistence to happen in resource competition theory (Tilman 1982, Leibold 1995, Chase and Leibold 2003).

For strongly antagonistic resources ($\alpha \gg 1$), resources can be substituted for each other, but a mixed uptake does not translate additively into growth because of the antagonism between the two resources. Such antagonism can for example happen when the two resources are patchily distributed. In this scenario, the consumers focus their uptake on the most profitable resource (based on its abundance) and avoid the least profitable (often rarest) one. Mixed uptake only occurs for a narrow range of availabilities corresponding to equally profitable resources, where the consumer is about to switch from consuming one resource to consuming the other. Switching uptake can be modelled by the following mathematical expression:

$$C_{ij}(R_1, R_2) = \frac{(x_{ij}R_i)^{\alpha}}{(x_{1j}R_1)^{\alpha} + (x_{2j}R_2)^{\alpha}}\, g_j(R_1, R_2) \qquad (4)$$

This choice of consumptive impact is consistent with mass balance of substitutable resources, where all the resources taken up translate into growth by an equivalent amount of biomass. This can formally be checked by summing the two uptake rates of a species j given by Equation (4) and retrieving its growth rate ($g_j = C_{1j} + C_{2j}$).

Finally, in the weakly complementary and weakly antagonistic resource scenarios ($\alpha \lesssim 1$ and $\alpha \gtrsim 1$, respectively), resources are close to being consumed proportionally to their availability, i.e $C_{ij}(R_1, R_2) \approx x_{ij}R_i$, consistent with their close proximity to the perfectly substitutable case ($\alpha = 1$; Schreiber and Tobiason 2003). Such a close-to-proportional consumptive impact is obtained by borrowing

equation (4) from the antagonistic resource scenario and using it for these two weakly interacting scenarios. Indeed, equation (4) tends towards linear consumption when α tends towards 1.

Fig. 1 illustrates how each of these resource types affects trade-off relations among minimum resource requirements for the two resources. Previous work has almost exclusively focused on the strongly complementary scenario (e.g. Tilman et al. 1997, Gross and Cardinale 2007, Hodapp et al. 2016); however, we found that different types of competition lead to different outcomes of community assembly, associated with distinctive signatures in terms of Complementarity-Selection Effects and Community Assembly components (see Box 1 and `Results' section).

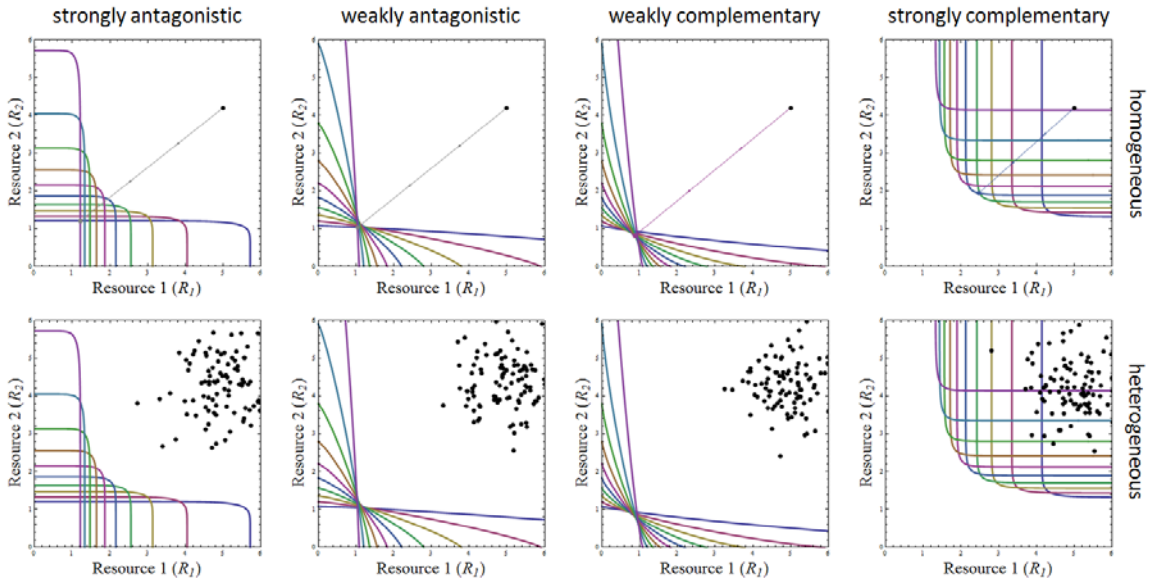


Figure 1. ZNGIs (colored, thick) and supply points (black dots) for each of the eight resource scenarios. ZNGIs are curved upwards or downwards when resources are complementary or antagonistic, respectively, and close to linear for the two weakly interacting scenarios. Within each scenario, the positions of the ZNGIs of the different species indicate where they lie along the generalist-specialist continuum: generalists have symmetrical ZNGIs at equal distance from the two resources axes while specialists on one resource have asymmetrical ZNGIs closer to that resource's axes. The superimposed supply points helps visualize environmental heterogeneity in each scenario. In the homogeneous case (top row), we also show how the consumption by the single species or pair of species (colored or gray arrows &

points, respectively) that dominates after community assembly decreases resource availability from the supply point to the resource levels at equilibrium, which is the outcome of invasion analysis (See main text and Box 1). Similar arrows could have been used to show the variation in outcome from each supply point in the heterogeneous case (lower panels) but are omitted for clarity purposes.

Resource Supply: We also considered how adding environmental heterogeneity at the local scale affects the results obtained from the Complementarity-Selection (Loreau and Hector 2001) and Community Assembly components methods. We ran our model assuming two different, spatially-implicit scenarios for the supply of resources over the local community: homogeneous resource supply and heterogeneous resource supply. The homogeneous environment was characterized by a constant influx of both resources (i.e. a single 'supply point', *sensu* Tilman 1982). The second scenario examines a heterogeneous environment, where species compete in an ecosystem characterized by 100 microhabitats, each with a separate supply point (see Fig. 1).

Generating a species pool: BEF experiments consist of replicate communities where species are selected from a common species pool. To generate this species pool, we created a global pool of ten species whose acquisition traits for the two resources are regularly spaced along a generalist-specialist continuum defined by a linear trade-off between the two acquisition traits. We implemented this with the Schreiber and Tobiason equation (equation 2) that we used to model species growth rates. In equation (2), x_{1j} and x_{2j} control how efficient species j is at acquiring resources. We assumed that these two traits are bound by a trade-off, which we assumed to be linear for simplicity ($x_{1j} + x_{2j} = 1$). Varying these two traits between 0 and 1 generates a generalist-specialist continuum.

Community Assembly Outcomes:

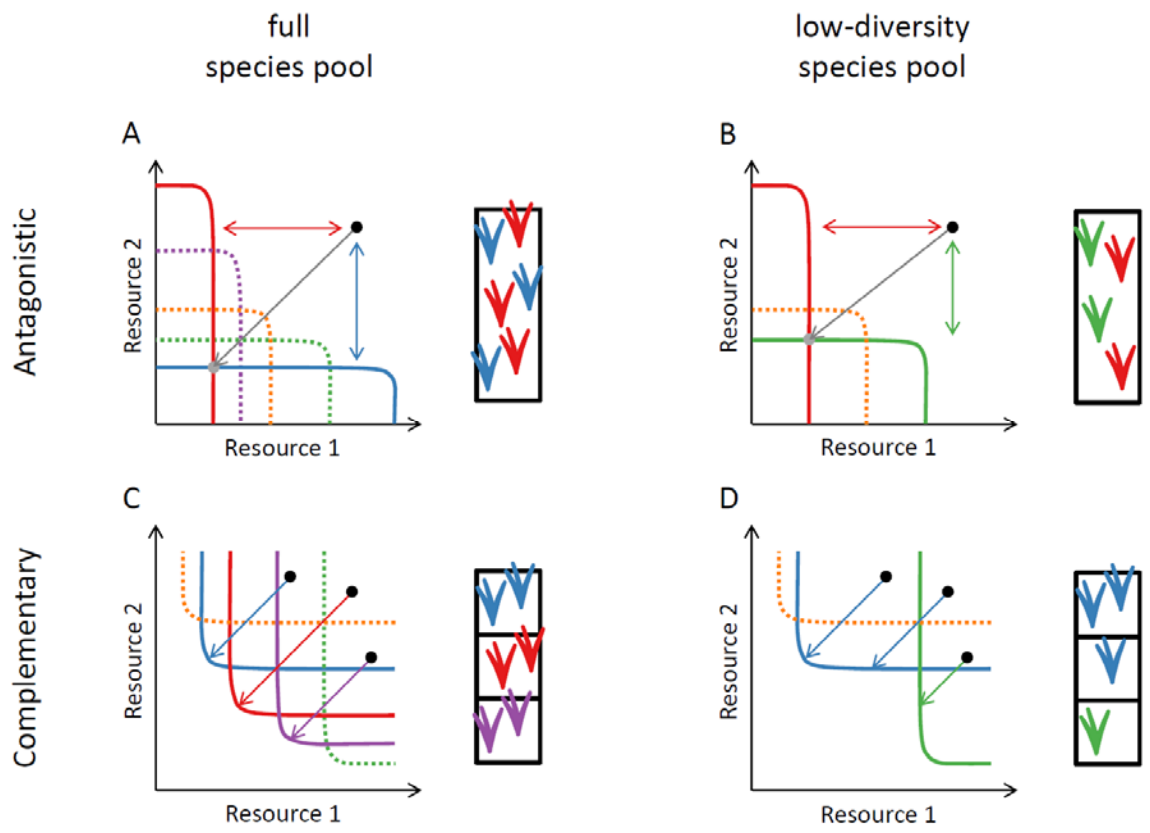
For each resource competition scenario (complementary vs. antagonistic, homogeneous vs. heterogeneous), we imitated the conditions used in standard BEF

