Dynamic Modelling Trophic Networks with Bio-PEPA

Ca'Foscari University
Department of Computer Science

MASTER THESIS

By: ALIYE KUERBAN

Supervisor: NICOLETTA COCCO

February 2016

Contents

Chapter 1

Introduction

A trophic network is a set of interconnected food chains within an ecosystem. It represents an ecological community of organisms and their feeding relations. Modelling a trophic network in a realistic way is a main goal for ecologists. There exist tools which are used for modelling and analysing trophic networks, but they give either a static model of the system at steady state or they reproduce the dynamics described through specific Ordinary Differential Equations (ODEs).

Modelling and computer simulation are vitally important in system biology for integrating knowledge and experimental data and making testable predictions about the behaviour of biological systems. In recent years there has been proposals to use process algebras to model and analyse biological networks. Process algebras are suitable for describing such systems in a precise way. They have several attractive features that are not available in other modelling frameworks. A process algebra can model a system as the interaction of its subsystems and it can build up complex models from detailed components. Different kinds of analysis can be performed on a process algebra model. Bio-PEPA is a process algebra for modelling and analysing biological systems, it is an extension of PEPA, a process algebra originally defined for performance analysis of systems with concurrent behaviour. Bio-PEPA is able to handle some features of biological networks such as stoichiometry coefficients and different kinds of kinetic laws.

This thesis focuses on modelling trophic networks by using Bio-PEPA and on their continuous dynamic analysis. Bio-PEPA combines several simulation and analysis techniques to describe and analyse a trophic network in a more accurate and realistic way. Discrete stochastic and continuous deterministic models can be automatically generated from a Bio-PEPA model. We recall Moscardo's previous works [10] on modelling trophic networks in Bio-PEPA and applying to them the discrete stochastic analysis. We extend his work by considering a continuous deterministic model. This allows us to see the variation of the species in a trophic network in their continuous dynamics and to discover the effects and

relationships between these species. At first we focus on a toy case of a simple trophic network called Wood, which was analysed by Moscardo in his master thesis. Among his experiments we consider Wood7, which gives the best result in terms of stochastic modelling. We analyse this model in continuous deterministic way and we do some modifications to see the effects of functional rates in this model. Then we apply the continuous dynamic analysis and we model the Venice Lagoon trophic network with Bio-PEPA. The model is based on the toy case Wood7 proposed by Moscardo. Besides, we discuss several experiments on its continuous dynamics.

The structure of thesis is as follows. Chapter 2 introduces the process algebra Bio-PEPA. Since it is an extension of PEPA, the main operators and the syntax of PEPA are also given. Then we describe the Bio-PEPA process algebra in detail, namely its system representation, derivation of rates and some auxiliary definitions of Bio-PEPA. We introduce the main analyses that can be applied in Bio-PEPA and in particular the relation between the two analysis techniques ODEs and CTMC. The software tools for PEPA and Bio-PEPA are also introduced. In Chapter 3 we focus on trophic networks. We give the basic definitions related to a trophic network and describe its representation as a matrix and in a graphical way. We introduce the static properties of a trophic network and the tools for analysing trophic networks such as Ecopath with Ecosim (EwE). In Chapter 4 we describe the Venice Lagoon trophic network which is the case that we want to model. We introduce the model structure, the main mass balance equations in the model of the Venice Lagoon and describe each individual term of the equations in detail. In Chapter 5 we discuss how to model trophic networks with Bio-PEPA. We recall the previous work of Moscardo and took his best model Wood7 to do our continuous dynamic analysis on this model. Then we modify the model in order to analyse the effects of the range of the functional rates of the species in a trophic network. In Chapter 6 we model a real case of trophic network with Bio-PEPA, namely the Venice Lagoon. We recall Moscardo's model and improve it. Then we analyse this model in a continuous way and discover some relationships among the species by doing several experiments on the continuous dynamics of the model.

Chapter 2

PEPA and Bio-PEPA

In this chapter we introduce two stochastic process algebras: PEPA and Bio-PEPA. In the first part we introduce PEPA process algebra briefly. In the second part we give the definition of a Bio-PEPA system and its components. Then some important properties useful in the analysis are described.

2.1 PEPA

The process algebra PEPA was originally defined for the performance analysis of systems with concurrent behaviour. PEPA uses powerful set of combinators to build up complex behaviour from simpler behaviour.

In this section we give an informal description of the language and the syntax. An example of PEPA model is also presented. All the information are taken from [2],[9].

2.1.1 Informal Description and Syntax

PEPA has a set of combinators in order to describe the possible reactions in a system. In PEPA the basic elements of the language are components and activities, these correspond to states and transitions in the underlying stochastic model. It allows to study quantitative properties of a system of interacting processes such as response time, utilization, throughput and it also permits to find the deadlocks if they exist.

PEPA represents systems with a composition of *components*, these components composed by *actions* and *rates*. In PEPA an action is assumed to have a duration. Thus the expression $(\alpha, r).P$ denotes a component with an action α, at rate r, then evolve into a component P. Here $\alpha \in \mathcal{A}$, where $\mathcal{A}$ is the *set of action types* and $P \in \mathcal{C}$, where $\mathcal{C}$ is the *set of components*.

Since an exponential distribution is uniquely determined by its parameter, the duration of an activity, which is an exponentially distributed random variable, may be represented by

a single real number parameter. This parameter is called the *activity rate*. It may be any positive real number, or the distinguished symbol $\top$, which should be read as *unspecified* rate. When a system is carrying out some action the identity of which is unknown or unimportant, here is a distinguished action type, τ, which can be regarded as the *unknown type*.

In this thesis we adopt the following conventions:

- Components are denoted by large letters; for example P, S, C or System.
- Action types are denoted by small greek letters; for example α, β, γ, or by names which start with a small roman letter, such as task, serve or use.
- Activity rates are denoted by single roman letters taken from the end of the alphabet, typically r or unknown action type $\top$.
- The character $\mathcal{L}$ is used to denote a subset of $\mathcal{A}$.

When the behaviour of the system is determined by a component P the system is said to *behave as* P. The action types which the component P may next engage in are *the current action types* of P, a set denoted $\mathcal{A}(P)$. The activities which the component P may next engage in are *the current activities* of P, a multiset denoted $\mathcal{A}ct(P)$.

Process algebras are defined by a simple syntax and semantics. The semantics is given by axioms and inference rules expressed in an operational way. A system is defined as a collection of agents which execute atomic actions. Some operators are allow for combining the agents. The main operators of PEPA are:

Prefix: $(\alpha, r).P$, it has the action of type α and an activity rate r and it subsequently behaves as P.

Choice: $P + Q$, it represents a system which may behave either as P or as Q. When P happens, then Q is discharged or vice versa. It reflects the competition between two components based on the race condition.

Hiding: $P/\mathcal{L}$, it behaves as P where activities of types within the set $\mathcal{L}$ are hidden. They appear as the unknown type $\top$ and can be regarded as an internal delay by the component.

Cooperation: $P \underset{\mathcal{L}}{\bowtie} Q$, it denotes the cooperation between the components P and Q over the cooperation set $\mathcal{L}$. The *cooperation set* $\mathcal{L}$ defines the actions on which the components must synchronize or cooperate. P and Q proceed independently with any activities not occurring in the cooperation set $\mathcal{L}$. If the set is empty, the processes work in parallel: $(P \parallel Q)$.

Constant: $C \stackrel{def}{=} P$. Constants are components whose meaning is given by a defining equation which associates to the constant C the behaviour of the component P. This is how one can assign names to components (behaviours).

The syntax for terms in PEPA is defined as follows:

$$P ::= (\alpha, r).P \mid P + Q \mid P \underset{\mathcal{L}}{\bowtie} Q \mid P/\mathcal{L} \mid C$$

The names of these operators and their intended interpretations have been already presented.

Definition 2.1.1. *A PEPA system $\mathcal{P}$ is a 4-tuple $\langle \mathcal{C}, \mathcal{A}ct, \mathcal{R}, P \rangle$ where:*

- *$\mathcal{C}$ is the set of components;*
- *$\mathcal{A}ct$ is the set of possible activities;*
- *$\mathcal{R}$ is the set of activity rates, $\mathcal{R} = \{x | x > 0; x \in \mathbb{R}\} \cup \{\top\}$; where $\top$ represents the unspecified activity rate which is the fastest possible rate.*
- *P is the sequential component describing the system.*

The rate is inversely proportional to the duration of the action, thus if an action have a small duration then it has a large rate because in a time period it happens many times. Let us suppose that we have one action with rate 2 and another action with rate 5. Then the first action has duration 1/2 and the second one has duration 1/5, thus the second action will happen more times than the first one in the same period. In PEPA, when two activities may occur in a choice composition, a race condition is considered, namely the activity with greater rate occurs before the activity with lower rate. In other words the activities with higher rate have a greater probability to be executed because they have higher priority than the activities with lower rate.

Definition 2.1.2. *The* `apparent rate` *of an action of type α in a component P denoted $r_\alpha(P)$ is the sum of the rates of all the activities of type α in $\mathcal{A}ct(P)$.*

1. $r_\alpha((\beta, r).P) = \begin{cases} r & if\ \beta = \alpha \\ 0 & if\ \beta \neq \alpha \end{cases}$
2. $r_\alpha(P + Q) = r_\alpha(P) + r_\alpha(Q)$
3. $r_\alpha(P/_{\mathcal{L}}) = \begin{cases} r_\alpha(P) & if\ \alpha \notin \mathcal{L} \\ 0 & if\ \alpha \in \mathcal{L} \end{cases}$

4. $r_\alpha(P \underset{\mathcal{L}}{\bowtie} Q) = \begin{cases} min(r_\alpha(P),\ r_\alpha(Q)) & if\ \alpha \in \mathcal{L} \\ r_\alpha(P) + r_\alpha(Q) & if\ \alpha \notin \mathcal{L} \end{cases}$

The unspecified rate $\top$ means that the component does not contribute to the execution of the activities. When two components cooperate, the *active component* (which does not include $\top$) behaves actively, the component which include $\top$ is *passive component* and behaves passively. In this cooperation the active component determines the rate. The unspecified activity rate $\top$ has the following properties:

$r < w\top$	for all $r \in \mathbb{R}$ and for all $w \in \mathbb{N}$
$w_1\top < w_2\top$	if $w_1 < w_2$ for all $w_1, w_2 \in \mathbb{N}$
$w_1\top + w_2\top = (w_1 + w_2)\top$	for all $w_1, w_2 \in \mathbb{N}$
$\dfrac{w_1\top}{w_2\top} = \dfrac{w_1}{w_2}$	for all $w_1, w_2 \in \mathbb{N}$

2.1.2 PEPA Labelled Transition System

A Labelled transition system (LTS) is an abstraction of the behaviour of a system and it consists of the graph representing the transition among its possible states. When a reaction happens in the system, it produces a transition from a state to another and such transitions are represented in the LTS of the system.Certain states may be distinguished: a start state and perhaps one or more final states.

PEPA provides a relation called *transition relation.* It represents the behaviour of a process, defined as:

$$\xrightarrow{t} \subseteq S \times S$$

where S is set of states of the process and t belongs to the set of transition labels $\mathcal{T}$ (the set of labels of action).

Definition 2.1.3. *let $\mathcal{P} = \langle \mathcal{C}, \mathcal{A}ct, \mathcal{R}, P \rangle$ be a PEPA system, a* ***labelled transition system LTS*** *for $\mathcal{P}$ is a triple $\Big(\mathcal{C}, \mathcal{A}ct, \{\xrightarrow{(\alpha,r)} | (\alpha, r) \in \mathcal{A}cT\}\Big)$.*

where the transition relation $\xrightarrow{(\alpha,r)}$ is given by the rules in figure 1.1.

When two componentS cooperate on the same activity with different rates, the system execute that activity with rate R. In the case where the apparent rate is unspecified in one component, the apparent rate will be completely determined by the other component.

Unlike other performance modelling algorithms, PEPA allows models and states to be regarded as equivalent entities, both are represented as components. There are properties, such as equivalence relation, isomorphism and bisimulation, that allows one to analyse both model-to-model and state-to-state relations.

Prefix

$$\frac{}{(\alpha, r).E \xrightarrow{(\alpha,r)} E}$$

Choice

$$\frac{E \xrightarrow{(\alpha,r)} E'}{E + F \xrightarrow{(\alpha,r)} E'} \qquad \frac{F \xrightarrow{(\alpha,r)} F'}{E + F \xrightarrow{(\alpha,r)} F'}$$

Cooperation

$$\frac{E \xrightarrow{(\alpha,r)} E'}{E \bowtie_L F \xrightarrow{(\alpha,r)} E' \bowtie_L F} (\alpha \notin L) \qquad \frac{F \xrightarrow{(\alpha,r)} F'}{E \bowtie_L F \xrightarrow{(\alpha,r)} E \bowtie_L F'} (\alpha \notin L)$$

$$\frac{E \xrightarrow{(\alpha,r_1)} E' \quad F \xrightarrow{(\alpha,r_2)} F'}{E \bowtie_L F \xrightarrow{(\alpha,R)} E' \bowtie_L F'} (\alpha \in L) \quad \text{where } R = \frac{r_1}{r_\alpha(E)} \frac{r_2}{r_\alpha(F)} \min(r_\alpha(E), r_\alpha(F))$$

Hiding

$$\frac{E \xrightarrow{(\alpha,r)} E'}{E/L \xrightarrow{(\alpha,r)} E'/L} (\alpha \notin L) \qquad \frac{E \xrightarrow{(\alpha,r)} E'}{E/L \xrightarrow{(\tau,r)} E'/L} (\alpha \in L)$$

Constant

$$\frac{E \xrightarrow{(\alpha,r)} E'}{A \xrightarrow{(\alpha,r)} E'} (A \stackrel{def}{=} E)$$

Figure 2.1: Operational Semantics of PEPA

2.1.3 Analysis techniques in PEPA and Applications to system Biology

Time series analysis can be performed in a PEPA model through two different techniques: *ODEs* and *SSA*. The former results from mapping a PEPA model to a set of Ordinary Differential Equations, the latter runs Stochastic Simulation Algorithms on a PEPA model. Moreover study specific properties of the modelled system by means of *model checking*.

In the biological field, recently PEPA has been applied to the context of signalling pathways based on reagents and based on pathways [1]. In the reagent-centric view the PEPA sequential components represent various concentration levels of the species. In the pathway-centric approach of PEPA, the processes represent sub-pathways.

PEPA offers a high level of abstraction and focuses on compositionality and on the various interactions. Even though it has proved useful in studying signalling pathways, it can

not express all the features of biosystems and biochemical networks. The main difficulties in this case are the definition of stoichiometry coefficients and the representation of kinetic laws. In PEPA the stoichiometry is not represented explicitly and the reactions are assumed to be elementary.

2.2 Bio-PEPA

Bio-PEPA is a process algebra for modelling and analysing biochemical networks, it is designed specifically for the description of biological phenomena and for their analysis through quantitative methods such as stochastic simulation and probabilistic model-checking. It is based on PEPA, and extends it in order to handle some features of biochemical networks, such as stoichiometry and the role of the species in a given reaction. Functional rates in Bio-PEPA can express general kinetic laws. Each action type represents a reaction and is associated with a functional rate.

One of the unique feature of Bio-PEPA is that it integrates several kinds of analysis techniques. Discrete, stochastic and continuous deterministic models can be automatically generated from Bio-PEPA models, thus it allows to perform approximate time-series analysis via stochastic simulation, exact small-scale Markovian analysis, or numerical integration of the underlying ordinary differential equations.

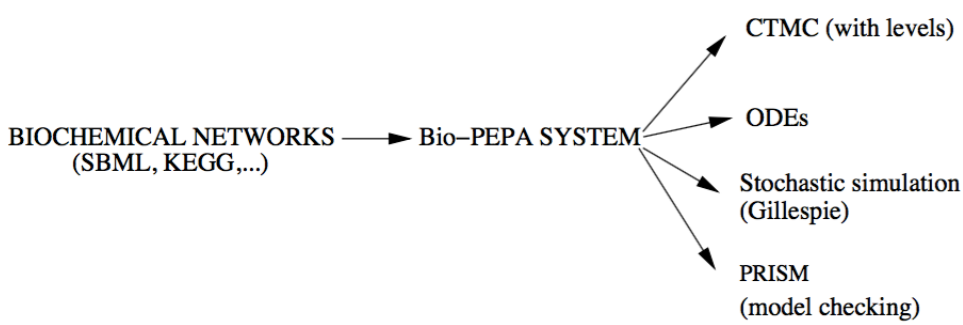


Figure 2.2: Schema of the Bio-PEPA framework [1]

These analyses can be very helpful for studying different aspects of biological models. In order to have a better understanding of the system they can be used in conjunction.

In Bio-PEPA there are species components, to represent biochemical species, and model components, to express how species components cooperate with each other. The model component contains the initial current concentration of each species. In addition to these components, a Bio-PEPA system is composed of the set of compartments, the set of functional rates, the set of constant parameters and auxiliary information for the analysis. All

the information in this section are taken from [1],[3],[8].

2.2.1 System Representation

The syntax of Bio-PEPA is designed in order to represent biological information :

$$S ::= (\alpha, k)\ op\ S \mid S + S \mid C \qquad\qquad P ::= P \underset{\mathcal{L}}{\bowtie} P \mid S(x)$$

where $op = \downarrow \mid \uparrow \mid \oplus \mid \ominus \mid \odot$.

Now we see the meaning of components.

- S is called *sequential component* (or species component) and represents a species.
- P is called a *model component*, describes the system and the interactions among components.
- (α, k) is the prefix, where $\alpha \in \mathcal{A}$ is the action type and k is *the stoichiometry coefficient* of the species in that reaction;
- C is a constant (as in PEPA).
- x is a positive real-valued parameter, usually interpreted as a concentration.
- *op* represents the role of the element in the reaction. Symbols $\downarrow \mid \uparrow \mid \oplus \mid \ominus \mid \odot$ indicate respectively a reactant, a product, an activator, an inhibitor and a generic modifier.

Since Bio-PEPA was developed based on PEPA, the choice operator, cooperation and definition of constant are unchanged. We have $\mathcal{L} \subseteq \mathcal{A}$ as in PEPA. In contrast to PEPA, the hiding operator is omitted because it is not necessary in Bio-PEPA.

In order to describe a biochemical network in Bio-PEPA, we need to define some structures like the compartments, the species, the constant parameters and the functional rates. In Bio-PEPA compartments are static (compartments are not actively involved in the reaction, they are simply containers) and reactions are irreversible reactions.

Definition 2.2.1. *Each compartment is described by " V: v unit". where:*

- *V is the compartment name,*
- *v is a positive real number expressing the compartment size,*
- *unit denotes the unit associated with the compartment size, it is optional.*

The set of compartments is denoted by $\mathcal{V}$.

When no information about compartments is available, we set a default compartment size 1 and the unit depends on the model.

Definition 2.2.2. *For each species we define the element "C : H = value_H, N = value_N, M = value_M, V = value_V, unit = value_u ", where:*

- *C is species component name,*
- *H is the step size and value_H $\in \mathbb{R}^+$,*
- *N is the maximum level and value_N $\in \mathbb{N}$,*
- *M is the maximum concentration and value_M $\in \mathbb{R}^+ \cup \{\epsilon\}$,*
- *V is the name of the enclosing compartment and value_V $\in name(\mathcal{V}) \cup \{\epsilon\}$,*
- *value_u represents the unit for concentration.*

where ϵ is the empty string. The set of all the elements described above is denoted as $\mathcal{N}$.

The elements described above are optional. If we have only the compartment name we can use the system for stochastic simulation and we can map it to ODEs, whereas if we are interested in the CTMC with levels or in model checking with PRISM, we also need the values for H and N (or, equivalently, H and M).

In order to describe the kinetics of a biological system, we can associate to each action α_i a functional rate f_{α_i}. It is possible to define a functional rate by means of simple mathematical expressions, using also constant components and parameters.

Definition 2.2.3. *The functional rates are expressed by the following grammar:*

$$f\_rate ::= f_\alpha(\overline{k}, \overline{C}) = sk \mid f_\alpha(\overline{k}) = sk2$$

$$sk ::= int \mid float \mid name \mid sk{+}sk \mid sk{\times}sk \mid sk/sk \mid sk{-}sk \mid sk^{sk} \mid exp(x) \mid log(sk) \mid sin(sk) \mid cos(sk)$$

$$sk2 ::= fMA(sk) \mid fMM(sk, sk) \mid fH(sk, sk, int)$$

The set of functional rates is denoted by $\mathcal{F}_R$.

The mathematical expressions are defined by some mathematical operators (sk). The predefined functions ($sk2$) are mass-action (fMA), Michaelis-Menten (fMM) and Hill kinetics (fH).

The functional rates are used to derive the transition rates of the system. In the functional rates some constant parameters can be used. These must be defined in the model by means of the set of parameter definitions $\mathcal{K}$.

Definition 2.2.4. *Each parameter is defined by "$k_{name} = value\ unit$", where:*

- $k_{name} \notin \mathcal{C}$ *is the parameter name,*
- *value denotes a positive real number,*
- *unit denotes the unit associated with the parameter, it is optional.*

The set of the parameters is denoted by $\mathcal{K}$.

Now we define the set of sequential components.

Definition 2.2.5. *The set Comp of sequential components is defined as:*

$$Comp ::= \{C \stackrel{\text{def}}{=} S, where\ S\ is\ a\ sequential\ component\}$$

According to the definitions above, we can define a Bio-PEPA system in the following way.

Definition 2.2.6. A Bio-PEPA system $\mathcal{P}$ is a 6-tuple $< \mathcal{V},\ \mathcal{N},\ \mathcal{K},\ \mathcal{F}_R,\ Comp,\ P >$,where:

- $\mathcal{V}$ is the set of compartments;
- $\mathcal{N}$ is the set of quantities describing each species;
- $\mathcal{K}$ is the set of parameter definitions;
- $\mathcal{F}_R$ is the set of functional rate definitions;
- Comp is the set of definitions of sequential components;
- P is the model component describing the system.

If each element of the Bio-PEPA system satisfies the following conditions, we call it a *well-defined* Bio-PEPA system:

1. $\mathcal{N}$ has to contain all the species component (at least a compartment must be defined). The optional elements satisfy $value\_H \in \mathbb{R}^+$ and $value\_H$ must be greater than 0, $value\_N$ must be equal or greater than 1 and it belongs to $\mathcal{N}$.
2. In $\mathcal{F}_R$, each variable in a functional rate refers to a species component in $\mathcal{N}$ or a constant parameter in $\mathcal{K}$.
3. Each element of $Comp$ must have prefix form $(\alpha, k)\ op\ S$ and the actions must be all distinct.
4. The model component P must be defined in terms of the species components ($Comp$), and each cooperation set $\mathcal{L}_i$ in P , is contained in $\mathcal{A}(P)$. (The set of all action types of the model component P is denoted by $\mathcal{A}(P)$).

We consider only *well-defined* Bio-PEPA systems.

2.2.2 Bio-PEPA Labelled Transition System

Bio-PEPA permits to consider two different binary relations over the processes depending from the kind of analysis that we want to do. They are the *Capability relation* and the *Stochastic relation*.

The Capability relation supports the derivation of quantitative information and it is auxilarity to the stochastic relation.

The Capability relation is defined as $\rightarrow_c \subseteq C \times \Theta \times C$, where the label Θ contains the quantitative information needed for the evaluation of the functional rates. Each element $\theta \in \Theta$ represents the pair (α, w) , where α is the label of the action, w is the list of the species that participate in the reaction defined as $[S : op(l,k)]|w :: w$ where l is level and k is stoichiometry coefficient.

The Capability relation $\rightarrow_c$ is the minimum relation satisfying the following rules:

PrefixReac $\quad ((\alpha,k) \downarrow S)(l) \xrightarrow{(\alpha,[S:\downarrow(l,k)])}_c S(l-k)\ , \qquad k \leq l \leq N$

PrefixProd $\quad ((\alpha,k) \uparrow S)(l) \xrightarrow{(\alpha,[S:\uparrow(l,k)])}_c S(l+k)\ , \qquad 0 \leq l \leq (N-k)$

PrefixMod $\quad ((\alpha,k)\ op\ S)(l) \xrightarrow{(\alpha,[S:\ op\ (l,k)])}_c S(l)\ , \qquad with\ op = \odot, \oplus, \ominus$
and

$$0 < l \leq (N),\ if\ op = \oplus,\ 0 \leq l \leq (N),\ otherwise$$

These three axioms describe the behaviour of the three different prefix terms. For a reactant ($\downarrow$) the level decreases, for a product ($\uparrow$) the level increases. while a modifier ($\oplus$, $\ominus$, $\odot$) level remains the same. If the enzyme is null, the rate of the enzymatic reaction with Michaelis - Menten kinetics is zero and the reaction is not possible.

Choice1 $\quad \dfrac{S_1(l) \xrightarrow{(\alpha,w)}_c S_1'(l')}{(S_1+S_2)(l) \xrightarrow{(\alpha,w)}_c S_1'(l')}$ $\qquad$ Choice2 $\quad \dfrac{S_2(l) \xrightarrow{(\alpha,w)}_c S_2'(l')}{(S_1+S_2)(l) \xrightarrow{(\alpha,w)}_c S_2'(l')}$

The rules Choice1 and cChoice2 have the usual meaning.

Constant $\quad \dfrac{S(l) \xrightarrow{(\alpha,[S:\ op\ (l,k)])}_c S'(l')}{C(l) \xrightarrow{(\alpha,[S:\ op\ (l,k)])}_c S'(l')} \qquad$ with $C \stackrel{\text{def}}{=} S$

The rule Constant is used to define the behaviour of a constant term, defined by one or more prefix terms in summation.

$$\text{Coop1} \qquad \frac{P_1 \xrightarrow{(\alpha,w)}_c P_1'}{(P_1 \bowtie_{\mathcal{L}} P_2) \xrightarrow{(\alpha,w)}_c P_1' \bowtie_{\mathcal{L}} P_2} \qquad \text{with } \alpha \notin \mathcal{L}$$

$$\text{Coop2} \qquad \frac{P_2 \xrightarrow{(\alpha,w)}_c P_2'}{(P_1 \bowtie_{\mathcal{L}} P_2) \xrightarrow{(\alpha,w)}_c P_1 \bowtie_{\mathcal{L}} P_2'} \qquad \text{with } \alpha \notin \mathcal{L}$$

$$\text{Coop3} \qquad \frac{P_1 \xrightarrow{(\alpha,w)}_c P_1' \qquad P_2 \xrightarrow{(\alpha,w)}_c P_2'}{(P_1 \bowtie_{\mathcal{L}} P_2) \xrightarrow{(\alpha,w)}_c P_1' \bowtie_{\mathcal{L}} P_2'} \qquad \text{with } \alpha \notin \mathcal{L}$$

The last three rules report the case of cooperation. The rules Coop1 and Coop2 concern the case when the enabled action does not belong to the cooperation set. In this case the label in the conclusion contains only the information about the component that fires the action. The rule Coop3 describes the case in which the two components synchronise and the label reports the information from both the components.

The Stochastic relation is the relation which specifies the rates. The rates are obtained by evaluating the functional rate associated with the action, divided by the step size, and by using the quantitative information derived from the capability relation.

The stochastic relation is defined as $\rightarrow_s \subseteq \widetilde{\mathcal{P}} \times \Gamma \times \widetilde{\mathcal{P}}$, where the label $\gamma \in \Gamma$ is defined as $\gamma = (\alpha, r_\alpha)$, with $r_\alpha \in \mathbb{R}^+$. r_α is computed through its functional rate and it represents the parameter of a negative exponential distribution. The dynamic behaviour of processes is determined by a race condition which is: all enabled activities attempt to proceed, but only the fastest one succeeds.

The stochastic relation $\rightarrow_s$ is defined as the minimal relation satisfying the rule:

$$\textbf{Final} \qquad \frac{P \xrightarrow{(\alpha_j,w)}_c P'}{\langle \mathcal{V}, \mathcal{N}, \mathcal{K}, \mathcal{F}, Comp, P \rangle \xrightarrow{(\alpha_j, r_\alpha[w,\mathcal{N},\mathcal{K}])}_s \langle \mathcal{V}, \mathcal{N}, \mathcal{K}, \mathcal{F}, Comp, P' \rangle}$$

The rate is calculated from the functional rate f_α in this way:

$$r_\alpha[w, \mathcal{N}, \mathcal{K}] = \frac{f_\alpha[w, \mathcal{N}, \mathcal{K}]}{H}$$

A stochastic labelled transition system can be derived from the Bio-PEPA system. Differently from other process algebras, each component is associated with a discrete level of concentration. Assume a finite maximum concentration and, given a concentration step size H, we can obtain a finite number of concentration levels for the species. The step size is assumed equal for all the species in a given compartment and it represents the granularity of the system. The smaller H, the finer the granularity.

Definition 2.2.7. *The Stochastic Labelled Transition System (SLTS) for a Bio-PEPA system is* $(\widetilde{\mathcal{P}}, \Gamma, \rightarrow_s)$*, where* $\rightarrow_s$ *is the minimal relation satisfying the rule* ***Final****.*

The states of the SLTS are defined in terms of the concentration levels of the system components and the transitions from one state to another represent reactions that cause changes in the concentration levels of some components. The number of levels depends on the stoichiometric coefficients of the species involved.

2.2.3 Derivation of Rates

The functional rate is important in a Bio-PEPA model . The stoichiometry coefficients are not enough to simulate and analyse a system and the range of the functional rates affects heavily the behaviour of the system. In [1] some examples are given, showing how to model the transition rate In the two main cases of reactions.

Case1: A reaction j is described by a kinetic law f_j and all its stoichiometric coefficients equal to one, then :

- $(\Delta t)^{-1}$ is the transition rate, where Δt is the time to have a variation in the concentration of one step for both the reactants and the products of the reaction.
- y is a variable describing product of the reaction.
- the rate equation for the species with respect to the given reaction is $dy/dt = f_j(\bar{x}(t))$, where $\bar{x}$ is the set of reactants/modifiers of the reaction.

