

---

## Development of spatiotemporal validation methods of an agent-based model of epithelial cells: analysis of a colony of keratinocytes

---

César Pichardo-Almarza\*

Unilever Research and Development,  
Colworth Park, Sharnbrook,  
MK44 1LQ, UK  
Fax: (+44) (0) 1234 222161  
E-mail: cesar.pichardo@unilever.com  
\*Corresponding author

Rod Smallwood

Department of Computer Science,  
University of Sheffield,  
Sheffield S1 4DP, UK  
E-mail: r.smallwood@sheffield.ac.uk

S.A. Billings

Department of Automatic Control and Systems Engineering,  
University of Sheffield,  
Sheffield S1 4DP, UK  
E-mail: s.billings@sheffield.ac.uk

**Abstract:** In this paper an Agent-Based Model (ABM) to study Normal Human Keratinocytes (NHK) (tissue cells) is investigated. ABM are widely used for the simulation of systems from several domains (biology, economics, meteorology, etc.). In biology, ABM are useful for predicting the social behaviour of systems; in particular, they seem well adapted to model the behaviour of a cell population. These models exhibit probabilistic behaviour and the validation of simulation results is often a qualitative analysis by experts (biologists). The aim of this paper is to propose new variables and metrics for the development of quantitative methods to validate the model.

**Keywords:** spatiotemporal analysis; epithelial cells; keratinocytes; ABM; agent-based models; validation; Monte-Carlo simulations.

**Reference** to this paper should be made as follows: Pichardo-Almarza, C., Smallwood, R. and Billings, S.A. (2008) 'Development of spatiotemporal validation methods of an agent-based model of epithelial cells: analysis of a colony of keratinocytes', *Int. J. Functional Informatics and Personalised Medicine*, Vol. 1, No. 4, pp.335–354.

**Biographical notes:** César Pichardo-Almarza is a research scientist at Unilever R&D Colworth, currently working in Systems Biology. He received his PhD in Automatic Control and Industrial Informatics from the 'Ecole Centrale de Lille'

(France), a MSc in Systems Engineering from Simon Bolivar University (Venezuela) and a MEng in Chemical Engineering from the same University. His research areas of interests include dynamical modelling and simulation, state and parameter estimation, system identification and other areas related to the application of control theory to study biological and physiological systems using different modelling approaches.

Rod Smallwood is a Professor of Computational Biology at the Department of Computer Science of the University of Sheffield. He received his PhD in Medical Physics from the University of Sheffield, a MSc in Solid State Physics from the University of Lancaster and a BSc in Physics from UCL (London). His current research focuses on an integrated approach to several key areas of computational systems biology including the development of agent-based models for the simulation of complex structure of biological tissues, cell populations, intracellular signal transduction and the development of a common computational framework for the simulation of agent-based models of different applications.

Stephen A. Billings was appointed as Professor in the Department of Automatic Control and Systems Engineering, University of Sheffield, UK in 1990 and leads the Signal Processing and Complex Systems research group. His research interests include system identification and information processing for nonlinear systems, narmax methods, model validation, prediction, spectral analysis, adaptive systems, nonlinear systems analysis and design, neural networks, wavelets, fractals, machine vision, cellular automata, spatio-temporal systems, fMRI and optical imagery of the brain, metabolic systems engineering, systems biology and related fields.

---

## 1 Introduction

A computational model has been developed in the Epitheliome project (<http://www.epitheliome.com>) based on biological rules that govern the self-organisation of Normal Human Keratinocytes (NHK) (Sun et al., 2007). This is the result of the combination of *in-vitro* and *in-vituo* models used to explore the behaviour of NHK. The model helps to predict the dynamic multi-cellular morphogenesis of NHK and of a keratinocyte cell line (HaCat cells) under varying extra-cellular Ca<sup>++</sup> concentrations. The use of this approach to model NHK and HaCat provides a description of the effects of extra-cellular calcium on NHK proliferation and differentiation and provides a simulation of the colony formation of keratinocytes. The aim of this study was to further develop the model to explore how NHKs self-organise into an epithelium, particularly how they form colonies.

To date, the quantitative validation of this *in-vituo* model has been confined to a comparison of real and simulated growth rates (using the total number of cells in the dish). Cell division and migration are stochastic processes, so the dynamical response of the *in-vitro* and *in-vituo* models would not be expected to yield identical outcomes after a period of growth under identical conditions. Nevertheless, it is reasonable to assume strong similarity and suitable metrics to compare *in-vitro* and *in-vituo* results need to be developed. The main objective of this work is to contribute to the development of these metrics. The aim is to define new variables with spatio-temporal information of the cell population in order to make a quantitative comparison of the

dynamical results of the models. Using a new approach based on a circular grid, this spatio-temporal analysis will allow the measurement (and comparison) of new variables associated with cell division, migration and apoptosis giving dynamical information about the growth rates in different regions of the dish.

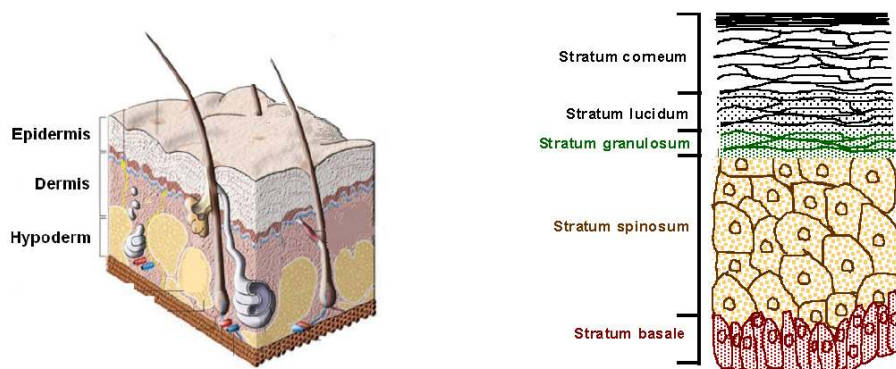
This paper is organised as follows: Section 2 gives a description of the agent-based model to be analysed. Section 3 and Section 4 introduce the circular approach for the spatiotemporal analysis and the new variables to be used. Section 5 shows an example where the circular approach is used to compare two simulations. Section 6 shows the application of Monte-Carlo Simulations to evaluate the dynamical responses of the new variables and, finally, Section 7 shows the conclusions of this work.

## 2 Description of the model

Normal Human Keratinocytes constitute over 80% of the cells in the inter-follicular epidermis. These cells are shed at the skin surface and are replaced by the division in the basal layers of the epidermis, known as the germinative compartment (Webb et al., 2004). Cells in the basal layer are heterogeneous in type and have a hierarchical population structure (Webb et al., 2004). They are pushed up through the layers of the epidermis, undergoing gradual differentiation until they reach the stratum corneum where they form a layer of enucleated, flattened, highly keratinised cells called squamous cells. This layer forms an effective barrier to the entry of foreign matter and infectious agents into the body and minimises moisture loss. Figure 1 shows different layers in human epidermis.

The model enables *in-virtuo* exploration of the relative importance of biological rules and was used to test hypotheses *in-virtuo* which were subsequently examined *in-vitro*.

