MODELLING THE POPULATION DYNAMICS IN A CELL CULTURE AT TWO DIFFERENT SCALES

by

M. Gökhan Habiboğlu

B.S., Electrical and Electronics Engineering, Boğaziçi University, 2004

M.S., Electrical and Electronics Engineering, Boğaziçi University, 2007

Submitted to the Institute for Graduate Studies in
Science and Engineering in partial fulfillment of
the requirements for the degree of
Doctor of Philosophy

Graduate Program in Electrical and Electronics Engineering

Boğaziçi University

2015

MODELLING THE POPULATION DYNAMICS IN A CELL CULTURE AT TWO DIFFERENT SCALES

APPROVED BY:

Prof. Yağmur Denizhan (Thesis Supervisor)	.....
Prof. Leyla Gören	.....
Assoc. Prof. Murat Saraçlar	.....
Prof. Yaman Barlas	.....
Prof. Oğuzhan Çiçekoğlu	.....

DATE OF APPROVAL: 16.01.2015

ACKNOWLEDGEMENTS

I am very indebted to my advisor Prof. Yağmur Denizhan, not only for her genuine interest and enlightening contributions to my thesis at its nearly every stage, but also for her inspirational way of thinking, attitude and personality. I am honoured to have been one of those who has witnessed the depth and diversity of her knowledge and wisdom. Without her, this thesis would not be possible.

It is hard not to mention the invaluable support and trust of my wife throughout my thesis, as well as the impact my son's birth near the end of my PhD journey. I will always remember them as a part of this thesis.

I am also grateful to my parents for supporting me whenever I needed.

Finally, I would like thank to Mehmet Ozansoy for his willingness to kindly answer every question related with the biological basis of this thesis.

This work is supported by the Turkish State Planning Organization (DPT) under the TAM Project, number 2007K120610.

ABSTRACT

MODELLING THE POPULATION DYNAMICS IN A CELL CULTURE AT TWO DIFFERENT SCALES

In this thesis the development of three novel models at two different scales are presented for population dynamics in cell cultures. Biological knowledge and empirical observations are used to design an agent-based discrete-time model at meso-scale, which then serves as a simulation environment and provides the necessary insight for lumped-parameter models at macro-scale. After demonstrating on basis of meso-scale simulation results that the flask-wide distribution of the population does not consistently become heterogeneous it is concluded that the population dynamics can also be represented at macro-scale. Two continuous time, differential equation-based, compact macro-scale models are developed. Both macro-scale models can be parameter-tuned and employed for predicting the evolution of the population size for given uniformly distributed initial populations. The thesis provides a procedure for estimating the parameter values of the macro-scale models via some simple tests to be conducted on the cell culture at hand. How well the macro-scale models can predict the evolution of the population size in comparison to the Meso-scale Model is evaluated on basis of four practically significant criteria. Furthermore; the robustness of the macro-scale models with respect to different initial energy distributions is evaluated. Finally, a philosophical perspective about modelling dynamic phenomena at different scales and how to deal with modelling challenges are presented.

ÖZET

HÜCRE KÜLTÜRÜNDE NÜFUS DİNAMİĞİNİN İKİ FARKLI ÖLÇEKTE MODELLENMESİ

Bu tezde, hücre kültürlerindeki nüfus dinamiğini iki farklı ölçekte betimleyen üç yeni modelin nasıl geliştirildiği sunulmaktadır. Biyolojik bilgi ve deneysel gözlemler kullanılarak, mezo-ölçekli ayrık zamanlı ve etmene dayalı bir model tasarlanmıştır. Bir benzetim ortamı işlevi gören bu model, makro-ölçekli toplu parametreli modellerin oluşturulması için gereken anlayışı sağlamıştır. Mezo-ölçekli modelin uzaysal dağılımının zaman içinde giderek daha heterojen hale gelmediği mezo-ölçekli benzetim sonuçlarıyla gösterildikten sonra, nüfus dinamiğinin makro-ölçekte de temsil edilebileceği sonucuna varılmıştır. İki adet sürekli zamanlı, diferansiyel denkleme dayalı ve kompakt makro-ölçekli model geliştirilmiştir. Bu iki makro-ölçekli model, parametreleri kestirildikten sonra, homojen dapılmış belli başlangıç nüfusundan yola çıkılarak nüfusun nasıl değişeceğini kestirmek için kullanılabilir. Tezde, eldeki hücre kültürü üzerinde yapılacak bazı basit testlerle makro-ölçekli modellerin parametrelerinin nasıl kestirileceğine dair bir prosedür sunulmuştur. Makro-ölçekli modellerin nüfusun değişimini Mezo-ölçekli Modele kıyasla ne kadar iyi öngörebildiği, pratik açıdan anlamlı dört kriter üzerinden değerlendirilmiştir. Ayrıca makro-ölçekli modellerin enerji dağılımlarındaki farklılıklara karşı ne ölçüde gürbüz olduğu belirlenmiştir. Son olarak, dinamik olguların farklı ölçeklerde modellemesine dair felsefi bir bakış açısı ve modellemede karşılaşılan zorluklarla nasıl başedildiği anlatılmıştır.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS

A	Rate for energy intake in Meso-scale Model
b	the average number of empty loci per cell.
D	Diffusion constant
E_{div}	Division energy in Meso-scale Mode
$\hat{E}_{div}$	Division energy in Macro-scale Models
$E_{i,j}$	Energy budget of the cell in locus [i j]
E_{rem}	Remediate energy of the cell in locus [i j]
f	Energy distribution of cells across the flask
g	Extended energy distribution of cells
H	Spatial entropy
k	Discrete time step in Meso-scale Model
M	Maximum number of cells a flask can hold.
n	Side of the n×n flask in Meso-scale Model
p	Total population
p_q	Onset population
q	Total quiescent population
R	Robustness measure
r	Coefficient of θ
t	Continuous time in Macro-scale Model
t_d	time when population starts to decrease
t_e	time when population goes extinct
$x_{i,j}$	Local toxicity in locus [i j]
x_{rel}	Toxicity release rate of a cell
α	Energy intake rate in Macro-scale Models
β	Toxicity release rate coefficient in Meso-scale Model
γ	Remediate energy coefficient in Meso-scale Model
$\hat{\gamma}$	Toxicity remediate rate coefficient in Macro-scale Models
Δk	Step size of physiological dynamics

Δk_D	Step size of diffusion dynamics
$d\triangle_{die}$	Fraction of cells, that will die at the next time instance
$d\triangle_{div}$	Fraction of cells, that will divide at the next time instance
ϵ_i^*	Average of errors over population size
$\overline{\epsilon}_i^*$	Initial-distribution-invariant average performance
$\%\epsilon_{pmax}$	Percentage error in the maximum population occupancy
ϵ_q	Average absolute quiescent estimation error
$\dot{\varepsilon}_{rem}$	Average toxicity remediate rate
$\%\epsilon_{RMS}$	Percentage error in RMS value of the Meso-scale population size
$\%\epsilon_{td}$	Percentage error in t_d
$\%\epsilon_{te}$	Percentage error in t_e
ζ	Total toxicity
η_H	Upper boundary of energy distribution
η_L	Lower boundary of energy distribution
η_S	Net energy intake of a 'virtual' cell
θ	Scaled version of total toxicity
κ	Slope of the quiescent-total population line
ξ	Amount of toxicity an average cell is exposed to
ρ_b	Birth rate
ρ_d	Death rate
ψ	Heterogeneity measure
ω	Toxicity release rate coefficient in Macro-scale Models

LIST OF ACRONYMS/ABBREVIATIONS

2D	Two Dimensional
3D	Three Dimensional
idiap	Initial-distribution-invariant average performance

1. INTRODUCTION

1.1. Overview

The aim of this thesis is to set out a modelling inquiry about a biological phenomenon by mathematical representation of observations at different organisational levels. For this purpose, we need to go back and forth between these different levels to develop compact analytical models. The contribution of this thesis consists of three models: a deterministic agent-based Meso-scale Model which is an improved version of an earlier work by Aksu [3], and two deterministic analytical macro-scale models which describe the cell population dynamics at flask level. In order to develop such macro-scale models, one still needs some insight about lower-scale dynamics. The Meso-scale Model has been developed first as a thought-experiment to fulfil this requirement and then has became a comprehensive simulation environment which allows the observation of cell-level dynamics.

In this thesis, the object of modelling inquiry is population dynamics in in-vitro cell cultures. A cell culture is the process of the removal of cells from an animal or plant and their subsequent growth in a favourable artificial environment [4]. Tissues can be dissociated into their component cells, from which individual cell types can be purified and used for biochemical analysis or for the establishment of cell cultures [1]. Many biological and medical laboratories use cell cultures in their studies, some examples of which are effects of new drugs, cancer research, impact of diseases on cells, gene therapy, vaccine production etc. Due to practical concerns, the study is confined to eukaryotic cell cultures. An analytic model of the population dynamics in cell cultures can be helpful in experiment design or in some cases might even let the scientist to reach some useful conclusions without conducting any real experiment. Lastly, it is not inaccurate to claim that the analytical models can serve as an inspirational tool for the generation of new Anstze about phenomena under investigation.

In the specific domain of cell culture dynamics most of the studies investigate

either the dynamics of individual cells or flask-wide population dynamics, commonly referred to as micro-scale and macro-scale respectively. However, there exists at least another scale that harbours dynamics worthy of scientific inquiry, namely the scale of intracellular dynamics. Although we will not attempt to model this dynamics and take it for granted when modelling the behaviour of individual cells, from the point of modelling philosophy we find it important to draw attention to its existence. Thus, in this study we will slightly diverge from the conventional terminology and refer to three different scales: micro-scale, the level of intracellular activities; meso-scale, the level of individual cells and their interactions; and macro-scale, the flask-wide level. In that respect, we want to warn the reader that the" meso-scale" in this contribution corresponds to what is commonly denoted as micro-scale in the literature.

The major contributions in this thesis as shown in Figure 1.1 can be summarised as follows:

(i) Improvement of a formerly developed agent-based distributed meso-scale model that mimics the population dynamics in a cell culture.
(ii) Two macro-scale analytical models (one of them being four dimensional and the other one five dimensional) that represent the flask-wide population dynamics in terms of macro state variables.
(iii) A heterogeneity measure is developed to investigate the evolution of clusters in the meso-scale model.

It is also worth to denote that although the apparent aim of this thesis is the prediction of the evolution of population size in a cell culture starting from an arbitrary initial value, the experience and the insight gained from the modelling challenges throughout the thesis has provided me with an invaluable knowledge and line of thought for future modelling inquiries.

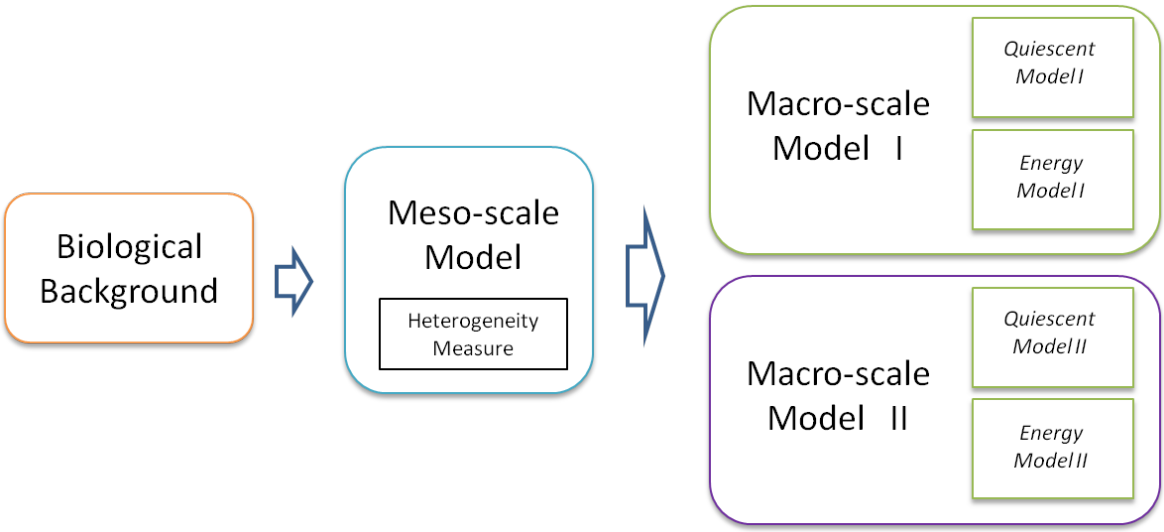

Figure 1.1. Major contributions of the thesis.

1.2. Literature Survey

The survey of the related literature can be mainly divided into two parts. The first part briefly summarizes the important major steps in the history of the biological basis of this thesis, namely cell culture studies. The second part is about the population dynamics, especially about the modelling studies of cell cultures.

Cell culture studies date back to the last decades of 19th century. The important landmarks in the development of cell cultures since then until the end of 20th century are well-summarised in [5] as follows: Roux shows that embryonic chick cells can be maintained alive in a saline solution outside the animal body (1885). Harrison cultivates amphibian spinal cord in a lymph clot, thereby demonstrating that axons are produced as extensions of single nerve cells (1907). Rous induces a tumor by using a filtered extract of chicken tumor cells, later shown to contain an RNA virus (1910). Carrel shows that cells can grow for long periods in culture provided they are fed regularly under aseptic conditions (1913). Earle and colleagues isolate single cells of the L cell line and show that they form clones of cells in tissue culture (1948). Gey and colleagues establish a continuous line of cells derived from a human cervical carcinoma, which later become the well-known HeLa cell line (1952). Levi-Montalcini and associates show that nerve growth factor (NGF) stimulates the growth of axons

in tissue culture (1954). Eagle makes the first systematic investigation of the essential nutritional requirements of cells in culture and finds that animal cells can propagate in a defined mixture of small molecules supplemented with a small proportion of serum proteins (1955). Puck and associates select mutants with altered growth requirements from cultures of HeLa cells (1956). Temin and Rubin develop a quantitative assay for the infection of chick cells in culture by purified Rous sarcoma virus. In the following decade the characteristics of this and other types of viral transformation are established by Stoker, Dulbecco, Green, and other virologists (1958). Hayflick and Moorhead show that human fibroblasts die after a finite number of divisions in culture (1961). Littlefield introduces HAT medium for the selective growth of somatic cell hybrids and Kato and Takeuchi obtain a complete carrot plant from a single carrot root cell in tissue culture (1964). Ham introduces a defined, serum-free medium able to support the clonal growth of certain mammalian cells and Harris and Watkins produce the first heterocaryons of mammalian cells by the virus-induced fusion of human and mouse cells. (1965). Augusti-Tocco and Sato adapt a mouse nerve cell tumor (neuroblastoma) to tissue culture and isolate clones that are electrically excitable and that extend nerve processes. A number of other differentiated cell lines are isolated at about this time, including skeletal muscle and liver cell lines (1968). Khler and Milstein produce the first monoclonal antibody-secreting hybridoma cell lines (1975). Sato and associates publish the first of a series of papers showing that different cell lines require different mixtures of hormones and growth factors to grow in serum-free medium (1976). Wigler and Axel and their associates develop an efficient method for introducing single-copy mammalian genes into cultured cells, adapting an earlier method developed by Graham and van der Eb (1977). Martin and Evans and colleagues isolate and culture pluripotent embryonic stem cells from mouse (1986). Thomson and Gearhart and their associates isolate human embryonic stem cells (1998).

The roots of population dynamics as a scientific inquiry goes back to more than two hundred years ago, to Malthus. In 1798, Malthus claimed that the increase in population and resources cannot grow with the same rate and elaborated his reasoning and one of his famous conclusions that the difference between their growth behaviour would eventually stop the population growth with his famous work "An Essay on the

Principle of Population" [6]. Since then, population dynamics has grown as a huge branch of mathematical biology, investigating how population size of a species change with respect to time and space and how it is affected by resources, other species and its environment with the help of various mathematical tools and approaches. Considering that this thesis aims to develop mathematical and computational models for cell culture populations, we should confine this literature survey only to cell culture dynamics.

Since 1960ies a wide range of approaches have been suggested in the literature for modelling population dynamics in cell cultures. The modelling approaches in these studies are typically based on either ordinary differential/difference equations or agent-based models that can only be realised by computer simulations.

In meso-scale, usually agent-based models are used but there exist also some Meso-scale Models that use differential equations as mentioned by M. A. Henson [7]. In this scale one can mention the studies of Zygourakis *et al.* [8] who have developed a cellular automaton model that represents contact inhibition in anchorage-dependent cells, and Cheng *et al.* [9] who describe three dimensional tissue growth including a stochastic model of cell migration. Some other studies on cellular automaton-like models have addressed multi-cellular spheroid tumour growth [10], proliferation dynamics of populations of migrating contact-inhibited cells [11], static and dynamic environments related to cell migration [12], and stochastic representation of cell migration, proliferation and differentiation [13].

On the other hand, macro-scale models involve almost exclusively differential or difference equations. M.A. Henson's survey paper [7] summarises different differential or difference equation-based approaches to cell population modelling in microbial cultures comprised of heterogeneous cells that differ in terms of their size and intracellular concentrations. Henson mainly distinguishes between two approaches, namely population balance equations which are characterized by one or two variables (usually cell age or cell mass) and cell ensemble models which use different state variables for concentrations of various species in all cellular processes. A. Ducrot *et al.* present a model to describe the spatial density and movement of cells in a mono-layer cell cul-

ture based on quiescence and contact inhibition phenomena [14]. There are also other studies that describe either a specific phenomenon in cellular activity [15] or propose general models for population dynamics in a cell line [16–19]. Finally, it is worth mentioning that P. Auger *et al.* review the transformation of a complex system with many variables to a reduced model with fewer parameters from a purely mathematical point of view, and present aggregation methods for ordinary differential equations, partial differential equations and discrete models [20].

However, there are only a few studies that include both agent-based and differential equation-based models. R. Serra *et al.* have combined these two modelling methods to describe the transformation foci in cell cultures, i.e. an in vitro analogue of tumors [21]. C. A. Chung *et al.* have developed hybrid mathematical models where the cell proliferation, aggregation, contact inhibition and random walk are described by a discrete cellular automaton, while the transport of nutrient is represented with partial differential equations [22, 23].

The outline of the thesis is as follows: In 1^{th} chapter, the biological background of the cell cultures is given briefly. The meso-scale model which is developed on the basis of the biological background will be presented in 2^{nd} chapter. An heterogeneity measure to evaluate the suitability of the meso-scale model for a representation with lumped-parameters is introduced also in this chapter. In 3^{th} chapter, two macro-scale models are developed and compared with the meso-scale model with respect to some performance criteria. Furthermore, a robustness test of macro-scale models to different initial energy distributions is conducted in this chapter. Finally, some conclusive remarks, including our modelling perspective, are presented in the 4^{th} chapter of the thesis.

1.3. Biological Background

Before introducing the models, we have to present a brief summary of the life cycle of a cell. Within its life-time of a eukaryotic cell one can distinguish roughly three different operational modes: active state, quiescence state and differentiated

state. A healthy cell, when it is in a sparse neighbourhood, is mitotically active, i.e. it can grow and after reaching the mature size and acquiring sufficient energy, undergo mitotic division. The life cycle of an active cell can be divided in four phases as shown in (Figure 1.2) : G1 phase, where cells grow and prepare for DNS synthesis; S phase, where the DNA is duplicated; G2 phase, where cells continue to grow and make final preparations for mitotic division and finally M phase where the nuclear division (mitosis) and cell division (cytokinesis) occur (Figure 1.3).

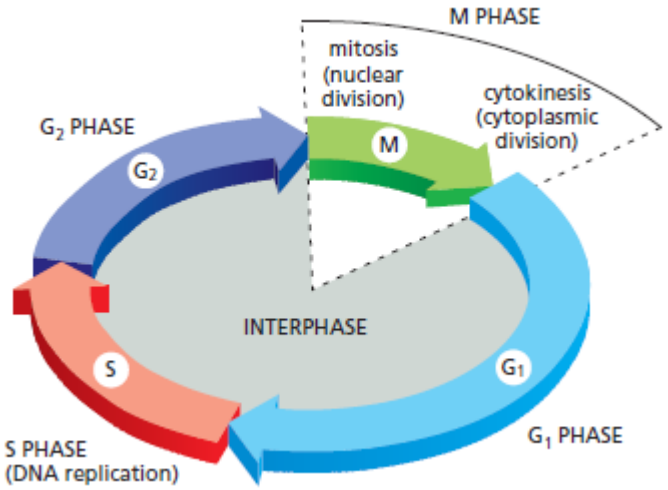

Figure 1.2. The four phases of the cell cycle [1].

However, if cell is in an overcrowded region, due to the tendency of contact inhibition, it enters the so-called quiescence state where it halts growth and reproduction activities. Cells of multi-cellular organisms can eventually enter the phase of senescence where they lose their reproductive ability, however cells used in in-vitro studies are generally immortalized by genetic modification such that senescence phase can be left of out of the scope in such studies. Cells of multi-cellular organisms can also undergo differentiation and transform themselves into more specialised cell types in an irreversible manner, however genetically modified cell lines used in in-vitro cell studies cannot differentiate unless externally induced, hence differentiation and differentiated cells are also left out of the scope of this study. A cell activity common to all operational modes is toxicity release; like all living beings cells pollute their environments with their toxic release, and are in turn affected by this toxicity [5].

The time profile of population growth for a wide range of cell cultures exhibit

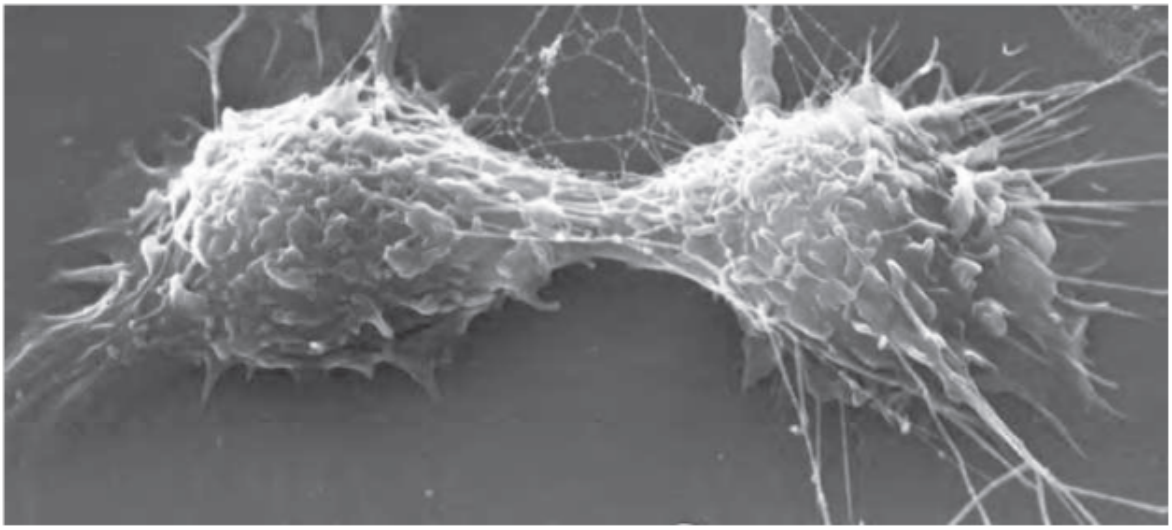

Figure 1.3. A scanning electron micrograph of a cultured animal cell dividing [1].

a common curve. Population increase is very slow or does not even occur at the beginning of the experiment. This phase, the duration of which depends on the culture conditions, is called "lag phase". Lag phase is followed first by an exponential and then by a declining population increase when toxicity is accumulated sufficiently (log phase). After this phase, they enter a stationary phase (plateau) where population growth completely stops and finally they rapidly start to die. In Figure 1.4, an example of a real data set collected from two different mammalian cell cultures are shown.

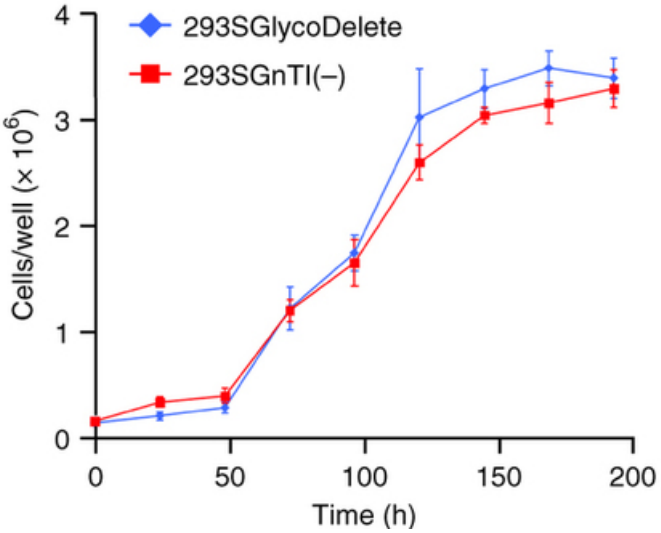

Figure 1.4. Real data of the growth curve for 293SGnTI(-) and 293SGlycoDelete cells in 6-well culture plates, counted every 24 h. [2].

2. MESO-SCALE MODEL

The Meso-scale Model is a discrete-time, agent-based model of the cell culture dynamics and based on the biological knowledge given in section 1.3. Here, each cell is considered as an agent' governed by a set of rules. The flask is represented as a n×n grid, where each mesh -referred to as 'locus'- can accommodate only one cell (Figure 2.1).

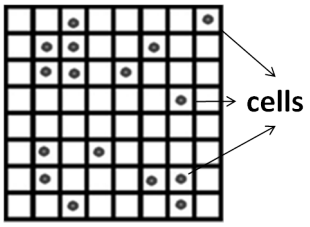

Figure 2.1. $n \times n$ grid representing the flask, each mesh referred to as 'locus'.

In this model the instantaneous (at the k$^{\text{th}}$ time instance) state of the system is represented by the toxicities at all loci ($x_{i,j}(k); \forall i, j = 1, \ldots, N$) and the energies of the cells occupying these loci ($(E_{i,j}(k); \forall i, j = 1, \ldots, N$).

The rules governing these variables are presented in Sections 2.1 and 2.2.