We can apply the *Taylor expansion* up to the second term, obtaining

$$\mathrm{y}_{n+1} \approx \mathrm{y}_n + f(\bar{x}) \times (t_{n+1} - t_n)$$

We can set $\mathrm{y}_{n+1} - \mathrm{y}_n = H$ and the time interval $(t_{n+1} - t_n) = \Delta t$ then $\Delta t \approx H/f(\bar{x}_n)$. From it we can obtain the transition rate $(\Delta t)^{-1} = f(\bar{x}_n)/H$.

Case2: A reaction has some stoichiometric coefficient different from one (Example 2.2.1 shows this case). Let y be a product of the reaction. The approximation gives

$$\mathrm{y}_{n+1} \approx \mathrm{y}_n + r \times k \times \prod_{i=1}^{n_r} {x^{k_i}}_{i,n} \times (t_{n+1} - t_n)$$

where

- r is the reaction constant rate.
- k is the stoichiometric coefficient of the product y.
- $x_i (i = 1...n_r)$ are the reactants and $k_i (i = 1...n_r)$ are the associated stoichiometric coefficients.
- n_r is number of distinct reactants.

We can set $\mathrm{y}_{n+1} - \mathrm{y}_n = k \times H$ and $(t_{n+1} - t_n) = \Delta t$. Then we obtain the transition rate from $\Delta t \approx H/(r \times \prod_{i=1}^{n_r} x^{k_i}{}_{i,n})$.

This approach depends on the time/concentration steps. It assumes the species can vary by one step size at a time and that reactants are decreased until 0, while products are increased.

We can summarize that the rate associated with a transition from one state u to another state v can be calculated as: $r_j = \frac{f_j[u]}{h}$, where h is the step size of the reactants and $f_j[u]$ is the evaluation of the functional rate in the state u. When the stoichiometric coefficient of a reagent is k then the reagent varies by k levels as a result of the transition.

Now we give an example of a biochemical reaction modelled in Bio-PEPA. The example is taken from [].

Example 2.2.1. The reaction $2X + Y \xrightarrow{f_M} 3Z$ is associated to the mass action kinetic law $f_M = r \times X^2 \times Y$. The three species can be described in Bio-PEPA by the components:

$$X \stackrel{def}{=} (\alpha, 2) \downarrow X \qquad Y \stackrel{def}{=} (\alpha, 1) \downarrow Y \qquad Z \stackrel{def}{=} (\alpha, 3) \uparrow Z$$

The system corresponding to the reaction is described by

$$(X(x_0) \underset{\alpha}{\bowtie} Y(\mathrm{y}_0)) \underset{\alpha}{\bowtie} Z(z_0)$$

where x_0, y_0 and z_0 are the initial concentration levels of the three components. In the case of the levels of concentration, the model component is described as

$$(X(l_{x_0}) \bowtie_{\{\alpha\}} Y(l_{\mathrm{y}_0})) \bowtie_{\{\alpha\}} Z(l_{z_0})$$

where $l_{x_0}, l_{\mathrm{y}_0}$ and l_{z_0} denote the initial levels of the three components and are derived from the initial concentrations. The functional rate is $f_\alpha = fMA(r)$

The rate associated with a transition α corresponds to the same case we illustrate previously and it is :

$$r_\alpha = \frac{r \times (l_X \times H)^2 \times (l_Y \times H)}{H}$$

where l_X, l_Y are the concentration levels for the species X and Y in a given state and H is the step size of all the species.

2.2.4 Auxiliary definitions of Bio-PEPA

We give the inductive definition of set of current action types in a species or in a model components of Bio-PEPA.

Definition 2.2.8. *The set of current action types enabled in the model component P, denoted $\mathcal{A}(P)$, is defined as:*

$$\mathcal{A}((\alpha, k)opS) = \{\alpha\}$$

$$\mathcal{A}(S_1 + S_2) = \mathcal{A}(S_1) \cup \mathcal{A}(S_2)$$

$$\mathcal{A}(C) = \mathcal{A}(S) \quad where\ C \stackrel{\text{def}}{=} S$$

$$\mathcal{A}(S(l)) = \mathcal{A}(S)$$

$$\mathcal{A}(P_1 \bowtie_{\mathcal{L}} P_2) = \mathcal{A}(P_1)\backslash\mathcal{L} \cup \mathcal{A}(P_2)\backslash\mathcal{L} \cup (\mathcal{A}(P_1) \cap \mathcal{A}(P_2) \cap \mathcal{L})$$

If $\mathcal{P}$ is a Bio-PEPA system with model component P, the set of current action types enabled in $\mathcal{P}$ is $\mathcal{A}(\mathcal{P}) = \mathcal{A}(P)$.

The only element that evolves in a Bio-PEPA system $\mathcal{P} = \langle \mathcal{V}, \mathcal{N}, \mathcal{K}, \mathcal{F}, Comp, P \rangle$ is the model component P, the other components remains unchanged, they collect information about the compartments, the species, and the definition of the species component. In some cases it is simpler to consider only the model component instead of considering whole system and use other components to derive the rates.

A function π_p given a Bio-PEPA system returns the model component, $P = \pi_p(\mathcal{P})$.

Definition 2.2.9. *Let $\mathcal{P}$ be a Bio-PEPA system and let $P = \pi_p(\mathcal{P})$. Let P_u, P_v be two derivatives of the model component P with P_v a one-step derivative of P_u. The set of action types associated with the transitions from the process P_u to the process P_v is denoted $\mathcal{A}(P_u|P_v)$.*

The next definition gives the complete action type set of a system $\mathcal{P}$ and of a component P.

Definition 2.2.10. *The complete action type set of a system $\mathcal{P}$ is defined as :*

$$\bar{\mathcal{A}} = \cup_{\mathcal{P}_i \in ds(\mathcal{P}} \mathcal{A}(\mathcal{P}_i)$$

The complete action type set of a component P is similarly defined.

2.2.5 Properties of Bio-PEPA

As we mentioned before, PEPA uses the equivalence relation, the isomorphism, and the bisimulation properties to analyse models and state relations. Similarly Bio-PEPA also has its meaningful properties, namely an isomorphism and a (strong) bisimulation.

Two Bio-PEPA systems $\mathcal{P}_1$ and $\mathcal{P}_2$ are equivalent if their respective model components are equivalent. We just recall some properties.

(1) Isomorphisim

Two components are Isomorphic if they generate derivation graphs with the same structure and capable of carrying out exactly the same activities. Some consequences of isomorphism are expressed in the *Equational laws*.

Choice The laws for choice are:

(a) $P + Q = Q + P$

(b) $P + (Q + R) = (P + Q) + R$

Cooperation The laws for cooperation are:

(a) $P \underset{\mathcal{L}}{\bowtie} Q = Q \underset{\mathcal{L}}{\bowtie} P$

(b) $P \underset{\mathcal{L}}{\bowtie} (Q \underset{\mathcal{L}}{\bowtie} R) = (P \underset{\mathcal{L}}{\bowtie} Q) \underset{\mathcal{L}}{\bowtie} R$

(c) $P \underset{\mathcal{K}}{\bowtie} Q = P \underset{\mathcal{L}}{\bowtie} Q \qquad if\ \mathcal{K} \cap (\bar{\mathcal{A}}(P) \cup \bar{\mathcal{A}}(Q)) = \mathcal{L}$

(d) $(P \underset{\mathcal{L}}{\bowtie} Q) \underset{\mathcal{K}}{\bowtie} R = \begin{cases} P \underset{\mathcal{L}}{\bowtie} (Q \underset{\mathcal{K}}{\bowtie} R) & if\ \bar{\mathcal{A}}(R) \cap (\mathcal{L} \backslash \mathcal{K}) = \emptyset \ \wedge\ \bar{\mathcal{A}}(P) \cap (\mathcal{K} \backslash \mathcal{L}) = \emptyset \\ Q \underset{\mathcal{L}}{\bowtie} (P \underset{\mathcal{K}}{\bowtie} R) & if\ \bar{\mathcal{A}}(R) \cap (\mathcal{L} \backslash \mathcal{K}) = \emptyset \ \wedge\ \bar{\mathcal{A}}(Q) \cap (\mathcal{K} \backslash \mathcal{L}) = \emptyset \end{cases}$

Constant: If $A \stackrel{def}{=} P$ then $A = P$

Bio-PEPA systems: Let $\mathcal{P}_1$ and $\mathcal{P}_2$ be two Bio-PEPA systems, with $P = \pi_P(\mathbb{P}_1)$ and $Q = \pi_P(\mathcal{P}_2)$. If $P = Q$ then $\mathcal{P}_1 = \mathcal{P}_2$.

(2) Strong bisimulation

Bisimilar components or systems are able to perform the same actions with the same rates resulting in derivatives that are themselves bisimilar. The definition of bisimulation is based on the *labelled transition system*.

Two components that are isomorphic are also strong bisimilar.

2.2.6 Analysis in Bio-PEPA

Bio-PEPA integrates several kinds of analysis techniques. From a Bio-PEPA system we can obtain:

- a CTMC (with and without levels),
- an ODE system for simulation and other kinds of analysis,
- a Gillespie model for stochastic simulation,
- a PRISM model for model checking.

Each of these analyses can be of help for studying different aspects of the biological model, they can be also used in conjunction.

(1) **From Bio-PEPA to CTMC with levels**

In Bio-PEPA the granularity of the system is expressed in terms of concentration levels of size H, when H decreases the number of levels increases. The states of the CTMC are characterised in terms of the concentration levels for each of the species of the system. A CTMC with levels is a CTMC whose states capture levels of concentration of the species and the transitions from one state to another reflect changes of these levels.

Definition 2.2.11. *A state of a CTMC with levels is defined as a vector of levels $\sigma = (l_1, l_2, ..., l_n)$, where l_i, for $i = 1, 2, ..., n$, is the level of the species i. The transitions of a CTMC with levels represent biochemical reactions. Each transition causes a change in the number of levels of one or more species, as determined by the stoichiometry.*

In the analysis of CTMC we consider two cases. If a reactions creates a new element of the kind " $\rightarrow A$" or "$C \rightarrow C + A$" that it is called *creation reaction.* We call a biochemical network without creation reactions *a bounded chemical network.* For a bounded chemical network we have the following proposition.

Proposition 1. *Let X_h be a CTMC corresponding to a bounded biochemical network with granularity h. Let $\sigma_0 = (l_{1,0}, l_{2,0}, ..., l_{n,0})$ be the initial state. If the values $l_{i,0} (i = 1, ..., n)$ are finite then X_h is finite and the maximum value of the level depends on σ_0 and on the stoichiometric coefficients of the reactions.*

In the case of creation reactions, starting from a finite number of levels, it is possible to obtain an infinite CTMC. We assume a maximum level also for the species that grows infinitely, in order to guarantee a finite CTMC.

Theorem 2.1. $(taken form[2])$ *For any finite Bio-PEPA system* $\mathcal{P} = \langle \mathcal{V}, \mathcal{N}, \mathcal{K}, \mathcal{F}_{\mathcal{R}}, Comp, P \rangle$, *let* $X(t)$ *be a stochastic process, such that* $X(t) = P_i$ *indicates that the system behaves as the component* P_i *at time* t. *Then* $X(t)$ *is a CTMC.*

A CTMC with levels has the following advantages:

- It is based on discrete levels of concentration instead of exact numbers of molecules. It allows one to deal with incomplete information about molecular concentrations.
- It leads to a reduction of the state space, leading to models which may be amenable to numerical solution and approaches such as stochastic model checking.
- The set of ODEs derived from a Bio-PEPA system is able to capture the limiting behaviour of the CTMC with levels representing the discretized system.

The complexity of CTMC is expressed in terms of the number of states. This number depends on the amount of levels of the species. The upper bound of number of states is :$\Pi_{i=1}^{n}(N_i + 1)$, where N_i is maximum level of species i and n is the number of species.

Now we see an example of CTMC, taken from [3].

Example 2.2.2. *Dimerization is the chemical reaction that joins two molecular subunits, resulting in the formation of a single dimer. In the dimerization of a protein we have following reaction:*

$$2A \xrightarrow{f_{MA}(k_1)} B \qquad\qquad B \xrightarrow{f_{MA}(k_2)} 2A$$

The dynamics are Mass Action Kinetics:

$$f_{MA}(k_1) = k_1.A^2 \qquad\qquad f_{MA}(k_2) = k_2.B$$

We assume $A = 10\ mol/L; B = 0\ mol/L;\ k_1 = k_2 = 1.0$. *Step size for both* A *and* B *is* $H = 5$. *Number of levels and initial concentration levels of* A *and* B *are* $N_A = 2$, $N_B = 1$ *and* $M_{A0} = 10$, $M_{B0} = 0$ *respectively. The compartment containing the species in both* A *and* B *is* `cell`. *Functional rates are* $f_{a1} = f_{MA}(k_1)$, $f_{a2} = f_{MA}(k_2)$.

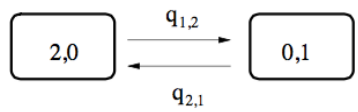
Set of species components:

$$A \stackrel{\text{def}}{=} (\alpha_1, 2) \downarrow A + (\alpha_2, 2) \uparrow A \qquad\qquad B \stackrel{\text{def}}{=} (\alpha_1, 1) \uparrow B + (\alpha_2, 1) \downarrow B$$

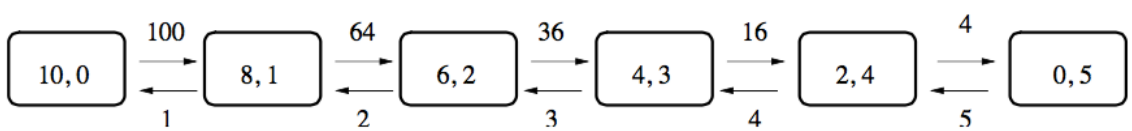
Model component:

$$A(2) \underset{\{\alpha_1,\alpha_2\}}{\bowtie} B(0)$$

In this case the initial state represents the initial concentration level, it is: $\left(\left\lfloor \frac{x_1(0)}{H} \right\rfloor, \left\lfloor \frac{x_2(0)}{H} \right\rfloor\right)$. *We derive CTMC with levels for H=5 and H=1. When H=5 the initial state is* $\left(\left\lfloor \frac{10}{5} \right\rfloor, \left\lfloor \frac{0}{5} \right\rfloor\right) = (2,0)$ *and the next state is (0,1). With rates* $q_{1,2} = f_{\alpha_1}/H = 20$ *and* $q_{2,1} = f_{\alpha_2}/H = 1$. *The CTMC for this case is:*



When H=1, the initial state is $\left(\left\lfloor \frac{10}{1} \right\rfloor, \left\lfloor \frac{0}{1} \right\rfloor\right) = (10,0)$, *the next states are* $(8,1), (6,2), (4,3), (2,4), (0,5)$. *The rates are calculated in the same way above. The CTMC for this case is :*



When H is smaller then there is a finer granularity. When H = 0.1 there are 51 states and 100 transitions.

(2) **From Bio-PEPA to ODEs**

We define the set of ODEs from a Bio-PEPA system as π_{ODE}. π_{ODE} entails three steps:

- Definition of the stoichiometry matrix $D(n \times m)$, where n is the number of species and m is the number of molecules;
- Definition of the kinetic law vector $(m \times 1) v_{KL}$ containing the kinetic law of each reaction;
- Definition of the vector $(n \times 1)$ x, with $x^T = (x_1, x_2, ..., x_n)$.

Definition 2.2.12. *The ODE system thus obtained has the form:*

$$\frac{d_x}{d_t} = D \times v_{KL}$$

the vector of initial concentration is x_0, with $x_{i,0} = l_{i,0} \cdot H$ where $i = 1, 2, ..., n$, and H is the concentration level.

Let us explain the construction of the stoichiometry matrix $D = \{d_{ij}\}$. The entries of the matrix are obtained as follows: for each sequential component C_i consider the subterms C_{ij} representing the contribution of the species i to the reaction j. If a subterm represents a reactant, we write the corresponding stoichiometry as $-k_{ij}$ in the entry d_{ij}. In the case of a product we write $+k_{ij}$. All other cases are null.

We continue Example 2.2.2, according to its assumptions and values derive an ODE system.

Example 2.2.3. *According to the given values, we obtain the stoichiometry matrix*

$$D = \begin{bmatrix} -2 & +2 \\ +1 & -1 \end{bmatrix}$$

We define the vector $x^T = (x_A,\ x_B)$. The Kinetic Law Vector is:

$$v_{KL}^T = (k_1 \cdot x_A^2,\ k_2 \cdot x_B)$$

The system of ODEs obtained by π_{ODE} is:

$$\frac{d_{x_A}}{d_t} = -2 \cdot k_1 \cdot x_A^2 + 2 \cdot k_2 \cdot x_B$$
$$\frac{d_{x_B}}{d_t} = k_1 \cdot x_A - k_2 \cdot x_B$$

with the initial condition: $(x_{A,0},\ x_{B,0})^T = (10, 0)$.

(3) **From Bio-PEPA to Stochastic Simulation**

Gillespie stochastic simulation algorithm is a widely-used method for the simulation of biochemical reactions. The derivation of the Gillespie model is made by creating molecules corresponding to each species and defining the possible reactions with appropriate adjustment of kinetic rates.

The translation of a Bio-PEPA model for Gillespie simulation is similar to the approach proposed for ODEs. The initial number of molecules for the species i can be calculated from the concentration as $X_{i,0} = x_{i,0} \times v \times N_A$, where v is the volume of the compartment

of the species and N_A is the Avogadro number, i.e the number of molecules in a mole of a substance. For details see [6].

Gillespie stochastic simulation algorithm supposes elementary reactions with at most two reactants and constant rates (mass-action kinetics). If the model contains only this kind of reactions the translation is straightforward. If there are non-elementary reactions and general kinetic laws, it is a widely used approach to consider them translated directly into a stochastic context.

(4) **From Bio-PEPA to PRISM**

PRISM [6] is a probabilistic model checker, a tool for the formal modelling and analysis of systems which exhibit random or probabilistic behaviour.

The PRISM language is composed of *modules* and *variables*.

- A model is composed of a number of modules which can interact with each other.
- A module contains a number of local variables.
- The values of these variables at any given time constitute the state of the module.
- The global state of the whole model is determined by the local state of all modules.
- The behaviour of each module is described by a set of commands. Each update describes a transition which the module can make if the guard is true.
- A transition is specified by giving the new values of the variables in the module, possibly as a function of other variables. Each update is also assigned a probability (or in some cases a rate) which will be assigned to the corresponding transition.

In PRISM models the variables express levels of concentration. It is possible to derive PRISM models where molecules are counted instead of levels. The values for the species are given in terms of levels or molecules and the rates must be chosen in order to take the interpretation into account.The maximum level/concentration for each species must be given in the specification of a Bio-PEPA system and, if necessary, the maximum number of molecules can be derived from it.

We focus on the dynamic analysis of the system and for this reason we mainly talk about CTMC and ODEs and the relation of these two models in a Bio-PEPA system.

2.2.7 Comparison of CTMC with levels and ODEs

The step size (granularity) h, is a very important factor to compare CTMC and ODEs models in a Bio-PEPA system. Appropriate value of granularity h is used to reduce the differences between these two models.

In order to compare these two models with acceptable granularity h, we apply Kurtz's theorem[3]. It requires the CTMC to be *Density Dependent Marcov Chain*. Kurtz's theorem indicates that, under some conditions, when the granularity h goes to 0, the limiting behaviour of Density dependent CTMC (with levels) is approaches to a set of ODEs.

For applying Kurtz's Theorem, we first show *Density dependent CTMC*, then we introduce all the conditions of Kurtz's theorem that must be satisfied (Appendix B).

Definition 2.2.13. *A family of CTMCs X_h, for the parameter h , is called* `density dependent`*, if the entry $q_{u,v}$ for the infinitesimal generator matrix is*

$$q_{u,v} = \sum_{\mathcal{A}(P_u|P_v)} f_j[u] \cdot h^{-1} \quad if \;\; u \neq v \qquad\qquad q_{u,v} = \sum_{u \neq v} q[u,v] \quad otherwise.$$

where $\mathcal{A}(P_u|P_v) = \{\alpha | P_u \xrightarrow{\alpha} P_v\}$ and $f_j[u]$ is the evaluation of the functional rate in state P_u.

Given a state σ of the CTMC, we denote by h_σ the vector $(h \cdot l_1, h \cdot l_2, ..., h \cdot l_n)$, where h is the step size and l_i is the level of the species i. Let D be the stoichiometry matrix obtained from the Bio-PEPA system and D^j the jth column of D. This vector represents the stoichiometric coefficients for all the species in a given reaction j. The kinetic law associated with the reaction j is denoted by $f_j(h_\sigma, D^j)$, where the rate is $f_j(h_\sigma, D^j) \cdot h^{-1}$.

There are three conditions of Kurtz's theorem .

Let x_0 be the initial concentration vector for the ODEs. The initial level vector is $l_0 = \lceil x_0/h \rceil$. Thus $\lim_{h \to 0} h \cdot l_0 = x_0$.

Consider the system of ODEs, $dX(t)/dt = F(x)$ where $F(x) = \sum D^j f_j(x, D^j)$ with initial condition $X(0) = x_0$. By hypothesis, the trajectory of $X(t)$ is bounded, so we can assume it is bounded by some open set E. Since each kinetic law is continuously differentiable (the first condition for the kinetic laws), it follows that f is Lipschitz. This the first condition of Kurtz's Theorem. The second and third conditions of Kurtz's Theorem state that for each transition the rate of change is bounded and that there is a bound for the whole state space so that the impact of each transition is bounded. For the third condition, we can observe that $f(x, D^j) = 0$ for all $|D^j| > C \; with \; C = \sum_{i,j} d_{ij}$.

Consider the ODE system $\pi_{ODE}(\mathcal{P})$, for a given Bio-PEPA system $\mathcal{P}$. We can observe that $F(x) = D \times v_{KL}$, as the kinetic law vector v_{KL} contains all the functions f_j for all the reactions. The ODE system $\pi_{ODE}(\mathcal{P})$ coincides with the one in Kurtz's Theorem, with initial condition $x_{i,0} = l_{i,0} \cdot h, for \; i = 1, 2, ..., n$.

This confirms that, in the limit, there is agreement between the ODEs and CTMC with levels derived from a Bio-PEPA system $\mathcal{P}$.

In [3] the authors in order to compare them propose two measures of distance between the two models. In both the cases, the distance between two models generally decreases with the step size h. In the case of dimerization example (Example 2.2.2) we show the behaviour of the distance with the step size in Figure 2.3. For very small H the number of states becomes large and even the simulation of the CTMC may become prohibitively expensive. Thus there is a trade-off between accuracy (in terms of both number of runs and step size) and tractability.

Step size	H=5	H=1	H=0.1	H=0.01
f_{dist}	53.69	3.1	1.02	1.00

Figure 2.3: Distance between two models in the dimerization example

2.2.8 Tools for PEPA and Bio-PEPA

(1) **Tools for PEPA**

The PEPA process algebra is supported by the PEPA Eclipse Plug-in project, a contribution to the Eclipse integrated development environment. The PEPA Eclipse Plug-in contains a PEPA editor and a performance analysers which uses Markov chain or ODE methods or simulation. Performance results are displayed graphically on the platform itself. It also contains an interface for abstracting PEPA models, and for model checking properties in the Continuous Stochastic Logic (CSL).

PEPA is supported by a suite of tools which include the PEPA Workbench and PEPAroni. The PEPA Workbench is available in two editions, one for ML and the other for Java. The ML edition of the PEPA Workbench transforms PEPA descriptions into a form suitable for solution by another solution tool such as Maple, Matlab and Mathematica. The Java edition of the PEPA Workbench can solve models without the need for a separate solution tool. PEPAroni is a discrete-event simulator for PEPA.

There are also other tools for obtaining different kind of analysis.

- The International PEPA compiler (IPC) is a stand-alone passage-time analysis tool for PEPA models.
- PEPA is also one of the languages supported by the *multi-paradigm* modelling tool *Möbius*.
- PEPA models can also be processed using the probabilistic model checker PRISM.

- The Imperial PEPA compiler (ipc) compiles PEPA models into the input language of Will Knottenbelt's *DNAmaca tool.*
- The Grouped PEPA analyser (GPAnalyser) allows fluid analysis of massively parallel systems expressed in the Grouped PEPA language.
- *EMPEPA* is a software that permits to find the most likely rates of a PEPA model according to a set of sample executions by using the *EM algorithm.*
- The *wflow2pepa* application program can be used to model performance of workflow systems. The user provides a description file by using a form of hierarchical pattern based language. From this, the application generates the corresponding PEPA performance model automatically.

(2) **Tools for Bio-PEPA**

Two software tools are available for modeling with Bio-PEPA: the Bio-PEPA Workbench and the Bio-PEPA Eclipse Plug-in.

The Bio-PEPA Workbench is a command-line tool which automatically translates Bio-PEPA models into equivalent representations for use in other tools (e.g. MATLAB, Dizzy). It is an implementation of Bio-PEPA which allows modellers to write models in the Bio-PEPA language and to animate them using stochastic simulation. The Bio-PEPA Workbench uses the StochKit stochastic simulation toolkit to perform exact stochastic simulations using Gillespie's SSA algorithm. The Bio-PEPA Eclipse Plug-in is a GUI-based modelling tool integrated in the Eclipse development environment. The syntax-aware editor assists users in model development and debugging, since it is able to detect syntactic errors and supports static analysis such as identification of sources and sinks and computation of invariants related to species populations. Dynamic analysis is supported via various stochastic simulators and ODE solvers that can be used for time-series analysis and model experimentation. Models can be also exported to other formats.

Let's see an example of a simple Bio-PEPA system, taken from [17], in this example we will briefly show the implementation of the system in Bio-PEPA Eclipse plug-in.

Example 2.2.4. In a reaction r_1 enzyme E combines with a substrate S to form a compound $E : S$. This compound in reaction rm_1 might degrade releasing the enzyme and the substrate or in reaction r_2 it might convert the substrate into a product P, releasing the enzyme. We can represent the reactions in this way:

$$E + S \xrightarrow{r_1} E : S \qquad E : S \xrightarrow{rm_1} E + S \qquad E : S \xrightarrow{r_2} E + P$$

In the Bio-PEPA Eclipse tool we use " $<<$ " *and* " $>>$ " to represent reactant and product. The cooperation symbol " $\bowtie$ " is denoted as " $< * >$ " in Eclipse. According to the species components definition:

- E is a reactant $(\downarrow, <<)$in reaction r_1 and a product $(\uparrow, >>)$ in reactions rm_1 and r_2;
- S is a reactant $(\downarrow, <<)$ in reaction r_1 and a product $(\uparrow, >>)$ in reaction rm_1;
- $E : S$ is a product $(\uparrow, <<)$ in reaction r_1 and a reactant $(\downarrow, >>)$ in reactions rm_1 and r_2;
- P is a product $(\uparrow, >>)$ in reaction r_2.

The Bio-PEPA system representing the tree reactions can be described in the following way.

Species Definitions:

$$\begin{gathered}\text{species } E : \text{ upper=100 , lower=0;}\\ \text{species } S : \text{ upper=100 , lower=0;}\\ \text{species } E : S : \text{ upper=100 , lower=0;}\\ \text{species } P : \text{ upper=100 , lower=0;}\end{gathered}$$

Parameter Definitions:

$$\begin{gathered}k_1 = 1.0;\\ km_1 = 0.1;\\ k_2 = 0.01;\end{gathered}$$

Functional Rates:

$$\begin{gathered}r_1 = k_1 * E * S;\\ rm_1 = km_1 * E : S;\\ r_2 = k_2 * E : S;\end{gathered}$$

Species Components:

$$\begin{gathered}E = (r_1, 1) << +(rm_1, 1) >> +(r_2, 1) >>;\\ S = (r_1, 1) << +(rm_1, 1) >>;\\ E : S = (r_1, 1) >> +(rm_1, 1) << +(r_2, 1) <<;\\ P = (r_2, 1) >>;\end{gathered}$$

Model Component(with initial levels):

$$E[40] < * > S[30] < * > \ E : S[0] \ < * > P[0]$$

Chapter 3

Trophic Networks

To date there have been various studies that investigate ecological problems and model the interactions in an ecosystem. In this chapter we introduce trophic networks and modelling of ecological systems. All the informations taken from [11],[12],[14],[15].

3.1 Basic Concepts in a Trophic networks

Ecology is the scientific analysis and study of interactions among organisms and their environment. It is truly different from other disciplines in its conceptions about how nature operates. In order to analyse an ecology system, it is very common to use trophic models. We introduce a trophic network and we give the main concepts related to it.

Definition 3.1.1. *Ecology is the study of the relationships of organisms with one another and with their non-living environment.*

In the real world there are also non-living components that affect the organisms. For this reason the concept of ecological community is extended to incorporate also the non-living components.

Definition 3.1.2. *An Ecosystem is a community of living organisms, such as plants, animal and microbes, in conjunction with the non-living components of their environment such as air, water and detritus, which interact as a system.*

We can see an ecosystem as a *trophic chain* or *pyramid*: Energy is fixed by autotrophs or primary producers at the first level, and some of those resources are transmitted to the herbivores at the second level. Transfer occurs to carnivores at the third trophic level and to top carnivores at the fourth.

Among an ecosystem we can find many reactions between the living and non-living organisms. For example the most common and simple reaction is 'eat', Carnivores eat herbivores, herbivores eat erbs, ect. And they form a chain, called *food chain*.

Definition 3.1.3. *A food chain is a sequence of one single reaction in a biological community (an ecosystem) to obtain nutrition.*

All the reactions in an ecosystem can be seen as a network where the species in the ecosystem are the elements of the network and the links between pair of elements are the species reactions. Such a network is called *ecological network*.

Definition 3.1.4. *An Ecological network is a representation of biotic reactions in the ecosystem, where the species are connected with pairwise link.*

There are two types of reactions in an ecological network. *Trophic* is the type of reactions which represent feeding connections. *Symbiotic* is the type of reactions such species depend on each other for survival.

Trophic describes everything relating to the nutrition of a living tissue or organ. Trophic relationship is the link that unites a predator and its prey in an ecosystem. A species trophic level indicates the number of times chemical energy is transformed from a consumers diet into a consumers biomass along the food chains that lead to the species.