**Figure 1** Human epidermis (see online version for colours)

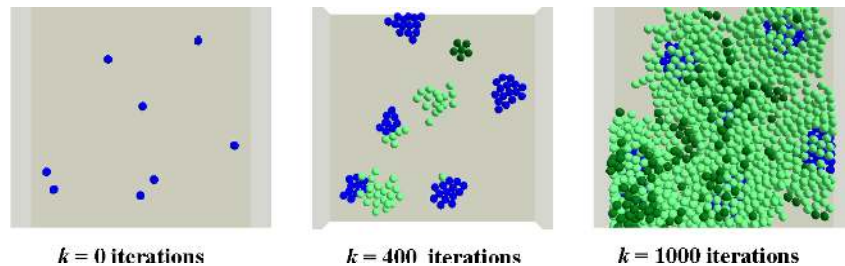


Each individual cell is represented in the computational model by a software agent. In this particular model, the software agents are a form of communicating X-machine (Balancescu et al., 1999; Kefalas et al., 2003). Each agent has a rule set, based on experimental cell biology, which determines the behaviour of the agent and its interaction with its neighbours. The interaction of a set of agents (which are equivalent to a cell population) can be used to model the organisation of multi-cellular aggregates (Walker et al., 2004; Grabe and Neuber, 2005; Sun et al., 2007).

Each cell is modelled as a non-deformable sphere of 20  $\mu\text{m}$  in diameter. Cells are capable of migration, proliferation and differentiation. The culture is modelled as a flat, square surface with ‘walls’ and the dimensions are user-defined. For the purpose of the experiments, the basement was 500  $\mu\text{m}$  in each dimension with a wall height of 100  $\mu\text{m}$ . The exogenous calcium level was set to 1.3 mM (physiological calcium level) but the model can be used for different calcium levels (i.e., 0.1 mM for the study of the system with a low calcium level).

Results obtained from the model include the type of cell (stem, Transit-Amplifying (TA), committed or corneocyte) and their location (Figure 2 shows some snapshots of simulations). Each cell then performs specific rules associated with the cell cycle. Following this, cells decide whether to change to another cell type based on the differentiation rules in the model. Cells then execute their migration rules, and finally execute physical rules. All rules are executed in the context of the agent’s own internal state and the states of the other cells around it.

**Figure 2** Snapshots of simulations of the agent-based model (see online version for colours)



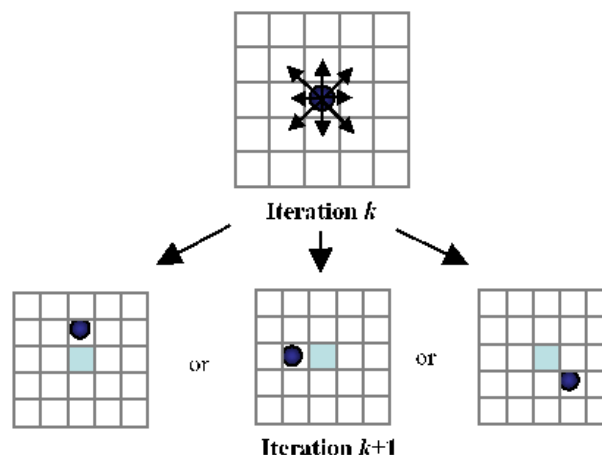
The computational model allows the user to access several variables associated with each cell in the dish (position, type, number of intercellular bonds, etc.) for each iteration  $k$ . Mean cell cycle time and migration rate are scaled so that each time step in the model represents approximately 30 min in real time. For the purpose of this paper, the dynamical response of the spatial location of each cell was the main variable.

### 3 Spatiotemporal analysis

#### 3.1 Why a spatiotemporal analysis and the development of validation methods?

A great amount of data can be extracted from individual agents in the simulations. Some global variables (i.e., the total number of cells in the dish, the total number of cell divisions, etc.) are accessible to the user for comparison with experimental data. However, there are some random processes associated with each individual generating a spatial probabilistic behaviour in the entire cell population.

Classical quantitative analysis could generate different spatial results for similar behaviours. For example, when a simulation is performed, the first observations that are available to the user are that individual cells can move in any direction from time  $k$  to time  $k + 1$  (Figure 3).

**Figure 3** Possible scenarios for iteration  $k + 1$  (see online version for colours)

This kind of reaction generates a stochastic behaviour for the entire cell population.

The analysis of the development of spatial distribution in stochastic processes where both type and position of the objects is involved has not been extensively developed in the literature. This kind of study is an important and necessary procedure for the validation of these models. The spatiotemporal analysis is particularly relevant to the understanding of the development of properly structured and functional tissue in multi-cellular organisms.

Figure 3 shows that the same migration behaviour could be considered as different for several experiments if a classic spatial analysis (i.e., discretising the dish with a 2D square grid) is used to study the spatial distribution of the cells. A cell can migrate to any case around its original location; thus, if a square grid is used, an occupied case in a given experiment could be empty in another experiment.

The aim of this work is to propose the calculation of new variables in order to develop methods to quantify similar spatio-temporal behaviours. In particular, these new variables will be useful to determine if two simulations are 'equal' or to verify if the model is 'right' (i.e., the dynamical response of the model is similar to the experimental data).

The ultimate goal of this approach is the development of validation methods to compare the simulation results with experimental data, possibly developing some approaches using correlation function tests.

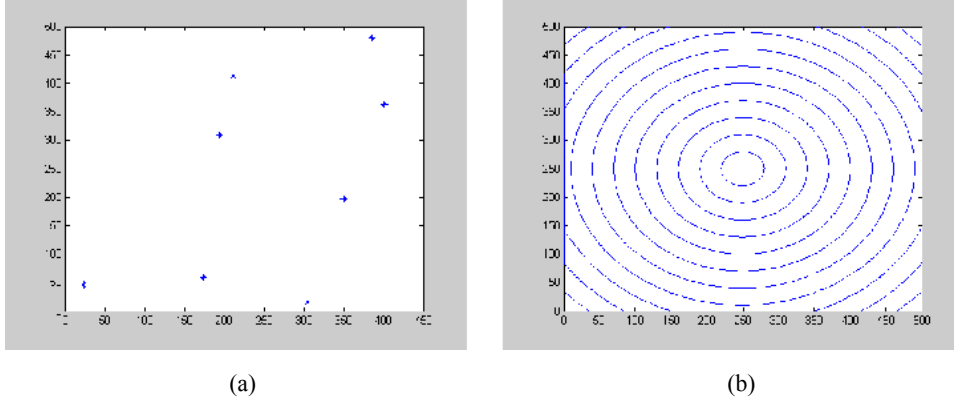
### 3.2 Selection of a grid for a quantitative analysis

As noted above, a square grid could generate different results for the same spatio-temporal behaviour.

Because of this effect, a circular approach, based on concentric circles around the centre of the dish ( $x = 250$ ,  $y = 250$ ) will be introduced. In this particular analysis, 12 circles with a radius variation of  $30 \mu\text{m}$  were employed (Figure 4).

Once the grid is defined, the objective is to calculate several variables associated with the spatial behaviour of the cell population. The calculation of these variables is made in each ring of the circular grid.