experiments to generate multiple individual species pools with different subsets of species from the global pool that differ in their species richness and composition. Our species pool of ten species yields 1023 possible subsets of this species pool, including the 10 different monocultures and the full species pool. We then considered the equilibrium states of communities initialized with all possible subsets of this global pool, arranged by their starting richness (e.g., 10 one-species communities, 45 two-species communities, etc.). This was performed using a global invasion analysis (Box 1), based on ZNGIs and their envelopes (Koffel et al. 2016). In the homogeneous case (a single supply point), only one or two species can subsist at equilibrium on two resources (Levin 1970) but more species can coexist if supply points are heterogeneous. In the heterogeneous case, 100 supply points were drawn from a multivariate normal distribution with the same mean as in the homogeneous case and a variance of 0.5 (See Fig. 1). The equilibrium community state was determined for each supply point. Because these equilibria for both the homogeneous and heterogeneous cases are deterministic, only one or two species will ultimately persist for any single supply point (more species can coexist for multiple supply points), and the identities of coexisting species and their abundances are predetermined by their growth functions and consumption vectors. Each resulting equilibrium community was characterized by the identity of the species remaining after community assembly and their equilibrium biomass. In the homogenous case, total ecosystem function is simply the sum of biomass for all species in the equilibrium community. For the heterogeneous case, equilibrium biomasses were obtained by averaging over the set of supply points, enabling a more direct comparison with the homogeneous case.

We then consider key properties of the communities that emerge under each resource scenario (across diversity levels and all initial compositions). These include the final biomass of each species, which together determine the total biomass or function of the ecosystem, and realized species richness (the final number of species present). This information for each of the subsets was then used in the Complementarity-Selection and Community Assembly calculations.

Box 1: Mechanisms

In this box, we give a mechanistic explanation of how different competitive scenarios determine the identity and function of persistent species. In turn, these will determine the sign and magnitude of the Community Assembly components that arise from comparing communities assembled from high and low diversity species pools. We consider both strongly antagonistic and complementary resource scenarios. We show a simplified example where the full species pool only includes five species, three of which appear in the low diversity subset. We focus on a homogenous environment (with one supply point; black dot) in the antagonistic example and a heterogeneous environment (with three supply points) in the complementary case.



These situations are represented and analyzed using the graphical approach to resource competition (MacArthur 1972, Tilman 1982, Leibold 1995, Chase and Leibold 2003, Koffel et al. 2016). We perform the invasion analysis by drawing the ZNGIs of the species in the community all together. The species whose ZNGIs are entirely located above other species' ZNGIs are systematically excluded (e.g. purple, orange, and green in A). When the impact vectors of the remaining species (red and blue in A) are combined with a supply point, we obtain the set of these species that actually end up persisting in the equilibrium community (e.g. red and blue coexist in A). In a heterogeneous environment (C and D), we perform the invasion analysis for each supply point independently and average the productivity of the resulting persisting species to obtain the equilibrium community.

In the antagonistic case, when all members of the full species pool are initially present (A), the three intermediate species (purple, orange, green; dotted lines) are excluded by the two extreme specialists (red and blue; solid lines), which coexist.

However, if one of these two specialists (e.g., blue) is randomly absent from a low-diversity community (B), it is replaced by its closest functional equivalent among the intermediate species (here, green). This species is not completely specialized on resource 2, so it does not perform as well as the lost specialist. Due to weak competition with this intermediate species, the performance of the remaining specialist (red) is lower than in the full species pool.

In the complementary case, the most diverse species pool (C) experiences the exclusion of the two most extreme species (orange and green), while the three intermediate species (blue, red, and purple) coexist, as each one is locally adapted to a particular supply point. If two of these intermediate species, which dominate the full species pool, are absent in a smaller community (e.g., red and purple), an extreme species (e.g., green) can survive and coexist with the remaining dominant species (blue). This extreme species (green) only partially compensates for the loss of function that occurs due to the absence of the purple species. Similarly, the remaining dominant species (blue) partially compensates for the loss of function due to the absence of the red species by increasing in abundance.

Complementarity-Selection analysis:

The outputs from our competition models serve as inputs to the Complementarity-Selection approach (Loreau and Hector 2001). For the purposes of this calculation, we assumed all species had equal 'initial' abundances. The decomposition divides total function into the Selection effect (SE) and Complementarity effect (CE) for each of the 1023 communities. SE and CE were averaged across all possible species pools at each level of introduced richness regardless of how many species actually existed at equilibrium.

Community Assembly analysis:

380 No supplementary assumption was needed to apply the Community Assembly
381 Components (Fox and Kerr 2012; Bannar-Martin et al. 2018). We compared the
382 community initialized with the entire pool of ten species (our baseline, *sensu*
383 Bannar-Martin et al. 2018) to each less diverse community. Using the Community
384 Assembly configuration of the Price equation as described in Bannar-Martin et al.
385 (2018), we calculated the impact on ecosystem function of species lost (SL), species
386 gained (SG) and the changes in function of species shared between the most diverse
387 (and hence, productive) community and each comparison community (CDE). We
388 then averaged each of these components across all possible species pools within
389 each level of introduced richness.
390
391 *Comparison with experimental studies:*
392
393 To illustrate how our approach can provide insights about actual patterns of
394 community assembly and the functioning of ecosystems in empirical systems, we
395 applied the same analyses to data from a well-studied, representative BEF
396 experiment conducted at Jena (Roscher et al. 2004; 2005; Weigelt et al. 2016).
397 Established in 2002, this experiment manipulated grassland communities,
398 controlling the seeded species richness (2, 4, 6, 8, and 16 species, as well as the full
399 60 species pool) and functional group representation (1 to 4 groups) of 66 plots;
400 additional small monocultures (3.5 x 3.5 m plots) were established for each species.
401 Randomly selected subplots were sampled in the spring and summer of each year:
402 above ground biomass was clipped, sorted to species, dried, and weighed (see
403 Roscher et al. 2004, 2005 for additional experimental details). Roscher et al. (2005)
404 found evidence of positive selection and complementarity effects across plots with 2
405 - 8 species, focusing on data from May of the second year of the experiment (2003).
406 Using data available from Weigelt et al. (2010), we expanded this Selection-
407 Complementarity analysis to cover 2003-2008 (focusing on spring data, as do
408 Roscher et al. 2005), and augmented it by applying the Community Assembly
409 approach. Prior to analysis, we averaged species' biomasses across replicate

subplots within each experimental plot. The results provide an empirical case study for comparison with our modeling exercise.

Complementarity-Selection results for each year are based on monoculture biomass data collected during the corresponding year. Generally, monoculture biomasses declined over the six-year horizon of this analysis (Marquard et al. 2013; declines in polyculture yield were weaker). Consequently, there is a tendency for the magnitude of Complementarity-Selection components to grow more pronounced over time, due to their mathematical definition (Loreau and Hector 2001). As the Community Assembly calculations are independent of monoculture data, they were not similarly affected.

Results:

Overall we studied eight different scenarios with our model. These varied in the type of competition and the degree of resource heterogeneity as shown in Table 1 and Fig. 1. Although ecosystem function (here total biomass or yield) was always a saturating function of initial species richness (size of the species pool), different scenarios produced qualitatively distinct relationships between realized species richness (at equilibrium) and initial species richness (Fig. 2). These and other results were often similar for different combinations of resource type and/or environmental heterogeneity. We focus on qualitatively distinct outcomes as follows:

Table 1. The eight modeling scenarios in our study, which differ in the type of resource competition and degree of environmental heterogeneity they feature, as described in text and illustrated in Fig. 1. Our results showed that both competition and heterogeneity produced variable effects on complementarity (CE), selection (SE), species loss (SL), species gain (SG) and context dependence (CDE). We also found that comparisons between SG and CDE were informative. The magnitude of these effects is summarized qualitatively through the categories “exactly zero” (0),

441 “approximately zero” (≈ 0), “negative” (−), “positive” (+) and “strongly positive”
442 (++).
443

Scenario	Resource type	Environmental heterogeneity	Predicted effect (see Results) on:					
			CE	SE	SL	SG	SG vs CDE	CDE
1	strongly antagonistic	homogeneous	++	−	−	++	>	≈ 0
2	strongly antagonistic	heterogeneous	++	−	−	++	>	≈ 0
3	weakly antagonistic	homogeneous	++	−	−	++	>	−
4	weakly antagonistic	heterogeneous	++	−	−	++	>	−
5	weakly complementary	homogeneous	≈ 0	+	−	++	>	0
6	weakly complementary	heterogeneous	≈ 0	+	−	+	<	+
7	strongly complementary	homogeneous	≈ 0	+	−	++	>	0
8	Strongly complementary	heterogeneous	+	+	−	+	<	+

