

Quantization of ecological interactions yields insights into food web assembly and dynamics

Justin D. Yeakel,¹ Mathias Pires, James O'Donnell, Marcus de Aguiar, Paulo Guimarães Jr, Dominique Gravel, and Thilo Gross

¹*School of Natural Sciences, University of California, Merced, Merced, CA 95340, USA*

Make this abstract great again.

Introduction

Model Description

The ENigMa Model We consider a framework incorporating multiple types of directional interactions between species, which, when paired, represent specific ecological relationships including trophic interactions, service-resource and service-service mutualisms, and commensalisms. We also introduce two types of nodes in our depiction of ecological networks: those representing *species* and those representing *objects*. Objects must be made by one or more species (here and henceforth referred to as engineers) and represent a modification to the available niche-space that can be utilized by other species in the community. Such alterations are to be considered in the abstract, but in empirical systems could represent an introduced compound, metabolite, or an alteration to the habitat/environment such as, for example, a burrow.

The ENigMa model consists of four directed interactions: e: eat, which specifies a dependency involving the exchange of biomass, n: need, which specifies a dependency that does not involve biomass flow (e.g. a reproductive service), i: ignore, the null interaction, and m: make, which connects a species to an object that it engineers. ‘Objects’ are interactive components that can be made by ≥ 1 species, and eaten, needed, or ignored by the others. The four directed interaction types describe specific dependencies that one species/object has on another, however it is the coupling of two opposing directed interactions that describe familiar ecological relationships (Table 1).

The $e \leftrightarrow i$ interaction describes a typical predator-prey relationship, where species 1 eats species 2, whereas aside from serving as a resource, species 2 does not interact with (ignores) species 1. Of course, a prey’s abundance does not *ignore* the effects of predation, however our framework operates at the scale of presence/absence rather than abundance, and we assume that if both species co-occur, they have positive population densities, such that the prey’s state (presence/absence) ignores the predator. A second type of trophic interaction is described by $e \leftrightarrow e$, where consumption is symmetric, which could be do to changing roles over an individual’s life-history. The $e \leftrightarrow n$ and $n \leftrightarrow n$ interactions describe service-resource and service-service mutualisms, respectively. In the case of the former, one species interacts by

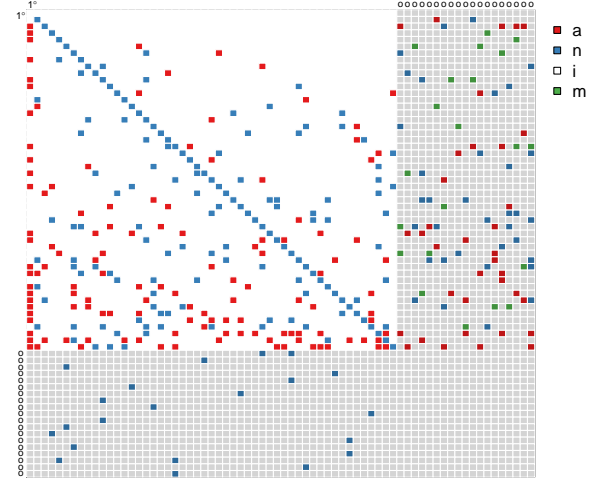


Figure 1: An example of the source pool interaction matrix where $S = 50$. Species and objects are aligned across the rows and columns; objects are shaded and labeled by ‘o’ to distinguish them from species. The interaction recorded in row i and column j describes the directed interaction from species/object i to species/object j . The first row/column represents the basal resource; species that assimilate the primary resource are capable of primary production. Species interact with other species and/or objects; objects only interact with their engineers by ‘needing’ them; objects do not interact with other objects.

way of a trophic interaction, whereas the other is provided a non-trophic need, such is the case in a plant-pollinator relationship. Unique to models of ecological networks, the $m \leftrightarrow n$ interaction describes ecosystem engineering, where a species makes an object, whereas the presence of the object ‘needs’ the presence of the species that makes it. Objects can be utilized by other species in the community, providing an indirect dependency that could be facilitated by multiple species (many engineers make the same object) and/or used by multiple species (many species eat or need the same object). As we will describe below, engineers can modify the niche space available to the community on timescales that last longer than the species themselves, which is an oft-assumed characteristic of engineering.

