

Modelling Artificial Ecosystem Selection: A Preliminary Investigation

Alexandra Penn

Centre for Computational Neuroscience and Robotics,
University of Sussex, U.K.
alexpc@cogs.susx.ac.uk

Abstract. The ability of whole ecosystems to respond to selection has recently been demonstrated in artificial selection experiments, [1, 2]. As well as having wide-ranging practical applications, this result significantly broadens the application of theoretical concepts of the mechanisms of heritability and variation in biological systems. Simulation models have the potential to be useful tools for the investigation of these issues. Whilst related simulation work exists [3, 4], ecosystem-level selection itself has yet to be modelled. This paper presents such a model, in which ecosystems are modelled as generalised Lotka-Volterra systems and are subject to a generational selection process. A positive response to selection for diversity is demonstrated, with the only sources of variation being sampling errors arising when ‘offspring’ ecosystems are produced.

1 Introduction

Artificial selection at the level of the ecosystem is a new field of research, recently demonstrated in laboratory experiments by Swenson et al [1, 2]. The possibility of creating ‘designer ecosystems’ has great potential practical usefulness. Swenson et al [2] have already demonstrated a response to selection in microbial communities evolved for improved biodegradation of the common pollutant, 3-chloroaniline. One could also envisage selection of microbial communities for biological sewage treatment, or selection of soil communities, such as mycorrhizal fungi, for increased above-ground biomass and productivity of crop plants.

Treating entire ecosystems as units of selection in their own right also raises interesting theoretical questions. Ecosystem-level selection is substantially different from individual-level selection. Firstly, selection can act on complex, ‘higher-level’ traits produced by the interactions between individuals and between species. Secondly, the mechanisms of variation and heritability which might allow selection at the level of the ecosystem are likely to be very different. As well as genetic variation (both within and between species), other potential sources of heritable variation include changes in the relative frequencies of different species and abiotic and stochastic factors. Swenson et al [1, 2] suggest that variation is primarily introduced through the sampling process by which ecosystems are reproduced, and that the effects of this variation can be significantly magnified by the dynamics of the ecosystem’s subsequent development. If this is the case,

then ecosystem dynamics must also play an important role in the inheritance of ecosystem-level traits.

Given the number of potential interdependent processes operating on a number of different levels and time-scales, understanding how ecosystems respond to selection presents a significant challenge. In this context, simulation may provide a useful tool for exploring intuitions and investigating hypotheses. This paper presents a preliminary investigation into modelling artificial ecosystem selection. Ecosystems, modelled using Lotka-Volterra competition equations, are successfully evolved to increase their species diversity. Two potential sources of heritable variation are modelled and compared, and the potential role of ecosystem dynamics is discussed.

2 Selection Above the Level of the Organism

The question of whether natural selection operates at a level above that of the individual organism remains a controversial one [5]. However, the requirements for selection to occur are general properties, not restricted to individual organisms [6]. Artificial selection experiments can sidestep the issues of whether selection happens in a natural setting and instead ask whether a response to selection can exist on a certain level when a particular population structure and selection pressure are imposed. A strong response to higher-level selection has been shown experimentally by a number of researchers, for example, single-species group selection with flour beetles [7], and cages of battery farm chickens [8, 9], and multispecies group selection with two species of flour beetle [10]. These experiments have typically shown a strong response to group selection, and poor, or even negative, response to individual selection, and are interesting illustrations of the mechanics of higher-level selection. All have demonstrated that group-level selection can act on *interactions* between individuals or species, not directly accessible to individual-level selection [11].

The possibility of artificial selection at the ecosystem level was first suggested by Sober and Wilson, [5], and then tested experimentally by Swenson, Wilson et al [1, 2]. In three separate experiments on soil and pond water microbial ecosystems, they showed statistically significant positive responses to selection for ‘ecosystem-level’ traits: above ground plant biomass, pH, and biodegradation of 3-chloroaniline. All experiments followed a similar procedure:

1. An initial ‘population’ of ecosystems was created by taking small, equally-sized, samples from the same ‘source’ ecosystem.
2. Samples used to inoculate a fixed amount of some sterile medium.
3. Ecosystems develop for a fixed amount of time, an ‘ecosystem generation’.
4. Ecosystems evaluated for the chosen ‘ecosystem-level’ phenotypic trait.
5. Top scoring N_{parent} ‘parent’ ecosystems used to found the next generation.
6. Ecosystems reproduced ‘sexually’, mixing all “parents” then taking new samples, or ‘asexually’, taking $N_{offspring}$ samples from each ‘parent’.
7. Steps then repeated from (2) for a set number of generations.

