

# Multiple states of environmental regulation in well-mixed model biospheres

Arwen E. Nicholson<sup>a</sup>, David M. Wilkinson<sup>b,c</sup>, Hywel T. P. Williams<sup>a</sup>, Timothy M. Lenton<sup>a</sup>

<sup>a</sup>*Earth System Science, University of Exeter, UK*

<sup>b</sup>*Natural Science and Psychology, Liverpool John Moores University, UK*

<sup>c</sup>*School of Life Sciences, University of Lincoln, UK*

---

## Abstract

The Gaia hypothesis postulates that life influences Earth's feedback mechanisms to form a self-regulating system. This provokes the question: how can global self-regulation evolve? Most models demonstrating environmental regulation involving life have relied on alignment between local selection and global regulation. In these models environment-improving individuals or communities spread to outcompete environment degrading individuals / communities, leading to global regulation, but this depends on local differences in environmental conditions. In contrast, well-mixed components of the Earth system, such as the atmosphere, lack local environmental differentiation. These previous models do not explain how global regulation can emerge in a system with no well defined local environment, or where the local environment is overwhelmed by global effects. We present a model of self-regulation by 'microbes' in an environment with no spatial structure. These microbes affect an abiotic 'temperature' as a byproduct of metabolism. We demonstrate that global self-regulation can arise in the absence of spatial structure in a diverse ecosystem without localised environmental effects. We find that systems can exhibit nutrient limitation and two temperature limitation regimes where the temperature is maintained at a near constant value. During temperature regulation, the total temperature change caused by the microbes is kept near constant by the total population expanding or contracting to absorb the impacts of new mutants on the average affect on the temperature per microbe. Dramatic shifts between low temperature regulation and high temperature regulation can occur when a mutant arises that causes the sign of the temperature effect to change. This result implies that self-regulating feedback loops can arise without the need for spatial structure, weakening criticisms of the Gaia hypothesis that state that with just one Earth, global regulation has no mechanism for developing because natural selection requires selection between multiple entities.

**Keywords:** Gaia, self-regulation, complexity, feedback, agent-based model

---

*Email address:* [arwen.e.nicholson@gmail.com](mailto:arwen.e.nicholson@gmail.com) (Arwen E. Nicholson)

*Preprint submitted to Elsevier*

*October 27, 2016*

## 1. Introduction

The Gaia hypothesis postulates that life on Earth interacts with abiotic processes to form a complex self regulating system that maintains habitable conditions on the planet [17] [13] [16]. This is evolutionary ecology at the very largest spatial and temporal scales [28]. Critics of the theory argue that any organism acting to improve the habitability of the planet would have to contend with “cheaters” who do not contribute to regulation, or that a system would be just as likely to drive itself extinct as it would to drive itself towards stability [6] [5]. This leads to the question: how can self-regulation evolve in a way consistent with evolutionary theory? With only one Earth, and thus a lack of data to analyse, this question has been addressed using theoretical models. Hence we describe other models to put this study into context.

The Daisyworld model [26] was the first model to present global regulation emerging by local selection of individual level traits that contribute to global regulation. In the original Daisyworld there are two species of daisy - black daisies that have a low albedo and white daisies that have a high albedo. The growth of daisies is a function of temperature and all daisies have the same ideal temperature for maximum growth rate. Incoming radiation from a ‘sun’ that evolves in the manner of a typical main sequence star, heats Daisyworld. Daisyworld initially starts off too cool for any daisy growth, but as the sun evolves the incoming solar radiation becomes high enough for the surface temperature to allow daisy growth. Black daisies are the first to appear. By absorbing more solar radiation they warm their local environment encouraging their own growth and warming the global environment. When the temperature increases enough, cooling high albedo white daisies appear. The balance between the number of white cooling daisies and the number of warming black daisies main-

tains a constant habitable temperature in Daisyworld. As the solar luminosity increases the white daisies take over and keep the planet cool, until the incoming radiation is too high and all daisies die.

The Guild model [7], also demonstrates global regulation arising from local selection. In the Guild Model, individuals consume and excrete chemicals that appear in the system via an inflow. Which chemicals they consume and excrete are determined by an individual’s genome. New ‘species’ (i.e. new genomes) arise via mutation during reproduction events. All individuals have maximum consumption levels when the ratio of chemicals is at a particular value. Individuals affect their local chemical ratios via their consumption and excretion and these effects diffuse to the global environment. As in Daisyworld, individuals that improve their local environment will be selected for, and this local selection contributes to global regulation. The Guild model finds that communities of individuals can exist together to create and regulate the preferred chemical ratios.

The Flask model [30] [29] [31] removed a limiting assumption of Daisyworld and the Guild model that traits selected for at the individual level always improve the global environment. Instead the organisms in the system affect the abiotic environment as a byproduct of their metabolism, making these effects selectively neutral at the individual level. Instead of each individual having a distinct local environment, groups of individuals share a common local environment. During reproduction there is a small constant probability of mutation per locus  $P_{mut}$  so that over time new species arise via mutation. A spatial version of the model connected multiple local environments by inflows and outflows [31] [29]. Stabilising environmental regulation still emerged and this model argues for spatial structure creating conditions where limited higher-level selection can take place. In a connected envi-

ronment, locations where local communities improve their environment achieve larger populations and thus can colonise and outcompete communities that degrade their environment leading to the spread of environment-improving communities and thus global regulation.

For local selection to take place on environment-related traits, local environments must be different. However, certain environments cannot be compartmentalised in a manner that seems conducive to local selection. The obvious example is the atmosphere (with its well mixed gases) but some aquatic environments are also potentially well mixed too. In this case it is not obvious where the local environments allowing for successful communities to develop would be, leading to motivation for a homogenous model of self-regulation.

Later versions of Daisyworld [18] and ‘Daisy-stat’ [8] removed the local environment and found regulation of the abiotic parameters. In these models ‘rein-control’ [4] [9] is responsible for the environmental regulation. In one version of these models [18] two main subgroups dominate the system - one group that acts to increase the abiotic parameter while preferring this parameter to be low, and another group that acts to lower the abiotic parameter, while preferring this parameter to be high. With these two groups pulling the system in opposite directions, environmental regulation is possible for significant periods of time. The Daisy-stat model [8] features the same ‘rein-control’ in this case regulating multiple abiotic parameters with a diverse array of species instead of the system being dominated by two main groups. In Flaskworld [31], the effect of allowing different microbe species to prefer different abiotic parameter values was explored and it was found that the system showed periods of stability where the abiotic parameter stayed near constant. These stable periods were interrupted with rapid transitions where the abiotic parameter would often then

stabilise at a different value to before. The system was stabilised by the ‘rein-control’ mechanism present in the Daisy-stat model.

The Daisy-stat model provides global regulation with a diverse population in the absence of spatial heterogeneity. However this model lacks mutation. Species begin reproducing when the environmental parameters allow them to, and all species are present at all times even if at vanishingly low levels. This means that as the environment changes, the system does not need to evolve new species to control or adapt to these changes, the species are already present and ready to start reproducing as soon as conditions allow. Therefore in Daisy-stat, the system cannot go extinct. This does not reflect real world biology where the existing population must evolve to cope with a changing environment and total extinction is a possibility. For this reason we follow the Flask model implementation of microbes with selectively neutral abiotic effects that reproduce and mutate allowing new species to appear in the system.

The atmosphere taken as a single entity has a flux of energy coming in as light from the sun, heat from the mantle and various chemicals spewed forth by volcanoes similar to the nutrient and abiotic parameter inflow in the Flask model. For something like the Earth’s atmosphere a single well mixed environment would be a more accurate representation than local environments interacting with a global environment.  $CO_2$  fluxes, for example, at various points on the Earth do not vary wildly (ignoring the very small scale i.e. surrounding a currently active volcano) making a single flask Flask model a good approximation to the system.

The original Flask model [30] was also a single flask environment, however the implementation of the model was quite different and the focus of the paper was on nutrient recycling and not abiotic regulation. In the original Flask model [30] instead of microbes all having the

same preferred value for a single abiotic parameter, there were two abiotic parameters and microbes had an encoded preference for a particular ratio of these two parameters. This ratio preference was not constant for all microbes and therefore not all microbes experienced the environment identically. When there is a universal preference for an abiotic parameter, this sets a constant target for regulation, where the preferences for abiotic parameter values differ, there is no such constant target. The target will change as the genetics within the population change. The microbes were able to evolve towards preferring the state of the current abiotic environment and exploit all the nutrients in the system. In this paper we instead focus on what happens in a system where the microbes cannot evolve towards preferring the current environment, and instead of a preferred ratio between two abiotic parameters that differs between different microbe species, we have a single abiotic parameter with a constant preferred value for this parameter,  $\beta$  that is the same for all microbes.

For our single flask model we closely follow the implementation detailed in [29] limiting the system to a single flask. We present a model of self-regulation of a purely global environment arising via evolution. This single Flask model allows for the possibility of rebel mutants disrupting the system, due to the lack of distinct environments and removes the issue of “cheater” species, due to the selectively neutral abiotic effects of the microbes. It is also possible for the system to drive itself to extinction - all scenarios being criticisms of the Gaia theory [6] [5]. The combination of assumptions presented here differs to what has been tried in previous models. The model has a shared preference for a single abiotic parameter, but lacks spatial structure as in the previous Flask models [30] [29] [31]. Mutation occurs in this model with a constant probability per reproduction event, and the system can suffer from total irreversible extinction, differing from the Daisystat [8] and models by

McDonald-Gibson [18]. Finally, the model lacks local environments, differing from the original Daisyworld [26], and the Guild model [7].

In Section 2 of this paper we give a brief outline of the model (an in depth description can be found in Appendix A). Section 3 details the behaviour of the model for various important regimes. We present results both of typical individual simulations and for trends in systems with the same parameter settings. In Section 4 we discuss the results and provide parallels to behaviour found in the real world.

## 2. Model

In the Flask model [31] [29], flasks contain an abiotic environment with parameters (that can be thought of as temperature, pH, salinity), and nutrients which are the substrates for metabolism, with a constant inflow and outflow of these abiotic parameters. The flasks are seeded with ‘microbes’ which consume the nutrients available and affect the abiotic parameters as a side effect of their metabolism. In turn the value of these abiotic parameters affects the microbes’ ability to metabolise. In this paper we limit the system to a single flask.

‘Microbes’ are characterised by a binary genome. This genome determines what nutrients a microbe will consume and what it excretes (with the limitation that nothing may eat what it excretes). Microbes with the same genome are considered to be the same ‘species’. As microbes consume nutrients and convert them to biomass, they are able to reproduce once their biomass reaches a reproduction threshold  $B_R$ . During reproduction there is a small constant probability of mutation per locus  $P_{mut}$  so that over time new species arise via mutation. Microbes die if their biomass drops to the starvation threshold  $B_D$  and there is also a probability of death by other causes

$P_D$ . The maintenance cost  $\lambda$  for each microbe is 1 biomass unit per timestep.

We refer to microbes with different genomes as being different ‘species’, however our model is essentially microbial, e.g. akin to the Earth during the Archean. In microbes, extensive horizontal gene transfer can make speciation a complex matter, and in the case of this model, it is the phenotype of the microbes that is important, rather than their genotype, which just determines which nutrients they eat and excrete. We have a rather small genome size in this model and so minor changes usually associated with specific variation are not possible. Our model is best understood in terms of the ‘genomes’ of the flask ‘microbes’ as trait vectors, whereby the ‘mutation’ operator is just a simple way of introducing variation. The model mutations can represent quite large changes in metabolism that would in reality most likely involve a longer sequence of smaller mutations. In this context it is relevant that results from an ecology model called the Tangled Nature model, used for investigating stability in ecosystems, found that allowing for gradual changes in the phenotype of the agents in the model, rather than large scale changes each mutation, simply lead to the same dynamics slowed down [1].