We explore the assembly of a novel community that emerges from a species source pool, which is represented by a source pool interaction matrix where all eat, need, ignore, and make interactions are established between

all species and objects. As such, a set of interactions for a particular species defines how it interacts with any other *a priori*, thereby establishing its potential interaction niche space. The source pool is used to seed a novel community, which arises as the result of colonization and extinction rules, the details of which we describe below.

Building the source pool The source pool interaction matrix \mathbf{P} is generated by first setting the number of species in the pool \mathcal{S}_P and determining the number of objects \mathcal{O}_P that are made by ecosystem engineers. The resulting matrix is $\mathcal{N}_P \times \mathcal{N}_P$ where $\mathcal{N}_P = \mathcal{S}_P + \mathcal{O}_P$, and is subdivided into four quadrants, only two of which play a role here: species-species interactions and species-object interactions (see Fig. 1). In these two quadrants, the expected frequency of eat interactions $E\{f_e\}$ and the expected frequency of need interactions $E\{f_n\}$ are free parameters, as is the expected number of objects made per species $E\{\mathcal{O}_i\} = \eta$. Here and throughout, we simplify this parameter space by assuming that the frequency of eat and need interactions for species-species (SS) interactions and species-object (SO) interactions scale as ϕ , such that $E_{SS}\{f_e\} = \phi E_{SO}\{f_e\}$ and $E_{SS}\{f_n\} = \phi E_{SO}\{f_n\}$. For each species, a set number of objects is drawn from $\text{Poiss}(\eta)$, such that the expected proportion of species that are engineers (species that make objects) is $1 - e^{-\eta}$. If a particular object is randomly and independently drawn for a given engineer from a complete list of all possible objects, such that multiple species – with some probability – can make the same object, the expected number of objects is

$$E\{\mathcal{O}_P\} = \mathcal{S}_P \eta \left(1 - \frac{1}{e}\right), \quad (1)$$

where e is Euler's number. The frequency of $m \leftrightarrow n$ interactions is then calculated as

$$E\{f_m\} = \frac{\eta}{\mathcal{S}_P (1 + \eta - \frac{\eta}{e})^2}. \quad (2)$$

Finally the frequency of the ignore interaction is calculated as $E_{SS}\{f_i\} = 1 - E_{SS}\{f_e\} + E_{SS}\{f_n\}$ and $E_{SO}\{f_i\} = 1 - E_{SO}\{f_e\} + E_{SO}\{f_n\} + E_{SO}\{f_m\}$ for species-species and species-object interactions, respectively. Pairwise interaction probabilities between both species and objects are then calculated as shown in Table I. These pairwise interactions are assigned randomly from these probabilities between species-species and species-objects independently in both quadrants, such that the source pool matrix has no imbued structure apart from the number of species, the number of objects, and the frequency of each directional interaction type. Each source pool is provided a *basal resource* (the first row/column). A species with a single trophic interaction to this resource is identified as a pure autotroph (Fig. 1), however the basal resource does not have eat, need, or make interactions itself.

Colonization and Extinction Assembly of a species community is the result of both local colonization and extinction of species that are drawn from the source pool.

The realized interactions within the assembled community \mathbf{A} are thus a subset of the potential interactions observed if every species were present (as recorded in the source pool \mathbf{P}). We determine the ability of a species to colonize a community as a function of two conditions: 1) the colonizing species must eat *at least one* species/object in the community, and 2) the colonizing species must satisfy all of its need interactions; if these conditions are both satisfied, colonization is possible. At each time-step, one potential colonizer that fulfills these conditions is selected at random and added to the community, as well as the objects that it makes if it is an engineer. Thus, in the first time-step, only species that consume the primary resource (row 1; figure 1) and do not have any 'need' interactions can initiate the assembly process.

Extinction occurs directly via competitive exclusion, or indirectly via the subsequent loss of a consumer's single resource or any of the species/objects it needs. Extinction due to competitive exclusion is determined by violation of a single condition: a species must be the strongest competitor for at least one of its food resources. We assess this by assuming that each species i in a community has a competition strength σ_i that is compared to every species j that shares each of its resources. If σ_i is not the highest σ for at least one its resources, then species i is competitively excluded from the community along with all unique objects that it makes. The competition strength for species i is set to increase as the sum of its potential need interactions, and decrease as both the sum of its potential eat interactions and the sum of its realized predators (its vulnerability) in the current community, such that

$$\sigma_i = \pi \sum_{j=1}^{\mathcal{N}_P} n_{P(i,j)} - \sqrt{2} \sum_{j=1}^{\mathcal{N}_P} e_{P(i,j)} - \sum_{i=1}^{\mathcal{N}_A} e_{A(i,j)}, \quad (3)$$

where the summations describe the number of need interactions, eat interactions, and predators, from left to right, and $\mathcal{N}_A = \mathcal{S}_A + \mathcal{O}_A$ is the sum of species and objects in the current community rather than the source pool. The coefficients serve only to prevent the substitution of different interaction types.

