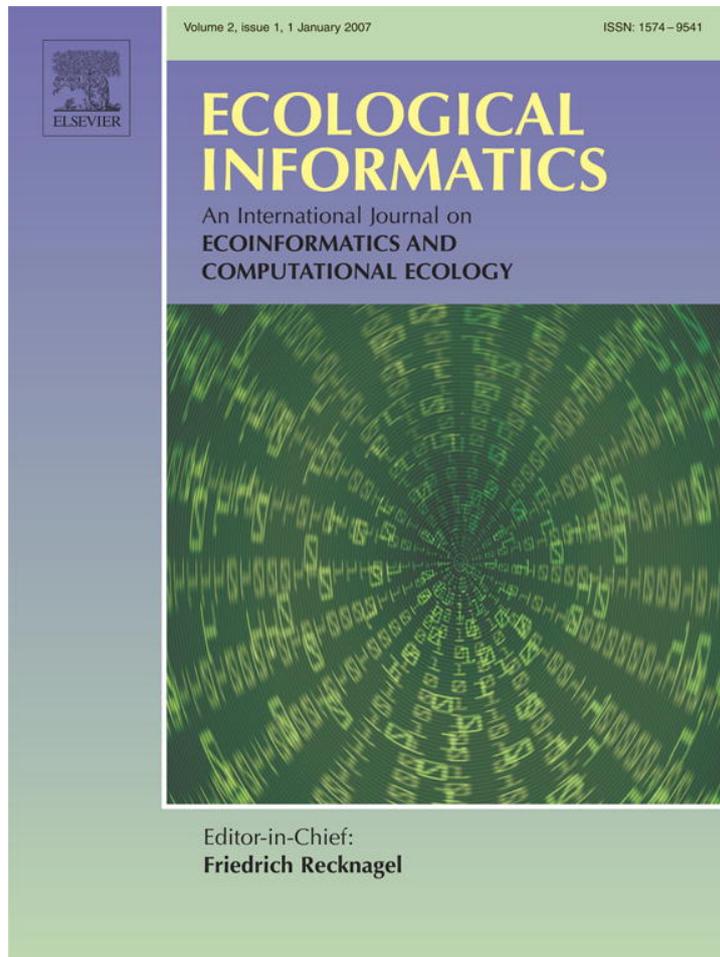


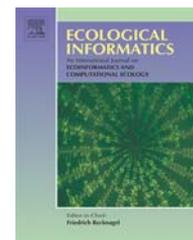
Provided for non-commercial research and educational use only.
Not for reproduction or distribution or commercial use.



This article was originally published in a journal published by Elsevier, and the attached copy is provided by Elsevier for the author's benefit and for the benefit of the author's institution, for non-commercial research and educational use including without limitation use in instruction at your institution, sending it to specific colleagues that you know, and providing a copy to your institution's administrator.

All other uses, reproduction and distribution, including without limitation commercial reprints, selling or licensing copies or access, or posting on open internet sites, your personal or institution's website or repository, are prohibited. For exceptions, permission may be sought for such use through Elsevier's permissions site at:

<http://www.elsevier.com/locate/permissionusematerial>

available at www.sciencedirect.comwww.elsevier.com/locate/ecolinf

Multi-objective optimization of an ecological assembly model

Pascal Côté^a, Lael Parrott^{a,*}, Robert Sabourin^b

^aComplex Systems Laboratory, Département de Géographie, Université de Montréal, C.P. 6128, Succursale Centre-Ville, Montréal, QC, Canada H3C 3J7

^bÉcole de technologie supérieure (ETS), Laboratoire d'Imagerie, de Vision et d'Intelligence Artificielle, 1100, rue Notre-Dame Ouest, Montréal, QC, Canada H3C 1K3

ARTICLE INFO

Article history:

Received 22 September 2006

Received in revised form

16 January 2007

Accepted 18 February 2007

Keywords:

Assembly models

Food webs

Community assembly

Genetic algorithms

Multi-objective optimization

Lotka–Volterra dynamics

ABSTRACT

The aim of the present work is to use multi-objective evolutionary algorithms (MOEA) to parameterise an ecological assembly model based on Lotka–Volterra dynamics. In community assembly models, species are introduced from a pool of species according to a sequence of invasion. By manipulating the assembly sequences, we look at the structure of the final communities obtained by a multi-objective process where the goal is to optimize the productivity of the final communities. The MOEA must also meet the constraint that the communities constructed in this fashion have a specified connectance. The Non-dominated Sorting Algorithm (NSGA-II) and the Strength Pareto Evolutionary Algorithm (SPEA2) were employed to optimize sequences according to the multi-objective optimization problem. The results show that the assembly process using optimized sequences generated different community structure than those generated via random sequences. First, the assembled communities are much more productive than those obtained from random sequences. We show that this increase of productivity is due to the degree distribution of the community food web, which was reshaped by the optimization process. In addition, using identical regional species pools the MOEAs were able to generate communities of different expected connectances. These results demonstrate the effectiveness of NSGA-II and SPEA2 for optimizing parameters in ecological models.

© 2007 Elsevier B.V. All rights reserved.

1. Introduction

Ecological assembly models are computer algorithms that use a set of rules or equations to represent the interaction among species in a community (Pimm, 1980; Drake, 1990; Case, 1990; Lockwood et al., 1997; Hewitt and Huxel, 2002; Fukami, 2005). The distinctive character of assembly models is that species in the community originate from a pool of species, identified as the regional species pool (RSP), and are introduced sequentially according to a random sequence of invasion. The possible consequences that the invasive species have on the community are the extinction of invasive species with or without secondary extinction(s) or successful invasion of the community with or without secondary extinction. Assembly

models show that, from the same RSP, different random sequences lead to a branching of possibilities for community assembly (Drake, 1990; Zimmermann et al., 2004; Lockwood and Samuels, 2004; but see Morton et al., 1996). This is a very important result since it emphasizes the role of historical events. Moreover, certain sequences can lead to communities with a complete resistance to invasion by other species in the RSP (Case, 1990; Morton and Law, 1997).