2.1. Cellular Dynamics

In order to obtain a compact set of rules governing a cell's behaviour as far as possible each cell has been characterised only by two properties: its locus (i, j) which serves as its identity and its energy $(E_{i,j}(k))$. In this model the cells are either in 'active' or in 'quiescent' mode, depending on how crowded their neighbourhood is. It should be noted that in this model the energy of a cell is an abstract variable that stands for a combination of its size and its physical/chemical energy. This energy is involved in the following processes:

- *Growth*: Any cell in the active mode takes in a constant amount of energy $E_{in} = A$ at each time step. The nutrition concentration in the flask is assumed to remain

constant[1] and uniform (equivalent to the assumption that the initial nutrient amount is high enough such that its decrease during the life of the cell culture is negligible).

- *Aging*: Although there are various factors contributing to the aging of a real cell, the dominating one in a cell culture is oxidative stress, i.e. the damage caused by exposure to toxicity. Therefore, in this model cell aging is represented in terms of the effect of local toxicity on the cell's energy. It is assumed that at each time step the cell needs to spend some energy (E_{rem}) in order to remediate the damage caused by local toxicity. As a first approach, this remediation energy spent by a cell at locus (i, j) at turn k is taken proportional to the local toxicity $x_{i,j}(k)$:

$$E_{rem}(k) = \gamma x_{i,j}(k) \tag{2.1}$$

  where γ is the remediation constant. If $E_{rem} > A$, the cell's energy starts to decrease.

- *Proliferation*: An active cell, when its energy attains the 'division energy' (E_{div}), undergoes mitotic division and the two daughter cells share this energy equally ($E_{div}/2$). One of the daughter cells is placed at the former locus of the mother, while the other one is randomly placed at any empty locus among the mother's 8-nearest-neighbourhood. It is worth noting that the random placement of the daughter cells within this close neighbourhood constitutes the only stochastic element in the Meso-scale Model.

- *Death*: Although biologically speaking death is a rather complex process, in this model it is represented as the cell's energy going down to zero after a gradual decrease when the remediation expenditures exceed the energy intake.

- *Contact inhibition and quiescence*: Cells' tendency of avoiding too much contact by halting growth and division is represented in our model as follows: If all 8 nearest loci around a cell are fully occupied it enters the 'quiescence mode', i.e. it limits the energy intake such that its energy cannot increase and thus reproduction becomes impossible. In this mode, if $E_{rem} > A$ the cell loses energy

[1] This assumption is not a necessary condition for the presented Meso-scale Model and can be easily be relaxed for modelling experiments of long duration where decrease of nutrition cannot be neglected.

although it does not limit the energy intake. Note also that a quiescent cell can return to active mode if some neighbouring loci are emptied due to deaths.

- *Initial Energy Distribution*: Considering that 'energy' of a cell is a virtual concept, it is unrealistic to try to measure the exact value of a cell's energy. Instead, the initial energies of cells have been assumed to obey a statistical distribution.

2.2. Toxicity Dynamics

Toxicity is coupled to the cell dynamics both via toxicity release from each cell and via cells' energy loss while trying to remediate the damages caused by exposure to toxicity. These interactions can be represented as follows:

- Each living cell releases a constant amount of toxicity per time step. For the sake of convenience, this toxicity release rate (x_{rel}) has been expressed as a fraction of the constant energy intake (A) per time step, i.e. $x_{rel} = \beta A$, where $0 < \beta < 1$.
- Released toxicity diffuses according to the well-known diffusion equation:

$$\frac{\partial x_{i,j}(t)}{\partial t} = D\left(\frac{\partial^2 x_{i,j}(t)}{\partial i^2} + \frac{\partial^2 x_{i,j}(t)}{\partial j^2}\right) \tag{2.2}$$

where D is the so-called diffusion constant. This equation can be discretised by the so called Forward-Time Centred-Space (FTCS) [24] method as follows:

$$\begin{aligned} x_{i,j}(k + \Delta k_D) =& D\Delta k_D(x_{i+1,j}(k) + x_{i-1,j}(k)) + (1 - 4D\Delta k_D)x_{i,j}(k) \\ &+ D\Delta k_D(x_{i,j+1}(k) + x_{i,j-1}(k)) \end{aligned} \tag{2.3}$$

where $\Delta i = \Delta j = 1$ due to discrete nature of our model and Δk_D represents the discrete diffusion time.

2.3. Choice of Parameters

In the simulations presented in this thesis, the parameters of the Meso-scale Model (D, β, γ, A and E_{div}) have been heuristically tuned to make the evolution of

the simulated population size resemble some empirical data obtained from a modest number of C2C12 cell culture experiments, as well as biological expectations. Thus, the parameters used in the Meso-scale Model simulations (Table 2.1) can be said to lie at least within realistic range.

Table 2.1. Parameters in the Meso-scale Model tuned heuristically to obtain biologically plausible results.

D	β	γ	A	E_{div}	Δk
4	0.8	0.008375	1	96	15 min

The assumptions and biological knowledge used in the setting of these parameters are summarised below:

- As a reference point to other parameters, the energy intake per time step A is selected to be equal to 1 and each time step in simulations is assumed to correspond to 15 minutes in real life experiments.
- At the beginning of experiments, cells have relatively small energies so that cell division is not observed immediately. In order to satisfy this condition and biological plausibility, we have selected the shape and boundaries of initial energy distribution as a Gaussian curve with mean $\mu = E_{div}/4$, standard deviation $\sigma = E_{div}/8$ and shifted downwards so that it takes positive values only between 0 and $E_{div}/2$ (Figure 2.2).
- Doubling time of a cell is typically given as approximately 12 hours by manufacturers. Using this knowledge with A and Δk values and assuming that toxicity is negligible, one can derive E_{div} as follows:

$$E(0) = \frac{E_{div}}{2}$$
$$E(k) = kA\Delta k + \frac{E_{div}}{2} = E_{div} \rightarrow E_{div} = 2kA\Delta k$$
$$k = \frac{12h}{15min} = 48 \rightarrow E_{div} = 96$$

- D, β and γ are tuned to obey the following experimental observation: When initial population size is approximately 10% of carrying capacity and toxicity is

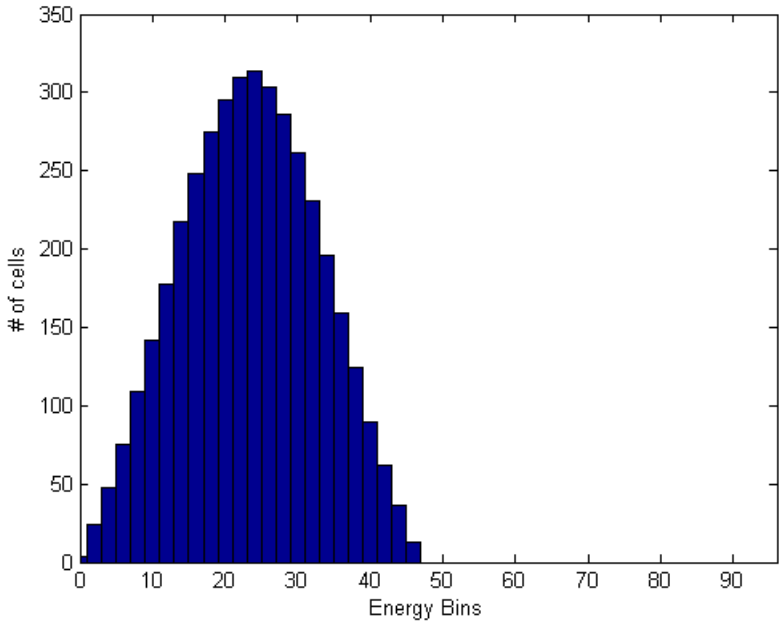

Figure 2.2. Gaussian-shaped initial energy distribution used to set the parameters of Meso-scale Model.

removed on the 3rd, 6th and 9th days of the experiment, cells fill the flask basin nearly completely and go extinct in a few days after the last removal. Simulation result of the Meso-scale Model in accordance with this observation is given in Figure 2.3.

However, it should be kept in mind that the correctness of the parameter values for a specific cell culture type and the method used in their tuning are irrelevant, because in this thesis the Meso-scale Model is meant to be an intermediary step for the development of a macro-scale model.

2.4. Simulation Results

Figures 2.4 to 2.10 show simulation results of the Meso-scale Model in terms of total population, total toxicity and flask views of these two variables starting from an initial population occupying 10% of the flask and zero toxicity. Cells are initially randomly uniformly distributed and the initial energy distribution is selected as in Figure 2.2. Toxicity is not removed from the flask on 3rd, 6th and 9th days which

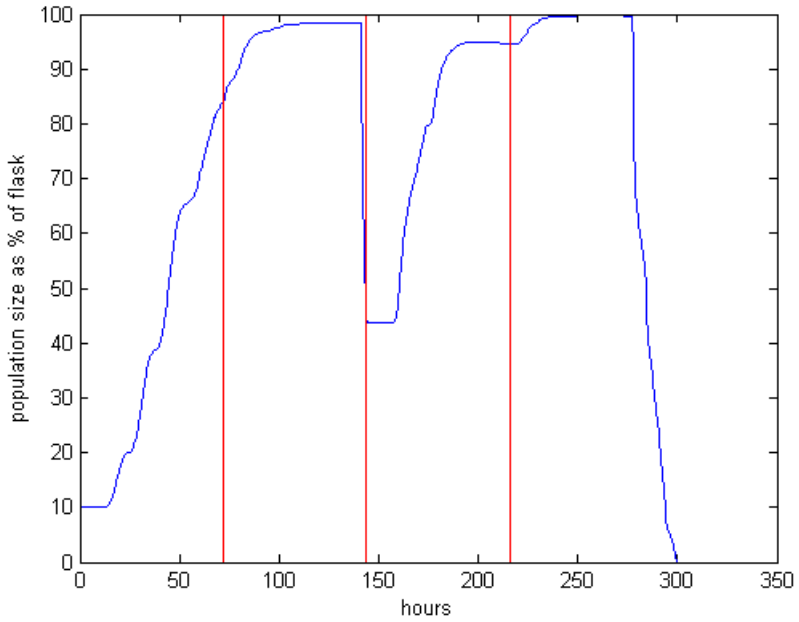

Figure 2.3. Time profile of total population. Red line represents toxicity removal on the 3^{rd}, 6^{th} and 9^{th} days.

results in a strictly increasing toxicity. Note that, the first population increase is observed when upper boundary of energy distribution reached E_{div} which occurs at around 50^{th} time step.

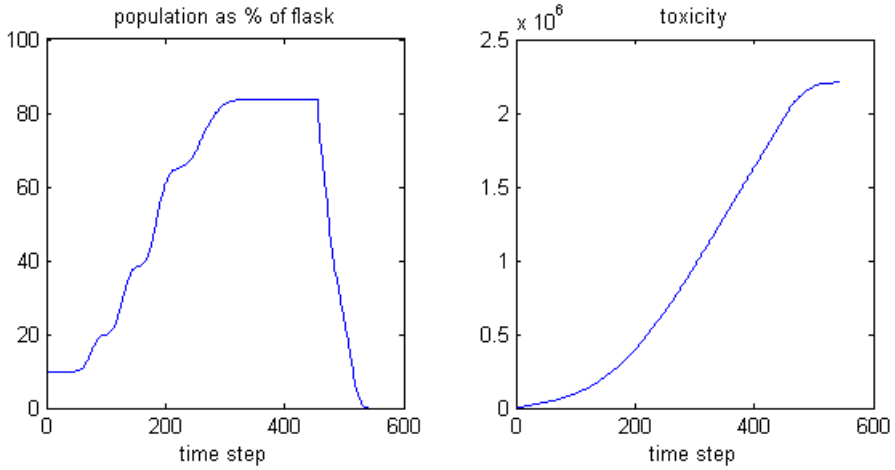

Figure 2.4. Time profiles of total population as % of flask and total toxicity.

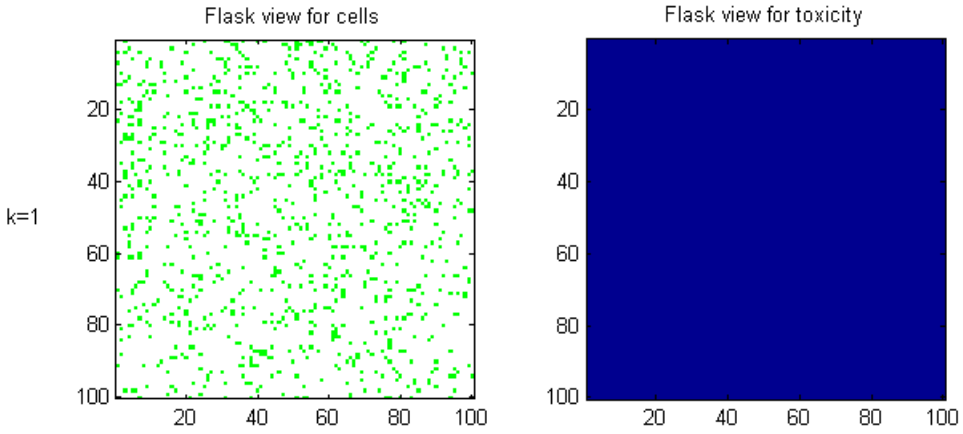

Figure 2.5. Flask views for cells and toxicity at 1st time step.

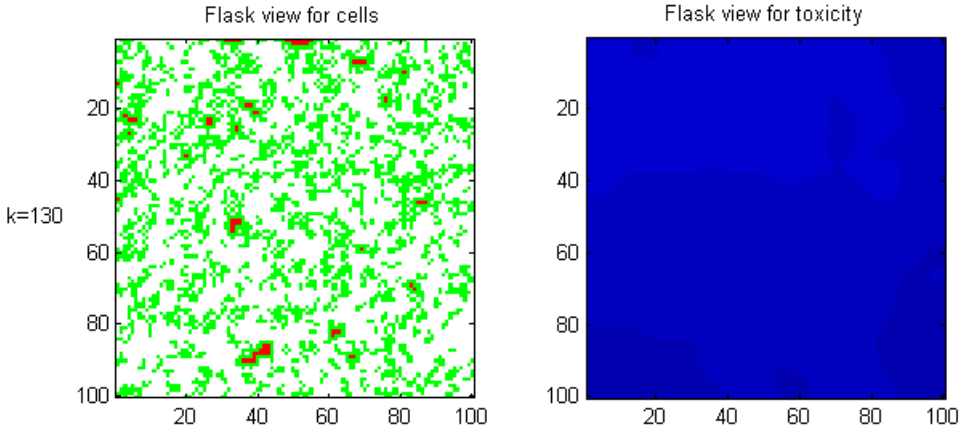

Figure 2.6. Flask views for cells and toxicity at 130th time step.

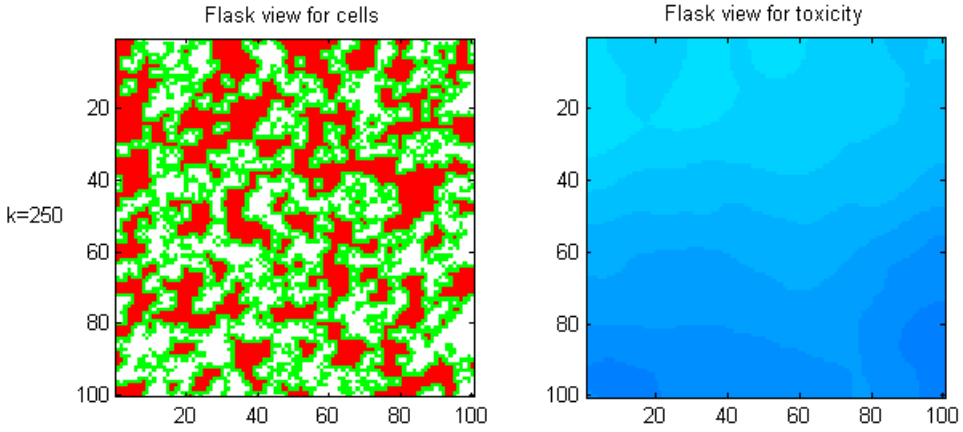

Figure 2.7. Flask views for cells and toxicity at 250th time step.

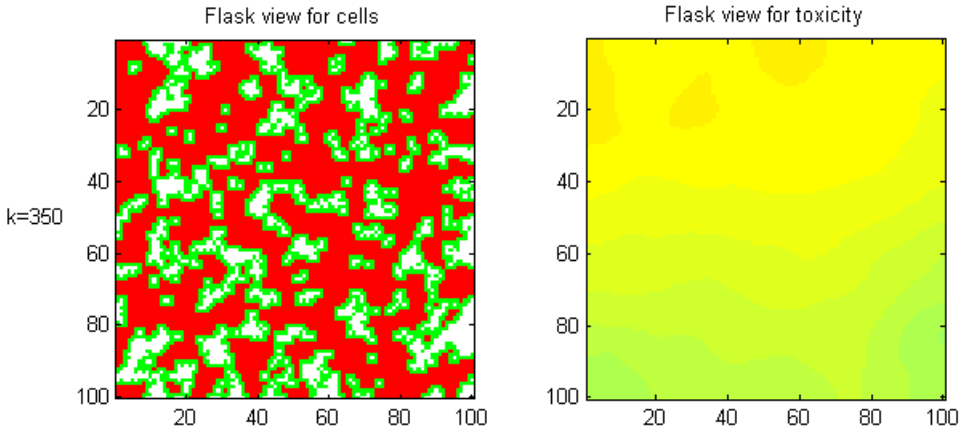

Figure 2.8. Flask views for cells and toxicity at 350$^{\text{th}}$ time step.

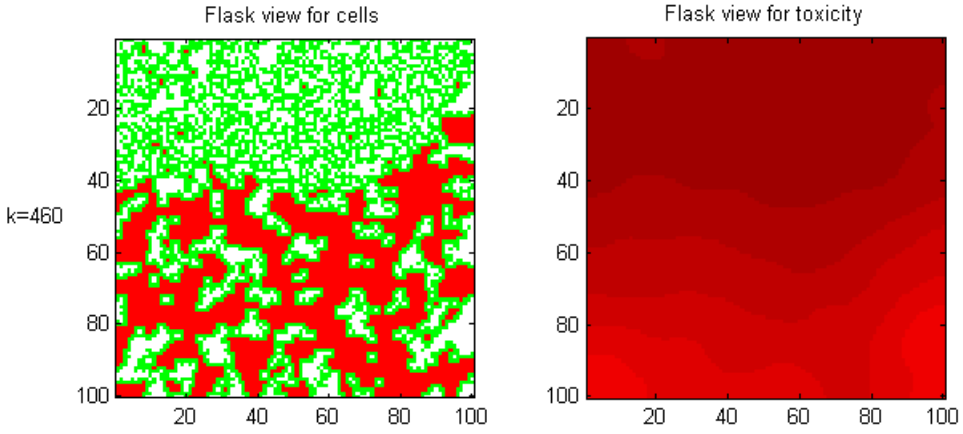

Figure 2.9. Flask views for cells and toxicity at 460$^{\text{th}}$ time step.

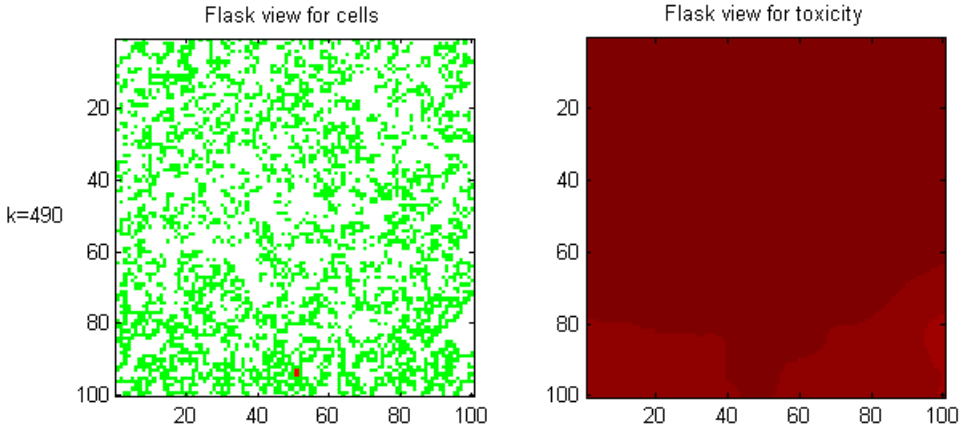

Figure 2.10. Flask views for cells and toxicity at 490$^{\text{th}}$ time step.

2.5. Assessment of Uniformity

In order to describe the dynamics of the Meso-scale Model via lumped parameters, the spatial distribution of individual entities should not affect the global dynamics of the ensemble. Emergent, complex aggregations and spatial patterns of individual entities might cause the dynamics of the ensemble to depend on local properties such that the concept of average cell loses its validity. Thus, our macro-scale models are valid only if cells in a cell culture are initially sufficiently uniformly distributed across the flask and remain so throughout the rest of the process. Sufficiently uniform initial distribution can be achieved by carefully spreading the population across the flask. But one still needs to verify that the dynamics does not amplify minor spatial heterogeneities. For that purpose a measure of heterogeneity is needed, the evolution of which under the given dynamics is to be investigated.

2.5.1. Heterogeneity Measurements

The evolution of the spatial heterogeneity of cells can be investigated via various methods but they either need a priori criteria or they have several weaknesses in different situations. In Sections 2.5.1.1 to 2.5.1.3, we summarise briefly which methods have been tried and developed to capture the evolution of heterogeneity. In Section 2.5.1.4, a novel approach will be explained which is free of the limitations of previous methods.

2.5.1.1. Entropy-based measure. The very first approach to consider for this purpose is the well known Shannon's entropy concept applied to spatial distributions. In this method, the flask is virtually divided into m boxes. The sum of likelihoods of cells being in these boxes gives the spatial entropy (Equation 2.4) and hence a measure for the uniformity of cell distribution.

$$H = -\sum_{i=1}^{m} p_i \ln p_i \tag{2.4}$$

where

$$p_i = \frac{number\ of\ cells\ in\ i^{th}\ box}{number\ of\ cells\ in\ flusk}$$

However, in order to apply this method, one must have a priori knowledge about the size of cell aggregations so that a proper box size can be selected for Equation 2.4.

2.5.1.2. Method of 'Hills'. In this thesis, we have proposed another approach for the spatial heterogeneity measurement by considering the flask as a rippled surface where 2D Gaussian distributions are placed onto each cell. The overall smoothness of this virtual topology gives us an indication for the uniformity of the cell distribution. Due to their nonlocal property, fractional derivatives are used to measure the ruggedness of the surface in all directions. However, standard deviations of 2D Gaussian distributions and the fractional degree of derivatives which should be determined before using this method affect this measure significantly. Selection of these two parameters needs further research and hence this approach has not been preferred as a measure of heterogeneity.

2.5.1.3. Variety of neighbours. Another consideration to measure the uniformity of a spatial distribution is to investigate how the number of neighbours of a cell varies across the population. This can be measured in terms of the standard deviation of number of neighbours of cells. A low standard deviation indicates a relatively uniform distribution whereas local heterogeneities give rise to high standard deviation. However, this approach focuses on the immediate neighbourhood of cells and thus misses the distribution of cell clusters across the flask.

2.5.1.4. Artificial Force Field. The situational performance of the above mentioned approaches make us to seek another method, which has relatively consistent performance and does not ask for a priori knowledge of some parameters.

In this approach, we assume that cells inside the flask are equicharged particles.

The aim is to express the heterogeneity of the distribution in terms of the overall tension that results from the special distribution of these 'charged' cells. In order to guarantee that perfectly uniform distribution across the flask corresponds to zero tension, the flask is assumed to be embedded in an infinite plane filled with uniform charge density equal to the average charge density inside the flask. If cells form clusters or deviate from their equilibrium distribution, then the net tension takes a positive value which can be used as a measure of uniformity of cell distribution.

Since we also want this measure to be independent of the population size, we can use the first moment (average) or second moment (standard deviation) of forces instead of the sum of the force magnitudes. The force acting on the i$^{\text{th}}$ cell applied by the j$^{\text{th}}$ cell is defined as:

$$\vec{F}_{ij} = \frac{\vec{x}_i - \vec{x}_j}{\|\vec{x}_i - \vec{x}_j\|} \; \frac{Q_i Q_j}{\|\vec{x}_i - \vec{x}_j\|^2} \tag{2.5}$$

where $\vec{x}_i$ is the location of the i$^{\text{th}}$ cell inside flask; $\vec{x}_j$ is the location of the j$^{\text{th}}$ cell inside flask or virtual space; $\vec{Q}_i$ is the charge of the i$^{\text{th}}$ cell inside flask; $\vec{Q}_j$ is the charge of the j$^{\text{th}}$ cell inside flask or virtual space and $\|.\|$ is the Euclidean norm of the two dimensional vector.

Although both first and second moments can be used as a heterogeneity measure, for the sake of simplicity, we decided to use the first moment of force magnitudes. The new heterogeneity measure ψ is proposed as the average of the magnitude of the net force applied on each cell, both by cells inside the flask and by the cells in the infinite plane enclosing the flask:

$$\psi = \frac{1}{N} \sum_{i=1}^{N} \left\| \sum_{\forall j \neq i \; ; \; j \in U} \vec{F}_{i,j} \right\| \tag{2.6}$$

where N is the total population of cells in the flask and U is the cells inside and outside the flask.

In case cells are distributed perfectly uniformly across the flask, all forces applied to each cell would negate each other and in such a scenario ψ becomes zero. Some trivial examples with their visualisations are shown below in order to obtain a better understanding of this method. Note that although the heterogeneity measure is selected as the mean of force magnitudes, in these examples we also provide the standard deviation of net forces ψ_{std} to show that both the mean and standard deviation are consistent within the scope of our purpose and can be used in place of each other.

- *Example 1*: Consider an 8×8 flask with 10 cells distributed as follows:

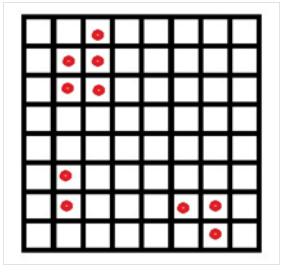

Figure 2.11. Example distribution with three clusters.

This flask is assumed to be placed inside an infinite virtual space full of charged particles as shown in Figure 2.12.