Each experiment showed wide divergence of different lines away from the value of the phenotypic trait of the original ecosystems. This was seen in both control and selected lines, however only the selected lines showed a systematic change in the direction of selection. Although the response to selection seemed somewhat erratic and variable, the difference between the first and last generations was found to be statistically significant in almost all cases; for example, a 4-fold increase in plant biomass, and 25-fold differences in H^+ concentration [1].

The idea of ecosystems as units of selection is an interesting one. The notion that selection can act at the level of ecosystems is not necessarily problematic. There are three general requirements for selection to take place amongst a population of units of any given type [6]:

1. Phenotypic variation amongst units.
2. Heritability of phenotypic differences.
3. Fitness consequences of phenotypic variation.

These requirements could, in principle, be satisfied at any level of biological organisation. In the case of artificial ecosystem selection experiments, both the creation of a population of units and the fitness consequences of phenotypic variation are imposed by the experimenters. However, for ecosystem-level selection to then occur still requires the first two conditions set out above to be met. That is, ecosystems sampled from the same original source must vary phenotypically, and that variation must be heritable. That these conditions should be fulfilled is far from intuitively obvious.

As ecosystems do not possess genomes it might be hard to see how any phenotypic variation between them could be inherited. Indeed, as all ecosystems are initially sampled from the same source ecosystem and develop in near-identical physical conditions, the source of their variation is not obvious. Swenson et al [1, 2], conjecture that the main source of phenotypic variation amongst ecosystems is due to their “sensitive dependence on initial conditions”. That is, that sampling results in small initial differences in species’ genetic composition or population sizes, and these are magnified by ecosystem dynamics to give rise to large differences in macroscopic phenotypic traits. However, this sensitivity is a potential problem for the heritability of phenotypic traits. Heritability requires that offspring resemble their parents, whereas variation appears to arise to because the process of sampling can lead to significant differences between parent and offspring. For selection to be successful, a fine balance must be achieved between the opposing forces of variation and heritability, both of which are consequences of the underlying ecosystem dynamics. Ecosystem selection must search therefore, not only for ecosystems with the required phenotypic trait, but also for systems which will come quickly to stable local equilibria, so that their properties can reliably be transmitted to the next generation.

The differences between this approach, and individual level selection should be noted. As well as species genetic variation, an entirely different source of (potentially) heritable variation is present: the species proportions in a particular ecosystem at the end of a generation. These can differ between two ecosystems

due to differences in initial species proportions, even in the absence of between-ecosystem genetic variation. Moreover, these differences may increase when the two ecosystems are themselves reproduced.

3 The Model

The model follows the same general procedure described above. Individual ecosystem dynamics were modelled using generalised Lotka-Volterra equations (described below). At the start of a run, a single source ecosystem is randomly generated and left to develop for 200 iterations. An initial population of 20 ecosystems is then created by repeatedly sampling the source ecosystem, using a fixed size ‘pipette’. Although each ecosystem is generated from the same source, sampling error can introduce between-ecosystem variation in both species genetic composition, and initial species population sizes. Each ecosystem in the population is then allowed to develop for 50 iterations, and then evaluated on an ecosystem-level phenotypic trait (specified below), and the top 5 ecosystems are selected to become parents for the next generation. Offspring ecosystems are produced asexually by taking 4 fixed-size samples from each parent, thereby generating a new population of 20 individuals. As before, sampling may introduce variation between offspring of the same parent. The process of sampling, evaluation and selection is repeated until the required number of generations is reached.

Within-ecosystem dynamics are modelled using the generalised Lotka-Volterra competition equations [4, 12]. For an ecosystem containing S species, the population size N_i of the i^{th} species at time $t + 1$ is given by:

$$N_{i,t+1} = N_{i,t} \left[1 + \frac{R}{K_i} \left(K_i - \sum_{j=1}^S N_{j,t} \alpha_{ij} \right) \right] \quad (1)$$

where K_i is the species’ carrying capacity, R is the growth rate (common to all species), and α_{ij} is an interaction coefficient representing the per capita effect of species j on species i . For this model, $S = 10$ and $R = 2$ are constants. When a source ecosystem is created, each K_i is set at uniform random in the range 100:1000, and each α_{ij} is set at uniform random in the range 0:2, unless $i = j$ and then $\alpha_{ij} = 1$. Note that although all direct interactions are competitive ($\alpha_{ij} > 0$), indirect effects may give rise to mutualisms or commensualisms.