**Figure 4** (a) Initial location of cells and (b) circular grid (see online version for colours)



#### 4 Defining quantitative variables for comparison

The first variable to be calculated is the number of cells in each ring of the grid. This will give information about the spatial distribution of the cells. Two variables were selected to evaluate the migration of cells: positive migration and negative migration. Positive migration is the number of cells entering the ring and negative migration represents the number of cells leaving the ring. The mitosis rate is calculated as the number of cells in the ring because of the division of cells. Finally, apoptosis is calculated as the number of cells dying in the ring at each iteration. These effects will be discussed in more detail in the following sections.

##### 4.1 Comparing the spatial distribution of cells

The spatial distribution can be evaluated as the variation of the total number of cells in each ring. With this objective in mind a vector with  $n$  rows ( $n$  = number of radii used in the circular grid) is built for each iteration  $k$ :

$$NCell_k = \begin{bmatrix} R_1 [NCell_k(1)] \\ \vdots \\ R_n [NCell_k(n)] \end{bmatrix} \quad (1)$$

##### 4.2 Comparing migration

The criterion to evaluate the migration of cells involves observations of the number of cells that move from ring ‘ $j$ ’ to ring ‘ $i$ ’. This behaviour can be achieved by filling a square matrix where each row and column corresponds to each ring.

$$\text{Mig}_k = \begin{matrix} & R_1 & & & R_n \\ R_1 & \left[ \begin{array}{ccc} \text{Mig}_k(1,1) & \cdots & \text{Mig}_k(1,n) \\ \vdots & & \\ R_n & \text{Mig}_k(n,1) & \text{Mig}_k(n,n) \end{array} \right] & & & \end{matrix} \quad (2)$$

The element  $(i, j)$  of the  $\text{Mig}_k$  matrix for each iteration corresponds to the number of cells that migrate from ring ' $i$ ' to ring ' $j$ '. Obviously the element  $(i, i)$  is the number of cells that remain in the ring ' $i$ '. In general, there will be a transfer of cells from/to ring ' $i$ ' to/from ring ' $i + 1$ ' and ' $i - 1$ ', which means that a tri-diagonal matrix is obtained at each iteration.

To calculate the total number of cells at time ' $k$ ' in ring ' $i$ ', a simple equation based on migration (from and to ring ' $i$ ') and mitosis can be used.

$$C_k^i = C_{k-1}^i + \text{Min}_{k-1}^i - \text{Mout}_{k-1}^i + \text{Mit}_{k-1}^i \quad (3)$$

where  $C_k^i$  is the total number of cells in ring ' $i$ ' at time  $k$ ,  $C_{k-1}^i$  is the total number of cells in ring ' $i$ ' at time  $k - 1$ ,  $\text{Min}_{k-1}^i$  is the number of cells that migrated from other rings to ring ' $i$ ',  $\text{Mout}_{k-1}^i$  is the number of cells that migrated from ring ' $i$ ' to other rings and  $\text{Mit}_{k-1}^i$  is the number of new cells in ring ' $i$ ' because of mitosis.

To compare migration the terms  $\text{Min}_k^i$  and  $\text{Mout}_k^i$  for each ring ' $i$ ' can be evaluated. These terms are useful for calculating the 'migratory' flow of cells; the objective is to evaluate the variation of the number of cells in ring ' $i$ ' ( $C_k^i$ ) because of  $\text{Min}_k^i$  and  $\text{Mout}_k^i$ . To illustrate this effect assume that the total number of cells in ring ' $i$ ' varies only because of the migration of cells from other rings,

$$C_k^i = C_{k-1}^i + \text{Min}_{k-1}^i. \quad (4)$$

Alternatively the total number of cells varies only because of the migration of cells from ring ' $i$ ' to other rings,

$$C_k^i = C_{k-1}^i - \text{Mout}_{k-1}^i. \quad (5)$$

Equations (4) and (5) can be used to compare the migration in each ring. The *positive migration* in ring ' $i$ ' ( $\text{PM}_k^i$ ) can then be defined as:

$$\text{PM}_k^i = \text{PM}_{k-1}^i + \text{Min}_{k-1}^i \quad (6)$$

and the *negative migration*  $\text{NM}_{k-1}^i$  as:

$$\text{NM}_k^i = \text{NM}_{k-1}^i - \text{Mout}_{k-1}^i \quad (7)$$

The term  $\text{Min}_{k-1}^i$  corresponds to the sum of the elements  $(i - 1, i)$  and  $(i + 1, i)$  of matrix  $\text{Mig}_{k-1}^i$  equation (8) and  $\text{Mout}_{k-1}^i$  corresponds to the sum of elements  $(i, i + 1)$  and  $(i, i - 1)$  of matrix  $\text{Mig}_{k-1}^i$  equation (9).

$$Min_{k-1}^i = Mig_k(i+1, i) + Mig_k(i-1, i) \quad (8)$$

$$Mout_{k-1}^i = Mig_k(i, i+1) + Mig_k(i, i-1). \quad (9)$$

### 4.3 Comparing mitosis

The same circular grid can also be used to evaluate the mitosis of cells in the dish (Figure 3(b)). Equation (3) can be used to calculate the term associated with mitosis ( $Mit_{k-1}^i$ ):

$$Mit_{k-1}^i = C_k^i - C_{k-1}^i - Min_{k-1}^i + Mout_{k-1}^i. \quad (10)$$

This will yield a vector for each iteration; thus, if there are 'n' radii,  $Mit_k$  will have n components:

$$Mit_k = \begin{bmatrix} M_k^1 \\ \vdots \\ M_k^n \end{bmatrix}. \quad (11)$$

Using a similar reasoning, a method to compare mitosis in each ring can be developed. This can be used to evaluate how the number of cells in ring 'i' ( $C_k^i$ ) increase because of mitosis occurring at time 'k'. Assuming that the total number of cells in ring 'i' varies only because of mitosis in this ring, gives

$$C_k^i = C_{k-1}^i + Mit_{k-1}^i. \quad (12)$$

Defining  $RMit_{k-1}^i$  as the *Mitosis rate* in ring 'i',

$$RMit_k^i = RMit_{k-1}^i + Mit_{k-1}^i \quad (13)$$

produces a variable which can be used to compare mitosis in different simulations.

### 4.4 Comparing apoptosis

The circular grid in this case is used to analyse apoptosis. The main idea is to evaluate the rate of cells dying in each ring of the grid. The calculation of the apoptosis rate uses a vector for each iteration,  $Apop_k$ , with the number of cells dying in each ring:

$$Apop_k = \begin{bmatrix} A_k^1 \\ \vdots \\ A_k^n \end{bmatrix}. \quad (14)$$

Using this vector a method to compare apoptosis in each ring can be developed. This can be used to evaluate how the number of dead cells in ring 'i' increases at time 'k'.

Defining  $RApop_{k-1}^i$  as the *Apoptosis rate* in ring 'i',

$$RApop_k^i = RApop_{k-1}^i + Apop_{k-1}^i \quad (15)$$

produces a variable which can be used to compare apoptosis in different simulations.



#### 4.5 Using statistics to compare results

As a first approach, the main objective of this paper is to propose accessible techniques to allow a quantitative comparison between the results from two (or several) simulations of the agent-based model. The purpose is to try to extend these methods for the validation of results of the agent-based model with experimental results obtained from the in vitro model.

