

Modelling Ecological Systems with the Calculus of Wrapped Compartments^{*}

Pablo Ramón¹ and Angelo Troina²

¹ Instituto de Ecología, Universidad Técnica Particular de Loja, Ecuador

² Dipartimento di Informatica, Università di Torino, Italy

Abstract. The Calculus of Wrapped Compartments is a framework based on stochastic multiset rewriting in a compartmentalised setting originally developed for the modelling and analysis of biological interactions. In this paper, we propose to use this calculus for the description of ecological systems and we provide the modelling guidelines to encode within the calculus some of the main interactions leading ecosystems evolution. As a case study, we model the distribution of height of *Croton wagneri*, a shrub constituting the endemic predominant species of the dry ecosystem in southern Ecuador. In particular, we consider the plant at different altitude gradients (i.e. at different temperature conditions), to study how it adapts under the effects of global climate change.

Keywords: Calculus of Wrapped Compartments, Stochastic Simulations, Computational Ecology

1 Introduction

Answers to ecological questions could rarely be formulated as general laws: ecologists deal with *in situ* methods and experiments which cannot be controlled in a precise way since the phenomena observed operate on much larger scales (in time and space) than man can effectively study. Actually, to carry on ecological analyses, there is the need of a “macroscope”!

Theoretical and Computational Ecology, the scientific disciplines devoted to the study of ecological systems using theoretical methodologies together with empirical data, could be considered as a fundamental component of such a macroscope. Within these disciplines, quantitative analysis, conceptual description techniques, mathematical models, and computational simulations are used to understand the fundamental biological conditions and processes that affect populations dynamics (given the underlying assumption that phenomena observable across species and ecological environments are generated by common, mechanistic processes) [39].

Ecological models can be deterministic or stochastic [18]. Given an initial system, deterministic simulations always evolve in the same way, producing a

^{*} This work has been partially sponsored by the BioBITs Project (*Converging Technologies 2007*, area: Biotechnology-ICT), Regione Piemonte. Angelo Troina is visiting the Instituto de Ecología at UTPL (Loja, Ecuador) within the Prometeo program founded by SENESCYT.

unique output [43]. Deterministic methods give a picture of the average, expected behaviour of a system, but do not incorporate random fluctuations. On the other hand, stochastic models allow to describe the random perturbations that may affect natural living systems, in particular when considering small populations evolving at slow interactions. Actually, while deterministic models are approximations of the real systems they describe, stochastic models, at the price of an higher computational cost, can describe exact scenarios.

A model in the Calculus of Wrapped Compartments (CWC for short) consists of a term, representing a (biological or ecological) system and a set of rewrite rules which model the transformations determining the system's evolution [27,24]. Terms are defined from a set of atomic elements via an operator of compartment construction. Each compartment is labelled with a nominal type which identifies the set of rewrite rules that may be applied into it. The CWC framework is based on a stochastic semantics and models an exact scenario able to capture the stochastic fluctuations that can arise in the system.

The calculus has been extensively used to model real biological scenarios, in particular related to the AM-symbiosis [24,19].³ An hybrid semantics for CWC, combining stochastic transitions with deterministic steps, modelled by Ordinary Differential Equations, has been proposed in [26,25].

While the calculus has been originally developed to deal with biomolecular interactions and cellular communications, it appears to be particularly well suited also to model and analyse interactions in ecology. In particular, we present in this paper some modelling guidelines to describe, within CWC, some of the main common features and models used to represent ecological interactions and population dynamics. A few generalising examples illustrate the abstract effectiveness of the application of CWC to ecological modelling.

As a real case study, we model the distribution of height of *Croton wagneri*, a shrub in the dry ecosystem of southern Ecuador, and investigate how it could adapt to global climate change.

2 The Calculus of Wrapped Compartments

The Calculus of Wrapped Compartments (CWC) (see [27,26,25]) is based on a nested structure of compartments delimited by wraps with specific proprieties.

2.1 Term Syntax

Let \mathcal{A} be a set of *atomic elements* (*atoms* for short), ranged over by a, b, \dots , and \mathcal{L} a set of *compartment types* represented as *labels* ranged over by $\ell, \ell', \ell_1, \dots$

³ Arbuscular Mycorrhiza (AM) is a class of fungi constituting a vital mutualistic interaction for terrestrial ecosystems. More than 48% of land plants actually rely on mycorrhizal relationships to get inorganic compounds, trace elements, and resistance to several kinds of pathogens.

Definition 1 (CWC terms). A CWC term is a multiset \bar{t} of simple terms t defined by the following grammar:

$$t ::= a \quad | \quad (\bar{a} \rfloor \bar{t}')^\ell$$

A simple term is either an atom or a compartment consisting of a *wrap* (represented by the multiset of atoms \bar{a}), a *content* (represented by the term \bar{t}') and a *type* (represented by the label ℓ). Multisets are identified modulo permutations of their elements. The notation $n * t$ denotes n occurrences of the simple term t . We denote an empty term with \bullet .

In applications to ecology, atoms can be used to describe the individuals of different species and compartments can be used to distinguish different ecosystems, habitats or ecological niches. Compartment wraps can be used to model geographical boundaries or abiotic components (like radiations, climate, atmospheric or soil conditions, etc.). In evolutionary ecology, individuals can also be described as compartments, showing characteristic features of their *phenotype* in the wrap and keeping their *genotype* (or particular *alleles* of interest) in the compartment content.

An example of CWC term is $20 * a \ 12 * b \ (c \ d \rfloor 6 * e \ 4 * f)^\ell$ representing a multiset (denoted by listing its elements separated by a space) consisting of 20 occurrences of a , 12 occurrence of b (e.g. 32 individuals of two different species) and an ℓ -type compartment $(c \ d \rfloor 6 * e \ 4 * f)^\ell$ which, in turn, consists of a wrap (a boundary) with two atoms c and d (e.g. two abiotic factors) on its surface, and containing 6 occurrences of the atom e and 4 occurrences of the atom f (e.g. 10 individuals of two other species). Compartments can be nested as in the term $(a \ b \ c \rfloor (d \ e \rfloor f)^{\ell'} \ g \ h)^\ell$.

2.2 Rewrite Rules

System transformations are defined by rewrite rules, defined by resorting to CWC terms that may contain variables.

Definition 2 (Patterns and Open terms). Simple patterns P and simple open terms O are given by the following grammar:

$$\begin{array}{lcl} P & ::= & a \quad | \quad (\bar{a} \ x \rfloor \bar{P} \ X)^\ell \\ O & ::= & a \quad | \quad (\bar{q} \rfloor \bar{O})^\ell \quad | \quad X \\ q & ::= & a \quad | \quad x \end{array}$$

where \bar{a} is a multiset of atoms, \bar{P} is a pattern (i.e., a, possibly empty, multiset of simple patterns), x is a wrap variable (can be instantiated by a multiset of atoms), X is a content variable (can be instantiated by a CWC term), \bar{q} is a multiset of atoms and wrap variables and \bar{O} is an open term (i.e., a, possibly empty, multiset of simple open terms).