Why do we assume that mutualisms increase a species' competitive strength? Although mutualisms serve to tie the existence of one species to another, which increases its risk exposure, we assume that this dependency evolved as the consequence of a fitness advantage inherent to the interaction, which we assume here translates to a competitive edge. Conversely, we assume that specialists (species with fewer trophic interactions) are competitively superior to generalists (ref), and that as a species spends more energy avoiding predation, it spends less energy competing, both of which serve to erode its competitive edge. Importantly, we emphasize that the role of mutualisms and trophic interactions in determining a species' competition strength is with respect to its *potential* interaction niche, and thus calculated from the source pool matrix \mathbf{P} , whereas its vulnerability to predation is determined

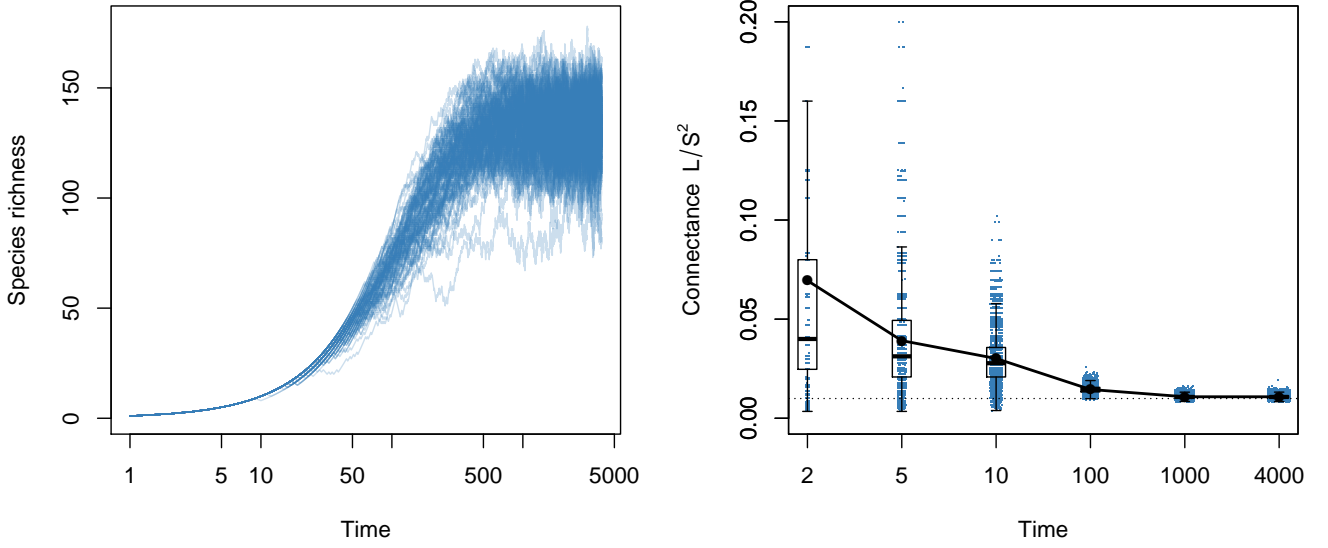


Figure 2: a) Species richness over time. b) Trophic connectance over time.

by a species' predators within the assembled community \mathbf{A} . Therefore, the first two factors of σ_i (the influence of its need and eat interactions) are independent of the community, whereas its vulnerability changes with community structure. We note that due to the threshold conditions for colonization, $\sum_{j=1}^{N_P} n_{P(i,j)} = \sum_{j=1}^{S_A} n_{A(i,j)}$.