As a byproduct of converting nutrients to biomass the microbes affect the abiotic parameters. Per unit of biomass created, the microbes change an abiotic parameter by a set amount (determined by their genome) in the range  $[-1, 1]$ . The environmental abiotic parameters in turn affect the rate at which microbes can consume nutrients. Each microbe,  $j$ , has a preferred level,  $\beta_j^i$ , for each abiotic parameter,  $i$ .  $\beta_{env}^i$  is the value of the  $i^{th}$  abiotic parameter.  $\tau$  controls how sensitive the microbes are to the abiotic parameters. If  $\tau = 0$ , the microbes are not influenced by the abiotic parameters. For  $\tau > 0$ , the abiotic environment affects metabolism. The higher  $\tau$  becomes the more sensitive the microbes become to their

environment and thus for a high  $\tau$  if the difference between each  $\beta_j^i$  and  $\beta_{env}^i$  is too large the microbes will be unable to consume nutrients. The quantity of nutrients a microbe is able to consume per timestep,  $C_j^{max}$ , depends on how closely each  $\beta_{env}^i$  matches  $\beta_j^i$  in the following way:

$$C_j^{max} = \psi_j C^{max} \quad (1)$$

$$\psi_j = e^{-(\tau p_j)^2} \quad (2)$$

$$p_j = \sqrt{\sum_{i=1}^A (\beta_{env}^i - \beta_j^i)^2} \quad (3)$$

where  $C^{max}$  is a constant determining the maximum rate of consumption for any microbe in ideal conditions,  $\psi_j$  is a microbe specific measure of the microbe’s satisfaction with the current abiotic environment.  $A$  is the total number of abiotic parameters. As the  $\beta_{env}^i$  values move away from the ideal  $\beta_j^i$  values,  $C_j^{max}$  will become smaller meaning the microbes are able to ingest fewer nutrients per timestep and for unfavourable enough conditions, they will be unable to consume anything.

The environment of the single flask is characterised by a constant inflow of nutrients and abiotic parameters,  $I_N$  and  $I_A$ , and a constant outflow  $O_N$  and  $O_A$ . In the absence of microbial activity the environment reaches a constant steady state with constant levels of nutrients and constant values for the abiotic parameters. After a preparation period  $t_{prep}$  to allow the system to come to equilibrium, the flask is seeded with 100 randomly generated individuals. For a single timestep the following actions are performed:

1. Influx / outflux of nutrients and abiotic parameters
2. Microbe selected randomly for a death event

3. Microbe selected randomly for a nutrient consumption event
4. Microbe selected randomly for a biomass creation event
5. Microbe selected randomly for a reproduction event
6. Repeat steps 2 - 5  $n$  times, where  $n$  is the total population of the system at the start of the current timestep.

In this way, on average, each microbe in the system will be selected for each event per timestep. We keep steps 1 - 5 separate in the code to introduce randomness into the model. We also neglect to consider any situations where particular genomes or consumption behaviours affect reproduction rates or death rates as can happen in real life.

We ran simulations of this single Flask model for various values for  $\tau$  to demonstrate that a single well mixed flask can exhibit environmental abiotic regulation with two stable regimes for a certain range of  $\tau$ . For each simulation we had  $N = 4$  nutrients and  $A = 1$  abiotic parameter, denoted as  $\beta$  (and referred to as ‘temperature’ throughout this paper) for our systems. We set  $\beta_j = \beta = 150$  for each microbe,  $j$ , and the abiotic ‘temperature’ without microbe activity to  $\beta_{env} = 100$ . Throughout this paper we will refer to the nutrients in the system and the abiotic parameter separately. Although the nutrients present in a system part of the abiotic environment, we reserve this label for the abiotic ‘temperature’.

For more details on the model presented in this paper see Appendix A.

### 3. Results

For a range of values for  $\tau$ , a key parameter that controls the strength of the feedback between the environmental state and life, we ran 100 simulations, all identical apart from their

initial random seed, and recorded how many of these 100 simulations survived (survival defined as having microbes alive at the end of the simulation) to  $10^5$  time steps. We then looked in closer detail at these surviving simulations.

Table 1 shows the survival rate of simulations for different values of  $\tau$ . We see that the survival rate of the system quickly starts to drop off above  $\tau = 0.015$ .

Table 1 also shows the mean lifespan for various  $\tau$  along with the standard deviation. For low  $\tau$  all the simulations survived to the end, but as  $\tau$  increases the survival rate decreases and so does the average lifespan.

We find that there are three ways in which the microbe population of the system can be limited - nutrient limited, high temperature limited and low temperature limited. Which of these regimes dominates the system depends on the value of  $\tau$ .

#### 3.1. Nutrient limitation for $\tau = 0$

In a nutrient limited regime the microbes consume all the available nutrients. Once the nutrients are depleted the population can no longer grow and the microbe population will stabilise such that the flow of incoming nutrients is enough to support the population, i.e. the system reaches the carrying capacity determined by the nutrient input. This characterises a key aspect of many real-world systems where the tendency for biological populations to proliferate if conditions are good leads to a reduction of resources which then becomes a stable state - a process called ‘biotic plunder’ by [22], and achieves a zero-net growth isocline (ZNGI) [21]. In a  $\tau = 0$  system, i.e. the microbes are indifferent to the abiotic temperature, nutrient limitation is the only way the system becomes limited.

Figure 1 shows a single simulation for  $\tau = 0$ . In Figure 1b the temperature has no overall

Table 1: Survival % and average lifespan as a fraction of the total simulation length ( $10^5$  timesteps) for a range of  $\tau$ .

$\tau$	Survival %	Average lifespan / $10^5$	Standard deviation
0.00	100	1	0
0.005	100	1	0
0.01	100	1	0
0.015	100	1	0
0.02	50	0.75	0.32
0.025	18	0.33	0.38
0.03	9	0.20	0.33

trend but is a random walk as the microbes are not affected by its value. The changes in temperature come from the byproducts of the microbes' metabolism. For every unit of biomass produced each microbe will add a set value to the temperature parameter as determined by their genome. Genetic mutation is occurring within the population and so new species with different affects on the temperature regularly appear. For  $\tau = 0$  the fitness, which we define as the rate of biomass production per microbe, per timestep, is at the maximum value for any value of T, meaning that the temperature is selectively neutral and so the system's temperature is effectively an unselected random walk, determined by the current populations genomes, and changing with genetic variation. When the system is nutrient limited, the system is still producing mutants at the same rate per microbe. This means that the total abiotic effect of the microbe ecosystem will be constantly changing as new mutants with differing abiotic impacts appear preventing the temperature from stabilising.

From Figures 1b and 1d we see that the population quickly reaches a maximum value and stays there and that the nutrient stocks are quickly reduced to near zero and also stay at that level. The total population possible in a system is determined by the nutrient flow.

Figure 1d shows a cartoon plot of temperature against fitness with respect to the abiotic temperature, all else being equal.  $f_{min}$  represents the minimum fitness, here defined as the number of offspring produced per individual per timestep, required for the microbes to maintain a constant population, i.e. the rate of reproduction matches the rate of death. As the value of the temperature does not affect fitness for  $\tau = 0$ , the fitness of the microbes is a constant independent on T. Note that the temperature in this model is arbitrary and does not correspond to real world temperatures, and so it is the behaviour of the temperature that is important, not the value.

### 3.2. Nutrient and temperature limitation for $\tau > 0$

For  $\tau > 0$  the microbes fitness is no longer constant for all temperature, see Equation (3). In  $\tau > 0$  systems, the system can still become nutrient limited if the average effect per microbe on the abiotic temperature,  $\epsilon_{avg}$ , is small enough to allow the microbes to exhaust the nutrient stocks. We calculate  $\epsilon_{avg}$  by summing up all the abiotic effects for all the microbes, and then dividing by the total population of the system to get the average metabolic effect per microbe. As  $\tau$  increases and the microbes become more and more sensitive to their environ-

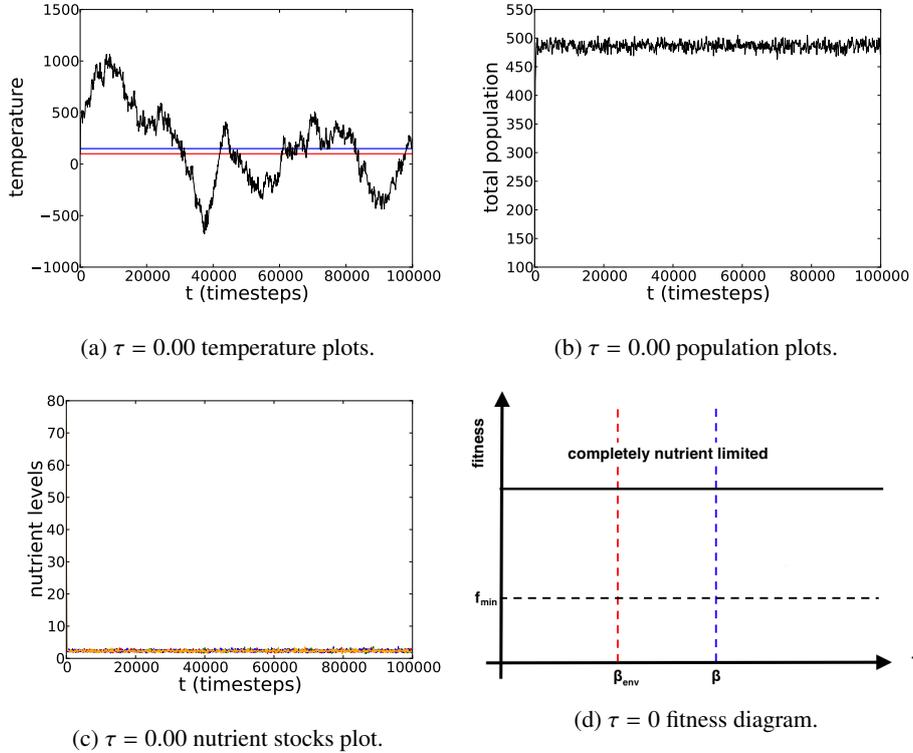


Figure 1: Example simulation of a single  $\tau = 0$  system. In a) and d) the blue line represents  $\beta$  and the red line represents  $\beta_{env}$ . The temperature a) has no general trend, the total population b) quickly rises and stabilises at the carrying capacity. The nutrient stocks c) quickly deplete and remain at near zero levels. The fitness doesn't depend on temperature so the fitness d) is a constant.

ment nutrient limitation becomes less likely, and when it does happen it quickly transitions to a temperature limited regime instead.

Figure 2 shows a  $\tau = 0.005$  simulation that demonstrates nutrient limitation and temperature limitation. There are periods in Figures 2b and 2c when the system is not nutrient limited as the total population falls below the maximum and the nutrient stocks are not completely exploited i.e. at  $t = 0 - 500$  where the system is high temperature limited, and  $t = 70,000 - 95,000$  where the system is low temperature limited. At low  $\tau$  the system is mainly nutrient limited, but as the temperature goes towards the extremes at which

the microbe's can survive, the system becomes temperature limited, and the fitness curve falls away from the maximum, shown in Figure 2d. When temperature limited, the system is in a negative feed back loop, with the stable point at the temperature that allows the minimum fitness required for a stable population. The green circles in Figure 2 represent the system at various temperatures. There are two points on the fitness curve where the temperature is in a semi-stable state - a state that persists for significant time spans but are prone to sudden transitions to another state. These semi-stable states occur where the fitness curve intersects the  $f_{min}$  line. Where these two lines cross we find the the upper and lower temperature lim-

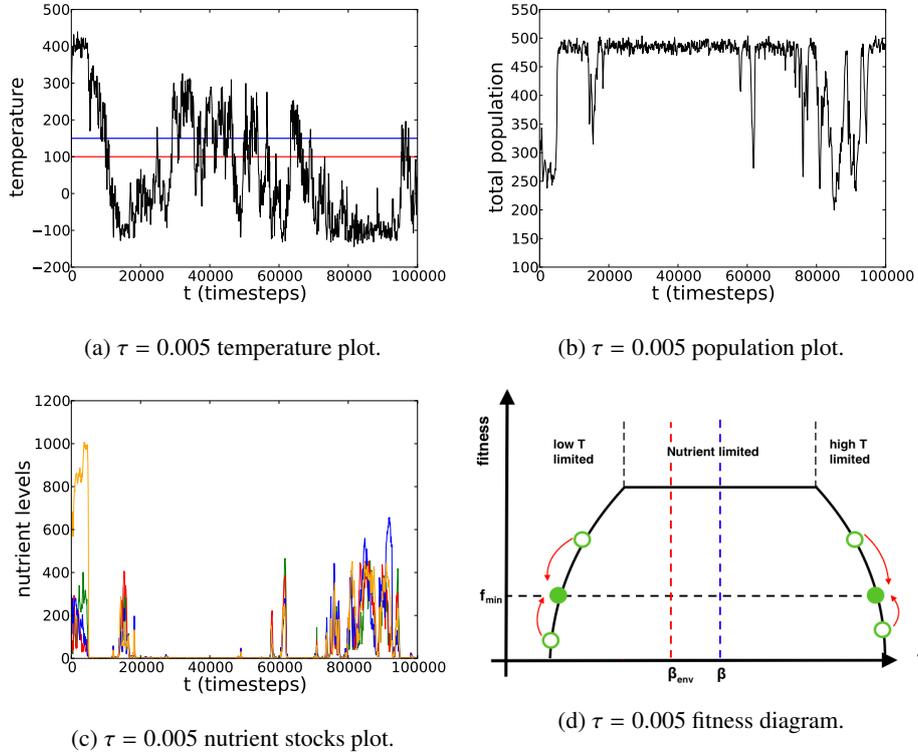


Figure 2: Temperature, population and nutrient stock plots for an individual  $\tau = 0.005$  system. In a) and d) the blue line represents  $\beta$  and the red line represents  $\beta_{env}$ . We see regions of temperature limitation a) where the population b) drops to lower than the carrying capacity, and the nutrient stock c) are higher. The fitness d) now depends on the temperature for extreme temperatures.

its. At these points green circles are filled to represent that the system temperature is semi-stable for this  $T$ . At other points on the fitness curve the system temperature is not stable, represented by non-filled circles.