Results from two simulations using the same parameters and initial conditions can be compared using classical statistical techniques, i.e., linear regressions, etc. Future work could include the use of specialised tools for the statistical analysis of circular data (Fisher, 1993; Jammalamadaka and SenGupta, 2001).

With the variables proposed in this paper a dynamical evolution in time is obtained for the different radii of the grid; thus, classical statistics can be used to compare this dynamical response from two simulations. The method to calculate the variables and a statistical analysis to compare simulation results are illustrated in the following example.

### 5 Example

In this example a comparison between the dynamical responses of two simulations is shown. In this case both simulations use the same initial positions of cells and parameters (calcium levels, etc.). Simulations are generated using physiological calcium levels (1.3 mM) and eight initial cells in the dish with different spatial locations (Figure 3(a)).

The spatial distribution, migration, mitosis and apoptosis are compared using the techniques proposed in sections above.

#### 5.1 Spatial distribution

The vectors associated with the spatial distribution of cells from these two simulations ( $N_{Cell}$ ) will be obtained. The main idea is to find this vector for each iteration by observing if the number of cells in each ring is similar in both simulations. For example, when the  $N_{Cell}$  vector is calculated for iteration 1000, the following results are obtained (Table 1):

**Table 1** Number of cells at iteration 1000 for both simulations

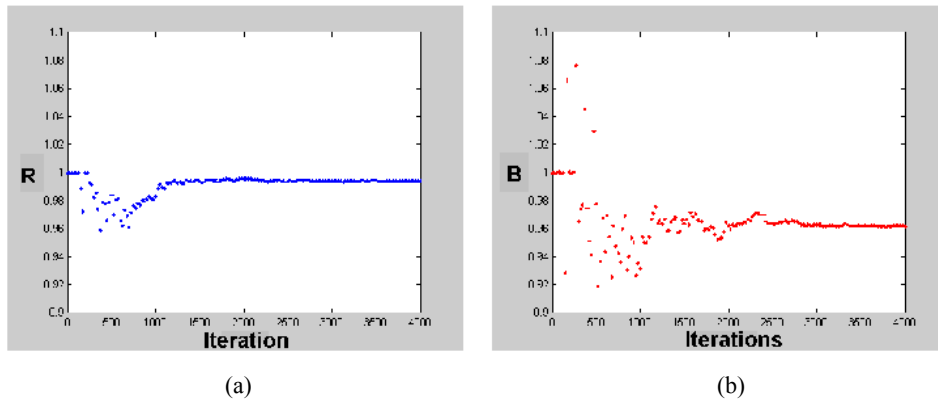
<i>Simulation 1</i>	<i>Simulation 2</i>
$N_{Cell}_{1000} = \begin{bmatrix} 11 \\ 40 \\ 57 \\ 74 \\ 93 \\ 110 \\ 137 \\ 144 \\ 122 \\ 59 \\ 24 \\ 7 \end{bmatrix}$	$N_{Cell}_{1000} = \begin{bmatrix} 8 \\ 25 \\ 38 \\ 64 \\ 101 \\ 123 \\ 133 \\ 147 \\ 123 \\ 50 \\ 20 \\ 7 \end{bmatrix}$

To compare these vectors a linear regression can be applied, observing the number of cells in each ring.

A regression factor ( $R$ ) equal to 0.98 and a constant of the linear model ( $B$ ;  $Y = BX$ ) equal to 0.94 are obtained for this iteration.

Calculating the  $NCell$  vector for each iteration and the linear correlation between the elements of both simulations, the obtained values of  $R$  and  $B$  are very close to 1. Figure 5 shows the temporal evolution of  $R$  and  $B$ .

**Figure 5** Spatial distribution: (a) correlation factor,  $R$  and (b) linear coefficient,  $B$  (see online version for colours)



### 5.2 Migration

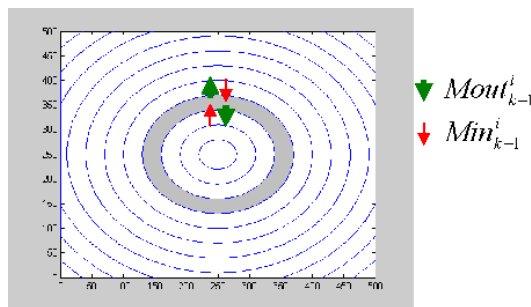
The migration of cells is evaluated using the circular grid of Figure 6. Calculating the migration matrix at iteration 1000, the results shown in Table 2 are obtained.

As mentioned above, the elements on the diagonal of matrices  $Mig_k$  are the number of cells that remain (from iteration  $k - 1$ ) in each ring. Thus, the evaluation of the positive migration is made by taking the elements of  $Mig_k$  in position  $(i + 1, i)$  and  $(i - 1, i)$  and the evaluation of the negative migration is made by taking the elements of  $Mig_k$  in position  $(i, i - 1)$  and  $(i, i + 1)$ .

$$Min_{k-1}^i = Mig_k(i + 1, i) + Mig_k(i - 1, i) \tag{8}$$

$$Mout_{k-1}^i = Mig_k(i, i + 1) + Mig_k(i, i - 1). \tag{9}$$

**Figure 6** Circular grid: evaluating migration and mitosis (see online version for colours)

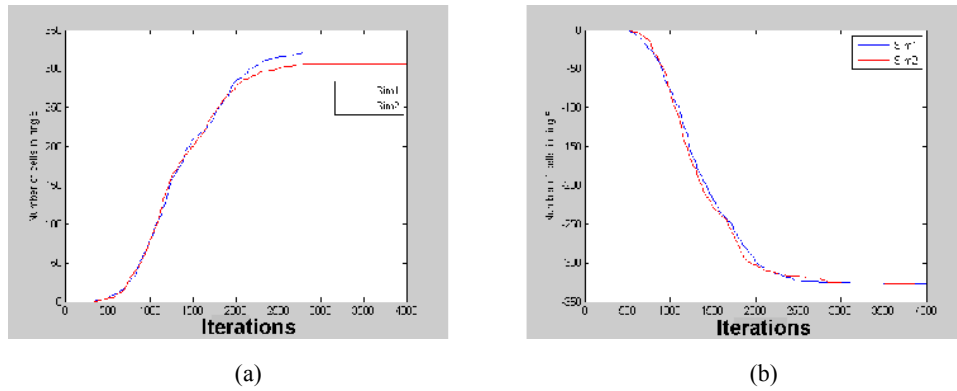


**Table 2** Migration matrices at iteration 1000 for both simulations

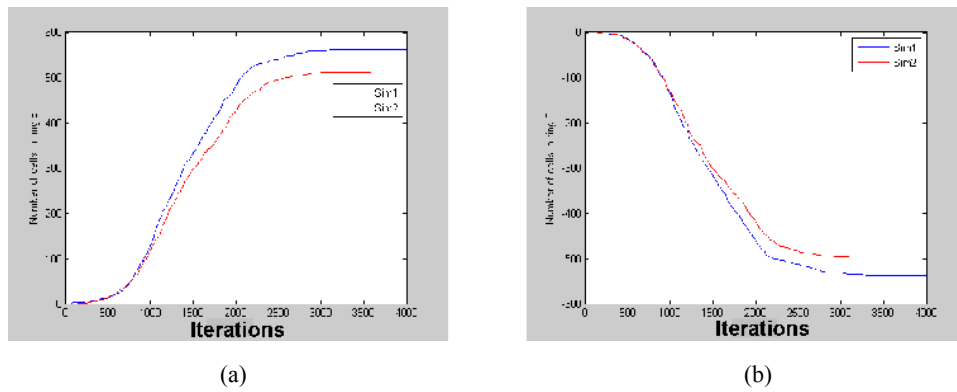
	Simulation 1	Simulation 2
$Mig_{1000} =$	$\begin{bmatrix} 10 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 25 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 40 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 57 & 6 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 83 & 4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 4 & 108 & 8 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 5 & 135 & 4 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 6 & 132 & 6 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 & 119 & 3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 5 & 48 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 23 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 6 \end{bmatrix}$	$\begin{bmatrix} 10 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 31 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 42 & 6 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 59 & 3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 4 & 83 & 10 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3 & 102 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 5 & 126 & 3 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 3 & 131 & 3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 & 111 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 35 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 17 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix}$