Most theoretical studies based on assembly models employ random sequences of invasion (Drake, 1990; Case, 1990; Hewitt and Huxel, 2002; Fukami, 2004, 2005; but see Côté and Parrott, 2006). This is due to the number of possible assembly sequences that can be generated. In fact, when the number of species in the RSP is large, random sequences represent

* Corresponding author.

E-mail address: lael.parrott@umontreal.ca (L. Parrott).

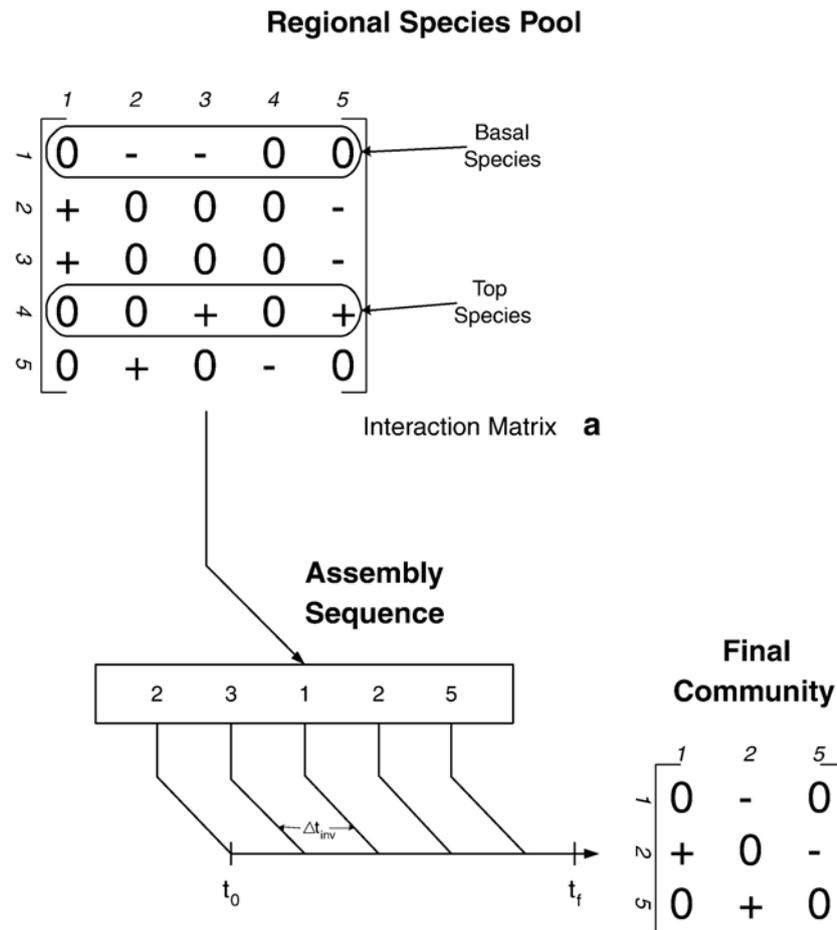


Fig. 1 – The ecological assembly model. Species are introduced in a community via a regional species pool according to a specific assembly sequence. The final community is composed of successful colonizers from the regional species pool.

only a small subset of the space composed of all the possible sequences. Hence, random ordering of assembly sequences does not constitute an appropriate and efficient tool to explore the entire parameter search space of these kinds of ecosystem models. In Côté and Parrott (2006), a single criteria optimization problem, where the objective was to maximize the diversity of a community, was used to demonstrate the effectiveness of evolutionary algorithms to explore the search space formed by all possible assembly sequences from a regional species pool. The results indicated that the ordering of species in community assembly sequences has a large impact on the community structure and that a single criteria optimization problem is an effective method to investigate the effect of historical events. Based on the same regional species pool, it has been shown that random sequences produce communities of inferior diversity compared to communities generated via optimized sequences.

We propose an application of evolutionary algorithms for searching the space of all possibilities in order to find assembly sequences that generate communities having certain desired characteristics. We apply an approach that uses multi-objective evolutionary algorithms (MOEA) to find assembly sequences satisfying a constrained multi-objective optimization problem. MOEAs are especially designed to explore large search spaces formed by several objective functions (Coello Coello, 1999;

Zitzler et al., 2000; Deb et al., 2002). MOEA have been successfully applied in a number of domains especially in engineering applications (see Coello Coello et al. (2002) for an exhaustive review and Coello Coello and Lamont (2004) for recent applications). To the best of our knowledge, only Reynold and Ford (1999) have implemented an MOEA for use in an ecological modeling context. They utilize a genetic algorithm, with a Pareto ranking fitness evaluation, to find a good parameter setting for a forest management model. In the field of ecological informatics, evolutionary computation has been applied to simulate the evolution of foraging strategies of individual species in a changing environment like in the Echo model (Holland, 1992; Hrabec and Milne, 1997) and in the EUZONE model (Downing, 1997), to model ecological time series (see Recknagel, 2002) or to optimize the stability of a food web model (Ruiz-Moreno et al., 2006) to name but a few applications.

In this paper we extend the results of Côté and Parrott (2006) by investigating further the role of assembly sequence on community structure by defining a multi-objective optimization problem. We show that it is possible to optimize community productivity while also producing a specific community structure. For example, in restoration projects, it could be of great help to evaluate which assembly sequences generate the more productive community having a specific food web structure. To address these questions, we introduce

a multi-objective optimization problem (MOOP) that deals with an assembly process to minimize the “assembly effort” and to maximize community productivity. The “assembly effort” in this specific context is the total sum of all population sizes of species introduced in the community. In other words, it represents the amount of biomass introduced in the community by the assembly process. The productivity of the final community is simply obtained by summing the population size of each species present in the community after the assembly process. We use both objective functions because if we optimize only the productivity of the final community by an optimization process, we must control the input biomass to be sure that this value does not increase as the optimization of the productivity goes on. We decide not to set the input biomass to a fixed value but rather minimize it to investigate if it is possible to maximize the productivity of a community assembly model by minimizing the input biomass. We solve the problem by using results obtained from two algorithms that are highly competitive: NSGA-II and SPEA2, to show that the results obtained are independent of the method used.