The two temperature regimes, high and low, work as follows:

- **High Temperature Limited Regime**

The high temperature regime supports a population of microbes whose collective effect on the environment is to heat it. As the microbes consume nutrients and create biomass the temperature increases until the microbes become unable to con-

sume nutrients due to the temperature being too extreme. At this point with no microbes creating biomass, and the constant flow creating a cooler environment, the temperature begins to fall until it reaches the point where microbes are able to consume nutrients again and the cycle repeats. In this way the system can be thought of as bouncing off an upper temperature limit set by how extreme a temperature the microbes can still metabolise in. In the high temperature limited regime  $\epsilon_{avg}$ , is positive - i.e. on average a microbe has a heating effect.

In a system limited by high temperature the total population and  $\epsilon_{avg}$  are highly

negatively correlated. If a population has a certain  $\epsilon_{avg}$  and a new microbe mutates into existence that causes  $\epsilon_{avg}$  to increase, then depending on the size of the temperature increase the microbes' metabolism may slow to levels too low to maintain a constant population, in which case random deaths will reduce the population, or, if the temperature increase is extreme enough, metabolism can halt entirely. This will lead to individuals starving and the population will drop. With a lower rate of metabolism or no metabolism happening at all the temperature of the system will start to drop due to the inflow and outflow of temperature to the system. At a certain point the temperature will drop enough that the microbes will be able to start consuming nutrients again and the system will continue at roughly the limiting high temperature but supporting a lower population.

Time scales are important in these events as if the temperature change is extreme enough that metabolism halts entirely, the whole population will very quickly die and so the system can only tolerate short lived excursions from habitable temperatures. For a less extreme temperature change that still allows metabolism to take place (albeit at a rate below the maintenance level) then the system can survive longer as it will take longer for the microbes to starve to death or, failing starvation, the random death events will reduce the population. In general, any affects that act to push the temperature beyond habitable limits must be counteracted quickly to avoid total extinction.

Conversely if the effect of the new mutant was to lower  $\epsilon_{avg}$ , the temperature would drop and thus the microbes would consume more nutrients, the population would increase, raising the temperature with it until it stabilised at around the limiting high temperature, this time support-

ing a higher population than before.

- **Low Temperature Limited Regime**

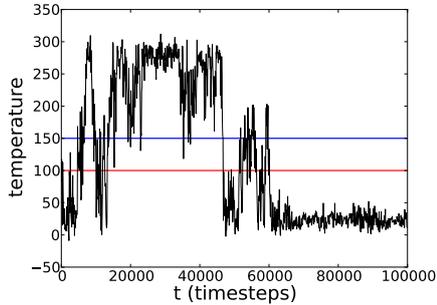
The low temperature regime is almost the mirror image of the high temperature regime. In the low temperature limited regime  $\epsilon_{avg}$  is negative - i.e. on average a microbe has a cooling effect. In this case the total population of microbes and  $\epsilon_{avg}$  are positively correlated. If a microbe mutates into existence that causes  $\epsilon_{avg}$  to decrease, the population will increase, and if the mutant acts to increase  $\epsilon_{avg}$  the population will decrease.

As  $\tau$  increases, temperature limitation becomes more important. For a higher  $\tau$ , the microbes are more sensitive to their abiotic temperature and nutrient limitation is possible for a smaller range of T. When nutrient limited, the system has a higher total population than when temperature limited, so mutants appear at a faster rate. This combined with the smaller nutrient limited T range means that the system can quickly random walk out of nutrient limitation and become temperature limited. So as  $\tau$  increases, nutrient limitation dominates systems less, and temperature limitation takes over. Figure 3 shows two systems, a  $\tau = 0.01$  system and a  $\tau = 0.015$  system, and shows a higher amount of temperature limitation for the higher  $\tau$ .

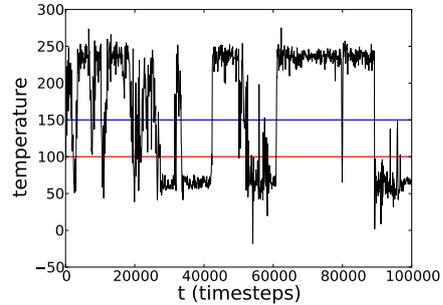
Comparing Figures 3c and 3d we see that for  $\tau = 0.015$ , the periods of nutrient limitation (seen where the nutrient stocks are close to zero) are less frequent and of shorter duration than they are for  $\tau = 0.01$ .

### 3.3. *Temperature limitation dominates for $\tau \geq 0.02$*

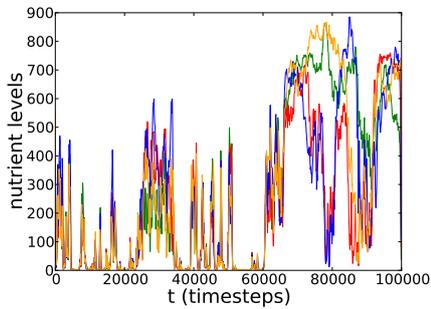
As  $\tau$  increases, the span of temperature where the microbes are nutrient limited shrinks further, so that the system becomes dominated by temperature limitation. The asymmetry in the



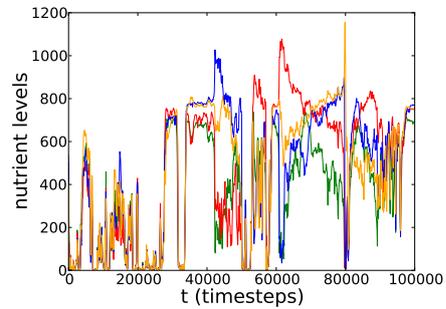
(a)  $\tau = 0.01$  temperature plot



(b)  $\tau = 0.015$  temperature plot



(c)  $\tau = 0.01$  nutrient stocks plot



(d)  $\tau = 0.015$  nutrient stocks plot

Figure 3: Plots for two individual systems, one at  $\tau = 0.01$  and the other at  $\tau = 0.015$ . In a) and b) the blue line represents  $\beta$  and the red line represents  $\beta_{env}$ . As  $\tau$  increases, temperature limitation becomes more likely than nutrient limitation.

model set up also begins to have a noticeable affect on the temperature limited regimes. As  $\beta_{env}$  is cooler than the microbes' preferred temperature  $\beta$  the total cooling effect needed by the microbes to become low-temperature limited is less than the total heating effect required to become high-temperature limited. For low  $\tau$  where the temperature range in which the microbes can function in is large, this does not have much effect, but as  $\tau$  decreases and the high and low temperature limits contract towards  $\beta$ , this starts to have an effect.

As the cooling needed to become low-temperature limited is less, the total population that this low temperature limited regime can support becomes lower and lower as the limiting low temperature increases. If a mu-

tant then arises that causes  $\epsilon_{avg}$  to cool more strongly, the population has to shrink to counteract this, and with an already small population this is more likely to drive the system to extinction than for the high temperature limiting regime, which can support a higher population. This means that as  $\tau$  increases the low temperature limiting regime becomes less stable and the non-extinct systems are far more likely to be found in the high-temperature limiting state. This behaviour is due purely to the fact that  $\beta_{env}$  is lower than  $\beta$ . Were  $\beta_{env}$  higher than  $\beta$ , we would see the same behaviour but flipped - the high temperature regime becoming less stable than the low temperature regime with increasing  $\tau$ .

Figure 4 shows a  $\tau = 0.02$  system that demon-

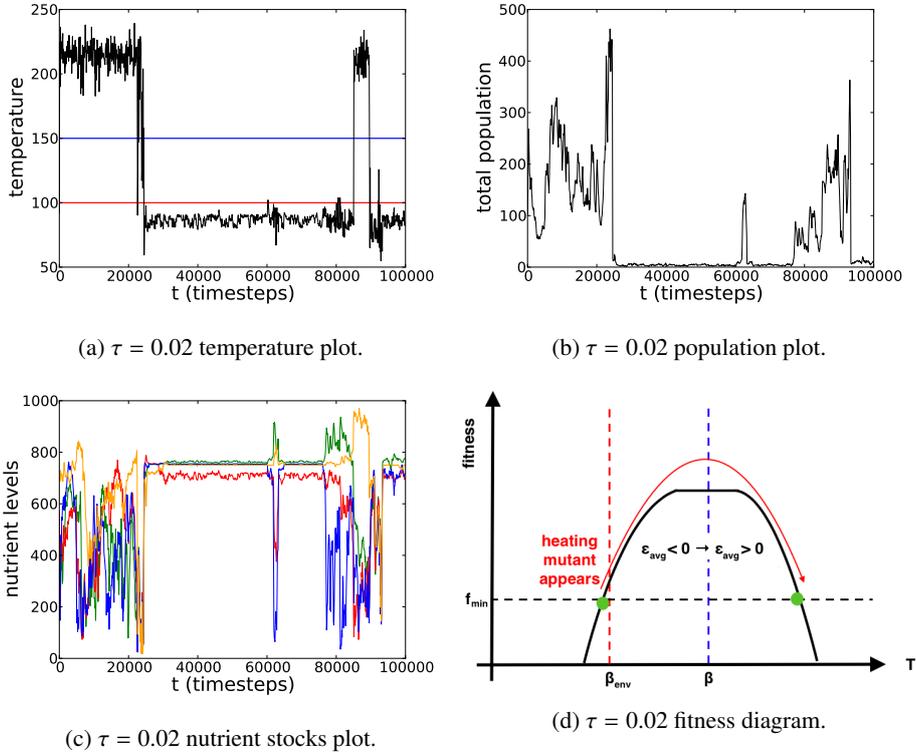


Figure 4: An individual  $\tau = 0.02$  system. In a) and d) the blue line represents  $\beta$  and the red line represents  $\beta_{env}$ . Note in a) and b) when the system is limited by low temperature, the total population is very low. d) shows the nutrient limiting range shrinking. The red arrow indicates the system moving straight from low temperature limitation to high temperature limitation.

strates this asymmetry. We see in the population graph, Figure 4b that the total population is much lower when the system is low temperature limited, than it is when the system is high temperature limited. We also see some very clear transitions between low temperature limitation and high temperature limitation without even a short a period of nutrient limitation in between. The higher  $\tau$  value means there is a much smaller range of temperature where the system can be nutrient limited, so a mutant microbe acting to change  $\epsilon_{avg} < 0$  to  $\epsilon_{avg} > 0$  doesn't have to have as strong an effect for the system to move through nutrient limitation temperature range and become temperature limited on the other side, as Fig-

ure 4d demonstrates. From Figure 4c we see that there are always nutrients available. None of the stocks are ever fully depleted, the microbe ecosystem is never able to fully exploit the nutrient resources. If  $\beta$  and  $\beta_{env}$  were close enough, the microbes would be able to consume all available nutrients and become nutrient limited, but for  $\tau = 0.025$  the temperature range that allows for nutrient limitation is very narrow so the system quickly gets knocked out due to mutants perturbing the system and pushing it to a temperature limited regime.

Increasing to  $\tau = 0.03$  and the asymmetry of the system now means that the high temperature becomes the only stable limitation regime.