We will use patterns as the l.h.s. components of a rewrite rule and open terms as the r.h.s. components of a rewrite rule. Patterns are intended to match, via

substitution of variables, with ground terms (containing no variables). Note that we force *exactly* one variable to occur in each compartment content and wrap of our patterns. This prevents ambiguities in the instantiations needed to match a given compartment.⁴

Definition 3 (Rewrite rules). A rewrite rule is a triple (ℓ, \bar{P}, \bar{O}) , denoted by $\ell : \bar{P} \mapsto \bar{O}$, where the pattern \bar{P} and the open term \bar{O} are such that the variables occurring in \bar{O} are a subset of the variables occurring in \bar{P} .

The rewrite rule $\ell : \bar{P} \mapsto \bar{O}$ can be applied to any compartment of type ℓ with \bar{P} in its content (that will be rewritten with \bar{O}). Namely, the application of $\ell : \bar{P} \mapsto \bar{O}$ to term \bar{t} is performed in the following way:

1. find in \bar{t} (if it exists) a compartment of type ℓ with content \bar{t}' and a substitution σ of variables by ground terms such that $\bar{t}' = \sigma(\bar{P} \ X)$;⁵
2. replace in \bar{t} the subterm \bar{t}' with $\sigma(\bar{O} \ X)$.

For instance, the rewrite rule $\ell : a \ b \mapsto c$ means that in compartments of type ℓ an occurrence of $a \ b$ can be replaced by c . We write $\bar{t} \mapsto \bar{t}'$ to denote a *reduction* obtained by applying a rewrite rule to \bar{t} resulting to \bar{t}' .

While a rewrite rule does not change the label ℓ of the compartment where it is applied, it may change the labels of the compartments occurring in its content. For instance, the rewrite rule $\ell : (a \ x \mid X)^{\ell_1} \mapsto (a \ x \mid X)^{\ell_2}$ means that, if contained in a compartment of type ℓ , a compartment of type ℓ_1 containing an a on its wrap can be changed to type ℓ_2 .

CWC Models. For uniformity reasons we assume that the whole system is always represented by a term consisting of a single (top level) compartment with distinguished label \top and empty wrap, i.e., any system is represented by a term of the shape $(\bullet \mid \bar{t})^\top$, which, for simplicity, will be written as \bar{t} . Note that while an infinite set of terms and rewrite rules can be defined from the syntactic definitions in this section, a *CWC model* consists of an initial system $(\bullet \mid \bar{t})^\top$ and a finite set of rewrite rules \mathcal{R} .

2.3 Stochastic Simulation

A stochastic simulation model for ecological systems can be defined by incorporating a collision-based framework along the lines of the one presented by Gillespie in [32], which is, *de facto*, the standard way to model quantitative aspects of biological systems. The basic idea of Gillespie's algorithm is that a rate is associated with each considered reaction which is used as the parameter of an exponential probability distribution modelling the time needed for the reaction

⁴ The linearity condition, in biological terms, corresponds to excluding that a transformation can depend on the presence of two (or more) identical (and generic) components in different compartments (see also [36]).

⁵ The implicit (distinguished) variable X matches with all the remaining part of the compartment content.

to take place. In the standard approach the reaction *propensity* is obtained by multiplying the rate of the reaction by the number of possible combinations of reactants in the compartment in which the reaction takes place, modelling the law of mass action.

Stochastic rewrite rules are thus enriched with a rate k (notation $\ell : \overline{P} \xrightarrow{k} \overline{O}$). Evaluating the propensity of the stochastic rewrite rule $R = \ell : a \ b \xrightarrow{k} c$ within the term $\bar{t} = a \ a \ a \ b \ b$, contained in the compartment $u = (\bullet \mid \bar{t})^\ell$, we must consider the number of the possible combinations of reactants of the form $a \ b$ in \bar{t} . Since each occurrence of a can react with each occurrence of b , this number is $3 \cdot 2$, and the propensity of R within u is $k \cdot 6$. A detailed method to compute the number of combinations of reactants can be found in [27].

The stochastic simulation algorithm produces essentially a *Continuous Time Markov Chain* (CTMC). Given a term \bar{t} , a set \mathcal{R} of rewrite rules, a global time δ and all the reductions e_1, \dots, e_M applicable in all the different compartments of \bar{t} with propensities r_1, \dots, r_M , Gillespie’s “direct method” determines:

- The exponential probability distribution (with parameter $r = \sum_{i=1}^M r_i$) of the time τ after which the next reduction will occur;
- The probability r_i/r that the reduction occurring at time $\delta + \tau$ will be e_i .⁶

The CWC simulator [2] is a tool under development at the Computer Science Department of the Turin University, based on Gillespie’s direct method algorithm [32]. It treats CWC models with different rating semantics (law of mass action, Michaelis-Menten kinetics, Hill equation) and it can run independent stochastic simulations over CWC models, featuring deep parallel optimizations for multi-core platforms on the top of FastFlow [5]. It also performs online analysis by a modular statistical framework [4,3].

3 Modelling Ecological Systems in CWC

We present some of the characteristic features leading the evolution of ecological systems, and we show how to encode it within CWC.

3.1 Population Dynamics

Models of population dynamics describe the changes in the size and composition of populations.

The *exponential growth model* is a common mathematical model for population dynamics, where, using r to represent the pro-capita growth rate of a population of size N , the change of the population is proportional to the size of the already existing population:

$$\frac{dN}{dt} = r \cdot N$$

⁶ Reductions are applied in a sequential way, one at each step.

CWC Modelling 1 (Exponential Growth Model) *We can encode within CWC the exponential growth model with rate r using a stochastic rewrite rule describing a reproduction event for a single individual at the given rate. Namely, given a population of species a living in an environment modelled by a compartment with label ℓ , the following CWC rule encodes the exponential growth model:*

$$\ell : a \xrightarrow{r} a \ a$$

Counting the number of possible reactants, the growth rate of the overall population is automatically obtained by the stochastic semantics underlying CWC.

A *metapopulation*⁷ is a group of populations of the same species distributed in different patches⁸ and interacting at some level. Thus, a metapopulation consists of several distinct populations and areas of suitable habitat.

Individual populations may tend to reach extinction as a consequence of demographic stochasticity (fluctuations in population size due to random demographic events); the smaller the population, the more prone it is to extinction. A metapopulation, as a whole, is often more stable: immigrants from one population (experiencing, e.g., a population boom) are likely to re-colonize the patches left open by the extinction of other populations. Also, by the *rescue effect*, individuals of more dense populations may emigrate towards small populations, rescuing them from extinction.

Populations are affected by births and deaths, by immigrations and emigrations (BIDE model [23]). The number of individuals at time $t + 1$ is given by:

$$N_{t+1} = N_t + B + I - D - E$$

where N_t is the number of individuals at time t and, between time t and $t + 1$, B is the number of births, I is the number of immigrations, D is the number of deaths and E is the number of emigrations.

CWC Modelling 2 (BIDE model) *We can encode within CWC the BIDE model for a compartment of type ℓ using stochastic rewrite rules describing the given events with their respective rates r, i, d, e :*

$$\begin{aligned} \ell : a &\xrightarrow{r} a \ a && (\text{birth}) \\ \top : a \ (x \mid X)^\ell &\xrightarrow{i} (x \mid a \ X)^\ell && (\text{immigration}) \\ \ell : a &\xrightarrow{d} \bullet && (\text{death}) \\ \top : (x \mid a \ X)^\ell &\xrightarrow{e} a \ (x \mid X)^\ell && (\text{emigration}) \end{aligned}$$

Starting from a population of N_t individuals at time t , the number N_{t+1} of individuals at time $t + 1$ is computed by successive simulation steps of the stochastic algorithm. The race conditions computed according to the propensities of the given rules assure that all of the BIDE events are correctly taken into account.