444

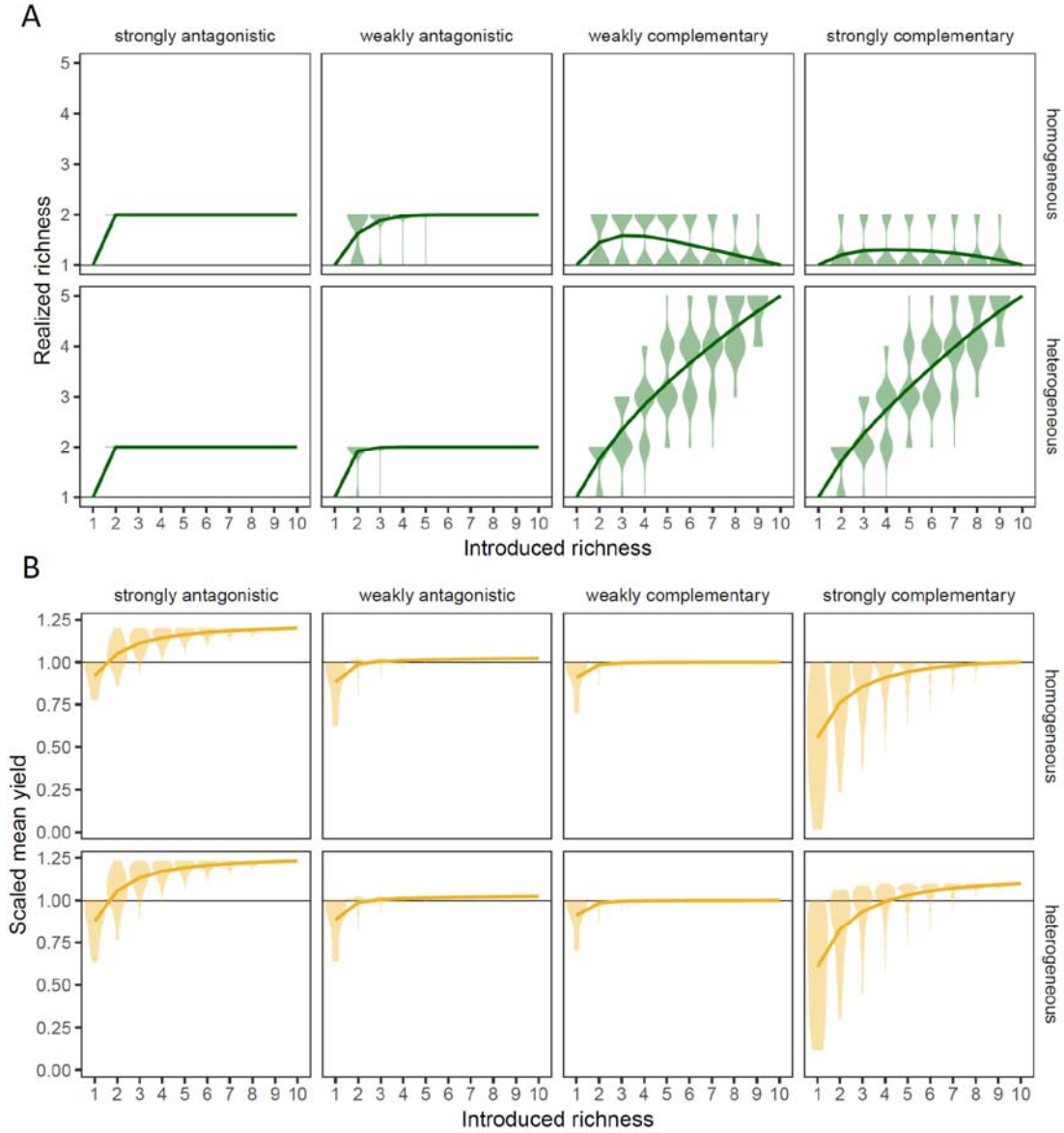


Figure 2. A) Realized richness obtained after species sorting, averaged across all possible species pools at each level of introduced richness. In the four antagonistic resource scenarios, realized richness quickly saturated at two species. In the complementary resource scenarios, average realized richness is hump-shaped in the homogeneous case, but increasing in the heterogeneous case. B) Total community yield obtained after species sorting, averaged across all possible species pools at each level of introduced richness; values are scaled relative to the yield of the most productive monoculture (solid horizontal line). All scenarios give saturating responses but in scenarios where resources are close to substitutable (central panels), moving from monocultures to polycultures causes weaker increases in yields.

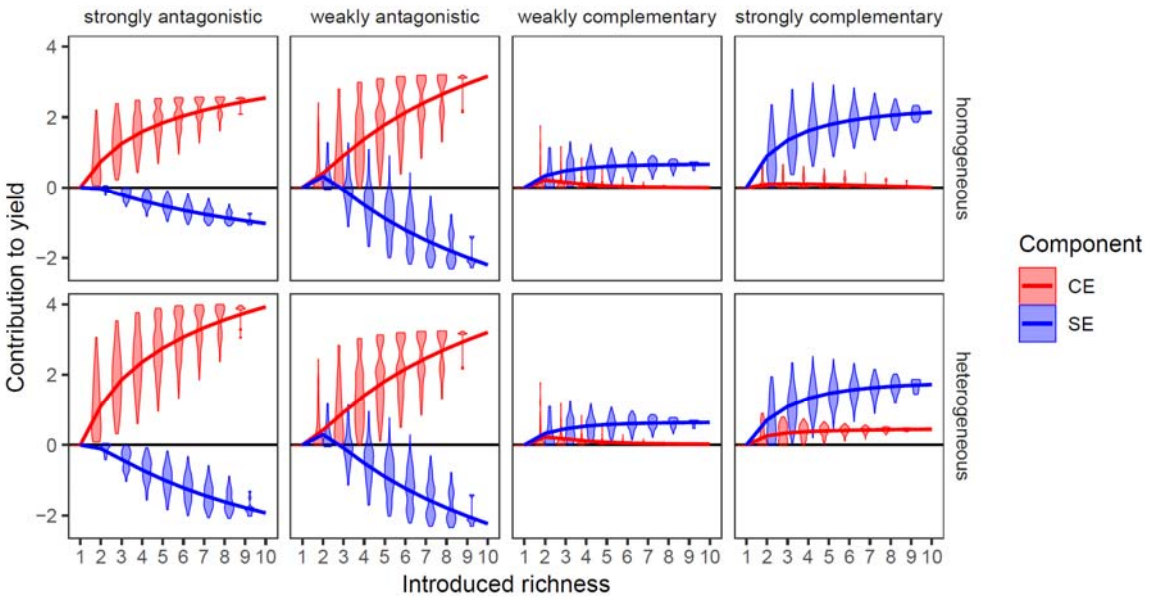


Figure 3. Complementarity (CE) and selection effects (SE) vary in magnitude and sign as a function of introduced richness, competitive scenario (columns), and supply point heterogeneity (rows). Antagonistic resources lead to CE > 0 and SE < 0, while complementary resources result in CE > 0 and SE > 0. As in Fig. 2, violin plots indicate averages and variability across all possible species pools at each level of introduced richness.

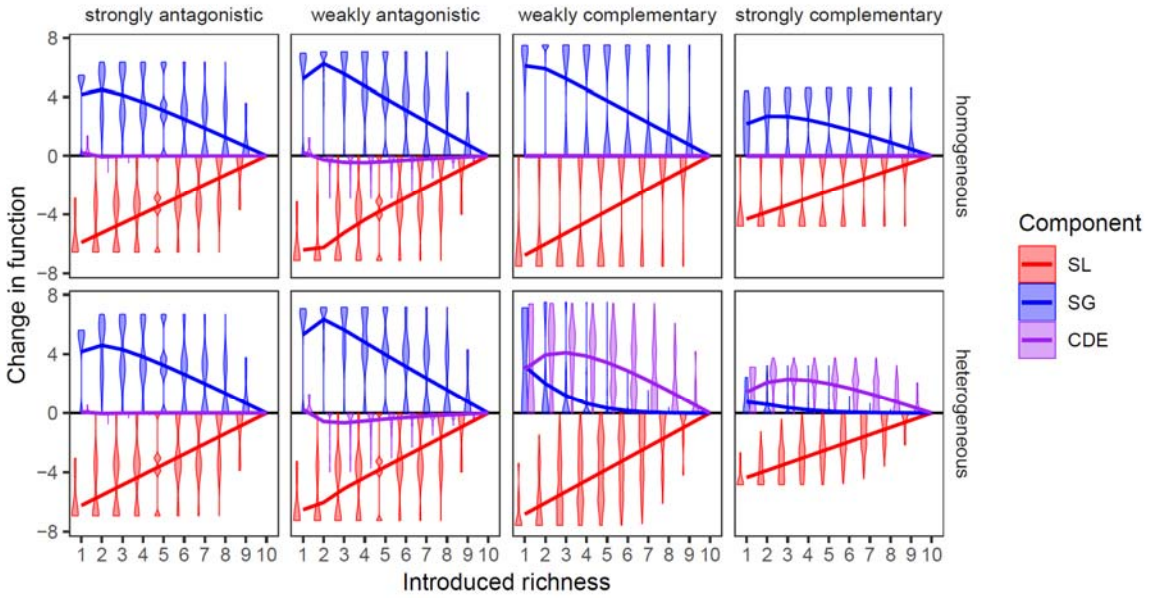


Figure 4. Community Assembly Components method, comparing the most diverse polyculture (initial richness of 10) against less diverse mixtures. CDE is nearly always zero or slightly negative, except when resources are complementary and have a heterogeneous supply. As in Fig. 3, violin plots indicate averages and variability across all possible species pools at each level of introduced richness.

Antagonistic resources:

Antagonistic resources, whether weak or strong, or supplied uniformly or heterogeneously, produced similar results. First, the maximum yield of polycultures was always greater than the maximum monoculture. Second the realized richness was always two species or fewer. In these scenarios, the average complementarity effect (CE) was always positive — setting aside the weakly antagonistic case with introduced richness of two — whereas the average selection effect (SE) was always negative (Fig. 3). Finally, the average Community Assembly Components always showed positive species gains (SG) and negative species loss (SL) components whereas the CDE was always small or zero (Fig. 4).