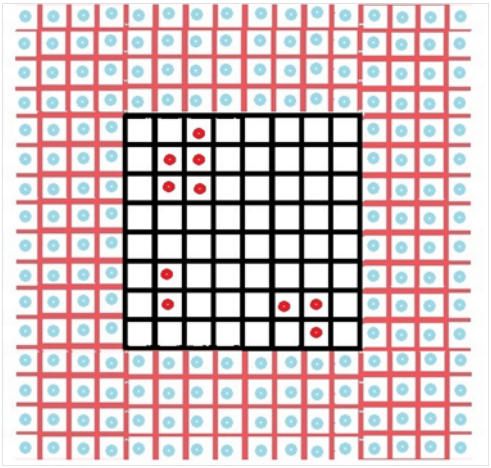

Figure 2.12. Flask is assumed to be embedded in an infinite plane.

The mean and standard deviation of force magnitudes are calculated as follows:

$$\psi_{mean} = 2.1370$$

$$\psi_{std} = 1.005$$

- *Example 2*: Now, assume that the particles have a more uniform distribution:

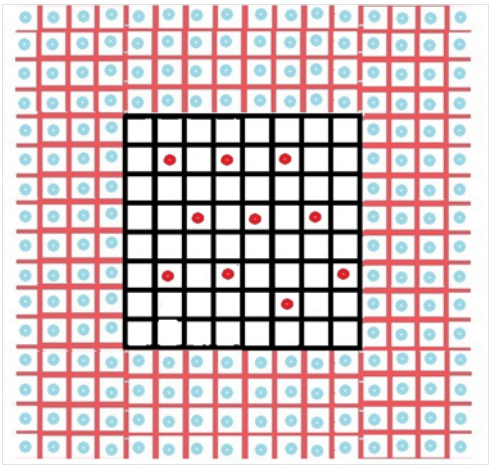

Figure 2.13. Relatively uniform distribution embedded in an infinite plane.

The heterogeneity measurements of this distribution are calculated below:

$$\psi_{mean} = 1.0175$$

$$\psi_{std} = 0.3978$$

As expected, both mean and standard deviation of net forces of Example 1 is less than the ones of Example 2, indicating a more uniform distribution.

For the comparison of heterogeneities of populations in flasks of different size, we need to normalize ψ such that its maximum value for any flask and population size is one. Although we worked on it, we have not yet concluded how the theoretical maximum heterogeneity can be calculated for a specific population and flask size and left it as an open question. However it is important to mention that such a comparison between different flask sizes is out of the scope of this thesis, which aims to describe the dynamics of a cell culture population in a given flask.

This method, despite being very computationally intensive, describes the spatial heterogeneity of a population distribution in a more reliable manner than the other methods mentioned, and does not depend on a priori knowledge of some criteria. Therefore, we will use this method to evaluate the evolution of heterogeneity.

2.5.2. Evolution of heterogeneity

The main purpose of the heterogeneity measure is to create a tool to analyse how the spatial heterogeneity evolves under the population dynamics specified by the Meso-scale Model. Figure 2.14 shows how ψ evolves in a meso-scale simulation with perfectly uniform distributed initial energy between $\frac{E_{div}}{2}$ and E_{div}.

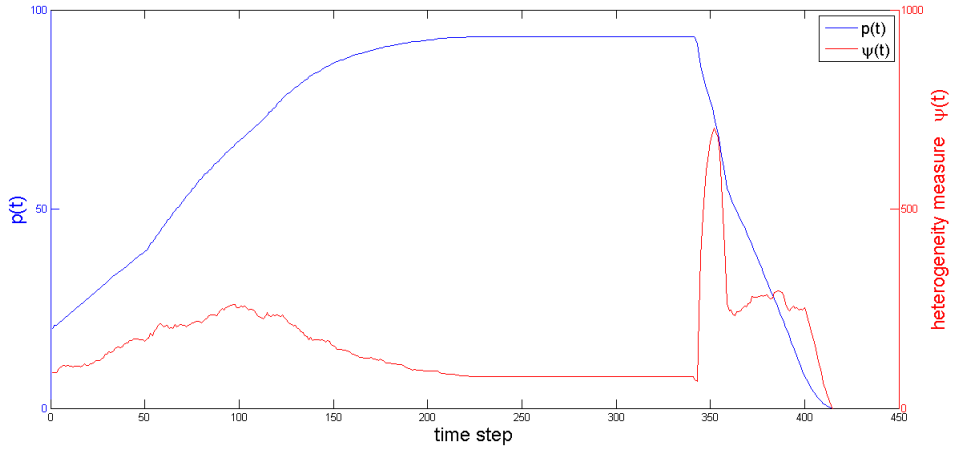

Figure 2.14. Total population occupation starting from 20% initial occupation and heterogeneity measure $\psi(t)$.

In order to have a more reliable evaluation of how heterogeneity evolves, the time profile of ψ is also obtained for initial concentrations 10%, 30% and 40% (Figure 2.15).

A close inspection of Figure 2.15 allows us to conclude that the dynamics represented by the Meso-scale Model does not give rise to a monotonous increase in the heterogeneity. All four initial populations exhibit a similar heterogeneity profile with a transient increase at the beginning, followed by a decline and then a sudden peaking behaviour when cells start to die. The temporal increase of ψ at the beginning of experiment is due to cell division which places the daughter cells in a close vicinity of the mother cell, thus creating local clusters. However, after a sufficiently many division, these clusters start to merge, thus contributing to the uniformity of the distribution and decreasing the heterogeneity. When cell deaths start at centres of former clusters due to local toxicity accumulation at those regions, large local heterogeneities are created at the beginning of the population decrease which is the main reason of the peaking

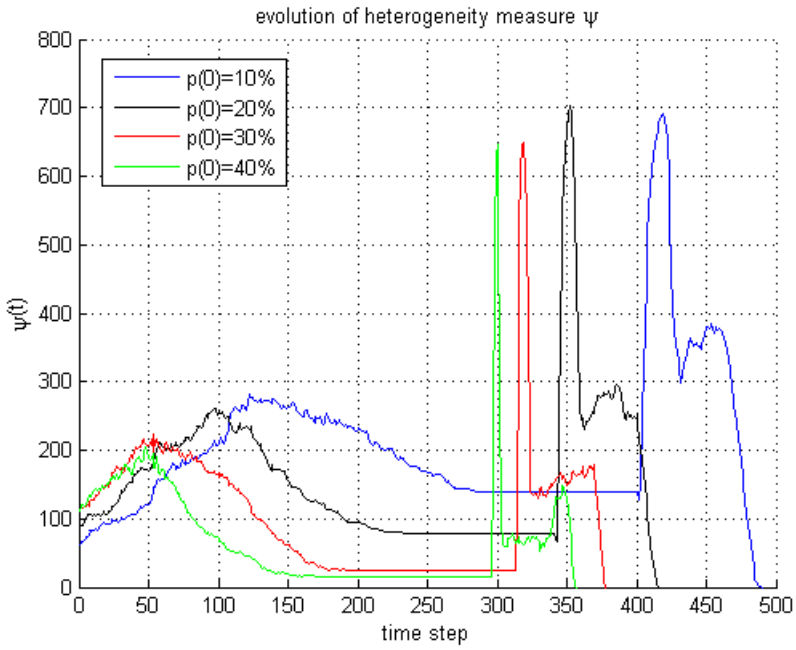

Figure 2.15. Time profile of heterogeneity measure ψ for four different initial occupations.

behaviour. After the effect of accumulated toxicity reaches to cells in relatively sparse areas and kill them, uniformity of distribution finally increases and all heterogeneities vanish, while population goes extinct.

In conclusion, we have observed that the dynamics of the Meso-scale Model does not amplify heterogeneities in the spatial distribution of cells continuously, making the Meso-scale Model feasible to describe in terms of macro-scale variables.

3. MACRO-SCALE MODEL

3.1. Introduction

Any attempt of modelling a real system has to deal with the problem of choosing the right resolution and perspective. The same system can be described in drastically different ways by different disciplines having different concerns, goals and interests. Even within the same discipline the same system can be modelled at different scales concentrating on different organisational levels. In this thesis, our main concern of predicting the population size in a cell culture asks for a flask-wide analytical macro-scale model. However, in order to develop such a model one has to have -either explicitly or implicitly- some hypotheses about the lower-scale dynamics. In the previous section, a biologically plausible agent-based Meso-scale Model has been developed for this purpose and in this section we will derive two macro-scale models from it by theoretical reasoning and some empirical observations from simulations of the Meso-scale Model.

In this chapter, we will provide the details of how these two macro-scale models are developed. The outline is as follows: The main concepts needed to develop macro-scale models will be presented in Section 3.2, the dynamics of the state variables in Section 3.3, their corresponding state equations in Section 3.4, how their parameters will be set in Section 3.5, the simulation results of both macro-scale models in Section 3.6, and finally the robustness of the macro-scale models to different initial conditions in Section 3.7.

3.2. Main Concepts of the Macro-scale Models

Here, a macro-scale model is supposed to provide a lumped-parameter, analytical representation for the cell cultures considered in case of the Meso-scale Model, i.e. cell cultures of the same type, under the assumptions that the food is inextinguishable, toxicity is never removed, toxicity diffusion rate is high enough, and living cells constitute a single layer adhering to the flask ground. In order to develop a model with lumped-

parameter concepts, it is assumed that these concepts are uniform abstractions of their virtual or physical counterparts, i.e. complex formations or emergent behaviours are not observed as an outcome of their dynamics. For the both macro-scale models that will be described in the following sections, five important common assumptions have been made about the biological phenomenon under investigation:

(i) Cells are initially randomly[2] uniformly distributed across the flask. Local heterogeneities formed due to statistical aggregations are not significantly amplified by cellular dynamics.

(ii) The energy distribution of cells across the population does not depend on the spatial distribution of cells. This assumption prevents a group of cells with the same energy range not to be accumulated in a specific region of the flask.

(iii) As a consequence of high toxicity diffusion rate in the Meso-scale Model, toxicity is assumed to be uniformly distributed across the flask at all times.

(iv) Macro-scale model is valid only after the first population increase is observed. So, it is assumed that the upper boundary of energy distribution across the population stands at the division energy.

(v) Quiescent cells and Non-quiescent cells have approximately a similar energy distribution.

These assumptions may not always reflect the true nature of cell cultures, nevertheless they are biologically plausible and can even be realized partially in laboratory environments. All the underlying concepts, state variables and approximations needed to describe the macro-scale model are developed under these assumptions and are presented in the following sections.

It is also worth to mention that in accordance with the common macroscopic perception of time, the macro-scale model will be represented in continuous-time (t) as opposed to the Meso-scale Model, which represents the dynamics of agents in discrete time (k).

[2]Unlike 'perfectly uniform', where cells are placed in equidistant intervals, 'randomly uniformly' means that cells are placed to their loci according to a uniform probability function.

3.2.1. State variable candidates

To represent the population dynamics in terms of global variables, one first needs to determine which set of global state variables can provide an operationally closed state space. Since the main modelling task is to predict how the population size changes, the total population $p(t)$ is an obvious candidate as a state variable. With the insight gained from the Meso-scale Model, the total toxicity in the flask and the distribution of cell energies across the population come into focus as important factors affecting the population dynamics. Total toxicity in the flask, $\zeta(t)$, constitutes a reasonable candidate as the second global variable, however it is not trivial what kind of state variables can sufficiently well represent the distribution of cell energies across the population. Meso-scale simulation results demonstrate that the distribution $f(E,t)$ of cell energies across the population[3] exhibits rather variable and complex characteristics during the life time of the cell culture. But it is also observed that the lower and upper boundaries of the energy distribution change in a relatively simple way.

Consequently, as a first approach it has been decided to represent the population-wide energy distribution of the first macro-scale model as a uniform one between two time-variable boundaries, thus obtaining the two additional state variables of the first macro-scale model: $\eta_L(t)$ and $\eta_H(t)$, the lower and upper boundaries of the energy distribution.

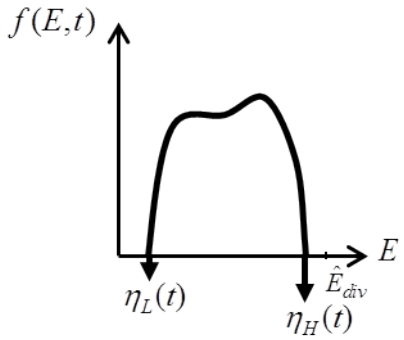

Figure 3.1. Energy distribution of cells across the population.

For the second macro-scale model, the assumption of uniform energy distribution

[3]Number of cells within energy range $[E_1, E_2]$ at any time t can be found using the energy distribution: $\int_{E=E_1}^{E_2} f(E,t)dE$

is relaxed and a supplementary state variable $\eta_S(t)$, a virtual energy value representing the total change in the extended energy distribution (which will be detailed in Section 3.3.3.2, is added to the model in order to improve its prediction performance.

As a last note, it is important to remark that throughout this thesis the population quantity $p(t)$ is given as the percentage occupancy of all cells in the flask instead of real number of cells.

3.2.2. Approximation of the Quiescent Cell Population

The quiescent population is by definition linked to the total number of cells. We have developed two approximations for the quiescent population, the first one being a simpler approximation based on pure reasoning and used in Macro-scale Model I while the other one is partially based on empirical results of the Meso-scale Model simulations and constitutes a better predictor for the time profile of quiescent population, which is a part of Macro-scale Model II. The first quiescent approximation is developed under the perfect uniform assumption of total population and hence describes the quiescent population only as a function of total population rather than as a variable with its own independent dynamics. However, the second quiescent approximation depends also on time as an independent variable the purpose and reasoning of which will be detailed in Section 3.2.2.2. In both approaches, quiescent population $q(t)$ is given in percentage occupancy of quiescent cells in the flask like $p(t)$.

3.2.2.1. Quiescent Model I. In order to derive an approximate formula for the $q(t)$, let us consider a flask that can hold a maximum number of M cells, i.e. to put it in the terminology of the Meso-scale Model, it has M loci. The assumption that all cells are uniformly distributed across the flask at all times implies that also the empty loci are uniformly distributed. Consequently, $b(t)$, the average number of empty loci per cell, can be calculated as:

$$b(t) = \frac{M - p(t)}{p(t)} \tag{3.1}$$

When there are $p(t)$ uniformly distributed cells in the flask, all of them will be in active mode if each cell has at least one empty neighbouring locus; thus

$$b(t) \geq 1 \Rightarrow q(t) = 0 \tag{3.2}$$

On the other hand, when the average number of empty loci per cell turns out to be between 0 and 1, one can interpret this situation in a statistical manner as follows: each living cell has either 0 or 1 empty locus in its neighbourhood, and $b(t)$ is equal to the fraction of cells with 1 empty neighbouring locus (i.e. the ratio of active cells to the total number of cells). Consequently, for $0 \leq b(t) < 1$ the quiescent cell population can be approximated as:

$$0 \leq b(t) < 1 \Rightarrow q(t) = [1 - b(t)]p(t) \tag{3.3}$$

Combining Equations 3.1, 3.2 and 3.3 the quiescent cell population can be approximated as follows (Figure 3.2):

$$q(t) = \begin{cases} 0 & if\ p(t) \leq \frac{M}{2} \\ 2p(t) - M & if\ \frac{M}{2} < p(t) \leq M \end{cases} \tag{3.4}$$

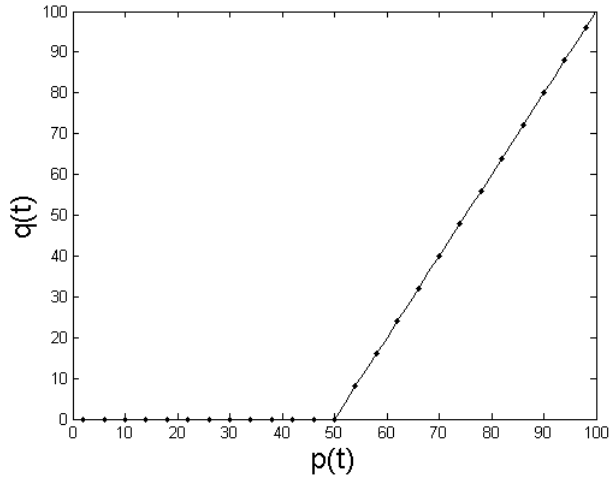

Figure 3.2. Approximate model of $q(t)$ as a function of $p(t)$ in Quiescent Model I.

As a summary, this quiescent population approximation exhibits two important

characteristics:

(i) If there is any quiescent population in the flask, its size increases linearly with the total population size.

(ii) The quiescent cells are observed only if at least half of the flask is occupied regardless of the initial population size, i.e. the intersection point of piecewise lines representing the quiescent population is at 50%.

The Quiescent Model I is developed under perfect uniformity assumption. Thus, its performance of estimating the quiescent occupancy, $q(t)$ on basis of $p(t)$ at any given time, should be correlated to some extent with the heterogeneity measure $\psi(t)$ given in Section 2.5.1.4.

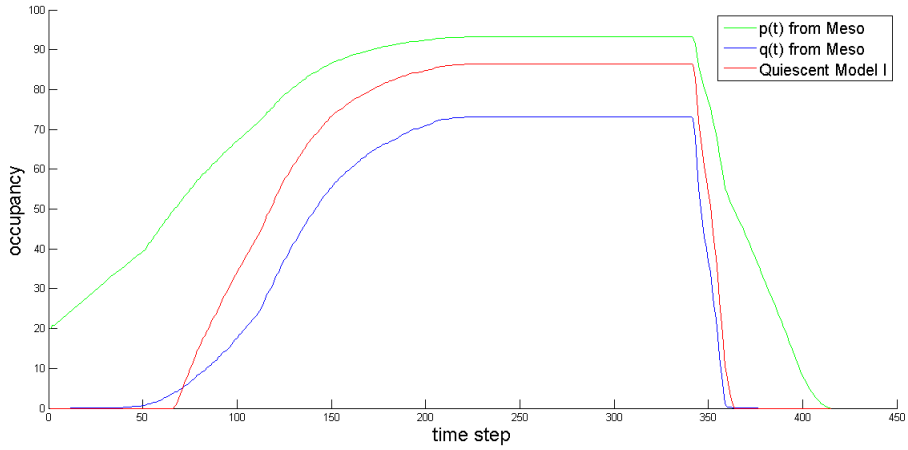

Figure 3.3. Quiescent population from the Meso-scale Model (blue) and its estimate by Quiescent Model I (red) for an initial population occupancy (green) of 20%.

As can be observed from Figure 3.3, Quiescent Model I slightly underestimates $q(t)$ for $p(t) < 50\%$ and overestimates for higher occupancies. As expected, the absolute estimation error of $q(t)$ exhibits qualitative similarity with the heterogeneity measure except for a short duration where Quiescent Model I transcends from underestimating to overestimating the quiescent occupancy (Figure 3.4).

3.2.2.2. Quiescent Model II. Before developing a better approximation for $q(t)$, it should be verified how appropriate the characteristics of the Quiescent Model I re-

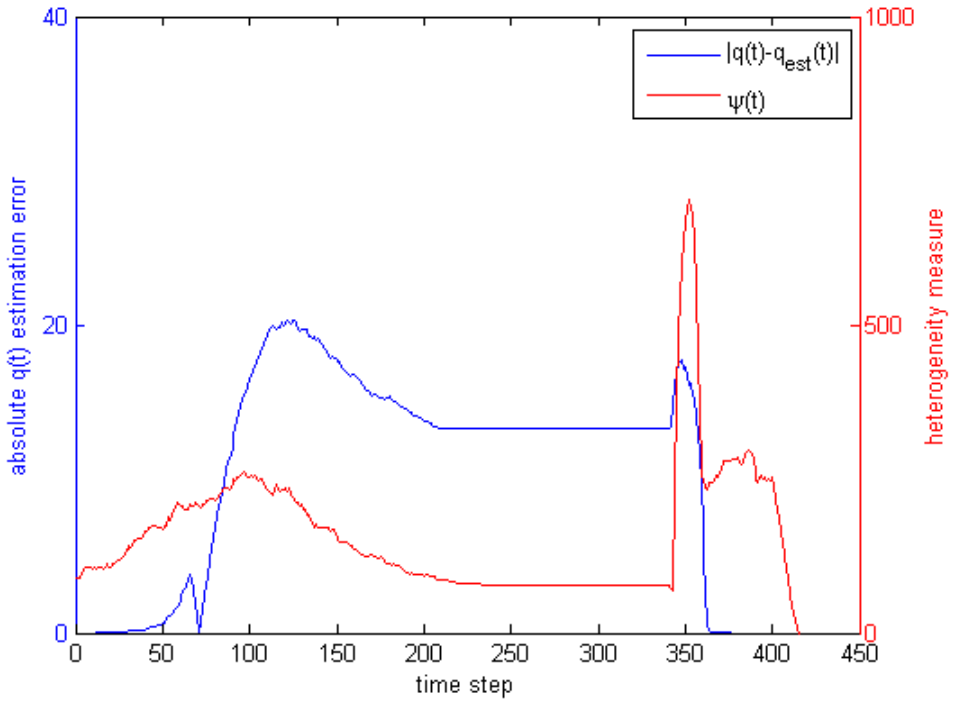

Figure 3.4. Absolute quiescent occupancy estimation error and the heterogeneity measure $\psi(t)$ exhibit qualitative similarity.

flect the true nature of quiescent population dependence on the total population. For this purpose, a further investigation of $q(t)$ is needed from the meso-scale simulations.

One can immediately observe that the linear relationship between $p(t)$ and $q(t)$ is quiet acceptable except at the initial duration when the first quiescent cells appear. However, this duration is rather short as compared to the rest of population's life span and can therefore be approximated with the same linear relationship. It can also be observed that $q(t)$ does not exhibit qualitatively different characteristics after $p(t)$ starts to decrease (represented with black dots in Figure 3.5). Furthermore, the sparsity of the data during this session indicates that cell deaths constitute only a small portion of total life span, allowing us to neglect small deviations from the main trend of the relationship . Hence, it is reasonable to approximate this part of the data also with the same linear model.

On the other hand, the onset point which indicates the appearance of the first quiescent cells is obviously not fixed at half of the flask like in the first approximation.

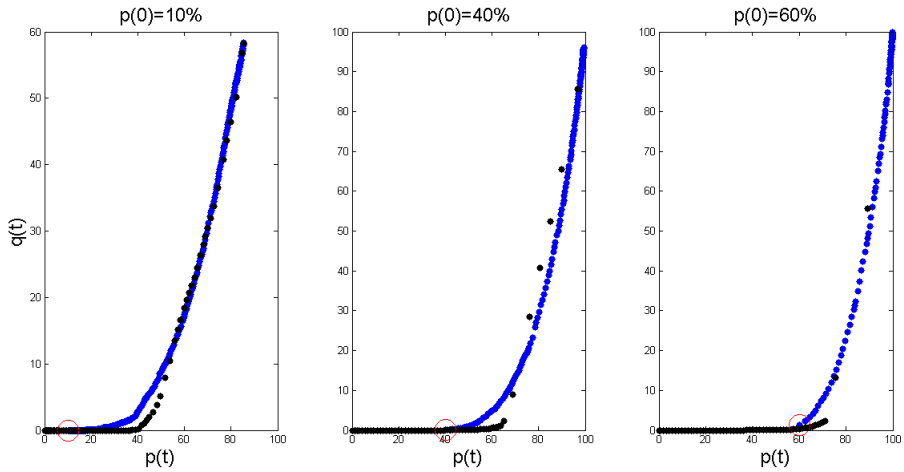

Figure 3.5. $q(t)$ as a function of $p(t)$ obtained from the meso-scale simulations with initially uniform energy distribution for three different initial populations. Red circle represents the initial population size, while the black dots indicate values after total population starts to decrease.

We can easily deduct its reasoning from the knowledge of the Meso-scale Model: The onset points should depend on the initial population size due to division process. When cells undergo division, the placing of daughter cells to the neighbourhood of mother cell creates local heterogeneities and after a sufficient time these aggregations cause some cells to become quiescent. Hence, the assumption of uniformity at all times in the first approximation is disturbed due to new born cells. The time needed for the appearance of the first quiescent cells varies with respect to $p(0)$ since the only factor determining this duration is obviously the initial crowdedness of cells. Therefore, this duration decreases monotonically with the increasing $p(0)$. On basis of these observations about how the onset points depend on $p(0)$, we decided to develop a more accurate representation of $q(t)$ by fitting simple piecewise lines to 50 different initial populations (Figure 3.6) and visualised them in a 3D space, the third dimension of which is selected as time (Figure 3.7).

The onset points shown with black circles in Figure 3.7 are defined as $[p_q \ t_q]$, where t_q represents the time when cells start to enter quiescence and p_q represents the minimum population occupancy where first quiescent cells appear. A closer inspection of Figure 3.7 reveals that as the initial population occupancy increases, the time t_q

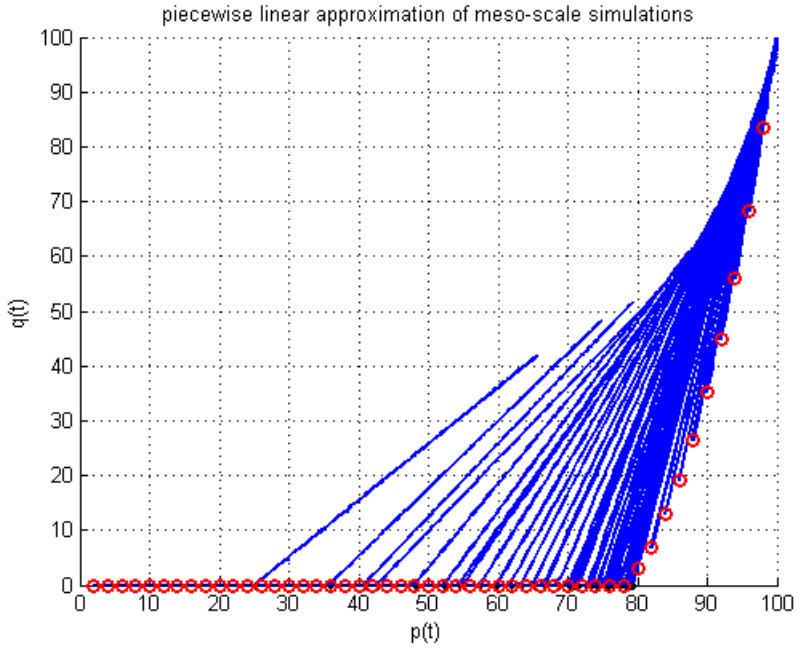

Figure 3.6. Piecewise linear approximations of meso-scale simulations starting from 50 different initial populations until cells start to die. Red circles indicate initial population size $p(0)$.