2. Ecological assembly model

We use an assembly model similar to the one designed by Montoya and Solé (2003). The first step of this process is the creation of a regional species pool (RSP) composed of basal species (prey only), intermediate species (both prey and predator) and top species (predator only). Secondly, an assembly sequence is defined to sequentially introduce species in a community. The species in this community interact through a system of equations based on Lotka–Volterra dynamics.

2.1. System dynamics

The population size of a species i in a community Y evolves through Lotka–Volterra community dynamics, defined as follows:

$$x_i(t+1) = x_i(t) + \Delta t \left[x_i(t) \left(\left(1 - \frac{x_i(t)}{K}\right) b_i + \sum_j^Y a_{ij} x_j(t) \right) \right] + \epsilon_i(t) \quad (1)$$

where $x_i(t)$ is the population size of species i at time t , a is the community matrix used to model the interaction of species i on species j in the community Y (see Fig. 1), $b_i > 0$ is the growth rate of basal species whereas $b_i < 0$ is the mortality rate of top or intermediate species, the term $1/K$ is related to the carrying capacity of the environment and is non-zero only for basal species and finally, $\epsilon_i(t)$ is the increase of the population size due to colonization events. Rates of mortality (or growth for basal species), b_i , are chosen from a uniform distribution varying from $0 < b_i \leq 1$ for basal species and $-1 \leq b_i < 0$ for other types of species. When the system of Eq. (1) is iterated, we consider a species to be extinct if its population size falls below x_{\min} .

2.2. Regional species pool

We model the interactions between species in the RSP by a community matrix a where a value of $a_{ij} > 0$ corresponds to the

interaction strength that predator i has on prey j while $a_{ij} < 0$ is the interaction strength that predator j has on prey i . A value of $a_{ij} = a_{ji} = 0$ means that there is no predator–prey relation between both species. Thus a basal species is defined by a row in the matrix a composed of only negative and zero values (no predation). Inversely, top predators are defined by a row of positive and zero values (see Fig. 1). The probability of connection (p_c) of species i to j in the regional species pool, i.e., the probability that species i eat species j , is fixed and identical for all species. Interaction strengths were derived from a uniform distribution on the interval $0 \leq a_{ij} \leq a_{\max}$ for top species, $-a_{\max} \leq a_{ij} \leq 0$ for basal species and $-a_{\max} \leq a_{ij} \leq a_{\max}$ for intermediate species (Montoya and Solé, 2003).

2.3. Assembly sequences

From this RSP, we select a series of species to be introduced in a community. All species from this list, named the assembly sequence l with an initial population size identified by d . Thus, the increases in the population size of a species i due to colonization events $\epsilon_i(t)$ (Eq. (1)) are set as follows:

$$\epsilon_i(t) = \begin{cases} d(k) & \text{if } \{I(k) = i \wedge k = t/\Delta t_{\text{inv}}\}, k = \{1 \dots n\} \\ 0 & \text{elsewhere} \end{cases} \quad (2)$$

where n is the number of colonization events, Δt_{inv} is the time between two invasion events. We iterate the assembly model from time t_0 to time t_f and each species is introduced sequentially at time t_k . We set the time between two invasion events equal for all species so that $t_{k+1} - t_k = \Delta t_{\text{inv}}$ (see Fig. 1). While this setting most probably influences the assembly process, we have chosen to use a fixed time between two invasion events since this approach is commonly used with this kind of model (Drake, 1990; Case, 1990; Montoya and Solé, 2003). We allow for the possibility of recolonization so if any species i in an assembly sequence l is introduced twice, we update its current population size by adding $d(k)$.

3. Multi-objective optimization

3.1. Optimization problem

With the assembly model at hand, we defined a multi-objective optimization problem (MOOP) in which the first objective corresponds to the minimization of the assembly effort and the second objective is to maximize the productivity of the final community. The MOOP is defined as follows:

$$\min f_1 = \sum_{k=1 \dots n} d(k) \quad (3)$$

$$\max f_2 = \sum_{i \in Y} x_i(t_f) \quad (4)$$

$$\text{s.t. } C_{\min} \leq C \leq C_{\max} \quad (5)$$

The first objective f_1 is calculated by the sum of all initial population sizes of species when they are introduced for the first time or reintroduced due to recolonization of the

community. After completing a simulation of the assembly model, that is when all species in a sequence l have been introduced, the second objective f_2 is calculated by the sum of the population sizes of each species present in the community at the end of the assembly process.

As we deal with a regional species pool that is defined as a predation network, we must consider the connectance of the food web. The connectance (C) is the number of active links divided by all possible links in a network. Here the network, also called a food web, is formed by all species in the community (nodes) having predation–prey relationships (links). If we exclude cannibalism, the connectance is defined by:

$$C = \frac{L}{S(S-1)} \quad (6)$$

where L is the number of links and S is the diversity of the community. Connectance and diversity play an important role in the characterization of food webs (Martinez, 1992; Dunne et al., 2002a,b). Compared to other networks such as random, scale-free (Barabasi and Albert, 1999) or small-world networks (Strogatz, 2001), food webs possess high connectance values for their relatively low sizes (Dunne et al., 2002a). While empirical data show that there is no specific value for C , we can assume that a real ecosystem should be restricted to a particular range (Dunne et al., 2002a). Thus, we add a constraint to the MOOP to ensure that C falls between a realistic minimum and a maximum value. The range is fixed at the beginning of the simulation. We hypothesize that, if the MOEA is able to respect this constraint, this new adaptable approach could help ecologists to virtually recreate a specific food web structure.

The evaluation of a solution requires complete simulation of the assembly process. A solution is defined as (i) an assembly sequence l and (ii) a vector of real values d used to initialize and/or update (in the case of recolonization) the population size of each species. In other words, a solution s is a pair of vectors $\langle l, d \rangle$ where l is a vector of integer values that represent the sequence of invasion in the community (i.e., an element $l(k)$ corresponds to the index i in the RSP of the species introduced at time t_k , see Fig. 1) and a vector of real values for the initial population sizes d .