For example, drawing the temporal evolution of the positive and the negative migration for rings 5 and 8, Figures 7 and 8 are obtained respectively.

**Figure 7** (a) Positive migration for ring 5 and (b) negative migration for ring 5 (see online version for colours)



**Figure 8** (a) Positive migration for ring 8 and (b) negative migration for ring 8 (see online version for colours)



Figures 7 and 8 show that the dynamical responses of the positive and negative migrations for both simulations are very similar. The same procedure can be applied for all rings and a similar behaviour will be obtained.

Tables 3 and 4 show the results of  $R$  and  $B$  of a linear regression between the results of both simulations:

**Table 3** Linear regression for positive migration

<i>Ring</i>	<i>Correlation factor, R</i>	<i>Linear factor, B (Y = BX)</i>
1	0.9897	0.7092
2	0.9977	1.0022
3	0.9963	1.1100
4	0.9961	1.0967
5	0.9961	1.0371
6	0.9998	0.9932
7	0.9988	1.0091
8	0.9995	1.1061
9	0.9990	1.2252
10	0.9974	1.2326
11	0.9984	1.0757
12	0.9867	0.7637

**Table 4** Linear regression for negative migration

<i>Ring</i>	<i>Correlation factor, R</i>	<i>Linear factor, B (Y = BX)</i>
1	0.9880	0.6248
2	0.9980	0.9864
3	0.9938	1.0393
4	0.9972	1.1836
5	0.9990	0.9935
6	0.9995	0.9687
7	0.9994	1.0292
8	0.9997	1.0768
9	0.9996	1.2635
10	0.9994	1.1855
11	0.9966	1.1348
12	0.9935	1.5598

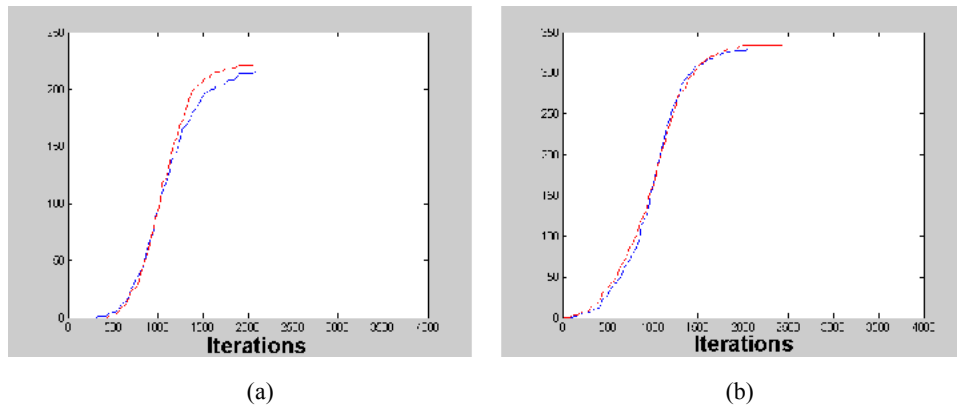
### 5.3 Mitosis

In this case, the evaluation of mitosis in each ring is made using the circular grid of Figure 4(b) and equation (13).

$$RMit_{k-1}^i = C_{k-1}^i + Mit_{k-1}^i. \quad (13)$$

With this equation the mitosis rate for both simulations is obtained. Figure 9 shows an example of the evolution of mitosis in time for rings 5 and 8. Similar results are obtained for the other radii. Table 5 shows the values of the linear correlation ( $R$  and  $B$ ) between the migration results of both simulations.

**Figure 9** (a) Mitosis rate in ring 5 and (b) mitosis rate in ring 8 (see online version for colours)

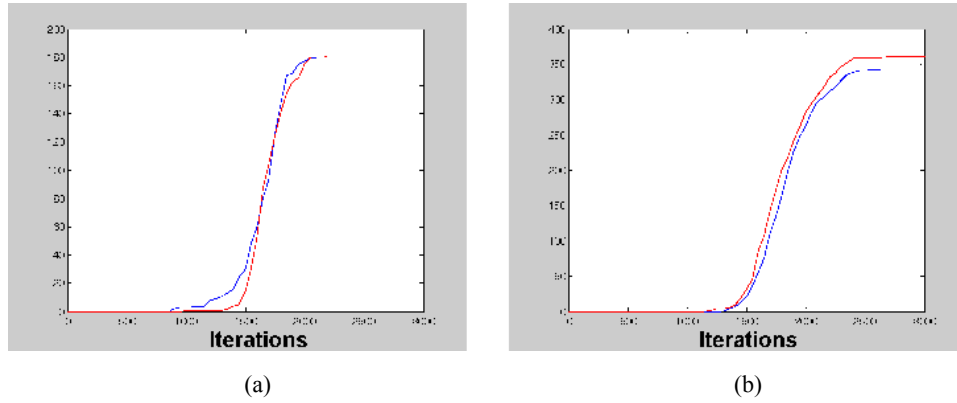


**Table 5** Linear regression for mitosis

Ring	Correlation factor, $R$	Linear factor, $B$ ( $Y = BX$ )
1	0.9915	0.8282
2	0.9969	0.92969
3	0.9984	0.8656
4	0.9991	1.1706
5	0.9992	0.96496
6	0.9997	1.0133
7	0.9998	1.06
8	0.98894	0.99933
9	0.99962	1.1045
10	0.9971	1.0419
11	0.99441	1.1697
12	0.92399	3.0375

#### 5.4 Apoptosis

Finally, the apoptosis in each ring of the grid is evaluated by calculating the apoptosis rate (Section 2.4). Figure 10 shows the dynamical response of the apoptosis rate in rings 5 and 8 of the grid. Table 6 shows the values of the linear correlation between both simulations.