Ecosystem reproduction involves taking a fixed-size sample from a selected parent ecosystem. In real ecosystem selection experiments, the process of sampling can introduce variation between offspring both in species genetic composition, and initial species population sizes. In the model, genetic variation due to sampling was modelled very simply. Genetic variation was assumed to affect interaction coefficients (α_{ij}) only. For each interaction coefficient, the offspring value was generated by taking a random deviate from a Gaussian distribution with a mean equal to the parent value and a standard deviation of 0.02 (i.e., 1% of the range). The initial population size for each species is calculated on

the assumption that a sample contains individuals chosen at random from the parent ecosystem, thus the expected frequency of a species in a sample is equal to its frequency within the sampled ecosystem. Since species population sizes are continuous variables, sampling was modelled using the standard Gaussian approximation to a binomial distribution. Thus, N_i , the size of the species in the new sample was generated at random from a Gaussian distribution with mean, Bp_i , and standard deviation, $\sqrt{Bp_i(1-p_i)}$, where p_i is the frequency of the species in the parent ecosystem, and $B = 100$ is the mean size of the sample.

Species diversity was chosen as the ecosystem-level selection criterion, (useful in this context as calculated from the population distribution only). The degree of diversity exhibited by an ecosystem is a function of species interactions (both direct and indirect), and as such, a property of the ecosystem as whole rather than being attributable to individual species. The Shannon-Weaver Diversity Index—a commonly used ecological measure—was employed to quantify diversity [13]. It is defined as follows:

$$H = - \sum_i^S p_i \ln p_i \quad (2)$$

where S is the total number of species, and p_i is the proportion of species i in the population.

4 Results

The first question addressed was simply whether or not the model ecosystems would respond to selection. The ecosystem selection procedure was run 50 times. Each run lasted 200 generations and was started from a different randomly generated source ecosystem. Figure 1 shows the change in fitness score of best ecosystem for each generation averaged over all 50 runs. The difference between the first and last generation mean best fitness is statistically significant (Mann-Whitney U-Test $p < 0.001$). To ensure that this improvement was a result of selection, another 50 runs were carried out using the same set of source ecosystems, but parents were selected at random. This time there was no overall directional trend (see fig. 1), and no significant difference between the first and last generation means (Mann-Whitney U-Test $p = 0.95$).

An obvious question to ask is whether both sources of ecosystem variation are contributing to the successful response to selection just described. Two further sets of runs were carried out, with each using only one source of variation. The first set, employing only genetic variation, also showed a significant improvement (Mann-Whitney U-Test $p < 0.001$), achieving only a marginally lower final-generation mean than in the original experiment (0.62 compared to 0.64), which was not a significant difference (Mann-Whitney U-Test $p = 0.45$). In contrast, the set of runs employing only population size variation showed little sign of responding to selection; there was only a marginal improvement with no significant difference between a first generation mean of 0.322 and a final generation mean of 0.344 (Mann Whitney U-Test $p = 0.74$).

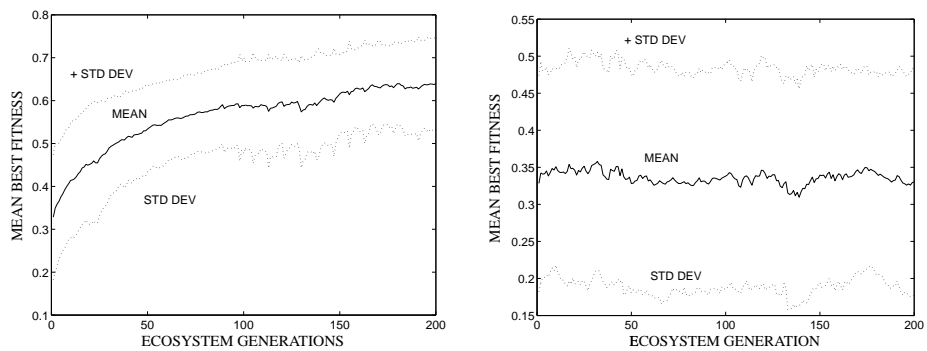


Fig. 1. Change in best ecosystem fitness over 200 generations, averaged over 50 different runs, with selection(left), random selection(right)