Strongly complementary resources in homogeneous environment:

In contrast to antagonistic resources, results varied substantially depending on just how strongly resources are complementary and on the degree of environmental heterogeneity. Starting with strongly complementary resources (almost equivalent to essential resources), we found that maximum yield in polycultures did not differ from that of the most productive monoculture. Both selection (SE) and complementarity effects (CE) were positive, in contrast with antagonistic resources (Fig. 3). The increase in yield with diversity was largely due to an increase in the mean selection effect: as the size of species pools grew, more productive competitors were more likely to be present. The relationship with complementarity was unimodal and reflected the increased role of complementarity on yield in small species pools where the better competitors were absent. The effects of complementarity were also generally small. Community Assembly Components showed that effects of CDE were zero and that all of the effects involved positive effects of SG and negative effects of SL (Fig. 4). We also found that mean realized species richness was a unimodal function of the size of the initial species pool in the homogeneous scenario (but still limited to two or fewer species; Fig. 2.A and S1).

This means that the relationship between realized diversity and yield appears complex (Figure S1).

Strongly complementary resources in heterogeneous environments:

The patterns differed from the homogenous environment results in several respects. First, the maximum productivity was typically higher than that of the most productive monoculture. Also, realized species richness was now positively related to the size of the species pool. In contrast with the homogeneous case, the positive complementarity effect (CE) persisted at high diversity levels (Fig. 3), while species gains (SG) had much smaller effects and CDE was positive, counteracting a similar SL effect (Fig. 4). We also examined an intermediate scenario, with smaller variance in supply point heterogeneity, and found intermediate patterns between the homogenous case and the full variance case: intermediate SE and CE, and CDE and SG of comparable magnitudes (not shown).

Weakly complementary resources:

With resources that were only weakly complementary, the increase in yield with diversity was less pronounced. Both CE and SE results were similar to the complementary case (Fig. 3), except that environmental heterogeneity did not support a sustained positive CE effect at high diversity levels.

Relations between realized species richness and biomass:

In each of our scenarios, we also examined how total biomass was related to the realized species richness (the number of species coexisting at equilibrium). Over all initial richness levels, there was generally a positive relationship between standing biomass and realized species richness (Fig. S1). However, scenarios differed in the strength of this relationship and in the structure of covariation between biomass and realized species richness when the size of the species pool was accounted for.

In most scenarios (scenarios 1-6 and 8 in Table 1), we found that biomass was positively related to realized species richness and that this was true both at the global scale of analysis and within subsets of communities that differed in the size of the introduced species pool. In contrast, when resources were strongly complementary and environmental heterogeneity low (scenario 7) we found that total biomass was negatively related to realized species richness for any given level of initial richness (Fig. S2).

Mechanisms underlying relationships between resource scenarios and Complementary-Selection and Community Assembly patterns

The striking differences between the antagonistic and complementary resource competition scenarios captured by the Community Assembly components can be linked to the influence of resource type on the outcomes of community assembly. In the antagonistic case, a pair of extreme specialists dominates all the other species for any supply point (Box 1, panels A, B). This strongly affects both the complementarity and selection effects (CE and SE), and the Community Assembly components (SL, SG, and CDE). Dominance by a pair of specialists leads to a realized richness that cannot exceed two even in the heterogeneous case. In monoculture, each specialist strategy generally performs worse than the intermediate generalist strategies, which explains why the selection effect is negative. Yet, the two specialists together perform better than any other set of one or two species, driving the positive CE pattern. When one of the two dominant specialist strategies is absent (causing a negative SL component), it is replaced by a less specialized species with lower function than the missing specialist (which results in a positive SG component but $|SL| > |SG|$). The less specialized species competes more strongly with the remaining specialist due to greater niche overlap, reducing the performance of the remaining specialist, which creates the negative CDE.

In the complementary cases, there is usually a unique generalist species from the full species pool that dominates for a given supply point (Box 1, panels C, D). This

562 dominant species is also the one that performs the best – better than any other set
563 of one or two species – for this supply point, which explains why selection is
564 positive and complementarity weak and hump-shaped. The CDE of exactly zero in
565 the homogeneous case is a direct consequence of having a single dominant strategy:
566 either this dominant strategy is absent from a low diversity community, which will
567 then ultimately share no species with the high diversity community, or the
568 dominant strategy is present, producing an endpoint identical to the high diversity
569 community. Finally, the positive CDE in the heterogeneous case arises because in the
570 absence of one or more of the dominant strategies, other dominant species that still
571 occur compensate for functional losses. Such CDE is larger than the SG component
572 on average because the functional contributions of new species (able to join a
573 community due to the loss of at least one dominant species) are smaller than
574 increases in function occurring through compensation by the remaining dominants.
575
576 *Comparisons with experimental data:*
577
578 Although all of our scenarios produce broadly saturating BEF patterns, our results
579 indicate that we can qualitatively discriminate among most of these eight different
580 community assembly scenarios (Table 1) by combining the Community Assembly
581 (Fox and Kerr 2012, Bannar-Martin et al. 2018) and the Complementarity-Selection
582 (Loreau and Hector 2001) analyses. This suggests that applying these methods to
583 empirical data may offer insights into which competitive scenarios drive assembly,
584 giving rise to empirically-obtained BEF relationships (or at least provide evidence
585 about which mechanisms may be ruled out). To assess this, we analyzed data from a
586 representative BEF experiment, conducted at Jena, Germany, to determine if we
587 could identify the community assembly mechanisms underpinning the functioning
588 of this grassland ecosystem. We used six years of data from the Jena experiment,
589 starting in 2003 (a year after the experiment was established), to consider whether
590 results were stable over time.
591

We found that the data from Jena did not match any of the scenarios shown in Table 1. On the one hand, we see a divergence between increasingly negative selection effects and increasingly positive complementarity effects from 2003 to 2008 (Fig. 5), creating a pattern consistent with the antagonistic resource scenario using the Complementarity-Selection approach (Fig. 3, see also Marquard et al. 2009). On the other hand, we see persistent trends in the Community Assembly components, including positive CDE effects across the entire time span (Fig. 6), which were only observed in our model scenarios given competition for heterogeneous, complementary resources (Fig. 4).

Consequently, the analysis of Complementarity-Selection effects (Fig. 5) is consistent with antagonistic resources, be it in homogeneous or heterogeneous supply points (compare with Fig. 3), whereas the Community Assembly Components suggest a scenario of complementary resources in the presence of environmental heterogeneity (comparing Fig. 5 with Fig. 4). This indicates, perhaps unsurprisingly, that the eight simplistic scenarios we considered probably lack additional mechanisms involved in regulating or modifying the effects of community assembly on ecosystem function (and hence the metrics used to study it).

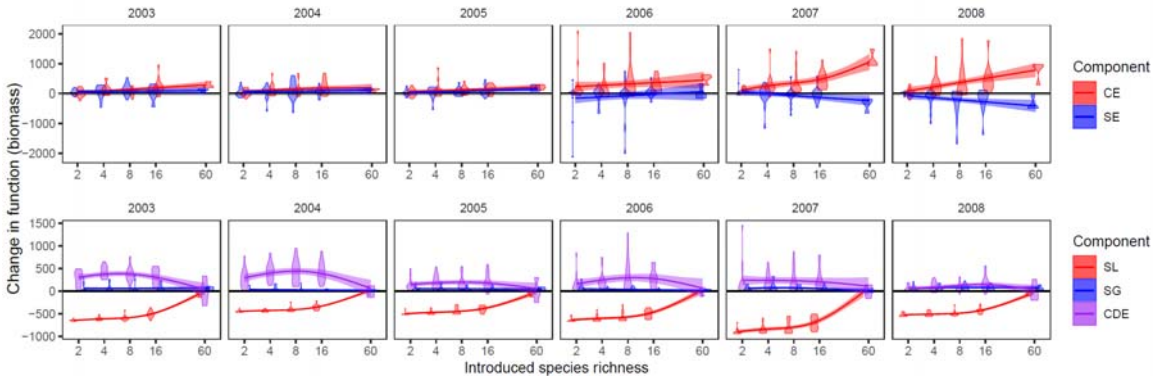


Figure 5. Complementarity (CE) and selection (SE) effects (top row), and Community Assembly components comparing the most diverse (60 species) community to less diverse communities (bottom row) over successive years at the Jena Experiment, as a function of introduced species richness (log scale). Early in the experiment, CE and SE were both positive, and weakly related to diversity (as found by Roscher et al. 2005). However, by the end of the period we examined, the CE

effect was positive and increased strongly with diversity, while the SE effect was negative and decreased with diversity. Although there is some variation in the magnitude of the Community Assembly components through time, at all lower diversity levels, CDE effects are consistently positive, SG is negligible, and SL strongly contributes to declines in function. Trend lines show Generalized Additive Models (GAMs) fit to the observed plot-level values; points jittered for visual clarity.

We thus considered other possible factors that might have affected community assembly at Jena. One of these is that the observed dynamics may not have reached steady state as assumed by our models above. To address this possibility, we simulated the temporal dynamics of community assembly for the heterogeneous, complementary resources scenario (because previous work has mostly emphasized that this scenario is more likely than others for plants; see Tilman 1982, Tilman et al. 1997, Gross and Cardinale 2007, Hodapp et al. 2016). Starting from an initial condition where all species are seeded at an equally low density, we let their populations grow and compete over time. When a species' density dropped under a certain minimal threshold (set equal to 10^{-4}), we considered it extinct and set its density equal to zero. We then extracted species densities at representative time points and applied both partitions as previously described. Although this exercise shows that there are important transient dynamics in the community assembly dynamics of BEF experiments (Fig. 6), none of the resulting patterns are consistent with the patterns we observed in Jena (Fig. 5). Intriguingly, we found that the dynamics in our model (Fig. 6) showed a strong qualitative shift in the relative importance of CE and SE through time and a somewhat subtler shift from linear to unimodal CDE and an increase in SG through time. However, the data from Jena show no such trends.