**Figure 10** (a) Apoptosis rate in ring 5 and (b) apoptosis rate in ring 8 (see online version for colours)**Table 6** Linear regression for apoptosis

Ring	Correlation factor, $R$	Linear factor, $B$ ( $Y = BX$ )
1	0.9386	0.9019
2	0.9699	0.9146
3	0.9793	1.0380
4	0.9933	1.1037
5	0.9981	0.9912
6	0.9995	0.9348
7	0.9972	0.9959
8	0.9988	1.0594
9	0.9975	1.0960
10	0.9995	1.1324
11	0.9926	1.0159
12	0.8682	1.6757

## 6 Monte-Carlo simulations

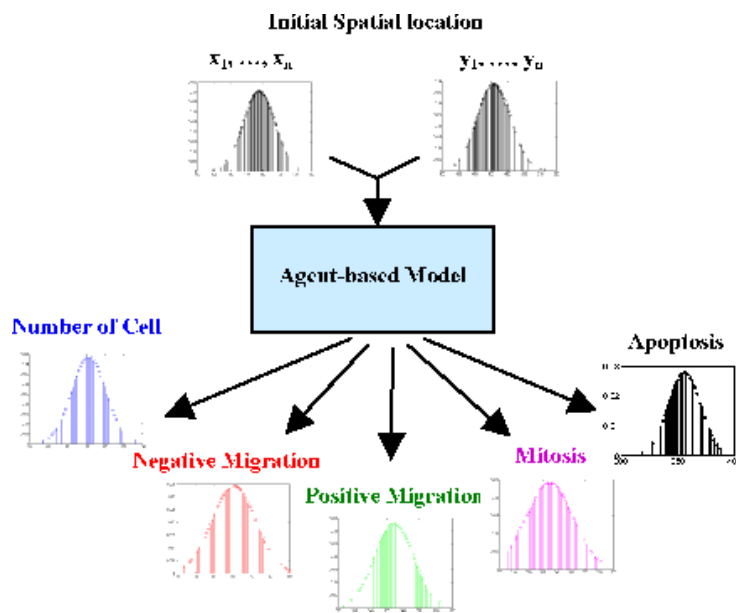
Because of the probabilistic behaviour of the agent-based model used to simulate the behaviour of NHK, there is a possibility of obtaining different responses for two simulations (using the same set of parameters). A good method which allows the evaluation of the behaviour of a random process is Monte-Carlo Simulations.

MonteCarlo (MC) methods are statistical simulation methods, using random numbers to give approximate solutions to mathematical problems. Developed by Neumann (Neumann, 1966), Ulam and Metropolis (Metropolis, 1987; Metropolis and Ulam, 1949) during World War Two, this method has been used to model a large variety of problems, from the estimation of pi (Mathews, 1972) to immune system simulations (Dasgupta, 1992; Mannion, 2002).

In the particular case of this work, MC simulations are used to observe the behaviour of the agent-based model. The procedure consisted of running 100 simulations of the

model. For each simulation a set of random positions  $(x, y)$  with a normal distribution was generated. After each simulation a dynamic response of the variables defined to evaluate the spatial behaviour of cells (number of cells, positive and negative migration and mitosis) was obtained. Finally the results of the 100 simulation were analysed using statistical tools (Figure 11 shows a block diagram for the MC simulations). Because of the complexity of the model of the entire population and all the calculations related to individual cells, big computational power was needed for this analysis: a cluster of 160 processors (2.4 GHz AMD Opteron) (<http://www.shef.ac.uk/wrgrid/iceberg>) was used to obtain simulation results in a reasonable time.

**Figure 11** Block diagram for the MC simulations (see online version for colours)



### 6.1 Analysing the Monte-Carlo simulation results

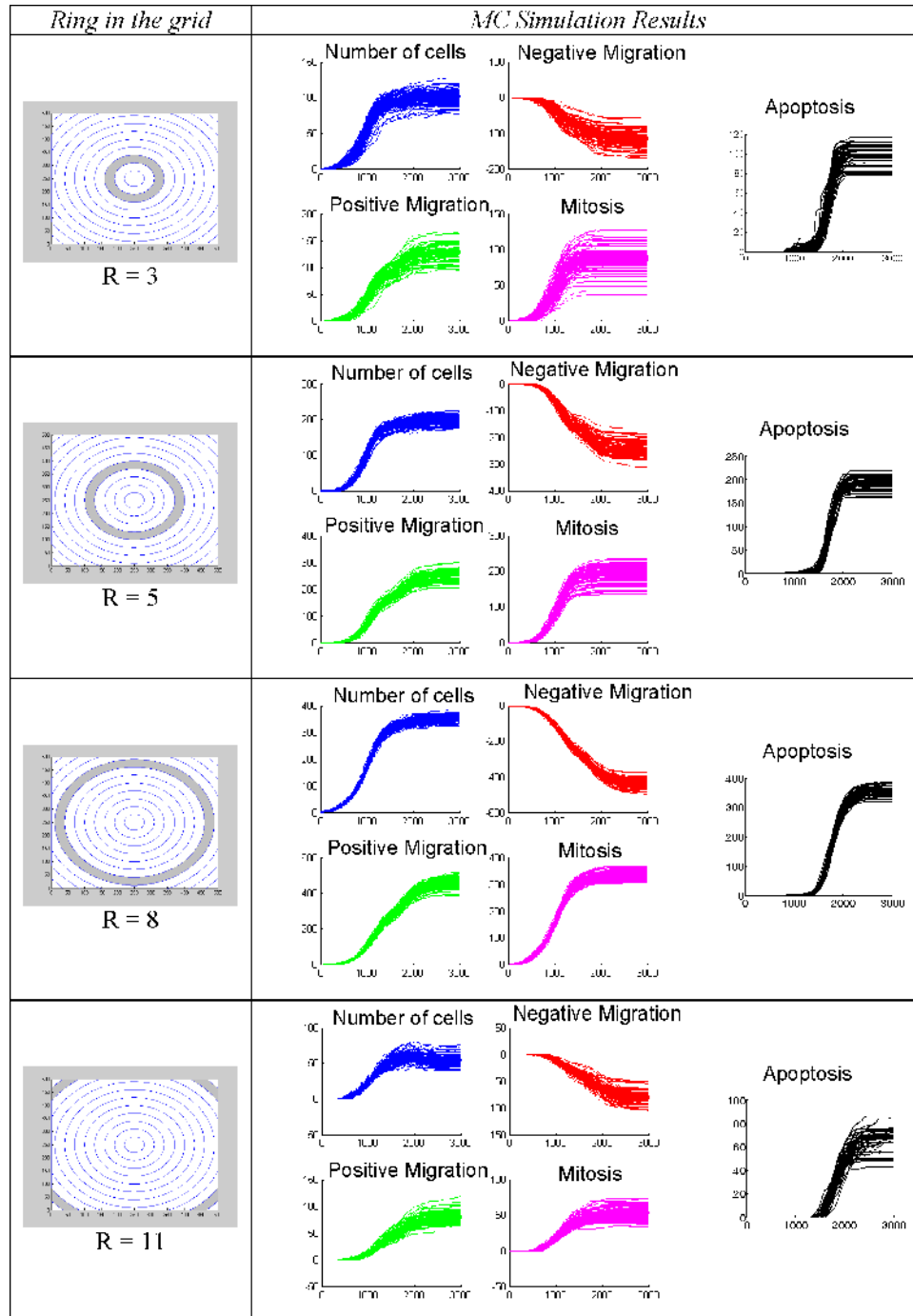
Classical statistics can be used to interpret the results of the Monte-Carlo simulations. The main idea is to calculate some statistical variables for the evaluation and testing of the model. In this paper, classical tools have been selected to evaluate the dynamic simulation results. The main objective of this statistical analysis is to determine whether similar behaviours are obtained for several simulations under similar conditions and parameter values.