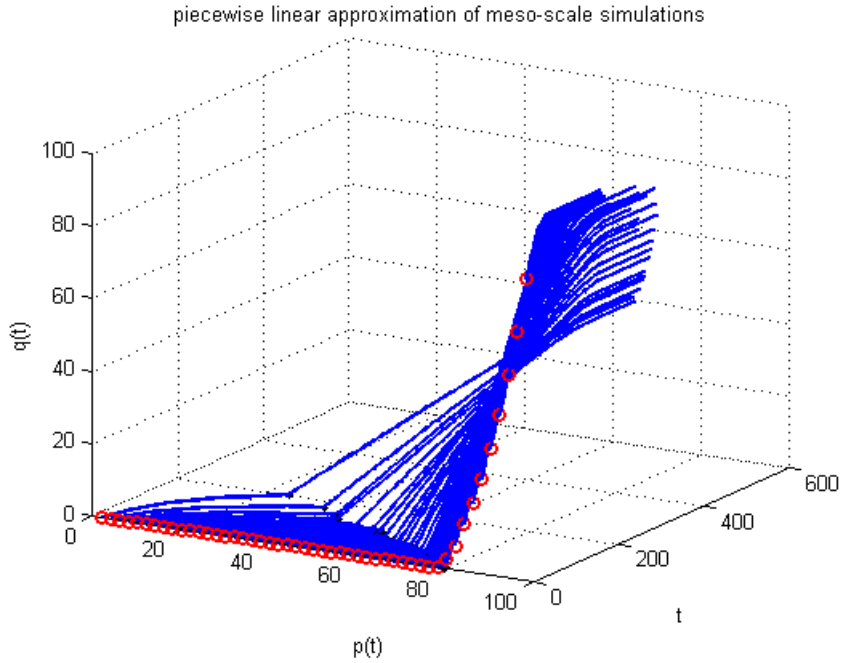

Figure 3.7. 3D visualisation of Figure 3.6 with time as third dimension. Black circles indicate the onset time t_q.

required for $p(t)$ to reach the onset point decreases as expected. So, our reasoning is supported by the shifting of onset points with respect to $p(0)$ and hence t_q (Figure 3.8).

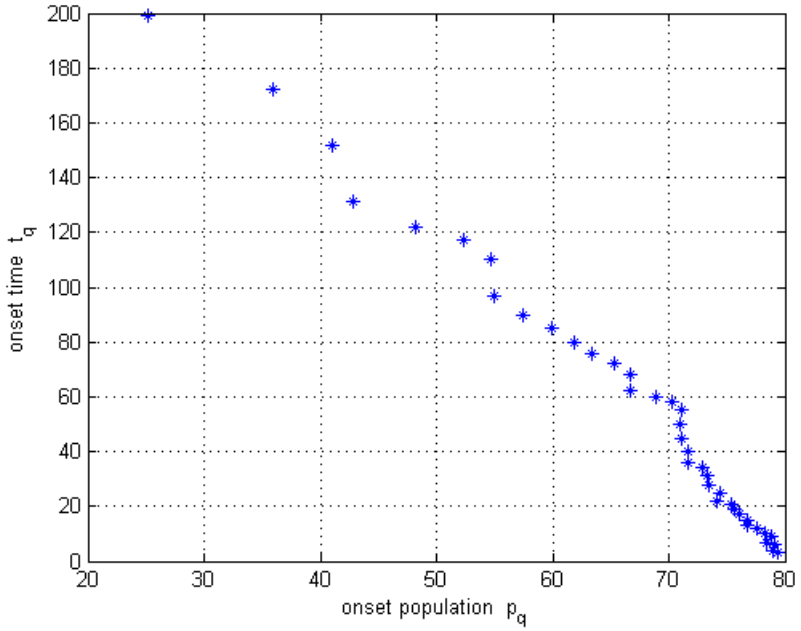

Figure 3.8. Blue stars indicate the onset points, i.e. at which population size and time step quiescent cells are observed first time.

Under further investigation, we have decided to fit a line to describe the relationship between population and time at onset points (Figure 3.9). This line represents when the population-time curve will leave the plane of $q = 0$. However this data is specific to the initial energy distribution of cells in the Meso-scale Model. Although Figures 3.6 to 3.9 are obtained for the uniform energy distribution, simulations from other distribution shapes reveal that the characteristics of this line do not change significantly. Consequently, the average line obtained from four set of different initial energy distributions[4] is as follows:

$$p_q(t) = -0.265t_q + 82.7 \tag{3.5}$$

[4]The energy distributions used to estimate the parameters of Equation 3.5 are selected to cover a wide range of biologicaly plausible distributions, i.e. Gaussian, linearly decaying, exponentially decaying and uniform one which will be shown in Section 3.7 and also used in order to measure the robustness of macro-scale models.

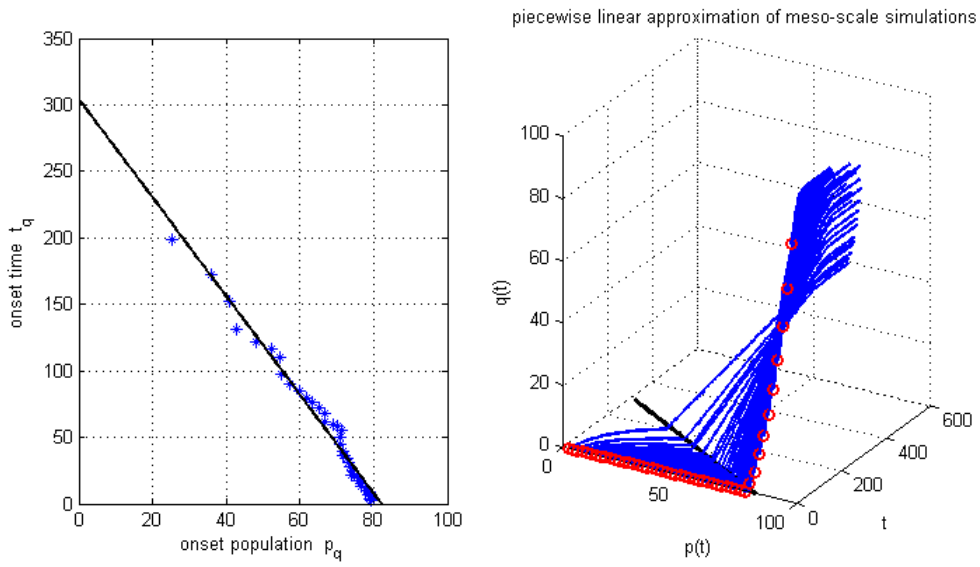

Figure 3.9. Fitted line representing the relationship between onset population and onset time.

Once $q(t)$ has a positive value, it will increase approximately linearly as stated before and hence will have the following form:

$$q(t) = \kappa p(t) + c$$

The parameters of this equation can be estimated considering two distinct conditions:

- Condition I : When flask is completely occupied, all cells are in quiescence phase, i.e. :

$$100 = 100\kappa + c$$

- Condition II : Quiescent population starts to have a presence when $p(t)$ reaches the onset point, i.e. when the line of $q(t)$ intersects $p(t)$-axis at onset point.

$$0 = \kappa p_q(t) + c$$

Using these two conditions, quiescent population is expressed as:

$$q(t) = \frac{100}{100 - p_q(t)} \left(p(t) - p_q(t)\right) \tag{3.6}$$

It is important to state that this function is meaningful only for $p(t)$ which do not initially contain any quiescent population. A sufficiently high initial population size will already contain quiescent cells at the beginning of the experiment and hence will have no onset point. The maximum $p(0)$ without any initially quiescent cell presence can be found from Equation 3.5 by taking t as zero:

$$p_q^* = p_q(0) = 82.7$$

$q(t)$ for initial populations higher than p_q^*, tends to exhibit a similar characteristics with $q(t)$ for p_q^*. Figure 3.10 also shows that it is reasonable to approximate $q(t)$ for those $p(0)$ values with the line starting from p_q^*.

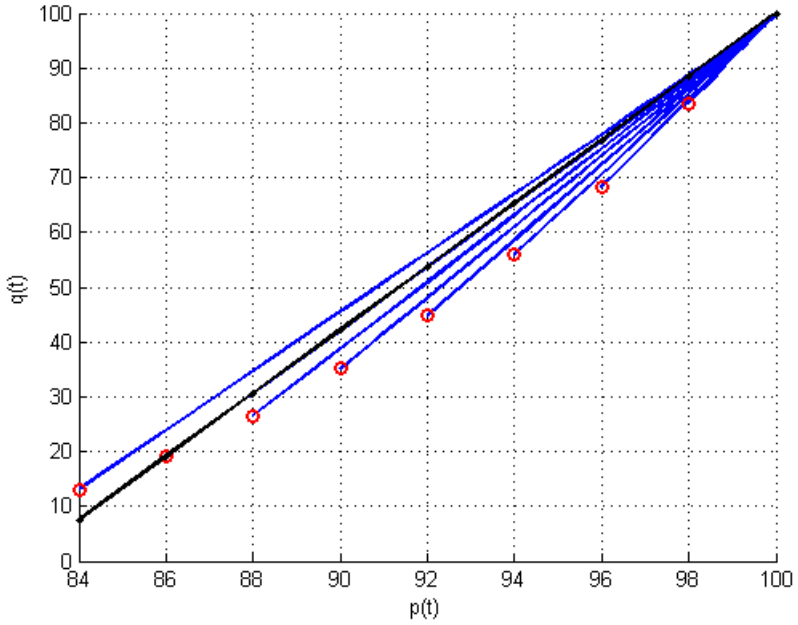

Figure 3.10. Black line is the quiescent occupancy approximation for initial occupancies higher than p_q^*.

Putting p_q^* into Equation 3.6 results in:

$$q(t) = 5.78p(t) - 478 \quad if\ p(0) \geq 82.7 \tag{3.7}$$

Consequently, populations starting from $p(0)$ smaller than 82.7% will contain zero quiescent population until the onset point and then $q(t)$ will linearly increase according to Equation 3.6. For $p(0)$ higher than 82.7% concentration, $q(t)$ will increase linearly according to Equation 3.7. Eventually, when populations start to die, $q(t)$ will decrease with the same rate until their onset point and then become zero. All of these behaviours can be summarised in the following single equation:

$$q(t) = \begin{cases} \left[\frac{100}{100-p_q(t)}\left(p(t) - p_q(t)\right)\right]_+ & where\ p_q(t) = -0.265t_q + 82.7 \quad if\ p(0) < 82.7 \\ [5.78p(t) - 478]_+ & if\ p(0) \geq 82.7 \end{cases} \tag{3.8}$$

where $[x]_+$ denotes $max(x, 0)$

3.2.3. Average Cell and Its Behaviour

In the macro-scale model the population is envisaged as consisting of identical cells all of which exhibit the average behaviour of individual cells in the Meso-scale Model. Throughout this chapter the term 'average cell' will be used in that sense.

Each cell releases toxicity at a constant rate ω. Those cells that are in the active mode (active cells) take in energy at a constant rate α and spend remediation energy at a rate $\dot{\varepsilon}_{rem}(t)$, where $\dot{\varepsilon}_{rem}(t)$ depends on the toxicity they are exposed to. On the other hand, all members of the quiescent population take in energy at a rate equal to $\dot{\varepsilon}_{rem}(t)$ as long as this value does not exceed α, such that their energies remain the same. If, however, $\dot{\varepsilon}_{rem}(t)$ is larger than α, the energy intake rate is kept at α, thus leading to energy decrease. The dependence of $\dot{\varepsilon}_{rem}(t)$ on toxicity will be explained in 3.2.4.

3.2.4. Average Toxicity Exposure and Remediation Rate

In order to calculate remediation rate in macro-scale model, a toxicity-related concept needs to be introduced: $\xi(t)$, the amount of toxicity an average cell is exposed to at time instance t. Let us call this variable 'average toxicity exposure'. Referring to the Meso-scale Model one can say that $\xi(t)$ is the toxicity present at the locus which the cell occupies. It should be noted that $\xi(t)$ actually depends on the toxicity diffusion rate and its exact calculation would require very complicated calculations. Nevertheless, its upper and lower bounds can be rather easily estimated considering the two extremes of the diffusion rate. If the diffusion rate is very high such that released toxicity instantly diffuses to the whole flask, each cell will be exposed to the same toxicity, which constitutes approximately the lower bound of the average toxicity exposure (although the average toxicity exposure can attain lower values for a short while):

$$\xi_{min}(t) \approx \frac{\zeta(t)}{M} \tag{3.9}$$

On the other hand, the upper bound of the average toxicity exposure at time t can be approximately calculated assuming that the toxicity diffusion rate is extremely low, and that toxicity is present only at occupied loci (there are $Mp(t)$ such loci):

$$\xi_{max}(t) \approx \frac{\zeta(t)}{Mp(t)} \tag{3.10}$$

The results (Figure 3.11) obtained from simulations of the Meso-scale Model with realistic parameter values reveal that for diffusion rates typical for small particles in the cytoplasm [25], $\xi(t)$ remains rather close to $\xi_{min}(t)$. Thus, average toxicity exposure has been taken equal to the lower bound:

$$\xi(t) = \xi_{min}(t) \approx \frac{\zeta(t)}{M} \tag{3.11}$$

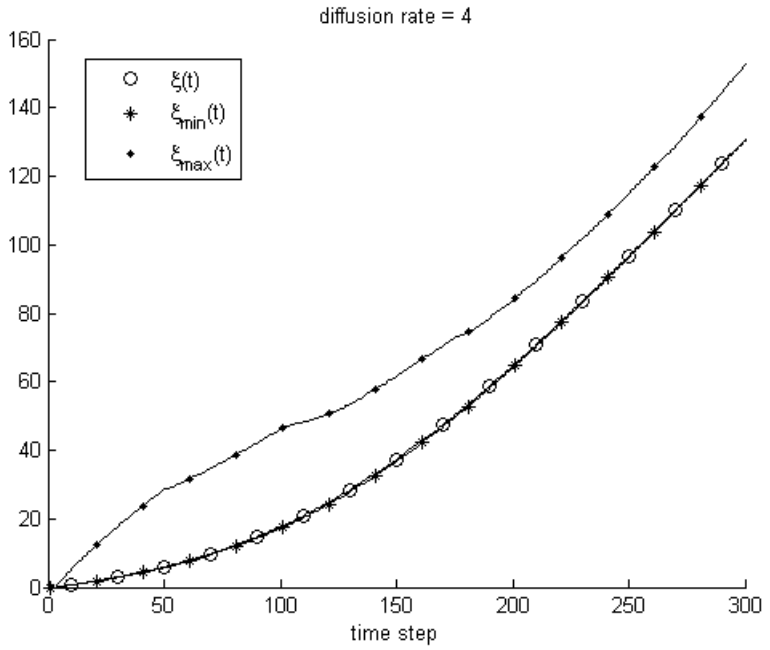

Figure 3.11. Meso-scale simulation results showing average toxicity $\xi(t)$, its approximate lower bound $\xi_{min}(t)$ and upper bound $\xi_{max}(t)$.

Adopting a continuous-time version of the logic used in the Meso-scale Model, one can say that an average cell that is exposed to the average toxicity $\xi(t)$ for a duration dt needs an energy $d\varepsilon$ for remediating the resulting damage:

$$d\varepsilon_{rem} = \hat{\gamma}\xi dt \tag{3.12}$$

where $\hat{\gamma}$ corresponds to the same concept as γ in the Meso-scale Model, i.e. the remediation constant, but a different notation has been introduced for formal clarity. Using the approximation from Equation 3.11, the remediation rate $\dot{\varepsilon}_{rem}$ can be expressed in terms of total toxicity ζ as follows:

$$\dot{\varepsilon}_{rem}(t) \approx \frac{\hat{\gamma}\zeta(t)}{M} \tag{3.13}$$

Here, it should be noted that γ has the dimension $\left[\frac{1}{time}\right]$.

3.3. Dynamics of State Variables

As stated in Section 3.2.1, total population $p(t)$, total toxicity $\zeta(t)$ and boundaries of energy distribution across cell population (η_L and η_H) are state variables of the Macro-scale Model I. For the Macro-scale Model II, an additional state variable, total energy shift η_S is used in order to obtain a better prediction of the Meso-scale Model. The dynamics of these state variables are explained in the following sections.

3.3.1. Total Population Dynamics

With simple reasoning, the change in population size can be expressed as:

$$\dot{p}(t) = (birth\ rate) \times (number\ of\ active\ cells) - (death\ rate) \times p(t) \tag{3.14}$$

In order to calculate the birth and death rates one needs to resort to the distribution of cell energies across the population.

Let $d\triangle_{div}$ denote the fraction of cells, the energies of which will reach $\hat{E}_{div}$ within next dt and thus will be able to divide if there is enough space. Here, $\hat{E}_{div}$ corresponds to the same concept as E_{div} in the Meso-scale Model, i.e. the division energy. $d\triangle_{div}$ (Figure 3.12) can be calculated as follows:

$$d\triangle_{div}(t) = \left[\frac{\int_{\hat{E}_{div}-(\alpha-\dot{\varepsilon}_{rem}(t))dt}^{\eta_H(t)} f(E,t)dE}{\int_{\eta_L(t)}^{\eta_H(t)} f(E,t)dE} \right]_+ \quad if\ \alpha > \dot{\varepsilon}_{rem}(t) \tag{3.15}$$

Equation 3.15 has a nonzero value only for the case where $\eta_H(t) > (\hat{E}_{div} - (\alpha - \dot{\varepsilon}_{rem}(t))dt)$. Furthermore keeping in mind that the division energy $\hat{E}_{div}$ is the maximum energy a cell can have, the condition $(\hat{E}_{div} - (\alpha - \dot{\varepsilon}_{rem}(t))dt) < \eta_H(t) \leq \hat{E}_{div}$ is tantamount to $\eta_H(t) = \hat{E}_{div}$, i.e. cells can divide only if the upper boundary of the energy distribution $\eta_H(t)$ reaches the estimated division energy $\hat{E}_{div}$. Hence, Equation 3.15 becomes:

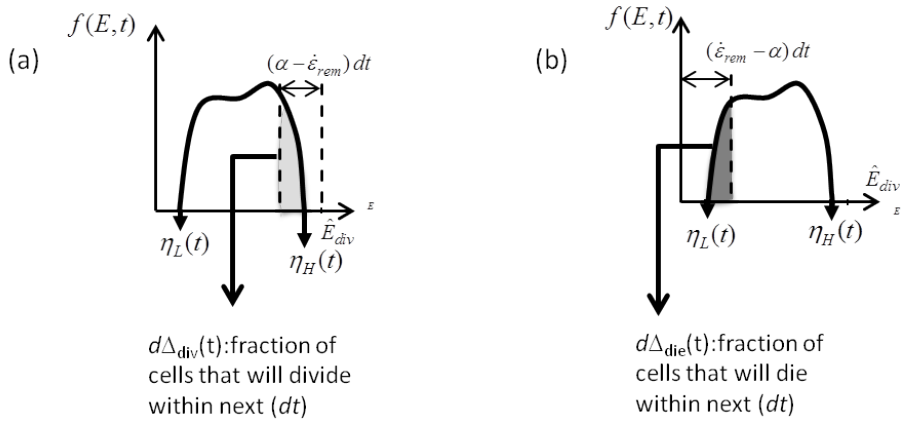

Figure 3.12. a) Fraction of cells ($d\triangle_{div}(t)$) that will undergo mitotic division within the next dt for $\alpha > \dot{\varepsilon}_{rem}(t)$, b) Fraction of cells ($d\triangle_{die}(t)$) that will die within the next dt for $\alpha < \dot{\varepsilon}_{rem}(t)$).

$$d\triangle_{div}(t) = \left[\frac{f(\hat{E}_{div}, t)(\eta_H(t) - \hat{E}_{div} + (\alpha - \dot{\varepsilon}_{rem}(t))dt}{\int_{\eta_L(t)}^{\eta_H(t)} f(E,t)dE} \right]_+ \tag{3.16}$$

which can be reformulated as birth rate (ρ_b):

$$\rho_b(t) = \frac{d\triangle_{div}(t)}{dt} = \dot{\triangle}_{div}(t) = \begin{cases} (\alpha - \dot{\varepsilon}_{rem}(t)) \frac{f(\hat{E}_{div},t)}{\int_{\eta_L(t)}^{\hat{E}_{div}} f(E,t)dE} & if\ \eta_H(t) = \hat{E}_{div} \\ 0 & otherwise \end{cases} \tag{3.17}$$

Under the further assumption that the energies of active cells are also uniformly distributed, the number of cells that will divide within next dt can be calculated as $[p(t) - q(t)]\rho_b$, where $[p(t) - q(t)]$ represents the number of active cells. It is worth keeping in mind that the birth rate is only meaningful if the energy intake ($\alpha\ dt$) is larger than the energy loss ($\dot{\varepsilon}_{rem}(t)dt$).

Similarly, let $d\triangle_{die}$ denote the fraction of cells, the energies of which will fall below zero within next dt because the energy intake ($\alpha\ dt$) is less than the energy loss

($\dot{\varepsilon}_{rem}(t)dt$). So $\triangle_{die}$ (Figure 3.12b) can be calculated as:

$$d\triangle_{die}(t) = \left[\frac{\int_{\eta_L(t)}^{(\dot{\varepsilon}_{rem}(t)-\alpha)dt} f(E,t)dE}{\int_{\eta_L(t)}^{\eta_H(t)} f(E,t)dE} \right]_+ \quad if\ \alpha < \dot{\varepsilon}_{rem}(t) \tag{3.18}$$

Equation 3.18 has a nonzero value only for the case where $\eta_L(t) < (\dot{\varepsilon}_{rem}(t) - \alpha)dt$. Since the minimum energy of a cell $\eta_L(t)$ cannot be negative, the condition $0 \leq \eta_L(t) < (\dot{\varepsilon}_{rem}(t) - \alpha)dt$ is equivalent to $\eta_L(t) = 0$, i.e. cell can die only if the lower boundary of the energy distribution $\eta_L(t)$ falls to zero. Hence, Equation 3.18 can be reformulated to find death rate (ρ_d):

$$\rho_d(t) = \frac{d\triangle_{die}(t)}{dt} = \dot{\triangle}_{die}(t) = \begin{cases} (\dot{\varepsilon}_{rem}(t) - \alpha)\frac{f(0,t)}{\int_0^{\eta_H(t)} f(E,t)dE} & if\ \eta_L(t) = 0 \\ 0 & otherwise \end{cases} \tag{3.19}$$

Due to the fact that the conditions in Equation 3.17 and 3.19 are opposite to each other, the birth and death rates cannot be non-zero at the same time, i.e. cells births and deaths cannot occur simultaneously. Hence, Equation 3.14 can be rewritten as:

$$\dot{p}(t) = \begin{cases} [p(t) - q(t)]\rho_b(t) & \text{with } \rho_b(t) = \left(\alpha - \frac{\hat{\gamma}\zeta(t)}{M}\right) \frac{f(\hat{E}_{div},t)}{\int_{\eta_L(t)}^{\hat{E}_{div}} f(E,t)dE} \\ & \text{if } \eta_H(t) = \hat{E}_{div}\ and\ \alpha > \frac{\hat{\gamma}\zeta(t)}{M} \\ -p(t)\rho_d(t) & \text{with } \rho_d(t) = \left(\frac{\hat{\gamma}\zeta(t)}{M} - \alpha\right) \frac{f(0,t)}{\int_0^{\eta_H(t)} f(E,t)dE} \\ & \text{if } \eta_L(t) = 0\ and\ \alpha < \frac{\hat{\gamma}\zeta(t)}{M} \\ 0 & \text{otherwise} \end{cases} \tag{3.20}$$

Note that $p(t)$ is naturally defined as a non-negative variable. Hence $\dot{p}(t)$ is taken as zero if $p(t) < 0$.

3.3.2. Total Toxicity

A basic assumption of the macro-scale model is that the toxicity diffuses fast enough so that the toxicity distribution across the flask can always be approximated by a uniform one.

In the macro-scale simulations given in this thesis it has been assumed that the flask is initially clean (i.e. $\zeta(0) = 0$) and that during an experiment toxicity is never removed from the flask. Furthermore active and quiescent cells are assumed to release toxicity at the same constant rate ω, such that the rate of change of total toxicity $\zeta(t)$ can be expressed as proportional to the total population:

$$\dot{\zeta}(t) = \omega p(t) \tag{3.21}$$

The lumped-parameter dynamics of the total toxicity can also be derived analytically step by step from the knowledge of local dynamics in Meso-scale Model as shown in Appendix A.

3.3.3. Energy Distribution and Its Dynamics

The energy of a cell is an important factor that affects both birth and death rates. Therefore, in a macro-scale model the distribution of cell energies across the cell population needs to be accounted for. Furthermore, since this distribution will change while the population keeps evolving, one also needs to model the dynamics of the energy distribution. The complexity of the model has to be kept as low as possible in the course of accomplishing this task.

In this thesis, two different models (Energy Model I and II) have been developed for the dynamics of $f(E,t)$, the energy distribution across cell population. How $f(E,t)$ is to be used to calculate instantaneous birth and death rates has been explained in Section 3.3.1.

As can be seen from Equation 3.20, the births are calculated as the product of the birth rate ($\rho_b(t)$) with the % occupancy of the non-quiescent (active) cells $[p(t) - q(t)]$ whereas deaths are calculated as the product of the death rate ($\rho_d(t)$) with the % occupancy $p(t)$ (i.e. all cells). Consequently, $\rho_b(t)$ needs to be calculated (Equation 3.17) using the energy distribution among active cells, while $\rho_d(t)$ has to be calculated (Equation 3.19) using the energy distribution among all cells. In this thesis, however, it has been assumed that the energy distribution among active cell population and total cell population have the same shape. Strictly speaking, this assumption is not correct because in dense populations new-born cells have higher probability of entering quiescence mode than 'elder' cells, which biases the shape of the energy distribution among quiescence cells more towards low energy levels. Why this approximation error is tolerable will be explained in Section 3.6.3.2.