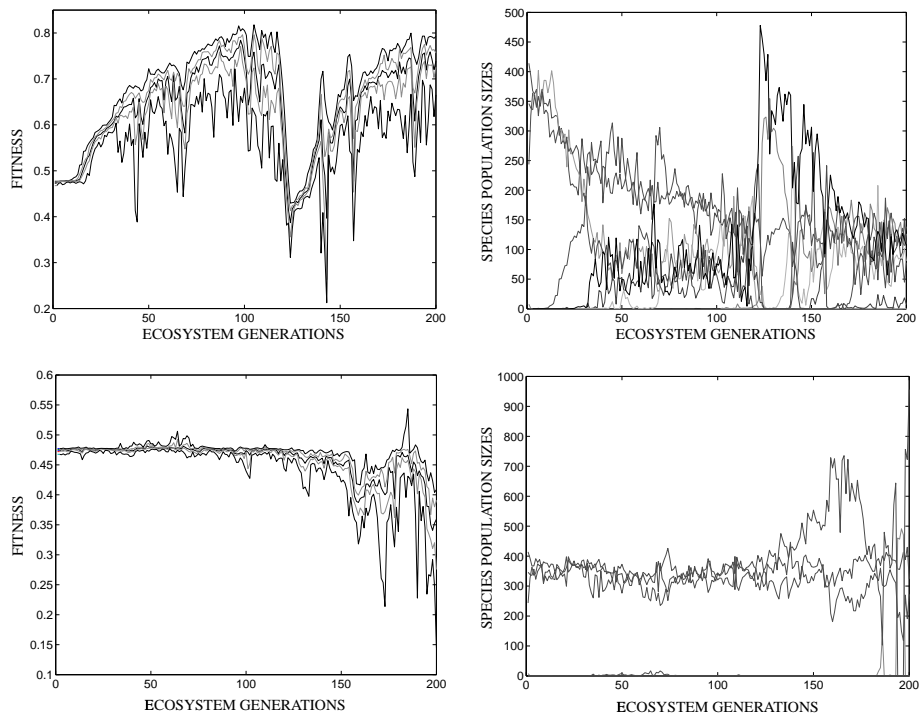


Fig. 2. Change in ecosystem population fitnesses over 200 generations for source ecosystem 4, with directed selection (top), and random selection (bottom) , showing best, worst, median, and upper and lower quartiles under random selection(left), corresponding changes in species frequencies at end of each generation for best ecosystem in population(right)

5 Discussion and Conclusions

Despite its simplicity, the model presented in this paper can illustrate some of the concepts of ecosystem-level variation and heritability introduced earlier. From the dynamical systems perspective set out in [1], phenotypic variability can arise in two ways: (i) sampling of ecosystems that are following different trajectories to the same attractor (instantaneous variation) or, (ii) ecosystems being knocked into different basins of attraction by sampling error. Figure 2 shows the dynamics of two runs starting from the same source ecosystem, one under directed and one under random selection. Phenotypic variability of both types arose under both selection procedures. This is particularly evident in the graph showing final species frequencies over 200 generations for the best ecosystem in the directed selection run (top right, fig.2). The species composition changes only slightly gen. 50-100 (accumulation of instantaneous variation), then undergoes sudden dramatic change gen. 100-150 with some species going extinct, others re-emerging from a low-level position. Finally it settles into a new ‘attractor’ with a totally different species distribution. These dynamics are accompanied by large changes in fitness (top left, fig.2). Fitness initially climbs, then collapses as the population dynamics suddenly change. It begins to climb again as they settle into a new equilibrium. This illustrates very well the previously discussed balance between variability and heritability. The population accumulates fitter variants produced by sampling error, but the fitness collapses as a small variation changes the underlying dynamics such that the ecosystem ‘attractor’ switches. This first ecosystem state is not robust enough to changes in initial conditions to be reliably heritable. Instantaneous variation is not sufficient for a reliable response to ecosystem selection. Any sustainable response to selection would require a network of ecosystem attractors (i.e. relatively stable ecosystem-level phenotypes) which could be reached via variations in species frequencies and genetic composition at the sampling phase. More heritable states would have wider basins of attraction, and come to equilibrium more quickly. In principle new ecosystem phenotypes can be reached with variation only in species proportions. However, in this model runs with population sampling error only, although displaying instantaneous variability in fitness, very rarely moved to new overall species compositions. Genetic sampling error alone did allow new species compositions to arise and fitness to increase significantly. This result may have been a consequence of the particular ecosystem dynamics used. Multispecies L-V competition systems can possess many stable attractors [13], however this does not necessarily mean that they are *sensitively* dependent on initial conditions. Ecosystems were started from small samples, and sampling error was small due to the assumption of perfect mixing. This might not have given a large enough range of possible initial species proportions for population sampling error to have played a significant role. Conversely, genetic sampling error may have been too large. This remains to be explored in future work.