Ecologists consider all life forms into one of two categories, the autotrophs and the heterotrophs. Autotrophs produce biomass energy, either chemically without the suns energy or by capturing the suns energy in photosynthesis, than they use during metabolic respiration. Heterotrophs consume rather than produce biomass energy, as they metabolize, grow, and add to levels of secondary production.

Definition 3.1.5. *The trophic level of an organism is the position it occupies in a food chain.*

Trophic chain is a synonym for food chain. At the bottom of the food chain are autotrophs, which produce organic matter from inorganic matter (plants for example) followed by primary consumers, herbivores. Next the predators come and finally the decomposers or detritivores. Each species in a food chain depends on the previous species for its survival.

Definition 3.1.6. *Trophic network is a set of interconnected food chains within an ecosystem.*

Trophic network can be studied in quantitative and systematic fashion at several levels.

Definition 3.1.7. *An ecological network is called food web when all the reaction in the system are trophic.*

Food webs within ecological systems, describe the food chains in these systems and generally a graphical representation of "what-eats-what" in an ecological community. The linkages in a food web illustrate the feeding pathways, such as where heterotrophs obtain organic matter by feeding on autotrophs and other heterotrophs. Food cycle is the antiquated term that is synonymous with food web.

3.2 Representation of Trophic Networks

An ecosystem is generally an open system, meaning that it exchanges material and energy with its surroundings and also with the outside world (in Figure 3.1 for example). Hence, for representing and analysing trophic networks, generally also the input and output flows are taken into account. Inputs can be primary production, immigration or incoming of detrital matter into the system, while outputs can be emigration, harvesting by humans and exit of detrital matter from the system. Some energy may be dissipated into heat (respiration) or some material may be degraded into its lowest energy form (detritus).

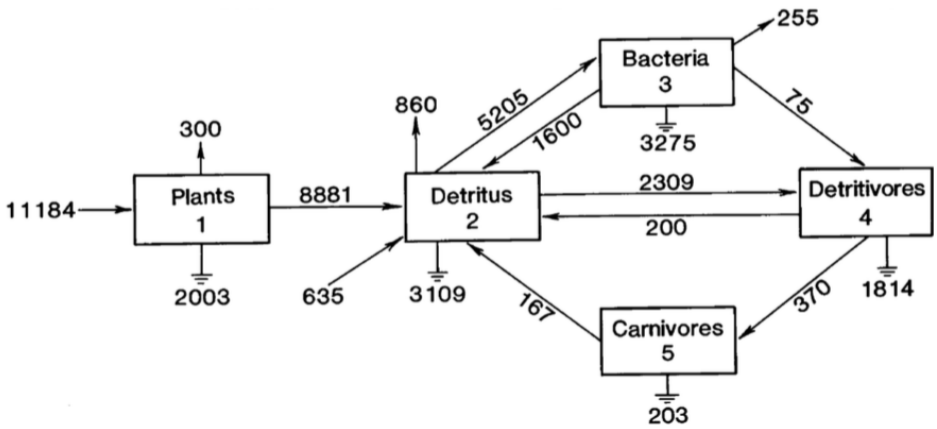


Figure 3.1: The trophic exchanges of energy in the Cone spring ecosystem (Tilly,1968). Arrows not originating from a box represent exogenous inputs. Ground symbols represent dissipations. Arrows not terminating in a box portray exogenous outputs [11].

In order to define an ecological network in a more realistic way, it becomes necessary to identify the significant taxa or nodes in the ecosystem. For each taxa, other nodes are present in its diet. Another important consideration is the flow rate among these taxas. Once these informations are known for all taxa, they can be represented in two ways, one is graphically as a directed graph and another is a matrix representation.

3.2.1 Graph representation

On a digraph the nodes are usually represented as boxes, and each transaction is represented as an arrow that originates out of the prey taxon and terminates with an arrowhead at the predator node. There are four kinds of graph representation.

- Undirected single graph

  A graph G can be defined as a pair (V, E) where V is a set of vertices representing the nodes and E is a set of edges representing the connections between the nodes. We

define as $E = \{(i,j)|i,j \in V\}$ the single connection between nodes i and j. In this case, we say that i and j are neighbors.

- Directed graph

  A directed graph is defined as an ordered triple $G = (V, E, f)$, where f is a function that maps each element in E to an ordered pair of vertices in V. The ordered pairs of vertices are called *directed edges*, arcs or arrows. An edge $E = (i, j)$ is considered to have direction from i to j. Directed graphs are mostly suitable for the representation of biological pathways or procedures which show the sequential interaction of elements at one or multiple time points and the flow of information throughout the network.

- Weighted graph

  A weighted graph is defined as a graph $G = (V, E)$ where V is a set of vertices and E is a set of edges between the vertices $E = \{(u, v)|u, v \in V\}$ associated with it there is a weight function $w : E \to R$, where R denotes the set of all real numbers. Most of the times, the weight w_{ij} of the edge between nodes i and j represents the relevance of the connection. Usually, a larger weight corresponds to higher reliability of a connection. Weighted graphs are currently the most widely used networks throughout the field of bioinformatics.

  There is another graph representation called *Bipartite graph*. This graph is used to the modelling of biological networks range from the representation of enzyme-reaction links in metabolic pathways to ontologies or ecological connections. We don't give the detailed information about this graph since in our case we are not using this representation.

3.2.2 Matrix Representation

There are two main data structures used to store network graph representations:

- Adjacency matrix

  Given a graph $G = (V, E)$, the adjacency matrix representation consists of a $|V| \times |V| = n \times n$ matrix $A = (a_{ij})$, such that $a_{ij} = 1$, if $(i, j) \in V$, or $a_{ij} = 0$, otherwise.

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & \ddots & \vdots * \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

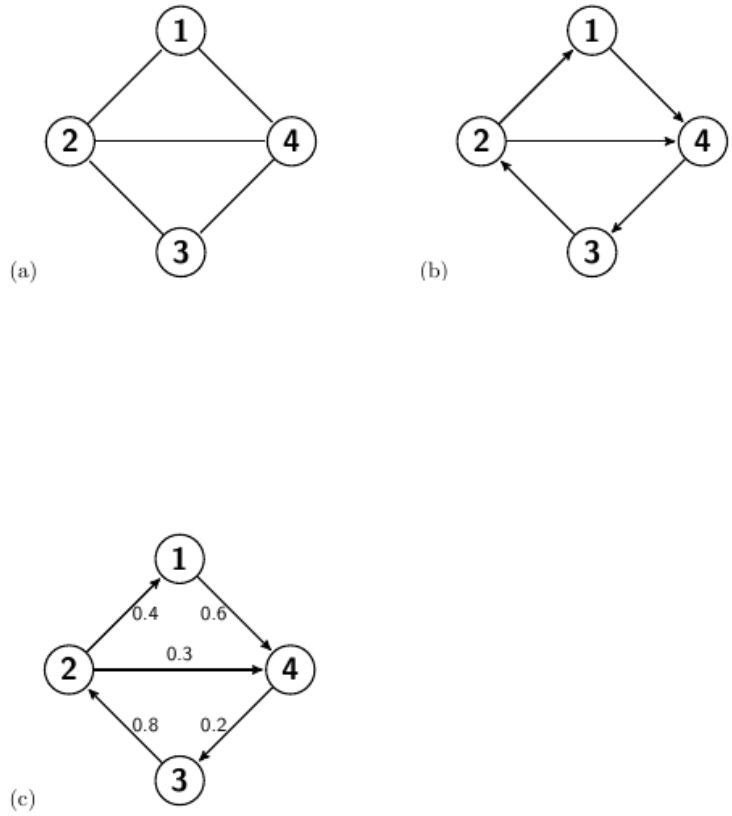


Figure 3.2: (a).Undirected Graph (b). Directed Graph (c). Weighted Graph

$n = |V|$. In the case where we have weighted graphs $a_{ij} = w_{ij}$ if $(i, j) \in V$ or $a_{ij} = 0$ otherwise. For undirected graphs the matrix is symmetric because $a_{ij} = a_{ji}$. We can apply linear algebra to this representation to analyse the network. That's why the matrix methods are widely used.

- Adjacent list

  Given a graph $G = (V, E)$, the adjacency list representation consists of an array Adj of $|E|$ elements, where for each $e \in E$, $Adj(0, e) = i \in V$. Adjacency lists require space $\Theta(|V| + |E|)$ and are preferable for sparse graphs with a low density of connections. An example of how these data structures represent a graph is given in Figure 3.3.

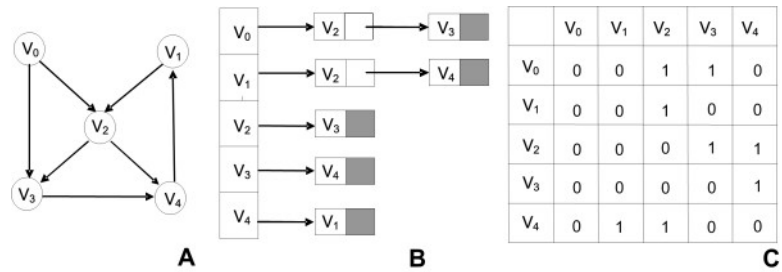


Figure 3.3: ***A. Directed Graph:*** *consisting of five nodes and six directed edges.* ***B. Adjacency List:*** *The data structure represents the directed graph using lists.* ***C. Adjacency Matrix:*** *represents the directed graph using a 2D matrix. The zeros represent the absence of the connection whereas the ones represent the existence of the connection between two nodes. The matrix is not symmetric since the graph is directed.*

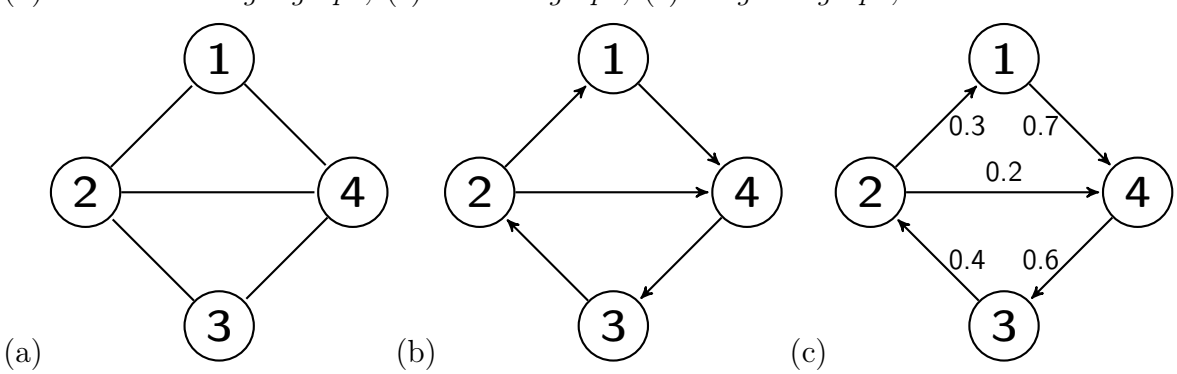
We give a simple example of an ecosystem and see how to represent the reactions in the graphical or in the matrix representation way.

Example 3.2.1. *Let us consider a simple ecosystem without connection with the external world. We suppose to have 4 species, where the reaction is eat:*

- *species 1 is eaten from species 2 and it eats species 4;*
- *species 2 is eaten from species 3 and it eats species 1 and 4;*
- *species 3 is eaten from species 4 and it eats species 2;*
- *species 4 is eaten from species 1 and 2, it eats species 3.*

1) **Graph representation:**

   (*a*) *undirected single graph,* (*b*) *directed graph,* (*c*) *weighted graph;*



2) **Matrix representation:**

$$A = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

3.3 Trophic Networks Analysis

Different network properties provides valuable informations about the internal organization of a biological network. In order to analyse an ecosystem efficiently, it is necessary to capture the important properties of the network. We give an brief introduction about the descriptive parameters of trophic networks.

Connectance is the proportion of possible links between species that are realized. In food webs, the level of connectance is related to the statistical distribution of the links per species. The distribution of links changes from (partial) power-law to exponential to uniform as the level of connectance increases. The definition of connectance is :

$$C = \frac{L}{S(S-1)} \tag{3.3.1}$$

where S is the number of species and L only the reactions of predation. In this case we have *connectance minimum*. Instead, if in L we consider also the reaction of competition (a reaction between organisms), we have *connectance maximum*. If we multiply the number of species S with the number of connectance maximum, we can obtain the *connectivity*:

$$S \cdot C = \frac{L}{S} = constant$$

3.3.1 Analysing Indirect Effects in Trophic Networks

Indirect effects are the effect of one node on another that is not directly adjacent to it in the network. Indirect effects are known to have significant influences in ecosystems and communities. The indirect effects at one node have the potential to have impacts on other components via paths of edges.

We can analyse the indirect influences in two categories following: *interaction chains* and *modifications*. Interaction chains occur when a change in one species affects another through a third species. For example, the death of tigers affects the erbs through the amount of hunting deer . Interaction modifications occur when a change in one species alters the interaction between two additional species. The born or death of tigers affects the amount of consumption between deer and erbs.

In general, indirect effects are known to be a powerful influence affecting ecosystem dynamics. They can result in overall positive interactions despite local negative direct interactions. Even when all direct interactions are known, indirect effects can complicate the outcome and reveal relationships between species. That's why understanding of indirect effects is necessary to understand ecosystems.

These indirect effects can be seen through the graphical or matrix representation, the latter is more efficient and easy to use. For analysing indirect effects we use a matrix called *matrix of dietary proportions* $[G]$. The elements of matrix are obtained from the elements of flow matrix $[T]$, and the input vector (X).

$$g_{ij} = \frac{T_{ij}}{T_{.j} + X_j} \tag{3.3.2}$$

g_{ij} represents that fraction which i comprises of the total intake by j. It considers the percentages that each dietary item i constitutes of the full intake by j. Each element T_{ij} represents the rate of a transfer from the species i to the species j . For the flow matrix $[T]$ we have following *mass balance equation*:

$$X_i + \sum_{j=1}^{n} T_{ji} = \sum_{k=1}^{n} T_{ik} + E_i + \Omega_i \tag{3.3.3}$$

where X_i represent the rate of the external input flow, E_i is the rate of output flow from taxon i to the outside world and Ω_i is the dissipation rate. Dissipation represents all the actions, as respiration or degradation, which generate a lower energetic form. For shorter time intervals, this balancing condition may not be verified, otherwise it is always announced in trophic networks.

The algebraic powers of matrix $[G]$ have a very important meaning. In Figure 3.4, we can see a simple network of four component.

The adjacency matrix and the flow matrix $[T]$ of this network are:

$$[A] = \begin{bmatrix} 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad [T] = \begin{bmatrix} 0 & t_{12} & t_{13} & t_{14} \\ 0 & 0 & t_{23} & t_{24} \\ 0 & 0 & 0 & t_{34} \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

The flow matrix $[T]$ is simply obtained from the adjacent matrix by changing the ones with the associated rates. Then we can obtain the dietary proportional matrix $[G]$.

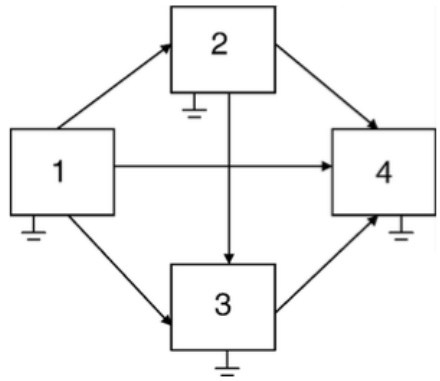


Figure 3.4: An example of network with four component

$$[G] = \begin{bmatrix} 0 & g_{12} & g_{13} & g_{14} \\ 0 & 0 & g_{23} & g_{24} \\ 0 & 0 & 0 & g_{34} \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Multiplying the matrix $[G]$ with it self, we get matrix $[G]^2$.

$$[G]^2 = \begin{bmatrix} 0 & 0 & g_{12}g_{23} & (g_{12}g_{24} + g_{13}g_{34}) \\ 0 & 0 & 0 & g_{23}g_{34} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

In the matrix $[G]^2$ each of the non-zero elements g_{ij}^2 corresponds to a pathway of length 2 that connect with i and j. In this case element (1,3) of $[G]^2$ reveals how much biomass gets from 1 to 3 over a single two-step pathway $1 \to 2 \to 3$. The element (1, 4) of $[G]^2$ is composed by two two-step pathways, the first term represents the path $1 \to 2 \to 4$ and the second one is the path $1 \to 3 \to 4$.

Next step we multiply matrix $[G]^2$ by matrix$[G]$ to obtain matrix$[G]^3$.

$$[G]^3 = \begin{bmatrix} 0 & 0 & 0 & (g_{12}g_{23}g_{34}) \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

There is just one non-zero elements in this matrix, and it corresponds to the three step pathway $1 \to 2 \to 3 \to 4$. If we continue to multiply matrix$[G]^3$ by matrix$[G]$ we can obtain

matrix $[G]^4$ where there are no non-zero elements (four step pathways).

In order analyse the indirect effects between the components in a trophic network, ecologists provided *Leontief structure matrix* $[S]$. It is a matrix which is thought to express the total direct and indirect flows between any two compartments of a system.

The elements of $[G]$ are normalized so that $g_{ij} \leq 1$, then the elements in the higher powers of [G] will grow progressively smaller. In 1949 Simon and Hawkins show that such series of matrices converges to a finite limit, and this limit is

$$\lim [I] + [G] + [G]^2 + [G]^3 + [G]^4 + \cdots \rightarrow [I - G]^{-1}$$

The limit, $[S] = [I - G]^{-1}$ called *Leontief structure matrix.* The (i, j)-th element of it reveals the fraction of the total input into j that left i and travelled over all paths of all lengths to satisfy a final demand upon j of one unit. The sum of each column of the structure matrix $[S]$ is related to the number of trophic transfers that a species i traverses during its flow to a given species j.

The matrix $[D]$, called *total dependency matrix* [11], is calculated using the Leontief structure matrix $[S]$ and the flow matrix $[T]$. The formula to calculate each element is:

$$d_{ij} = (s_{ij} - \delta_{ij}) \left(\frac{\sum_{k=1}^{n+2} T_{ik}}{s_{ij} \sum_{m=0}^{n} T_{mj}} \right)$$

where δ_{ij} is the element of the identity matrix I. The element d_{ij} represents how much of what arrives at j can be traced to a particular species i. The element T_{0j} represents the external inputs to j, $T_{i,(n+1)}$ are the usable exports from i to other comparable systems, and $T_{i,(n+2})$ are the dissipative losses from i. To represent all the information in the system, the indexes run from 0 to n+2.

The matrix $[D]$ in the column j contains the indirect diet of j, namely how much j eats from each element i in the ecosystem. It is useful for differentiating trophic roles.

All matrices and indexes described so far are used to describe only steady-state or temporally averaged networks. Several attempts have been made to expand I/O theory to treat time-varying systems. The main idea is to apply information-theoretic methods to a time series of network snapshots depicting the dynamics over an interval [10].

The trophic level of an organism is the position it occupies in a food chain. The concept of trophic level arise from the very simplistic image of an ecosystem as a trophic chain or pyramid. That is, trophic level 1 is characterised by primary producers such as plants, level 2 by herbivores, predators are at level 3 and carnivores or apex predators at level 4 or 5. In the real world the cases are more complicated than in a pyramid or in a chain, there are elements that occur in more than one trophic level, and no simple one-to-one mapping of taxa to integer trophic levels appears to be feasible.

In the previous section, it was mentioned how the sums of the columns of the structure matrix $[S]$, were related to the number of trophic transfers that the medium had incurred along its way to the target compartment. Although this sum is generally not an integer, Levine [11] suggested that it should be regarded as the average or effective trophic level at which that particular taxon is feeding. In the following we give a simple example to demonstrate this idea.

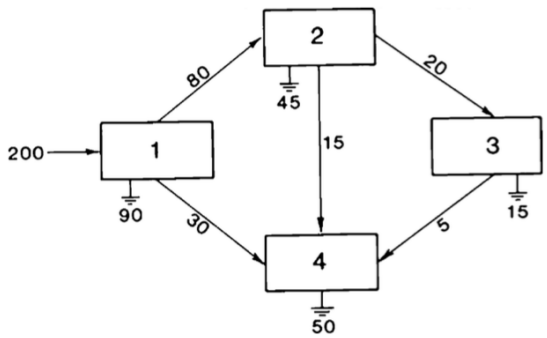


Figure 3.5: A hypothetical trophic network

Example 3.3.1. *(Taken from* [11]*) Let us consider an ecosystem represented in Figure 3.5. There are the compartments 1, 2, and 3 connected in chain-like fashion. Compartment 4 receives only 5 of its 50 units of total flow at the fourth trophic level. It receives* 30% *of its sustenance from the third level and* 60% *at the second level.*

From the given information we can infer that the compartment 4 is in the fourth trophic level only in the chain $1 \to 2 \to 3 \to 4$*. The compartment 4 receives 50 units and 5 given from the compartment 3. It represents the* 10% *of its sustenance. From the compartment 2, 15 units arrives as third trophic level in the chain* $1 \to 2 \to 4$ *and from the compartment 1, 30 units arrives as second trophic level in the chain* $1 \to 4$*. Its effective trophic level becomes :*

$$(0.6 \times 2) + (0.3 \times 3) + (0.1 \times 4) = 2.5.$$

In the structural matrix $[S]$*, the first three columns sum to 1.0, 2.0 and 3.0, respectively. The sum down the fourth column is 2.5 as just calculated.*

3.3.2 Cycling in Trophic networks

Cycling is one of the most important phenomena of ecosystems, in various researches it studied by the meaning of emery accounting. Cycling represents the energy fluxes between the species in a food chain, when there is a reutilization of a medium more than one time by

the same species in the system. In this case, a predator may become a prey. For example, in the real world Oxygen cycles indirectly through an ecosystem by the cycling of respiration, Carbon dioxide and Photosynthesis between the plants, air and humans. For finding the cycles one may perform a depth-first search with backtracking.

In trophic networks once mainly considers the energy fluxes generated by cycling. This index can be computed by combining the matrices that we introduced in the previous section. It calls *Finn Cycling index* FCI, it is used to compute the percentage of all fluxes in a cycling trophic network.

FCI computed by multiplying all the diagonal elements of the Leontief structure matrix $[S]$ with the rows of the flow matrix $[T]$ then sum it and divide it by the double summation of the flow matrix $[T]$ that is the total system through flow "$T_{..}$", the formula as following:

$$FCI = \frac{\sum_i [T_{i.} S_{ii}]}{T_{..}}$$

One of the problems of FCI is it does not account all flows engaged in recycling. Especially when there is a modification of the ecosystem, it is also changed even if it is a constant, this can not reflect the new status of an ecosystem. For this reason, during the analysis, the cycles are removed and then studied separately. Recently ecologists developed a new version of such index called the *comprehensive cycling index* (CCI), it considers all the fluxes generated by cycling [18]. CCI requires a large amount of time for its computation. It is a linear transformation of the FCI into the CCI (it is not present here for detail).

3.3.3 Dynamic Considerations

The time-series analysis can describe the actual dynamics of the species in a trophic network. A single snap-shot of the system can not describe the system behaviour in a more realistic way. Since the dynamics of the species vary over time, their composition also alters accordingly. In order to analyse the dynamics of the species in the system, it is necessary to consider the temporal dynamics of the species in a proper time scale.

The main approach to study dynamics of the trophic networks is based on Ordinary Differential Equation (ODEs). They describe all the interactions in an ecological system, modelling the variation of the species in a dynamic way. Some mathematical techniques are used to solve the system ODEs, the Lotka-Volterra predator prey model is a well known example for modelling system dynamics.

The Lotca-Volterra equations is a pair of first order, non-linear, differential equations. It used to analyse the dynamics of the species when they are interact with each other in an ecological system, namely describes a reaction "eat" between the predator and prey. It described as following:

$$\frac{dx}{dt} = \alpha x - \beta xy;$$
$$\frac{dy}{dt} = \delta xy - \gamma y;$$

where x, y are the number of prey and predator respectively, the derivatives $\frac{dx}{dt}$ and $\frac{dy}{dt}$ represent the growth rate of the species over the time t, the constant parameters α, β, δ, γ are represent the interaction of the species (prey and predator).

3.3.4 Tools for Analysing Trophic Networks

The most used tool for analysing ecological networks is Ecopath with Ecosim (EwE). It combines software for ecosystem mass balance analysis (Ecopath) with a dynamic modelling capability (Ecosim) to evaluate ecosystem effects of fishing, and to model effects of environmental changes.

With this software it is possible to do analyses in spite of missing data since it permits to calculate the missing data by using specific ecological equations representing properties of the data.

The Ecopath software has been under development for two decades, with Ecosim emerging in 1995, and Ecospace in 1998, leading to an integrated and widely applied package. Ecopath describes the system by using mass balance equations, which gives a static mass-balanced snapshot of the system. Ecosim studies the system behaviour using differential equations. Ecospace allows for a spatial and temporal dynamic modelling primarily designed for exploring impact and placement of protected areas.

Chapter 4

The Venice Lagoon Trophic Network

Marine ecosystems are fundamental food webs in the ecology system, they may have different structures. In this section we introduce the marine ecosystem of the Venice Lagoon which is a complex system in terms of its biological ecosystem and hydrodynamics. All the information in this section are taken from [11], [16].

4.1 Model structure of Venice Lagoon

The Venice Lagoon is situated in the north west of the Adriatic Sea. It covers an area of 50 by 20 km, with the major axis in the north-east, south-west direction. It is characterised by a few deep channels (maximum depth around 20 meters) surrounded by shallow water with an average depth of the order of 1 meter. The exchanges between the lagoon and the Adriatic Sea take place through three mouths, Lido, Malamocco and Chioggia, from 500 to 1000 meters wide and up to 20 meters deep.

The energy and mass flux estimation is a very important consideration in a marine ecosystem. It allows to calculate ecological indexes, it provides usable observations about the ecosystem behaviour and it may also evaluate the status of a marine ecosystem. The estimation of energy and mass flux is determined through steady-state linear food web models. They permit to derive ecological indexes from incomplete data sets.

The steady-state food web model of the Venice Lagoon presented in Sorokin and Giovanardi [11]. Field data used to construct the food web model were collected by Sorokin, from September 24 to October 11, 1993, at 36 sampling stations located in the lagoon. Sampling and analytical methodologies are described in detail in the original paper. The energy balance presented here was carried out for station 34, located in the Venice-Lido area. The compartmental structure of the food web model is shown in Figure 4.1.

The model represents a simplified benthic-pelagic food web, composed by 10 species, by

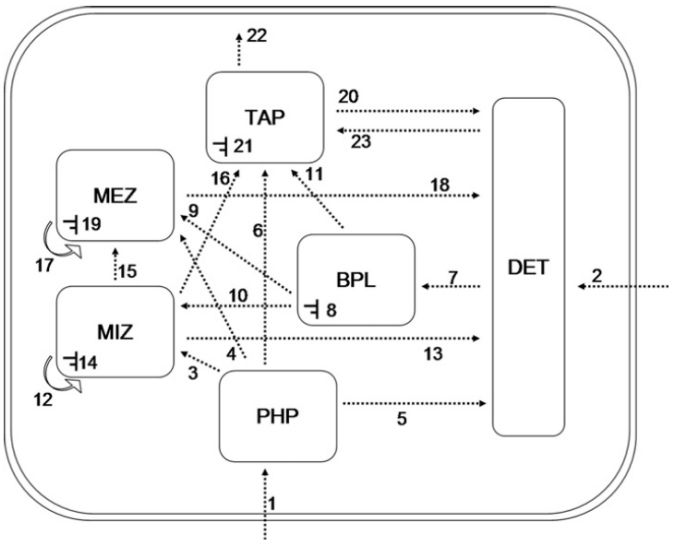


Figure 4.1: The food web model of Venice Lagoon [11]

the usage and effects of such species, we can divide them in two categories: internal and external species. Internal species are:

- **R. philippinarum**, it is the top predator in this food web, it is a predator of MIZ, PHP, BPL, DET. Represented by `TAP` in the model.
- **Phytoplankton**, it is at the lowest level of the model, eaten by MIZ, MEZ, TAP, represented by `PHP` in the model.
- **Mesozooplankton**, it is a predator of MIZ, PHP, BPL, represented by `MEZ` in the model.
- **Microzooplankton**, is a predator of PHP and BPL, eaten by MEZ and TAP. In the model is represented by `MIZ`.
- **Bacterioplankton**, it is a predator of DET and eaten by TAP, MEZ, MIZ, represented by `BPL` in the model

External species of the model are:

- **Organic Detritus**, these are the dead bodies of all other internal species, represented by `DET` in the model.
- **CO2**, it represents the respiration of all the internal species. Except for the species PHP, all the respiration reactions of the internal species are represented by the sympol ⊣ in the graphical model in figure 3.1.

- **Input**, it gives an input for the DET from the external world. Represented by `input` in the interactions of the model.
- **Harvesting**, it is the human interaction with the ecosystem, it harvests the species TAP. Represented by `Harvesting` in the interactions of the model.
- **Export**, it is the exit of the species DET from the ecosystem to the external world. For the simplicity of the model it is not represented in the graph but in the interactions of the model it is represented by `Export`.

The numbers on the arcs in the graph in Figure 4.1 represent the interaction numbers not the weight of the flows like in the toy case. There are 24 interactions within these species. These interactions are represented by the flow numbers in the model.