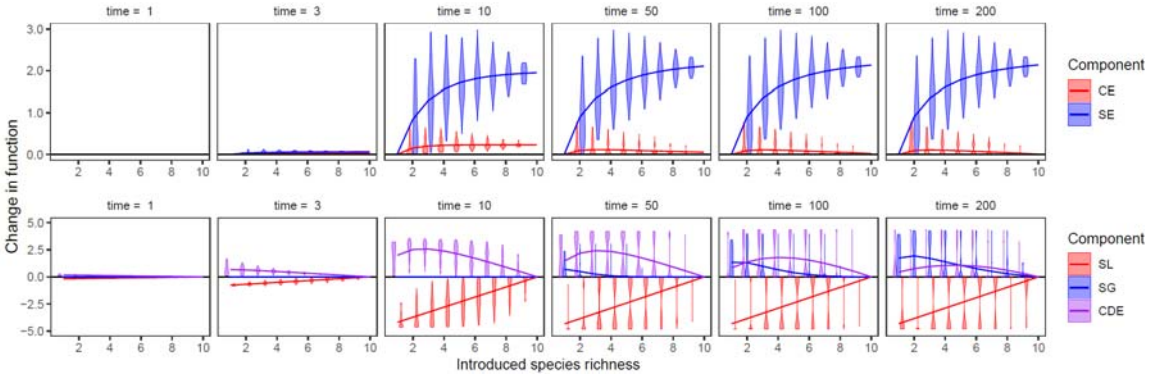


Figure 6: Transient dynamics in Complementarity-Selection effects and Community Assembly components as assembly proceeds towards equilibrium in the scenario where species compete for heterogeneous, strongly complementary resources.

We also considered that facilitation, perhaps due to the nitrogen (N) fixing effects of legumes, may play an important role in assembly, and consequently affect ecosystem functioning. To address this, we adapted the model of Koffel et al. (2018), and again applied Complementarity-Selection and Community Assembly analyses (see SI for details). For the case with high environmental heterogeneity and strong facilitation under high initial N-limitation, our results (Fig. 7) correspond at least qualitatively to our observations at Jena. Both the facilitation model and Jena data demonstrate greater complementarity effects (CE) in comparison to selection (SE) and, at least for Jena's 2007 and 2008 data, a negative SE in the Complementary-Selection Effects analysis (Fig. 5,7A). This seems consistent with observations from the Jena Experiment that showed that N fixation rates by legumes, and thus potentially facilitation, increased from 2004 to 2008 (Roscher et al. 2011). The Community Assembly Components also show similarities, with a substantial and humped CDE effect that emerges from the model (Fig. 8B) also being seen in the Jena data (Fig. 6). The addition of facilitation creates a combination of Complementarity-Selection Effects and Community Assembly Components patterns that otherwise appear to point to divergent resource competition models in the absence of facilitation.

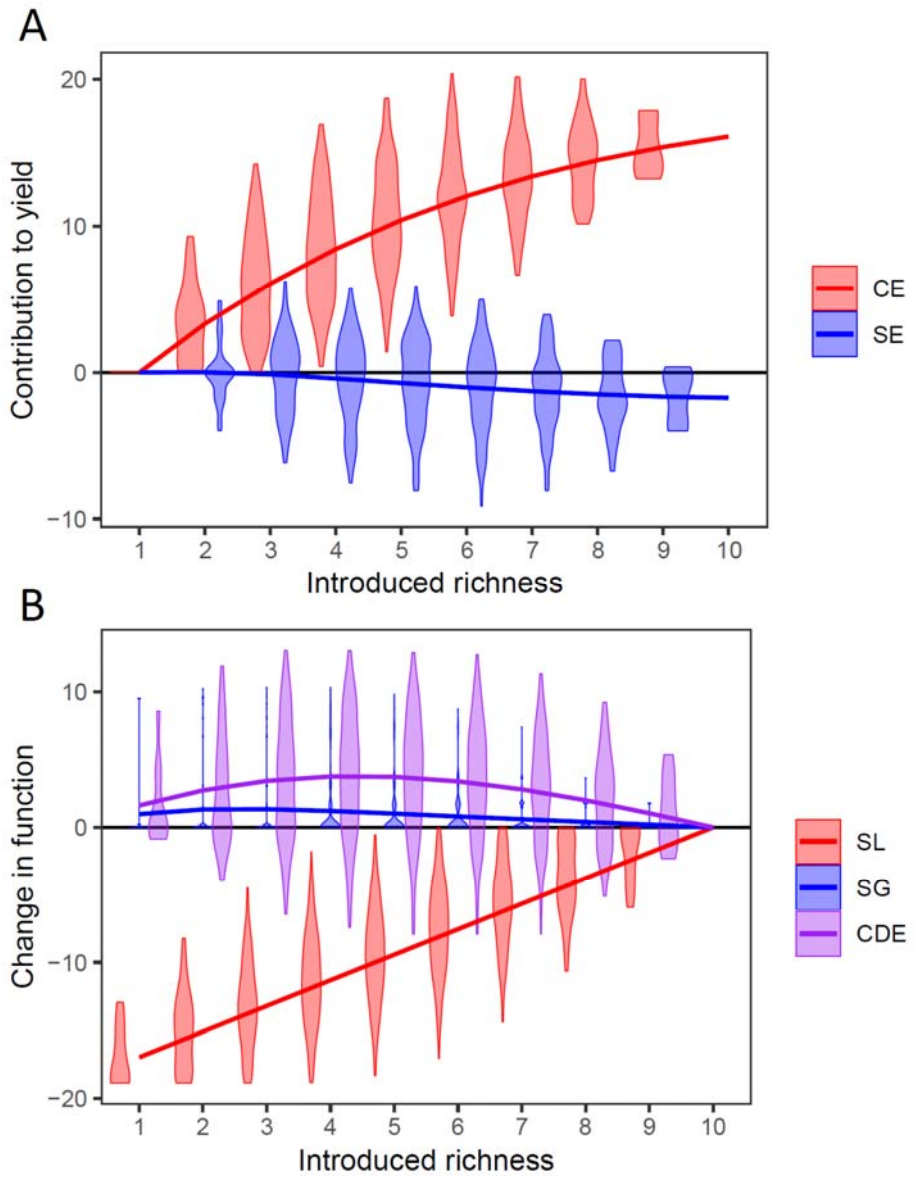


Figure 7: Model results from including nitrogen fixing plants that compete with non-fixers. Complementary-Selection Effects (panel A) and Community Assembly Components (panel B).

<u>Discussion:</u>

There are two major implications of our work. First, we find that the joint use of the Community Assembly (Bannar-Martin et al. 2018) and the Complementarity-

679 Selection (Loreau and Hector 2001) approaches can be a powerful tool in inferring
680 mechanisms involved in linking community dynamics to the functioning of
681 ecosystems. Second, we find that empirical results (at least in the case of the Jena
682 BEF experiment) do not match any of the scenarios we examined for purely
683 competitive mechanistic models but that they did when we incorporated facilitation
684 by N-fixing plants. This suggests that the joint evaluation of Complementarity-
685 Selection and Community Assembly components could be a powerful tool for
686 identifying plausible mechanisms that relate community assembly to the
687 functioning of ecosystems.
688
689 *Validating the joint use of Community Assembly and Complementarity-Selection*
690 *partitions to interpret mechanisms involved in BEF dynamics:*
691
692 Numerous studies in controlled and replicated conditions (such as those that prevail
693 in the large literature involving BEF studies) have shown that there is generally a
694 positive and directional effect between the size of the species pool with access to a
695 community and the function of at least some aspects of ecosystems, and in
696 particular the total biomass of the organisms involved. Previous work on
697 relationships between species pool diversity and ecosystem function has largely
698 focused on two very general effects that can explain this pattern: selection (the
699 species pool, when large, is more likely to include highly competitive and highly
700 productive species) and complementarity (the species pool, when large, is more
701 likely to include species that coexist and whose production is less mutually
702 inhibitory). To date, theoretical explanations for these two general effects have
703 focused on resource competition and used models based on complementary
704 resources (e.g., Tilman et al. 1997, Gross and Cardinale 2007, Leibold et al. 2017).
705 However, other mechanisms may also be involved. Here we explored the joint
706 consequences of manipulating the mechanisms governing species competition
707 (ranging from strongly antagonistic to strongly complementary resources), resource
708 heterogeneity and the size of the species pool on realized species richness,
709 composition, and production. We found that selection enhanced production when

resources were complementary, but undermined production when resources were antagonistic. Complementarity between species had little effect on production except when resources were antagonistic and/or heterogeneous.

Consequently, the mechanisms involved are not necessarily clearly resolved by the study of complementarity and selection alone. We found that additional insights could be gained by also considering the Community Assembly components (Bannar-Martin et al. 2018). Our work shows that these components can provide additional and complementary insights into the mechanisms structuring the relationships among species pool size, ecosystem function, realized species richness and community composition. In our case, the joint use of Community Assembly and Complementarity-Selection approaches provided well-defined criteria by which we could distinguish between competition for complementary vs. antagonistic resources as well as evaluate the role of environmental heterogeneity and facilitation.