⁷ The term metapopulation was coined by Richard Levins in 1970. In Levins' own words, it consists of "a population of populations" [34].

⁸ A patch is a relatively homogeneous area differing from its surroundings.

Example 1. Immigration and extinction are key components of island biogeography. We model a metapopulation of species a in a context of 5 different patches: 4 of which are relatively close, e.g. different ecological regions within a small continent, the last one is far away and difficult to reach, e.g. an island. The continental patches are modelled as CWC compartments of type ℓ_c , the island is modelled as a compartment of type ℓ_i . Births, deaths and migrations in the continental patches are modelled by the following CWC rules:

$$\begin{array}{l} \ell_c : a \xrightarrow{0.005} a \quad \ell_c : a \xrightarrow{0.005} \bullet \\ \top : (x \mid a \mid X)^{\ell_c} \xrightarrow{0.01} a \mid (x \mid X)^{\ell_c} \quad \top : a \mid (x \mid X)^{\ell_c} \xrightarrow{0.5} (x \mid a \mid X)^{\ell_c} \end{array}$$

These rates are drawn considering days as time unites and an average of life expectancy and reproduction time for the individuals of the species a of 200 days ($\frac{1}{0.005}$). For the modelling of real case studies, these rates could be estimated from data collected *in situ* by tagging individuals.⁹ In this model, when an individual emigrates from its previous patch it moves to the top-level compartment from where it may reach one of the close continental patches (might also be the old one) or start a journey through the sea (modelled as a rewrite rule putting the individual on the wrapping of the island compartment):

$$\top : a \mid (x \mid X)^{\ell_i} \xrightarrow{0.2} (x \mid a \mid X)^{\ell_i}$$

Crossing the ocean is a long and difficult task and individuals trying it will probably die during the cruise; the luckiest ones, however, might actually reach the island, where they could eventually benefit of a better life expectancy for them and their descendants:

$$\begin{array}{l} \top : (x \mid a \mid X)^{\ell_i} \xrightarrow{0.333} (x \mid X)^{\ell_i} \quad \top : (x \mid a \mid X)^{\ell_i} \xrightarrow{0.0005} (x \mid a \mid X)^{\ell_i} \\ \ell_i : a \xrightarrow{0.007} a \quad \ell_i : a \xrightarrow{0.003} \bullet \end{array}$$

Considering the initial system modelled by the CWC term:

$$\bar{t} = (\bullet \mid 30 * a)^{\ell_c} (\bullet \mid 30 * a)^{\ell_c} (\bullet \mid 30 * a)^{\ell_c} (\bullet \mid 30 * a)^{\ell_c} (\bullet \mid \bullet)^{\ell_i}$$

we can simulate the possible evolutions of the overall diffusion of individuals of species a in the different patches. Notice that, on average, one over $\frac{0.333}{0.0005}$ individuals that try the ocean journey, actually reach the island. In Figure 1 we show the result of a simulation plotting the number of individuals in the different patches in a time range of approximatively 10 years. Note how, in the final part of the simulation, empty patches get recolonised. In this particular simulation, also, an exponential growth begins after the colonisation of the island. The full CWC model describing this example can be found at: <http://www.di.unito.it/~troina/cmc13/metapopulation.cwc>.

In ecology, using r to represent the pro-capita growth rate of a population and K the *carrying capacity* of the hosting environment,¹⁰ r/K selection theory [38]

⁹ In the remaining examples we will omit a detailed time description.

¹⁰ I.e., the population size at equilibrium.

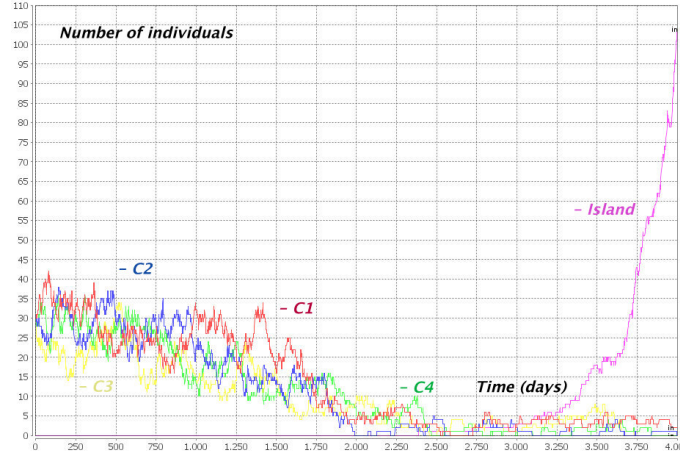


Fig. 1. Metapopulation dynamics.

describes a selective pressure driving populations evolution through the *logistic model* [47]:

$$\frac{dN}{dt} = r \cdot N \cdot \left(1 - \frac{N}{K}\right)$$

where N represents the number of individuals in the population.

CWC Modelling 3 (Logistic Model) *The logistic model with growth rate r and carrying capacity K , for an environment modelled by a compartment with label ℓ , can be encoded within CWC using two stochastic rewrite rules describing (i) a reproduction event for a single individual at the given rate and (ii) a death event modelled by a fight between two individuals at a rate that is inversely proportional to the carrying capacity:*

$$\begin{aligned} \ell : a &\xrightarrow{r} a \ a \\ \ell : a \ a &\xrightarrow{\frac{2 \cdot r}{K-1}} a \end{aligned}$$

If N is the number of individuals of species a , the number of possible reactants for the first rule is N and the number of possible reactants for the second rule is, in the exact stochastic model, $\binom{N}{2} = \frac{N \cdot (N-1)}{2}$, i.e. the number of distinct pairs of individuals of species a . Multiplying this values by the respective rates we get the propensities of the two rules and can compute the value of N when the equilibrium is reached (i.e., when the propensities of the two rules are equal): $r \cdot N = \frac{2 \cdot r}{K-1} \cdot \frac{N \cdot (N-1)}{2}$, that is when $N = 0$ or $N = K$.

For a given species, this model allows to describe different growth rates and carrying capacities in different ecological regions. Identifying a CWC compartment type (through its label) with an ecological region, we can define rules describing the growth rate and carrying capacity for each region of interest.

Species showing a high growth rate are selected by the r factor, they usually exploit low-crowded environments and produce many offspring, each of which has a relatively low probability of surviving to adulthood. By contrast, K -selected species adapt to densities close to the carrying capacity, tend to strongly compete in high-crowded environments and produce fewer offspring, each of which has a relatively high probability of surviving to adulthood.