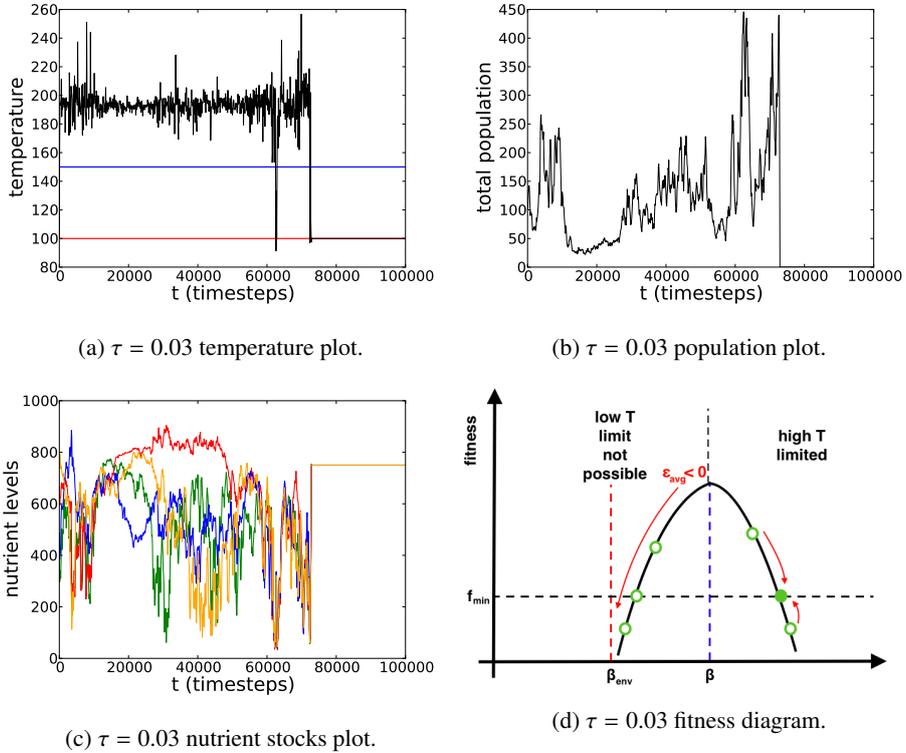


Figure 5: Plots for an individual  $\tau = 0.03$  system that goes extinct. In a) and d) the blue line represents  $\beta$  and the red line represents  $\beta_{env}$ . d) shows that the low temperature regime is now not possible.

The temperature span for the nutrient limitation regime has reduced so much that the system will very quickly random walk away from this regime, and the lower temperature limiting regime  $T$  value is now above  $\beta_{env}$  (where the fitness curve intersects the  $f_{min}$  line in Figure 5d). This means there is no negative feedback mechanism for maintaining the system for a temperature below  $\beta$ . If the temperature falls below  $\beta$  and the microbes act to cool, then the temperature will decrease until the microbes begin to die off. As  $\beta_{env} < \beta$  however, now there is nothing to pull the system back towards  $\beta$ , the system will be pulled towards  $\beta_{env}$ , which is now too cool for any microbes to survive and the system will go extinct. The only way for the system to avoid extinction for  $T < \beta$  is if  $\epsilon_{avg} > 0$ . In this case as the microbes heat the

system, the temperature will increase towards  $\beta$ , increasing the fitness, increasing the total population and therefore accelerating the heating in a positive feedback loop, until  $T > \beta$ , and the system becomes high temperature limited in a negative feedback loop.

We can see in Figure 5a one time at  $t = 65,000$  where the system is able to recover from  $T < \beta$ . In this case a mutant acting to change  $\epsilon_{avg} < 0$  to  $\epsilon_{avg} > 0$  appears and prevents the system from going extinct, but the system is not so lucky a second time, and goes extinct the next time  $T < \beta$ . The likelihood of producing a mutant depends on the reproduction rate and the population size and hence when the microbes' metabolism is constrained by temperature the reproduction rate is low. The high

nutrient stocks present in the system however create a potential for rapid growth if a ‘good’ mutant appears, i.e. one that moves the temperature closer to  $\beta$ . Referring back to Table 1 we see that for  $\tau = 0.03$  the survival rate for  $10^5$  timesteps was 9%, so systems that are able to recover are the minority. Again, Figure 5c shows that the nutrient stocks are never depleted, the microbes are unable to stay within the very narrow nutrient limited temperature range.

### 3.4. Waiting time for ‘Quakes’

We can look at waiting time statistics for ‘quakes’ - a period where the system’s temperature is not stable - to get an idea of how long the quasi-stable periods last within systems with various  $\tau$ . For the purposes of the following analysis we define a quake in the following way:

$$quake = \begin{cases} 1, & \text{if } |T_{avg}^{prev} - T_{avg}^{cur}| > 5 \\ 0, & \text{otherwise} \end{cases} \quad (4)$$

where a value of 1 means a quake took place, and a value of 0 means the temperature is remaining stable - no quake.  $T_{avg}^{prev}$  is the temperature averaged over the previous 100 timesteps, and  $T_{avg}^{cur}$  is the temperature averaged over the next 100 timesteps. We compare  $T_{avg}^{prev}$  to  $T_{avg}^{cur} \pm 5$  as some small temperature fluctuation does occur during stable periods so to compare the two with no buffer would lead to an artificially high number of quakes being recorded. Using this method we can record the times at which quakes occurred in a system.  $\pm 5$  is chosen as it is large enough to take into account fluctuations that happen within a stable period, but small enough that quakes are noticed. Changing  $\pm 5$  to some other limit does not qualitatively change the results much, but quantitatively the recorded number of quakes for all systems will increase if the limit is reduced, and will

decrease a little if increased until the limit gets so large that quakes become unidentifiable.

The value of each waiting time bin in the histograms is divided by the number of simulations included.

Figure 6 shows histograms of the average frequency of waiting times for quakes for 4 values of  $\tau = 0.00, 0.01, 0.02, 0.03$ . Figure 6a shows a histogram for all simulations while Figure 6b shows a histogram including only data from non extinct simulations. This allows us to look for characteristics in the ‘successful’ systems for various  $\tau$ . We want to understand what behaviours a system needs to have in order to avoid extinction. We are looking at the the Gaia hypothesis from the point of view of a planet that has successfully had uninterrupted life for billions of years. We could be incredibly lucky, and our planet might, if ‘reset’ and run a 100 times, usually be doomed to total extinction, or perhaps every 100 times life would emerge and successfully regulate the planet to maintain habitable conditions. As we don’t know which scenario we are in, it is useful to look for signatures in ‘successful’ systems for both scenarios (those likely to survive and those unlikely to survive) to see how we might be able to tell them apart. For each non extinct simulation for a particular  $\tau$  the waiting times for quakes are measured and binned, and then these bins are divided by the number of non-extinct simulations to give an estimate of how many times per simulation for a particular  $\tau$  we can expect to wait a certain period of time for a quake. Note in Figure 6 that both the  $x$  and the  $y$  axis are plotted to log scale. Also note that because each simulation ends at  $t = 10^5$  if a simulation lasts in a stable period for the whole simulation, the number of quakes would be 0 and thus in the plots in Figure 6 a lower frequency of waiting times means a more stable system as it indicates the stable periods of the system have remained mostly uninterrupted.

Figure 6 shows a histogram of waiting times

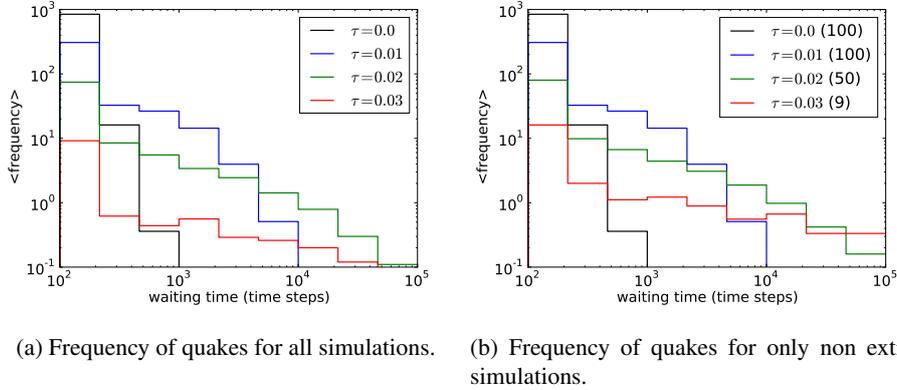


Figure 6: Histograms showing the average frequency of waiting times for quakes for different  $\tau$ . The number of non-extinct simulations in b) for each  $\tau$  is given in brackets in the legend. Note that both the  $x$  and  $y$  axis are logarithmic.

for  $\tau = 0, 0.01, 0.02, 0.03$ . We see that the frequency of short waiting times for  $\tau = 0$  is high. This frequency drops until just after  $10^3$  timesteps the frequency is 0. This agrees with previous plots showing no abiotic temperature regulation for  $\tau = 0$ . When  $\tau = 0$  the microbes are not regulating the environmental temperature, so the temperature is free to wander. ‘Quakes’ in this scenario regarding the temperature no longer make sense as the temperature is never really stable but measuring for ‘quakes’ we would expect them to be frequent and for there to be short waiting times between them, which is what Figure 6 confirms.

For  $\tau = 0.01$  the microbes are regulating the environmental temperature and the system can exhibit both temperature limitation and nutrient limitation. Here we see that longer waiting times occur and the shorter waiting times are less frequent than for the  $\tau = 0$  case. This demonstrates that on average, for simulations with  $\tau = 0.01$  we can expect longer periods of temperature stability with fewer quakes, however frequency of quakes drops to 0 for waiting times longer than roughly  $10^4$  timesteps.

For  $\tau = 0.02$  we see that short waiting times for quakes happen at an even smaller frequency than for  $\tau = 0.01$ . Across almost all wait-

ing times, the frequency of quakes is lower for  $\tau = 0.02$  than for  $\tau = 0.01$ . This tells the story of a system with longer periods of stability and fewer quakes. From Table 1 recall that the survival rate for  $\tau = 0.02$  is low at 28%. It is not therefore that at  $\tau = 0.02$  the microbes are far better at keeping the environmental temperature from fluctuating than they are at  $\tau = 0.01$ , but that for a  $\tau = 0.01$  system the more lenient restrictions on the microbes means that the system is better able to recover from a quake, but in a  $\tau = 0.02$  case, quakes come with a larger probability of total extinction. So for  $\tau = 0.02$  simulations with a lower level of quakes will have a greater probability of surviving. The relationship between frequency and waiting time for  $\tau = 0.02$  in Figure 6 shows a roughly linear relationship suggesting that there is a power law relationship. If the frequency of quakes goes as  $f_q \propto t^{-\alpha}$  where  $t$  represents time, and  $\alpha$  is some constant, we can take the log of both sides to find a linear relationship in log-log space which is indicative of a power law.

For  $\tau = 0.03$  we see that the frequency of all waiting times for quakes is very low, and with only 9% of the simulations avoiding extinction, this shows a more extreme version of the

$\tau = 0.02$  case.

Comparing Figures 6a and 6b we can see that including only non-extinct simulation has the effect of ‘flattening’ the frequency curve, making it less steep and increasing the frequency of quakes for longer waiting times. When we include the data from simulations that went extinct, it lowers the quake frequency as an extinct system cannot quake, and it makes the high  $\tau$  systems appear less prone to quakes than they are in reality.

### 3.5. Population, Temperature and $\epsilon_{avg}$

To get a clearer understanding of which regimes (nutrient limiting, high temperature limiting or low temperature limiting) are dominating the systems we plot the average microbe effect on the temperature,  $\epsilon_{avg}$ , vs the temperature and vs the total population of the system for various  $\tau$ .  $\epsilon_{avg}$ , and the total population of the system are correlated when the system is in a temperature limiting regime - negatively correlated in the high temperature limiting regime and positively correlated in the low temperature limiting regime. Thus we expect to see (in a high temperature limiting scenario) that as  $\epsilon_{avg}$  increases, the total population decreases.

Figure 7 shows  $\epsilon_{avg}$  vs total population for the data from all non extinct simulations over a range of  $\tau$ . We see for  $\tau = 0$ , the total population remains constant for any value of  $\epsilon_{avg}$  which agrees with previous results. We can clearly see the nutrient limited regime for very low  $\tau$  start initially wide and become increasingly narrower as  $\tau$  increases. For  $\tau = 0.005$  we clearly see both the low and high temperature regimes, the left curve showing the total population increases for an increase in (negative)  $\epsilon_{avg}$  and the right curve showing the total population decreasing for increasing (positive)  $\epsilon_{avg}$ . These two curves are slightly asymmetrical and this is due to  $\beta_{env}$  being cooler than  $\beta$ . This means that to become higher temper-

ature limited a higher population for any  $\epsilon_{avg}$  is needed than for the corresponding negative  $\epsilon_{avg}$ .