Insights about mechanisms that drive community-ecosystem function relations:

Our study also highlights the fact that we do not necessarily understand all the mechanisms that drive community-ecosystem relationships. Patterns in the data from the Jena experiment did not correspond to any of the eight scenarios we modeled and allowed us to reject all eight of these hypotheses. On the one hand, this shows how powerful the joint application of Community Assembly and Complementarity-Selection metrics can be. However, this also shows that the mechanisms governing the diversity and function of communities must involve other processes besides simple resource competition. This is not totally surprising since our theoretical models are quite simplistic caricatures of how species interact. Here we used the joint analysis of Community Assembly and Complementarity-Selection metrics to evaluate two other possible explanations. In one, we evaluated the possibility that community assembly at Jena was not in final equilibrium — a hypothesis supported by the directional changes in Complementarity-Selection

741 components across years (Fig. 5) but found that this did not explain the data at Jena.
742 In the other, we derived predictions that included the role of N-fixing plants and
743 found that this model with facilitation did qualitatively correspond to the
744 observations at Jena. Critically, we note that a correspondence between theory and
745 observation is not definitive proof that a particular mechanism is operating. For
746 example, even though our model was explicitly based on N fixation, there is no
747 reason to believe that the facilitation mechanism that emerge in our model is
748 restricted to this ecological situation. Results similar to ours could arise from a
749 broader range of facilitation mechanisms at play in BEF experiments, such as
750 facilitation related to other abiotic factors (Wright et al. 2017) or based on soil
751 organisms (Eisenhauer et al. 2012). Of course, there are probably a number of other
752 mechanisms that could also be modeled, including interactions mediated via other
753 organisms such as grazers, pathogens and other mutualists.
754
755 *Community assembly and the functioning of ecosystems in other scenarios:*
756
757 How then does community assembly determine the functioning of ecosystems in
758 general? We focused on a very simple scenario for community assembly that
759 mimics the experimental design of 'classic' BEF experiments. This scenario consists
760 of looking at all possible subsets of introduced species combinations for a given
761 'total species pool'. This can be thought of as a filter on potential to 'arrive' in a
762 community via dispersal limitation. Assembly is then followed by species
763 interactions in which some species within this subset of potential species are
764 excluded by competition for resources. We find that this process can be related to
765 analysis by both the Complementarity-Selection and the Community Assembly
766 approaches. However, community assembly in real communities can involve
767 numerous factors not typically incorporated in BEF experiments, including species
768 invasions, disturbance and dispersal dynamics, environmental change, and
769 stochastic processes (e.g. Thompson and Gonzalez 2016, Hodapp et al. 2016, Leibold
770 et al. 2017, Bannar-Martin et al. 2018). The approach developed here on the well-
771 known BEF experimental setup could potentially be adapted to the study of these

other natural conditions. This would, however, require careful thinking about how to model these mechanisms and their interaction with resource competition.

While using the Community Assembly and the Complementarity-Selection approaches together can be a useful way to assess mechanisms, it is important to note that Complementarity-Selection partitions require estimates of ecosystem function in monocultures. These are readily available in most 'classic' BEF experiments as a pro-forma aspect of the experimental design. However, these are almost always absent in other experiments that examine how community assembly affects ecosystem functions (e.g. disturbance, invasions, extinctions, environmental change), let alone in observational data. The logistical challenge is even greater in studies examining the effects of environmental change on ecosystem function, as monoculture treatments would be required for all the relevant environmental conditions. Additionally, many organisms are not easily grown in monocultures due, for example, to dependencies on biotic resources, or movement. Thus, the scenarios where the Complementarity-Selection approach can be used are limited. In contrast, the Community Assembly approach can be calculated for any pair of communities. Given the design of BEF studies, we opted to use the most diverse species pool as our baseline for comparisons, but this is not required (see Bannar-Martin et al. 2018 for examples). While the Complementarity-Selection approach is constrained by the requirement of monoculture data, it did appear to more clearly distinguish between antagonistic and complementary resource scenarios. This ability in part may rest in the definition of these metrics: both CE and SE are affected by changes in the abundance of a species. In contrast, of the three Community Assembly components, SL and SG are mostly influenced by the actual loss or gain of species from a community; all effects of less extreme changes in abundance are relegated to the CDE component alone. For this reason, while it is possible to apply a Community Assembly approach in a manner directly analogous to Complementarity-Selection (contrasting an observed polyculture with an idealized polyculture where species perform proportional to their monocultures), the results are somewhat less informative.

803

804 Our work shows that, while both methods exhibit their own advantages and
805 limitations, together they provide a more powerful means of diagnosing
806 mechanisms. Future work could focus on continuing to refine these metrics. This
807 might include developing analogs of Complementarity-Selection partitions that do
808 not require monoculture data. Alternatively, new or existing extensions of the
809 Community Assembly approach (e.g., 5-part Price equation; Fox & Kerr 2012;
810 Bannar-Martin et al. 2018) could be applied to these or other model scenarios. For
811 example, while empirical applications of the trait-based version of the Price
812 equation (Fox & Harpole 2008) can be hampered by the availability of trait data, in
813 modeling exercises these traits are explicitly known (e.g., species' ZNGIs in our
814 scenarios). Further work could elucidate not only the mechanisms structuring
815 ecosystem function across communities of different diversity levels but relate
816 ecosystem properties to the functional composition of communities.
817

818 Our study also illustrates that one can discriminate between *a-priori* hypotheses and
819 *ad-hoc* hypotheses in ways that relate to the entire issue of 'replicability' in science
820 (e.g. Nichols et al. 2019). Here we set out to test 8 different hypotheses about how
821 resource competition affects the relations between community assembly and the
822 functioning of ecosystems. We generated these predictions with *a-priori* models
823 based on mechanistic resource competition theory (Tilman et al. 1997, Gross and
824 Cardinale 2007, Hodapp et al. 2016). We subsequently tested these predictions on
825 data from a previously conducted experiment that had not been previously studied
826 in this way and found that we rejected all of our a priori hypotheses. Of course, we
827 might have found that one or more of these hypotheses were not rejected and
828 stopped there with statistically valid inferences. However, the rejection of all
829 hypotheses begs the question 'what could the explanation instead be?'. Here we did
830 some *post-hoc* modeling based on two additional mechanisms and found that one of
831 these was consistent with the data. However, the statistical significance of this
832 hypothesis is not clear since we might have rejected it and continued on to testing
833 other hypotheses until we found one that was consistent with the data whereupon

we would have perhaps stopped. In our discussion we tried to indicate that other mechanisms could also have been similarly successful. The point is not that we don't believe that our tests are useful (they did in fact reject 9/10 hypotheses), but rather that the strength of our inferences about the hypothesis that was supported by the data should be appropriately weighed by the fact that it was an *ad-hoc* hypothesis rather than an *a-priori* one.

Implications for future work:

These encouraging results suggest a number of ways whereby the analysis of community assembly components could be improved:

a) Most obviously, one could continue to propose and test other mechanisms, such as plant-soil feedbacks and multi-trophic interactions, that could have produced the patterns we observed in the Jena data, eventually potentially creating a 'periodic table' for how the Complementarity-Selection and Community Assembly metrics map onto a diverse array of community assembly scenarios. While there is some allure to this suggestion, there are a considerable number of interacting mechanisms to test requiring careful thinking about how to construct and interpret models (e.g., considering transient dynamics), or how to parse varying contrasts in the data (e.g., considering facilitation by legumes).
b) Discriminating among the wide array of mechanisms driving community assembly and its effects on ecosystem function likely requires moving beyond simple dichotomous tests with one or two metrics since many alternative models/mechanisms are capable of reproducing similar, low-dimensional patterns. Here we demonstrated how the addition of several carefully considered metrics could enhance our understanding of these mechanisms, as well as our ability to differentiate them. It is worth considering whether other additional metrics could similarly advance understanding. The degree to which this is so would depend on finding

865 metrics that reveal distinct aspects of mechanisms and that, thus, are not
866 simply redundant with each other. It is also useful when the link between
867 mechanism and metrics is relatively transparent (as we explored in Box 1).
868 In cases where this is not the case this is perhaps less true and other
869 modeling or experimental studies would be needed to consolidate any
870 conclusions. The increase in the number of metrics, however, may be
871 necessary if theoretical models increase in complexity by including missing
872 processes or by assessing different assumptions in order to decrease
873 equifinality and increase inference.

874 c) Another reason to consider additional metrics is that not all data will be
875 suitable for the same metrics. In our case, while Complementarity-Selection
876 was particularly useful in discriminating between different types of resource
877 competition, the approach is constrained by the need for monoculture data.
878 We will need new techniques if we wish to apply these types of approaches
879 to a wider array of experimental conditions, with designs that do not closely
880 target BEF relations or that involve organisms where such monocultures are
881 logistically difficult (e.g. birds, migrating organisms, species-rich tropical
882 communities).

883 d) While we need to explore other metrics to use in studies other than BEF
884 experiments, it would still be interesting to apply the methods in this paper
885 to other BEF experiments. We suspect that this may likely show substantial
886 variation in results across locations, habitats, and organisms. For example, it
887 would be interesting to compare results at Jena with data from sites that are
888 not as likely to be nitrogen limited. Adequately considering such variation
889 would require a far more comprehensive theoretical and analytical effort
890 than is warranted in this paper.

891 e) Even more challenging is to consider how the size of the species pool jointly
892 affects biodiversity, community composition, and ecosystem functioning
893 under other scenarios than those at play in classical BEF experiments.
894 Numerous challenges to BEF research have focused on the degree to which
895 BEF experiments may be unrealistic representations of how biodiversity and

ecosystem function are likely to interrelate because the experimental setup for community assembly in BEF experiments is so artificial (Wardle 2016). We believe the Community Assembly approach can shed light on other situations including arrivals and removals of invasive species as well as environmental change (see Bannar-Martin et al. 2018, Kremer et al. in prep). Linking these scenarios to explicit theoretical models (as we did here by asking how resource types and environmental heterogeneity affect a BEF experiment) represents an exciting avenue for future devoted work.