Example 2. There is little, or no advantage at all, in evolving traits that permit successful competition with other organisms in an environment that is very likely to change rapidly, often in disruptive ways. Unstable environments thus favour species that reproduce quickly (r -selected species). Stable environments, by contrast, favour the ability to compete successfully for limited resources (K -selected species). We consider individuals of two species, a and b . Individuals of species a are modelled with an higher growth rate with respect to individuals of species b ($r_a > r_b$). Carrying capacity for species a is, instead, lower than the carrying capacity for species b ($K_a < K_b$). The following CWC rules describe the r/K selection model for $r_a = 5$, $r_b = 0.00125$, $K_a = 100$ and $K_b = 1000$:

$$\begin{array}{ll} \ell : a \xrightarrow{5} a & \ell : b \xrightarrow{0.00125} b \\ \ell : a \xrightarrow{0.1} a & \ell : b \xrightarrow{0.0000025} b \end{array}$$

We might consider a disruptive event occurring on average every 4000 years with the rule:

$$\top : (x \rfloor X)^\ell \xrightarrow{0.00025} (x \rfloor a \ b)^\ell$$

devastating the whole content of the compartment (modelled with the variable X) and just leaving one individual of each species. In Figure 2 we show a 10000 years simulation for an initial system containing just one individual for each species. Notice how individuals of species b are disadvantaged with respect to individuals of species a who reach the carrying capacity very soon. A curve showing the growth of individuals of species b in a stable (non disruptive) environment is also shown. The full CWC model describing this example can be found at: <http://www.di.unito.it/~troina/cmc13/rK.cwc>.

3.2 Competition and Mutualism

In ecology, *competition* is a contest for resources between organisms: animals, e.g., compete for water supplies, food, mates, and other biological resources. In the long term period, competition among individuals of the same species (*intraspecific competition*) and among individuals of different species (*interspecific competition*) operates as a driving force of adaptation, and, eventually, by natural selection, of evolution. Competition, reducing the fitness of the individuals involved,¹¹ has a great potential in altering the structure of populations,

¹¹ By fitness it is intended the ability of surviving and reproducing. A reduction in the fitness of an individual implies a reduction in the reproductive output. On the opposite side, a fitness benefit implies an improvement in the reproductive output.

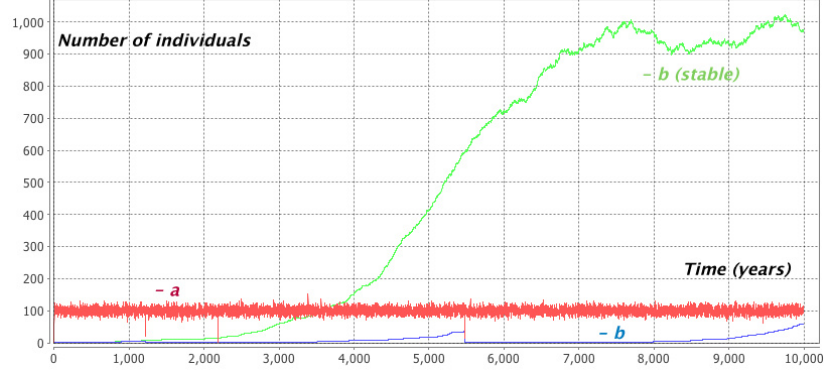


Fig. 2. r/K selection in a disruptive environment.

communities and the evolution of interacting species. It results in the ultimate survival, and dominance, of the best suited variants of species: species less suited to compete for resources either adapt or die out. We already depicted a form of competition in the context of the logistic model, where individuals of the same species compete for vital space (limited by the carrying capacity K).

Quite an apposite force is *mutualism*, contest in which organisms of different species biologically interact in a relationship where each of the individuals involved obtain a fitness benefit. Similar interactions between individuals of the same species are known as *co-operation*. Mutualism belongs to the category of symbiotic relationships, including also *commensalism* (in which one species benefits and the other is neutral, i.e. has no harm nor benefits) and *parasitism* (in which one species benefits at the expense of the other).

The general model for competition and mutualism between two species a and b is defined by the following equations [44]:

$$\begin{aligned}\frac{dN_a}{dt} &= \frac{r_a \cdot N_a}{K_a} \cdot (K_a - N_a + \alpha_{ab} \cdot N_b) \\ \frac{dN_b}{dt} &= \frac{r_b \cdot N_b}{K_b} \cdot (K_b - N_b + \alpha_{ba} \cdot N_a)\end{aligned}$$

where the r and K factors model the growth rates and the carrying capacities for the two species, and the α coefficients describe the nature of the relationship between the two species: if α_{ij} is negative, species N_j has negative effects on species N_i (i.e., by competing or preying it), if α_{ij} is positive, species N_j has positive effects on species N_i (i.e., through some kind of mutualistic interaction).

The logistic model, already discussed, is included in the differential equations above. Here we abstract away from it and just focus on the components which describe the effects of competition and mutualism we are now interested in.

CWC Modelling 4 (Competition and Mutualism) *For a compartment of type ℓ , we can encode within CWC the model about competition and mutualism*

for individuals of two species a and b using the following stochastic rewrite rules:

$$\ell : a \ b \xrightarrow{f_a \cdot |\alpha_{ab}|} \begin{cases} a \ a \ b & \text{if } \alpha_{ab} > 0 \\ b & \text{if } \alpha_{ab} < 0 \end{cases} \quad \ell : a \ b \xrightarrow{f_b \cdot |\alpha_{ba}|} \begin{cases} a \ b \ b & \text{if } \alpha_{ba} > 0 \\ a & \text{if } \alpha_{ba} < 0 \end{cases}$$

where $f_i = \frac{r_i}{K_i}$ is obtained from the usual growth rate and carrying capacity. The α coefficients are put in absolute value to compute the rate of the rule, their signs affect the right hand part of the rewrite rule.

Example 3. Mutualism has driven the evolution of much of the biological diversity we see today, such as flower forms (important to attract mutualistic pollinators) and co-evolution between groups of species [45]. We consider two different species of pollinators, a and b , and two different species of angiosperms (flowering plants), c and d . The two pollinators compete between each other, and so do the angiosperms. Both species of pollinators have a mutualistic relation with both angiosperms, even if a slightly prefers c and b slightly prefers d . For each of the species involved we consider the rules for the logistic model and for each pair of species we consider the rules for competition and mutualism. The parameters used for this model are in Table 1. So, for example, the mutualistic relations between a and c are expressed by the following CWC rules

$$\top : a \ c \xrightarrow{\frac{r_a}{K_a} \cdot \alpha_{ac}} a \ a \ c \quad \top : a \ c \xrightarrow{\frac{r_c}{K_c} \cdot \alpha_{ca}} a \ c \ c$$

Figure 3 shows a simulation obtained starting from a system with 100 individuals of species a and b and 20 individuals of species c and d . Note the initially balanced competition between pollinators a and b . This random fluctuations are resolved by the “long run” competition between the angiosperms c and d : when d predominates over c it starts favouring the pollinator b that now can win its own competition with pollinator a . The model is completely symmetrical: in other runs, a faster casual predominance of a pollinator may lead the evolution of its preferred angiosperm. The CWC model describing this example can be found at: <http://www.di.unito.it/~troina/cmc13/compmutu.cwc>.

Species (i)	r_i	K_i	α_{ai}	α_{bi}	α_{ci}	α_{di}
a	0.2	1000	•	-1	+0.03	+0.01
b	0.2	1000	-1	•	+0.01	+0.03
c	0.0002	200	+0.25	+0.1	•	-6
d	0.0002	200	+0.1	+0.25	-6	•

Table 1. Parameters for the model of competition and mutualism.