As  $\tau$  increases and the microbes become more sensitive to their environment the left hand side of the curves in Figure 7 become less populated. The net cooling needed to become low temperature regulated is smaller than the net heating to become high temperature limited and as  $\tau$  increases the microbes need to keep a tighter control on their abiotic temperature and the upper and lower temperature bounds contract towards their ideal temperature  $\beta$ . Therefore, in order to survive, as  $\tau$  increases, heating their abiotic temperature becomes a better strategy for the microbes as the high temperature limiting regime can support a higher number of microbes increasing their ability to adapt to new mutants making this regime more stable than the low temperature limiting regime. Thus we see that surviving simulations tend to have adopted a high temperature limited regime.

Figure 8 shows similar plots this time for  $\epsilon_{avg}$  vs temperature. Here we see for  $\tau = 0$ , there is a linear relationship as expected - the total population remains constant and so changing  $\epsilon_{avg}$  has a linear effect on the temperature. As  $\tau$  increases we see a step like function, where for low and high  $\epsilon_{avg}$  the temperature remains constant, and for a region of  $\epsilon_{avg}$  around 0, there is a linear relationship - this is the region where the system is nutrient limited. As  $\tau$  increases, this transition between the low and the high temperature limits becomes steeper showing that as  $\tau$  increases, the system becomes increasingly less likely to find itself nutrient limited. We also see that increasing  $\tau$  leads to the high and low temperature limits to contract towards  $\beta$  as the microbes habitable temperature range shrinks. Again we can see that for high  $\tau$ , the system is more likely to be in the high temperature limited regime.

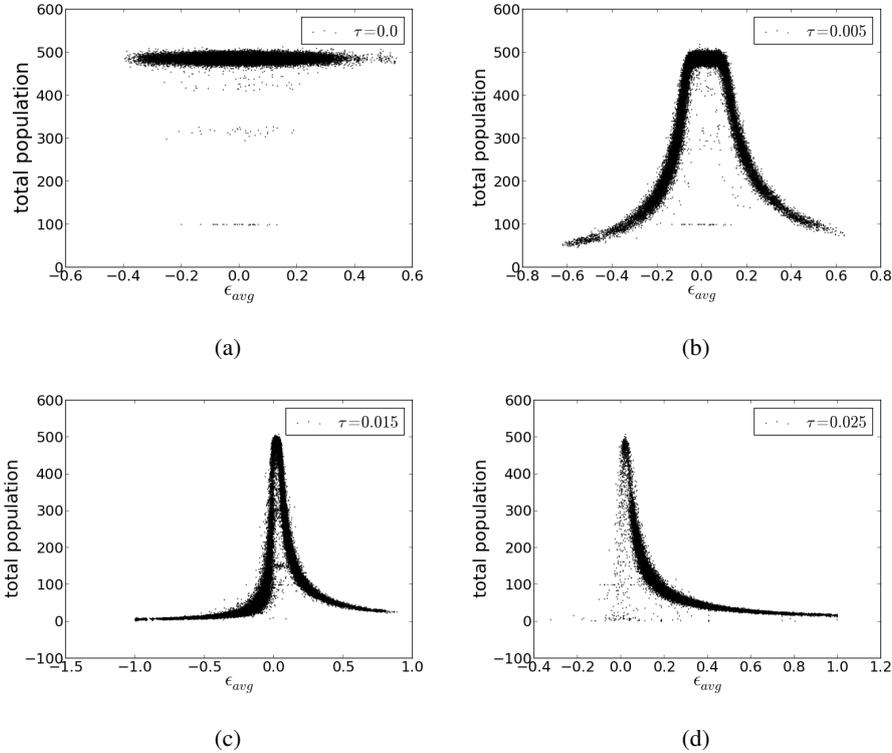


Figure 7: Plots showing  $\epsilon_{avg}$  against total population  $\tau$ . Total population is constant for varying  $\epsilon_{avg}$  when the system is nutrient limited, however for temperature limited, the population must adjust as  $\epsilon_{avg}$  changes to keep the total effect constant.

### 3.6. Changing the environmental abiotic temperature

We investigated the effect of gradually increasing or decreasing the temperature for a range of  $\tau$ . We found that temperature regulation is maintained in the face of a changing temperature, and microbes are able to keep the temperature habitable after the environmental equilibrium temperature would have become uninhabitable. However when quakes occurred, the system was highly susceptible to extinction once the equilibrium temperature was no longer habitable. Table 2 shows the survival and lifespan statistics for cooling the system from  $\beta_{env} = 100$  to  $\beta_{env} = 50$ , and heating the system from  $\beta_{env} = 100$  to  $\beta_{env} = 200$ .

When changing  $\beta_{env}$  from 100 to 200, the environmental temperature is closer to  $\beta = 150$ , the preferred temperature of the microbes, during the experiment than when  $\beta_{env} = 100$  for the entire experiment. This allows the system to become nutrient limited more often. Typically the temperature limited regime with the largest distance to  $\beta_{env}$  will support a higher population, as more microbes are required to achieve the required heating / cooling for temperature limitation. Recall Figure 4 where the population supported during the low temperature limited regime was significantly lower than the population supported during the high temperature limited regime. While  $\beta_{env} < \beta$  the high temperature limiting regime can on average support a higher population, however when

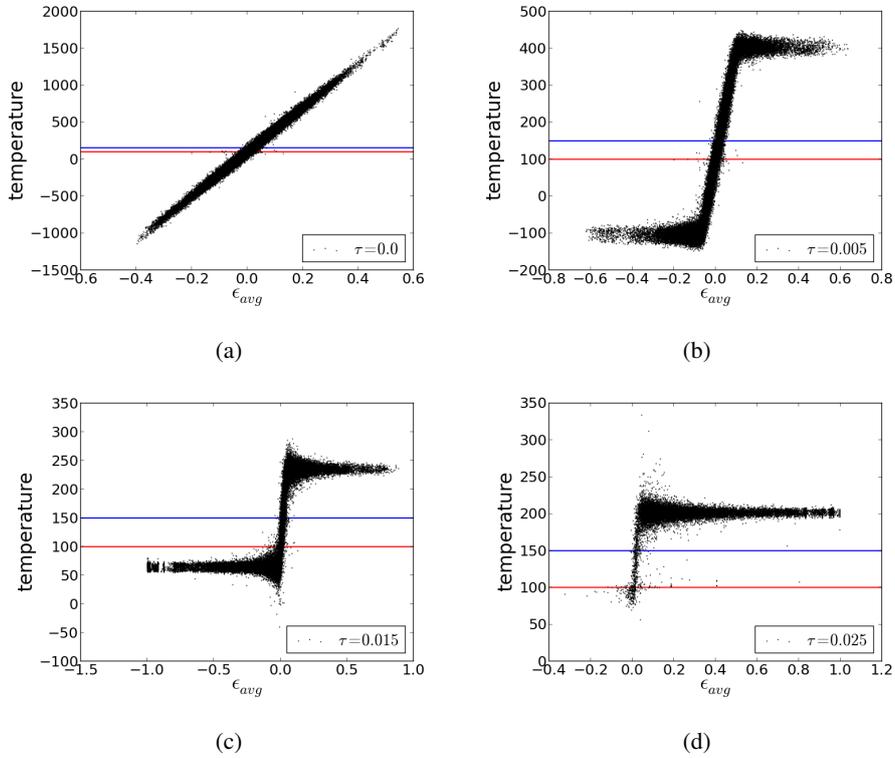


Figure 8: Plots showing  $\epsilon_{avg}$  against the abiotic temperature  $\tau$ . The blue line represents  $\beta$  and the red line represents  $\beta_{env}$ . When the system is nutrient limited the temperature changes linearly with  $\epsilon_{avg}$ , however during temperature limitation, the temperature remains constant for changing  $\epsilon_{avg}$ .

Table 2: Heating and cooling survival % and average lifespans as a fraction of the total simulation length ( $10^5$  timesteps) for a range of  $\tau$ . Comparing with Table 1 we see that heating increases the survival % and cooling decreases it.

$\tau$	Heating			Cooling		
	Survival %	Average lifespan / $10^5$	Standard deviation	Survival %	Average lifespan / $10^5$	Standard deviation
0.00	100	1	0	100	1	0
0.005	100	1	0	100	1	0
0.01	100	1	0	96	1	0.01
0.015	100	1	0	13	0.60	0.26
0.02	87	0.95	0.21	9	0.32	0.30
0.025	19	0.46	0.47	9	0.20	0.32
0.03	2	0.15	0.31	4	0.01	0.22

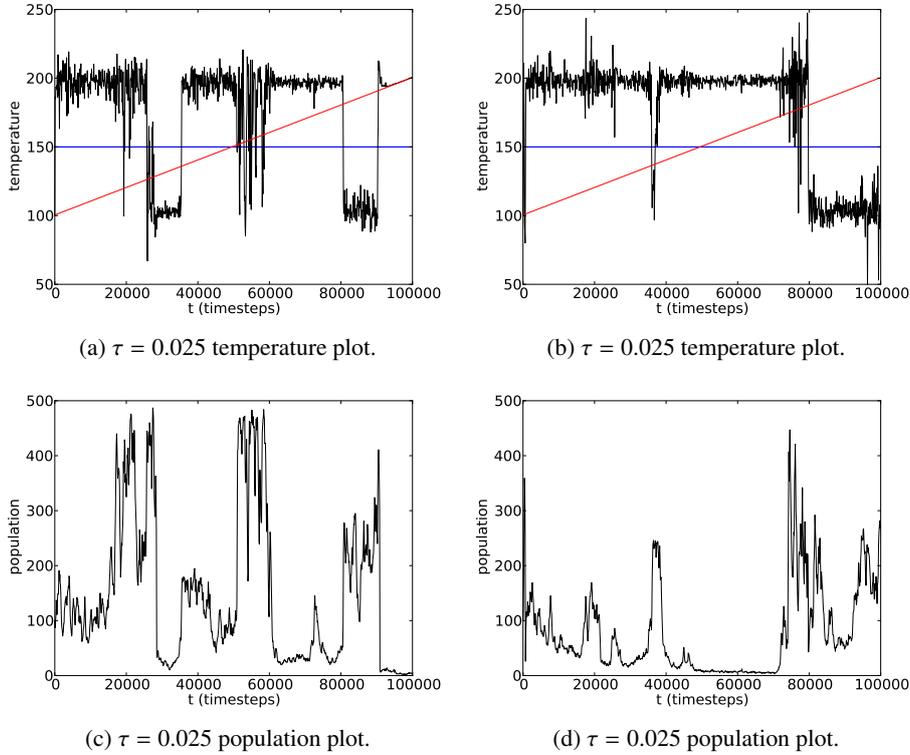


Figure 9: Heating the system with  $\tau = 0.025$ .

the environment has warmed so that  $\beta_{env} > \beta$ , the low temperature regime becomes able to support a higher population. As  $\tau$  increases, the system becomes highly susceptible to quakes. Recall from Figure 5 that for  $\tau = 0.03$  the low temperature limiting regime was no longer possible. When we change from  $\beta_{env} = \beta - 50$  to  $\beta_{env} = \beta + 50$  we change which of the temperature limiting regimes is possible. In order for a  $\tau = 0.03$  system to survive heating, it must transition at a correct time from high temperature limitation to low temperature limitation.  $\tau = 0.025$  and  $\tau = 0.03$  systems are already highly susceptible to extinction during quake events so adding a necessary quake in order to survive further reduces the probability of survival.

Figure 9 shows two  $\tau = 0.025$  systems under-

going heating. In Figure 9a we can see that the system transitions to the high temperature limitation regime towards the end of the experiment and the microbes lose control of the temperature regulation, and temperature starts to follow  $\beta_{env}$ . We can see in Figure 9c that the system has not yet gone extinct at the end of the experiment, however it seems likely to do so.

When reducing  $\beta_{env}$  the system is now further from the microbes' preferred  $\beta = 150$  than for the  $\beta_{env} = 100$  experiments. This has the effect of making the low temperature limiting regime unviable earlier than before. The survival rate starts to drop off for lower  $\tau$ .