In this case, the mean value, standard deviation and confidence intervals were calculated for each iteration. After these calculations a 'mean' dynamical response is obtained for each variable with additional information about the temporal evolution of the confidence intervals.

Additionally to evaluate the probability of obtaining similar final values of the simulations a probability density function (pdf) was calculated at each radius.

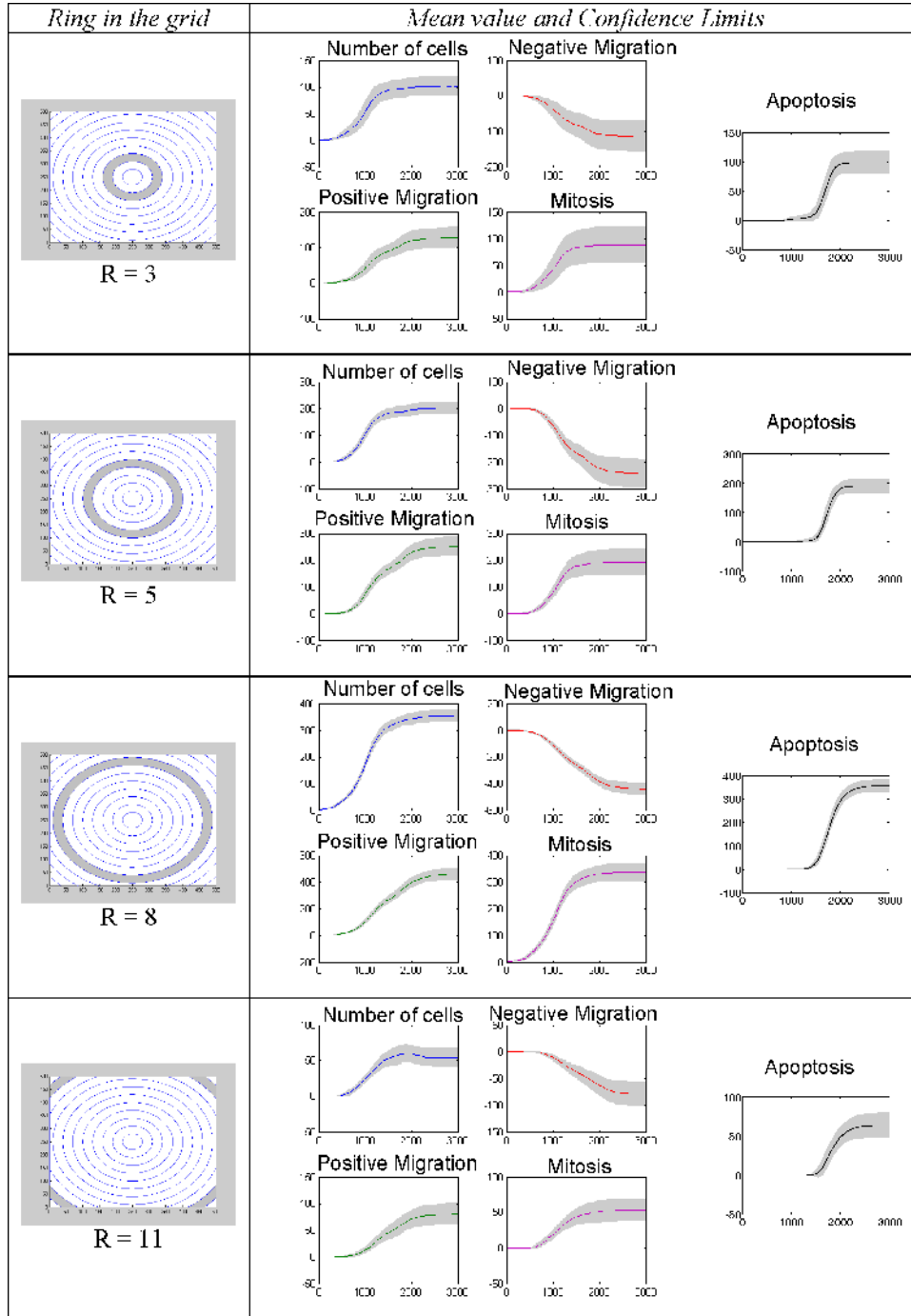
Figures 12 and 13 show the simulation results, the mean value and the confidence intervals for several radii in the grid. Figure 14 shows the normal pdf for the final values of number of cells, negative migration, positive migration and mitosis.