$1.CO_2 \rightarrow PHP$	$2.input \rightarrow Det$	$3.PHP \rightarrow MIZ$
$4.PHP \rightarrow MEZ$	$5.PHP \rightarrow DET$	$6.PHP \rightarrow TAP$
$7.DET \rightarrow BPL$	$8.BPL \rightarrow CO_2$	$9.BPL \rightarrow MEZ$
$10.BPL \rightarrow MIZ$	$11.BPL \rightarrow TAP$	$12.MIZ \rightarrow MIZ$
$13.MIZ \rightarrow DET$	$14.MIZ \rightarrow CO_2$	$15.MIZ \rightarrow MEZ$
$16.MIZ \rightarrow TAP$	$17.MEZ \rightarrow MEZ$	$18.MEZ \rightarrow Det$
$19.MEZ \rightarrow CO_2$	$20.TAP \rightarrow DET$	$21.TAP \rightarrow CO_2$
$22.TAP \rightarrow Harvesting$	$23.DET \rightarrow TAP$	$24.Det \rightarrow Export$

4.2 Model equations of Venice Lagoon

Ecologists reported the main equations of this model in [11]. These equations are listed in Figure 4.2 and they are based on the predator-prey relations and on the interactions of internal and external species in the system. In these equations the species at the left side of the right arrow ($\rightarrow$) indicates the prey (reactant), the species on the right side represents the predators (products). The plus ($+$) symbols indicate the different kinds of interactions in the equations. For example the equation

$$BPLLoss = BPL \rightarrow MEZ + BPL \rightarrow MIZ + BPL \rightarrow TAP$$

indicates that there are three different interactions in the loss of species BPL, the first one is MEZ eating BPL, the second one is MIZ eating BPL and the third one is TAP eating BPL.

We can see that there are at least two equations for each species in Figure 3.2. There are six terms in these equations namely Consumption, Loss, Feaces, Assimilation, Respiration and Production. From these equations we can conclude that:

Equations**

PHP

$PHPProduction = CO_2 \rightarrow PHP$

$PHPLoss = PHP \rightarrow MIZ + PHP \rightarrow MEZ + PHP \rightarrow DET + PHP \rightarrow TAP$

BPL

$BPLAssimilation = BPLConsumption$

$BPLRespiration = BPL \rightarrow CO_2$

$BPLProduction = BPLAssimilation - BPLRespiration$

$BPLLoss = BPL \rightarrow MEZ + BPL \rightarrow MIZ + BPL \rightarrow TAP$

MIZ

$MIZConsumption = PHP \rightarrow MIZ + BPL \rightarrow MIZ + MIZ \rightarrow MIZ$

$MIZFaeces = MIZ \rightarrow DET$

$MIZAssimilation = MIZConsumption - MIZFaeces$

$MIZRespiration = MIZ \rightarrow CO_2$

$MIZProduction = MIZAssimilation - MIZRespiration$

$MIZLoss = MIZ \rightarrow MIZ + MIZ \rightarrow MEZ + MIZ \rightarrow TAP$

MEZ

$MEZConsumption = PHP \rightarrow MEZ + BPL \rightarrow MEZ + MIZ \rightarrow MEZ + MEZ \rightarrow MEZ$

$MEZFaeces = MEZ \rightarrow DET$

$MEZAssimilation = MEZConsumption - MEZFaeces$

$MEZRespiration = MEZ \rightarrow CO_2$

$MEZProduction = MEZAssimilation - MEZRespiration$

$MEZLoss = MEZ \rightarrow MEZ$

DET

$DETConsumption = PHP \rightarrow DET + MIZ \rightarrow DET + MEZ \rightarrow DET + input \rightarrow DET$

$DETloss = DET \rightarrow BPL + DET \rightarrow TAP + DET \rightarrow Export$

TAP

$TAPConsumption = PHP \rightarrow TAP + BPL \rightarrow TAP + MIZ \rightarrow TAP + DET \rightarrow TAP$

$TAPFaeces = TAP \rightarrow DET$

$TAPAssimilation = TAPConsumption - TAPFaeces$

$TAPRespiration = TAP \rightarrow CO_2$

$TAPProduction = TAPAssimilation - TAPRespiration$

$TAPLoss = TAP \rightarrow Harvesting$

Figure 4.2: Model equations of food web model of Venice lagoon

- *Consumption* represents the input flow of the species that increases its biomass.
- *Loss* is the output flow of the species that decreases its biomass. It is usually represented by the interactions: eaten by other species, export to the external world or harvesting.
- *Feaces* is one of the output flows of the species MIZ and MEZ. It is represented by becoming detritus in the system. $XFeaces = X \rightarrow DET$ (where X represents a species).
- *Assimilation* is the amount of input biomass of the species. For the species X= BPL,

$$XAssimilation = XConsumption$$

For the species MIZ and MEZ it is the amount of biomass without the Feaces, it has the equation

$$XAssimilation = XConsumption - XFeaces$$

- *Respiration* of the species indicates the output flow of the species, usually represented as $XRespiration = X \rightarrow CO_2$. In the model, species PHP and DET do not have the Respiration.

- *Production* of a species indicate the interactions that produces such species, without respiration. It has the equation $XProduction = XAssimilation - XRespiration$.

One major problem in an ecological model is due to missing parameters. By composing the equations for each species one gets *the mass balance equation* of the system. This allows one to compute the missing data by using the residual error ϵ:

$$XProduction - XLoss = \epsilon$$

In the next chapter we use these equations to construct our Bio-PEPA model of Venice Lagoon.

4.3 Parameter estimation

The nine mass-balance equations (constraint equalities and their constraints) in Figure 4.3, are under determined, because they include 24 unknown energy fluxes. The system of equations and inequations is abstractly represented by the following formulation given in [11] .

$$\begin{cases} Ax = b + \epsilon \\ Ex = f \\ Gx \geq h \end{cases} \tag{4.3.1}$$

where the first row expresses the mass balance, second and third rows express respectively set equality and inequality constrains and ϵ is an error vector.

The ecologists developed a package in R called LIM [11], to infer the missing data in an ecological model, in our case the flows of biomass between the species. LIM package gets as input the linear equality and the inequality conditions which can be solved by least squares or by linear programming techniques [16]. The list of flows calculated by LIM for the food

Compartments	Constraint Equalities [cal m^{-2} day^{-1}]	Constraint Inequalities
PHP = phytoplankton	PHPProduction = 460	MIZRespiration > 0.2 * MIZConsumption
BPL = bacterioplankton	BPLProduction = 900	MEZRespiration > 0.2 * MEZConsumption
MIZ = microzooplankton	BPLRespiration = 2,800	TAPRespiration > 0.2 * TAPConsumption
MEZ = mesozooplankton	MIZProduction = MIZBiomass*0.8	TAPProduction < 0.3 * TAPConsumption
TAP = *R. philippinarum*	MEZProduction = MEZBiomass*0.04	TAPProduction > 0.09 * TAPConsumption
DET = Organic detritus	TAPProduction = TAPBiomass*P/B_{Rp}	PHP → MIZ > 0.1 * MIZConsumption
	MIZAssimilation = MIZConsumption*0.5	PHP → MEZ > 0.1 * MEZConsumption
Biomasses [cal m^{-2}]	MEZAssimilation = MEZConsumption*0.5	PHP → TAP > 0.1 * TAPConsumption
PHPBiomass = 160	TAPAssimilation = TAPConsumption*0.5	MIZ → MEZ > 0.1 * MEZConsumption
BPLBiomass = 1,600		MIZ → TAP > 0.1 * TAPConsumption
MIZBiomass = 170		BPL → MIZ > 0.1 * MIZConsumption
MEZBiomass = 3,000		BPL → MEZ > 0.1 * MEZConsumption
TAPBiomass = 1,463		BPL → TAP > 0.1 * TAPConsumption
		DET → TAP > 0.1 * TAPConsumption

Figure 4.3: Constraints of food web model of Venice lagoon

web model of Venice Lagoon is shows in Figure 4.4. The data obtained from LIM are used to built our dynamic Bio-PEPA model of the Venice Lagoon.

flux "X"	"unconstrained.solution"	"residualNorm"	"solutionNorm"	"IsError"	"type"
"CO2->PHP"	459.999999999999	460	1.18503681534321e-07	29947988.5043255	FALSE "ldei"
"input->DET"	2730.73765549052	1163.10298256538	1.18503681534321e-07	29947988.5043255	FALSE "ldei"
"PHP->MIZ"	57.0174942328742	-58.8070672478215	1.18503681534321e-07	29947988.5043255	FALSE "ldei"
"PHP->MEZ"	64.2022590519998	-86.0070672478185	1.18503681534321e-07	29947988.5043255	FALSE "ldei"
"PHP->DET"	330.652468950682	578.800734744705	1.18503681534321e-07	29947988.5043255	FALSE "ldei"
"PHP->TAP"	8.12777776444178	26.0133997509326	1.18503681534321e-07	29947988.5043255	FALSE "ldei"
"DET->BPL"	3700	3700	1.18503681534321e-07	29947988.5043255	FALSE "ldei"
"BPL->CO2"	2800	2800	1.18503681534321e-07	29947988.5043255	FALSE "ldei"
"BPL->MEZ"	393.618071624319	253.593177667083	1.18503681534321e-07	29947988.5043255	FALSE "ldei"
"BPL->MIZ"	449.487484161665	280.793177667081	1.18503681534321e-07	29947988.5043255	FALSE "ldei"
"BPL->TAP"	56.8944442140181	365.613644665836	1.18503681534321e-07	29947988.5043255	FALSE "ldei"
"MIZ->MIZ"	63.6699633102337	26.1265110004131	1.18503681534321e-07	29947988.5043255	FALSE "ldei"
"MIZ->DET"	285.087470852387	124.056310709837	1.18503681534321e-07	29947988.5043255	FALSE "ldei"
"MIZ->CO2"	149.087470852386	-11.9436892901624	1.18503681534321e-07	29947988.5043255	FALSE "ldei"
"MIZ->MEZ"	64.2022589886741	-1.07348899958419	1.18503681534321e-07	29947988.5043255	FALSE "ldei"
"MIZ->TAP"	8.12777770109122	110.946977999167	1.18503681534321e-07	29947988.5043255	FALSE "ldei"
"MEZ->MEZ"	119.999999999986	119.999999999986	1.18503681534321e-07	29947988.5043255	FALSE "ldei"
"MEZ->DET"	321.011294832502	143.256310709846	1.18503681534321e-07	29947988.5043255	FALSE "ldei"
"MEZ->CO2"	201.011294832481	23.2563107098243	1.18503681534321e-07	29947988.5043255	FALSE "ldei"
"TAP->DET"	40.6388889379052	-25.1066562889177	1.18503681534321e-07	29947988.5043255	FALSE "ldei"
"TAP->CO2"	33.3238889379071	-32.4216562889162	1.18503681534321e-07	29947988.5043255	FALSE "ldei"
"TAP->Harvesting"	7.31499999999366	7.31499999999413	1.18503681534321e-07	29947988.5043255	FALSE "ldei"
"DET->Export"	8.67733888298972e-07	-1163.10298256538	1.18503681534321e-07	29947988.5043255	FALSE "ldei"
"DET->TAP"	8.12777819625808	-552.787334993773	1.18503681534321e-07	29947988.5043255	FALSE "ldei"

Figure 4.4: Energy fluxes of Venice lagoon computed by LIM package

The model was calibrated and validated on field data collected at three Northern Adriatic lagoons, including the lagoon of Venice. The relationships between the anabolic and catabolic processes and water temperature were also modelled in [11].

Specific production rate is a very important input parameter of a food web model. Thermal condition have heavy effects on the metabolic rates. In order to estimate the specific

production of single species component, ecologists proposed dynamic individual bioenergetic models given by ODE's and they use "P/B" to represent the production rate [11]. In the steady-state food web model of the Venice Lagoon, they consider R.philippinarum (P/B_{RP}) as a specific species to estimate the effect of changes in metabolic rates induced by water temperature fluctuations on the food web model.

Chapter 5

Modelling Trophic Networks with Bio-PEPA

In the previous chapters we presented trophic networks and Bio-PEPA, a process algebra which we intend to use for modelling and analysing trophic networks in a continuous way.

In this chapter we recall and discuss the previous work of Moscardo, whose thesis [10] explores how to model trophic networks in a stochastic way with Bio-PEPA. He considered first Wood, a toy case of trophic networks, and then the Venice Lagoon.

5.1 Description of Wood

Wood is a toy ecosystem. It is composed by three internal species, namely *carnivores*, *herbivores* and *grass*. There are also three external species that interact with those internal species: respiration of the internal species produces CO_2, their death produces *Detritus* and they can be hunted by humans becoming *Hunt*.

5.1.1 Interactions between the species

The interactions between the species represent the flow of biomass. All the interactions within the internal species are expressed by the action "eat", interactions between the internal species and CO_2 are expressed by the action "breath" and interactions between the internal species with Detritus are expressed by the action "die". Now we introduce all the interactions among species in this ecology system.

- Carnivores are the top predator of this ecosystem. They `Eat` herbivores and `Breath` to produce CO_2 and after `Death` become Detritus, hence there are three flow of biomass related to Carnivores, indicated with c_Eat, c_Breath and c_Death:

$$Herbivores \xrightarrow{c\_Eat} Carnivores;$$
$$Carnivores \xrightarrow{c\_Breath} CO_2;$$
$$Carnivores \xrightarrow{c\_Death} Detritus;$$

- Herbivores can `Eat` both grass and detritus, and also have actions `Breath` to produce CO_2 and `Death` to become Detritus. In order to distinguish the two kinds of eating, in the interaction they are represented with the respective number 1 and 2, hence there are four flow of biomass related to Herbivores e_Eat1, e_Eat2, e_Breath and e_Death:

$$Detritus \xrightarrow{e\_Eat1} Herbivores;$$
$$Grass \xrightarrow{e\_Eat2} Herbivores;$$
$$Herbivores \xrightarrow{e\_Breath} CO_2;$$
$$Herbivores \xrightarrow{e\_Death} Detritus;$$

- Grass is eaten by other internal species, it has the interaction `Growth` by consuming CO_2 and it becomes Detritus after its `Death`, hence there are two flows of biomass starting from the species Grass:

$$CO_2 \xrightarrow{g\_Growth} Grass;$$
$$Grass \xrightarrow{g\_Death} Detritus;$$

- Hunt is an external entity in this ecosystem, `Hunt` can capture the carnivores and herbivores. There are two flows of biomass, from the carnivores to Hunt and from the herbivores to Hunt.

$$Carnivores \xrightarrow{c\_Hunt} Hunt;$$
$$Herbivores \xrightarrow{e\_Hunt} Hunt;$$

CO_2 and Detritus are non living external compartments of the system. Since they can't perform any action actively, they just increase their mass with the breathing or dying of other internal species, and decrease their mass consumed by internal species. For example CO_2 produced by the breathing of carnivores and herbivores, is consumed by the grass. Detritus is consumed by herbivores or grass and produced by their death.

5.1.2 Parameters of the system

From the description above, we see that there are 11 interactions in this ecosystem. Each species has its initial biomass and each action has its weight representing the flow of biomass.

All the input and output flows of biomass are balanced in each internal species because of the steady-state properties of a food web model. The weight of each interaction is given in the following:

- The weight of action c_Eat is 8;
- The weight of action c_Hunt is 1;
- The weight of action c_Breath is 2;
- The weight of action c_Death is 5;
- The weight of action e_Eat1 is 9;
- The weight of action e_Eat2 is 10;
- The weight of action e_Hunt is 1;
- The weight of action e_Breath is 4;
- The weight of action e_Death is 6;
- The weight of action g_Death is 10;
- The weight of action $growth$ is 20.

The initial biomass of the internal species is proportional to the metabolic flows. For the external species, the initial biomass is determined by their mass balance situations in the system. For example, if the output flow is greater than the input flow, the biomass should be very large. If the species have only input flows the initial biomass should be 1. We have the following initial biomass for the species:

- The initial biomass of Carnivores is 10;
- The initial biomass of Herbivores is 20;
- The initial biomass of Grass is 100;
- The initial biomass of CO_2 is 10000;
- The initial biomass of Detritus is 100;
- The initial biomass of Hunt is 1.

5.1.3 Graphical representation of Wood

The graphical representation of Wood, uses a directed weighted graph that we introduced in section 3.2.1. Nodes represent the internal and external species in the system, arcs represent the interactions between the species and weights represent the fluxes. The direction of the arcs represent the direction of the flow of biomass through the species when an interaction happens.

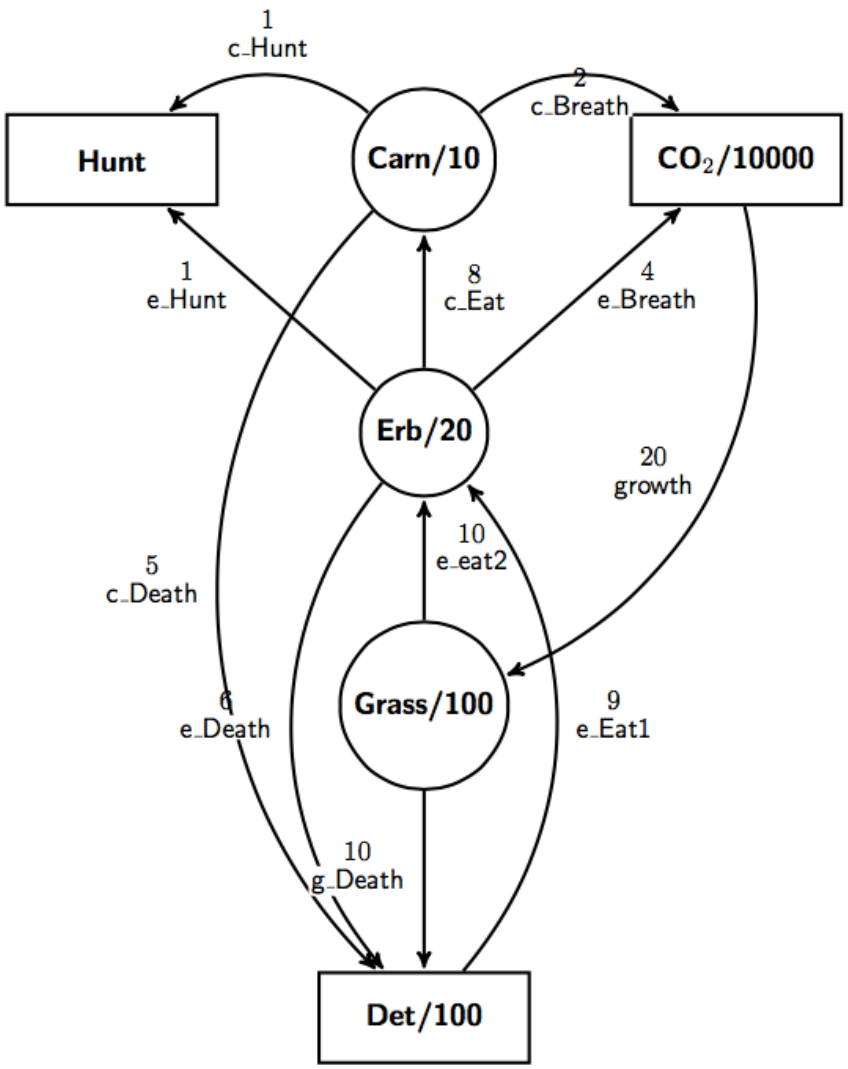


Figure 5.1: Wood graph representation

For simplicity the nodes are labelled by abbreviations of the names of the species. `Carn` represents the species Carnivores, `Erb` represents the species Herbivores, `Det` represents the species Detritus.

5.2 Previous works on Wood

Moscardo used the Wood toy case in his thesis [10]. He defined many PEPA and Bio-PEPA stochastic models of Wood and discussed their properties. In his thesis there are 8 models of Wood, four of them are modelled in PEPA and four in Bio-PEPA. In this thesis we just recall his Bio-PEPA models, since our work is mainly focused on modelling in Bio-PEPA.

In a Bio-PEPA model functional rates and stoichiometry coefficients can be used. Moscardo defined the functional rates in his Wood models in two ways. One considers them as constants and the other computes them as a function of the amounts of the species. When the functional rate is constant, the action happens always at the same rate, independently from the amount of species involved. But in the real world this is impossible, because when the population of a species increases or decreases there must be effect on the eating rate of this species. A more realistic case is when the functional rate is a function related to the amounts of the species, the probability of the action α depends on the product of the amounts of the involved populations and a constant γ_α. This corresponds to use the predefined Bio-PEPA function fMA.

Example 5.2.1. (taken from [10]) *If we have two species and species $S1$ eats species $S2$, in Bio-PEPA we can model them with the sequential components:*

$$S1 \stackrel{\text{def}}{=} (eat, 1) \uparrow S1$$
$$S2 \stackrel{\text{def}}{=} (eat, 1) \downarrow S2$$

and the model component

$$M \stackrel{\text{def}}{=} S1 \underset{*}{\bowtie} S2$$

If we consider the functional rate as a constant, we have

$$f_{eat} = k_{eat}$$

If we consider the functional rate as a function of the amounts of the species we have

$$f_{eat} = \gamma_{eat} \cdot S1 \cdot S2$$

where γ_{eat} represents the probability that the action eat happens when $S1$ and $S2$ meet. $S1 \cdot S2$ represents the number of possible meetings of the two species. If there are more than two predators for one prey, there is a competition between this predators based on their functional rates.

The stoichiometry coefficients are the second important factors of a Bio-PEPA model. These coefficients are constant. They may represent the flow of biomass in a trophic network.

The stoichiometry coefficients can describe the amount of each different kind of species produced from actions.

In Moscardo's thesis there are four stochastic Bio-PEPA models for the Wood ecosystem that differ on the stoichiometry and functional rates.

- The Bio-PEPA model Wood5 is defined with constant rate and with unitary stoichiometry coefficients. The functional rates are defined as $f_\alpha = W_\alpha$, α is the name of action and W_α is the weight of the flow. There is no set of parameter $\mathcal{K}$. The model component is a cooperation between the species with their initial amount of biomasses.

- The Bio-PEPA model Wood6 is defined with a constant rate equal to 1 and the stoichiometry coefficients represent the flows in Figure 4.1. The set of parameter $\mathcal{K}$ is defined as $\{f = 1\}$, since all the rates are equal to one.

- The Bio-PEPA model Wood7 is defined with unitary stoichiometry, functional rates are function of the flows and of the amounts of biomass of the species, that is they follow the Mass Action law.

  The set of parameters $\mathcal{K}$ is defined by the constant values γ_α, computed as the ratio between the flow and the initial amounts of biomass of the internal species involved in each action. CO_2 and Detritus are modelled as infinitely available. The model component is defined as the cooperation of the six components in the ecosystem with their initial biomasses.

  There exist an unbalance problem in this model, but it has more realistic sequences of actions in comparison to other models. Static analysis of this model shows that, there are no invariant state because the amount of biomass of the internal species does not remain constant. In the dynamic analysis, i.e. stochastic simulation of the system, the species Grass continues to grow and other species are not do any action. This is because the functional rate of species Grass is much greater than the functional rate of other species. This indicates that the ranges of functional rates affect heavily the behaviour of the system.

- In the model Wood8, the functional rates follow the Mass Action law and depend on the biomasses of the species. The stoichiometry coefficients correspond to the flows. The formula of the functional rates are modified accordingly.

  All the components are defined as in Wood7 and the static and dynamic analysis of the model is similar to the one of Wood7 with the same problems caused by the range of the functional rates.

In the case of Wood8, although the resulting analysis is similar to the one of Wood7, the stoichiometry coefficients in the model strongly limit the execution of actions. Actions to be executable require the availability of the biomass indicated by the stoichiometry coefficients for each interacting component. Then the initial biomasses of species could be not sufficient for the system dynamics. If we use stoichiometry coefficients to represent the flows we must change the model component.

Moscardo in his thesis concludes that the best stochastic model of Wood in Bio-PEPA is Wood7, because it models the system in a more realistic way.

In the following sections we start from Wood7, since it is the more realistic model within these models, we introduce new ideas to improve it and we analyse the model in a continuous way.

5.3 Bio-PEPA model of Wood7

We recall here the model Wood7 given in [10]. We also recall the results of the stochastic analysis given in [10]. Then we apply the continuous analysis to Wood7.

5.3.1 Set of compartment $\mathcal{V}$

Since in this model we represent only the biomass, there is just one compartment, defined as:

$$Wood : 1\ cal\ m^{-2}$$

5.3.2 Set of quantities $\mathcal{N}$

The set of quantities $\mathcal{N}$ contains the definition of all the species in the ecosystem. We have three internal and three external species in our system. For simplicity we illustrate only the species Carnivores.

- Species component name $C = Carn$
- Step size $H = 1\ cal\ m^{-2}$
- Minimum level $N = 0\ cal\ m^{-2}$
- Maximum concentration $M = 50000\ cal\ m^{-2}$

- Name of the enclosing compartment $Wood$
- The unit for concentration is 1 $cal\ m^{-2}$

The definition of the species Carnivores is:

$$Carn : 1, 0, 50000, Wood, cal\ m^2$$

Similarly we can define all other elements:

$$\begin{gathered}\mathcal{N} = \{Carn : 1, 0, 50000, Wood, cal\ m^2 \\ Erb : 1, 0, 50000, Wood, cal\ m^2 \\ Grass : 1, 0, 50000, Wood, cal\ m^2 \\ Hunt : 1, 0, 50000, Wood, cal\ m^2 \\ CO : 1, 0, 50000, Wood, cal\ m^2 \\ Det : 1, 0, 50000, Wood, cal\ m^2\}\end{gathered}$$

5.3.3 Set of functional rates $\mathcal{F}_{\mathcal{R}}$

The functional rate $\mathcal{F}_{\mathcal{R}}$ is defined as a function of species level. In order to compute the functional rates, the species that affect an interaction must be identified. The functional rate of an interaction involving internal species depends on the probability to meet between the species involved. We have the following functional rates for actions involving the internal species:

- the functional rate of action c_Eat: $f_{c\_Eat} = \gamma_{c\_Eat} \cdot Carn \cdot Erb$;
- the functional rate of action c_Breath: $f_{c\_Breath} = \gamma_{c\_Breath} \cdot Carn$;
- the functional rate of action c_Death: $f_{c\_Death} = \gamma_{c\_Death} \cdot Carn$;
- the functional rate of action e_Eat2: $f_{e\_Eat2} = \gamma_{e\_Eat2} \cdot Erb \cdot Grass$;
- the functional rate of action e_Breath: $f_{e\_Breath} = \gamma_{e\_Breath} \cdot Erb$;
- the functional rate of action e_Death: $f_{e\_Death} = \gamma_{e\_Death} \cdot Erb$;
- the functional rate of action g_Death: $f_{g\_Death} = \gamma_{g\_Death} \cdot Grass$;

For actions involving the external species we have following functional rates:

- the functional rate of action $growth$: $f_{growth} = \gamma_{growth} \cdot Grass$;

- the functional rate of action c_Hunt: $f_{c\_Hunt} = \gamma_{c\_Hunt} \cdot Carn$;
- the functional rate of action e_Hunt: $f_{e\_Hunt} = \gamma_{e\_Hunt} \cdot Erb$;
- the functional rate of action e_Eat1: $f_{e\_Eat1} = \gamma_{e\_Eat} \cdot Erb$;

5.3.4 Set of parameters $\mathcal{K}$

The constant γ_α represents the probability that the action α happens. If the action involves two internal species, the constant value is computed as the ratio between the biomass flow of the action and the product between the initial biomass of the two species, for example r_{c_Eat} is defined as:

$$\gamma_{c\_Eat} = \frac{w_{c\_Eat}}{Carn_0 \cdot Erb_0}$$

If there is just one internal species involved, the constant is defined as the ratio between the flow of the action and the initial amount of the species, for example r_{c_Breath} is defined as:

$$\gamma_{c\_Breath} = \frac{w_{c\_Breath}}{Carn_0}$$

We have the following constant parameters:

- the probability that the Carnivores eat the Herbivores is equal to: $\gamma_{c\_Eat} = \frac{8}{10 * 20} = 0.04day$;
- the probability that Carnivores breat is equal to: $\gamma_{c\_Breath} = \frac{2}{10} = 0.2day$;
- the probability that Herbivores breath is equal to: $\gamma_{e\_Breath} = \frac{4}{20} = 0.2day$;
- the probability that the hunter hunts the Carnivores is equal to: $\gamma_{c\_Hunt} = \frac{1}{10} = 0.1day$;
- the probability that the hunter hunts the Herbivores is equal to: $\gamma_{e\_Hunt} = \frac{1}{20} = 0.05day$;
- the probability that Carnivores die is equal to: $\gamma_{c\_Death} = \frac{5}{10} = 0.5day$;
- the probability that Herbivores die is equal to: $\gamma_{e\_Death} = \frac{6}{20} = 0.3day$;
- the probability that Grass die is equal to: $\gamma_{g\_Death} = \frac{10}{100} = 0.1day$;
- the probability that Grass grows is equal to: $\gamma_{growth} = \frac{20}{100} = 0.2day$;

- the probability that Herbivores eat Detritus is equal to: $\gamma_{e\_Eat1} = \dfrac{9}{20} = 0.45day$;
- the probability that Herbivores eat Grass is equal to: $\gamma_{e\_Eat2} = \dfrac{10}{20 * 100} = 0.005day$.

Hence the set of parameters is defined as: $\mathcal{K} = \{\gamma_{c\_Eat} = 0.04day,\ \gamma_{c\_Hunt} = 0.1day,\ \gamma_{c\_Breath} = 0.2day,\ \gamma_{c\_Death} = 0.5day,\ \gamma_{e\_Eat1} = 0.45day,\ \gamma_{e\_Eat2} = 0.005day,\ \gamma_{e\_Hunt} = 0.05day,\ \gamma_{e\_Breath} = 0.2day,\ \gamma_{e\_Death} = 0.3day,\ \gamma_{g\_Death} = 0.1day,\ \gamma_{g\_growth} = 0.2day\}$

5.3.5 The model component P

The model component is defined as:

$$Wood7 \stackrel{\text{def}}{=} Carn[10] \underset{*}{\bowtie} Erb[20] \underset{*}{\bowtie} Grass[100] \underset{*}{\bowtie} CO[100] \underset{*}{\bowtie} Hunt[1] \underset{*}{\bowtie} Det[100]$$

5.3.6 Set of sequential component $Comp$

Since we assume that there is infinite availability of CO_2 and Detritus, they con not decrease in the interaction, $\odot$ operator is used to represent it and the initial level of CO_2 can be reduced to 100. The set of sequential components is defined as in Figure 5.2.