Figure 10 shows two  $\tau = 0.015$  systems undergoing cooling. One system survives the ex-

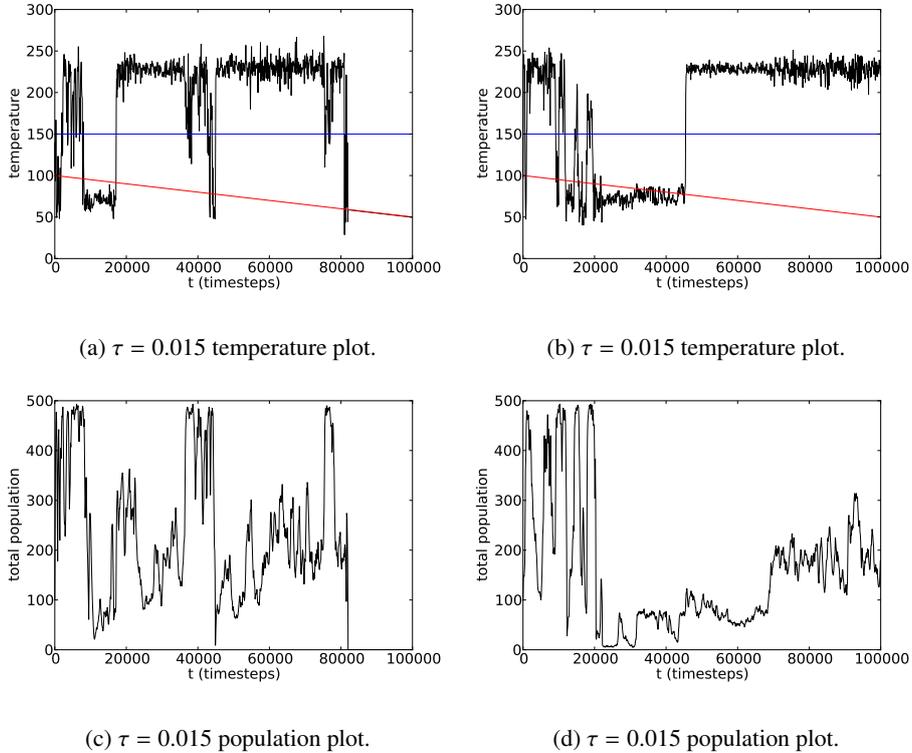


Figure 10: Cooling the system.

periment and one goes extinct. We can see that around roughly 80,000 timesteps,  $\beta_{env}$  becomes too low to support the low temperature limiting regime. For a  $\tau = 0.015$  system to survive therefore it must be in the high temperature limiting regime and remain there. Towards the end of the cooling experiments we have a similar situation to the one we had for high  $\tau$  when  $\beta_{env} = 100$ ; the system must remain high temperature limited to survive, and quake events carry a high probability of total extinction, and these factors combine to reduce the survival probability for systems. Once  $\beta_{env}$  has fallen to below the temperature where low temperature limiting takes place  $T_L$  low temperature limitation becomes impossible. From Figures 10a and 10b we can see that  $T_L \approx 70$ .

If the system can remain in the high temper-

ature limiting regime, then as  $\beta_{env}$  drops the microbes in the system can compensate by increasing their population and thus increasing their heating on the environment. We can see towards the end of Figure 10d that the population is increasing as  $\beta_{env}$  is dropping. If we decreased  $\beta_{env}$  enough, the microbes would reach a limit beyond which they would be unable to heat their environment sufficiently to reach the high temperature limiting regime. Beyond that point no form of temperature limitation is possible anymore and the temperature will fluctuate until the system is pushed to extinction, which given the extreme  $\beta_{env}$  would not take long.

### 3.7. Changing the strength of the microbes' abiotic effects

The results so far have been for microbes with byproduct affects on the abiotic parameter generated from the range  $[-1, 1]$ . To investigate how the strength of these byproduct affects impact temperature regulation in the model, for  $\tau = 0.015$  we performed experiments with microbes with half-strength abiotic byproducts, i.e. taken from the range  $[-0.5, 0.5]$ , denoted experiment  $S_H$ , and microbes with double strength abiotic byproducts, taken from the range  $[-2, 2]$ , denoted experiment  $S_D$ . All 100 simulations survived for  $S_H$ , and 72 survived for  $S_D$ . Both systems showed temperature limitation, however  $S_H$  systems show more nutrient limitation than  $S_D$  systems. The frequency of waiting times for quakes was affected by changing the abiotic byproduct strengths as shown in Figure 11. The definition for a quake is the same as in Equation 4.

With weaker byproduct affects, the system can on average support a higher population of microbes while temperature limited, and this higher population means that there is a higher rate of mutants appearing in the system and thus a higher rate of destabilising mutants. We would then expect to see a higher frequency of quakes for  $S_H$  systems. From Figure 11 it does appear that  $S_H$  do not have to wait as long for quakes, as they have a higher frequency of quakes at small waiting times. For  $S_D$  systems we find the opposite, at small waiting times there is a lower frequency of quakes. With stronger abiotic byproducts, on average there will be a smaller population supported during temperature limitation and thus a slower rate of mutants appearing, thus a smaller chance of destabilising mutants.  $S_D$  systems are also likely to become temperature limited faster than  $S_H$  systems as microbes have a stronger affect on the environment, it is easier for them to move the system away from nutrient limitation to temperature limitation, meaning that the system will spend less time with the tempera-

ture random walking during nutrient limitation, causing fewer rapid changes in temperature appearing in short time spans.

Otherwise the shapes of all three curves in Figure 11 are quite similar, demonstrating that changing the strength of the abiotic byproducts, although affecting the frequency of quakes for short waiting times, it does not have a large impact on the frequency of quakes for long waiting times.

We again have both a histogram from all experiments in Figure 11a and from only non extinct experiments Figure 11b. As  $S_D$  systems are the only systems to have any experiments go extinct, this is the only histogram that differs between the two, and we can see that for only non-extinct experiments, the frequency of quakes for longer waiting times is slightly higher than when we group extinct and non-extinct simulations all together.

### 3.8. Changing $P_{mut}$

We investigated the effect of changing  $P_{mut}$  for  $\tau = 0.015$  systems, and found that the qualitative dynamics of the system remained the same. The survival rate for all systems was 100%. We found that a decreased rate of mutation  $P_{mut} = 0.005$  resulted in fewer quakes at short waiting times, and more at longer waiting times showing the overall dynamics of the system to be slightly slowed down. Increasing the mutation rate to  $P_{mut} = 0.02$  did not largely affect the expected waiting time for quakes. Figure 12 shows a histogram for the waiting times of quakes for each mutation rate. Importantly, temperature regulation was found to be present regardless of the value of  $P_{mut}$ .

## 4. Discussion

The single Flask model demonstrates environmental regulation arising from the byproducts

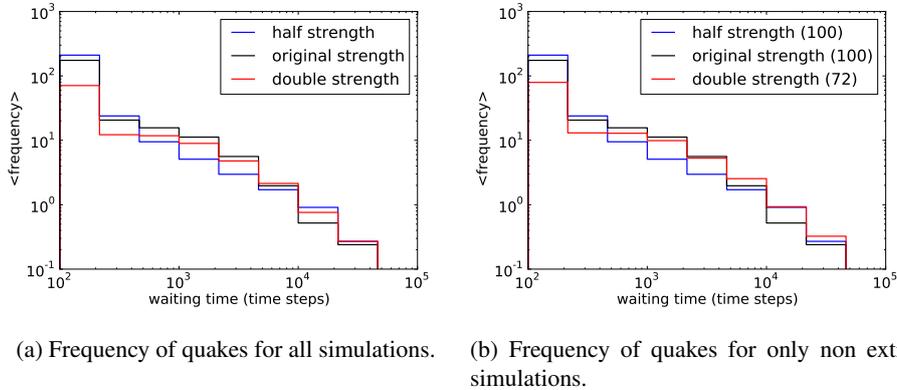


Figure 11: Histograms showing the average frequency of waiting times for quakes for  $\tau = 0.015$  with different microbe byproduct strengths. The number of non-extinct simulations in b) is given in brackets in the legend. Note that both the  $x$  and  $y$  axis are logarithmic.

of biota consistent with evolutionary theory, with a constant rate of mutation per reproduction event, and total extinction a possibility. Microbes in the model share a preference for the abiotic parameters but no spatial structure is present. This combination of assumptions differs from previous Gaian models demonstrating environmental regulation.

This model provides a quantitative illustration of more qualitative ideas from the late 1990s - that the most obvious way to make Gaian ideas compatible with evolutionary theory was for Gaian processes to be based on byproducts of processes that had been selected for other reasons [23][27]. A criticism of the Daisyworld model is that it is set up so that local adaptations of daisies in the system are also beneficial to global regulation. A black daisy is able to survive at a lower solar output than a white daisy due to its low albedo affect. It absorbs more energy thus heating its local environment and in turn the global environment pushing the Daisyworld towards habitability. As the solar output increases, white daisies start to take over to act in the reverse, cooling the planet. The daisies alter both the local and global temperature in the same direction meaning that what is selected for at the individual level di-

rectly impacts its global effects making Daisyworld a special case [32]. The original Multi-Flask model addressed this criticism by having the abiotic effects a byproduct rather than something to be selected for, as is the case in early Daisyworld models [26] and the Guild model [7], allowing environment improving local communities to develop and colonise and outcompete environment degrading communities leading to global regulation.

We have taken a step further in the single Flask model by having no local environment, only global. Any abiotic effect the microbes have acts on this global environment, and so a microbe is no more affected by the temperature increases / decreases due to its own metabolic processes than it is to the increases / decreases of others. This means all microbes feel the abiotic temperature identically at all times and none can gain an individual advantage due to its abiotic effects and competing communities cannot arise. Despite this, clear temperature regulation still appears in the system demonstrating that no spatial environment parameters are needed for temperature regulation to occur.

No single well adapted species emerges but the microbe ecosystem as a whole adjusts its to-

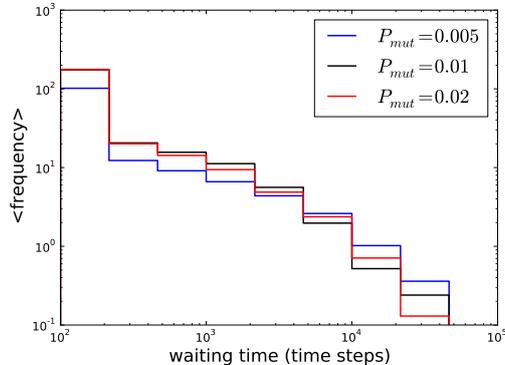


Figure 12: Histograms showing the average frequency of waiting times for quakes per non extinct simulation for  $\tau = 0.015$  systems with different mutation rates. Note that both the  $x$  and  $y$  axis are logarithmic.

tal population to absorb the effects of mutants arising so that the total impact on the environmental parameter is kept constant. Internal perturbations arising from new mutants can knock the system from one temperature limiting regime to another and such transitions are rapid. In this way during temperature limitation we have a single negative feedback loop regulating the system, different from previous abiotic regulation mechanisms, such as the two ‘rein’ feedback in [18]. As all microbes share a preference for the abiotic parameter, subgroups that pull the environment in different directions are not able to form. Instead, our single negative feedback mechanism can be thought of as a single ‘rein’ pulling against the abiotic parameter, with the strength of rein kept constant by a balance between  $\epsilon_{avg}$  and the total population. The single rein can, for suitable  $\tau$  pull in either direction to cause temperature limitation, with occasional regime shifts. Rapid regime shifts are a pattern also seen in nature [25]. A real-world example of these large regime shifts is the evolution of oxygenic photosynthesis in the late Archean ultimately causing a transition from a reducing to an oxidising atmosphere [3]. However as there was more than a 300 Myr delay between the evolution of oxygenic photosynthesis and the rise in oxygen levels this real

world example is more complex than the behaviour portrayed in this model.

Destabilising mutants - so called ‘Ghengis Khan’ species [10] - that greatly upset the current status quo, do not cause the system to go extinct for low and intermediate  $\tau$ , rather they cause the system to quake and return to its previous temperature regulation, or can cause the system to flip from one temperature regulation to another. Dramatic change can take place in the system and yet the system can continue to exist. For high  $\tau$  however, these ‘Ghengis Khan’ species can drive the system to extinction, as shown by the low survival rate for  $\tau > 0.015$ . The rate of mutation in the model will determine how often these large scale destabilisations occur. Each time there is a mutation, the system will need to adjust to absorb the effects of the mutants metabolism. For a slower rate of mutation, the system would be more stable with large transitions occurring at longer intervals, and for a higher rate of mutation, the system would be destabilised more regularly.