Software availability:

An R package is available to facilitate future work (including analyzing other biodiversity experiments) https://github.com/ctkremer/priceTools

Acknowledgements:

This paper is a joint effort of the working group sCAFE and an outcome of a workshop kindly supported by sDiv, the Synthesis Centre of the German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig (DFG FZT 118). The Jena Experiment is funded by the German Research Foundation (FOR456/1451) with additional support from the University of Jena and the Max Planck Society.

References

Bannar-Martin, K. H., Kremer, C. T., Ernest, S. K. M., Leibold, M. A., Auge, H., Chase, J., Declerck, S. A. J., Eisenhauer, N., Harpole, S., Hillebrand, H., Isbell, F., Koffel, T., Larsen, S., Narwani, A., Petermann, J. S., Roscher, C., Cabral, J. S., and Supp, S. R. (2018). Integrating community assembly and biodiversity to better understand ecosystem function: the Community Assembly and the Functioning of Ecosystems (CAFE) approach. *Ecology Letters*, 21:167–180.

Bell, T., Lilley, A. K., Hector, A., Schmid, B., King, L., and Newman, J. A. (2009). A linear model method for biodiversity-ecosystem functioning experiments. *The American Naturalis*t, 174(6):836–849.

Cardinale, B.J., Duffy, J.E., Gonzalez, A., Hooper, D.U., Perrings, C., Venail, P., et al. (2012). Biodiversity loss and its impact on humanity. *Nature*, 486, 59-67.

Chase, J. M. and Leibold, M. A. (2003). *Ecological Niches: Linking Classical and Contemporary Approaches*. University of Chicago Press, Chicago.

Eisenhauer, N., Reich, P.B. and Scheu, S., 2012. Increasing plant diversity effects on productivity with time due to delayed soil biota effects on plants. Basic and Applied Ecology, 13(7), pp.571-578.

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., & Harpole, W.S. (2008). Revealing how species loss affects ecosystem function: the trait-based Price equation partition. *Ecology*, 89, 269–279.

Fox, J.W., & Kerr, B. (2012). Analyzing the effects of species gain and loss on ecosystem function using the extended Price equation partition. *Oikos*, 121, 290-298.

Gross, K., & Cardinale, B.J. (2007). Does species richness drive community production or vice versa? Reconciling historical and contemporary paradigms in competitive communities. *The American Naturalis*t., 170, 207-220.

Hector, A., Bell, T., Connolly, J., Finn, J., Fox, J. Kirwan, L., Loreau, M., McLaren, J., Schmid, B., and Weigelt, A. (2009). The analysis of biodiversity experiments: from pattern toward mechanism. pp. 94-104 in Naeem, S., Bunker, D.E., Hector, A.,

Loreau, M., and Perrings, C. (eds.). Biodiversity, Ecosystem Functioning, and Human Wellbeing: An Ecological and Economic Perspective. Oxford University Press. Oxford.

Hooper, D. U., Adair, E. C., Cardinale, B. J., Byrnes, J. E. K., Hungate, B. A., Matulich, K. L., Gonzalez, A., Duffy, J. E., Gamfeldt, L., and O'Connor, M. I. (2012). A global synthesis reveals biodiversity loss as a major driver of ecosystem change. *Nature*, 486(7401):105–108.

Kirwan, L., Connolly, J., Finn, J. A., Brophy, C., Lüscher, A., Nyfeler, D., and Sebastià, M.-T. (2009). Diversity-interaction modeling: estimating contributions of species identities and interactions to ecosystem function. *Ecology*, 90(8):2032–2038.

Koffel, T., Daufresne, T., Massol, F., and Klausmeier, C. A. (2016). Geometrical envelopes: extending graphical contemporary niche theory to communities and eco-evolutionary dynamics. *Journal of Theoretical Biology*, 407:271–289.

Leibold, M. A. (1995). The niche concept revisited: mechanistic models and community context. *Ecology*, 76(5):1371–1382.

Leibold, M. A., Chase, J. M., & Ernest, S. K. M. (2017). Community assembly and the functioning of ecosystems: how metacommunity processes alter ecosystem attributes. *Ecology*, 98, 909–919.

León, J. A. and Tumpson, D. B. (1975). Competition between two species for two complementary or substitutable resources. *Journal of Theoretical Biology*, 50(1):185–201.

Levin, S. A. (1970). Community equilibria and stability, and an extension of the competitive exclusion principle. *The American Naturalist*, 104(939):413–423.

Loreau, M. & Hector, A. (2001). Partitioning selection and complementarity in biodiversity experiments. *Nature*, 412, 72-76.

MacArthur, R. H. (1972). *Geographical Ecology: Patterns in the Distribution of Species*. Harper & Row, New York.

Marquard, E., Weigelt, A., Temperton, V.M., Roscher, C., Schumacher J., Buchmann, N., Fischer, M., Weisser, W.W., Schmid, B., (2009). Plant species richness and functional composition drive overyielding in a six-year grassland experiment. *Ecology* 90(12), pp. 3290-3302.

Nichols, J.D., Kendall, W.L., Boomer, G.S. (2019) Accumulating evidence in ecology: Once is not enough. *Ecology and Evolution* 2019; 00:1-14.

Ptacnik, R., Moorthi, S.D., & Hillebrand, H. (2010). *Hutchinson Reversed, or Why There Need to Be So Many Species*. In G. Woodward (Ed.), Advances in Ecological Research (Vol. 43, pp. 1–43). Burlington: Academic Press.

Roscher, C., Temperton, V.M., Scherer-Lorenzen, M., Schmitz, M., Schumacher, J., Schmid, B., Buchmann, N., Weisser, W.W. and Schulze, E.D., (2005). Overyielding in experimental grassland communities–irrespective of species pool or spatial scale. *Ecology Letters*, *8*(4), pp.419-429.

Roscher, C., Schumacher, J., Baade, J., Wilcke, W., Gleixner, G., Weisser, W.W., Schmid, B. and Schulze, E.D., (2004). The role of biodiversity for element cycling and trophic interactions: an experimental approach in a grassland community. *Basic and Applied Ecology*, *5*(2), pp.107-121.

Roscher, C., Thein, S., Weigelt, A., Temperton, V.M., Buchmann, N. and Schulze, E.D., 2011. N 2 fixation and performance of 12 legume species in a 6-year grassland biodiversity experiment. Plant and Soil, 341(1-2), pp.333-348.

Rychtecká, T., Lanta, V., Weiterová, I., and Lepš, J. (2014). Sown species richness and realized diversity can influence functioning of plant communities differently. *Naturwissenschaften*, 101(8):637–644.

Schreiber, S. J. and Tobiason, G. A. (2003). The evolution of resource use. *Journal of Mathematical Biology*, 47:56–78.

Thompson, P.L., and Gonzalez, A. (2016). Ecosystem multifunctionality in metacommunities. *Ecology*, doi: 10.1002/ecy.1502.

Tilman, D. (1982). *Resource Competition and Community Structure*. Princeton University Press, Princeton, NJ.

Tilman, D., Lehman, C. L., and Thomson, K. T. (1997). Plant diversity and ecosystem productivity: theoretical considerations. *Proceedings of the National Academy of Sciences (USA),* 94:1857–1861.

Tilman, D., Reich, P. B., and Isbell, F. (2012). Biodiversity impacts ecosystem productivity as much as resources, disturbance, or herbivory. *Proceedings of the National Academy of Sciences*, 109(26), 10394-10397.

Wardle, D.A. (2016). Do experiments exploring plant diversity-ecosystem functioning relationships inform how biodiversity loss impacts natural ecosystems? *J. Veg. Sci.*, 27, 646-653.

Weigelt, A., Marquard, E., Temperton, V.M., Roscher, C., Scherber, C., Mwangi, P.N., Felten, S., Buchmann, N., Schmid, B., Schulze, E.D. and Weisser, W.W. (2010). The Jena Experiment: six years of data from a grassland biodiversity experiment. *Ecology*, *91*(3), pp.930-931.

Weigelt, A., de Luca, E., Roscher, C., Temperton, V., Buchmann, N., Fischer, M., Scher-Lorenzen, M., Schmid, B., Schulze, E.-D., Weisser, W.W., Guo, L., Meyer, S.T. (2016). Collection of aboveground community and species-specific plant biomass from the Jena Experiment (time series since 2002). PANGAEA. https://doi.org/10.1594/PANGAEA.866358

Wright, A.J., Wardle, D.A., Callaway, R. and Gaxiola, A., 2017. The overlooked role of facilitation in biodiversity experiments. Trends in Ecology & Evolution, 32(5), pp.383-390.