3.3 Trophic Networks

A *food web* is a network mapping different species according to their alimentary habits. The edges of the network, called *trophic links*, depict the feeding pathways

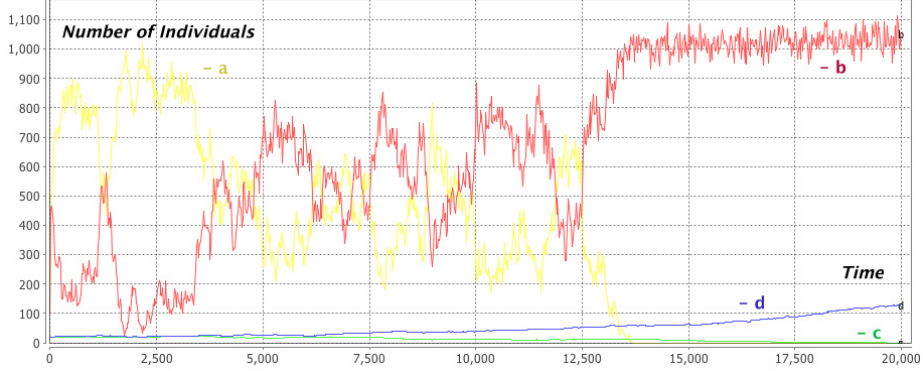


Fig. 3. Competition and Mutualism.

(“who eats who”) in an ecological community [30]. At the base of the food web there are autotroph species¹², also called basal species. A *food chain* is a linear feeding pathway that links monophagous consumers (with only one exiting trophic link) from a top consumer, usually a larger predator, to a basal species. The length of a chain is given by the number of links between the top consumer and the base of the web. The influence that the elements of a food web have on each other determine important features of an ecosystem like the presence of strong interactors (or *keystone species*), the total number of species, and the structure, functionality and stability of the ecological community.

To model quantitatively a trophic link between species a and b (i.e., a particular kind of competition) we might use Lotka-Volterra equations [48]:

$$\begin{aligned}\frac{dN_b}{dt} &= N_b \cdot (r_b - \alpha \cdot N_a) \\ \frac{dN_a}{dt} &= N_a \cdot (\beta \cdot N_b - d)\end{aligned}$$

where N_a and N_b are the numbers of predators and preys, respectively, r_b is the rate for prey growth, α is the prey mortality rate for per-capita predation, β models the efficiency of conversion from prey to predator and d is the mortality rate for predators.

CWC Modelling 5 (Trophic Links) *Within a compartment of type ℓ , given a predation mortality α and conversion from prey to predator β , we can encode in CWC a trophic link between individuals of species a (predator) and b (prey) by the following rules:*

$$\begin{aligned}\ell : a \ b &\xrightarrow{\alpha} a \\ \ell : a \ b &\xrightarrow{\beta} a \ a \ b\end{aligned}$$

¹² Self-feeding: able to produce complex organic compounds from simple inorganic molecules and light (by *photosynthesis*) or inorganic chemical reactions (*chemosynthesis*).

Here we omitted the rules for the prey exponential growth (absent predators) and predators exponential death (absent preys). These factors are present in the Lotka-Volterra model between two species, but could be substituted by the effects of other trophic links within the food web. In a more general scenario, a trophic link between species a and b could be expressed condensing the two rules within the single rule:

$$\ell : a \ b \xrightarrow{\gamma} a \ a$$

with a rate γ modelling both the prey mortality rate and the predator conversion factor.

Example 4. Trophic cascades occur when predators in a food web suppress the abundance of their prey, thus limiting the predation of the next lower trophic level. For example, an herbivore species could be considered in an intermediate trophic level between a basal species and an higher predator. Trophic cascades are important for understanding the effects of removing top predators from food webs, as humans have done in many ecosystems through hunting or fishing activities. We consider a three-level food chain between species a , b and c . The basal species a reproduces with the logistic model, the intermediate species b feeds on a , species c predaes species b :

$$\ell : a \xrightarrow{0.4} a \ a \quad \ell : a \ a \xrightarrow{0.0002} a \quad \ell : a \ b \xrightarrow{0.0004} b \ b \quad \ell : b \ c \xrightarrow{0.0008} c \ c$$

Individuals of species c die naturally, until an hunting species enters the ecosystem. At a rate lower than predation, b may also die naturally (absent predator). An atom h may enter the ecosystem and start hunting individuals of species c :

$$\ell : c \xrightarrow{0.52} \bullet \quad \ell : b \xrightarrow{0.03} \bullet \quad \top : h \ (x \mid X) \xrightarrow{\ell \ 0.003} (x \mid X \ h)^\ell \quad \ell : h \ c \xrightarrow{0.5} h$$

Figure 4 shows a simulation for the initial term $h \ (\bullet \mid 1000 * a \ 100 * b \ 10 * c)^\ell$. When the hunting activity starts, by removing the top predator, a top-down cascade destroys the whole community. The CWC model describing this example can be found at: <http://www.di.unito.it/~troina/cmc13/trophic.cwc>.

4 An application: *Croton wagneri* and Climate Change

Dry ecosystems are characterised by the presence of discontinuous vegetation that may reflect less than 60% of the available landscape. The main pattern in arid ecosystems is a vegetation mosaic composed of patches and clear sites. In [31] about 1300 different species belonging to the dry ecosystems in Northwest South America have been identified.

For this study we focused on the species *Croton wagneri* Müll. Arg., belonging to the Euphorbiaceae family. This species, particularly widespread in tropical regions, can be identified by the combination of latex, alternate simple leaves, a pair of glands at the apex of the petiole, and the presence of stipules. *C. wagneri* is the dominant endemic shrub in the dry scrub of Ecuador and has been listed as Near Threatened (NT) in the Red Book of Endemic Plants of Ecuador [46]. This kind of shrub could be considered as a nurse species¹³ and is particularly

¹³ A nurse plant is one with an established canopy, beneath which germination and survival are more likely due to increased shade, soil moisture, and nutrients.

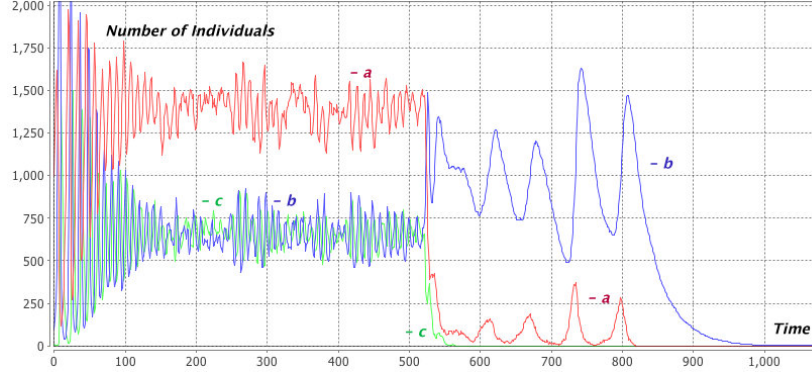


Fig. 4. A Trophic Cascade.

important for its ability to maintain the physical structure of the landscape and for its contribution to the functioning of the ecosystem (observing a marked mosaic pattern of patches having a relatively high biomass dispersed in a matrix of poor soil vegetation) [33].