Before going into the details of Energy Model I and II, let us explain what is common to both of them. Both models of energy distribution have to be compatible with the following observations made on basis of the Meso-scale Model:

(i) Since the energy of any cell at any time must be between 0 and $\hat{E}_{div}$, the energy distribution $f(E,t)$ must be confined to the range $[0,\ \hat{E}_{div}]$.

(ii) Even if the initial range of energy distribution is between 0 and $\hat{E}_{div}$, it quickly converges to the range $[\hat{E}_{div}/2,\ \hat{E}_{div}]$ because cells steadily take in the energy and initially toxicity related losses are negligible, while the division process produces cells with $\hat{E}_{div}/2$, such that soon there remain no cells with energy below $\hat{E}_{div}/2$. An exception to this behaviour occurs if a meso-scale simulation starts with very high populations such that some cells in initially overcrowded regions with energies below $\hat{E}_{div}/2$ enter quiescence mode right from the beginning. Excluding this exceptional case, it is reasonable to start the energy distribution from range $[\hat{E}_{div}/2,\ \hat{E}_{div}]$, i.e. $\eta_L(0) = \hat{E}_{div}/2$ and $\eta_H(0) = \hat{E}_{div}$.

(iii) Two simple and relevant variables characterising the energy distribution $f(E,t)$ will be taken as the upper and lower boundaries of the distribution, i.e. $\eta_L(t)$ and $\eta_H(t)$, respectively. In order to guarantee the continuity of these two boundaries, the initial energy distribution is required to be nonzero within $[\eta_L(0), \eta_H(0)]$.

(iv) Provided that condition (iii) is satisfied, and the toxicity in the flask is low enough to guarantee that remediation energies can be covered by the constant energy intake , the energies of all cells in the Meso-scale Model keep on increasing up to the division energy E_{div}. This phenomenon can be represented in the macro-scale model with an upper boundary increasing up to $\hat{E}_{div}$. However, since $\eta_H(0) = \hat{E}_{div}$, $\eta_H(t)$ will stay at $\hat{E}_{div}$ until sufficient toxicity accumulates. Likewise, $\eta_L(t)$ will start and stay at $\hat{E}_{div}/2$ until the same condition is satisfied.

(v) When the total toxicity becomes high enough such that the constant energy intake α cannot cover anymore the remediation expenditures, all cells in the Meso-scale Model start losing energy. This corresponds to the decrease of both lower and upper boundaries in the macro-scale model.

Combining the observations (i)-(v), the approximate dynamics of $\eta_L(t)$ and $\eta_H(t)$ can be represented mathematically as follows:

$$\dot{\eta}_H = \begin{cases} 0 & if \;\; \dot{\varepsilon}_{rem}(t) \leq \alpha \;\; or \;\; \eta_H(t) \leq 0 \\ \alpha - \dot{\varepsilon}_{rem}(t) & if \;\; \dot{\varepsilon}_{rem}(t) > \alpha \end{cases} \qquad \eta_H(0) = \hat{E}_{div} \tag{3.22}$$

$$\dot{\eta}_L = \begin{cases} 0 & if \;\; \dot{\varepsilon}_{rem}(t) \leq \alpha \;\; or \;\; \eta_L(t) \leq 0 \\ \alpha - \dot{\varepsilon}_{rem}(t) & if \;\; \dot{\varepsilon}_{rem}(t) > \alpha \end{cases} \qquad \eta_L(0) = \frac{\hat{E}_{div}}{2} \tag{3.23}$$

It should be denoted that $\eta_H(t)$ and $\eta_L(t)$ are obviously defined as non-negative variables.

3.3.3.1. Energy Model I. As a first approach, the shape of energy distribution, which is bounded between $\eta_L(t)$ and $\eta_H(t)$, is assumed to be uniform at all times.

Combining the uniformity assumption of the energy distribution along with observation (ii) given above, the denominator of birth rate (Equation 3.17) for $\eta_H(t) = \hat{E}_{div}$

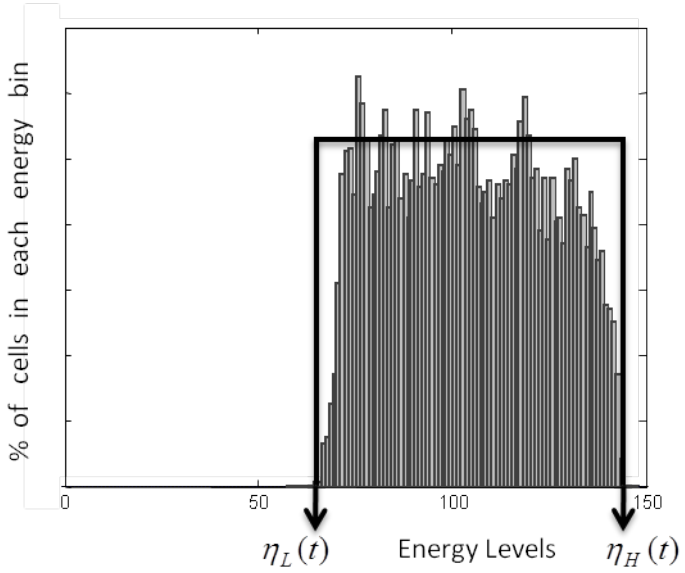

Figure 3.13. Energy histogram of the cells and the uniform distribution approximation.

can be simplified as follows:

$$\int_{\eta_L(t)}^{\hat{E}_{div}} f(E,t)dE = f(\hat{E}_{div},t)(\hat{E}_{div} - \eta_L(t)) = f(\hat{E}_{div},t)\frac{\hat{E}_{div}}{2}$$

Hence, Equation 3.17 becomes:

$$\rho_b(t) = \begin{cases} \frac{2(\alpha - \dot{\varepsilon}_{rem}(t)}{\hat{E}_{div}} & if\ \eta_H(t) = \hat{E}_{div} \\ 0 & otherwise \end{cases} \tag{3.24}$$

With the same line of thought, the death rate (Equation 3.19) becomes:

$$\rho_d(t) = \begin{cases} \frac{2(\dot{\varepsilon}_{rem}(t) - \alpha}{\eta_H(t)} & if\ \eta_L(t) = 0 \\ 0 & otherwise \end{cases} \tag{3.25}$$

3.3.3.2. Energy Model II. Energy Model I discards the effect of division process on the shape of the energy distribution and assumes that this distribution remains uniform

at all times. For a more realistic representation let us consider the two factors that modify the energy distribution:

(i) The net energy gain: If $[\alpha - \dot{\varepsilon}_{rem}(t)]$ is positive, cell energies keep increasing at that rate (as long as they remain below $\hat{E}_{div}$), which means that the energy distribution 'shifts' towards right along the energy axis. If toxicity and thus the corresponding remediation energy $\dot{\varepsilon}_{rem}(t)$ are high enough, the net energy gain becomes negative and the energy distribution shifts towards left (Figure 3.14).

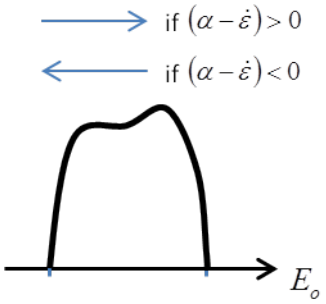

Figure 3.14. Shift of the energy distribution depending on the net energy gain.

(ii) If the net energy gain is positive, the shift of energy distribution towards right causes some portion of the cells to pass the $\hat{E}_{div}$ threshold. But actually these cells divide into two daughter cells with energy $\hat{E}_{div}/2$. From the perspective of the energy distribution this can be described as "the portion of the distribution shifting beyond $\hat{E}_{div}$ folding back to $\hat{E}_{div}/2$, while doubling the magnitude of the distribution at that value (Figure 3.15). As a result of this, we observe a magnified and qualitatively preserved version of the distribution after each cell at that distribution undergoes division.

The combined operation of these two factors can be represented in a circular energy-space. Shifting with a constant speed in such a space corresponds to a periodic behaviour. However, the shifting speed of the energy distribution changes with respect to remediation energy, whereas the magnitude of the distribution is doubled after a cycle in that space is completed. Therefore, it would be more appropriate to describe the dynamics of the energy distribution in a helix-like space, where shifting of the energy distribution would eventually result in the doubling of its magnitude. In order

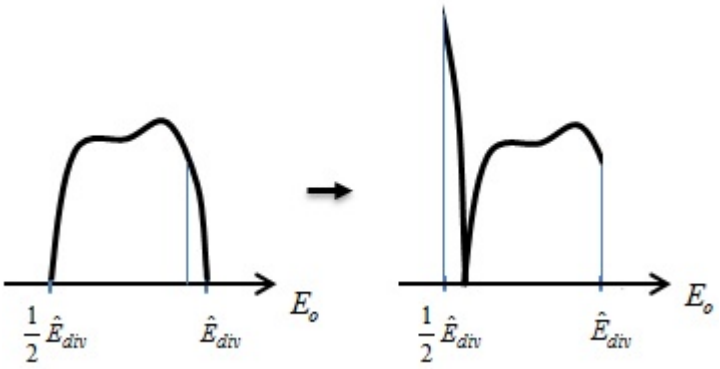

Figure 3.15. The dividing part of the energy distribution is doubled and 'folded' after each division.

eliminate the representational difficulty of both of these factors in such a space, we have decided to make use of an 'extended energy distribution' in a linear space which 'mimics' the folding and doubling operation of the energy distribution by cascading the future projections of the initial energy distribution (Figure 3.16).

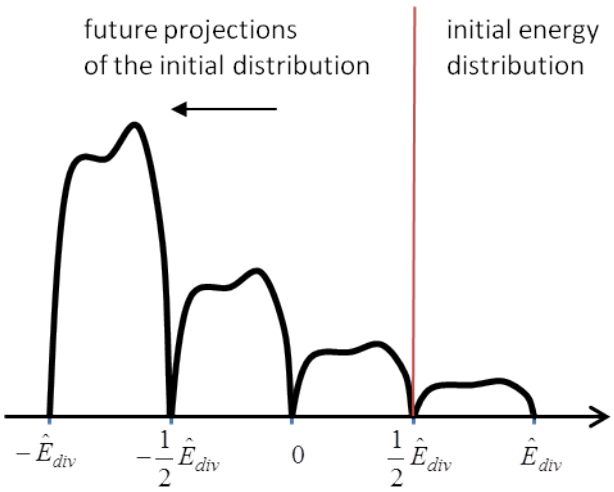

Figure 3.16. Initial extended energy distribution $g(E, t = 0)$.

The extended energy distribution at $t = 0$, $g(E, t = 0)$, can be constructed as the convolution of the initial energy distribution which is confined to $[\frac{\hat{E}_{div}}{2}\hat{E}_{div}]$ with a special Dirac train (D_{doub}) which accounts for the repetitive magnitude doubling. It is

defined at $t = 0$ as:

$$\begin{aligned} g(E, t=0) &= f(E, t=0) * D_{doub}(E, t=0) \\ &= f(E, t=0) * \sum_{k=0}^{\infty} 2^k \delta\left(E + k\frac{\hat{E}_{div}}{2}\right) \end{aligned} \tag{3.26}$$

where $\delta(x)$ is the Dirac function defined as $\delta(x) = \begin{cases} \infty & if\ x = 0 \\ 0 & if\ x \neq 0 \end{cases}$ and $\int_{-\infty}^{\infty} \delta(x)dx = 1$.

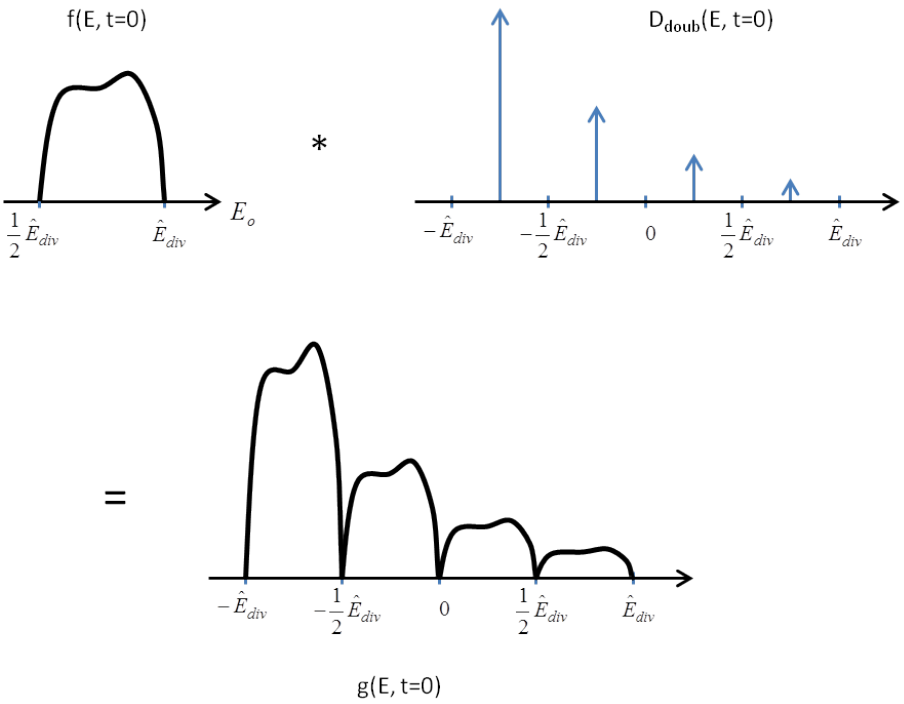

Figure 3.17. Graphical representation of how the extended energy distribution is constructed.

Shifting of the energy distribution due to net energy gain is introduced with the help of an additional state variable $\eta_S(t)$. It is placed inside the equation of extended energy distribution as a shift operator for impulse train function:

$$\begin{aligned} g(E, \eta_S(t)) &= f(E, 0) * D_{doub}(E, \eta_S(t)) \\ &= f(E, 0) * \sum_{k=0}^{\infty} 2^k \delta\left(E - \eta_S(t) + k\frac{\hat{E}_{div}}{2}\right) \end{aligned} \tag{3.27}$$

where

$$\dot{\eta}_S(t) = \alpha - \dot{\varepsilon}_{rem}(t) \quad ; \eta_S(0) = 0 \tag{3.28}$$

$\eta_S(t)$ can be thought as the energy of a 'virtual' cell which does not undergo division at $\hat{E}_{div}$ and its energy can pass beyond $\hat{E}_{div}$ if $\alpha > \dot{\varepsilon}_{rem}(t)$, and also does not die when its energy drops to zero, resulting its energy to become negative if $\alpha < \dot{\varepsilon}_{rem}(t)$. Therefore the dynamics of $\eta_S(t)$ is governed exactly by the same equation as two energy-related state variables, $\eta_L(t)$ and $\eta_H(t)$, change. But since it is a virtual concept and does not have a physical correspondence, unlike $\eta_L(t)$ and $\eta_H(t)$, it is not confined to 0 and $\hat{E}_{div}$. However, when calculating the division and death rates, the extended energy distribution is evaluated only for the region between $\eta_L(t)$ and $\eta_H(t)$.

With the mathematical representation of the energy distribution dynamics, the birth and death rates become:

$$\rho_b(t) = \begin{cases} (\alpha - \dot{\varepsilon}_{rem}(t)) \frac{g(\hat{E}_{div}, \eta_S(t))}{\int_{\frac{\hat{E}_{div}}{2}}^{\hat{E}_{div}} g(E, \eta_S(t)) dE} & if\ \eta_H(t) = \hat{E}_{div} \\ 0 & otherwise \end{cases} \tag{3.29}$$

$$\rho_d(t) = \begin{cases} (\dot{\varepsilon}_{rem}(t) - \alpha) \frac{g(0, \eta_S(t))}{\int_0^{\eta_H(t)} g(E, \eta_S(t)) dE} & if\ \eta_L(t) = 0 \\ 0 & otherwise \end{cases} \tag{3.30}$$

As a summary, the Energy Model II introduces a new state variable $\eta_S(t)$ among former $\eta_L(t)$ and $\eta_H(t)$ and describes the dynamics of the energy distribution, hence the birth and death rates, in a more realistic manner.

It should be noted that the initial energy distribution in this model is left undetermined on purpose and will be selected in Section 3.7 after presenting robustness

test results to different initial conditions.

3.4. State Equations of the Macro-scale Models

For the sake of convenience the total toxicity $\zeta(t)$ is replaced by another state variable proportional to it, $\theta(t) = \hat{\gamma}\zeta(t)$ and in the state equation of this new variable $\dot{\theta} = \hat{\gamma}\dot{\zeta}(t) = \omega\hat{\gamma}p(t)$ the parameters are lumped into $r = \theta\hat{\gamma}$. Consequently, $\dot{\varepsilon}_{rem}(t) = \frac{\hat{\gamma}\zeta(t)}{M}$ as given in Equation 3.13 can be expressed as $\dot{\varepsilon}_{rem}(t) = \frac{\theta(t)}{M}$. How the appropriate values of parameters M, α, $\hat{E}_{div}$ and r can be chosen, will be explained in Section 3.5.

It is also important to mention that the dynamics of birth rate $\rho_b(t)$ and death rate $\rho_d(t)$ are time dependent due to $\theta(t)$ in Macro-scale Model I and due to $\theta(t)$ and $\eta_S(t)$ in Macro-scale Model II, whereas the conditions of these rates are dependent on $\eta_H(t)$ and $\eta_L(t)$ in both models. Therefore, it is more appropriate to represent them in terms of state variables as follows:

For Macro-scale Model I :

$$\rho_b(t) = \rho_b\left(\zeta(t), \eta_H(t), \eta_L(t)\right)$$
$$\rho_d(t) = \rho_d\left(\zeta(t), \eta_H(t), \eta_L(t)\right)$$

For Macro-scale Model II :

$$\rho_b(t) = \rho_b\left(\zeta(t), \eta_H(t), \eta_L(t), \eta_S(t)\right)$$
$$\rho_d(t) = \rho_d\left(\zeta(t), \eta_H(t), \eta_L(t), \eta_S(t)\right)$$

Finally, let us remind that $p(t)$, $\zeta(t)$, $\eta_L(t)$ and $\eta_H(t)$ are obviously defined as non-negative variables since these variables have a physical correlate.

3.4.1. Macro-scale Model I

Macro-scale Model I uses the Quiescent Model I (Section 3.2.2.1) for quiescent population approximation and Energy Model I (Section 3.3.3.1) for energy distribution dynamics. With the change of variables mentioned above and in a compact mathematical notation the 4-dimensional Macro-scale Model I is given in Equations (3.31) to (3.34):

$$\dot{p}(t) = \begin{cases} [p(t) - q(t)]\rho_b(t) & if\ \eta_H(t) = \hat{E}_{div}\ and\ \alpha > \frac{\theta(t)}{M} \\ -p(t)\rho_d(t) & if\ \eta_L(t) = 0\ and\ \alpha < \frac{\theta(t)}{M} \\ 0 & otherwise \end{cases} \quad ;\ p(0) = p_0 \tag{3.31}$$

$$\dot{\theta}(t) = rp(t)\ ;\ \theta(0) = 0 \tag{3.32}$$

$$\dot{\eta_H}(t) = \left[\alpha - \frac{\theta(t)}{M}\right]_- \ ;\ \eta_H(0) = \hat{E}_{div} \tag{3.33}$$

$$\dot{\eta_L}(t) = \left[\alpha - \frac{\theta(t)}{M}\right]_- \ ;\ \eta_L(0) = \frac{\hat{E}_{div}}{2} \tag{3.34}$$

where $\rho_b(t) = \frac{2}{\hat{E}_{div}}\left(\alpha - \frac{\theta(t)}{M}\right)$; $\rho_d(t) = \frac{1}{\eta_H(t)}\left(\frac{\theta(t)}{M} - \alpha\right)$ and $[x]_-$ denotes $min(x, 0)$.

Although the macro-scale model is 4-dimensional, a close inspection reveals that $\eta_H(t)$ and $\eta_L(t)$ change in the same manner and therefore keep the initial difference until cells start to die. Hence until then, one can say that the system remains within a 3-dimensional manifold of the 4-dimensional state space.

3.4.2. Macro-scale Model II

In this model, quiescent population is approximated by the Quiescent Model II (Section 3.2.2.2) and the dynamics of the energy distribution is represented with Energy Model II (Section 3.3.3.2) . With the addition of η_S, 5-dimensional Macro-scale Model

II is presented as follows:

$$\dot{p}(t) = \begin{cases} [p(t) - q(t)]\rho_b(t) & if\ \eta_H(t) = \hat{E}_{div}\ and\ \alpha > \frac{\theta(t)}{M} \\ -p(t)\rho_d(t) & if\ \eta_L(t) = 0\ and\ \alpha < \frac{\theta(t)}{M} \\ 0 & otherwise \end{cases} \quad ;\ p(0) = p_0 \tag{3.35}$$

$$\dot{\theta}(t) = rp(t)\ ;\ \theta(0) = 0 \tag{3.36}$$

$$\dot{\eta_H}(t) = \left[\alpha - \frac{\theta(t)}{M}\right]_- \ ;\ \eta_H(0) = \hat{E}_{div} \tag{3.37}$$

$$\dot{\eta_L}(t) = \left[\alpha - \frac{\theta(t)}{M}\right]_- \ ;\ \eta_L(0) = \frac{\hat{E}_{div}}{2} \tag{3.38}$$

$$\dot{\eta_S}(t) = \alpha - \frac{\theta(t)}{M}\ ;\ \eta_S(0) = 0 \tag{3.39}$$

where $\rho_b(t) = \left(\alpha - \frac{\theta(t)}{M}\right) \frac{g(\hat{E}_{div}, \eta_S(t))}{\int_{\frac{\hat{E}_{div}}{2}}^{\hat{E}_{div}} g(E, \eta_S(t))dE}$; $\rho_d(t) = \left(\frac{\theta(t)}{M} - \alpha\right) \frac{g(0, \eta_S(t))}{\int_0^{\eta_H(t)} g(E, \eta_S(t))dE}$ and $g(E, \eta_S(t)) = f(E, 0) * \sum_{k=0}^{\infty} 2^k \delta\left(E - \eta_S(t) + k\frac{\hat{E}_{div}}{2}\right)$

A careful inspection reveals that the Macro-scale Model II, despite being 5-dimensional, does not occupy all 5 dimensions at all times. $\eta_S(t)$ changes in the same manner with $\eta_H(t)$ when population is increasing and also with $\eta_L(t)$ when population does not change. Therefore, one can state that the system remains within 4-dimensional manifold during population increase, shifts even to a reduced 3-dimensional manifold when population increase stops and enters to a different 4-dimensional manifold of its 5-dimensional state space when cells start to die.

As a conclusion, we can say that the system never occupies all dimensions of the five dimensional model simultaneously.

3.5. Choice of Parameters

Some parameters of the macro-scale model can be properly chosen on the basis of available biological knowledge, while others have to be estimated from empirical data gathered via relatively simple measurements on the cell culture at hand. It is worth

remembering that the only measurable quantity in this system is the population size.

3.5.1. Number of Loci in the Flask (M)

A common feature of mammalian cell cultures is that cells adhere to the flask ground as long as they are alive. Although in crowded regions cells can sometimes live on top each other, such allocations eventually lead to physiological changes in the cells and are therefore avoided by experimenters. Hence, in this thesis we stick to the biologically rather plausible assumption that living cells can only form a single layer on the flask bottom. With this assumption, the maximum number (M) of cells the flask can hold, can be easily estimated as $M = (flask\ area)/(average\ cell\ area)$, where the average cell area can be measured under the microscope.

3.5.2. Energy Intake Rate (α)

A careful inspection of the macro-scale equations shows that the energy intake rate of a cell (α) cannot be easily estimated because $p(t)$, the only observable variable, does not depend on it. Therefore, α can be chosen arbitrarily without affecting the dynamics of the observable variable, but any choice of α requires some scaling in the non-observable dynamics. In this sense, α serves for a degree of freedom in system dynamics and in this thesis it has been taken as 1 for the sake of convenience.

3.5.3. Division Energy ($\hat{E}_{div}$)

For the sake of distinguishing the same concepts at meso- and macro-scales, a 'hat' symbol is used with the variable name when referring to the concepts at the macro-scale model if there is a counterpart for that concept in the Meso-scale Model. So, as stated in Section 3.3.1 $\hat{E}_{div}$ corresponds to the same concept as E_{div} in the Meso-scale Model, but interpreted from a macroscopic point of view.

In order to estimate the division energy $\hat{E}_{div}$ one needs to conduct a test with an initially clean flask and a relatively diluted and uniformly distributed initial population,

so that quiescent population is still zero or can be neglected. After the first cell division is observed, population is recorded and the flask is kept clean by removing toxicity regularly (every 15 minutes $\equiv$ 1 time step is reasonable) until the initial population is doubled. One can deduce from reasoning that cells with lowest energy at the beginning of recording session, i.e. cells with energy $\hat{E}_{div}$, give birth at the doubling time T_d, which can be expressed mathematically as:

$$\frac{\hat{E}_{div}}{2} + \alpha T_d = \hat{E}_{div} \tag{3.40}$$

Equation 3.40 can be rearranged to find out $\hat{E}_{div}$:

$$\hat{E}_{div} = 2\ \alpha\ T_d \tag{3.41}$$

In this thesis, we have estimated $\hat{E}_{div}$ from 20% initial population concentration. For a more robust estimate, $\hat{E}_{div}$ can be calculated for a few diluted initial populations and their average can be taken.

3.5.4. Toxicity Coefficient (r)

The toxicity release coefficient ω (defined in Equation 3.21) and remediation energy coefficient $\hat{\gamma}$ (defined in Equation 3.13) cannot be estimated separately using the only observable variable $p(t)$. But this is not necessary anyway because these two parameters can be merged in to a single one $r = \hat{\gamma}\omega$.

$$\dot{\varepsilon}_{rem}(t) \approx \frac{\hat{\gamma}\zeta(t)}{M} = \frac{\hat{\gamma}\omega}{M}\int_0^t p(\tau)d\tau = \frac{r}{M}\int_0^t p(\tau)d\tau \tag{3.42}$$

Careful inspection of the Meso-scale Model reveals that the population increases until the remediation rate ($\dot{\varepsilon}_{rem}(t)$) reaches the energy intake rate (α). From that point

(t^*) onwards the energy of an average cell starts to decrease. So, at this time instance:

$$\dot{\varepsilon}_{rem}(t^*) = \alpha = \frac{r}{M} \int_0^{t^*} p(\tau) d\tau \tag{3.43}$$

For the estimation of r, a test (real or simulated) is to be conducted and the population size is recorded until t^*, i.e. when the population stops increasing. The estimate r can thus be extracted from Equation 3.43 as follows:

$$r = \frac{\alpha M}{\int_0^{t^*} p(\tau) d\tau} \tag{3.44}$$

In this thesis, r is estimated also from a test with 20% initial population occupation, just like the estimation of $\hat{E}_{div}$.