Hence the sequential components of the system *Wood7* are defined as:

$$\begin{aligned}
Carn &\stackrel{def}{=} (c_Eat, 1) \uparrow Carn + (c_Hunt, 1) \downarrow Carn + \\
&\quad + (c_Breath, 1) \downarrow Carn + (c_Death, 1) \downarrow Carn \\
Erb &\stackrel{def}{=} (c_Eat, 1) \downarrow Erb + (e_Hunt, 1) \downarrow Erb + (e_Breath, 1) \downarrow Erb + \\
&\quad + (e_Death, 1) \downarrow Erb + (e_Eat1, 1) \uparrow Erb + (e_Eat2, 1) \uparrow Erb \\
Grass &\stackrel{def}{=} (e_Eat2, 1) \downarrow Grass + (g_Death, 1) \downarrow Grass + \\
&\quad + (growth, 1) \uparrow Grass \\
Hunt &\stackrel{def}{=} (c_Hunt, 1) \odot Hunt + (e_Hunt, 1) \odot Hunt \\
Det &\stackrel{def}{=} (c_Death, 1) \uparrow Det + (e_Death, 1) \uparrow Det + (g_Death, 1) \uparrow Det + \\
&\quad + (e_Eat1, 1) \odot Det \\
CO &\stackrel{def}{=} (c_Breath, 1) \uparrow CO + (e_Breath, 1) \uparrow CO + (growth, 1) \odot CO
\end{aligned}$$

and the model component is:

$$Wood7 \stackrel{def}{=} Carn[10] \underset{*}{\bowtie} Erb[20] \underset{*}{\bowtie} Grass[100] \underset{*}{\bowtie} CO[100] \underset{*}{\bowtie} Hunt[1] \underset{*}{\bowtie} Det[100]$$

Figure 5.2: The sequential components of Wood7

In Moscardo's experiment, he used the Gillespie's stochastic simulation algorithm of Bio-PEPA tool to do the discrete stochastic analysis. The model was simulated 100 times in order to reduce the differences between the simulations. He obtained the following results:

- Static analysis. The Invariant Inference was performed. From the state invariant analysis he obtained that there is no invariant states, because the amount of biomass of the internal species does not remain constant when the reactions growth and e_Eat1 are performed. In fact when they are performed, the sum of all the amount of species increases of one unit of biomass. The interaction growth causes an unitary growth of biomass for Grass and the amount of CO remains constant. This holds also for the interaction e_Eat1. Instead for the interaction c_Hunt and e_Hunt it is the contrary, namely the total amount decreases. The activities invariants do not exist.

- Dynamic analysis. The time-series analysis was performed, namely the species Grass continues to grow and the other species go to zero (figure 5.3). This is due to the fact that the functional rate of growth increases because the amount of Grass increases and the other functional rates, corresponding to the other interactions, are not comparable. The other functional rates are much smaller than growth and for this reason the other interactions are performed rarely.

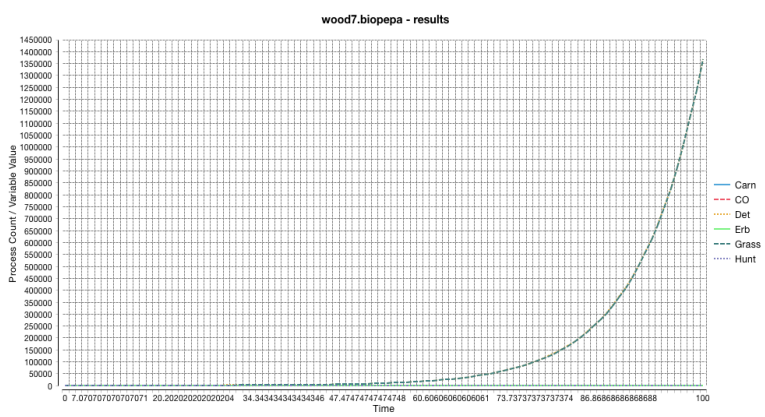


Figure 5.3: The stochastic analysis of Wood7

5.3.7 Continuous analysis of Wood7

Now we use Ordinary Differential Equation (ODEs) generated by Bio-PEPA for Wood7 to do *continuous analysis* of this model. The time-series analysis have been performed and the results are obtained as in Figure 5.4. All the external species, namely CO, Det, Hunt are continuously increasing in the time space and all the internal species are stable. If we focus

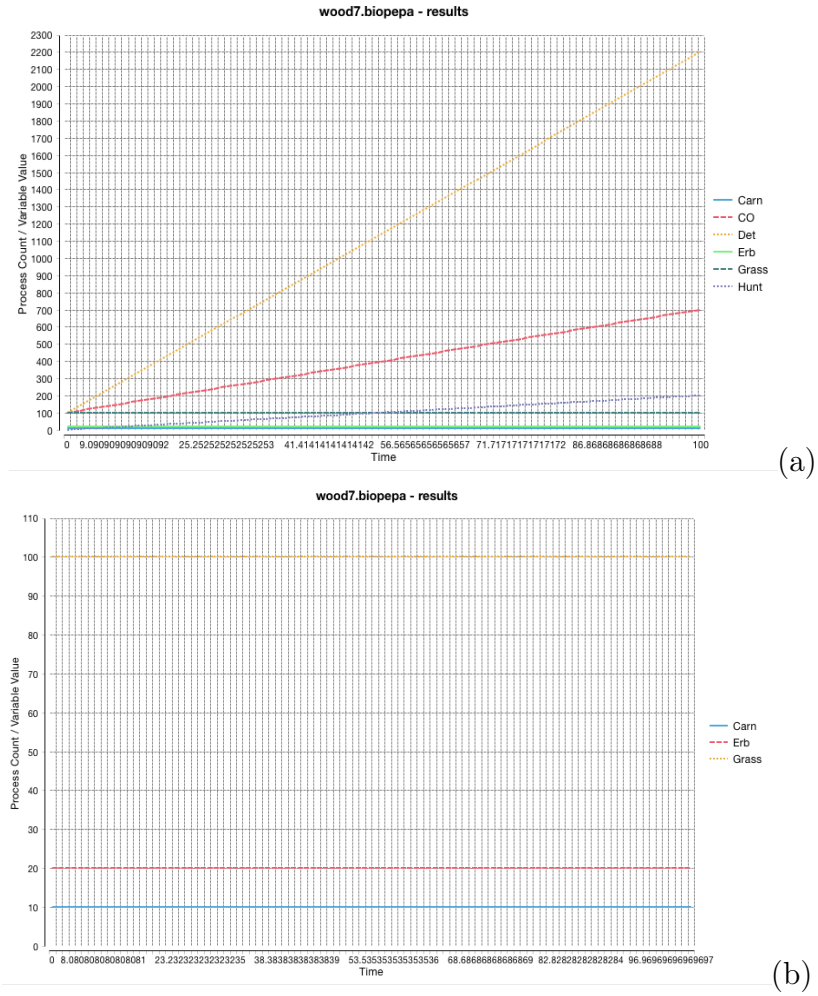


Figure 5.4: Continuous dynamic analysis of Wood7

only on the internal species we can see the result more clearly in Figure 5.4 (b).

From the results above we can see that Wood7 behaves reasonably in the continuous simulation. In particular internal species are at steady state, as they are expected to be. This is due to the fact that the continuous simulation corresponds to averaging many stochastic simulations and it avoids the variation due to the stochastic behaviour. In particular we can see that even if the functional rate of the species Grass is much greater than the one of other species, it doesn't affect the system behaviour at all. From this we can conclude that, the range of the functional rates does not have a dramatic impact in the continuous dynamic analysis of the system as much as in the stochastic case.

5.4 Modified model of Wood7: decomposing Grass

In order to get a more realistic result from the stochastic simulation, we decided to modify Wood7 and decompose the species with rates ranging in a larger interval. The functional rate of Grass is much higher than the one of other internal species and in the time-series analysis Grass is continuously increasing, inhibiting other actions. Hence in our modified model we split the species Grass in ten species.

All the the initial values of biomass remain the same except for the species Grass. Since its initial biomass was 100 and we decompose it in ten equal parts, each initial value of these decomposed parts is equal to 10. Figure 5.5 shows the weighted graph of this model, called wood7". For simplicity, we denote "G1,G2,...G10" to be the names of splitted species of Grass in the graph.

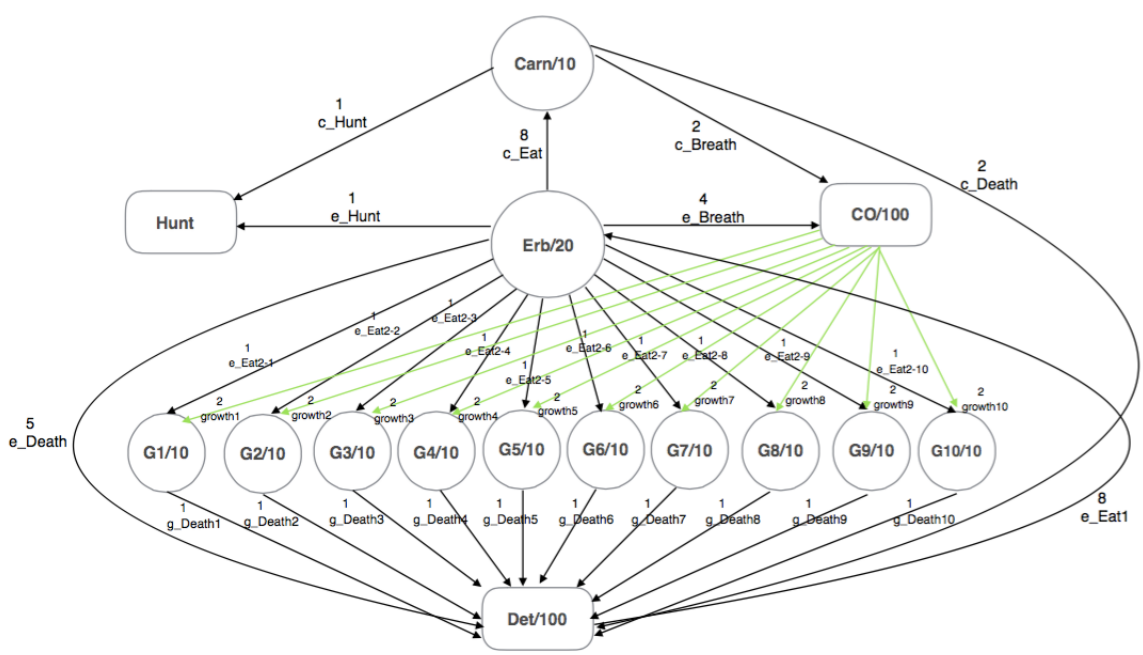


Figure 5.5: Weighted graph of modified model Wood7'

Let us present the modified Bio-PEPA model Wood7'.

5.4.1 Set of compartments $\mathcal{V}$

The set of compartment $\mathcal{V}$ doesn't change.

$$Wood : 1\ cal\ m^{-2}$$

5.4.2 Set of quantities $\mathcal{N}$

The set of quantities defined for all the other species does not change, but there are ten species of Grass. We have the set of quantities $\mathcal{N}$ =

$$\begin{gathered}
\{Carn : 1, 0, 50000, Wood, cal\ m^2 \\
Erb : 1, 0, 50000, Wood, cal\ m^2 \\
Grass1 : 1, 0, 50000, Wood, cal\ m^2 \\
Grass2 : 1, 0, 50000, Wood, cal\ m^2 \\
Grass3 : 1, 0, 50000, Wood, cal\ m^2 \\
Grass4 : 1, 0, 50000, Wood, cal\ m^2 \\
Grass5 : 1, 0, 50000, Wood, cal\ m^2 \\
Grass6 : 1, 0, 50000, Wood, cal\ m^2 \\
Grass7 : 1, 0, 50000, Wood, cal\ m^2 \\
Grass8 : 1, 0, 50000, Wood, cal\ m^2 \\
Grass9 : 1, 0, 50000, Wood, cal\ m^2 \\
Grass10 : 1, 0, 50000, Wood, cal\ m^2 \\
Hunt : 1, 0, 50000, Wood, cal\ m^2 \\
CO : 1, 0, 50000, Wood, cal\ m^2 \\
Det : 1, 0, 50000, Wood, cal\ m^2\}
\end{gathered}$$

5.4.3 Set of functional rates

The set of functional rates of Wood7' is modified due to the splitting of biomass of Grass but it is based on the set of functional rates of Wood7. It is defined as follows:

$f_{c\_Eat} = \gamma_{c\_Eat} \cdot Carn \cdot Erb;$	$f_{c\_Breath} = \gamma_{c\_Breath} \cdot Carn;$
$f_{c\_Death} = \gamma_{c\_Death} \cdot Carn;$	$f_{e\_Eat2-1} = \gamma_{e\_Eat2-1} \cdot Erb \cdot Grass1;$
$f_{e\_Eat2-2} = \gamma_{e\_Eat2-2} \cdot Erb \cdot Grass2;$	$f_{e\_Eat2-3} = \gamma_{e\_Eat2-3} \cdot Erb \cdot Grass3;$
$f_{e\_Eat2-4} = \gamma_{e\_Eat2-4} \cdot Erb \cdot Grass4;$	$f_{e\_Eat2-5} = \gamma_{e\_Eat2-5} \cdot Erb \cdot Grass5;$
$f_{e\_Eat2-6} = \gamma_{e\_Eat2-6} \cdot Erb \cdot Grass6;$	$f_{e\_Eat2-7} = \gamma_{e\_Eat2-7} \cdot Erb \cdot Grass7;$
$f_{e\_Eat2-8} = \gamma_{e\_Eat2-8} \cdot Erb \cdot Grass8;$	$f_{e\_Eat2-9} = \gamma_{e\_Eat2-9} \cdot Erb \cdot Grass9;$
$f_{e\_Eat2-10} = \gamma_{e\_Eat2-10} \cdot Erb \cdot Grass10;$	$f_{e\_Breath} = \gamma_{e\_Breath} \cdot Erb;$
$f_{e\_Death} = \gamma_{e\_Death} \cdot Erb;$	$f_{g\_Death1} = \gamma_{g\_Death1} \cdot Grass1;$
$f_{g\_Death2} = \gamma_{g\_Death2} \cdot Grass2;$	$f_{g\_Death3} = \gamma_{g\_Death3} \cdot Grass3;$
$f_{g\_Death4} = \gamma_{g\_Death4} \cdot Grass4;$	$f_{g\_Death5} = \gamma_{g\_Death5} \cdot Grass5;$
$f_{g\_Death6} = \gamma_{g\_Death6} \cdot Grass6;$	$f_{g\_Death7} = \gamma_{g\_Death7} \cdot Grass7;$
$f_{g\_Death8} = \gamma_{g\_Death8} \cdot Grass8;$	$f_{g\_Death9} = \gamma_{g\_Death9} \cdot Grass9;$
$f_{g\_Death10} = \gamma_{g\_Death10} \cdot Grass10;$	

For the external species we have following functional rates:

$f_{growth1} = \gamma_{growth1} \cdot Grass1;$	$f_{growth2} = \gamma_{growth2} \cdot Grass2;$
$f_{growth3} = \gamma_{growth3} \cdot Grass3;$	$f_{growth4} = \gamma_{growth4} \cdot Grass4;$
$f_{growth5} = \gamma_{growth5} \cdot Grass5;$	$f_{growth6} = \gamma_{growth6} \cdot Grass6;$
$f_{growth7} = \gamma_{growth7} \cdot Grass7;$	$f_{growth8} = \gamma_{growth8} \cdot Grass8;$
$f_{growth9} = \gamma_{growth9} \cdot Grass9;$	$f_{growth10} = \gamma_{growth10} \cdot Grass10;$
$f_{c\_Hunt} = \gamma_{c\_Hunt} \cdot Carn;$	$f_{e\_Hunt} = \gamma_{e\_Hunt} \cdot Erb;$
$f_{e\_Eat1} = \gamma_{e\_Eat} \cdot Erb;$	

5.4.4 Set of parameters $\mathcal{K}$

The computation of the constant γ_α of the modified model follows the same ideas as the one in Wood7. Hence we have the following constant parameters:

$\gamma_{c\_Eat} = 8/(10 * 20) = 0.04day;$	$\gamma_{c\_Breath} = 2/10 = 0.2day;$
$\gamma_{e\_Breath} = 4/20 = 0.2day;$	$\gamma_{c\_Hunt} = 1/10 = 0.1day;$
$\gamma_{e\_Hunt} = 1/20 = 0.05day;$	$\gamma_{c\_Death} = 5/10 = 0.5day;$
$\gamma_{e\_Death} = 6/20 = 0.3day;$	$\gamma_{g\_Death1} = 1/10 = 0.1day;$
$\gamma_{g\_Death2} = 1/10 = 0.1day;$	$\gamma_{g\_Death3} = 1/10 = 0.1day;$
$\gamma_{g\_Death4} = 1/10 = 0.1day;$	$\gamma_{g\_Death5} = 1/10 = 0.1day;$
$\gamma_{g\_Death6} = 1/10 = 0.1day;$	$\gamma_{g\_Death7} = 1/10 = 0.1day;$
$\gamma_{g\_Death8} = 1/10 = 0.1day;$	$\gamma_{g\_Death9} = 1/10 = 0.1day;$
$\gamma_{g\_Death10} = 1/10 = 0.1day;$	$\gamma_{growth1} = 2/10 = 0.2day;$
$\gamma_{growth2} = 2/10 = 0.2day;$	$\gamma_{growth3} = 2/10 = 0.2day;$
$\gamma_{growth4} = 2/10 = 0.2day;$	$\gamma_{growth5} = 2/10 = 0.2day;$
$\gamma_{growth6} = 2/10 = 0.2day;$	$\gamma_{growth7} = 2/10 = 0.2day;$
$\gamma_{growth8} = 2/10 = 0.2day;$	$\gamma_{growth9} = 2/10 = 0.2day;$
$\gamma_{growth10} = 2/10 = 0.2day;$	$\gamma_{e\_Eat1} = 9/20 = 0.45day;$
$\gamma_{e\_Eat2-1} = 1/(20 * 10) = 0.005day;$	$\gamma_{e\_Eat2-2} = 1/(20 * 10) = 0.005day;$
$\gamma_{e\_Eat2-3} = 1/(20 * 10) = 0.005day;$	$\gamma_{e\_Eat2-4} = 1/(20 * 10) = 0.005day;$
$\gamma_{e\_Eat2-5} = 1/(20 * 10) = 0.005day;$	$\gamma_{e\_Eat2-6} = 1/(20 * 10) = 0.005day;$
$\gamma_{e\_Eat2-7} = 1/(20 * 10) = 0.005day;$	$\gamma_{e\_Eat2-8} = 1/(20 * 10) = 0.005day;$
$\gamma_{e\_Eat2-9} = 1/(20 * 10) = 0.005day;$	$\gamma_{e\_Eat2-10} = 1/(20 * 10) = 0.005day;$

The set of parameters is defined as: $\mathcal{K} = \{\gamma_{c\_Eat} = 0.04day,\ \gamma_{c\_Hunt} = 0.1day,\ \gamma_{c\_Breath} = 0.2day,\ \gamma_{c\_Death} = 0.5day,\ \gamma_{e\_Eat1} = 0.45day,\ \gamma_{e\_Eat2-1} = 0.005day,\ \gamma_{e\_Eat2-2} = 0.005day,\ \gamma_{e\_Eat2-3} =$

$0.005day$, $\gamma_{e\_Eat2-4} = 0.005day$, $\gamma_{e\_Eat2-5} = 0.005day$, $\gamma_{e\_Eat2-6} = 0.005day$, $\gamma_{e\_Eat2-7} = 0.005day$, $\gamma_{e\_Eat2-8} = 0.005day$, $\gamma_{e\_Eat2-9} = 0.005day$, $\gamma_{e\_Eat2-10} = 0.005day$ $\gamma_{e\_Hunt} = 0.05day$, $\gamma_{e\_Breath} = 0.2day$, $\gamma_{e\_Death1} = 0.3day$, $\gamma_{g\_Death2} = 0.1day$, $\gamma_{g\_Death3} = 0.1day$, $\gamma_{g\_Death4} = 0.1day$, $\gamma_{g\_Death5} = 0.1day$, $\gamma_{g\_Death6} = 0.1day$, $\gamma_{g\_Death7} = 0.1day$, $\gamma_{g\_Death8} = 0.1day$, $\gamma_{g\_Death9} = 0.1day$, $\gamma_{g\_Death10} = 0.1day$, $\gamma_{g\_growth1} = 0.2day$ $\gamma_{g\_growth2} = 0.2day$ $\gamma_{g\_growth3} = 0.2day$ $\gamma_{g\_growth4} = 0.2day$ $\gamma_{g\_growth5} = 0.2day$ $\gamma_{g\_growth6} = 0.2day$ $\gamma_{g\_growth7} = 0.2day$ $\gamma_{g\_growth8} = 0.2day$ $\gamma_{g\_growth9} = 0.2day$ $\gamma_{g\_growth10} = 0.2day\}$

5.4.5 Set of sequential components $Comp$

The sequential components of the modified system are also based on the ideas in Wood7, we just add the splitted species into them, the resulting components are more complex.

- $Carn \stackrel{\text{def}}{=} (c\_eat, 1) \uparrow Carn + (c\_breath, 1) \downarrow Carn + (c\_hunt, 1) \downarrow Carn + (c\_death, 1) \downarrow Carn$;
- $Erb \stackrel{\text{def}}{=} (e\_eat1, 1) \uparrow Erb + (e\_eat2 - 1, 1) \uparrow Erb + (e\_eat2 - 2, 1) \uparrow Erb + (e\_eat2 - 3, 1) \uparrow Erb + (e\_eat2 - 4, 1) \uparrow Erb + (e\_eat2 - 5, 1) \uparrow Erb + (e\_eat2 - 6, 1) \uparrow Erb + (e\_eat2 - 7, 1) \uparrow Erb + (e\_eat2 - 8, 1) \uparrow Erb + (e\_eat2 - 9, 1) \uparrow Erb + (e\_eat2 - 10, 1) \uparrow Erb + (e\_breath, 1) \downarrow Erb + (c\_eat, 1) \downarrow Erb + (e\_hunt, 1) \downarrow Erb + (e\_death, 1) \downarrow Erb$;
- $Grass1 \stackrel{\text{def}}{=} (growth1, 1) \uparrow Grass1 + (e\_eat2 - 1, 1) \downarrow Grass1 + (g\_death1, 1) \downarrow Grass1$;
- $Grass2 \stackrel{\text{def}}{=} (growth2, 1) \uparrow Grass2 + (e\_eat2 - 2, 1) \downarrow Grass2 + (g\_death2, 1) \downarrow Grass2$;
- $Grass3 \stackrel{\text{def}}{=} (growth3, 1) \uparrow Grass3 + (e\_eat2 - 3, 1) \downarrow Grass3 + (g\_death3, 1) \downarrow Grass3$;
- $Grass4 \stackrel{\text{def}}{=} (growth4, 1) \uparrow Grass4 + (e\_eat2 - 4, 1) \downarrow Grass4 + (g\_death4, 1) \downarrow Grass4$;
- $Grass5 \stackrel{\text{def}}{=} (growth5, 1) \uparrow Grass5 + (e\_eat2 - 5, 1) \downarrow Grass5 + (g\_death5, 1) \downarrow Grass5$;
- $Grass6 \stackrel{\text{def}}{=} (growth6, 1) \uparrow Grass6 + (e\_eat2 - 6, 1) \downarrow Grass6 + (g\_death6, 1) \downarrow Grass6$;
- $Grass7 \stackrel{\text{def}}{=} (growth7, 1) \uparrow Grass7 + (e\_eat2 - 7, 1) \downarrow Grass7 + (g\_death7, 1) \downarrow Grass7$;

- $Grass8 \stackrel{\text{def}}{=} (growth8, 1) \uparrow Grass8 + (e\_eat2 - 8, 1) \downarrow Grass8 + (g\_death8, 1) \downarrow Grass8;$

- $Grass9 \stackrel{\text{def}}{=} (growth9, 1) \uparrow Grass9 + (e\_eat2 - 9, 1) \downarrow Grass9 + (g\_death9, 1) \downarrow Grass9;$

- $Grass10 \stackrel{\text{def}}{=} (growth10, 1) \uparrow Grass10 + (e\_eat2 - 10, 1) \downarrow Grass10 + (g\_death10, 1) \downarrow Grass10;$

- $Det \stackrel{\text{def}}{=} (g\_death1, 1) \uparrow Det + (g\_death2, 1) \uparrow Det + (g\_death3, 1) \uparrow Det + (g\_death4, 1) \uparrow Det + (g\_death5, 1) \uparrow Det + (g\_death6, 1) \uparrow Det + (g\_death7, 1) \uparrow Det + (g\_death8, 1) \uparrow Det + (g\_death9, 1) \uparrow Det + (g\_death10, 1) \uparrow Det + (c\_death, 1) \uparrow Det + (e\_death, 1) \uparrow Det + (e\_eat1, 1) \odot Det;$

- $CO \stackrel{\text{def}}{=} (growth1, 1) \odot CO + (growth2, 1) \odot CO + (growth3, 1) \odot CO + (growth4, 1) \odot CO + (growth5, 1) \odot CO + (growth6, 1) \odot CO + (growth7, 1) \odot CO + (growth8, 1) \odot CO + (growth9, 1) \odot CO + (growth10, 1) \odot CO + (e\_breath, 1) \uparrow CO + (c\_breath, 1) \uparrow CO;$

- $Hunt \stackrel{\text{def}}{=} (c\_hunt, 1) \odot Hunt + (e\_hunt, 1) \odot Hunt;$

5.4.6 The model component P

The model component is defined as:

$Wood7' \stackrel{\text{def}}{=} Carn[10] \underset{*}{\bowtie} Erb[20] \underset{*}{\bowtie} Grass1[10] \underset{*}{\bowtie} Grass2[10] \underset{*}{\bowtie} Grass3[10] \underset{*}{\bowtie} Grass4[10] \underset{*}{\bowtie} Grass5[10] \underset{*}{\bowtie} Grass6[10] \underset{*}{\bowtie} Grass7[10] \underset{*}{\bowtie} Grass8[10] \underset{*}{\bowtie} Grass9[10] \underset{*}{\bowtie} Grass10[10] \underset{*}{\bowtie} CO[100] \underset{*}{\bowtie} Hunt[1] \underset{*}{\bowtie} Det[100]$

5.4.7 Analysis on Wood7'

In the dynamic analysis, Gillespie's Stochastic simulation is performed and we obtain the result shown in Figure 5.6. In Figure 5.6 (b) we omit the Detritus which grows very fast in order to stress the other species behaviour. Although the splitted species obtained from Grass continue to increase, the other internal species are not becoming extinct. This is due to the fact that we changed the range of the functional rates. In order to see the other internal species behaviour more clearly, in Figure 5.7 we show only the species Carn and Erb.

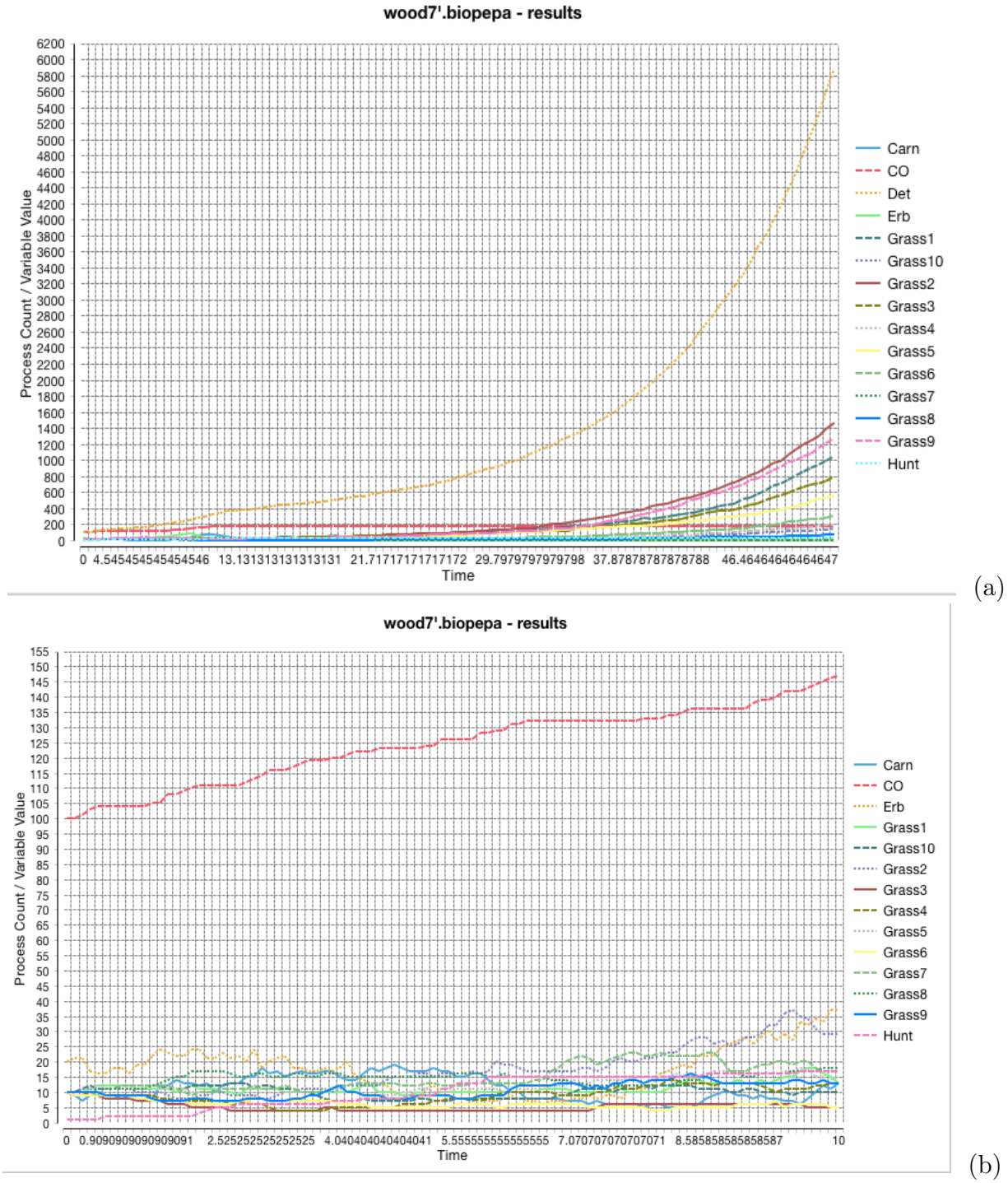


Figure 5.6: Stochastic analysis of modified model Wood7'

The continuous dynamic analysis of Wood7' performed with the Ordinary Differential Equation simulator, is shown in Figure 5.8. The behaviour of the internal species in the modified model is shown in Figure 5,9. In the continuous simulation there is not much difference with the continuous simulation of Wood7 shown in Figure 5.4. We can conclude

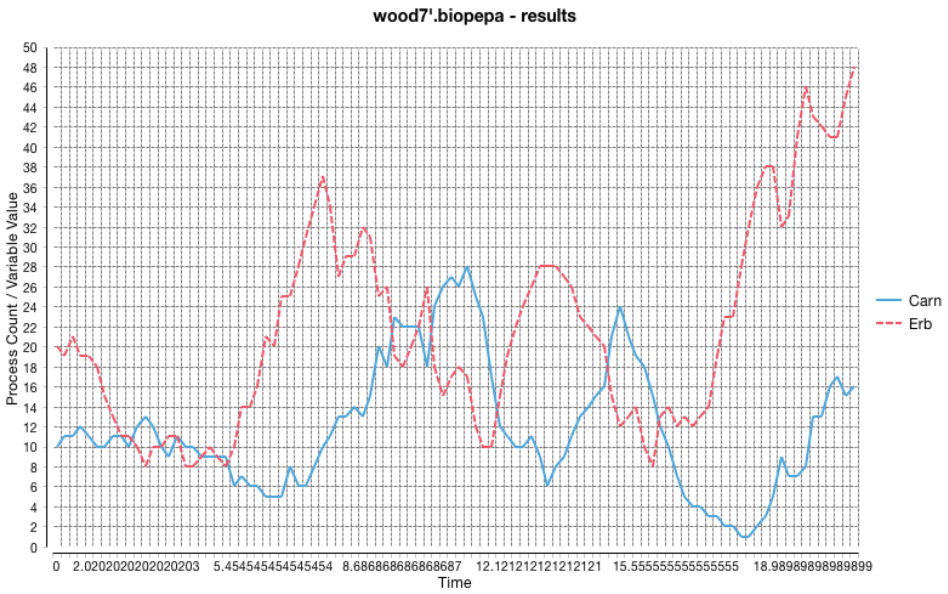


Figure 5.7: Dynamics of internal species in the stochastic simulation of the modified model Wood7'

that the range of the functional rates does not affect so much the continuous dynamics of the system behaviour as in the stochastic case.