The study site is located in a dry scrub in the south of Ecuador ($03^{\circ}58'29''$ S, $01^{\circ}25'22''$ W) near the Catamayo Valley, with altitude ranging from 1400m to 1900m over the sea level. Floristically, in this site we can find typical species of xerophytic areas (about 107 different species and 41 botanical families). The seasonality of the area directly affects the species richness: about the 50% of the species reported in the study site emerge only in the rainy season. Most species are shrubs (including *C. wagneri*) although there are at least 12 species of trees with widely scattered individuals, at least 50% of the species are herbs. The average temperature is 20° C with an annual rainfall around 600 mm, the most of the precipitation occurs between December and March. Generally, this area is composed by clay, rocky and sandy soils [1].

In the area, 16 plots have been installed along four levels of altitude gradients (1400m, 1550m, 1700m and 1900m): two 30mx30m plots per gradient in plane terrain and two 30mx30m plots per gradient in a slope surface (with slope greater than 10°). The data collection survey consisted in enumerating all of the *C. wagneri* shrubs in the 16 plots: the spatial location of each individual was registered using a digital laser hypsometer. Additionally, plant heights were measured directly for each individual and the crown areas were calculated according to the method in [42]. Weather stations collect data about temperatures and rainfall for each altitude gradient. An extract of data collected from the field can be found at: http://www.di.unito.it/~troina/cmc13/croton_data_extract.xlsx. This data show a morphological response of the shrub to two factors: temperature and terrain slope. A decrease of the plant height is observed at lower temperatures (corresponding to higher altitude gradients), or at higher slopes.

4.1 The CWC model

A simulation plot is modelled by a compartment with label P . Atoms g , representing the plot gradient (one g for each metre of altitude over the level of the sea), describe an abiotic factor put in the compartment wrap.

According to the temperature data collected by the weather stations we correlate the mean temperatures in the different plots with their respective gradients. In the content of a simulation plot, atoms t , representing 1°C each, model its temperature. Remember that, in this case, the higher the gradient, the lower the temperature. Thus, we model a constant increase of temperature within the simulation plot compartment, controlled by the gradient elements g on its wrap:

$$\top : (x \mid X)^P \xrightarrow{1} (x \mid t \mid X)^P \quad \top : (g \mid x \mid t \mid X)^P \xrightarrow{0.000024} (g \mid x \mid X)^P$$

Atoms i are also contained within compartments of type P , representing the complementary angle of the plot's slope (e.g., $90 * i$ for a plane plot or $66 * i$ for a 24° slope).

We model *C. wagneri* as a CWC compartment with label c . Its observed trait, namely the plant height, is specified by atomic elements h (representing one mm each) on the compartment wrap.

To model the shrub heights distribution within a parcel, we consider the plant in two different states: a “young” and an “adult” state. Atomic elements y and a are exclusively, and uniquely, present within the plant compartment in such a way that the shrub height increases only when the shrub is in the young state (y in its content). The following rules describe (i) the passage of the plant from y to a state with a rate corresponding to a 1 year average value, and (ii) the growth of the plant, affected by temperature and slope, with a rate estimated to fit the field collected data:

$$c : y \xrightarrow{0.00274} a \quad P : t \mid i \mid (x \mid y \mid X)^c \xrightarrow{0.000718} t \mid i \mid (x \mid h \mid y \mid X)^c$$

4.2 Simulation results

Now we have a model to describe the distribution of *C. wagneri* height using as parameters the plot's gradient ($n * g$) and slope ($m * i$). Since we do not model explicitly interactions that might occur between *C. wagneri* individuals, we consider plots containing a single shrub. Carrying on multiple simulations, through the two phase model of the plant growth, after 1500 time units (here represented as days), we get a snapshot of the distribution of the shrubs heights within a parcel. The CWC model describing this application can be found at: <http://www.di.unito.it/~troina/cmc13/croton.cwc>.

Each of the graphs in Figure 5 is obtained by plotting the height deviation of 100 simulations with initial term $(n * g \mid m * i \mid (\bullet \mid y)^c)^P$. The simulations in Figures 5 (a) and (c) reflect the conditions of real plots and the results give a good approximation of the real distribution of plant heights. Figures 5 (b) and (d) are produced considering an higher slope than the ones on the real plots

from where the data has been collected. These simulation results can be used for further validation of the model by collecting data on new plots corresponding to the parameters of the simulation.

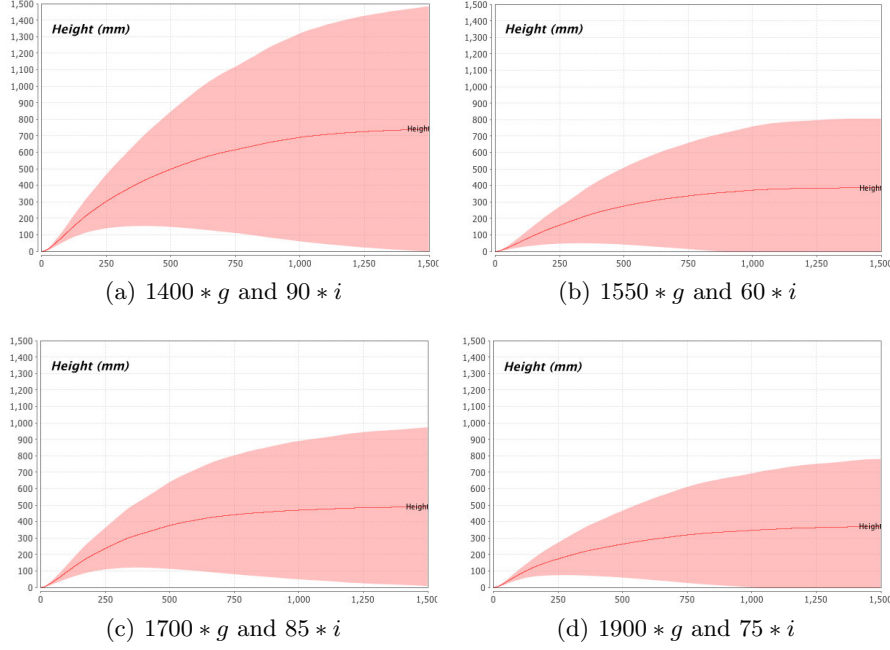


Fig. 5. Deviation of the height of *Croton wagneri* for 100 simulations.

If we already trust the validity of our model, we can remove the correlation between the gradient and the temperature, and directly express the latter. Predictions can thus be made about the shrub height at different temperatures, and how it could adapt to global climate change. Figure 6 shows two possible distributions of the shrub height at lower temperatures (given it will actually survive these more extreme conditions and follow the same trend).