3.6. Simulation Results and Their Evaluation

The meso- and macro-scale models presented in this theses are general representations of the dynamics of cell culture populations, suitable for most eukaryotic cell types, provided that the parameters are appropriately chosen on basis of real data.

The Meso-scale Model, the simulation results of which are presented in this section, has parameters within realistic ranges, chosen using a combination of some partial data (obtained from a C2C12 cell line) and general recommendations of an expert. Therefore, the presented meso-scale simulation results can be considered as a realistic example although they do not necessarily correspond to a specific real cell type. The most speculative part of the Meso-scale Model is the initial energy distribution of cells which is very hard to make a realistic assumption about. This stems partially from the fact that 'energy' as used in the Meso-scale Model (and also adopted by the macro-scale model) is actually an abstract combination of the size and metabolic energy of a cell. During the evaluation and analysis of the macro-scale results (Sections 3.6.1 to 3.6.3) we assumed that both the Meso-scale Model and the macro-scale models have

initially the same energy distribution, namely a uniform one between $E_{div}/2$ and E_{div} because we do not want the shape of the initial energy distribution has an impact on the performance of the macro scale models. However, in Section 3.7 we will also provide a compared analysis of the prediction performances of the macro-scale models for different (and more realistic) initial energy distributions and an evaluation of the robustness of the initial energy distribution to be used in the Macro-scale Model II.

The difference between the progress speeds of toxicity diffusion and physiological changes (including cellular changes and toxicity release) is rather high. In the Meso-scale Model this difference is accounted for by using a much shorter step size in the toxicity diffusion simulation as compared to the one used in the simulation of physiological dynamics. Taking the frequency of population size measurements under realistic experimental conditions into consideration, the step size of the physiological dynamics is heuristically chosen as equivalent to 15 minutes, while the step size of the diffusion dynamics is taken as 1/20 of this value, namely 45 seconds. The diffusion coefficient used in Equation 2.2 as $D = 4$ is low enough to satisfy the Courant-Friedrichs-Lewy (CFL) condition[5] ($D \leq D_{max} = \frac{1}{4\Delta k_D} = \frac{1}{4\frac{\Delta k}{20}} = 5$) but by far high enough (as justified by meso-scale simulation results) to legitimize the assumption about the ongoing uniformity of the toxicity distribution. The exact choice of D is not critical as long as it remains around the given value. Taking into account that the area of a regular cell in our experiment is around 1500 μm^2, $D = 4$ corresponds approximately to a diffusion rate of $66.7 \times 10^{-11} \frac{m}{s}$, which is in the realistic range of diffusion rates for small particles in a cytoplasm [25].

The macro-scale models are lumped-parameter representations of the actual dynamics and thus constitute a higher level mathematical abstraction. Their performances are evaluated on basis of several criteria that measure how efficiently they predict the population occupancy generated by the Meso-scale Model. For such a performance evaluation, the macro-scale parameters $\hat{E}_{div}$ and r have been estimated from meso-scale simulation results (obtained for an initial population occupancy of 20%) by

[5]Courant-Friedrichs-Lewy (CFL) condition states the uniqueness and stability conditions for numerical solutions of certain partial differential equations [26].

using the methods explained in Sections 3.5.3 and 3.5.3, respectively.

The time profile of the population occupancy obtained from these meso-scale simulations exhibit the typical behaviour of real cell cultures for a wide range of initial occupancies; i.e. increasing up to a certain level (plateau), staying there for a while and then decreasing rapidly until complete extinction. When the parameters of the macro-scale model are appropriately estimated from meso-scale simulation results using the methods presented in Section 3.5, the macro-scale model results exhibit a similar behaviour. Figure 3.18a depicts the population occupancy as a function of time for an initial occupancy of 10%. Although the other variables are empirically not observable, they are shown in Figure 3.18b-c-d for the sake of a general evaluation of the macro-scale model results against the meso-scale data. Here, instead of the macro-scale state variable θ, a physically more meaningful variable, the total toxicity ζ, is presented. Note that the simulation results in Figure 3.18b-c-d are obtained for energy intake rate $\alpha = 1$, which is equivalent to A (the energy intake per step) used in the Meso-scale Model. For any other choice of α, the population profile predicted by the macro-scale model is still the same, while the other (non-observable) state variables are scaled by a function of α.

3.6.1. Performance Evaluation of the Macro-scale Models

Since the macro-scale models are developed with the purpose of predicting the population size, their performance are evaluated on the basis of population related criteria alone. In this thesis, the following practically relevant performance criteria have been used:

- percentage error in the maximum population occupancy p_{max}:

$$\%\epsilon_{pmax} = 100\% \ \left| p_{max}^{macro} - p_{max}^{meso} \right| / p_{max}^{meso} \tag{3.45}$$

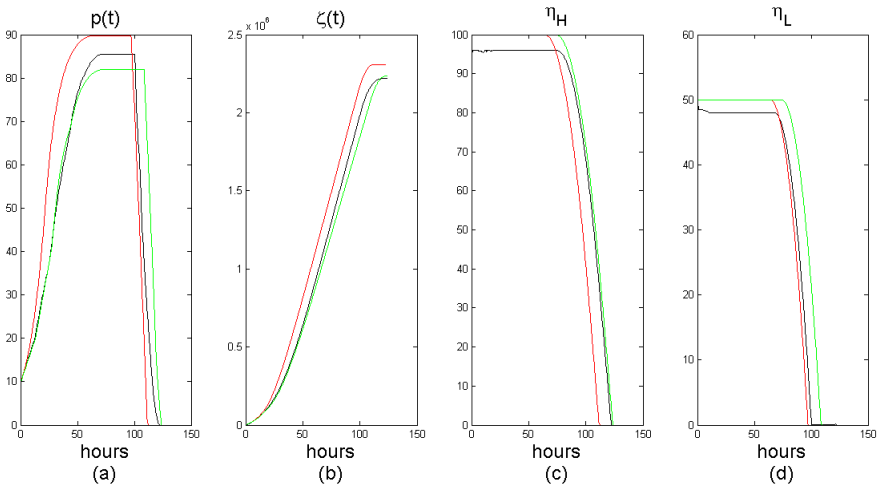

Figure 3.18. Time profile of (a) the population occupancy $p(t)$, (b) the total toxicity $\zeta(t)$, (c) the upper boundary of energy distribution $\eta_H(t)$ and (d) the lower boundary of energy distribution $\eta_L(t)$. These results are obtained for an initial occupancy of 10% (Black: Meso-scale model; Red: Macro-scale Model I; Green: Macro-scale Model II).

- percentage error in t_d, the beginning of population decrease after reaching p_{max}:

$$\%\epsilon_{td} = 100\% \ \left| t_d^{macro} - t_d^{meso} \right| / t_d^{meso} \tag{3.46}$$

- percentage error in t_e, the extinction time of the population:

$$\%\epsilon_{te} = 100\% \ \left| t_e^{macro} - t_e^{meso} \right| / t_e^{meso} \tag{3.47}$$

- the RMS error in the population size as a percentage of the RMS value of the meso-scale population size:

$$\%\epsilon_{RMS} = 100\% \ \frac{\sqrt{\frac{1}{t_{e,m}} \int_0^{t_{e,m}} [p^{macro}(t) - p^{meso}(t)]^2 dt}}{\frac{1}{t_{e,m}} \int_0^{t_{e,m}} [p^{meso}(t)]^2 dt} \tag{3.48}$$

where $t_{e,m} = max(t_e^{macro}, t_e^{meso})$

Percentage performance measures of both Macro-scale Model I and II in terms of these criteria are shown in Figure 3.19 for all possible values of initial occupancies.

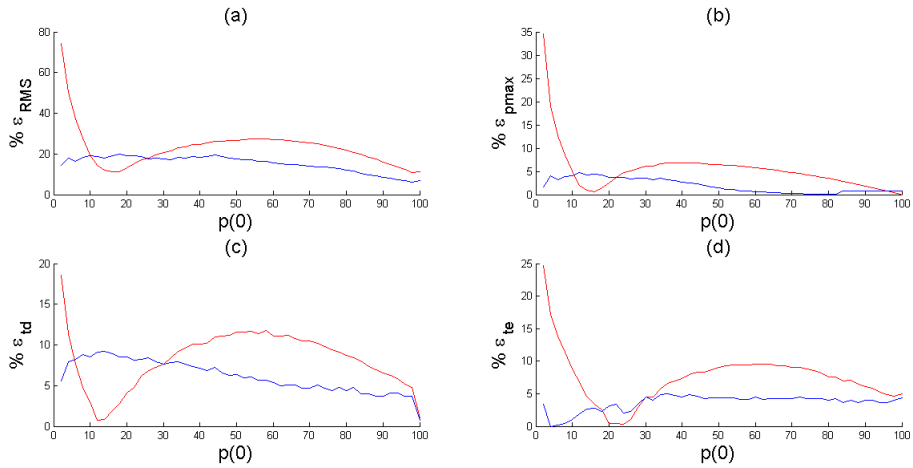

Figure 3.19. The performance criteria which evaluate how well the macro-scale model results (with α=1) resemble the Meso-scale Model results: (a) $\%\epsilon_{RMS}$, (b) $\%\epsilon_{pmax}$, (c) $\%\epsilon_{td}$ and (d) $\%\epsilon_{te}$ as a function of initial occupancy of the flask $p(0)$ (Red: Macro-scale Model I, Blue: Macro-scale Model II).

For Macro-scale Model I, population prediction performance measured in terms of all four criteria shows a common qualitative dependence on initial occupancy: poor performance for very low initial occupancies, a rapid improvement of performance when the considered initial occupancy approaches an intermediate value around 10-30%, a slight degradation of performance when the considered initial occupancy approaches a value around 50-60% and increasingly better performance for higher initial occupancies. On the other hand, Macro-scale Model II provides a much more consistent performance for all four criteria and even a better one than the Macro-scale Model I in general except for approximately 10-30% initial occupancy. Although Macro-scale Model II is developed with better approximations for quiescent occupancy and energy distribution dynamics, it has an inferior performance for this specific initial occupancy range. The performance of Macro-scale Model I and II will be analysed more closely in the following sections.

3.6.2. Detailed Performance Analysis for Macro-scale Model I

There are basically two approximations in the Macro-scale Model I which affect the population prediction performance seen in Figure 3.19: (i) approximation of the quiescent population according to Equation 3.4 and (ii) approximation of the energy distribution as a uniform one at all times as described in Section 3.3.3.1.

3.6.2.1. Effect of Quiescent Model I. $q(t)$ is estimated under the assumption that cells are uniformly distributed across the flask at all times. However, when the system starts with a uniformly distributed low initial occupancy and cell divisions start separate clusters are formed around the loci of the initial cells and thus the uniform distribution assumption is slightly violated until these s grow and start to merge. But in these cases -although the percentage estimation error is not very small- the actual quiescent occupancy is so small that approximating it as zero still provides a rather acceptable macro-scale model performance.

On the other hand, this approximation results in an overestimation of $q(t)$ for populations occupying more than 50% of the flask for nearly all initial occupancies except very low ones. Even for these initial occupancies, quiescent occupancy is overestimated after population reaches 50%-60%. However, as the population approaches maximum capacity nearly all cells enter quiescence mode and the overestimation in Quiescent Model I becomes negligible (Figure 3.20).

In summary, the estimation of the active population (p-q) performs well for $p(t)$ below 50% and for very high population sizes, while in between the active population is underestimated by Macro-scale Model I.

3.6.2.2. Effect of Energy Model I. Both in the meso-scale and macro-scale models it is assumed that the distribution of cell energies across the population is initially uniform, as explained in previous chapters. Macro-scale Model I further assumes that the energy distribution will remain uniform at all times, which is actually in conflict with the meso-

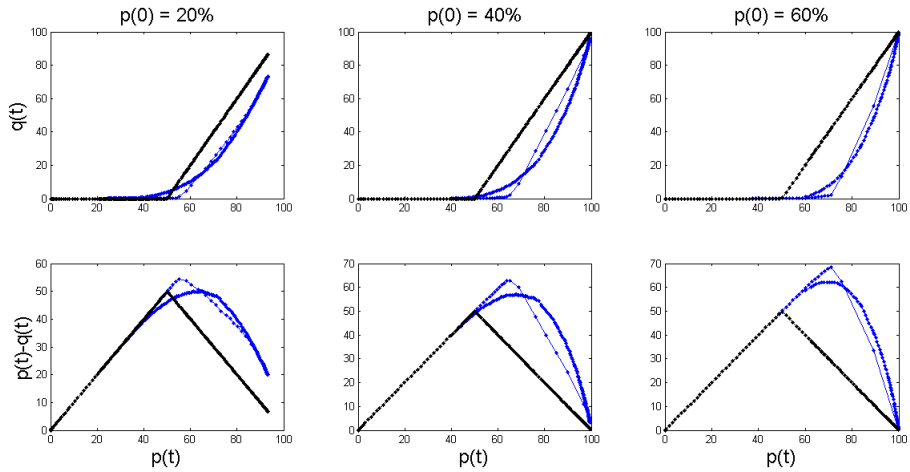

Figure 3.20. Quiescent occupancy $q(t)$ versus $p(t)$ and active population occupancy $(p(t) - q(t))$ versus $p(t)$ for three different initial occupancies (Black: Quiescent Model I, Blue: Meso-scale model). The sparse secondary part of the blue curve corresponds to the data after cells start to die.

scale simulation results. The impact of this rough approximation on the prediction performance of the macro-scale model needs to be evaluated.

For this purpose, let us first consider the actual characteristics of the energy distribution as observed in meso-scale results (Figure 3.21) for the early stages when the effect of toxicity is negligible. As the cell energies keep on increasing due to constant food intake, the histogram shifts steadily towards higher energy values (as long as the population is not very high and thus most of the cells are in active mode). On the other hand, since the cells that reach E_{div} undergo mitotic division, their number is doubled while their energies are halved, which folds the part of the histogram exceeding E_{div} back to $E_{div}/2$. Under the influence of this 'shifting and folding', in despite of the increasing toxicity effect and quiescent population presence, the actual energy distribution loses and regains a relatively uniform shape over and over again as can be seen in Figure 3.21.

The Macro-scale Model I calculates the number of cells that will undergo mitotic division within the next time step under the assumption that the energies of active cells are uniformly distributed. However, when it becomes non-uniform (see Figure 3.21b-c-

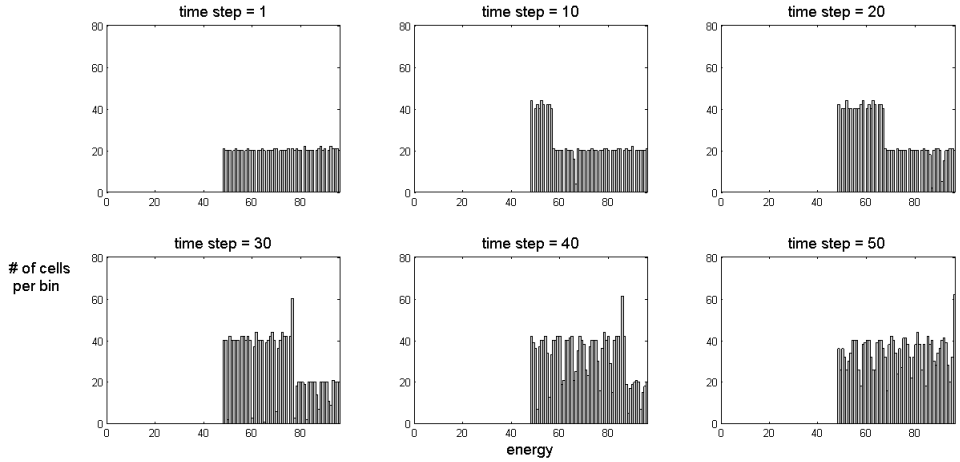

Figure 3.21. Energy histograms at the (a) 1st , (b) 10th , (c) 20th , (d) 30th, (e) 40th and (f) 50th time steps for an initial occupancy of 10%.

d-e) the energy distribution is typically biased in such a way that higher energy levels are less occupied. In those phases the uniform energy distribution assumption overestimates $\rho_b(t)$ (Equation 3.24) and consequently the overall population (Figure 3.22).

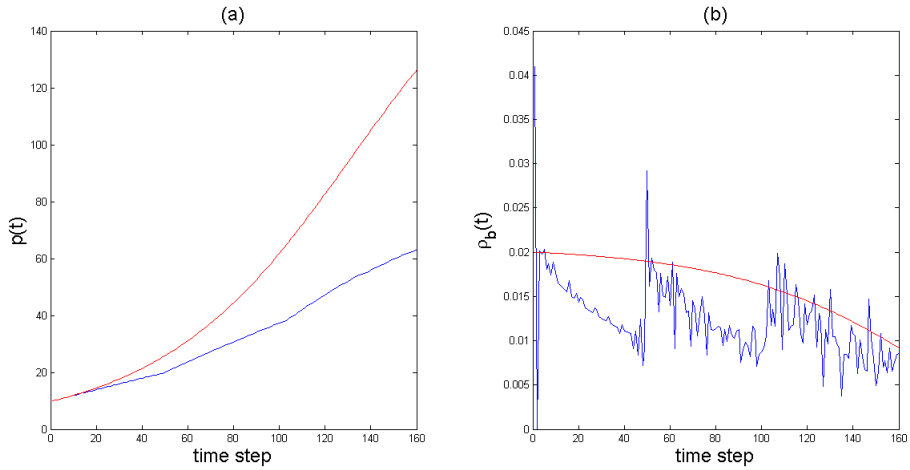

Figure 3.22. (a) $p(t)$ and (b) birth rates $\rho_b(t)$ of the Meso-scale Model (blue) and Macro-scale Model I (red) starting from 10% initial occupancy during the early stage of population growth. Macro-scale model I has been simulated using real quiescent occupancy values from Meso-scale Model to make a reliable comparison.

It should be noted that when the population is very high, most of the cells are in quiescent mode, thus fewer divisions occur and the uniformity of energy distribution is

disturbed less, resulting in a smaller overestimation of $\rho_b(t)$.

3.6.2.3. The combined effect of Quiescent Model I and Energy Model I. In summary, these two approximations exhibit relatively poor performance for some initial occupancy ranges: active (non-quiescent) population occupancy is underestimated when $p(t) > M/2$, while $\rho_b(t)$ is overestimated for all $p(t)$ values. According to the first line of Equation 3.31 these two factors with opposite errors compensate each other for populations occupying more than half of the flask, such that the macro-scale model approximates $\dot{p}(t)$ rather well in that population range, while for lower populations $\dot{p}(t)$ is overestimated.

It should be noted that the approximations in Quiescent Model I and Energy Model I directly affect $\dot{p}(t)$, whereas the performance criteria are evaluated on basis of $p(t)$. Let us consider how the approximation error affect $p(t)$ for different initial occupancy ranges.

Simulation results obtained from the Meso-scale Model reveal that for $p(0) \geq M/2$ the error due to underestimation of the active population occupancy dominates the error due to overestimation of the energy distributions. Hence for these population occupancies, the population increase rate and thus also the population occupancy are initially underestimated until cells start to die. Since underestimated population occupancy implies also an underestimation of the total toxicity, population decrease due to toxicity in the macro-scale model occurs later than in the meso-scale. However, once the population decrease occurs, the decrease rate predicted by the macro-scale model is very close to that of the Meso-scale Model as can be seen in Figure 3.23a. Keeping in mind that under these conditions the macro-scale model is operating according to the second line of Equation 3.31 ($\eta_L(t) = 0$ and $\alpha < \frac{\theta(t)}{M}$) the parallel decrease of $p(t)$ in meso- and macro-scale models indicates that the assumption about uniform energy distribution holds to a great extent.

However; for $p(0) < M/2$, as long as $p(t)$ remains below $M/2$, the uniform energy

distribution assumption results in an overestimation of $p(t)$. As soon as $p(t)$ exceeds $M/2$, the error due to underestimation of the active population occupancy gains significance and starts to suppress the accumulated overestimation of $p(t)$. Consequently, the Macro-scale Model I approximates the course of $p(t)$ best for $p(0)$ around 10-20% (Figure 3.23b), which can also be observed in Figure 3.19 where all four population-based evaluation criteria exhibit the best performance for this initial population range.

With much lower $p(0)$ it takes longer for the population to reach $M/2$, and thus the overestimation in $\dot{p}(t)$ (due to the uniform energy distribution approximation) has more time to accumulate. By the time $p(t) = M/2$, the population occupancy estimated by the macro-scale model is much higher than in the meso-scale. Consequently, the underestimation of the active population occupancy (which occurs from the time onwards when $p(t)$ exceeds $M/2$) cannot sufficiently compensate the overestimation of $p(t)$ as shown in Figure 3.23c.

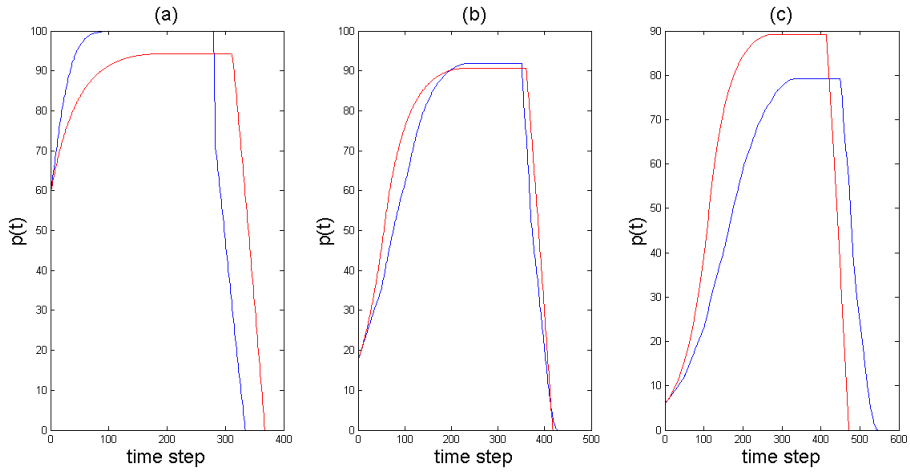

Figure 3.23. Time profile of $p(t)$ according to the meso-scale(blue) and Macro-scale Model I (red) for different initial occupancies, $p(0)$: a)60% b)18% and c)6% of the flask.

3.6.3. Detailed Performance Analysis for Macro-scale Model II

3.6.3.1. Effect of the Quiescent Model II. Both Quiescent Models I and II approximate $q(t)$ in a piecewise linear manner; namely as zero up to a certain quiescence onset point, p_q, and as a line with a positive slope κ. In Quiescent Model I, the quiescence

onset point p_q and κ are fixed at 50% and 2, respectively (Figure 3.20). In Quiescent Model II, however, p_q depends on how much time has passed since the beginning of births until the rise of quiescent population. This also renders the slope κ defined for $p > p_q$ dependent on time. A closer inspection from the meso-scale simulations reveals that Quiescent Model II underestimates the active population occupancy until the onset point is reached, and thereafter it tends to overestimate it slightly (Figure 3.24).

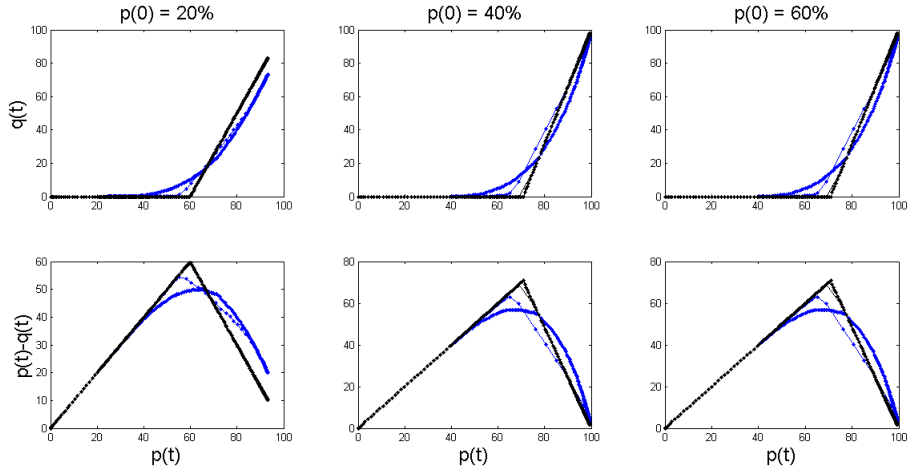

Figure 3.24. Quiescent occupancy $q(t)$ versus $p(t)$ and active population occupancy $(p(t) - q(t))$ versus $p(t)$ for three different initial occupancies (Black: Quiescent Model II, Blue: Meso-scale model). The sparse secondary part of the blue curve corresponds to the data after cells start to die.

Since the onset point in Quiescent Model II is not fixed, we observe that it estimates the quiescent occupancy on average better than Quiescent Model I for an arbitrary initial occupancy. In Table 3.1, we show the average absolute quiescent estimation error (Equation 3.49) for three different initial populations:

$$\epsilon_q = \frac{1}{t_e} \int_{t=0}^{t_e} |q_{meso}(t) - q_{macro}(t)| \, dt \tag{3.49}$$

3.6.3.2. Effect of the Energy Model II. Energy Model II is developed to account for the shifting and folding effects of energy distribution mentioned in Section 3.3.3.2. As a

Table 3.1. Average quiescent estimation error for 3 different initial occupancies.

	ϵ_q for $p(0) = 20\%$	ϵ_q for $p(0) = 40\%$	ϵ_q for $p(0) = 60\%$
Quiescent Model I	10.188	8.3429	6.1873
Quiescent Model II	6.4636	3.2988	1.4048

results of this, birth rate is neither overestimated nor underestimated significantly and gives a much better performance for predicting the time profile of $p(t)$ (Figure 3.25).