We integrate these colonization and extinction rules to simulate community assembly over time using a Gillespie algorithm. At each time-step a single event is chosen at random to iterate the simulation forward, where possible event types include: 1) species colonization, 2) species extinction, and 3) object extinction. The likelihood of drawing each event increases with the number of potential colonizers (n_c) or the number of species (n_s) and objects (n_o) that meet the conditions required for extinction. The change in time dt at each step in the simulation is therefore dynamic, where $dt = (n_c(t) + n_s(t) + n_o(t))^{-1}$. This algorithm allows, with some probability, both species and objects to remain in the system after they are selected for extinction, however this probability declines as the rate of colonization (number of potential colonization events, n_c) decreases and the rate of extinction (number of potential extinction events, $n_s + n_o$) increases. To incorporate the idea that the timescale of objects in a system should be longer than the timescale of the engineers that made them, we bias the algorithm to select a colonization or species extinction event with weight τ . If $\pi_s = n_s/n$ and $\pi_o = n_o/n$, species extinction occurs with probability $\tau\pi_s/(\tau\pi_s + (1-\tau)n_o)$ vs. an object extinction event that occurs with probability $(1-\tau)n_o/(\tau n_s + (1-\tau)n_o)$: as τ increases from $\tau = 1/2$, the likelihood that an object is eliminated from the sys-

tem declines; if $\tau = 1/2$, there is no imposed delay in the extinction of objects.

Results

Diverse interactions without engineers Community assembly in the absence of engineers ($\eta = 0$) reveals the emergence of food web and mutualistic network properties consistent with observations of assembling and steady state empirical systems. Unless otherwise specified, we set $S_P = 200$, $f_e = 0.01$, $f_n = 0.002$, and $t_{\max} = 4000$. Because only primary producers that do not have outgoing need interactions can colonize initially, a diverse base of autotrophic species typically constitutes the early assembly process. (We note that in the ENIgMa framework, generalist consumers can function as specialists if few of their resources are available initially, and may become more generalist over the course of assembly.) In order for communities to have > 1 pure autotroph, we do not consider competitive exclusion of the basal resource, such that all non-mutualistic pure autotrophs have the potential to coexist. Following the establishment of a suite of autotrophs, both mixotrophs and lower trophic-level heterotrophs assemble into the community (Fig. 2a). As species richness increases, available resources accumulate consumers and competitive exclusion leads to an increase in the extinction rate until a steady state is reached at $S_A^* = 130$ species. This community steady state increases as the number of mutualisms established in the source pool decreases (lower $E\{f_n\}$) because mutualisms introduce dependencies that inhibit coloniza-

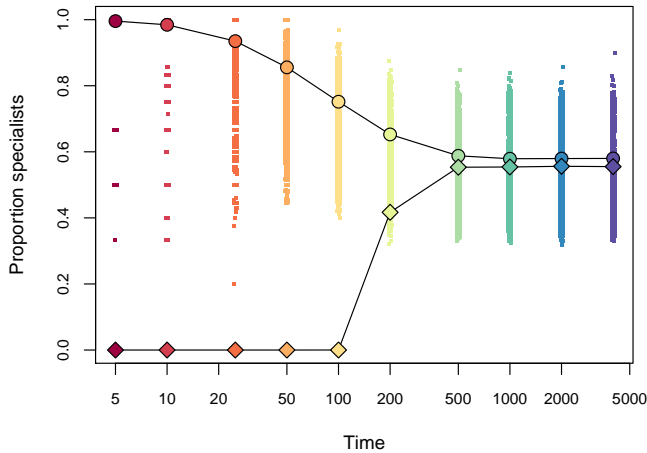


Figure 3: The proportion of specialists as a function of assembly time, where a specialist is defined as a species with a generality index $G_i < 1$. All measures of G_i are scaled by the average number of links per species L/S , and we consider different values of L/S on G_i : Circles: G_i^{all} where L accounts for all links in the food web and S accounts for all species relative to each time interval in the assembly process (averaged across replicates); Points: G_i^{hetero} , where we consider only the links and species richness of heterotrophs, excluding autotrophs (each point shows an individual replicate); Diamonds: G_i^* , where L and S are measured with respect to the communities at steady state, which is most similar to the measure used to evaluate assembling mangrove food webs (averaged across replicates).

tion.

As the community assembles, we find that the connectance of trophic interactions ($C = L/S^2$) where in our case $L = \sum_{i,j} e_{A(i,j)}$ and $S = S_A$) follows a decay-like trajectory to values similar to – but on average 9% greater than – the connectance of the source pool P (Fig. 2b). Decaying connectance has been documented in the assembly of mangrove communities (Piechnik), however this decay is a statistical inevitability, as a growing food web early in the assembly process inevitably has high link density (few species that are nearly fully connected), and over the course of increasing species richness and the establishment of different trophic levels and compartments, a decline in link density. That the connectance of assembled communities is greater than the source pool is due to the fact that only species connected by trophic interactions can enter the community to begin with, increasing expected link density compared to the overall pool.