5 Conclusions and Related Works

The long-term goal of Computational Ecology is the development of methods to predict the response of ecosystems to changes in their physical, chemical and biological components. Computational models, and their ability to understand and predict the biological world, could be used to express the mechanisms governing the structure and function of natural populations, communities, and ecosystems. Until recent times, there was insufficient computational power to run stochas-

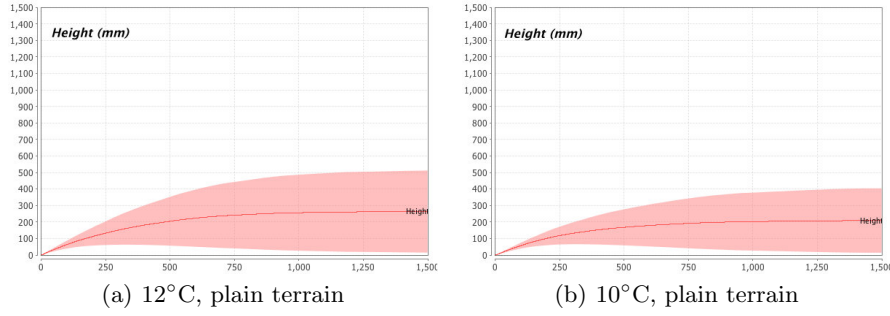


Fig. 6. Deviation of the height of *Croton wagneri* for 100 simulations.

tic, individually-based, spatially explicit models. Today, however, some of these techniques could be investigated [37].

Calculi developed to describe process interactions in a compartmentalised setting are well suited for the description and analysis of the evolution of ecological systems. The topology of the ecosystem can be directly encoded within the nested structure of the compartments. These calculi can be used to represent structured natural processes in a greater detail, when compared to purely numerical analysis. As an example, food webs can give rise to combinatorial interactions resulting in the formation of complex systems with emergent properties (as signalling pathways do in cellular biology), and, in some cases, giving rise to chaotic behaviour.

As a final remark about ecological modelling with a framework based on stochastic rewrite rules, we underline an important compositional feature. How can we test an hypothetical scenario where a grazing species is introduced in the model of our case study? A possibility could be to represent the grazing species with a new CWC atom (e.g. s) and then just add the new competitive rules to the previously validated model (e.g. the rule $P : s (h \ x \mid X)^c \mapsto s (x \mid X)^c$). Changing in the same sense a model based on ordinary differential equations would, instead, result in a complete new model were all previous equations should be rewritten.

5.1 Related works

As P-Systems [40,41] and the Calculus of Looping Sequences (CLS, for short) [11], the Calculus of Wrapped Compartments is a framework modelling topological compartmentalisation inspired by biological membranes, and with a semantics given in terms of rewrite rules.

CWC has been developed as a simplification of CLS, focusing on stochastic multiset rewriting. The main difference between CWC and CLS consists in the exclusion of the sequence operator, that constructs ordered strings out of the atomic elements of the calculus. While the two calculi keep the same expressiveness, some differences arise on the way systems are described. On the one

hand, the Calculus of Looping Sequences allows to define ordered sequences in a more succinct way (for examples when describing sequences of genes in DNA or sequences of amino acids in proteins).¹⁴ On the other hand, CWC reflects in a more realistic way the fluid mosaic model of the lipid bilayer (for example in the case of cellular membrane description, where proteins are free to float), and, the addition of compartment labels allows to characterise the properties peculiar to given classes of compartments. Ultimately, focusing on multisets and avoiding to deal explicitly with ordered sequences (and, thus, variables for sequences) strongly simplifies the pattern matching procedure in the development of a simulation tool.

The Calculus of Looping Sequences has been extended with type systems in [6,28,29,8,16]. As an application to ecology, stochastic CLS (see [7]) is used in [12] to model population dynamics.

P-Systems have been proposed as a computational model inspired by biological structures. They are defined as a nesting of membranes in which multisets of objects can react according to pre defined rewrite rules. Maximal-parallelism is the key feature of P-Systems: at each evolution step all rewrite rules, in all membranes, are applied as many times as possible. Such a feature makes P-Systems a very powerful computational model and a versatile instrument to evaluate expressiveness of languages. However, it is not practical to describe stochastic systems with a maximally-parallel evolution: exact stochastic simulations based on race conditions model systems evolutions as a sequence of successive steps, each of which with a particular duration modelled by an exponential probability distribution.

There is a large body of literature about applications of P-Systems to ecological modelling. In [20,21,22], P-Systems are enriched with a probabilistic semantics to model different ecological systems in the Catalan Pyrenees. Rules could still be applied in a parallel fashion since reduction durations are not explicitly taken into account. In [13,14,15], P-Systems are enriched with a stochastic semantics and used to model metapopulation dynamics. The addition of *mute rules* allows to keep a form of parallelism reducing the maximal consumption of objects.

While all these calculi allow to manage systems topology through nesting and compartmentalisation, explicit spatial models are able to depict more precise localities and *ecological niches*, describing how organisms or populations respond to the distribution of resources and competitors [35]. The spatial extensions of CWC [17], CLS [9] and P-Systems [10] could be used to express this kind of analysis allowing to deal with spatial coordinates.

References

1. Aguirre, Z., Kvist, P., Sánchez, O.: Floristic composition and conservation status of the dry forests in ecuador. *Lyonia* 8(2) (2005)

¹⁴ An ordered sequence can be expressed in CWC as a series of nested compartments, ordered from the outermost compartment to the innermost one.

2. Aldinucci, M., Coppo, M., Damiani, F., Drocco, M., Giovannetti, E., Grassi, E., Sciacca, E., Spinella, S., Troina, A.: CWC Simulator. Dipartimento di Informatica, Università di Torino (2010), <http://cwcsimulator.sourceforge.net/>
3. Aldinucci, M., Coppo, M., Damiani, F., Drocco, M., Sciacca, E., Spinella, S., Torquati, M., Troina, A.: On parallelizing on-line statistics for stochastic biological simulations. In: HiBB'11. Lecture Notes in Computer Science, vol. 7156, pp. 3–12. Springer (2011)
4. Aldinucci, M., Coppo, M., Damiani, F., Drocco, M., Torquati, M., Troina, A.: On designing multicore-aware simulators for biological systems. In: Proc. of Intl. Euromicro PDP 2011: Parallel Distributed and network-based Processing. pp. 318–325. IEEE Computer Society (2011)
5. Aldinucci, M., Torquati, M.: FastFlow website. FastFlow (Oct 2009), <http://mc-fastflow.sourceforge.net/>
6. Aman, B., Dezani-Ciancaglini, M., Troina, A.: Type disciplines for analysing biologically relevant properties. *Electr. Notes Theor. Comput. Sci.* 227, 97–111 (2009)
7. Barbuti, R., Maggiolo-Schettini, A., Milazzo, P., Tiberi, P., Troina, A.: Stochastic calculus of looping sequences for the modelling and simulation of cellular pathways. *Transactions on Computational Systems Biology IX*, 86–113 (2008)
8. Barbuti, R., Dezani-Ciancaglini, M., Maggiolo-Schettini, A., Milazzo, P., Troina, A.: A formalism for the description of protein interaction. *Fundam. Inform.* 103(1–4), 1–29 (2010)
9. Barbuti, R., Maggiolo-Schettini, A., Milazzo, P., Pardini, G.: Spatial calculus of looping sequences. *Theoretical Computer Science* 412(43), 5976 – 6001 (2011)
10. Barbuti, R., Maggiolo-Schettini, A., Milazzo, P., Pardini, G., Tesi, L.: Spatial p systems. *Natural Computing* 10(1), 3–16 (2011)
11. Barbuti, R., Maggiolo-Schettini, A., Milazzo, P., Troina, A.: The calculus of looping sequences for modeling biological membranes. In: Workshop on Membrane Computing. Lecture Notes in Computer Science, vol. 4860, pp. 54–76. Springer (2007)
12. Basuki, T.A., Cerone, A., Barbuti, R., Maggiolo-Schettini, A., Milazzo, P., Rossi, E.: Modelling the dynamics of an aedes albopictus population. In: AMCA-POP. vol. 33, pp. 18–36. EPTCS (2010)
13. Besozzi, D., Cazzaniga, P., Pescini, D., Mauri, G.: Seasonal variance in p system models for metapopulations. *Progress in Natural Science* 17(4), 392–400 (2007)
14. Besozzi, D., Cazzaniga, P., Pescini, D., Mauri, G.: Modelling metapopulations with stochastic membrane systems. *Biosystems* 91(3), 499–514 (2008)
15. Besozzi, D., Cazzaniga, P., Pescini, D., Mauri, G.: An analysis on the influence of network topologies on local and global dynamics of metapopulation systems. In: AMCA-POP. vol. 33, pp. 1–17. EPTCS (2010)
16. Bioglio, L., Dezani-Ciancaglini, M., Giannini, P., Troina, A.: Typed stochastic semantics for the calculus of looping sequences. *Theor. Comp. Sci.* 431, 165–180
17. Bioglio, L., Calcagno, C., Coppo, M., Damiani, F., Sciacca, E., Spinella, S., Troina, A.: A spatial calculus of wrapped compartments. In: MeCBIC. vol. abs/1108.3426. CoRR (2011)
18. Bolker, B.: *Ecological models and data* in R. Princeton University Press (2008)
19. Calcagno, C., Coppo, M., Damiani, F., Drocco, M., Sciacca, E., Spinella, S., Troina, A.: Modelling spatial interactions in the arbuscular mycorrhizal symbiosis using the calculus of wrapped compartments. In: CompMod 2011. vol. 67, pp. 3–18. EPTCS (2011)
20. Cardona, M., Colomer, M.A., Margalida, A., Palau, A., Pérez-Hurtado, I., Pérez-Jiménez, M.J., Sanuy, D.: A computational modeling for real ecosystems based on p systems. *Natural Computing* 10(1), 39–53 (2011)