3.2. Optimization techniques

Most multi-objective algorithms are based on the concept of domination to evaluate the fitness of a solution. The concept of domination allows evaluating if a solution is better than another solution without using a weighted sum method. Let us assume that we have to deal with M objective functions to be minimized. A solution s_1 dominates a solution s_2 ($s_1 \succeq s_2$) if the following conditions are true:

$$f_m(s_1) \leq f_m(s_2), \quad \forall m = 1 \dots M \quad (7)$$

$$\exists m \{ f_m(s_1) < f_m(s_2) \} m = 1 \dots M \quad (8)$$

All members of a population P that are not dominated by any other member of this population form the non-dominated set F or the first Pareto front defined as:

$$F = \{s_1 \succeq s_2 \mid \forall s_1 \in F \wedge s_2 \in P \setminus F\} \quad (9)$$

The NSGA-II (Deb et al., 2002) is an elitism version of NSGA (Srinivas and Deb, 1994) where the fitness of a solution is obtained by a Pareto ranking procedure. Population $R_{\tau+1}$ in NSGA-II is constructed as follows: after the calculation of the objective functions of all solutions, the population R_τ is sorted using the crowded Pareto ranking procedure. The first half of the solutions are directly introduced in a population $P_{\tau+1}$. Selection, crossover and mutation operators are applied on $P_{\tau+1}$ in order to create the offspring population $Q_{\tau+1}$. Iteration terminates when the size of $Q_{\tau+1}$ is equal to the size of $P_{\tau+1}$ with the construction of $R_{\tau+1} = P_{\tau+1} \cup Q_{\tau+1}$. Complete details of this algorithm can be found in Deb (2001) and Deb et al. (2002).

The algorithm SPEA2 (Zitzler et al., 2002b) is a revised version of the SPEA proposed by Zitzler and Thiele (1999). To calculate the fitness of a solution, the Pareto ranking procedure is not required since the fitness is based on a strength value. Basically, the strength value of a solution corresponds to the number of solutions that dominated the solution to be evaluated. Hence, a zero value indicates a non-dominated solution. The SPEA2 uses two populations, P_τ and E_τ , where the E_τ is an external population set to keep the best non-dominated solutions found so far. Population $P_{\tau+1}$ is constructed as follows: all non-dominated solutions in P_τ and E_τ are copied in $E_{\tau+1}$. As the size of the external set is fixed, if the number of solutions exceeds the size of $E_{\tau+1}$ a truncation operator removes solutions according to a k -nearest neighbor distance. The iteration is completed by the application of selection, crossover and mutation on $E_{\tau+1}$ to create $P_{\tau+1}$. Complete details of this algorithm can be found in Zitzler and Thiele (1999) and Zitzler et al. (2002a).

3.3. Genetic operators

For the NSGA-II, the constraint handling is assumed by the “constrained crowded tournament” (see Deb et al., 2002, chapter 6). This selection procedure uses the constraint dominance operator (\succeq_c) where a solution s_1 “constrained-dominated” a solution s_2 ($s_1 \succeq_c s_2$) if one of the following conditions is true:

- (1) Solution s_1 is feasible and solution s_2 is infeasible;
- (2) Solutions s_1 and s_2 are infeasible but s_1 is “least” infeasible;
- (3) Both solutions s_1 and s_2 are feasible but $s_1 \succeq s_2$.

where \succeq is the dominance operator (Eqs. (7) and (8)). A solution s_1 wins the tournament if:

- (1) Solution $s_1 \succeq_c s_2$ or
- (2) Solution $s_1 = s_2$ but the “crowded” distance measure of s_1 is greater than s_2 .

In SPEA2, the fitness of a solution is directly based on the constraint dominance operator including a density measure. Therefore, solution s_1 wins the tournament if the fitness value (strength of a solution) is lower than that of solution s_2 .

The definition of a solution of our MOOP requires that we deal with both integer and real valued variables. So, we implement uniform crossover with random mutation for integer vectors (assembly sequence l) and the simulated binary crossover (SBX) (Deb and Kumar, 1995) for real vectors

(initial population sizes d). In SBX, two offspring are generated from two parents according to the equations:

$$d_1(k) = 0.5[(1 + \beta_{qi})d_i(k) + (1 - \beta_{qi})d_2(k)] \quad (10)$$

$$d_2(k) = 0.5[(1 - \beta_{qi})d_i(k) + (1 + \beta_{qi})d_2(k)] \quad (11)$$

where

$$\beta_{qi} = \begin{cases} \frac{1}{(2u_i)\eta_c + 1}, & \text{if } u_i \leq 0.5 \\ \left(\frac{1}{2(1-u_i)}\right)\eta_c + 1, & \text{otherwise} \end{cases} \quad (12)$$

with u_i , a random uniform number between $[0; 1]$. A high value of η_c produces “near-parent” offspring whereas a low value of η_c produces offspring far from their parents.

4. Simulations

Empirical data concerning natural food webs reveal that the connectance of these networks varies between communities from roughly $0.05 \leq C \leq 0.2$. Thus, we investigated if the MOEAs were able to assemble communities that met the objective functions while at the same time having realistic food web connectance. The first series of simulations, identified as “Low Connectance”, uses an expected connectance $\{C_{\min}, C_{\max}\} = \{0.05, 0.075\}$ and the second series of simulations, identified as “High Connectance”, uses an expected connectance $\{C_{\min}, C_{\max}\} = \{0.15, 0.175\}$. For low and high connectance series we executed 30 simulations with both algorithms. All simulations were done with a regional species pool defined as follows:

- (1) Regional species pool (RSP) size = 300 species
- (2) Basal species in the RSP = 100
- (3) Top species in the RSP = 100
- (4) Intermediate species in the RSP = 100
- (5) Expected connectance for the RSP $p_c = 0.2$
- (6) Maximum interaction strength $a_{\max} = 0.1$

We decided to set proportions of basal, top and intermediate species equal to eliminate any bias during the manipulation of the assembly sequences made by the algorithm. The size of the RSP is arbitrary and does not significantly affect the conclusions drawn from the results obtained, so long as the size is relatively greater than the length of the assembly sequence. The maximum interaction strength was set to a common value used with this kind of model (Montoya and Solé, 2003).