Recent empirical work has indicated that generalist species may play an important role early in community assembly, whereas specialists tend to colonize after a diverse resource base has accumulated. Because the definition of a specialist or generalist to some degree depends on the size and connectance of the larger

food web, trophic generality can be defined as $G_i = \sum_j e_{A(i,j)} (L/S)^{-1}$, such that the number of trophic interactions for a consumer is scaled by the average number of trophic interactions per species in the community L/S (Piechnik, others). A species is classified as a generalist if the number of its trophic interactions is greater than the average number of links per species, or $G_i > 1$, and a specialist if $G_i < 1$, where a community can be described by the proportion of specialists found therein.

For interaction networks that are assembling over time, generality can be scaled by a number of different measures of L/S , and this has a large effect on our interpretation of the role of generality in community assembly. For instance, L/S may be quantified by either including all autotrophic species or only autotrophic functional groups. Furthermore, the scaling of generality may be made with respect to the current state of the community at each point in time, or with respect to the community at steady state. For instance, in their investigation of assembling mangrove food webs (originally described by Simberloff, xxx), Piechnik et al. (2008) scaled trophic breadth to a standard steady state value of $L^*/S^* = 0.2$ averaged across 102 food webs. To examine how our assessment of the role of generalism over the course of assembly changes based on the application of different scalings, we employ three different measures of L/S to calculate G_i : 1) G_i^{all} , where L accounts for all links in the food web and S accounts for all species relative to each time interval in the assembly process (circles; Fig. 3b); 2) G_i^{hetero} , where we consider only the links and species richness of heterotrophs, excluding autotrophs (points; Fig. 3b); 3) G_i^* , where L and S are measured with respect to the communities at steady state, which is most similar to the measure used to evaluate assembling mangrove food webs (diamonds; Fig. 3b).

Whether trophic breadth is scaled to the current state of L/S or the steady state value of L^*/S^* has a large influence on the estimated proportion of generalists in the community, particularly when the size of the system is small. We observe that for G_i^{all} , the system is initially assembled by specialist species, though over the course of assembly the proportion of specialists relative to generalists declines to intermediate values (circles representing the average over replicates in Fig. 3). If only the trophic links between non-autotrophs are considered as in G_i^{hetero} , specialists still dominate early in assembly, but there is a greater range, such that some systems can be described by a mixed proportion of specialists and generalists (individual points representing independent replicates in Fig. 3). If generalism is measured with respect to the steady state L^*/S^* as in G_i^* , we observe that generalists dominate early in assembly, with an increase in specialists as assembly progresses (diamonds representing the average over replicates in Fig. 3). At steady state, all measures of L/S are approximately equivalent, and the proportion of specialists levels out at ca. 60%, which is similar to the empirical observations for Simberloff’s mangrove communities in Piechnik et al. (2008).

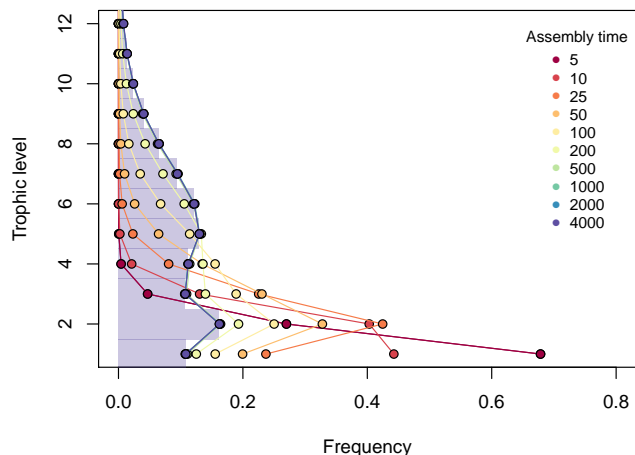


Figure 4: The frequency distribution of trophic levels over the course of assembly taken across 10^4 replicates. Autotrophs occupy trophic level 1.