We can think of the “regulator” of the system as being the total biomass production rate, rather than genetic distribution. The increase or decrease in the total biomass production rate, and in consequence the expansion and contrac-

tion of the total population is what regulates the temperature, while the genetic variation, determined by the reproduction and mutation rates, acts more like a perturbation the system must adapt to. In the event of a perturbation, the total biomass production rate required to maintain temperature regulation will change as  $\epsilon_{avg}$  will now be different, and this will cause the actual total biomass production rate performed by the system to change. The population will then increase or decrease until the average biomass production rate per microbe, reaches the replacement threshold,  $R_r$ , - the rate at which a stable population can be maintained, i.e. on average each microbe can reproduce once before its death.  $R_r$  is a constant throughout the simulation, so as the total biomass production rate changes as mutants are introduced, the population must adjust. There are no specific ecosystem engineers [12] present in the model. There are no individual species that provide the regulation for other species to benefit from; all species collectively provide the temperature regulation. Some species may contribute towards regulation more than others at certain times, i.e. those with the largest population or the strongest abiotic affects per microbe, however we see from population and nutrient stock graphs that the genetic population of an ecosystem can be rapidly changing while temperature regulation continues uninterrupted. This model demonstrates how non-evolutionary mechanisms, i.e. feedbacks on growth [13], can change the fitness landscape. In the multi-Flask world models, this mechanism would also be present, however the connected flasks allow for a higher-level selection to reduce the harmful perturbations of the ‘wrong’ kind of mutation, adding a second layer of regulation. ‘Key-stone’ species [19], species with a large effect on the environment per biomass, can occur and the death of such microbe’s can also be a trigger for ‘quakes’. If a microbe contributing strongly to the regulation dies, the regulation might be disrupted enough to allow for a transition to another regime.

The single Flaskmodel presents situations in which microbes sit in a nutrient-rich state but are unable to exploit them. With a nutrient rich environment we might expect the arrival of microbes able to exploit these abundant nutrients, however as it is the total biomass production that controls regulation in the temperature regulation state, all microbes will have their biomass production rates, and thus their reproduction rates, limited to the same value. In a shared environment, there is no way to select for a microbe that neutrally affects the current temperature value and thus can grow to a large population, exploiting the nutrients without affecting the temperature regulation. Even if the system could support a higher number of certain species and maintain regulation, there is no mechanism by which those species can reproduce at a faster rate, while ones that would destabilise the system do not. While it would be advantageous for the microbes to ‘remove’ their feedback on the environment, as it would remove extinction causing quake events, no organism can be independent of the physical environment, so the feedback will always exist. Life must take the resources it requires from the local environment and must dump its waste products into the environment [28].

The first Flaskmodel [30] focused on syntrophy - cross-feeding or producer-consumer relationships, in a single flask. Robust nutrient recycling loops were found and we find this in our model too; if the microbe waste is removed immediately from the system after excretion, the total population of the system is reduced. The environment, however, in a single well mixed flask is the same for all microbes, and all resources are exchanged via the environment. Hence no ‘exclusive’ syntrophic relationships can emerge as any relationship is open to exploitation or parasitism and no multi-strain assemblages can be distinguished at a scale smaller than the single well-mixed flask.

The same system for a different value of  $\tau$  can be nutrient limited or temperature limited. Nat-

ural systems can switch between nutrient limitation and some other abiotic environmental limitation, i.e. nutrient runoff from farmland into lakes leading to eutrophication - in this case the system goes from nutrient limited to some other limiting regime [20] [11], or potentially the response of some plants in the arctic tundra in response to warming where the plant may go from temperature limitation to nutrient limitation in areas of the High Arctic where nutrient levels are low [24]. Therefore a model that can present both behaviours is useful although these smaller real-world examples are not thought to be regulating their environments as strongly as this model demonstrates.

A longstanding argument against the Gaia hypothesis is that with just one Earth global regulation has no mechanism for developing because natural selection requires selection between multiple entities [5]. This model shows that an ecosystem of 'temperature' sensitive microbes reacting in a simple way to changes in a global temperature can lead to robust temperature regulation. The system can be thought of as bouncing off an upper or lower bound, similar to oxygen levels on Earth being upwardly bound by fire in the Phanerozoic [14]. This result weakens this criticism of the Gaia hypothesis. This temperature-regulation occurs only when the microbes are sensitive to the abiotic temperature. For regulation to occur there must be a feedback on the biota from the environment, without this, i.e. for low  $\tau$ , the temperature cannot be regulated. When the microbes are sensitive to their environment however, temperature regulation robustly arises. For low to intermediate values of  $\tau$ , systems have a high survival rate despite quakes upsetting the system. This suggests that for a range of  $\tau$  between roughly  $0 \leq \tau \leq 0.015$ , we have what is known as a 'probable Gaia' [15]. Systems tend towards stability with total extinction being a rare event. For higher values of around  $\tau \geq 0.02$  we start to see extinctions becoming more probable; systems are less able to cope with quakes. The systems that survive

do so due to the low number of quakes experienced during the experiment. In this scenario surviving systems are known as 'lucky Gaia' systems [15]. Those that survive do so due to the low frequency quake inducing mutants arising.

The single Flask model has a number of limitations, the largest being that the abiotic effects by the microbes are a direct by-product of microbe metabolism. For each biomass created, a constant value determined by the microbes' genetics is added to the temperature. In the real world, direct heating effects from respiration have very little effect on parameters such as global temperatures. The effects come from the chemicals in the system, and these are influenced by life by what the life removes and provides to the system. In the example of global temperature, the concentrations of  $CO_2$  and  $CH_4$  are key in determining how insulating the Earth's atmosphere is. Another limitation is that the inflow and outflow rates providing and removing nutrients and abiotic parameters to the flask are rather rapid. This means that once a population of microbes slow or cease their metabolic activity, the abiotic parameters are quickly pulled back towards the equilibrium value, the value it would have in the absence of life. This means that microbes do not have to suffer the consequences of destabilising mutants for too long and it enables the system to recover quickly - quickly here meaning a shorter length of time than the average lifespan of a microbe. The results of destabilising mutants are also almost immediately felt by the system allowing for rapid adaptation as soon as destabilising mutants arise. In many real world systems the feedback from a perturbation can take timescales that are much longer than the lifespan of the organisms creating the perturbation, for example the 300 Myr delay between the evolution of oxygenic photosynthesis and the Great Oxidation Event [2]. Therefore the effect of a destabilising mutant might take many generations to be felt preventing a rapid adjustment in the system before the mu-

tant has left many descendants. This might prevent the system from successfully adapting to absorb perturbations of new mutants and might weaken the regulation mechanism. Another assumption the Flaskmodel makes is that types of metabolism are not correlated with certain effects on the environment, for example heterotrophy and the production of CO<sub>2</sub>, which warms the planet. In the Flaskmodel the same metabolism can have different affects on the abiotic parameters depending on the genome of the microbe with that metabolism. This is more general than in the real world, and a limitation of the model.

We speculate that the identified regulation mechanism could in principle operate in natural well mixed environments, such as the Earth's atmosphere - i.e. regulation by fire near the upper bound for Oxygen [14]. Historically, massive regime shifts have occurred in global temperature and atmospheric composition without interrupting the existence of life on the planet. Our model shows similar patterns, and suggests a mechanism for how organisms might collectively interact via global parameters to regulate their environment.

## Acknowledgements

We thank the Gaia Charity and the University of Exeter for their support of this work.

## References

- [1] Andersen, C. W. and Sibani, P. (2016). Tangled nature model of evolutionary dynamics reconsidered: Structural and dynamical effects of trait inheritance. *Physical Review E*, 93(5):052410–.
- [2] Brocks, J. J., Logan, G. A., Buick, R., and Summons, R. E. (1999). Archean molecular fossils and the early rise of eukaryotes. *Science*, 285(5430):1033–1036.
- [3] Catling, D. C., Zahnle, K. J., and McKay, C. P. (2001). Biogenic methane, hydrogen escape, and the irreversible oxidation of early earth. *Science*, 293:839–843.
- [4] Clynes, M. (1969). Cybernetic implications of rein control in perceptual and conceptual organization. *Annals of the New York Academy of Sciences*, 156(2):629–664.
- [5] Dawkins, R. (1982). *The Extended Phenotype*. Oxford University Press.
- [6] Doolittle, W. (1981). Is nature really motherly? *Coevol Q*, pages 58–63.
- [7] Downing, K. and Zvirinsky, P. (1999). The simulated evolution of biochemical guilds: reconciling gaia theory and natural selection. *Artificial Life*, 5(4):291–318.
- [8] Dyke, J. G. (2010). The daisystat: A model to explore multidimensional homeostasis.
- [9] Dyke, J. G. and Weaver, I. (2013). The emergence of environmental homeostasis in complex ecosystems. *PLoS Computational Biology*, 9(5).
- [10] Hamilton, W. D. (1995). Ecology in the large: Gaia and genghis khan. *Journal of Applied Ecology*, 32(3):451–453.
- [11] Janssen, A. B., Teurlincx, S., An, S., Janse, J. H., Paerl, H. W., and Mooij, W. M. (2014). Alternative stable states in large shallow lakes? *Journal of Great Lakes Research*, 40:813–826.
- [12] Jones, C. G., Lawton, J. H., and Shachak, M. (1994). Organisms as ecosystem engineers. *Oikos*, 69(3):373–386.
- [13] Lenton, T. M. (1998). Gaia and natural selection. *Nature*, 394(6692):439–447.
- [14] Lenton, T. M. (2001). The role of land plants, phosphorus weathering and fire in the rise and regulation of atmospheric oxygen. *Global Change Biology*, 7:613–629.
- [15] Lenton, T. M. and Wilkinson, D. M. (2003). Developing the gaia theory. a response to the criticisms of kirchner and volk. *Climatic Change*, 58(1):1–12.
- [16] Lovelock, J. E. (2000). *The Ages of Gaia: A Biography of Our Living Earth*, 2nd edition. OUP Oxford, Oxford.
- [17] Lovelock, J. E. and Margulis, L. (1974). Atmospheric homeostatis by and for the biosphere: the gaia hypothesis. *Tellus*, 26:2–10.
- [18] McDonald-Gibson, J., Dyke, J. G., Di Paolo, E., and Harvey, I. (2008). Environmental regulation can arise under minimal assumptions. *Journal of Theoretical Biology*, 251(4):653–666.
- [19] Paine, R. T. (1969). A note on trophic complexity and community stability. *The American Naturalist*, 103(929):91–93.
- [20] Scheffer, M., Hosper, S., Meijer, M.-L., Moss, B., and Jeppesen, E. (1993). Alternative equilibria in shallow lakes. *Trends in Ecology and Evolution*, 8:275–279.
- [21] Tilman, D. (1980). Resources: A graphical-mechanistic approach to competition and predation. *The American Naturalist*, 116(3):362–393.
- [22] Tyrrell, T. (2004). Biotic plunder: control of the environment by biological exhaustion of resources. *Schneider, S.H., Miller, J.R., Crist, E. and Boston, P.J. (eds.) Scientists debate Gaia: the next Century*. Cambridge MA, USA, MIT Press, pages 137–147.

- [23] Volk, T. (1998). *Gaia's Body: Towards a Physiology of Earth*. Springer New York.
- [24] Walker, M. D., Wahren, C. H., Hollister, R. D., Henry, G. H. R., Ahlquist, L. E., Alatalo, J. M., Bret-Harte, M. S., Calef, M. P., Callaghan, T. V., Carroll, A. B., Epstein, H. E., Jónsdóttir, I. S., Klein, J. A., Magnússon, B., Molau, U., Oberbauer, S. F., Rewa, S. P., Robinson, C. H., Shaver, G. R., Suding, K. N., Thompson, C. C., Tolvanen, A., Totland, Ø., Turner, P. L., Tweedie, C. E., Webber, P. J., and Wookey, P. A. (2006). Plant community responses to experimental warming across the tundra biome. *Proceedings of the National Academy of Sciences*, 103(5):1342–1346.
- [25] Wang, R., Dearing, J. A., Langdon, P. G., Zhang, E., Yang, X., Dakos, V., and Scheffer, M. (2012). Flickering gives early warning signals of a critical transition to a eutrophic lake state. *Nature*, 492(7429):419–422.
- [26] Watson, A. J. and Lovelock, J. E. (1983). Biological homeostasis of the global environment: the parable of daisyworld. *Tellus*, 35B:284–289.
- [27] Wilkinson, D. M. (1999). Is gaia really conventional ecology? *Oikos*, 84:533–536.
- [28] Wilkinson, D. M. (2006). *Fundamental Processes in Ecology: An Earth Systems Approach*. Oxford University Press.
- [29] Williams, H. T. and Lenton, T. M. (2008). Environmental regulation in a network of simulated microbial ecosystems. *PNAS*, 105(30):10432–10437.
- [30] Williams, H. T. P. and Lenton, T. M. (2007). The flask model: emergence of nutrient-recycling microbial ecosystems and their disruption by environment-altering 'rebel' organisms. *Oikos*, 116(7):1087–1105.
- [31] Williams, H. T. P. and Lenton, T. M. (2010). Evolutionary regime shifts in simulated ecosystems. *Oikos*, 119(12):1887–1899.
- [32] Wood, A. J., Ackland, G. J., Dyke, J. G., Williams, H. T. P., and Lenton, T. M. (2008). Daisyworld: A review. *Review of Geophysics*, 46(1).