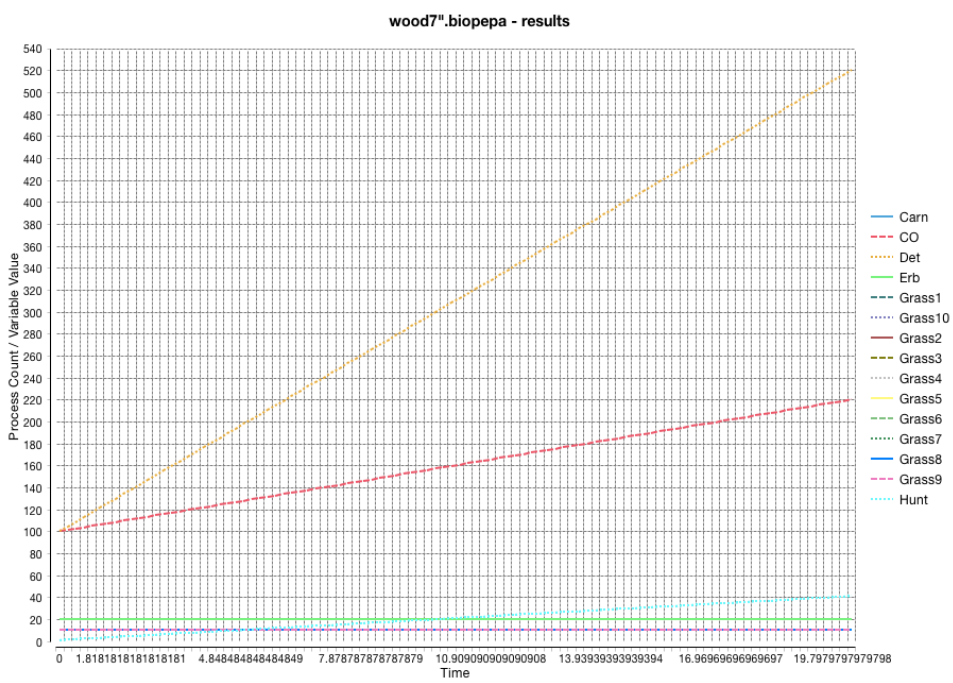


Figure 5.8: Continuous dynamic analysis of the modified model Wood7'

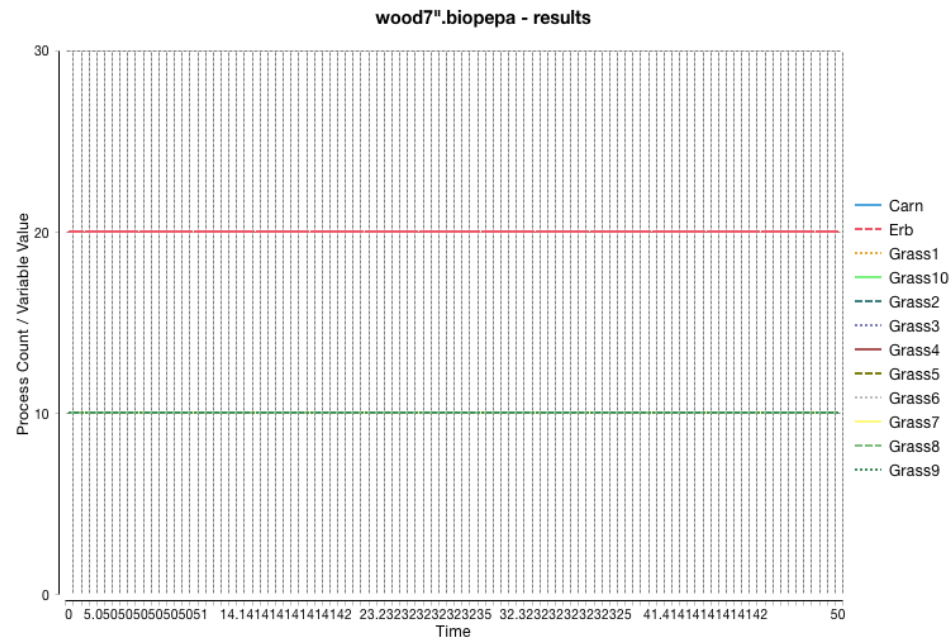


Figure 5.9: Continuous dynamic analysis of the modified model Wood7': behaviour of internal species. Since the initial values of the internal species Carn and Grasss are qual to 10, they are overlapping at level 10.

5.4.8 Experiment on modified model Wood7': CO_2 can decrease

As in Wood7, in our modified model wood7' the species CO_2 is infinitely available and never decreasing. In this experiment, the initial biomass of CO_2 is not changed and it is set to 100. We set that CO_2 can be decreased by the Grass growth and we analyse the system behaviour on both the stochastic and the continuous simulation. We changed the sequential component CO as follows :

$$CO \stackrel{def}{=} (growth1, 1) \downarrow CO + (growth2, 1) \downarrow CO + (growth3, 1) \downarrow CO + (growth4, 1) \downarrow +(growth5, 1) \downarrow CO + (growth6, 1) \downarrow +(growth7, 1) \downarrow CO + (growth8, 1) \downarrow CO + (growth9, 1) \downarrow CO + (growth10, 1) \downarrow CO + (e\_breath, 1) \uparrow CO + (c\_breath, 1) \uparrow CO;$$

- Stochastic analysis is performed by Gillespie stochastic algorithm. After running the simulator 50 times, we obtain the result in Figure 5.10. The figure shows that, decreasing the species CO affects the behaviour of other internal species. Species Erb is more dependent from CO, when CO decreases the species Erb also decreases. While the other species are oscillating until CO gets to zero. Figure 5.10 (b) shows the behaviour of the internal species when CO decreases.

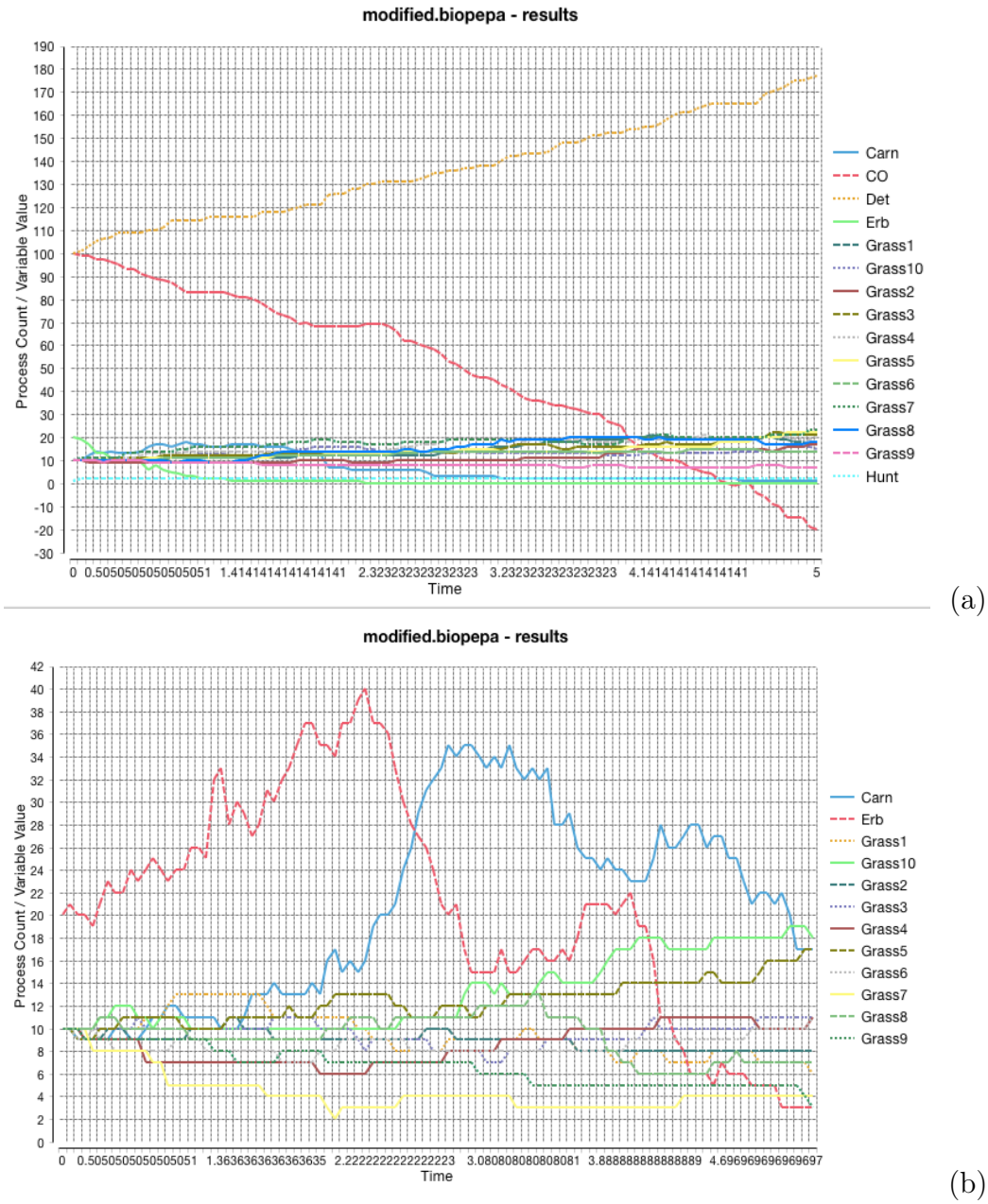
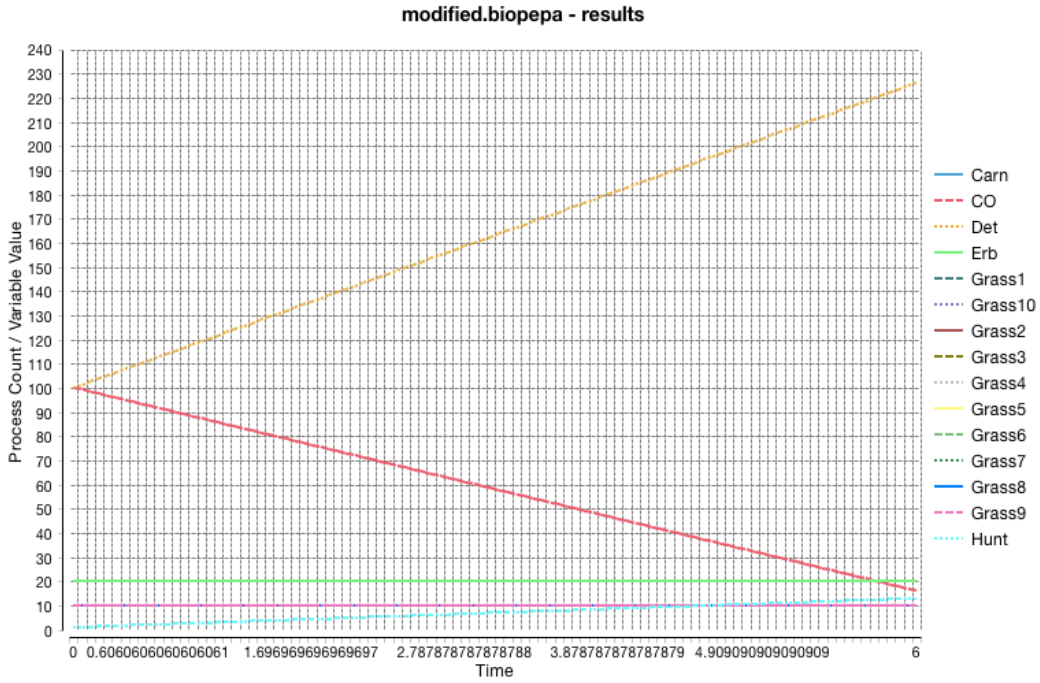


Figure 5.10: Stochastic analysis in the modified model Wood7' when CO_2 can be decreased

- Continuous analysis performed by translation into ODEs. The system behaviour is shown in Figure 5.11. The figure shows that, until the species CO reaches 0, all the internal species are stable. This indicates that, in the continuous dynamics of the system, the amount if CO does not affect the behaviour of the internal species until it goes to zero. Figure 5.11 (b) shows the behaviour of internal species, the result is same as in Figure 5.9.



(a)

(b)

Figure 5.11: Continuous analysis of system behaviour when CO_2 can be decreased in modified model Wood7"

Chapter 6

Modelling the Venice Lagoon with Bio-PEPA

From the experiments on Wood that we presented in the previous chapter, we can summarize that defining the functional rates and the stoichiometry coefficients is a very crucial work in Bio-PEPA modelling of a trophic network. The range of the functional rates has a great impact on the stochastic simulation, but it is not so destructive on the continuous dynamics of the system.

In this chapter, the Venice Lagoon ecosystem is modelled with Bio-PEPA. We base our model on the previous work done by Moscardo in his master thesis [10]. We choose the model Wood7 as a template for modelling a trophic network, since it is the more realistic model among the cases studied in Moscardo's thesis.

In particular we intend to explore continuous modelling with ODE's as offered by Bio-PEPA. In the case of Wood, continuous modelling was rather arbitrary, since a discrete representation of the system seems to be more appropriate and realistic, due to the small number of individuals in the considered species. On the contrary, when dealing with the Venice Lagoon, a continuous representation of the fluxes is totally justified and natural.

6.1 Model of Venice Lagoon in Bio-PEPA

In Chapter 3, we introduced the equations and parameter estimations of the Venice Lagoon ecosystem. In this section we use those information to construct our Bio-PEPA model.

We give the list of flows and the initial biomasses of the species as specified in [11] and

shown in Figure 4.3. The flows are taken from Figure 4.4 as computed by LIM. In the model we use action names instead of the interaction numbers shown in the graphic representation in Figure 4.1. "1" is represented by the action name "growth", ”2" is represented by the action name "input" and so on, action names of the species MIZ and MEZ are represented by m_α and M_α respectively (α is action name).

Action name	Interaction	flow $(cal\ m^{-2} day^{-1})$	Action name	Interaction	flow $(cal\ m^{-2} day^{-}$
growth	$CO_2 \rightarrow PHP$	460	m_death	$MIZ \rightarrow DET$	285
input	$input \rightarrow DET$	2731	m_breath	$MIZ \rightarrow CO_2$	149
m_eat1	$PHP \rightarrow MIZ$	57	M_eat3	$MIZ \rightarrow MEZ$	64
M_eat1	$PHP \rightarrow MEZ$	64	t_eat3	$MIZ \rightarrow TAP$	8
p_death	$PHP \rightarrow DET$	331	M_eat4	$MEZ \rightarrow MEZ$	120
t_eat1	$PHP \rightarrow TAP$	8	M_death	$MEZ \rightarrow DET$	321
b_eat1	$DET \rightarrow BPL$	3700	M_breath	$MEZ \rightarrow CO_2$	201
b_breath	$BPL \rightarrow CO_2$	2800	t_death	$TAP \rightarrow DET$	41
M_eat2	$BPL \rightarrow MEZ$	394	t_breath	$TAP \rightarrow CO_2$	33
m_eat2	$BPL \rightarrow MIZ$	449	harvest	$TAP \rightarrow Harvesting$	7
t_eat2	$BPL \rightarrow TAP$	57	t_eat4	$DET \rightarrow TAP$	8
m_eat3	$MIZ \rightarrow MIZ$	64	export	$DET \rightarrow Export$	$8/10^7$

The initial values of the species biomass in the ecosystem are:

$$
\begin{array}{cc}
PHP_0 = 160\ cal\ m^{-2} & BPL_0 = 1600\ cal\ m^{-2} \\
MIZ_0 = 170\ cal\ m^{-2} & MEZ_0 = 3000\ cal\ m^{-2} \\
TAP_0 = 1463\ cal\ m^{-2} & DET_0 = 5000\ cal\ m^{-2} \\
CO2_0 = 5000\ cal]m^{-2} & Harvesting_0 = 1\ cal\ m^{-2} \\
Input_0 = 3000\ cal\ m^{-2} & Export_0 = 1\ cal\ m^{-2}
\end{array}
$$

Internal species biomass is as specified in Figure 4.3. We fixed DET_0, $CO2_0$ and $Input_0$ to large amount of biomass, since they are external species in input to the system and we assume to have enough of them to feed the system.

6.1.1 Set of Compartments $\mathcal{V}$

We just state only one compartment:

$$Lagoon : 1\ cal\ m^{-2}$$

6.1.2 Set of Quantities $\mathcal{N}$

The set of quantities $\mathcal{N}$ contains the definition of all the species in the ecosystem. We have six internal and four external species in our system. For the sake of simplicity, we show the detail only for the species PHP.

- Species component name $C = PHP$
- Step size $H = 1\ cal\ m^{-2}$
- Minimum level $N = 0\ cal\ m^{-2}$
- Maximum concentration $M = 50000\ cal\ m^{-2}$
- Name of the enclosing compartment $Lagoon$
- The unit for concentration is $1\ cal\ m^{-2}$

The definition of species PHP is:

$$PHP : 1, 0, 50000, Lagoon, cal\ m^2$$

Accordingly we can set all other elements and we have the set of quantities:

$$\begin{gathered}\mathcal{N} = \{PHP : 1, 0, 50000, Lagoon, cal\ m^2 \\ BPL : 1, 0, 50000, Lagoon, cal\ m^2 \\ MIZ : 1, 0, 50000, Lagoon, cal\ m^2 \\ MEZ : 1, 0, 50000, Lagoon, cal\ m^2 \\ TAP : 1, 0, 50000, Lagoon, cal\ m^2 \\ CO_2 : 1, 0, 50000, Lagoon, cal\ m^2 \\ Harvesting : 1, 0, 50000, Lagoon, cal\ m^2 \\ Export : 1, 0, 50000, Lagoon, cal\ m^2 \\ Input : 1, 0, 50000, Lagoon, cal\ m^2 \\ DET : 1, 0, 50000, Lagoon, cal\ m^2\}\end{gathered}$$

6.1.3 Set of Functional Rates $\mathcal{F}_{\mathcal{R}}$

Functional rates are computed in the same style of Wood7. We consider the mass action law and the functional rates depend on the internal species biomass. The external species biomass do not have any effect on the rates. For each reaction we have the following functional rates:

Interaction	Functional rate	Interaction	Functional rate
growth	$f_{growth} = \gamma_{growth} \cdot PHP$	m_death	$f_{m\_death} = \gamma_{m\_death} \cdot MIZ$
input	$f_{input} = \gamma_{input}$	m_breath	$f_{m\_breath} = \gamma_{m\_breath} \cdot MIZ$
m_eat1	$f_{m\_eat1} = \gamma_{m\_eat1} \cdot PHP \cdot MIZ$	M_eat3	$f_{M\_eat3} = \gamma_{M\_eat3} \cdot MEZ \cdot MIZ$
Meat1	$f_{M\_eat1} = \gamma_{Meat1} \cdot PHP \cdot MEZ$	t_eat3	$f_{t\_eat3} = \gamma_{t\_eat3} \cdot MIZ \cdot TAP$
p_death	$f_{p\_death} = \gamma_{p\_death} \cdot PHP$	M_eat4	$f_{M\_eat4} = \gamma_{M\_eat4} \cdot MEZ \cdot MEZ$
t_eat1	$f_{t\_eat1} = \gamma_{t\_eat1} \cdot PHP \cdot TAP$	M_death	$f_{M\_death} = \gamma_{M\_death} \cdot MEZ$
b_eat1	$f_{b\_eat1} = \gamma_{b\_eat1} \cdot BPL$	M_breath	$f_{M\_breath} = \gamma_{M\_breath} \cdot MEZ$
b_breath	$f_{b\_breath} = \gamma_{b\_breath} \cdot BPL$	t_death	$f_{t\_death} = \gamma_{t\_death} \cdot TAP$
M_eat2	$f_{Meat2} = \gamma_{Meat2} \cdot BPL \cdot MEZ$	t_breath	$f_{t\_breath} = \gamma_{t\_breath} \cdot TAP$
m_eat2	$f_{m\_eat2} = \gamma_{m\_eat2} \cdot BPL \cdot MIZ$	harvest	$f_{Harvest} = \gamma_{Harvest} \cdot TAP$
t_eat2	$f_{t\_eat2} = \gamma_{t\_eat2} \cdot BPL \cdot TAP$	t_eat4	$f_{t\_eat4} = \gamma_{t\_eat4} \cdot TAP$
m_eat3	$f_{m\_eat3} = \gamma_{m\_eat3} \cdot MIZ \cdot MIZ$	export	$f_{export} = \gamma_{export}$

6.1.4 Set of Parameters $\mathcal{K}$

The computation of the constants γ_α is based on the same ideas as in Wood7.

Interaction	γ_α	Interaction	γ_α
growth	460/160	m_death	285/170
input	2731	m_breath	149/170
m_eat1	57/160*170	M_eat3	64/170*3000
M_eat1	64/160*3000	t_eat3	8/170*1463
p_death	331/160	M_eat4	120/3000*3000
t_eat1	8/160*1463	M_death	321/3000
b_eat1	3700/1600	M_breath	201/3000
b_breath	2800/1600	t_death	40/1463
M_eat2	394/1600*3000	t_breath	33/1463
m_eat2	449/1600*170	harvest	7/1463
t_eat2	57/1600*1463	t_eat4	8/1463
m_eat3	64/170*170	export	8/10000000

6.1.5 Set of Sequential Components $Comp$

In the set of sequential components of Venice Lagoon, all the stoichiometric coefficients are set to 1 as in Wood7. Moreover CO_2, DET and Input are assumed to be infinitely available. They can not decrease, thus we model their interactions with $\odot$ operator.

- $CO2 = (growth, 1) \odot CO2 + (t\_breath, 1) \uparrow CO2 + (m\_breath, 1) \uparrow CO2 + (M\_breath, 1) \uparrow CO2 + (b\_breath, 1) \uparrow CO2$;
- $PHP = (growth, 1) \uparrow PHP + (m\_eat1, 1) \downarrow PHP + (M\_eat1, 1) \downarrow PHP + (p\_death, 1) \downarrow PHP + (t\_eat1, 1) \downarrow PHP$;
- $MIZ = (m\_eat1, 1) \uparrow MIZ + (m\_eat2, 1) \uparrow MIZ + (m\_eat3, 1) \odot MIZ + (m\_breath, 1) \downarrow MIZ + (m\_death, 1) \downarrow MIZ + (M\_eat3, 1) \downarrow MIZ + (t\_eat3, 1) \downarrow MIZ$;
- $MEZ = (M\_eat1, 1) \uparrow MEZ + (M\_eat2, 1) \uparrow MEZ + (M\_eat3, 1) \uparrow MEZ + (M\_eat4, 1) \odot MEZ + (M\_death, 1) \downarrow MEZ + (M\_breath, 1) \downarrow MEZ$;
- $TAP = (t\_eat1, 1) \uparrow TAP + (t\_eat2, 1) \uparrow TAP + (t\_eat3, 1) \uparrow TAP + (t\_death, 1) \downarrow TAP + (t\_breath, 1) \downarrow TAP + (t\_eat4, 1) \uparrow TAP + (Harvest, 1) \downarrow TAP$;
- $BPL = (b\_eat1, 1) \uparrow BPL + (M\_eat2, 1) \downarrow BPL + (t\_eat2, 1) \downarrow BPL + (m\_eat2, 1) \downarrow BPL + (b\_breath, 1) \downarrow BPL$;
- $DET = (m\_death, 1) \uparrow DET + (p\_death, 1) \uparrow DET + (b\_eat1, 1) \odot DET + (input, 1) \uparrow DET + (M\_death, 1) \uparrow DET + (t\_death, 1) \uparrow DET + (t\_eat4, 1) \odot DET + (export, 1) \odot DET$;
- $Input = (input, 1) \odot Input$;
- $Harvesting = (Harvest, 1) \uparrow Harvesting$;
- $Export = (export, 1) \odot Export$;

6.1.6 The model component P

The model component of Venice Lagoon is defined as in [10]:

$$P = PHP[160] \underset{*}{\bowtie} MIZ[170] \underset{*}{\bowtie} MEZ[3000] \underset{*}{\bowtie} BPL[1600] \underset{*}{\bowtie} TAP[1463] \underset{*}{\bowtie} DET[5000] \underset{*}{\bowtie} Export[1] \underset{*}{\bowtie} Harvesting[1] \underset{*}{\bowtie} Input[3000] \underset{*}{\bowtie} CO2[5000]$$

In Moscardo's thesis [10] a very similar model for the Venice Lagoon was proposed. In this present model the Export has been introduced as in [11] and some errors have been corrected.

6.1.7 Analysis of the model

In Moscardo's thesis the model was simulated stochastically and the result was reasonable: the amount of internal species had small variation in time but the system was rather stable.

We are interested in analysing the continuous dynamics of the system, thus we applied the ODEs solver to the Bio-PEPA model of Venice Lagoon and get the result in Figure 6.1. The figure shows that the external species continuously increase and the other species are stable. DET and CO_2 infinitely grow, if we focus on the other species like in Figure 6.1 (b), we can see that all the internal species are stable and the species Harvesting is slightly increasing.

This shows that the internal species in the continuous simulation of Venice Lagoon are at steady state, as they should be, since the Bio-PEPA model intends to represent the mass balance equation given in [11].

6.2 Experiments on the model of the Venice Lagoon: external species

We can do some experiments on the Bio-PEPA model of the Venice Lagoon. We can modify the model to have a more realistic representation of the Venice Lagoon and analyse the continuous simulation of the modified model. As we said, in the model we assume there is infinite availability of CO_2, DET and Input, for this reason they were never allowed to be decreased through interactions with other species. We can consider more realistic models, namely we can explore:

- what happens if CO_2 can also decreases in the interactions.
- what happens if DET can also be decreased in the interactions.
- what happens if both of them can be decreased in the interactions.