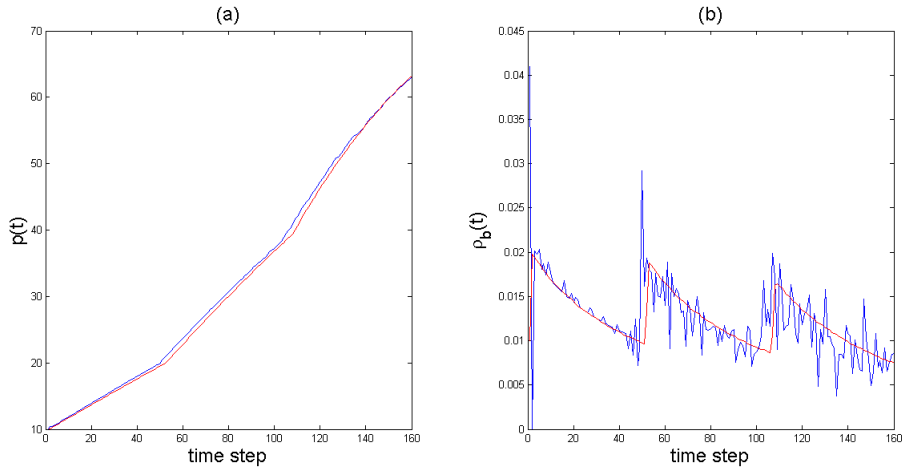

Figure 3.25. (a) $p(t)$ and (b) birth rates of the Meso-scale Model (blue) and Macro-scale Model II (red) starting from 10% initial occupancy. Macro-scale model has been simulated using real quiescent occupancy values from Meso-scale Model to make a reliable comparison.

However, it should be noted that the performance of the Energy Model I decreases for birth rates when $p(t)$ increases and reaches the plateau. We have observed from the meso-scale simulations that during that phase an important assumption made of the macro-scale models is violated, namely the assumption that the energy distributions of the active and total populations have the same shape. This assumption loses its validity when $p(t)$ increases significantly, since in a sufficiently crowded region new born cells enter immediately quiescence mode, giving rise to an unexpected accumulation of quiescent cells at energy level $E_{div}/2$. This excessive population of quiescent cells at a specific energy level deforms the energy distribution of active cells and thus renders the shapes of the energy distributions of active cells and all cells dissimilar. However, during

this phase the number of active cells decreases dramatically due to high population, and likewise the birth rate shrinks drastically due to toxicity (and reaches zero in the plateau region) which together reduce the effect of the discrepancy between the real and modelled energy distributions. Hence, we can assume that the Energy Model II provides in general a qualitatively good estimation for the evolution of energy distribution.

3.6.3.3. The Combined Effect of Quiescent Model II and Energy Model II. The performance of Macro-scale Model II is much more straight forward than Macro-scale Model I. The macro-scale Model II performs better than Macro-scale Model I except a specific region. The superiority of Macro-scale Model II is due to the better estimation capabilities of Quiescent Model II and Energy Model II.

Energy Model II predicts the total population significantly better than Energy Model I but still lacks modelling of the discrepancy between the energy distributions of active and quiescent cells which arises when population increases sufficiently. On the other hand, Quiescent Model II also produces estimation errors due to its piecewise linear nature, although not as high as Quiescent Model I. The combined effect of these slight over and under estimations constitutes a more consistent and better macro-scale model which is outperformed by the Macro-scale Model I only for initial occupations where the errors of Quiescent Model I and Energy Model I cancel each other quiet well.

3.7. Robustness to Different Initial Conditions

In Section 3.6.1, performance of the macro-scale models have been evaluated assuming that the meso- and macro-scale models have initially the same uniform energy distribution between $E_{div}/2$ and E_{div}. However, energy concept in this thesis is a virtual concept that combines the physical size and internal energy of cells, hence practically unmeasurable. Therefore, we expect that the physical and biological correlates in real cell cultures may correspond to initial energy distributions other than the assumed uniform distribution between $\hat{E}_{div}/2$ and $\hat{E}_{div}$. Therefore it is important to know how robust our macro-scale models are to variations in the real correlates of initial energy

distribution. For this purpose meso-scale simulations have been conducted using initial energy distributions of different shapes, the macro-scale models have been parameter-tuned for each of them and then the macro-scale performances have been evaluated. For a given macro-scale model, comparison of these performances gives an idea about its robustness to variations in the real correlates of the initial energy distribution. For this purpose a robustness measure has been defined. It is used to evaluate the robustness of Macros-scale Model I and II, as well as to select the most adequate initial energy distribution, for Macro-scale Model II among the ones considered. The details of this procedure are explained below.

As stated in Section 2.3, cells have initially low energies at the beginning of the experiment. So, it is reasonable to assume that the initial energy distribution is spread between 0 and $E_{div}/2$. The shape of the distributions used to evaluate the performance of the macro-scale model are selected biologically plausible in the sense that majority of cells do not occupy high levels of the energy distribution. Therefore, we decided to choose such four alternative energy distribution candidates which are mathematical realizations of four simple functions, namely a shifted Gaussian curve, a decaying line, a decaying exponential and finally the uniform function (Figure 3.26).

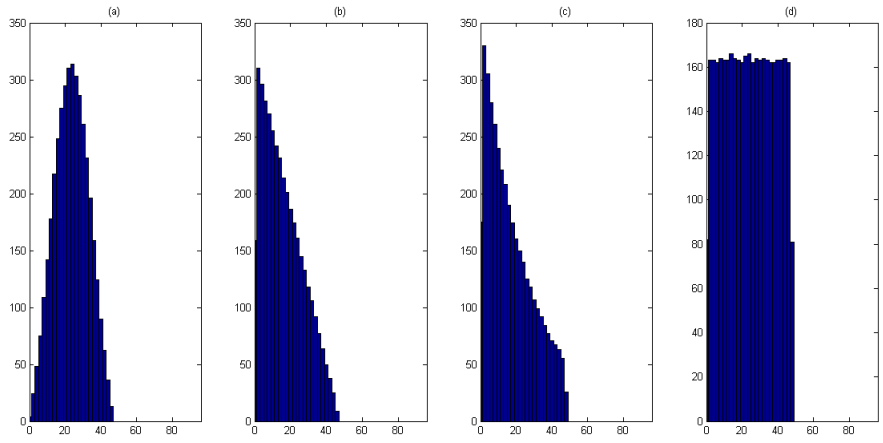

Figure 3.26. Initial energy distributions considered for robustness tests.

The macro-scale models assume that the flask is initially clean and the upper boundary of energy distribution starts at E_{div}. Regardless of the shape of the energy

distribution, these two assumptions can be easily satisfied if toxicity is removed immediately upon the observation of the first population increase. Thus, the performances of the macro-scale models (and also the estimation of their parameters) are evaluated only after the system has reached a state where the flask is clean and the population occupancy has started to increase. It is worth noting that waiting for the onset of cell divisions and cleaning the flask afterwards is experimentally feasible.

In this regard, the macro-scale parameters E_div and r have been estimated for each energy distribution shown in Figure 3.26a-d (using meso-scale simulation data gathered for $p(0) = 10\%$) as shown in Table 3.2.

Table 3.2. Estimates of E_{div} and r for different meso-scale initial energy distributions as described in Sections 3.5.3. and 3.5.4 and shown in Figure 3.26a-d.

	Meso-scale value	estimate for (a)	estimate for (b)	estimate for (c)	estimate for (d)
E_{div}	96	102	102	100	100
r	0.0067	0.0069801	0.0067016	0.0066218	0.0063334

In this thesis the following robustness measure R is proposed for macro-scale models:

- 4 different initial energy distributions between 0 and $E_{div}/2$ are considered:
  - (i) Gaussian: $f_a(E,0) = e^{-\frac{1}{2}\left(\frac{E-\mu}{\sigma}\right)^2} - e^{-\frac{1}{32}\left(\frac{\hat{E}_{div}}{\sigma}\right)^2}$ where $\mu = \frac{\hat{E}_{div}}{4}$ and $\sigma = \frac{\hat{E}_{div}}{8}$
  - (ii) Linearly decaying : $f_b(E,0) = -E + \frac{\hat{E}_{div}}{2}$
  - (iii) Exponentially decaying : $f_c(E,0) = e^{-0.04E}$
  - (iv) Uniform : $f_d(E,0) = 1$

    The numerical values of the parameters in these distributions have been heuristically selected.
- 4 different performance measures ϵ_i (i=1,...,4) have been defined for each initial

energy distribution as explained Section 3.6.1 :

$$\epsilon_1 = \epsilon_{RMS} \ ; \ \epsilon_2 = \epsilon_{pmax} \ ; \ \epsilon_3 = \epsilon_{td} \ ; \ \epsilon_4 = \epsilon_{te}$$

$\epsilon_i(p_0, k)$: i^{th} performance measure of a macro-scale model starting with $p(0) = p_0$ for predicting the results of meso-scale simulations with initial energy distribution $f_k(E, 0)$, where $k \in \{a, b, c, d\}$ as given above. Here, 50 different p_0 values have been considered: $p_0 \in S = \{2, 4, \ldots, 100\}$.

- $\epsilon_i^*(k)$, the average of $\epsilon_i(p_0, k)$ over $\forall p_0 \in S$, gives an idea about the average performance of a macro-scale model in estimating the results of a meso-scale simulation with initial energy distribution $f_k(E, 0)$:

$$\epsilon_i^*(k) = \frac{1}{cardinality(S)} \sum_{p_0 \in S} \epsilon_i(p_0, k) \tag{3.50}$$

- Finally R, the robustness of the macro-scale model under consideration, is defined as the relative percentage standard deviation of $\epsilon_i^*(k)$.

$$R = \frac{\sqrt{\frac{1}{4}\sum_k (\epsilon_i^*(k) - \overline{\epsilon}_i^*)^2}}{\overline{\epsilon}_i^*} * 100\% \tag{3.51}$$

where, $\overline{\epsilon}_i^*$ is the initial-distribution-invariant average performance (*idiap*) of a macro-scale model in estimating meso-scale simulation results, which is calculated as the average of $\epsilon_i^*(k)$ over $k \in \{a, b, c, d\}$:

$$\overline{\epsilon}_i^* = \frac{1}{4} \sum_{k \in \{a,b,c,d\}} \epsilon_i^*(k) \tag{3.52}$$

3.7.1. Macro-scale Model I

As can be observed from Figure 3.27 that all four performance criteria give qualitatively similar results except $\%\epsilon_{td}$, the error indicating the beginning of population decrease, which exhibits very bad performance for high initial populations. This is due to the fact that some cells at low energy levels become quiescent already at the

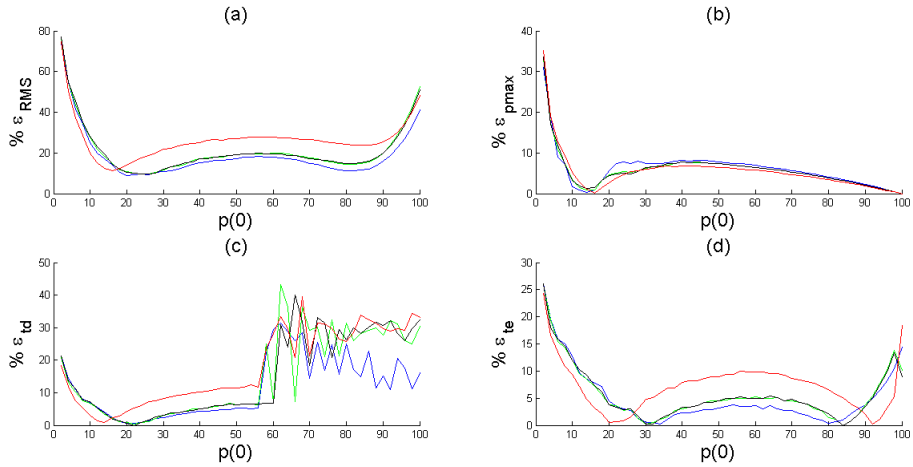

Figure 3.27. The performance of Macro-scale Model I in terms of four performance criteria for the meso-scale simulations with four different initial energy distributions: Gaussian curve (blue), decaying line (green), decaying exponential (black), uniform (red).

beginning of the experiment if the initial occupation is sufficiently high. These cells are not able to gain further energy and as a result of this, after total toxicity reaches a sufficient level, they die much earlier than the Macro-scale Model I anticipates the first death, which eventually results in a poor $\%\epsilon_{td}$ performance.

Furthermore, the performance of the Macro-scale Model I is also affected by the error of estimated parameters r and E_{div}, the best example of which is the visible discrepancy between the uniform (red line) and other energy distributions. The slightly worse estimation of the parameter r in uniform energy distribution gives rise to a worse performance for a majority of initial populations when predicting the overall shape of population curve, the death time and the extinction time. The effect of the parameter estimation error can be observed in more detail, if the macro-scale model is simulated with the same parameters of the Meso-scale Model as shown in Figure 3.28.

It is also intriguing for modelling intentions to observe that worse parameters might even create better results for some specific initial concentrations where the errors of modelled dynamics suppress the parameter estimation errors. This outcome will be discussed in forth chapter of the thesis in further detail.

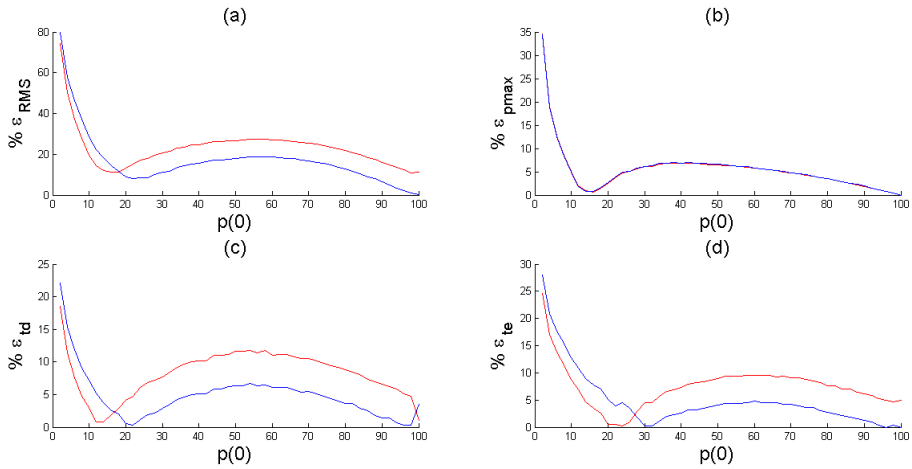

Figure 3.28. The performance of Macro-scale Model I in terms of the four performance criteria with estimated r and E_{div} (red line) and with real r and E_{div} from Meso-scale Model (blue line).

Table 3.3. Robustness measure for Macro-scale Model I.

Meso-scale Initial energy distribution	ϵ^*_{RMS}	ϵ^*_{pmax}	ϵ^*_{td}	ϵ^*_{te}
(a) Gaussian curve	18,79	6,12	11,72	5,03
(b) Decaying line	21,32	5,75	15,14	5,61
(c) Decaying exponential	21,34	5,79	15,04	5,64
(d) Uniform	26,45	5,47	17,67	7,21
idiap ($\overline{\epsilon}^*_i$)	21,97	5,78	14,89	5,87
Robustness measure	14,64	4,62	16,37	15,89

3.7.2. Macro-scale Model II

The Macro-scale Model II is not developed with a pre-selected initial energy distribution. Instead, it employs Energy Model II which describes how the energy distribution evolves. As a result, any shape can be used for the initial energy distribution in Macro-scale Model II. However, it is more reliable to select this shape according to the robustness and overall performance results of Macro-scale Model II.

For this purpose, we have assigned four different candidates as the initial energy distribution of Macro-scale Model II. Those candidates are the same energy distributions (Figure 3.26) used in robustness tests. The performance of each candidate is given in terms of four performance criteria for meso-scale simulations with four initial distributions.

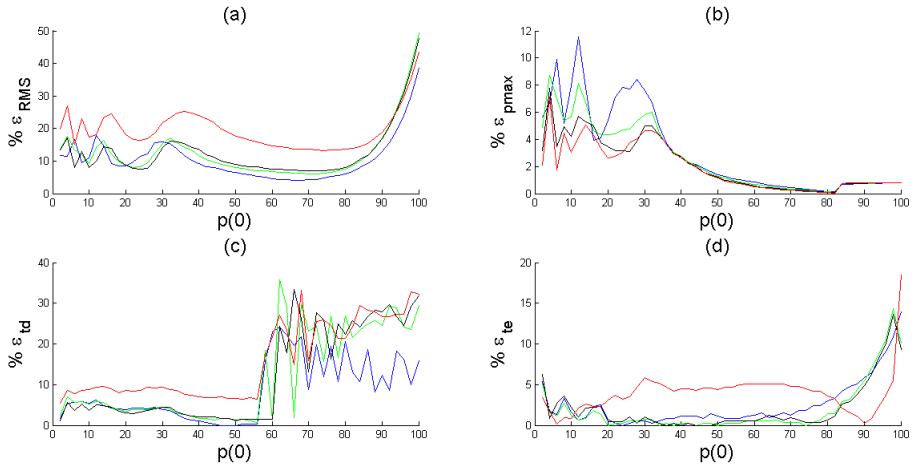

Figure 3.29. The performance of Macro-scale Model II with Gaussian initial distribution in terms of four performance criteria for the meso-scale simulations with four different initial energy distributions: Gaussian curve (blue), decaying line (green), decaying exponential (black), uniform (red).

In order to evaluate which initial energy distribution in Macro-scale Model II performs better, the *idiap* values for different macro-scale initial energy distributions are regrouped in Table 3.8. One can easily observe that the *idiap* values of linearly and exponentially decaying energy distributions are worse than Gaussian and uniform ones. Thus linearly and exponentially decaying energy distributions can be eliminated.

Table 3.4. Robustness measure for Macro-scale Model II with Gaussian initial energy distribution.

Meso-scale Initial energy distribution	ϵ^*_{RMS}	ϵ^*_{pmax}	ϵ^*_{td}	ϵ^*_{te}
(a) Gaussian curve	10,61	3,07	8,58	2,50
(b) Decaying line	12,55	2,57	11,98	1,92
(c) Decaying exponential	12,58	2,12	11,91	2,02
(d) Uniform	19,44	1,92	15,51	3,93
idiap ($\bar{\epsilon}^*_i$)	13,80	2,42	12,00	2,60
Robustness measure	28,11	21,24	23,56	35,71

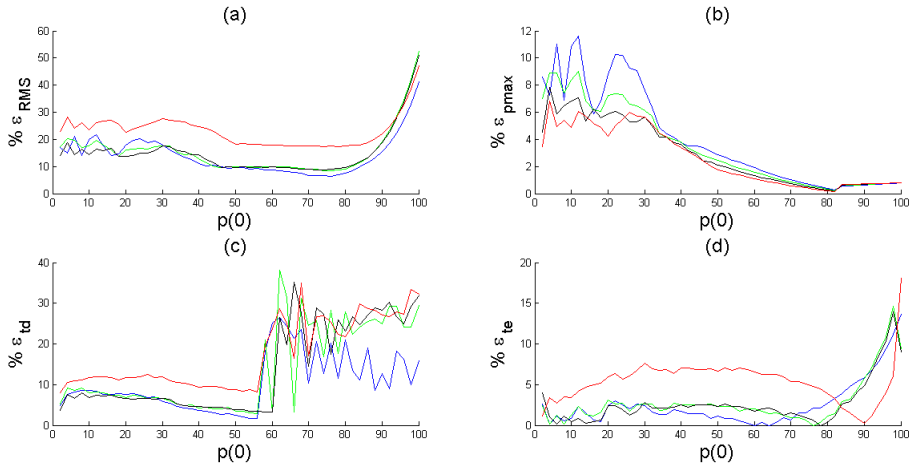

Figure 3.30. The performance of Macro-scale Model II with linearly decaying initial distribution in terms of four performance criteria for the meso-scale simulations with four different initial energy distributions: Gaussian curve (blue), decaying line (green), decaying exponential (black), uniform (red).

Table 3.5. Robustness measure for Macro-scale Model II with linearly decaying initial energy distribution.

Meso-scale Initial energy distribution	ϵ^*_{RMS}	ϵ^*_{pmax}	ϵ^*_{td}	ϵ^*_{te}
(a) Gaussian curve	14,08	3,99	10,46	2,52
(b) Decaying line	15,51	3,41	13,91	2,80
(c) Decaying exponential	14,98	2,90	13,79	2,63
(d) Uniform	23,03	2,63	17,35	5,36
idiap ($\overline{\epsilon}^*_i$)	16,90	3,23	13,88	3,33
Robustness measure	24,42	18,64	20,27	40,95

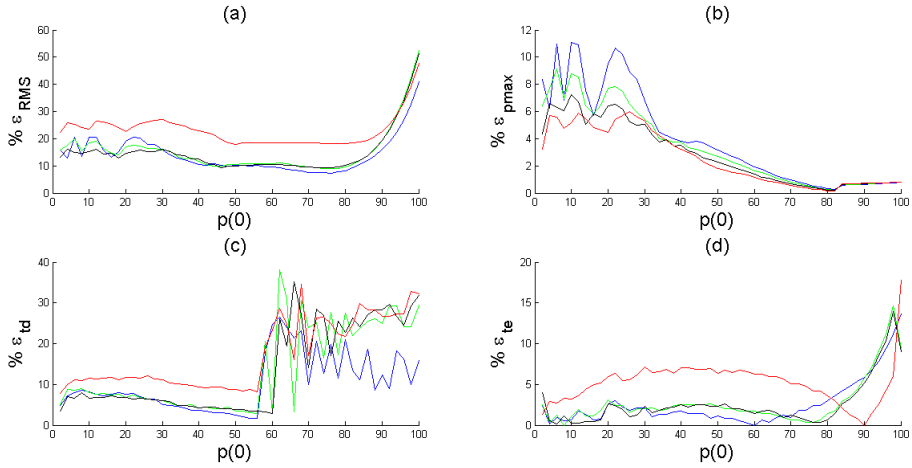

Figure 3.31. The performance of Macro-scale Model II with exponentially decaying initial distribution in terms of four performance criteria for the meso-scale simulations with four different initial energy distributions: Gaussian curve (blue), decaying line (green), decaying exponential (black), uniform (red).

Table 3.6. Robustness measure for Macro-scale Model II with exponentially decaying initial energy distribution.

Meso-scale Initial energy distribution	ϵ^*_{RMS}	ϵ^*_{pmax}	ϵ^*_{td}	ϵ^*_{te}
(a) Gaussian curve	14,10	3,94	10,31	2,54
(b) Decaying line	15,48	3,35	13,73	2,71
(c) Decaying exponential	14,86	2,85	13,56	2,52
(d) Uniform	22,94	2,56	17,13	5,17
idiap ($\overline{\epsilon}^*_i$)	16,85	3,17	13,68	3,23
Robustness measure	24,36	19,09	20,37	39,96

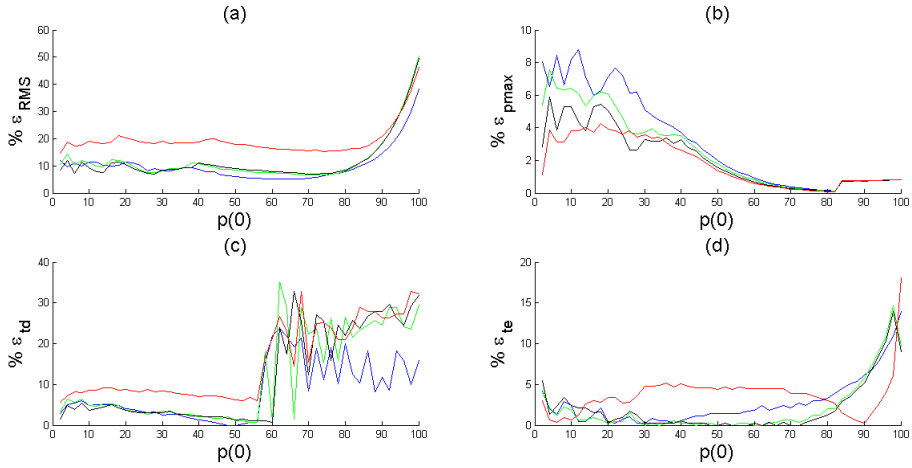

Figure 3.32. The performance of Macro-scale Model II with uniform initial distribution in terms of four performance criteria for the meso-scale simulations with four different initial energy distributions: Gaussian curve (blue), decaying line (green), decaying exponential (black), uniform (red).

Table 3.7. Robustness measure for Macro-scale Model II with uniform initial energy distribution.

Meso-scale Initial energy distribution	ϵ^*_{RMS}	ϵ^*_{pmax}	ϵ^*_{td}	ϵ^*_{te}
(a) Gaussian curve	10,16	3,14	8,37	2,71
(b) Decaying line	11,97	2,58	11,76	1,98
(c) Decaying exponential	11,70	2,11	11,69	2,01
(d) Uniform	19,30	1,84	15,24	3,68
idiap ($\overline{\epsilon}^*_i$)	13,28	2,42	11,76	2,60
Robustness measure	30,79	23,70	23,82	30,69

The remaining two distributions have rather close performances, but the Gaussian distribution exhibits higher robustness in three out of four criteria as shown in Table 3.9.

Table 3.8. *idiap* values for different initial energy distributions of the Macro-scale Model II.

Initial Distribution of Macro-scale Model II	ϵ^*_{RMS}	ϵ^*_{pmax}	ϵ^*_{td}	ϵ^*_{te}
Gaussian curve	13,80	2,42	12,00	2,60
Decaying line	16,90	3,23	13,88	3,33
Decaying exponential	16,85	3,17	13,68	3,23
Uniform	13,28	2,42	11,76	2,60

Hence, it seems appropriate (albeit not imperative or crucial due to its small difference) to assign the shifted Gaussian distribution as the initial energy distribution of the Macro-scale Model II (Equation 3.53).

$$f(E,0) = e^{-\frac{1}{2}\left(\frac{E-\mu}{\sigma}\right)^2} - e^{-\frac{1}{32}\left(\frac{\hat{E}_{div}}{\sigma}\right)^2} \tag{3.53}$$

where $\mu = \frac{\hat{E}_{div}}{4}$ and $\sigma = \frac{\hat{E}_{div}}{8}$

Table 3.9. Robustness of Gaussian and uniform macro-scale initial energy distributions.