# Appendices

## A. Model Description

### A.1. The Flask Environment

We have a single well mixed environment with no spatial element - we assume that in the flask the liquid medium is well mixed so that the composition of the flask is in a homogeneous steady state. The flask is characterised by nutrient levels and the abiotic parameters. The nutrients present may be consumed by microbes and converted into biomass. The abiotic parameters are affected by and can affect the microbe activity.

The state of the flask is given by a vector  $V$ :

$$V = (n_1, \dots, n_N, a_1, \dots, a_A) = (v_1, \dots, v_{N+A}) \quad (5)$$

where  $n_i$  is the concentration of nutrient  $i$ ,  $\beta_{env}^i$ , is the level of abiotic parameter  $i$ , or equivalently,  $v_i$ , is the level of the  $i^{th}$  environmental state variable.  $N$  is the number of nutrients and  $A$  is the number of abiotic parameters.

As we break down each timestep into a number of iterations  $n$  where  $n$  is the total population of the system at the start of the timestep, we break down the inflow and outflow of nutrients and other abiotic parameters to prevent sudden changes at the the start of each timestep. The steps within a timestep would ideally all be computed in parallel but computational limitations prevent this, and so for agent based dynamics we effectively freeze the system while the selected microbes performs an action (being nutrient consumption / biomass production / reproduction / death). If we simply added and deducted the flow amounts at the start of each timestep, microbes selected at the beginning of a timestep could see a very different world to those selected at the end of a timestep if the population is large due to the microbes effect on the environment (nutrient consumption reducing nutrient levels and biomass creation affecting the abiotic parameters). Although these effects would largely average out due to the random selection of microbes during each timestep, a single large influx per timestep could be thought of as a periodic perturbation on the system which could affect the results seen. To counter this, we calculate the net influx of nutrients  $N^{net}$  and abiotic parameters at the start of each timestep:

$$N_{net} = I_N - O_N N^{current} \quad (6)$$

where  $I_N$  is the number of units of nutrient inflow per timestep,  $O_N$  is the percentage outflow, and  $N^{current}$  is the current nutrient levels in the system at the start of the timestep. We can then do  $N^i = N^{net} / K^{current}$  where  $K^{current}$  is the total population of the system at the start of the timestep, and then for each iteration within a timestep we increment the nutrient levels by  $N^i$ . This results in the same quantity of nutrients being added / removed from the system as if there was just one update at the start of the timestep, but it results in a much smoother transition and means that

microbes selected at the start and end of a timestep will see much more similar worlds. In doing this, we treat nutrient levels as continuous but the microbes can only ever treat the nutrients as units. So while each iteration we might be adding 10.7 nutrient units per iteration, any microbes in the system can only act on the integer amounts of nutrients present.

We calculate the abiotic parameter changes by diluting the current abiotic parameters by a certain percentage of fresh influx  $I_A$ . So for the abiotic parameters we update each iteration by  $A^{net}$ :

$$A^{net} = A^{source} I_A - A^{current} I_A \quad (7)$$

where  $A^{source}$  is the abiotic parameters of the source, and  $A^{current}$  is the abiotic parameters of the current environment in the flask.

### A.2. Microbes

The microbes consume and excrete nutrients in fixed proportions and affect the levels of abiotic parameters in their environment as a side effect of biomass creation. The ratios of nutrient consumption / excretion and the byproduct effect on the abiotic parameters are genetically encoded for each microbe species. All microbes share the same preferred abiotic conditions (i.e. the state of the abiotic environment which results in the maximum growth rate). Microbes grow by consuming nutrients and converting them to biomass, and they reproduce asexually by splitting once their biomass reaches a threshold. Biomass is reduced by a fixed amount per timestep to represent the cost of staying alive. Microbes die if their biomass drops to a fixed threshold, which can happen during nutrient limitation or temperature limitation causing the microbes being unable to consume the nutrients present.

In the code we do not record microbes of the same species individually as doing so would slow the code considerably. Instead we group microbes of the same species together and record the species total biomass. Thus each species can be thought of as a vector  $S$ :

$$S = (N, B, \lambda, \mu, \alpha, \beta) \quad (8)$$

where  $N$  is the population of the species,  $B$  is the total biomass of the species,  $\lambda = (1, \dots, N)$  represents the ratio in which nutrients are consumed,  $\mu = (1, \dots, N)$  represents the ratio in which excreta are returned to the environment as nutrients,  $\alpha = (1, \dots, A)$  represents the effect from one microbe in the species on the abiotic parameters and  $\beta = (1, \dots, A)$  represents the environmental abiotic parameter levels that maximise the growth for microbes in species  $S$ .

### A.3. Genotype

The genotype of a microbe is recorded as the decimal representation of an 8 bit binary string, and this is used to group microbes into species. Microbes that share the same genome are of the same species. We create tables for microbe nutrient / excretion rules and abiotic effects and this

genome is used as the reference to look up the particular metabolism rules for a microbe. With an 8 bit long binary genome there are 256 possible species (as each gene in a genome can have the value 0 or 1).

#### A.4. *Reproduction and Mutation*

If the microbe is able to consume enough nutrients to reach the reproduction threshold  $T_R$  it will reproduce asexually, splitting in half. Half of the biomass will go to the new microbe and the parent microbe will half its biomass. The new microbe will have the same genome as the parent unless a mutation occurred during the reproduction. There is a small constant probability of mutation for each locus. If a mutation occurs at a locus then the gene at that point will be 'flipped', turning it to 0 if it were previously 1, or to 1 if it were previously 0.

#### A.5. *Maintenance Cost and Death*

There is a fixed biomass cost of staying alive for each microbe. This reduces a microbes biomass by a constant rate. This cost represents the energy costs of maintaining cellular machinery and metabolic inefficiency. This cost is assumed to be lost from the flask environment as unrecoverable heat radiation. This ensures that the nutrients cannot be infinitely recycled and it sets the carrying capacity of the system. This carry capacity is reached when the total heat dissipation matches the energy supplied in the form of nutrients.

If the biomass falls to a starvation threshold  $T_D$  the microbe will starve to death. There is also a small probability of death by natural causes  $P_D$  that represents death by predation, apoptosis etc. When a microbe dies its biomass is be removed from the system, as if the dead microbe were washed out of the flask.

#### A.6. *Nutrient Consumption / Excretion*

During a single timestep a different microbe is selected  $n$  times for a nutrient consumption event, where  $n$  is the total population of the system at the start of the timestep. This means that on average every microbe will be selected for nutrient consumption once per timestep. When a microbe is selected it will attempt to eat its  $C_j^{max}$  of nutrients (the value of  $C_j^{max}$  depending on how closely the abiotic parameters meet the microbes' preferred values and the microbes' sensitivity to its environment ), and if the nutrients are available, and in the correct ratios, the microbe will consume them. The nutrient ratios are fixed at the start of each simulation for each genome and remain constant.

The nutrient consumption / excretion vectors for each genome are of  $N$  length, there  $N$  is the number of nutrients. If we assume we have 3 nutrients we would then have 2 vectors of length 3. We populate these vectors with random numbers generated between  $[-1, 1]$  and then sum. For example if our two vectors were  $[-0.3, 0.5, 0.6]$  and  $[-0.2, -0.2, 0.1]$  then summed we would have:  $[-0.5, 0.3, 0.7]$ . We take negative values to mean that nutrient is excreted and positive values that that nutrient is consumed. Therefore any case where all values in the vector are

positive or all are negative are instantly disqualified as a microbe must eat and excrete. For our example above we see that our microbe consumes nutrients 2 and 3 and excretes nutrient 1. When consuming nutrients this microbe must eat 3 units of nutrient 2 with 7 nutrients of nutrient 3 (a unit of nutrient is non divisible), or the microbe cannot consume anything. This particular metabolism is limiting the microbe to be only able to survive in ideal abiotic conditions, if we take our  $C^{max} = 10$  (the maximum consumption rate for any microbe) as if the abiotic conditions move away, we get  $C_j^{max} < C^{max}$  and so  $C_j^{max} < 10$  and with our specific nutrient ratio, if the microbe cannot eat 10 units of nutrient, it cannot consume at all or it would violate its metabolic nutrient ratio rules.

#### A.7. Effect of abiotic factor on metabolic rate

The state of the abiotic environment affects the rate at which microbes can consume nutrients which in turn affects the rate of biomass production and thus the growth of the microbes. A microbe will attempt to consume a maximum amount  $C_j^{max}$  of nutrients each timestep with the demand being met depending on nutrient availability. The  $C_j^{max}$  is calculated for each microbe  $j$  as a function of the match between the microbes' genetically specified preferred conditions and the current abiotic state of the environment. This function is has a Gaussian form and falls away smoothly from its maximum as the distance between the optimum and the current environment increases. Mathematically we write this as:

$$C_j^{max} = \psi_j C^{max} \quad (9)$$

$$\psi_j = e^{-(\tau p_j)^2} \quad (10)$$

$$p_j = \sqrt{\sum_{i=1}^A (\beta_{env}^i - \beta_j^i)^2} \quad (11)$$

where  $C_{max}$  is a constant determining the maximum rate of consumption for any microbe,  $\psi_j$  is a microbe specific measure of the microbe's satisfaction with the current abiotic environment,  $\tau$  is a universal constant parameter that determines how sensitive the microbes are to their environment ( $\tau = 0$  means the microbes are not affected by the abiotic environment at all, and a higher  $\tau$  means the microbes become more sensitive to the abiotic conditions).  $p_j$  is a measure of the distance between the current environmental level for each abiotic factor  $\beta_{env}^i$  and the microbe's preferred level  $\beta_j^i$ .

#### A.8. Effect of microbial activity on environment

Microbes can affect their abiotic environment as a side effect of biomass creation. The effect the microbe has is proportional to its rate of biomass creation and thus its growth rate, so faster

growing species will have a larger effect than slower growing species. Through the consumption of nutrients and excretion of waste products microbes also affect the nutrient levels in the environment.

Each microbe has an effect on the abiotic parameters per unit of biomass created, and these effects are numbers in the range  $[-1, 1]$ . These numbers are randomly generated in this range at the beginning of each simulation for each species and remains constant throughout the simulation. Thus each member of a species has the same effect on the abiotic environment for the duration of the simulation.

#### A.9. Parameters

Parameter	Value	Description
$N$	4	Number of nutrients
$A$	1	Number of abiotic parameters
$B_R$	120	Reproduction threshold (biomass units)
$B_D$	50	Starvation threshold (biomass units)
$P_{mut}$	0.01	Probability of mutation at each locus during reproduction
$P_D$	0.002	Probability of death by natural causes (other than starvation) at each timestep
$\lambda$	1	Maintenance cost (biomass units / timestep)
$\theta$	0.6	Nutrient conversion efficiency
$\tau$	[0.00, 0.005, 0.01, 0.015, 0.02, 0.025, 0.03]	Level of influence of abiotic environment on metabolism
$I_N$	150	Rate of nutrient influx (units / timestep)
$O_N$	0.25	Rate of nutrient outflux (percentage / timestep)
$I_A$	0.2	Rate of abiotic factor influx (percentage / timestep)
$O_A$	0.2	Rate of abiotic factor outflux (percentage / timestep)
$K_M$	100	Number of individuals in flask inoculum
$t_{prep}$	500	Flask equilibration time prior to seeding (timesteps)
$t_{run}$	$10^5$	Duration of run (timesteps)
$\beta$	150	Abiotic environmental preference
$\beta_{env}$	100	Environmental ‘temperature’ in the absence of microbe activity

#### A.10. Method

We used agent based dynamics to run the simulation. A timestep is broken down into iterations, the number of iterations matches the number of microbes alive in the system at the start of the timestep. For each iteration we perform the following steps:

- Influx / outflux of abiotic parameters and nutrients via trickle

- An individual selected randomly for a death event
- An individual selected randomly for a nutrient consumption event
- An individual selected randomly for a biomass creation event
- An individual selected randomly for a reproduction event

We repeated this process  $n$  times for one timestep.

Each simulation in this paper was run for  $10^5$  timesteps.