We initialize d with n random uniform numbers between $0 < d_n \leq 1$ and l by n randomly chosen species from the regional species pool. As recolonization is possible, there is no restriction on the presence of the same species twice in l . The parameters for the assembly model are defined by:

- (1) Number of invasion events = 50
- (2) Time between invasion events $\Delta t_{\text{inv}} = 10$
- (3) Threshold for extinction $x_{\min} = 0.001$
- (4) Carrying capacity of the environment $K = 10$
- (5) Simulation time $t_0 = 0, t_f = 500$ with $\Delta t = 0.1$

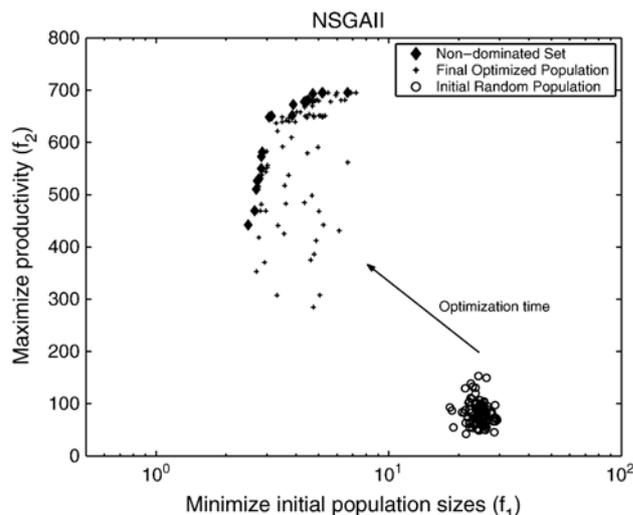


Fig. 2 – Initial and final solutions of an NSGA-II run. This is a representative example of one of the thirty simulations done with this algorithm.

We also fixed the MOEA parameters for all simulations, for both NSGA-II and SPEA2, equal to:

- (1) Population size = 50;
- (2) Maximum number of iterations $\tau_{\max} = 250$;
- (3) Probability of mutation = 1/50;
- (4) Probability of crossover = 1.

The probability of mutation corresponds to the inverse of the length of a chromosome (in this case, the length of the assembly sequences), a value commonly used (Eiben et al., 1999). Simulations were performed to evaluate the impact of the parameter η_c and it was found that $\eta_c = 0.5$ gave consistently good results in terms of the diversity of solutions found and the convergence performance.

When comparing the optimized sequences and/or their community properties with the random sequences and/or their community properties, we select all non-dominated solutions of the final population (optimized) of an MOEA’s run and we randomly chose 2 solutions in the initial population (random).

5. Results

5.1. MOEA’s performance for the objective functions

We first investigated if the MOEAs were able to generate assembly sequences to optimize the MOOP introduced in Section 3. The comparison of the initial solutions (i.e., random sequences) and the final solutions of NSGA-II (Fig. 2) and SPEA2 (Fig. 3) runs shows that it is possible to order sequences of invasion in a manner so that the assembly process is more efficient according to the MOOP. The form of the Pareto front (the limits reached by the non-dominated set) of the final solutions gives some information about the assembly process. A small variation in objective one, the initial population sizes, has a great influence on the population sizes of all species in

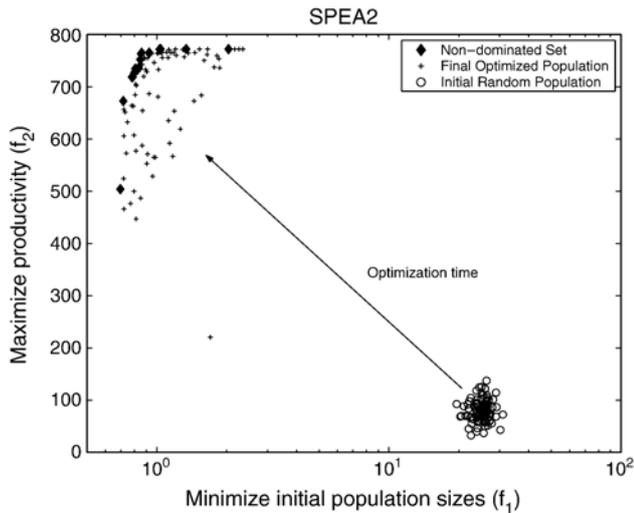


Fig. 3 – Initial and final solutions of a SPEA2 run. This is a representative example of one of the thirty simulations done with this algorithm.

the final community. It is also interesting to note that the Pareto front curves radically as the initial population sizes increase. This is probably due to the carrying capacity of the environment that limits the maximum productivity of the community. We return to the ecological implications of these results in the Discussion section. We also found that the SPEA2 algorithm presents a better convergence performance compared to the NSGA-II algorithm (Figs. 2 and 3). This observation is true for the majority of the 30 simulations done with both low and high connectance series.

5.2. Constraint handling and the variation in the expected connectance of food webs

We imposed a constraint on the expected connectance of the assembled food webs to be sure that the communities evolve to “realistic” food webs not only composed of basal species (zero connectance). As shown in Fig. 4, all solutions in the final non-dominated sets have a food web connectance that falls within the range imposed by the constraint in the MOOP. Control of the food web properties of an assembly process by the invasion sequences is certainly more feasible for real applications than modifying parameters such as the interaction matrix or the connectance of the regional species pool.

6. Discussion

6.1. The optimized assembly sequences

We present a new approach based on multi-objective evolutionary algorithms to optimize the assembly process of an ecological community model. The results show that it is possible to reorder the assembly sequences to maximize the productivity of the community and also control the expected food web connectance. But the important question is why historical events influence community properties. The differ-

ence between random solutions and optimized solutions (random sequences and optimized sequences) is shown in Fig. 5. These graphs show the initial population sizes and the species richness of basal and non-basal (intermediate+top species) species for random and optimized solutions. In random sequences, the number of basal and non-basal species and their initial population sizes is randomly distributed according to a linear relationship (Fig. 5a–c). This relationship is a result of the initial uniformly distributed values chosen for the population size of species. Note that the number of basal species in an assembly sequence equals one minus the number of non-basal species. But with the optimized sequences, there is a clear difference between basal and non-basal species (Fig. 5b–d). As the MOEAs have to optimize the productivity of the assembled community, solutions include a greater number of basal species than the number of non-basal species. This is not surprising since the persistence of non-basal species depends on the presence of these basal species. We show in Fig. 5b–d that the initial population sizes of basal species are also greater than the initial population sizes of non-basal species. We hypothesize that the initial population sizes of basal species are higher in order to resist predation pressure by intermediate or top predators.