The role of specialists early in assembly is primarily due to the accumulation of autotrophs specializing on the basal resource. This is evident when we observe that the trophic level distribution early in assembly is peaked at the lowest trophic level (trophic level 1). Four trophic levels are typically established by $t = 50$, where colonization is still the dominant dynamic, and by the time communities reach steady state the interaction networks are characterized by ca. 10 trophic levels (Fig. 4). The distribution of trophic levels changes shape over the course of assembly: early on, we observe that the community exhibits a pyramidal structure, where the vast majority of species inhabit low-trophic positions. At steady state, we observe that intermediate trophic levels (2-7) dominate, with frequencies that reveal an hour-glass structure. We emphasize that these structures are diversity-weighted rather than biomass or abundance-weighted as is often the case (Trebilco et al. 2013, Gibert & Yeakel 2019). Trophic levels higher than 7 do occur, but are increasingly rare.

Because the ENIgMa framework includes multiple types of interactions, we must examine whether structures characteristic of mutualistic networks are observed. Empirical observations of mutualistic systems reveal that such interactions tend to be nested (where specialist interactions are subsets of generalist interactions). Increasing the frequency of need interactions increases the fre-

quency of both service-resource ($e \leftrightarrow n$) and service-service ($n \leftrightarrow n$) mutualisms. In the ENIgMa framework, the probability of competitive exclusion is reduced with an increase in the prevalence mutualistic interactions, and this should lead more stable nested bipartite motifs over the course of assembly (see Supplementary Information). Our expectation then, is that nestedness should

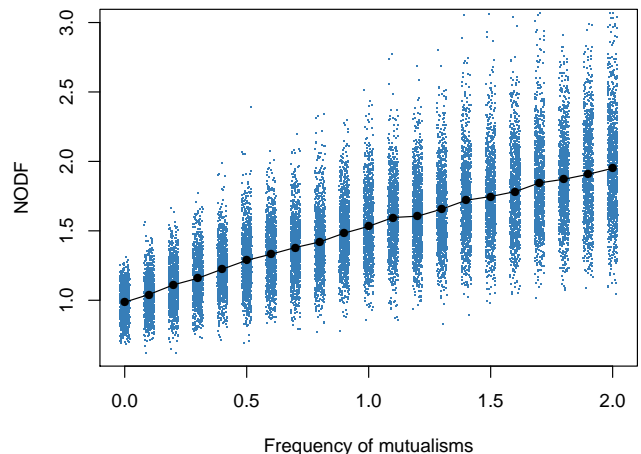


Figure 5: Nestedness (measured as NODF) as a function of the frequency of need interactions in the species pool interaction matrix P .

increase with the frequency of mutualisms, though this is difficult to predict a priori. As we increase the frequency of need interactions in the source pool, we indeed observe an increase in nestedness (measured as NODF; Fig. 5). That the absolute values of nestedness are low compared to those measured for empirical mutualistic networks is unsurprising: observations of mutualistic interactions are generally for bipartite networks and isolated to specific systems (e.g. ant-plant mutualisms). Here the NODF metric is taken across both eat and need interactions across the entire assembled community.

The role of engineers

Discussion

Appendices

Parameter	Definition	Value/Range
\vec{a}	assimilate	
\vec{n}	need	
\vec{i}	ignore	
\vec{m}	make	
$e \leftrightarrow i$	Asymmetric consumption	$p_{ei} = p_i(p_e/(p_e + p_n + p_i)) + p_e(p_i/(p_a + p_i + p_n))$
$e \leftrightarrow e$	Symmetric consumption	$p_{ee} = p_e(p_e/(p_i + p_n + p_e))$
$e \leftrightarrow n$	Trophic mutualism	$p_{en} = p_n(p_e/(p_e + p_n + p_i + p_m)) + p_e(p_n/(p_a + p_i + p_n))$
$n \leftrightarrow n$	Non-trophic mutualism	$p_{nn} = p_n(p_n/(p_e + p_n + p_i + p_m))$
$n \leftrightarrow i$	Commensalism	$p_{ni} = p_n(p_i/(p_e + p_n + p_i + p_m)) + p_i(p_n/(p_e + p_n + p_i))$
$m \leftrightarrow n$	Engineering	$p_{mn} = p_n(p_m/(p_e + p_n + p_i + p_m)) + p_m$
$i \leftrightarrow i$	Null	$p_{ii} = p_i(p_i/(p_e + p_n + p_i))$
\mathcal{N}	Number of species + objects	dyn.
\mathcal{S}	Number of species	dyn.
\mathcal{O}	Number of objects	dyn.

Table I: Table of parameters, definitions, and assigned values or ranges.