6.2.1 Experiment 1: CO_2 can decrease

In the real world, the species CO_2 is consumed by PHP. In order to model also the decreasing of CO_2, we modify the sequential components CO_2:

$$CO2 = (growth, 1) \downarrow CO2 + (t\_breath, 1) \uparrow CO2 + (m\_breath, 1) \uparrow \\ CO2 + (M\_breath, 1) \uparrow CO2 + (b\_breath, 1) \uparrow CO2;$$

Figure 6.2 and Figure 6.3 shows the behaviour of the species with the old definition of CO_2 and with the new definition of the continuous simulation. We can see that, when we

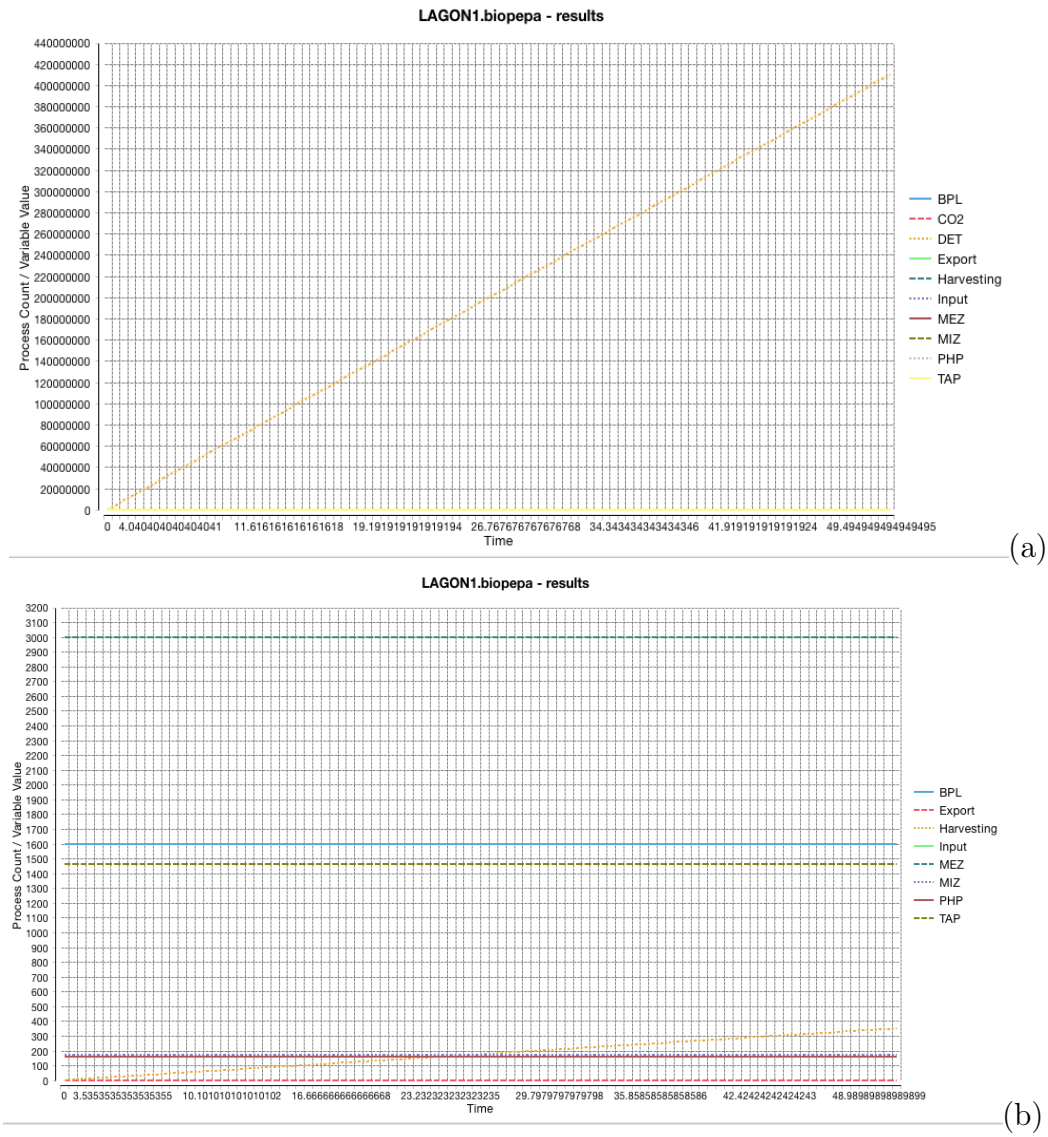
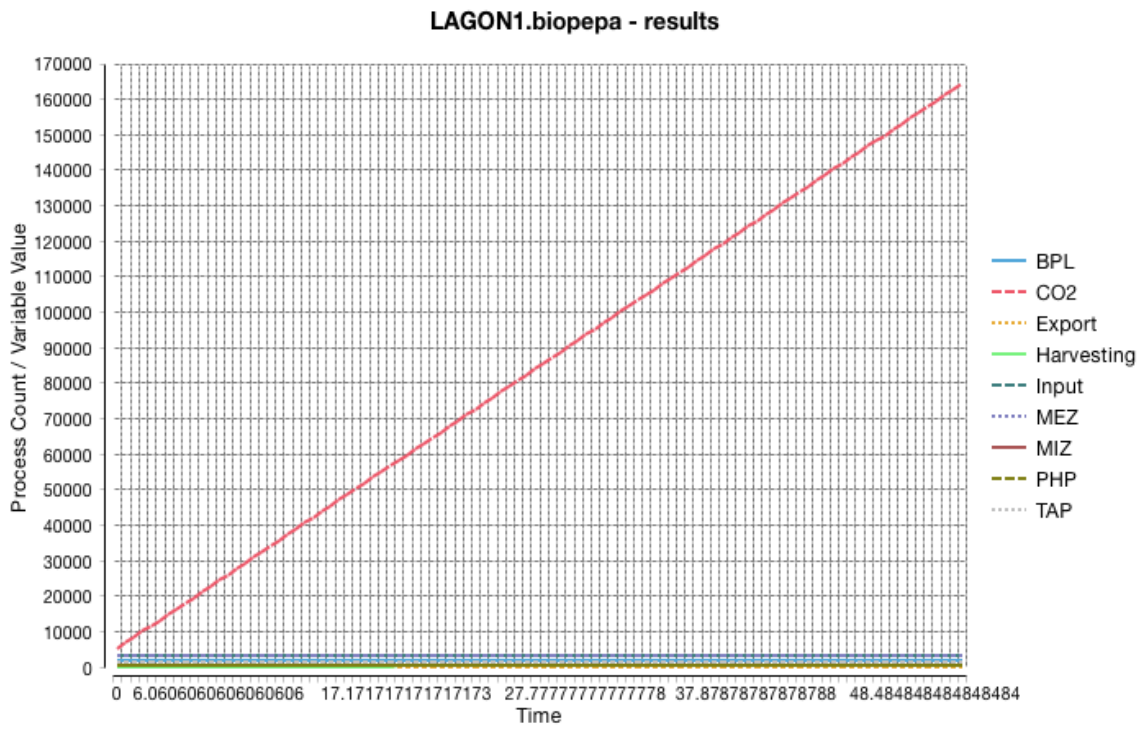


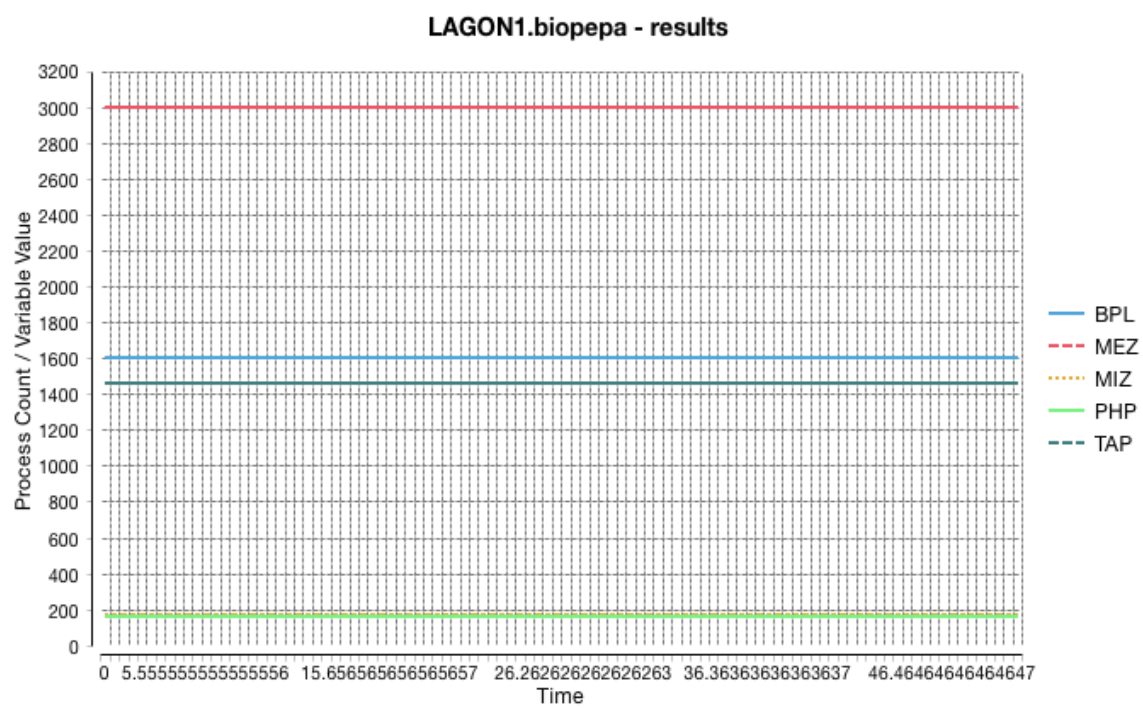
Figure 6.1: Continuous analysis of Venice Lagoon Bio-PEPA model

set that CO_2 can be decreased in the interaction with PHP, the behaviour of the system does not change, it does not have any effect on the internal species. Notice that the CO_2 increases in both cases, but in the second case, the increase is lower as shown by the values of the biomass axes.

This result shows that the decrease of the external species CO_2 due to PHP growth does not affect the behaviour of the internal species in the continuous simulation of the Venice Lagoon.

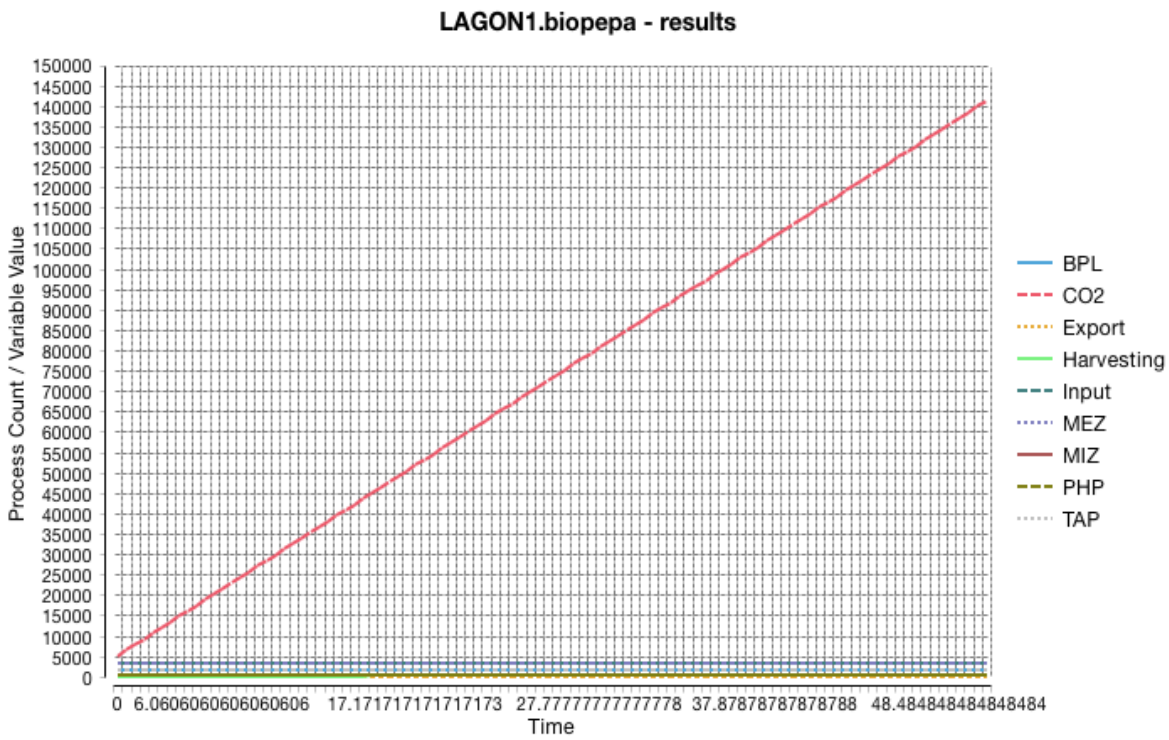


(a)The system when CO_2 can not be decreased

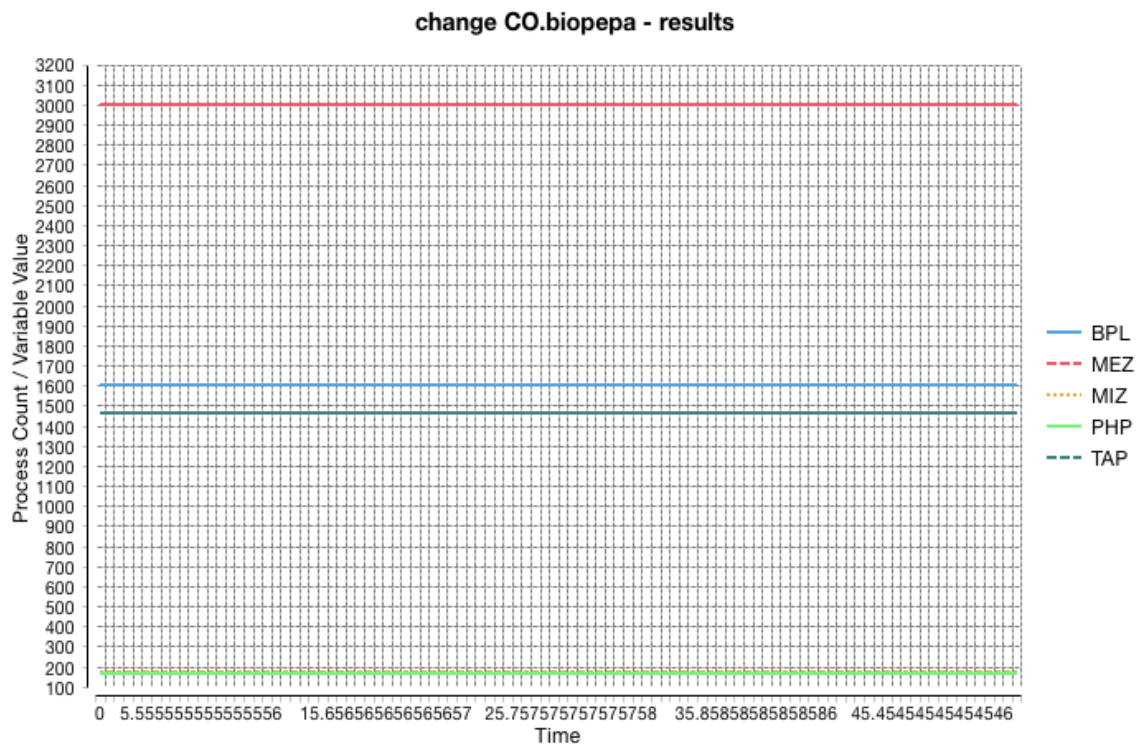


(b)The system when CO_2 can not be decreased: internal species are stable.

Figure 6.2: Continuous analysis of Venice Lagoon Bio-PEPA model



(a)The system when CO_2 can be decreased because of PHP growth



(b)The system when CO_2 can be decreased: internal species are still stable.

Figure 6.3: Continuous analysis of Venice Lagoon, Experiment 1.

6.2.2 Experiment 2: DET can decrease

The species DET is eaten by BPL and by TAP, moreover it depends also on the Export from the lagoon. Thus we set more realistically that the species DET can be decreased, by changing the sequential component of DET as follows:

$$DET = (m\_death, 1) \uparrow DET + (p\_death, 1) \uparrow DET + (b\_eat1, 1) \downarrow DET + (input, 1) \uparrow DET + (M\_death, 1) \uparrow DET + (t\_death, 1) \uparrow DET + (t\_eat4, 1) \downarrow DET + (export, 1) \downarrow DET;$$

When we set that the Detritus can be decreased, the result is like in Figure 6.4 (a) and (b). Species DET continuously increases and all the internal species are stable. This modification does not affect the behaviour of the internal species in the system. We can see the behaviour of the internal species in Figure 6.4 (b) more clearly, they are stable in the time interval. Notice that the increasing speed of DET does not change with respect to the original model.

This result shows that the decrease of the external species DET, does not affect the behaviour of the system in the continuous simulation of the Venice Lagoon.

6.2.3 Experiment 3: both CO_2 and DET can decrease

In this experiment we set that both CO_2 and DET can be decreased. The two corresponding sequential components have the following definitions:

$$CO2 = (growth, 1) \downarrow CO2 + (t\_breath, 1) \uparrow CO2 + (m\_breath, 1) \uparrow CO2 + (M\_breath, 1) \uparrow CO2 + (b\_breath, 1) \uparrow CO2;$$

$$DET = (m\_death, 1) \uparrow DET + (p\_death, 1) \uparrow DET + (b\_eat1, 1) \downarrow DET + (input, 1) \uparrow DET + (M\_death, 1) \uparrow DET + (t\_death, 1) \uparrow DET + (t\_eat4, 1) \downarrow DET + (export, 1) \downarrow DET;$$

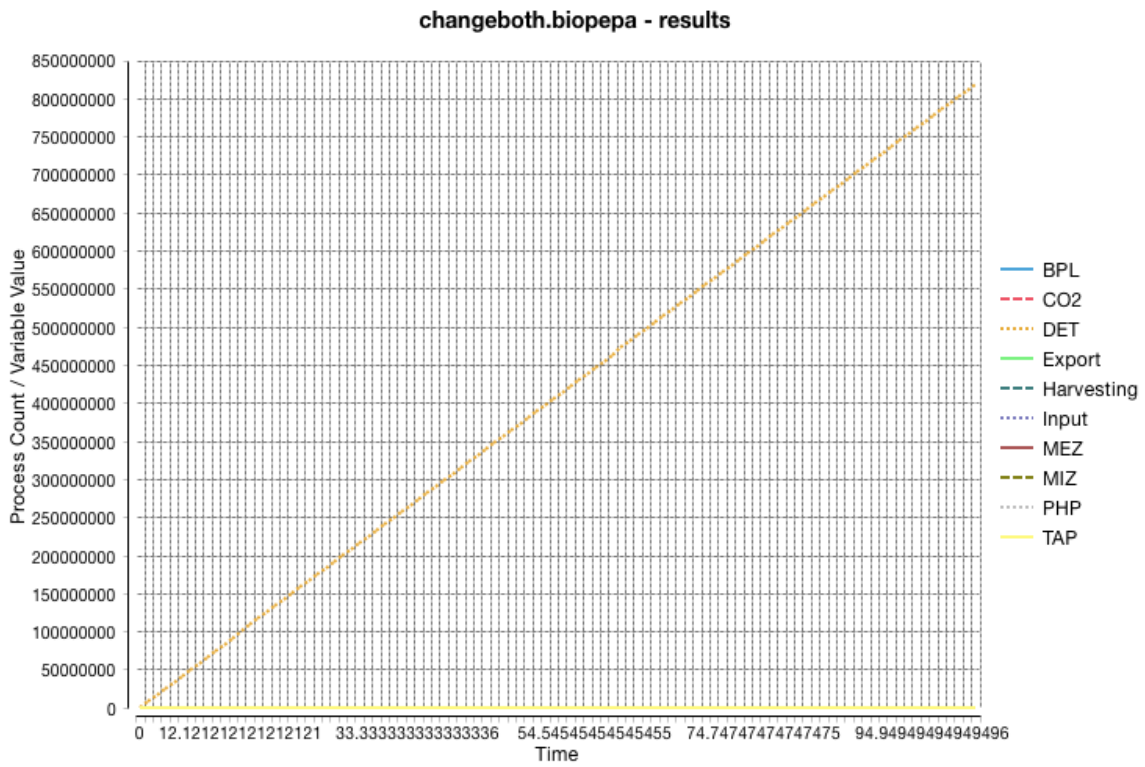
By continuous simulation we get the result in Figure 6.5 (a) and (b). Since the amount of CO_2 and DET is much greater than the amount of other species, in order to see the effect of the modification on the other species, in Figure 6.5 (b) we show only the internal species. CO_2 and DET continuously increase, the other species are stable and only Harvesting is slightly increasing.

This result shows that external species CO_2 and Det can be modelled more realistically by allowing them to be decreased. In the continuous behaviour their modification does not affect the internal species in the system.

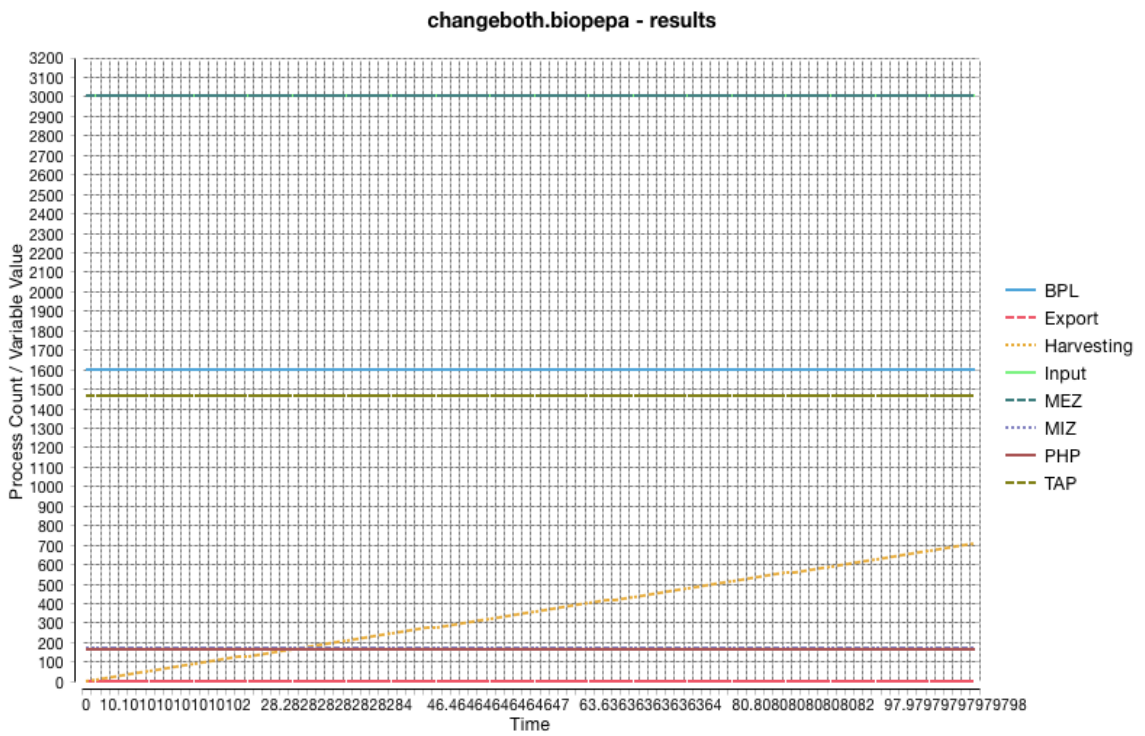
(a) The whole system behaviour when DET can be decreased: they are all behave the same as before.

(b) The internal species behaviour when Det can be decreased: internal species are all stable.

Figure 6.4: Continuous analysis of Venice Lagoon, Experiment 2



(a) The whole system behaviour when Det and CO_2 both changed.



(b) The system behaviour when Det and CO_2 both changed: the internal species are stable.

Figure 6.5: Continuous analysis of Venice Lagoon, Experiment 3

6.3 Experiments on the model of the Venice Lagoon: internal species

There are 5 internal species in the Venice Lagoon ecosystem. These species are placed at different trophic levels. From the food web model of Venice Lagoon (Figure 4.1), we can visually see the direct and indirect relations between the species.

PHP and BPL are the internal species at the first trophic level and they are independent one from each other. We can study the influence of the internal species PHP and BPL in the Venice Lagoon.

We perturbate their amounts by affecting their functional rates in the interactions growth for PHP and b_eat1 for BPL. In both the species we increase and decrease their amount.

6.3.1 Experiment 4: PHP grows or decreases

We change the amount of PHP by modifying the rate of action growth. We tried different values and we determine the smallest variation of the constant γ_{growth} in the functional rate which affects the system. The constant γ_{growth} represents the probability of the action growth. In the model, it is equal to 460/160 (introduced in section 6.1.5). We modify it by adding or subtracting 30/160. In this way, the functional rate of growth also is perturbated.

Figure 6.6, shows the behaviour of internal species in the interactions. Figure 6.6 (a) shows that, when we increase PHP (γ_{growth}= 460/160 + 30/160), in the long time the species MEZ increases accordingly. All the other internal species are decreasing. This may be due to the fact that the functional rate of "MEZ eats PHP" is greater than "MEZ eats other species in the system". While the amount of MEZ increases, the amount of species MIZ and BPL decreases because they are both eaten by species MEZ. The reason why TAP decreases, may be that its preys are all eaten by MEZ and for TAP there may be lack of food.

Figure 6.6 (b) shows the behaviour of the internal species when we decrease PHP (γ_{growth}= 460/160 - 30/160), in the long time, PHP goes to zero, the species MEZ decreases accordingly, the species MIZ slightly increases, the species BPL slightly decreases and the species TAP decreases also in this case. We can notice that except for the species MIZ, all other species are decreasing. The variation of species MEZ is proportional to the variation of PHP, this strong relation between MEZ and PHP is due to the functional rate of "MEZ eats PHP" (M_eat1). Since they are both decreasing, the species MIZ slightly increase, because the functional rate of the action "MEZ eats MIZ" is larger than the rates of other predators eating MIZ. In this case the predator MEZ is decreasing then, MIZ has a chance to increase.

The species BPL does not have direct interaction with the species PHP but it is affected by PHP, it is slightly decreases. Species TAP decreases since it is a predator of PHP and PHP decreases.

From these results, we can conclude that:

(1) there is a strong direct dependency between the species MEZ and the species PHP (MEZ increases when PHP increases and when PHP decreases, MEZ also decreases).

(2) even if there is no direct interaction between the species PHP and the species BPL, there is a dependency relation between these two species (BPL decreases whether PHP increases or decreases).

(3) this modifications also affected the other species among the system. The species MIZ and TAP have a weaker direct relation with the species PHP.

6.3.2 Experiment 5: BPL increases or decreases

We can change the amount of BPL by modifying b_eat1. We tried different values and we determine the smallest variation of the constant γ_{b_eat1} in the functional rate, which affects the system. The constant γ_{b_eat1} represents the probability of the action b_eat1. In the model, it is equal to 3700/1600 (introduced in section 6.1.5). We modify it by adding or subtracting 30/1600. In this way the functional rate of b_eat1 also is perturbated.

Figure 6.7 (a) shows that, in the long time, when we increase the amount of BPL (γ_{b_eat1}= 3700/1600 + 30/1600), the species MEZ oscillates according to the species BPL. Also other species in the system are affected. The species TAP and MEZ have a direct correlation and an inverse correlation with the species MIZ and PHP.

From Figure 6.7 (b) we can see that, when we decrease the amount of BPL (γ_{b_eat1}= 3700/1600 - 30/1600), the species MEZ oscillates proportional to the species BPL, also the species TAP slightly oscillates according to the species BPL. The species PHP oscillates inversely to the species BPL, also the species MIZ slightly oscillates inversely to the species BPL.

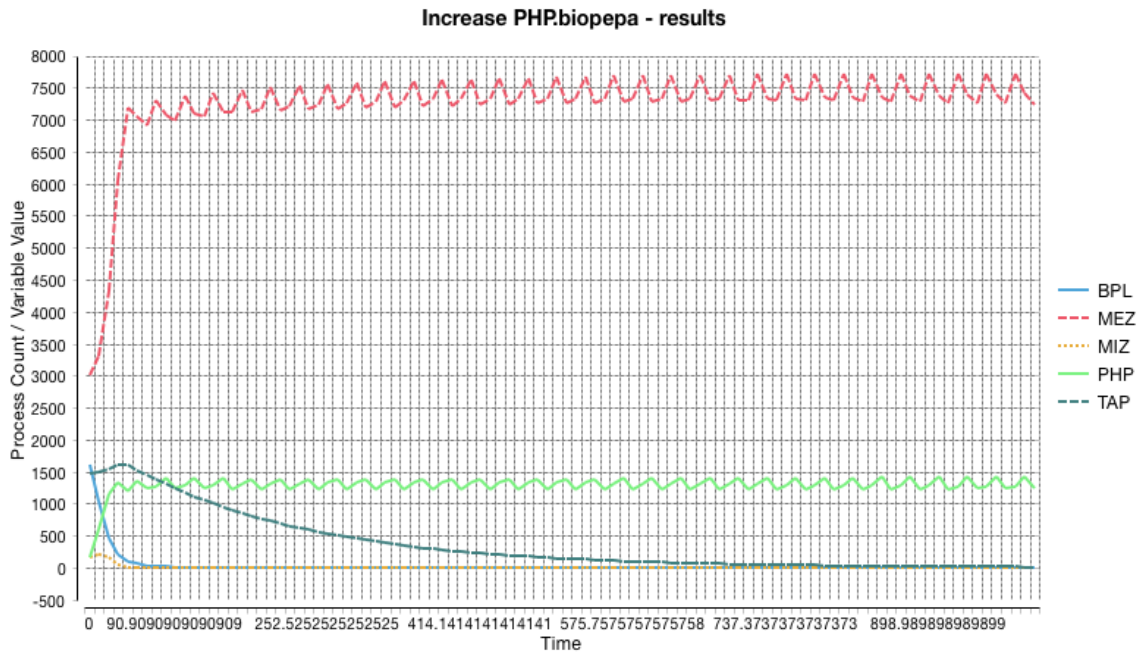
From these results we can conclude that:

(1) the species BPL has strong direct correlation with the species MEZ and weaker direct correlation with the species TAP.