However, the patterns in the assembly sequences are not only a question of a difference in the number of basal and non-basal species. As imposed by the MOOP, the MOEA have to reduce the initial population sizes of all species in the community. Thus, the algorithm sets the initial population sizes of non-basal species at low values. With the assembly model dynamics at hand, the only way that non-basal species increase their productivity is to establish a significant number of links with basal species. The productivity of basal species is limited by the carrying capacity of the environment. This is not the case for the non-basal species, which are limited by the basal species productivity. There is a slight difference between these limitations because non-basal species may be

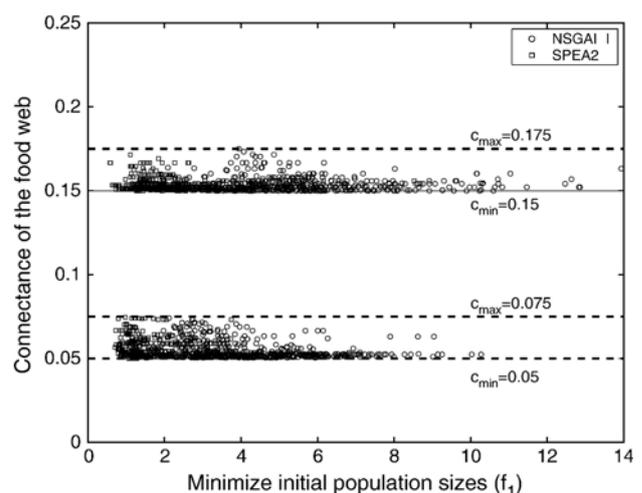


Fig. 4 – Final non-dominated solutions obtained from NSGA-II (\square) and SPEA2 (\circ). Graph shows the food web connectance versus initial population sizes (objective one) for final assembled communities of all non-dominated sets obtained from the 30 simulations done with “low connectance” and “high connectance” series.

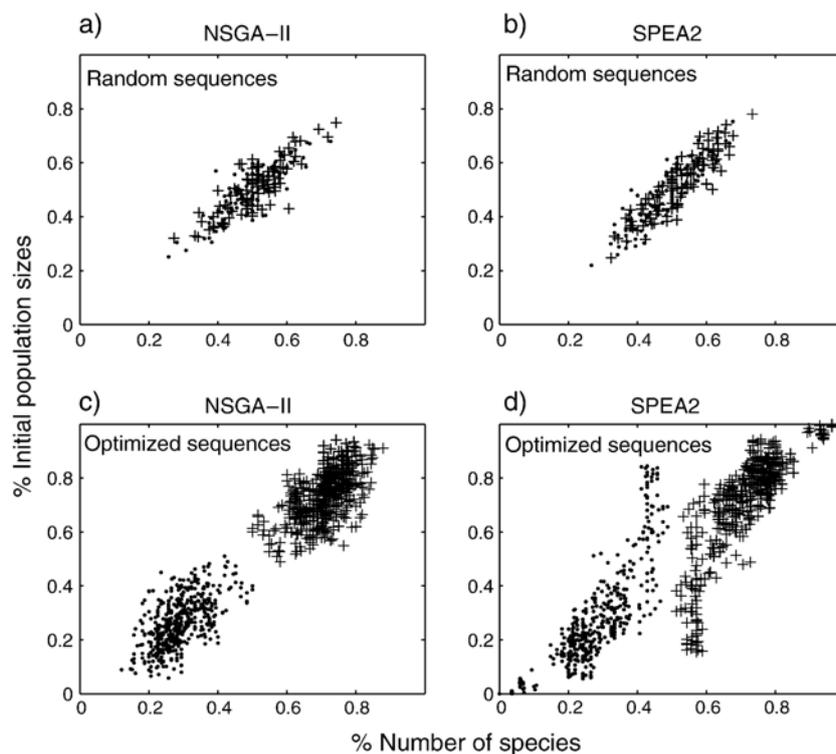


Fig. 5 – Proportion of the number of basal species (+) and non-basal species (–) versus the proportion of their initial population sizes in the random sequences (a)–(c) and the optimized sequences (b)–(d). The graphs show the results of all assembly sequences of the optimized non-dominated solutions obtained from the “high connectance” series.

linked to more than one basal species. This is exactly what we observe in the data. The optimized sequences are not only composed of a small proportion of non-basal species but also only a small proportion establish a significant number of links with basal species. As the expected connectance is fixed, the number of highly connected non-basal species has to be low compared to other species to keep the expected connectance in the specified range. Thus, as a result the degree distribution of communities generated via optimized sequences is different than those generated via random sequences (Fig. 6).

The degree distribution of a random network like that used in this paper for the regional species pool is fit by a Poisson distribution (Bollobas, 1985). The food webs generated via random sequences also produce a Poisson degree distribution (Fig. 6a–d). This is not a surprising result since the random sequences are in fact a random sampling of the regional species pool. For the communities generated by high connectance simulations, the degree distribution is also fit by a Poisson distribution but the tail of the distribution is expanded compared to random solutions (Fig. 6b–e). This provides a confirmation of the fact that the non-basal species have to establish more links if the productivity of the community has to be higher. This observation is much more evident if we look at the optimized sequences from the low connectance series (Fig. 6c–f). In this case, as the MOEAs have to increase the productivity by keeping a low connectance and the non-basal species have to establish many more links. Moreover, the diversity of the community increases to balance the effect on the connectance of adding links.

As we see in Fig. 6c–f, the degree distribution turns to an exponential distribution when the number of poorly connected species (the hump of the Poisson distribution) decreases and the number of highly connected species (the tail of the distribution) increases. Thus, it is possible to generate exponential degree distribution food webs from a random regional species pool that is fit with a Poissonian distribution. As pointed out by Arii and Parrott (2004) who generated non-random degree distributions from randomly structured networks exposed to frequent immigration–extinction events, the fact that a local community presents an exponential degree distribution does not necessarily imply that the regional pool is also exponentially distributed. This is an important result since in empirical data (Dunne et al., 2002a) or in statistical models like the niche model (Williams and Martinez, 2000; Camacho et al., 2002) no food web is fit by a Poissonian distribution. Moreover, with the optimization process that we defined we conclude that the exponential degree distribution of food webs corresponds to a more productive community that was generated via an efficient assembly process.