**Figure 12** MC Simulation results (see online version for colours)

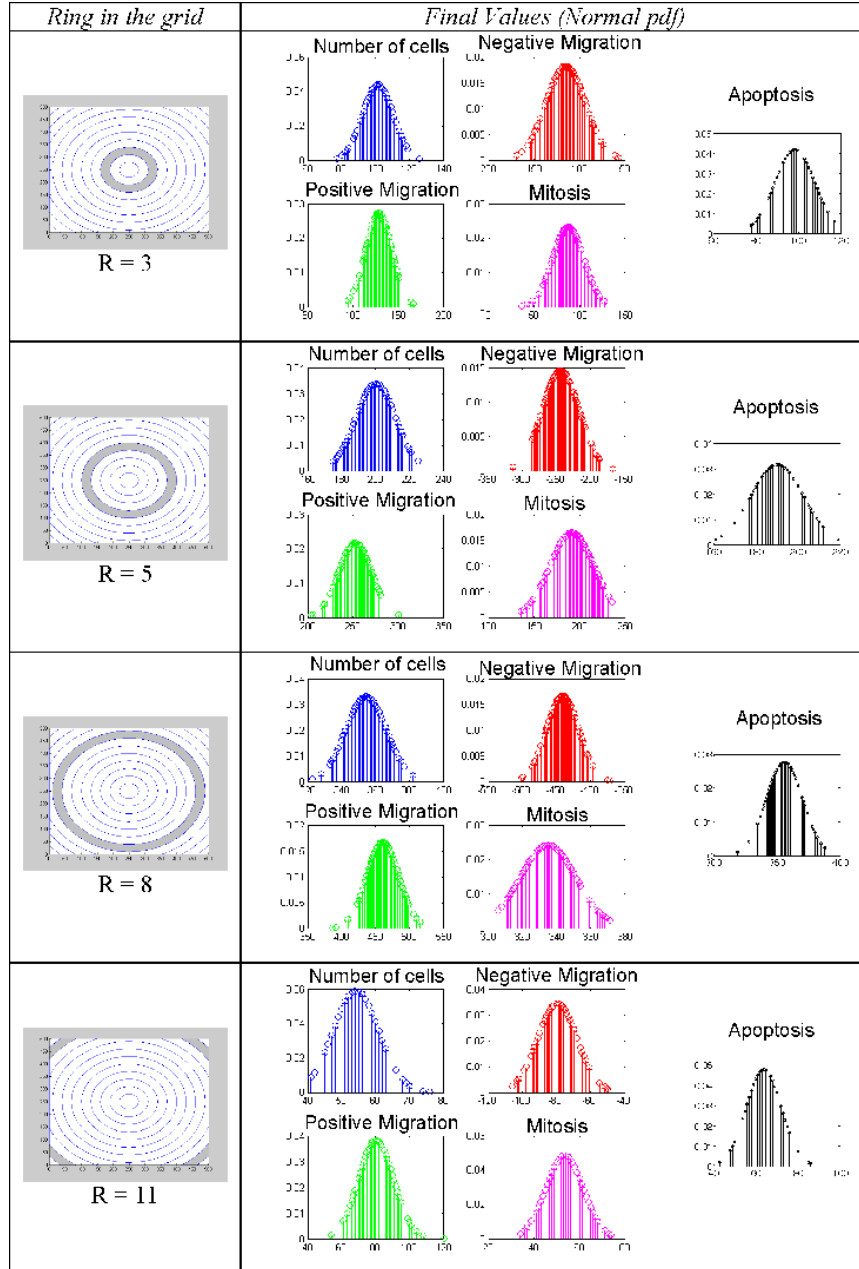




**Figure 13** Mean values and confidence intervals obtained for the MC simulations (see online version for colours)



**Figure 14** Normal pdf of the final values of the MC simulations (see online version for colours)



These results show similar behaviours for several simulations using similar conditions. The confidence intervals were consistent with the dynamical behaviour of the cell population; a more random behaviour is obtained in small areas (i.e., the centre and the corner of the dish), thus, larger confidence intervals are obtained for these radii. However, the obtained normal pdfs indicate that the results from simulations had the highest probability to be near the mean values.

## 7 Conclusions

This work is an extension of a first approach to the spatio-temporal analysis of an agent-based model of NHK (Pichardo-Almarza et al., 2007). This analysis represents an important contribution for the study of the spatial distribution in stochastic processes where both type and position of the objects are involved.

This kind of study is an important and necessary procedure for the validation of these models. A good method for the spatio-temporal analysis is particularly relevant to the understanding of the development of properly structured and functional tissue in multi-cellular organisms.

The new variables defined in this paper allow more spatio-temporal information to be obtained from the computational model of NHK; these variables represent useful measurements in quantitative comparisons.

The methods developed, allow a quantitative comparison of two simulations with respect to the spatial distribution of cells, how the cells move in the dish (migration) and the reproduction of cells (mitosis) using classical statistics methods.

The spatial distribution has been compared by building a vector for each iteration with the number of cells in each ring of the grid. Thus, a linear regression can be calculated and a temporal evolution of the correlation factors can be obtained.

Migration has been compared by observing how cells move from one ring to another. In this case, the 'positive' and the 'negative' migration can be compared in each ring for the total simulation time. Mitosis is compared using a method very similar to the migration method and simulations can be compared using a 'Mitosis rate'. Finally, apoptosis is compared evaluating the number of cells dying in each iteration.

In this work the behaviour of cells was analysed using the same calcium level (physiological calcium level). However, the calcium level is a parameter which can have a large influence on the behaviour of the cell population (Walker et al., 2006). Future research will include similar spatiotemporal analysis concerning the variation of this parameter and the influence on the dynamical response of the agent-based model.

Algorithms are being formulated in Matlab (<http://mathworks.com>) for application in the analysis and comparison of data from different models and/or different experiments. The results obtained from this work are encouraging and it is expected that they will be useful in comparing experimental results from an accurate cell tracking system, which is currently under development as part of the Epitheliome project (<http://www.epitheliome.com>).

## Acknowledgements

The authors gratefully acknowledge support from the Engineering Physical Sciences Research Council (EPSRC) and the Royal Academy of Engineering for Dr. Pichardo-Almarza for this project. Also, the authors have to express a special acknowledgement to Dr. Sun Tao and Dr. Phil McMinn for providing support with the computational and *in-vitro* models of NHK.

## References

- Balanescu, T., Cowling, A.J., Georgescu, H., Gheorghe, M., Holcombe, M. and Vertan, C. (1999) 'Communicating stream X-machines systems are no more than X-machines', *J. Univ. Comput. Sci.*, Vol. 5, pp.494–507.
- Dasgupta, S. (1992) 'Monte-Carlo simulation of the shape space model of immunology', *Physica A: Stat. Theor. Phys.*, Vol. 189, Nos. 3–4, pp.403–410.
- Fisher, N.I. (1993) *Statistical Analysis of Circular Data*, Cambridge University Press, Cambridge.
- Grabe, N. and Neuber, K. (2005) 'A multicellular systems biology model predicts epidermal morphology, kinetics and  $Ca^{2+}$  flow', *Bioinformatics*, Vol. 21, pp.3541–3547.
- Jammalamadaka, S.R. and SenGupta, A. (2001) *Topics in Circular Statistics*, World Scientific Publishing, Singapore.
- Kefalas, P., Holcombe, M., Eleftherakis, G. and Gheorghe, M. (2003) 'A formal method for the development of agent-based systems', in Plekhavona, V. (Ed.): *Intelligent Agent Software Engineering*, Idea Group Publishing, Hershey, PA, pp.68–98.
- Mannion, R., Ruskin, H.J. and Pandey, R.B. (2002) 'Effects of viral mutation on cellular dynamics in a Monte Carlo simulation of HIV immune response model in three dimensions', *Theory Biosci.*, Vol. 121, No. 2, pp.237–245.
- Mathews, J.H. (1972) 'Monte Carlo estimate for pi', *Pi Mu. Epsilon J.*, Vol. 5, pp.281–282.
- Metropolis, N. and Ulam, S. (1949) 'The Monte Carlo method', *J. Am. Stat. Assoc.*, Vol. 44, pp.335–341.
- Metropolis, N. (1987) 'The beginning of the Monte Carlo method', *Los Alamos Sci.*, Vol. 15, p.125.
- Neumann, J.V. (1966) *Theory of Self-Reproducing Automata*, in Burks, A.W. (Ed.), University of Illinois Press, Champaign, IL.
- Pichardo-Almarza, C., Smallwood, R. and Billings, S.A. (2007) 'Spatiotemporal analysis of an agent-based model of a colony of keratinocytes: a first approach for the development of validation methods', *Proceedings of the 7th IEEE International Conference on Bioinformatics and Bioengineering*, Vol. II, pp.1173–1177.
- Sun, T., McMinn, P., Coakley, S., Holcombe, M., Smallwood, R. and MacNeil, S. (2007) 'An integrated systems biology approach to understanding the rules of keratinocyte colony formation', *Journal of the Royal Society Interface*, Vol. 4, pp.1077–1092.
- Walker, D.C., Southgate, J., Hill, G., Holcombe, M., Hose, D.R., Wood, S.M., MacNeil, S. and Smallwood, R.H. (2004) 'The epitheliome: agent-based modelling of the social behaviour of cells', *BioSystems*, Vol. 76, pp.89–100.
- Walker, D., Sun, T., Macneil, S. and Smallwood, R. (2006) 'Modelling the effect of exogenous calcium on keratinocyte and hacaat cell proliferation and differentiation using an agent-based computational paradigm', *Tissue Eng.*, Vol. 12, No. 8, pp.2301–2309.
- Webb, A., Li, A. and Kaur, P. (2004) 'Location and phenotype of human adult keratinocyte stem cells of the skin', *Differentiation*, Vol. 72, pp.387–395.

## Websites

<http://www.shef.ac.uk/wrgrid/iceberg>  
 The Epitheliome Project: <http://www.epitheliome.com>  
 The Mathworks, Inc., <http://www.mathworks.com>