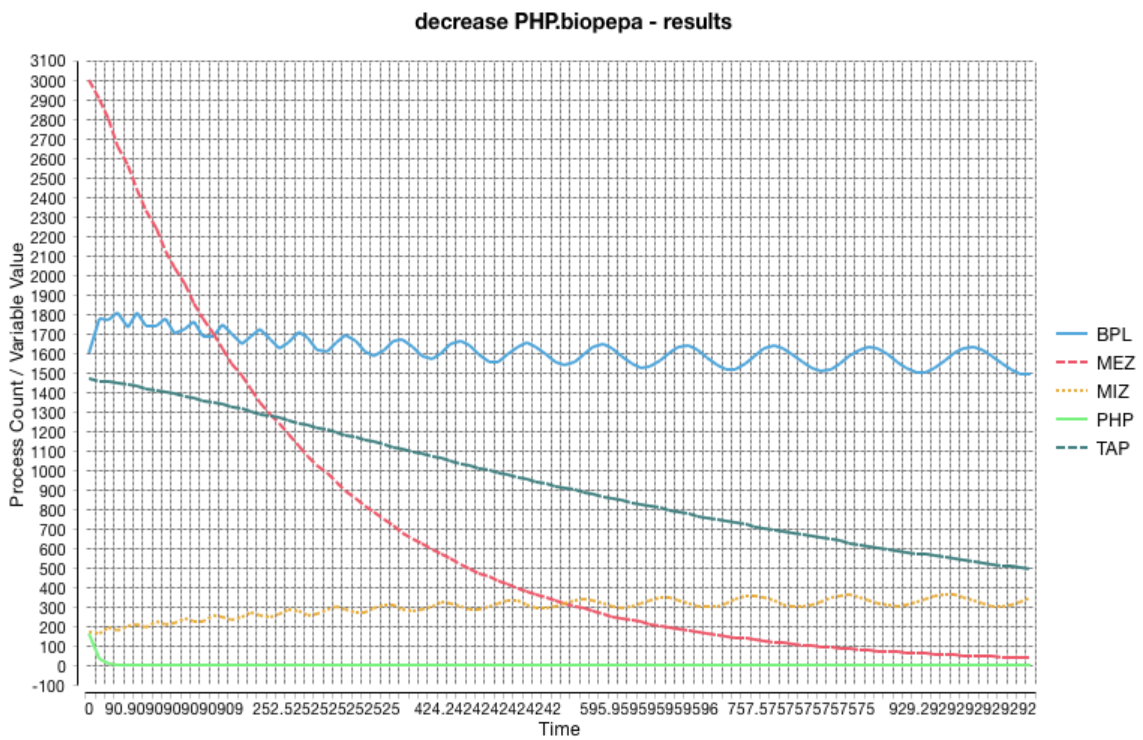
(2) the species BPL has an inverse correlation with the species PHP. In both cases the inverse correlation between BPL and PHP is very clear.

(3) the variation of the species BPL affects weakly on the behaviour of the speices MIZ.

From these experiments we can conclude that, correlations are not trivial to infer and by perturbating the species at the first trophic level non-trivial dependency relations between species can be revealed.

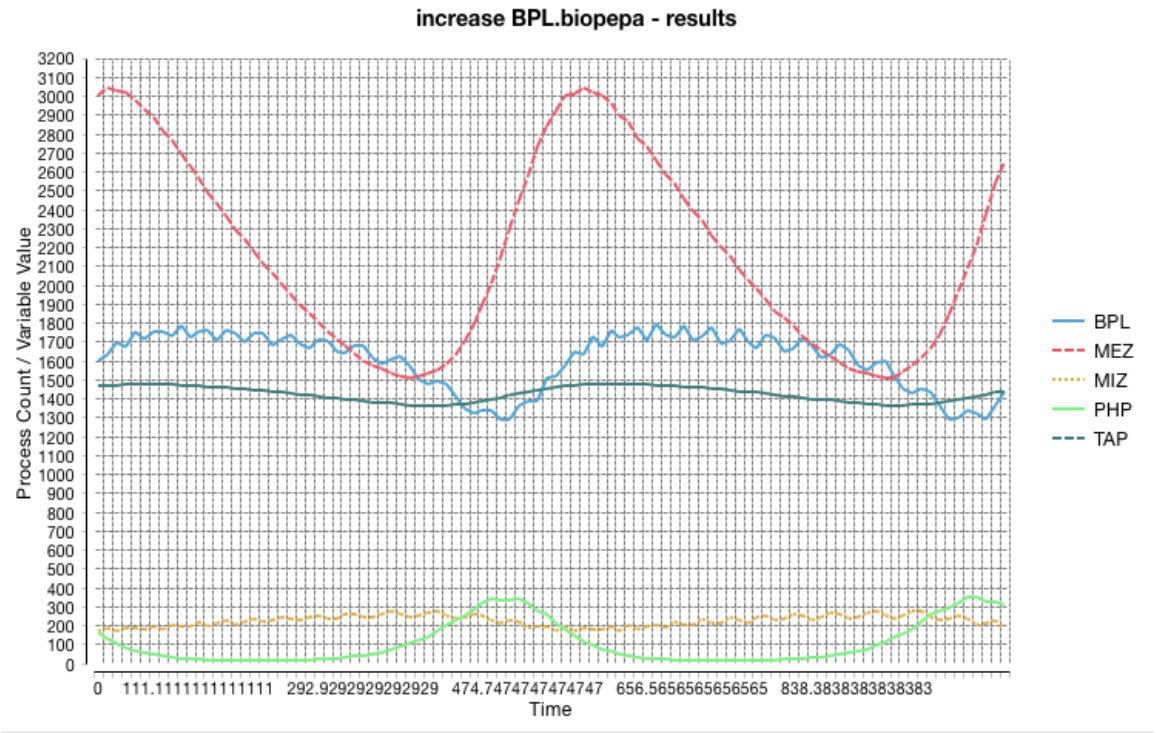


(a)The system behaviour when PHP increases.

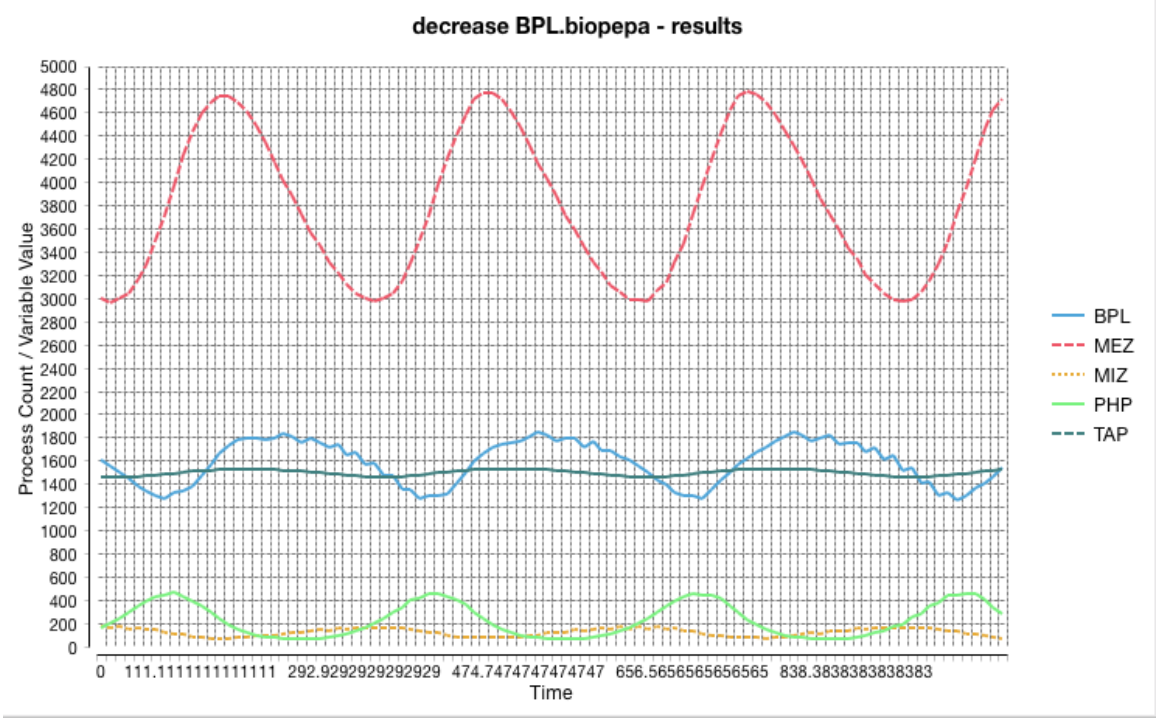


(b)The system behaviour when PHP decreases.

Figure 6.6: Continuous analysis of Venice Lagoon: Experiment 4



(a)The system behaviour when BPL increases.



(b)The system behaviour when BPL decreases.

Figure 6.7: Continuous analysis of Venice Lagoon: Experiment 5.

Chapter 7

Conclusion

This thesis focuses on the continuous modelling and dynamic analysis of trophic networks by using the process algebra Bio-PEPA. We base our analysis on the previous work of Moscardo, who explored the stochastic modelling of trophic networks with Bio-PEPA. He considered a steady state snapshot of Wood, a toy trophic network. We took his best experiment result, Wood7, and apply the continuous dynamic analysis to it.

Wood7 is modelled with unitary stoichiometry coefficients and its functional rates are computed based on the Mass Action law. The stochastic analysis of the model shows that in time, all the species get extinct except for the species Grass. This happens for the reason that the functional rate of growth increases because the amount of Grass increases and the other functional rates in the system are much smaller than this functional rate. This shows that the range of the functional rate affects heavily the behaviour of the system in the stochastic modelling. In order to handle this problem, we modify Wood7 by splitting the species Grass in ten equal units and call this model Wood7'. In Wood7', the stoichiometry coefficients are unitary and functional rates are modified accordingly and based on the Mass Action law as in Wood7. Unlike Wood7, in the stochastic analysis of Wood7', the internal species do not become extinct. This is because the range of the functional rates of the species has been reduced and this keeps the system more stable. We apply the continuous analysis to Wood7, in this case the internal species are all stable. This is due to the fact that the continuous simulation corresponds to averaging many stochastic simulations and it avoids the variation due to the stochastic model. The continuous dynamics of Wood7', is very similar to the one of Wood7. This shows that the range of the functional rates does not have such a dramatic impact on the continuous system behaviour as it has in the stochastic one.

We did some experiments on the modified model Wood7'. Since the species CO_2 should be infinitely available in the system, it was not decreases in the interactions. We model

it more reasonably by allowing CO_2 to be decreased in the interactions. When CO_2 may decrease, the analysis shows that, there is no effect on the internal species in the continuous simulation of the system. But in the stochastic simulation it affects the internal species: Erb decreases accordingly and the other internal species are oscillating until CO_2 goes to zero.

Based on Moscardo's work, we use the knowledge gained from the experiments on the toy case to built a Bio-PEPA model of a real trophic network, namely the Venice Lagoon. A first Bio-PEPA model of the Venice Lagoon was given by Moscardo in his master thesis. We did some improvements on his model of the Venice Lagoon and we apply to it the continuous dynamic simulation. The analysis shows that the Bio-PEPA model of the Venice Lagoon is in steady state with internal species, as it should be, when we consider the continuous dynamics of the system.

Some experiments have been performed on the external and internal species of the system. The species CO_2 and the species DET are assumed to be infinitely available in the system, and they can not be decreased. In our first set of experiments we modify the model to allow them to decrease. CO_2 may decrease due to the growth of PHP: the analysis shows that CO_2 still increases, but slower than the first model and it does not affect the stability of the internal species in the continuous dynamics of the system. DET can be decreased because it is eaten by some internal species or it can exit the system. We modified the model accordingly. The new model of the Venice Lagoon continuous simulation shows that DET still increases and it does not affect the stability of the internal species of the system. We modify the model so that both CO_2 and DET are allowed to be decreased, the result of continuous simulation shows that the internal species of the system still remain stable. This indicates that we can model the decrease of the external species CO_2 and DET without altering their infinite availability and the model is more realistic.

In order to analyse the relationships between the internal species, we did a series of experiments by slightly changing the amounts of the internal species which are at the first trophic level, namely PHP and BPL. We change the flow which determines the biomass of each species separately, by increasing or decreasing its original value. The experiments show that the species MEZ is strongly correlated to PHP and BPL. The species MIZ and TAP are also correlated to PHP and BPL, but with a weaker relation. In particular, when PHP increases, the species MEZ increases with a strong direct relation, TAP, MIZ and BPL all decrease. When PHP decreases, the species MEZ decreases accordingly, the species BPL, TAP decrease and the species MIZ slightly increases. In the experiment on the internal species BPL, when BPL increases, we can observe that TAP and MEZ have a direct correlation between them and an inverse correlation with PHP and MIZ. When BPL decreases

there is a weaker direct correlation with TAP and a strong correlation with MEZ, and has the inverse correlation with MIZ and PHP. When altering the biomass of BPL the inverse correlation between BPL and PHP is very clear.

These experiments show that, when modelling continuous behaviour of the system, we can avoid unrealistic assumptions on the external species. Moreover, by perturbating the species at the first trophic lavel we can observe that the internal species of the system are all correlated with strong and weaker, direct and indirect relations. These correlations are not trivial to infer by the system description. Hence by perturbating the species at the first trophic level, non-trivial dependency relations between species can be revealed.

Bibliography

[1] F. Ciocchetta, and J. Hillston. *Bio-PEPA: a framework for the modelling and analysis of biological systems*. Technical report EDI-INF-RR-1231, University of Edinburgh, 2008.

[2] J. Hillston. *A compositional approach to performance modelling*. Cambridge University Press, 1996.

[3] Federica Ciocchetta, Andrea Degasperi, Jane Hillston, and Muffy Calder. *Some investigations concerning the CTMC and the ODE model derived from Bio-PEPA*. Laboratory for Foundations of Computer Science. The University of Edinburgh, Scotland. Department of computer science, University of Glasgow, Scotland.

[4] J.Hillston. *Process algebras for quantitative analysis*. School of Informatics, University of Edinburgh, UK.

[5] M. Calder, V. Vyshemirsky, D. Gilbert, and R. Orton, *Analysis of signalling pathways using continuous time markov chains*. T.Comp, Sys.Biology, VI, LNCS 4220 (2006) 44-67, Springer.

[6] D.T. Gillespie. *Exact stochastic simulation of coupled chemical reactions*. Journal of Physical Chemistry 81 (1977) 2340-2361.

[7] A.P. Arkin, C.V. Rao. *Stochastic chemical kinetics and the quasi-steady-state assumption: application to the Gillespie algorithm*, Journal of Chemical Physics 11 (2003) 4999-5010.

[8] Prism web site. *http://www.prismmodelchecker.org* .

[9] Federica Ciocchetta, and Jane Hillston *Process algerbras in system biology*. Laboratory for Foundations of Computer Science. The University of Edinburgh, Edinburgh EH93JZ, Scotland.

[10] Moscardo Marco, Cocco Nicoletta. *Modelling trophic networks in PEPA and Bio-PEPA*. University of Ca'Foscari, thesis for MSc in Computer Science, A.A. 2013-2014.

[11] Robert E. Ulanowicz. *Quantitative methods for ecological network analysis*. Chesapeake Biological Laboratory, University of Maryland Center for Environmental Science, Solomons, MD 20688-0038, USA Received 2 September 2004, accepted 5 September 2004.

[12] Raymond L.Lindeman. *The Trophic-Dynamic aspect of ecology*. Ecology, Vol.23, No.4 (Oct., 1942), 399-417.

[13] Villy Christensen, Carl J. Walters *Ecopath with Ecosim: methods, capabilities and limitations*. Fisheries Centre, University of British Columbia, 2204 Main Mall, Vancouver, BC, Canada V6T1Z4.

[14] Andria K. Salas. *Indirect effects in trophic and evolutionary Networks*. Center of Marine Science University of North Carolina Wilmington 2011.

[15] Brigolin Daniele, Pastres Roberto. *Influence of Intra-Seasonal variability of metabolic rates on the output of a steady-state food web model* . Department of environmental sciences, informatics and statistics, universita Ca'Foscari, Venezia, Dorsoduro 2137, 30123 Venezia, Italy.

[16] Karline Soetaert and Dick Van Oevelen. *Package LIM , implementing linear inverse models in R*. Royal Netherlands Institute of Sea Research Yerseke, the Netherlands.

[17] *The Bio-PEPA eclipce plug in user manual*. http://www.biopepa.org. April 20, 2012.

[18] Stefano Allesinaa and Robert E Ulanowiczb. *Cycling in ecological networks: Finn's index revisited*. Computational Biology and Chemistry. Volume 28, Issue 3, Pages 227-233, July 2004.

Appendix A

Bio-PEPA Codes of Models

A.1 Bio-PEPA Code of Wood7

$speciesCarn : upper = 5000, lower = 1;$
$speciesErb : upper = 5000, lower = 1;$
$speciesGrass : upper = 5000, lower = 1;$
$speciesDet : upper = 5000, lower = 0;$
$speciesCO : upper = 500000, lower = 1;$
$speciesHunt : upper = 500000, lower = 0;$
$c\_breath = [0.2 * Carn];$
$c\_hunt = [0.1 * Carn];$
$c\_death = [0.5 * Carn];$
$c\_eat = [0.04 * Carn * Erb];$
$e\_breath = [0.2 * Erb];$
$e\_hunt = [0.05 * Erb];$
$e\_death = [0.3 * Erb];$
$e\_eat2 = [0.005 * Grass * Erb];$
$g\_death = [0.1 * Grass];$
$e\_eat1 = [0.45 * Erb];$
$growth = [0.2 * Grass];$
$Carn = (c\_eat, 1) >> +(c\_breath, 1) << +(c\_hunt, 1) << +(c\_death, 1) <<; Erb = (e\_eat1, 1) >> +(e\_eat2, 1) >> +(e\_breath, 1) << +(c\_eat, 1) << +(e\_hunt, 1) << +(e\_death, 1) <<;$
$Grass = (growth, 1) >> +(e\_eat2, 1) << +(g\_death, 1) <<;$
$Det = (g\_death, 1) >> +(c\_death, 1) >> +(e\_death, 1) >> +(e\_eat1, 1)(.);$
$CO = (growth, 1)(.) + (e\_breath, 1) >> +(c\_breath, 1) >>;$
$Hunt = (c\_hunt, 1) >> +(e\_hunt, 1) >>;$
$Grass[100] < * > Erb[20] < * > Det[100] < * > Carn[10] < * > CO[100] < * > Hunt[1]$

A.2 Bio-PEPA Code of Wood7'

species Carn : upper=5000, lower=1;
species Erb : upper=5000, lower=1;
species Grass1 : upper=5000, lower=1;
species Grass2 : upper=5000, lower=1;
species Grass3 : upper=5000, lower=1;
species Grass4 : upper=5000, lower=1;
species Grass5 : upper=5000, lower=1;
species Grass6 : upper=5000, lower=1;
species Grass7 : upper=5000, lower=1;
species Grass8 : upper=5000, lower=1;
species Grass9 : upper=5000, lower=1;
species Grass10 : upper=5000, lower=1;
species Det : upper=5000, lower=0;
species CO : upper=500000, lower=1;
species Hunt : upper=500000, lower=0;
$c\_breath = [0.2 * Carn]$;
$c\_hunt = [0.1 * Carn]$;
$c\_death = [0.5 * Carn]$;
$c\_eat = [0.04 * Carn * Erb]$;
$e\_breath = [0.2 * Erb]$;
$e\_hunt = [0.05 * Erb]$;
$e\_death = [0.3 * Erb]$;
$e\_eat2\_1 = [0.005 * Grass1 * Erb]$;
$e\_eat2\_2 = [0.005 * Grass2 * Erb]$;
$e\_eat2\_3 = [0.005 * Grass3 * Erb]$;
$e\_eat2\_4 = [0.005 * Grass4 * Erb]$;
$e\_eat2\_5 = [0.005 * Grass5 * Erb]$;
$e\_eat2\_6 = [0.005 * Grass6 * Erb]$;
$e\_eat2\_7 = [0.005 * Grass7 * Erb]$;
$e\_eat2\_8 = [0.005 * Grass8 * Erb]$;
$e\_eat2\_9 = [0.005 * Grass9 * Erb]$;
$e\_eat2\_10 = [0.005 * Grass10 * Erb]$;
$g\_death1 = [0.1 * Grass1]$;
$g\_death2 = [0.1 * Grass2]$;
$g\_death3 = [0.1 * Grass3]$;
$g\_death4 = [0.1 * Grass4]$;
$g\_death5 = [0.1 * Grass5]$;
$g\_death6 = [0.1 * Grass6]$;
$g\_death7 = [0.1 * Grass7]$;
$g\_death8 = [0.1 * Grass8]$;
$g\_death9 = [0.1 * Grass9]$;
$g\_death10 = [0.1 * Grass10]$;
$e\_eat1 = [0.45 * Erb]$;

$growth1 = [0.2 * Grass1];$
$growth2 = [0.2 * Grass2];$
$growth3 = [0.2 * Grass3];$
$growth4 = [0.2 * Grass4];$
$growth5 = [0.2 * Grass5];$
$growth6 = [0.2 * Grass6];$
$growth7 = [0.2 * Grass7];$
$growth8 = [0.2 * Grass8];$
$growth9 = [0.2 * Grass9];$
$growth10 = [0.2 * Grass10];$
$Carn = (c\_eat, 1) >> +(c\_breath, 1) << +(c\_hunt, 1) << +(c\_death, 1) <<;$
$Erb = (e\_eat1, 1) >> +(e\_eat2\_1, 1) >> +(e\_eat2\_2, 1) >> +(e\_eat2\_3, 1) >> +(e\_eat2\_4, 1) >>$
$+(e\_eat2\_5, 1) >> +(e\_eat2\_6, 1) >> +(e\_eat2\_7, 1) >> +(e\_eat2\_8, 1) >> +(e\_eat2\_9, 1) >>$
$+(e\_eat2\_10, 1) >> +(e\_breath, 1) << +(c\_eat, 1) << +(e\_hunt, 1) << +(e\_death, 1) <<$
$;$
$Grass1 = (growth1, 1) >> +(e\_eat2\_1, 1) << +(g\_death1, 1) <<;$
$Grass2 = (growth2, 1) >> +(e\_eat2\_2, 1) << +(g\_death2, 1) <<;$
$Grass3 = (growth3, 1) >> +(e\_eat2\_3, 1) << +(g\_death3, 1) <<;$
$Grass4 = (growth4, 1) >> +(e\_eat2\_4, 1) << +(g\_death4, 1) <<;$
$Grass5 = (growth5, 1) >> +(e\_eat2\_5, 1) << +(g\_death5, 1) <<;$
$Grass6 = (growth6, 1) >> +(e\_eat2\_6, 1) << +(g\_death6, 1) <<;$
$Grass7 = (growth7, 1) >> +(e\_eat2\_7, 1) << +(g\_death7, 1) <<;$
$Grass8 = (growth8, 1) >> +(e\_eat2\_8, 1) << +(g\_death8, 1) <<;$
$Grass9 = (growth9, 1) >> +(e\_eat2\_9, 1) << +(g\_death9, 1) <<;$
$Grass10 = (growth10, 1) >> +(e\_eat2\_10, 1) << +(g\_death10, 1) <<;$
$Det = (g\_death1, 1) >> +(g\_death2, 1) >> +(g\_death3, 1) >> +(g\_death4, 1) >>$
$+(g\_death5, 1) >> +(g\_death6, 1) >> +(g\_death7, 1) >> +(g\_death8, 1) >> +(g\_death9, 1) >>$
$+(g\_death10, 1) >> +(c\_death, 1) >> +(e\_death, 1) >> +(e\_eat1, 1)(.);$
$CO = (growth1, 1)(.) + (growth2, 1)(.) + (growth3, 1)(.) + (growth4, 1)(.) + (growth5, 1)(.) +$
$(growth6, 1)(.) + (growth7, 1)(.) + (growth8, 1)(.) + (growth9, 1)(.) + (growth10, 1)(.) + (e\_breath, 1) >>$
$+(c\_breath, 1) >>;$
$Hunt = (c\_hunt, 1) >> +(e\_hunt, 1) >>;$
$Grass1[10] <*> Grass2[10] <*> Grass3[10] <*> Grass4[10] <*> Grass5[10] <*>$
$Grass6[10] <*> Grass7[10] <*> Grass8[10] <*> Grass9[10] <*> Grass10[10] <$
$*> Erb[20] <*> Det[100] <*> Carn[10] <*> CO[100] <*> Hunt[1]$

Bio-PEPA code of the experiments on the modified model Wood7' is not present here. The sequential component CO changed as in Section 5.4.8 and others are remained same .

A.3 Bio-PEPA Code of Venice Lagoon

species PHP : upper=50000, lower = 0;
species BPL : upper=50000, lower = 0;
species MIZ : upper=50000, lower = 0;
species MEZ : upper=50000, lower = 0;
species TAP : upper=50000, lower = 0;
species CO2 : upper=50000, lower = 0;
species Harvesting : upper=50000, lower = 0;
species Export : upper=50000, lower = 0;
species Input : upper=50000, lower = 0;
species DET : upper=50000, lower = 0;
$g\_growth = 460/160;$
$g\_input = 2731;$
$g\_m\_eat1 = 57/(160 * 170);$
$g\_M\_eat1 = 64/(160 * 3000);$
$g\_p\_death = 331/160;$
$g\_t\_eat1 = 8/(160 * 1463);$
$g\_b\_eat1 = 3700/1600;$
$g\_b\_breath = 2800/1600;$
$g\_M\_eat2 = 394/(1600 * 3000);$
$g\_m\_eat2 = 449/(1600 * 170);$
$g\_t\_eat2 = 57/(1600 * 1463);$
$g\_m\_eat3 = 64/(170 * 170);$
$g\_m\_death = 285/170;$
$g\_m\_breath = 149/170;$
$g\_M\_eat3 = 64/(170 * 3000);$
$g\_t\_eat3 = 8/(170 * 1463);$
$g\_M\_eat4 = 120/(3000 * 3000);$
$g\_M\_death = 321/3000;$
$g\_M\_breath = 201/3000;$
$g\_t\_death = 41/1463;$
$g\_t\_breath = 33/1463;$
$g\_Harvest = 7/1463;$
$g\_t\_eat4 = 8/1463;$
$g\_export = 8/10000000;$
$growth = [g\_growth * PHP];$
$input = [g\_input];$
$m\_eat1 = [g\_m\_eat1 * PHP * MIZ];$
$M\_eat1 = [g\_M\_eat1 * PHP * MEZ];$
$p\_death = [g\_p\_death * PHP];$
$t\_eat1 = [g\_t\_eat1 * PHP * TAP];$
$b\_eat1 = [g\_b\_eat1 * BPL];$
$b\_breath = [g\_b\_breath * BPL];$
$M\_eat2 = [g\_M\_eat2 * BPL * MEZ];$

$m\_eat2 = [g\_m\_eat2 * BPL * MIZ];$
$t\_eat2 = [g\_t\_eat2 * BPL * TAP];$
$m\_eat3 = [g\_m\_eat3 * MIZ * MIZ];$
$m\_death = [g\_m\_death * MIZ];$
$m\_breath = [g\_m\_breath * MIZ];$
$M\_eat3 = [g\_M\_eat3 * MIZ * MEZ];$
$t\_eat3 = [g\_t\_eat3 * MIZ * TAP];$
$M\_eat4 = [g\_M\_eat4 * MEZ * MEZ];$
$M\_death = [g\_M\_death * MEZ];$
$M\_breath = [g\_M\_breath * MEZ];$
$t\_death = [g\_t\_death * TAP];$
$t\_breath = [g\_t\_breath * TAP];$
$Harvest = [g\_Harvest * TAP];$
$t\_eat4 = [g\_t\_eat4 * TAP];$
$export = [g\_export];$
$CO2 = (growth, 1)(.)+(t\_breath, 1) >> +(m\_breath, 1) >> +(M\_breath, 1) >> +(b\_breath, 1) >>$
$;$
$PHP = (growth, 1) >> +(m\_eat1, 1) << +(M\_eat1, 1) << +(p\_death, 1) << +(t\_eat1, 1) <<$
$;$
$MIZ = (m\_eat1, 1) >> +(m\_eat2, 1) >> +(m\_eat3, 1)(.)+(m\_breath, 1) << +(m\_death, 1) <<$
$+(M\_eat3, 1) << +(t\_eat3, 1) <<;$
$MEZ = (M\_eat1, 1) >> +(M\_eat2, 1) >> +(M\_eat3, 1) >> +(M\_eat4, 1)(.)+(M\_death, 1) <<$
$+(M\_breath, 1) <<;$
$TAP = (t\_eat1, 1) >> +(t\_eat2, 1) >> +(t\_eat3, 1) >> +(t\_death, 1) << +(t\_breath, 1) <<$
$+(t\_eat4, 1) >> +(Harvest, 1) <<;$
$BPL = (b\_eat1, 1) >> +(M\_eat2, 1) << +(t\_eat2, 1) << +(m\_eat2, 1) << +(b\_breath, 1) <<$
$;$
$DET = (m\_death, 1) >> +(p\_death, 1) >> +(b\_eat1, 1)(.)+(input, 1) >> +(M\_death, 1) >>$
$+(t\_death, 1) >> +(t\_eat4, 1)(.) + (export, 1)(.);$
$Input = (input, 1)(.);$
$Harvesting = (Harvest, 1) >>;$
$Export = (export, 1)(.);$
$PHP[160] < * > MIZ[170] < * > MEZ[3000] < * > BPL[1600] < * > TAP[1463] < * >$
$DET[5000] < * > Export[1] < * > Harvesting[1] < * > Input[3000] < * > CO2[5000]$

Bio-PEPA codes of the experiments on the Venice Lagoon are not present. The sequential components CO_2, DET and the value of the functional rates of PHP and BPL are changed as in Section 6.2 and Section 6.3.

Appendix B

Kurtz's Theorem

In section 2.2.7 we gave the definition of density dependent CTMC for applying Kurtz's theorem. Here is the Kurtz's theorem.

Theorem B.1. *Let X_V be a family of density dependent CTMCs with the infinitesimal generator matrix as in the definition above. Assume $X(t)$ is the solution of the ODE system $dX/dt = F(X)$, where $F(X) = \sum_s sf(x,s)$ and let $X(0) = x_0$. If there exists an open set $E \subset \mathbb{R}^n$ such that $X(t) \in E$ and*

- $\exists M, \forall x, y \in E \mid F(x) - F(y) \mid < M \mid x - y \mid$;
- $sup \sum_s \mid s \mid f(x,s) < \infty$;
- $\lim_{d\to\infty} sup_{x\in E} \sum_{|s|>d} |s| \; f(x,s) = 0$.

then

$$\lim_{V\to\infty} V^{-1}X_V(0) = X_0 \Rightarrow \forall\delta > 0, \forall t > 0 \; \lim_{V\to\infty} \mathbb{P}\,(\sup_{z<t} \mid V^{-1}X_V(z) - X(z) \mid > \delta) = 0.$$

In this theorem, the states represent numbers of individuals and they are normalised with respect to a parameter V. Hence $V^1X_V(z)$ represents the scaled Markov process with concentrations. In our case the scaling factor is in terms of h instead of V.