With these results, we want to emphasize the role of the assembly process on the properties of ecological communities. We hypothesize that the order of arrival of species in a community has a large impact on the network structure of the community and, moreover, if the sequence produces a more productive community than those produced via random sequences, the degree distribution of the community food web fit by a degree distribution similar to those observed in real food webs. Testing this hypothesis empirically might be very difficult

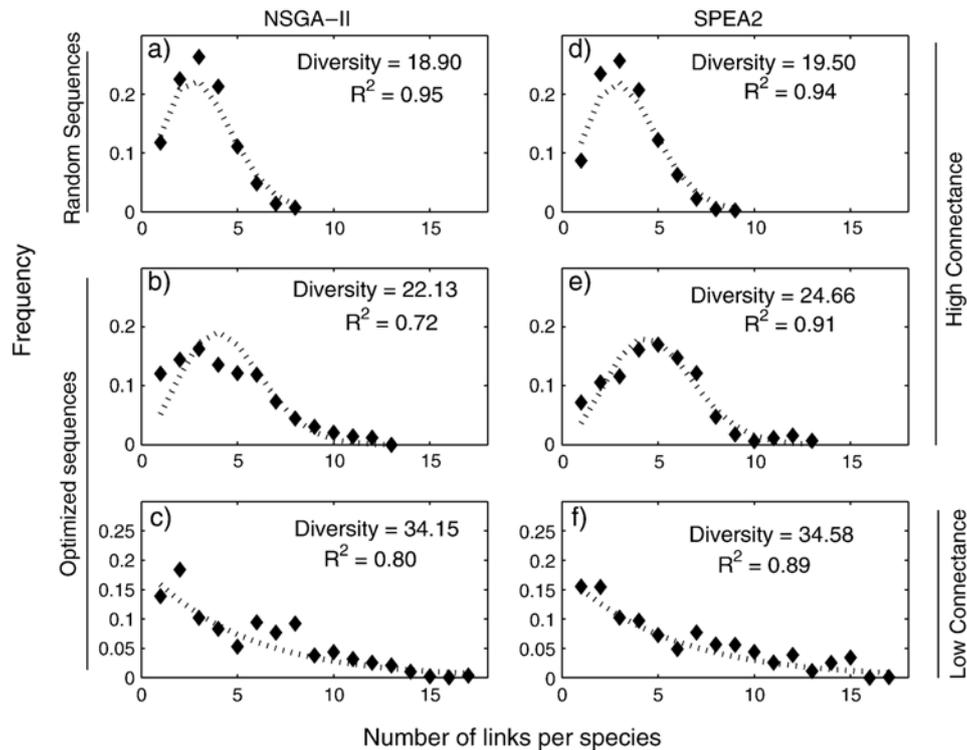


Fig. 6 – Linear plot of the degree distribution of community food webs. Dashed line and the R^2 represent the best fit of Poissonian or exponential distributions. Each observation is the mean of the frequency of the number of links from the 30 runs done with one series of simulations: NSGA-II with high connectance (a–b) and low connectance (c) and SPEA2 with high connectance (d–e) and low connectance (f). For each series the mean is also computed over all non-dominated solutions of the final population (b–c–e–f). The degree distributions of food webs of communities generated via random sequences are a mean over the solutions chosen from the initial population for each of the 30 runs of NSGA-II (a) and SPEA2 (d). Diversity is the mean diversity of the food web (i.e., species richness of the community) of all solutions that served to compute the degree distribution.

because of the lack of data concerning natural assembly sequences but microcosm experiments could be considered.

7. Conclusion

We present a new approach based on multi-objective evolutionary algorithms to analyze the community productivity and structure of an ecological assembly model based on Lotka–Volterra dynamics. We propose a constrained multi-objective optimization problem where the objective functions are (i) minimize the initial population sizes of all species introduced in the community and (ii) maximize the productivity of the community generated by the assembly process. We also add a constraint on the final community: the connectance of the food web of the community has to fall in a specific range (e.g. high connectance or low connectance).

The results show that the optimized sequences have a large influence on community properties, such as food web structure and the productivity of species. We also demonstrate that it is possible to control the connectance of food webs by generating optimized assembly sequences of species from the regional species pool. We emphasize the importance of historical events on community structure by showing that the optimized sequences lead to a rewiring of the food web that

enhances the productivity of the community. This is a consequence of the multi-objective optimization problem that we imposed on the assembly process since the algorithms have to minimize the population sizes of all species introduced in the community to maximize the productivity of the entire community. We conclude that being able to manipulate invasion sequences to obtain a desired community may have a number of practical applications in ecological engineering and ecological restoration projects. Evolutionary algorithms are a very interesting approach to explore the large search space formed by the parameters of this kind of complex system.

Acknowledgements

Simulations were run on equipment provided by the Réseau Québécois de Calcul de Haute Performance (RQCHP). The authors wish to thank team members from the CCS (Center of Computational Science), Université de Sherbrooke, Sherbrooke, Canada, for their technical support for the MPI C++ simulations executed on the mammoth-series cluster. Financial support for this project was provided by the Canadian Foundation for Innovation and les Fonds québécois de la recherche sur la nature et les technologies.