Evidently the model presented here is an extreme simplification, used as a first step to explore ideas based on a dynamical systems perspective on ecosystems [1]. Real ecosystems exhibit certain patterns of connectivity, are constrained

by energetic restrictions, and show higher-order interactions with the presence of some species modifying interactions between others. More realistic ecosystem dynamics may need to be incorporated, as well as stochasticity, and an abiotic environment. The abstract fitness measure chosen for these experiments satisfied the criterion of being an ecosystem-level trait, produced by the interactions of all species present. However, other fitness measures need to be investigated, including those which could be influenced by either individual or ecosystem selection. The potential for comparing an ecosystem's response to particular fitness criteria with selection at these two levels is particularly interesting. One of the next steps will be to introduce evolution within species during ecosystem generations, allowing both individual and ecosystem-level dynamics to act. Given that experiments so far have concentrated on fast-breeding, microbial ecosystems, the interplay between these dynamics is likely to be important. This model is only a first step towards the effective use of simulation tools in understanding and exploring the mechanisms of ecosystem-level selection. Ultimately it is hoped that simulations such as these can be used to help answer important practical questions on the mechanisms and conditions underlying ecosystem selection.

Acknowledgements

Thanks to Matt Quinn and Inman Harvey for constructive discussion.

References

1. Swenson, W., Wilson, D.S, Elias, R.: Artificial Ecosystem Selection, *Proc. Natl. Acad. Sci. USA*, 97, (2000), 9110–9114
2. Swenson, W., Arendt, J., Wilson, D.S.: Artificial Selection of Microbial Ecosystems for 3-chloroaniline Biodegradation, *Environ. Microbiology*, 2(5), (2000), 564–571
3. Ikegami, T, and Hashimoto, K.: Dynamical Systems Approach to Higher-level Heritability, *J. Biol. Phys.*, 28(4),(2002), 799-804
4. Wilson, D.S.: Complex Interactions in Metacommunities, with Implications for Biodiversity and Higher Levels of Selection. *Ecology*, 73(6),(1997) 1984–2000
5. Sober, E., and Wilson, D.S.: *Unto Others: The Evolution and Psychology of Unselfish Behaviour*, Harvard University Press, Cambridge, Massachusetts (1998)
6. Lewontin, R.C.: The Units of Selection, *Annu.Rev.Ecol.Syst.*, 1, (1970), 1–18
7. Wade, M.J.: Group Selection Among Laboratory Populations of *Tribolium*, *Proc. Natl. Acad. Sci. USA*, 73, 4604–4607
8. Craig, J.V., and Muir, W.M.: Group Selection for Adaptation to Multiple-Hen Cages: Beak-related Mortality, Feathering, and Body Weight Responses., *Poultry Science* 75 ,(1995), 294–302
9. Muir, W.M.: Group Selection for Adaptation to Multiple-Hen Cages: Selection Program and Direct Responses., *Poultry Science* 75 ,(1995), 447–458
10. Goodnight, C.J.: Experimental Studies of Community Evolution, I: The Response to Selection at the Community Level., *Evolution* 44,(1990), 1614–1624
11. Goodnight, C.J.: Heritability at the Ecosystem Level, *Proc. Natl. Acad. Sci. USA*, 97(17), (2000), 9365–9366
12. MacArthur, R.H.: *Geographical Ecology*, Harper and Row, N.Y., USA (1972)
13. Morin, P.J.: *Community Ecology*, Blackwell Science (1999)