Supplementary Informations

Figures

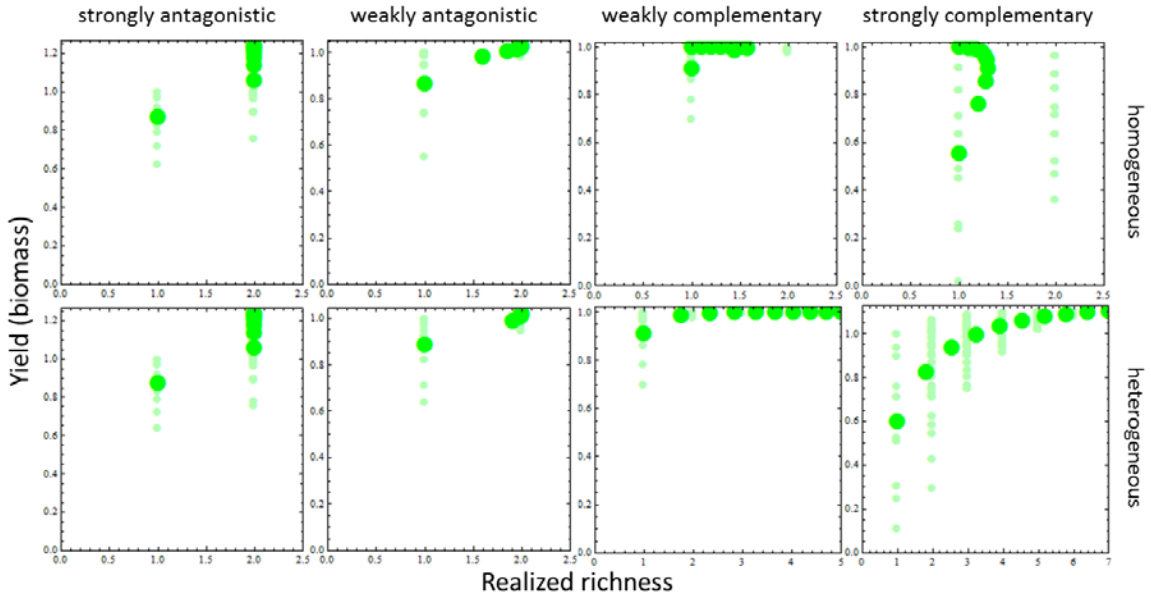


Figure S1. Relationship between yield and realized richness for the 8 scenarios. The small light points represent the equilibrium community resulting from seeding all possible subsets of the global species pool. The large darker points give the average yield and realized richness, where the average is taken across all possible communities for a given level of introduced richness.

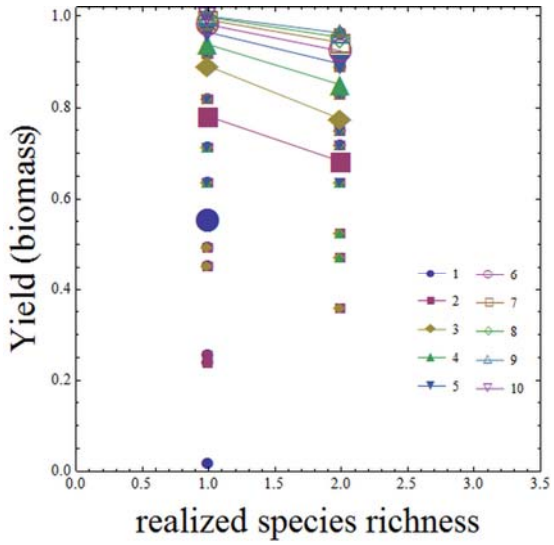


Figure S2. Relationship between realized richness and yield for the strongly complementary case in a homogeneous environment (scenario 7, Table 1). Each small point represents an equilibrium community resulting from seeding a given subset of the global species pool, coded by levels of introduced richness (from 1 to 10). For each level of introduced richness, the large points give the average yield across each level of realized richness. The lines connecting these points highlight the fact that within scenario 7, total biomass is negatively related to realized species richness for any given level of introduced richness.

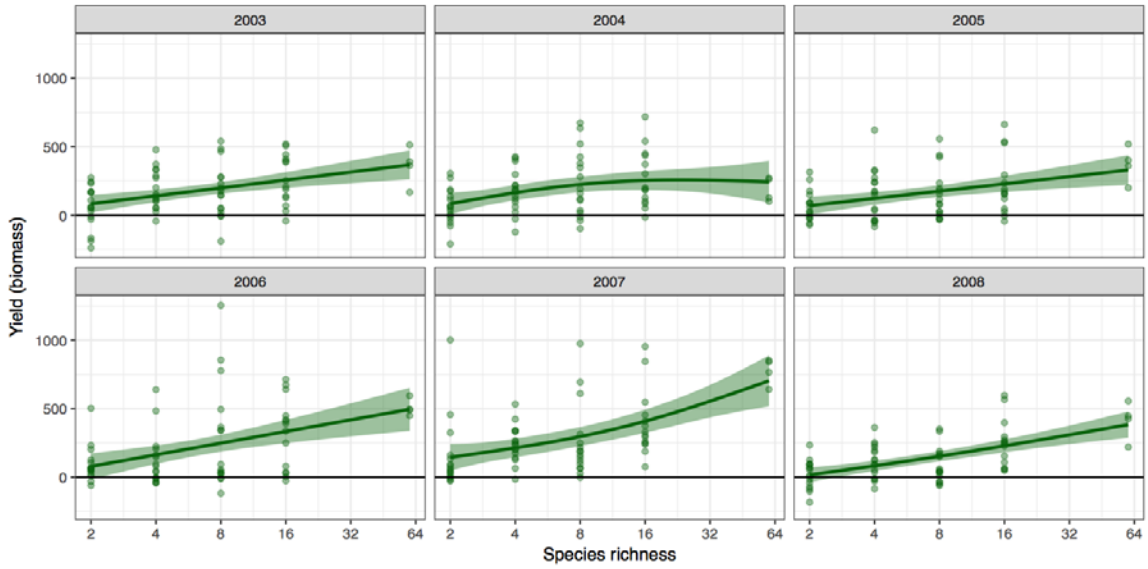


Figure S3. Jena yield (biomass) over time, relative to hypothetical community based on monoculture performance, as a function of introduced richness (log scale). Complementarity-Selection analysis (Fig. 4) partitions these responses.

Facilitation model

To study the effects of facilitation on the Complementarity-Selection Effects and the Community Assembly components, we developed a variant of the resource competition model from the main text specifically tailored to model the interactions between nitrogen-fixing and non-fixing terrestrial plants. This model is very similar to the model of Koffel et al. (2018), with a few differences that will be highlighted below. This model still fits the general form of Eq. (1), with now R_1 and R_2 respectively being soil Phosphorous (P) and Nitrogen (N). Yet, the expressions of both the per capita growth rate and the consumptive impacts differ from the general form from Schreiber and Tobiason (2003) presented in the main text. The growth rate $g_j(R_1, R_2)$ here reads:

$$g_j(R_1, R_2) = \mathrm{Min}(x_{1j}R_1, x_{2j}R_2 + f_j)$$

As the two resources R_1 and R_2 correspond to P and N, they are modeled as being strictly essential ($\alpha = -\infty$ in the approach presented in the main text) which results in taking the minimum of the growth rates on the two resources via the "Min" function.

The extra term that is not accounted for by the Schreiber and Tobiason (2003) approach is the constant fixation rate f_j that boosts growth when N is limiting. The fixation rate f_j is the focal trait here, and the species pool was generated by selecting species with varying values of f_j. A species with $f_j = 0$ is a non-fixer, while species with non-zero values of f_j describe N-fixers with various intrinsic abilities to fix N. We included a cost to N-fixation in the form of a decreasing ability to acquire phosphorous through x_{1j} as f_j increased, mathematically given by $x_{1j} = h(f_j)$, with h a decreasing function. Conversely, soil nitrogen acquisition x_{2j} was assumed to be the same for all species, and thus taken constant.

For facilitation to happen, N fixed by the N-fixers has to become available to the other species. This is done by modifying the consumptive impacts to account for N restitution through recycling, potentially leading to net negative consumptive impacts when N restitution by an N-fixer is greater than N uptake. This leads to the following consumptive impacts on N:

$$C_{2j}(R_1, R_2) = q_{2j}\left[\mathrm{Min}\left(x_{1j}R_1, x_{2j}R_2\right) - \lambda_{2j}m\right]$$

The positive term in $C_{2j}(R_1, R_2)$ accounts for soil N uptake through growth, and is thus equal to $g_j(R_1, R_2)$ without the fixation term f_j because fixed N comes from the atmosphere, not the soil. The negative term is N restitution through recycling from dead plant biomass (generated at the rate m, assuming the same mortality rate for all species for simplicity) with an efficiency given by λ_{2j}. We assumed that the efficiency $\lambda_{2j}$was larger for species with large fixation rate f_j, translating the fact that N-fixing species tend to see their dead biomass more readily recycled. The coefficient q_{2j} is a stoichiometric coefficient similar to the ones introduced in the models from the main text.

On the P side, the consumptive impact reads:

$$C_{1j}(R_1, R_2) = q_{1j}[g_j(R_1, R_2) - \lambda_{1j}m]$$

The first term of the sum is here exactly equal to growth $g_j(R_1, R_2)$, which translates the fact that unlike N, all the P entering plant biomass has been taken up from the soil. We assume that P is also recycled, through an efficiency λ_{1j} assumed to be the same for all species for simplicity. The coefficient q_{1j} is again a stoichiometric coefficient.

The global species pool was generated by selecting 10 equally spaced values for the fixation rate f_j, and the corresponding phosphorous acquisition rates x_{1j} and N-

recycling efficiencies λ_{2j} given by the trade-offs detailed above. All the other parameters were chosen to be the same for these 10 species.

For every given subset of the species pool, community assembly was simulated numerically by seeding the species present in that subset altogether and letting them grow until an equilibrium was reached.

Because N-fixing species with large f_j can, under certain circumstances, have negative consumptive impacts $C_{2j}(R_1, R_2)$, their presence leads to the net enrichment of the N pool. This results in strong fixers facilitating non-fixers and weak fixers (species that have small fixation rates f_j) through two mechanisms: 1) a strong N-fixer can coexist at equilibrium with a weak fixer, supplying the latter with limiting N, preventing the weak fixer from going extinct or enabling it to reach a higher biomass than in the absence of the strong fixer, and 2) in a initially very N-limited environment that prevents the weak-fixer from establishing, a strong N-fixer can establish and lead to soil N accumulation. This makes establishment of the weak fixer possible, but eventually leads the strong fixer to its own demise as the weak fixer excludes it. This second mechanism is a pure effect of community assembly that can generate facilitation-driven succession, as described in detail in Koffel et al. (2018).