REFERENCES

- Arii, K., Parrott, L., 2004. Emergence of non-random structure in local food webs generated from randomly structured regional webs. *Journal of Theoretical Biology* 227, 327–333.
- Barabasi, A.L., Albert, R., 1999. Emergence of scaling in random networks. *Science* 286, 509–512.
- Bollobas, B., 1985. *Random Graphs*. Academic Press, London.
- Camacho, J., Guimerá, R., Amaral, L.A.N., 2002. Robust patterns in food web structure. *Physical Review Letters* 88, 228102.
- Case, T.J., 1990. Invasion resistance arises in strongly interacting species-rich model competition communities. *PNAS* 87, 9610–9614.
- Coello Coello, C.A., 1999. A comprehensive survey of evolutionary-based multiobjective optimization techniques. *Knowledge and Information Systems* 1 (3), 269–308.
- Coello Coello, C.A., Lamont, G.A., 2004. *Applications of Multi-objective Evolutionary Algorithms*. World Scientific, Singapore.
- Coello Coello, C.A., Van Veldhuizen, D.A., Lamont, G.A., 2002. *Evolutionary Algorithms for Solving Multi-objective Problems*. Kluwer Academic Publishers, New York.
- Côté, P., Parrott, L., 2006. Controlling food web structure by optimization of a community assembly model. *Ecological Informatics* 1 (2), 125–131.
- Deb, K., 2001. *Multi-objective Optimization Using Evolutionary Algorithms*. John Wiley and Sons, England.
- Deb, K., Kumar, A., 1995. Real-coded genetic algorithms with simulated binary crossover: studies on multi-modal and multi-objective problems. *Complex Systems* 9 (6), 431–454.
- Deb, K., Pratap, A., Agarwal, S., Meyarivan, T., 2002. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation* 6 (2), 182–197.
- Downing, K., 1997. Euzone: simulating the evolution of aquatic ecosystem. *Artificial Life* 3, 307–333.
- Drake, J.A., 1990. The mechanics of community assembly and succession. *Journal of Theoretical Biology* 147, 213–233.
- Dunne, J.A., Williams, R.J., Martinez, N.D., 2002a. Food-web structure and network theory: the role of connectance and size. *PNAS* 99 (20).
- Dunne, J.A., Williams, R.J., Martinez, N.D., 2002b. Network structure and biodiversity loss in food webs: robustness increases with connectance. *Ecology Letters* 5, 558–567.
- Eiben, A.E., Hinterding, R., Michalewicz, Z., 1999. Parameter control in evolutionary algorithms. *IEEE Trans on Evolutionary computation* 3 (2), 124–141.
- Fukami, T., 2004. Assembly history interacts with ecosystem size to influence species diversity. *Ecology* 85, 3234–3242.
- Fukami, T., 2005. Integrating internal and external dispersal in metacommunity assembly: preliminary theoretical analyses. *Ecological Research* 20, 623–631.
- Hewitt, C.L., Huxel, G.R., 2002. Invasion resistance and community resistance in single and multiple invasion models: do the models support the conclusion? *Biological Invasion* 4, 263–272.
- Holland, J.H., 1992. *Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control and artificial intelligence*, First MIT press edition. Massachusetts Institute of Technology.
- Hraber, P.T., Milne, B.T., 1997. Community assembly in a model ecosystem. *Ecological Modelling* 103, 267–285.
- Lockwood, J.L., Samuels, C.L., 2004. Assembly models and the practice of restoration. In: Temperton, V.M., Hobbs, R.J., Halle, T.N. (Eds.), *Assembly Rules and Restoration Ecology: Bridging the Gap Between Theory and Practice*. Island Press, 1718 Connecticut Avenue, NW, Suite 300, Washington, DC 20009. S.
- Lockwood, J.L., Powell, R., Nott, M.P., Pimm, S.L., 1997. Assembling ecological communities in time and space. *Oikos* 80, 549–553.
- Martinez, N.D., 1992. Constant connectance in food webs. *American Naturalist* 139, 1208–1218.
- Montoya, J.M., Solé, R.V., 2003. Topological properties of food webs: from real data to community assembly models. *Oikos* 102, 614–622.
- Morton, D.R., Law, R., 1997. Regional species pools and the assembly of local ecological communities. *Journal of Theoretical Biology* 187, 321–331.
- Morton, D.R., Law, R., Pimm, S.L., Drake, J.A., 1996. On models for assembling ecological communities. *Oikos* 75, 493–499.
- Pimm, S.L., 1980. Food web design and the effect of species deletion. *Oikos* 35, 139–149.
- Recknagel, F., 2002. *Ecological Informatics. Understanding Ecology by Biologically-Inspired Computation*. Springer, Berlin.
- Reynold, J.H., Ford, E.D., 1999. Multi-criteria assessment of ecological process models. *Ecology* 80 (2), 538–553.
- Ruiz-Moreno, D., Pascual, M., Riolo, R., 2006. Exploring network space with genetic algorithms: modularity, resilience, and reactivity. In: Pascual, M., Dunne, J.A. (Eds.), *Ecological Networks: Linking Structure to Dynamics in Food Webs*. Oxford University Press.
- Srinivas, N., Deb, K., 1994. Multiobjective optimization using nondominated sorting in genetic algorithms. *Evolutionary Computation* 2 (3), 221–248.
- Strogatz, S.H., 2001. Exploring complex networks. *Nature* 410, 268–276.
- Williams, R.J., Martinez, N.D., 2000. Simple rules yield complex food webs. *Nature* 404 (9), 180–183.
- Zimmermann, C.R., Fukami, T., Drake, J.A., 2004. An experimentally-derived map of community assembly space. In: Bar-Yam, Y., Minai, A.A. (Eds.), *Unifying Themes in Complex Systems II: Proceedings of the Second International Conference on Complex Systems*. Perseus Books.
- Zitzler, E., Thiele, L., 1999. Multiobjective evolutionary algorithms: a comparative case study and the strength Pareto approach. *IEEE Transactions on Evolutionary Computation* 3 (4), 257–271.
- Zitzler, E., Deb, K., Thiele, L., 2000. Comparison of multiobjective evolutionary algorithms: empirical results. *Evolutionary Computation* 8 (2), 173–195.
- Zitzler, E., Laumanns, M., Thiele, L., 2002a. SPEA2: Improving the strength Pareto evolutionary algorithm. In: Ursem, R.K. (Ed.), *Proceedings of the Evolutionary Methods for Design, Optimization and Control with Applications to Industrial Problems (EUROGEN2001)*. Morgan Kaufmann, pp. 95–100.
- Zitzler, E., Laumanns, M., Thiele, L., Fonseca, C.M., da Fonseca, V.G., 2002b. Why quality assessment of multiobjective optimizers is difficult. In: Langdon, M.B., et al. (Ed.), *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO2002)*. Morgan Kaufmann, pp. 666–674.