Robustness measure	ϵ^*_{RMS}	ϵ^*_{pmax}	ϵ^*_{td}	ϵ^*_{te}
Gaussian curve	28,11	21,24	23,56	35,71
Uniform	30,79	23,70	23,82	30,69

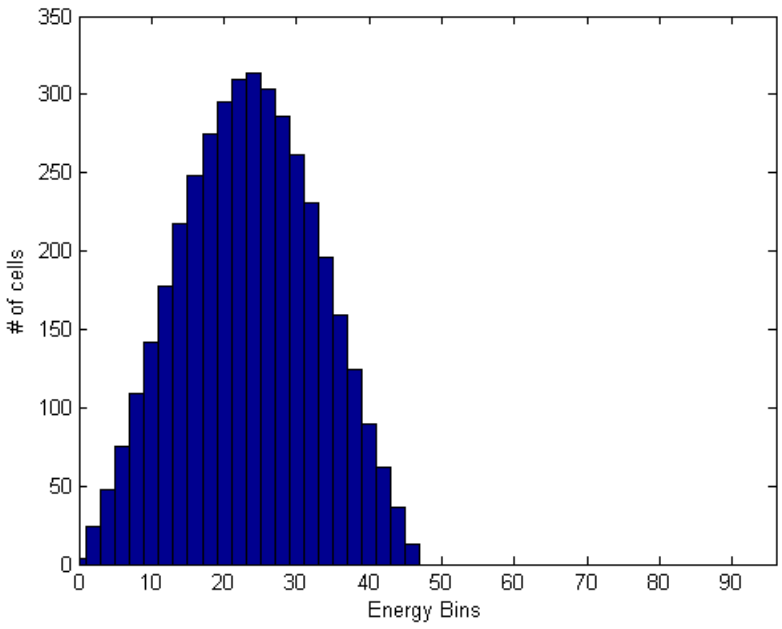

Figure 3.33. Gaussian-shaped energy distribution between 0 and $E_{div}/2$ is selected as the initial energy distribution of the Macro-scale Model II.

Finally, it is important to note that regardless of its initial energy distribution selection, Macro-scale Model II outperforms Macro-scale Model I in all four criteria. The comparison of overall performances (*idiap*) of Macro-scale Model I and Macro-scale Model II with the selected initial energy distribution is given in Table 3.10.

Table 3.10. Comparison of overall performances of Macro-scale Model I and Macro-scale Model II with the Gaussian-shaped initial energy distribution.

idiap ($\overline{\epsilon}_i^*$)	$\overline{\epsilon}_{RMS}^*$	$\overline{\epsilon}_{pmax}^*$	$\overline{\epsilon}_{td}^*$	$\overline{\epsilon}_{te}^*$
Macro-scale Model I	21,97	5,78	14,89	5,87
Macro-scale Model II	13,80	2,42	12,00	2,60

4. CONCLUDING REMARKS

The aim of this thesis was to develop a model that allows the user to predict how the population size of a eukaryotic cell culture will evolve starting from an arbitrary initial value. In this thesis two novel analytical models that can serve such a purpose have been obtained in two steps: First, a biologically plausible agent-based Meso-scale Model has been designed that makes use of the available knowledge about the behaviour of individual cells. Next, two autonomous macro-scale models have been derived from this Meso-scale Model.

The knowledge that constitutes the biological basis of this Meso-scale Model is mainly the recommendations of an expert and the insight gained from some empirical studies on a C2C12 cell culture. This Meso-scale Model is assumed to be customisable by properly selecting its parameters using some results of simple tests. We hope that the agent-based discrete-time Meso-scale Model, which mimics the dynamics of individual cells, provides a sufficiently detailed simulation environment where the evolution of the cell population can be computed starting from any initial condition. However, it should be noted that the phenomenon of cell differentiation has been excluded from the Meso-scale Model. Therefore, the Meso-scale Model can be used either for cell cultures that are incapable of differentiation or for differentiable cell cultures until cells undergo differentiation (which usually happens when the flask is completely occupied).

Under some uniformity assumptions, the flask-wide dynamics of the cell culture can be represented in a much more compact form, as a set of coupled nonlinear differential equations with some parameters that need to be customized using empirical data. In this regard, two analytical macro-scale models have been developed from the meso-scale agent-based model and by the intuition gained from its simulations. During the development of these models, some biologically plausible approximations have been made to represent lumped-parameter concepts. Here, one of the most important assumptions is that cells are spatially sufficiently uniformly distributed at all times. In order to evaluate the uniformity of the cell distribution across the flask in the Meso-

scale Model, a measure of heterogeneity has been proposed. Using this measure, it has been shown that the dynamics of the Meso-scale Model does not significantly amplify local heterogeneities, which supports the employment of the uniformity assumption.

Macro-scale model I is developed with further uniformity assumptions about the two important factors affecting the population dynamics, namely the dynamics of the energy distribution and the quiescent population. These rather simplistic assumptions lead to rather poor approximations in the associated parts of the population dynamics but these errors largely cancel each other, resulting in a relatively successful prediction of the population profile as obtained from the Meso-scale Model. Figure 3.19 demonstrates that the prediction performance of Macro-scale Model I is satisfactory for initial populations occupying more than 10% of the flask, and especially well for 10-30% occupancy.

Macro-scale Model II is developed in order to improve the approximations for both quiescent occupancy and energy distribution dynamics. With the insight gained from a closer investigation of meso-scale simulation results, Quiescent Model II and Energy Model II have been developed, both of which have a superior performance compared to their counterparts in Macro-scale Model I. Macro-scale Model II estimates the population occupancy in a more consistent manner for different initial population ranges.

Although Macro-scale Model II in general performs better than Macro-scale Model I in estimating the maximum population occupancy p_{max}, the death time t_d, the extinction time t_e and finally the overall qualitative shape of the population evolution, Macro-scale Model I surprisingly outperforms Macro-scale Model II for a specific initial occupancy range as can be seen from Figure 3.19. Therefore, for experiments with initial occupancy within this range the usage of Macro-scale Model I can be recommended. Here, it is worth to remark that from a modelling point of view the performance of the components in a model might be unimportant as long as the combined effect of these elements constitute a sufficiently well working system. This conclusion can be valid even for a single wrongly set parameter which has a positive impact on the outcome of

the model as in our case (Figure 3.28).

In that regard, it is impossible not to mention the quotation of G.E.P. Box: "Essentially, all models are wrong, but some are useful." and accordingly suggest that "Some are even more useful when they become 'wronger' ".

4.1. Future Contribution

The most important weakness of the models developed in this thesis is the lack of experimental validation. In fact, the Meso-scale Model has been developed to constitute a simulation environment for cell culture dynamics and provide a basis for lumped-parameter macro-scale models. In this thesis, the validation of macro-scale models has been done. according to the results of the meso-scale simulations. However, the Meso-scale Model, which is based on biological knowledge and qualitative empirical observations, needs to be better grounded and fine-tuned via detailed laboratory experiments that can provide quantitative validation. Therefore, the most important future work for this thesis would involve experimental works with different cell culture types and the validation of the meso-scale accordingly.

4.2. Modelling Challenge I : virtual entities

Besides providing a customisable model useful for predicting the evolution of the population size, this study also sheds light on how a compact analytic model of the collective behaviour of a large number of lower-scale components can be constructed on basis of the individual dynamics of these components. However the development of a compact and lucid macro-scale model from the knowledge of lower-scale dynamics may not always be achieved in a trivial manner by averaging, coarse-graining, aggregating etc. As a matter of fact, exactly such examples constitute interesting and challenging modelling problems and let the modeller go back and forth between choosing the adequate state variables and formulating their dynamics until an autonomous system representation is obtained. In the selection of state variables, observable variables of interest play the leading role. Generally during the formulation of the dynamics of

these observables other state variables turn out to be necessary.

The observable variables of interest are usually global variables representing the total or average behaviour of the collective, while the deviations from the average - i.e. the diversity within the collective- need to be represented by introducing further variables (if they have any contribution to the dynamics of the chosen observable variables). Such 'diversities' can be accounted for by introducing stochastic variables, but this would result in a stochastic macro-scale model. If one wants to remain within the realm of deterministic models, one can choose to express the distribution of the quantity of interest across the collective as a distribution function as in our case.

When constructing a macro-scale model, observable state variables and those that represent the distribution of quantities of interest across the collective may still not suffice to achieve the autonomy of the system. In this case, one may need to envisage 'virtual' variables, which do not necessarily have a lower-scale counterpart. As a matter of fact, the 'virtuality' of such variables is a matter of philosophical interpretation, because as long as they are formulated appropriately and meet pragmatic needs of modelling they can be considered rather 'real' from the macroscopic perspective; a very obvious example being the temperature in thermodynamics.

4.2.1. Modelling Challenge II : excessive reductions

Analytical macro-scale models are mathematical abstractions of Ontos [6] and like all models they must contain some reductions of the properties of phenomena under investigation. These reductions made in favour of pragmatic needs constitute the essence of act of modelling. However, if the reductionist operations are not carried out carefully or conducted excessively, some of the important features of the modelled phenomenon can be missed, distorted or misrepresented. Furthermore, in some cases complex or unexpected outcomes are observed as a reflection of notions not inherited in the developed model.

[6] $o\nu\tau\omega\varsigma$ (Greek) : the thing, the reality, the authentic essence of being.

After the Macro-scale Model I has been developed, we have observed such a phenomenon in the form of an oscillatory behaviour in cell energies during population growth which immediately drew our attention since no oscillatory part was intended in the meso-scale dynamics, let alone in the energy dynamics. Likewise, Macro-scale Model has no component which could reflect this oscillatory behaviour in some mathematical sense.

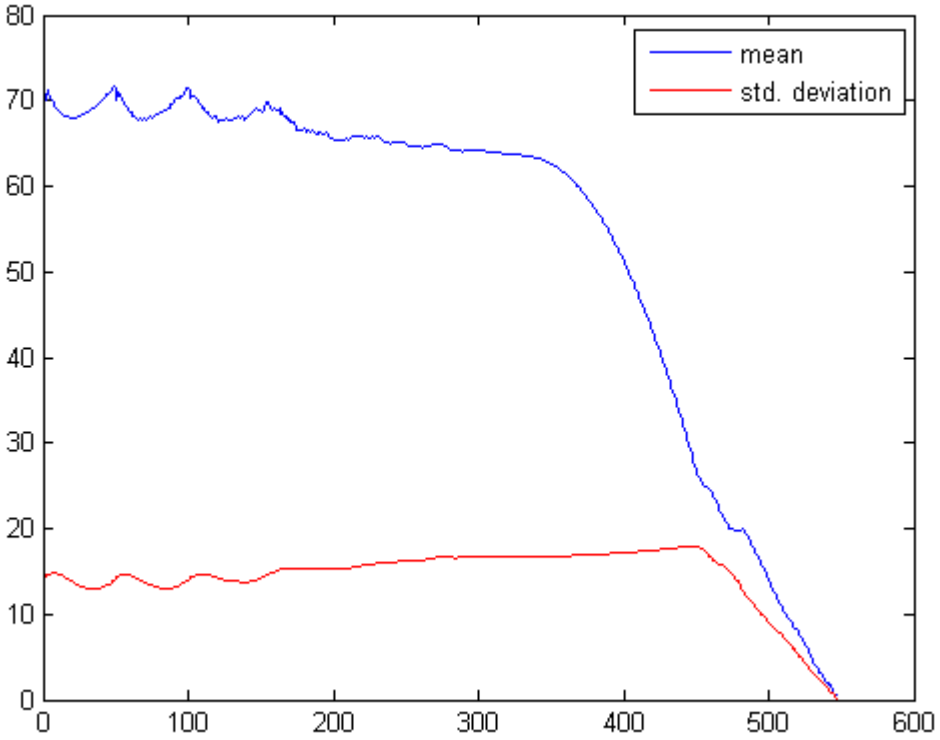

Figure 4.1. Oscillating behaviour in the first and second moments of cell energies.

This oscillating behaviour was indicating us that an essential feature has been missed in the dynamics of cell energies and thus encouraged us to investigate in more detail how the energy distribution in the meso-scale dynamics evolved. As a result of this examination, we realized that the real reason behind this oscillatory behaviour is the shifting and folding effect in the dynamics of the energy distribution as explained in Section 3.3.3.2. The disclosure of the origin of this phenomenon shed light on the development of the Macro-scale Model II and thus this oscillatory behaviour is represented inherently inside the equations of birth rate.

Although the details are beyond the scope of this thesis, it is worth indicating the close relation between the invention of entities (mentioned in previous subsection), which appear virtual from a lower-scale perspective, and the concept of 'emergence as one of the liveliest areas of discussion in philosophy of science [27].

Finally, with the insight gained from this thesis, I want to state that I find mathematical modelling to constitute an invaluable tool in the sense that it allows us to make abstractions and conceptualizations of 'virtual' entities without of which we would not be able to describe the phenomenon of interest. However, at the same time it also encloses the danger of making excessive reductions, resulting in losing of essential attributes of the phenomenon which can have a direct impact on the performance of the model.

In my opinion, this is exactly why a modeller should always keep an eye open to grasp or envisage the notion of possible concepts at any scale, may it have a real correlate or not, while also preserving the right resolution for his perspective to keep the necessary amount of reductions for the development of the model as low as possible.

APPENDIX A: MACRO-SCALE DERIVATION OF TOTAL TOXICITY FROM MESO-SCALE KNOWLEDGE

When we investigate the dynamics of local toxicity, we realize that the total toxicity in flask can be derived analytically from it. Two factors affecting local toxicity change at a specific locus are (i) toxicity release from a cell at that locus - if there is any at all - and (ii) toxicity diffusion into or out of that locus. Hence, its dynamics can be described as follows:

$$x_{i,j}(t+\Delta t) = x_{i,j}(t) + \Delta t \beta A s\left(E_{i,j}(t)\right) + \Delta t D\left(\frac{\partial^2 x_{i,j}(t)}{\partial i^2} + \frac{\partial^2 x_{i,j}(t)}{\partial j^2}\right) \tag{A.1}$$

where $s(x) = \begin{cases} 1 & if\ x > 0 \\ 0 & if\ x \leq 0 \end{cases}$

To obtain total toxicity, Equation A.1 is summed over all loci:

$$\zeta(t) = \sum_{i=1}^{n}\sum_{j=1}^{n} x_{i,j}(t) \tag{A.2}$$

The following series of calculations result in a compact form of dynamics of total toxicity:

$$\zeta(t+\Delta t) = \sum_{i=1}^{n}\sum_{j=1}^{n} x_{i,j}(t+\Delta t) \tag{A.3}$$

$$\sum_{i=1}^{n}\sum_{j=1}^{n} x_{i,j}(t+\Delta t) = \sum_{i=1}^{n}\sum_{j=1}^{n}\left[x_{i,j}(t) + \Delta t \beta A s\left(E_{i,j}(t)\right) + \Delta t D\left(\frac{\partial^2 x_{i,j}(t)}{\partial i^2} + \frac{\partial^2 x_{i,j}(t)}{\partial j^2}\right)\right] \tag{A.4}$$

$$\sum_{i=1}^{n}\sum_{j=1}^{n} x_{i,j}(t+\Delta t) = \sum_{i=1}^{n}\sum_{j=1}^{n} x_{i,j}(t) + \sum_{i=1}^{n}\sum_{j=1}^{n} \Delta t\beta As\left(E_{i,j}(t)\right) + \sum_{i=1}^{n}\sum_{j=1}^{n}\left[\Delta tD\left(\frac{\partial^2 x_{i,j}(t)}{\partial i^2} + \frac{\partial^2 x_{i,j}(t)}{\partial j^2}\right)\right] \quad \text{(A.5)}$$

A close inspection of terms in Equation A.5 reveals that:

$\sum_{i=1}^{n}\sum_{j=1}^{n} x_{i,j}(t) \longrightarrow total\ toxicity = \zeta(t)$

$\sum_{i=1}^{n}\sum_{j=1}^{n} \Delta t\beta As\left(E_{i,j}(t)\right) \longrightarrow fraction\ of\ total\ population = \Delta t\beta Ap(t)$

$\sum_{i=1}^{n}\sum_{j=1}^{n}\left[\Delta tD\left(\frac{\partial^2 x_{i,j}(t)}{\partial i^2} + \frac{\partial^2 x_{i,j}(t)}{\partial j^2}\right)\right] \longrightarrow total\ toxicity\ change\ due\ to\ diffusion = 0$

Hence, Equation A.3 becomes :

$$\zeta(t+\Delta t) = \zeta(t) + \Delta t\beta Ap(t) \quad \text{(A.6)}$$

Rearranging terms and taking limit of time to zero implies:

$$\dot{\zeta(t)} = \beta Ap(t) \quad \text{(A.7)}$$

So, dynamics of total toxicity depends only on total population, which can also be verified by simulations of Meso-scale Model as shown in Figure A.1.

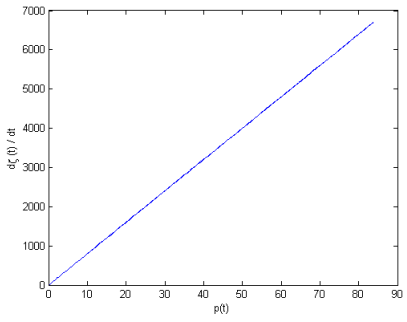

Figure A.1. A linear relationship is observed between total population and derivative of total toxicity. The slope of the line in this figure is equal to βA as in Equation A.7.

APPENDIX B: INSTRUCTIONS FOR THE OPERATOR

The macro-scale models are developed with the intention to estimate the time profile of cell culture populations and eliminate the need to conduct real experiments as far as possible. For this purpose, the macro-scale models have to be initially customized to the specific cell culture at hand. This can be achieved by estimating the macro-scale parameters M, r and $\hat{E}_{div}$ on basis of some simple tests applying the following procedures (for practical purposes it is recommended to use a smaller flask):

(i) Estimation of M:

- Measure the average cell size under the microscope.
- Calculate:

$$M = \frac{flask\ area}{average\ cell\ area}$$

(ii) Estimation of r:

- Conduct an experiment with a relatively small initial population size (10%-20% occupancy is reasonable) and ensure that cells are approximately uniformly distributed.
- Record $p(t)$ every 15 minutes until t^*, when the population stops increasing.
- Assume the constant energy intake $\alpha = 1$.
- Using t^* and $p(t)$ records calculate r:

$$r = \frac{\alpha M}{\int_0^{t^*} p(t)dt}$$

(iii) Estimation $\hat{E}_{div}$:

- Start the experiment with a relatively small initial population size (10%-20% occupancy is reasonable) and ensure that cells are approximately uniformly distributed.
- Clean the flask every 15 minutes, so that no significant toxicity is contained

within the flask.

- Record $p(t)$ every 15 minutes until T_d when the initial population is doubled.
- Assuming $\alpha = 1$, calculate:

$$\hat{E}_{div} = 2\alpha T_d$$

After the parameters are estimated, the differential equations of the Macro-scale Model I and II can be numerically solved the given initial values of the other state variables as specified in Section 3.4. and for the recommended $p(0)$ as specified in Chapter 4. Thus a predicted population profile $p(t)$ is obtained, from which one can extract any relevant information.

Warnings for users of the macro-scale models:

- Cell should not undergo differentiation. If the cells under investigation starts to differentiate at some stage of the experiment, the macro-scale model will be valid only until that stage.
- Cells should be placed within the flask as uniformly as possible. Operations that can cause clustering of cells across the flask should be avoided.

REFERENCES

1. Alberts, B., A. Johnson, J. Lewis, D. Morgan, M. Raff, K. Roberts, and P. Walter, *Molecular Biology of the Cell*, Garland Science, New York, 6th edition, 2014.

2. Meuris, L., F. Santens, G. Elson, N. Festjens, M. Boone, A. D. Santos, S. Devos, F. Rousseau, E. Plets, E. Houthuys, P. Malinge, G. Magistrelli, L. Cons, L. Chatel, B. Devreese, and N. Callewaert, "Glycodelete Engineering of Mammalian Cells Simplifies N-glycosylation of Recombinant Proteins", *Nature Biotechnology*, Vol. 32, No. 5, pp. 485–489, 2014.

3. Çiğdem Aksu, *Modelling of Population Dynamics in Cell Cultures*, M.S. Thesis, Boğaziçi University, 2010.

4. Invitrogen, "Cell Culture Basics", 2014, `www.invitrogen.com/cellculturebasics`, [Accessed: 2014 December].

5. Alberts, B., A. Johnson, J. Lewis, M. Raff, K. Roberts, and P. Walter, *Molecular Biology of the Cell*, Garland Science, New York, 4th edition, 2002.

6. Malthus, T., *An Essay on the Principle of Population*, J. Johnson, London, 1798.

7. Henson, M. A., "Dynamic Modeling of Microbial Cell Populations", *Current Opinion In Biotechnology*, Vol. 14, No. 5, pp. 460–467, 2003.

8. Zygourakis, K., R. Bizios, and P. Markenscoff, "Proliferation of Anchorage-dependent Contact-inhibited Cells: I. Development of Theoretical Models Based on Cellular Automata", *Biotechnology and Bioengineering*, Vol. 38, No. 5, pp. 459–470, 1991.

9. Cheng, G., B. B. Youssef, P. Markenscoff, and K. Zygourakis, "Cell Population Dynamics Modulate the Rates of Tissue Growth Processes", *Biophysical Journal*,

Vol. 90, No. 3, pp. 713 – 724, 2006.

10. Piotrowska, M. J. and S. D. Angus, "A Quantitative Cellular Automaton Model of In Vitro Multicellular Spheroid Tumour Growth", *Journal Of Theoretical Biology*, Vol. 258, No. 2, pp. 165–178, 2009.

11. Lee, Y., S. Kouvroukoglou, L. V. McIntire, and K. Zygourakis, "A Cellular Automaton Model for the Proliferation of Migrating Contact-inhibited Cells.", *Biophysical Journal*, Vol. 69, No. 4, pp. 1284–1298, 1995.

12. Hatzikirou, H. and A. Deutsch, "Cellular Automata as Microscopic Models of Cell Migration in Heterogeneous Environments.", *Current Topics in Developmental Biology*, Vol. 81, pp. 401–34, 2008.

13. Garijo, N., R. Manzano, R. Osta, and M. Perez, "Stochastic Cellular Automata Model of Cell Migration, Proliferation and Differentiation: Validation With in Vitro Cultures of Muscle Satellite Cells", *Journal of Theoretical Biology*, Vol. 314, No. 0, pp. 1 – 9, 2012.

14. Ducrot, A., F. L. Foll, P. Magal, H. Murakawa, J. Pasquier, and G. F. Webb, "An In Vitro Cell Population Dynamics Model Incorporating Cell Size, Quiescence, and Contact Inhibition", *Mathematical Models and Methods in Applied Sciences*, Vol. 21, No. supp01, pp. 871–892, 2011.

15. Armstrong, N. J, Painter, K. J, Sherratt, and J. A, "A Continuum Approach to Modelling Cell-Cell Adhesion", *Journal Of Theoretical Biology*, Vol. 243, No. 1, pp. 98–113, 2006.

16. Yin, K., H. Yang, P. Daoutidis, and G. Yin, "Simulation of Population Dynamics Using Continuous-time Finite-state Markov Chains", *Computers & Chemical Engineering*, Vol. 27, No. 2, pp. 235 – 249, 2003.

17. Kurtanjek, Z., "Simulation of Mammalian Cell Population Dynamics", *26th Inter-*

national Conference on Information Technology Interfaces - ITI 2004, Vol. 1, pp. 35–39, 2004.

18. Lim, Y., "Modeling and Prediction of Cell Population Dynamics", Puigjaner, L. and A. Espua (editors), *European Symposium on Computer-Aided Process Engineering-15, 38th European Symposium of the Working Party on Computer Aided Process Engineering*, Vol. 20 of *Computer Aided Chemical Engineering*, pp. 517 – 522, Elsevier, 2005.

19. Mhaskar, P., M. A. Hjorts, and M. A. Henson, "Cell Population Modeling and Parameter Estimation for Continuous Cultures of Saccharomyces cerevisiae", *Biotechnology Progress*, Vol. 18, No. 5, pp. 1010–1026, 2002.

20. Auger, P., R. B. de la Parra, J. Poggiale, E. Sánchez, and L. Sanz, "Aggregation Methods in Dynamical Systems and Applications in Population and Community Dynamics", *Physics of Life Reviews*, Vol. 5, No. 2, pp. 79 – 105, 2008.

21. Serra, R., M. Villani, and A. Colacci, "Differential Equations and Cellular Automata Models of the Growth of Cell Cultures and Transformation Foci", *Complex Systems*, Vol. 13, No. 4, pp. 347–380, 2001.

22. Chung, C., T.-H. Lin, S.-D. Chen, and H.-I. Huang, "Hybrid Cellular Automaton Modeling of Nutrient Modulated Cell Growth in Tissue Engineering Constructs", *Journal of Theoretical Biology*, Vol. 262, No. 2, pp. 267 – 278, 2010.

23. Cheng, G., P. Markenscoff, and K. Zygourakis, "A 3D Hybrid Model for Tissue Growth: The Interplay between Cell Population and Mass Transport Dynamics", *Biophysical Journal*, Vol. 97, No. 2, pp. 401–414, 2009.

24. Roache, P., *Fundamentals of Computational Fluid Dynamics*, Hermosa Publishers, Albuquerque, New Mexico, 1998.

25. Milo, R., P. Jorgensen, and M. Springer, "Reference Links for Key Numbers in

Biology", 2010, http://bionumbers.hms.harvard.edu/KeyNumbers.aspx, [Accessed: 2014 October].

26. Pletcher, R. H., J. C. Tannehill, and D. Anderson, *Computational Fluid Mechanics and Heat Transfer*, Taylor & Francis, Washington, DC, 2nd edition, 1997.

27. Bedau, M. A. and P. Humphreys, *Emergence: Contemporary Readings in Philosophy and Science*, A Bradford Book, Cambridge, Massachusetts, 1th edition, 2008.