21. Cardona, M., Colomer, M.A., Margalida, A., Pérez-Hurtado, I., Pérez-Jiménez, M.J., Sanuy, D.: A p system based model of an ecosystem of some scavenger birds. In: 10th Workshop on Membrane Computing. Lecture Notes in Computer Science, vol. 5957, pp. 182–195. Springer (2010)
22. Cardona, M., Colomer, M.A., Pérez-Jiménez, M.J., Sanuy, D., Margalida, A.: Modeling ecosystems using p systems: The bearded vulture, a case study. In: Membrane Computing, pp. 137–156. Springer (2009)
23. Caswell, H.: Matrix population models: Construction, analysis and interpretation, 2nd Edition. Sinauer Associates, Sunderland, Massachusetts. (2001)
24. Coppo, M., Damiani, F., Drocco, M., Grassi, E., Guether, M., Troina, A.: Modelling ammonium transporters in arbuscular mycorrhiza symbiosis. Transactions on Computational Systems Biology XIII, 85–109 (2011)
25. Coppo, M., Damiani, F., Drocco, M., Grassi, E., Sciacca, E., Spinella, S., Troina, A.: Simulation techniques for the calculus of wrapped compartments. Theor. Comp. Sci. 431, 75–95
26. Coppo, M., Damiani, F., Drocco, M., Grassi, E., Sciacca, E., Spinella, S., Troina, A.: Hybrid calculus of wrapped compartments. In: MeCBIC. vol. 40, pp. 103–121. EPTCS (2010)
27. Coppo, M., Damiani, F., Drocco, M., Grassi, E., Troina, A.: Stochastic Calculus of Wrapped Compartments. In: QAPL. vol. 28, pp. 82–98. EPTCS (2010)
28. Dezani-Ciancaglini, M., Giannini, P., Troina, A.: A type system for required/excluded elements in CLS. In: DCM’09. vol. 9, pp. 38–48. EPTCS (2009)
29. Dezani-Ciancaglini, M., Giannini, P., Troina, A.: A type system for a stochastic cls. In: MeCBIC’09. vol. 11, pp. 91–105. EPTCS (2009)
30. Elton, C.: Animal Ecology. Sidgwick and Jackson (1927)
31. Gentry, A.: A Field Guide to the Families and Genera of Woody Plants of North-west South America: (Colombia, Ecuador, Peru): with supplementary notes on herbaceous taxa. Washington DC, Conservation International (1993)
32. Gillespie, D.: Exact stochastic simulation of coupled chemical reactions. J. Phys. Chem. 81, 2340–2361 (1977)
33. Gutiérrez, J.: Importancia de los Arbustos Leñosos en los Ecosistemas de la IV Región, Libro Rojo de la Flora Nativa y de los Sitios Prioritarios para su Conservación: Región de Coquimbo, vol. 16. Ediciones Universidad de La Serena, Chile (2001)
34. Levins, R.: Some demographic and genetic consequences of environmental heterogeneity for biological control. Bulletin of the Entomological Society of America 15, 237–240 (1969)
35. Lomolino, M.V., Brown, J.W.: Biogeography. Sunderland, Mass: Sinauer Associates. (1998)
36. Oury, N., Plotkin, G.: Multi-level modelling via stochastic multi-level multiset rewriting. Mathematical Structures in Computer Science (2012)
37. Petrovskii, S., Petrovskaya, N.: Computational ecology as an emerging science. Interface Focus 2(2), 241–254 (2012)
38. Pianka, E.: On r and k selection. American Naturalist 104(940), 592–597 (1970)
39. Pielou, E.: Mathematical ecology. Wiley (1977)
40. Păun, G.: Computing with membranes. Journal of Computer and System Sciences 61(1), 108–143 (2000)
41. Păun, G.: Membrane Computing. An Introduction. Springer (2002)
42. Shiponeni, N., Allsopp, N., Carrick, P., Hoffman, M.: Competitive interactions between grass and succulent shrubs at the ecotone between an arid grassland and succulent shrubland in the karoo. Plant. Ecol. 212(5), 795–808 (2011)

43. Sugihara, G., May, R.: Nonlinear forecasting as a way of distinguishing chaos from measurement error in time series. *Nature* 344(6268), 734–741 (1990)
44. Takeuchi, Y.: Cooperative systems theory and global stability of diffusion models. *Acta Applicandae Mathematicae* 14, 49–57 (1989)
45. Thompson, J.: The geographic mosaic of coevolution. University of Chicago Press (2005)
46. Valencia, R., Pitman, N., León-Yáñez, S., Jorgensen, P.: Libro Rojo de las Plantas Endémicas del Ecuador. Herbario QCA, Pontificia Universidad Católica del Ecuador, Quito (2000)
47. Verhulst, P.: Notice sur la loi que la population poursuit dans son accroissement. *Corresp. Math. Phys.* 10, 113–121 (1838)
48. Volterra, V.: Variazioni e fluttuazioni del numero dindividui in specie animali conviventi. *Mem. Acad. Lincei Roma* 2, 31–113 (1926)