DISCRETE AND CONTINUOUS
DYNAMICAL SYSTEMS
Volume **34**, Number **12**, December **2014**

doi:10.3934/dcds.2014.34.5123

pp. **5123–5133**

# MODELLING COLLECTIVE CELL BEHAVIOUR

Deborah C. Markham, Ruth E. Baker and Philip K. Maini

Wolfson Centre for Mathematical Biology
Mathematical Institute, University of Oxford
Andrew Wiles Building, Radcliffe Observatory Quarter, Woodstock Road
Oxford, OX2 6GG, United Kingdom

Abstract. The classical mean-field approach to modelling biological systems makes a number of simplifying assumptions which typically lead to coupled systems of reaction-diffusion partial differential equations. While these models have been very useful in allowing us to gain important insights into the behaviour of many biological systems, recent experimental advances in our ability to track and quantify cell behaviour now allow us to build more realistic models which relax some of the assumptions previously made. This brief review aims to illustrate the type of models obtained using this approach.

1. **Introduction.** Collective movement of individuals occurs in many areas in ecology, for example, fish schooling, locust swarming, birds flocking. Cell biology is no different. For example, in early embryonic development cells move in sheets or migrate long distances individually. In wound healing, cell movement is essential to close the wound, while in cancer cell movement leads to often fatal secondary tumours (metatases). In this paper we briefly review some recent results on developing models for collective cell motion and cell-cell interactions. In Section 2 we outline the classical approach to modelling cell invasion where biological assumptions on behaviour are incorporated at the macroscopic, partial differential equation level. We then advance the modelling approach in Section 3 where the macroscopic model equations are derived by making assumptions on cell behaviour at the local level. In this case, we focus on the modifications to the diffusion term. In Section 4 we consider a revamp of the classical approach where we show that many of the implicit assumptions in the derivation of mean-field macroscopic models can lead to results which are inconsistent with the behaviour of the system purported to be modelled. We conclude in Section 5 by commenting on the exciting challenges that face theoreticians as the result of the recent technological advances in experimental biology.

2. **Classical modelling approach.** The classical modelling approach to cell movement and interaction is to assume that cells interact as if in a 'well-stirred' system, so that their kinetics can be modelled using the law of mass action, and their motion can be modelled as simple Fickian diffusion. However, cells have volume so this needs to be taken into account and there are a number of ways to do this. As one example, we take the work of Gatenby and Gawlinski [12], who proposed a novel

2010 *Mathematics Subject Classification.* 35K57, 92B99.

*Key words and phrases.* Discrete models, continuum models, volume exclusion, moment dynamics, cell biology.

model for tumour cell invasion. It is known that, in the absence of sufficient oxygen, cells will use glycolytic metabolism. However, tumour cells will sometimes utilise glycolytic metabolism even in the presence of oxygen (the Warburg effect [38]). This seems somewhat of a mystery as the glycolytic metabolic pathway is not as efficient at producing energy as the normal, aerobic respiration pathway. Obviously, this puts the tumour cells at a disadvantage. Gatenby and Gawlinski proposed that tumour cells actually gain a competitive advantage because a byproduct of glycolysis, lactic acid, is more toxic to normal cells than to cancer cells. They proposed a model which consisted of three coupled partial differential equations of reaction-diffusion type and showed that it exhibited travelling waves of invasion which, in certain parameter regimes, predicted an acellular gap between the invading tumour cell front and the regressing normal cells. They validated this model prediction in the laboratory. The model behaviour was mathematically analysed in [11] where it was shown that the system exhibited both slow and fast waves.

More recently, the paper by McGillen *et al.* [26] considered an extended version of the above model in which it was proposed that the tumour cells were not totally resistant to acid (a more realistic description). The model, in one spatial dimension for simplicity, takes the form

$$\frac{\partial U}{\partial t} = \rho_1 U\left(1-\frac{U}{\kappa_1}-\alpha_2\frac{V}{\kappa_1}\right)-\delta_1 UW, \tag{1}$$

$$\frac{\partial V}{\partial t} = \rho_2 V\left(1-\frac{V}{\kappa_2}-\alpha_1\frac{U}{\kappa_2}\right)-\delta_2 VW+\frac{\partial}{\partial x}\left[D_2\left(1-\frac{U}{\kappa_1}\right)\frac{\partial V}{\partial x}\right], \tag{2}$$

$$\frac{\partial W}{\partial t} = \rho_3 V-\delta_3 W+D_3\frac{\partial^2 W}{\partial x^2}, \tag{3}$$

where $U(x,t), V(x,t)$ and $W(x,t)$ are, respectively, normal tissue density, cancer cell density, and excess acid concentration at position $x$ and time $t$. All the parameters are non-negative and constant.

The interactions encoded in these equations are essentially based on the law of mass action. For example, equation (1) assumes that normal cells grow logistically in the absence of tumour cells, compete with the tumour cells and die due to excess acid. It is known that normal cells typically do not move in the adult and therefore there is no transport term in this equation. Equation (2) models growth of tumour cells in a way similar to that of normal cells, but takes into account the fact that tumour cells are invasive by including a diffusion term with a diffusion coefficent that depends on normal cell density. This assumes that tumour cells only move when the normal cell density falls below confluence. Equation (3) models excess lactic acid concentration as being produced linearly by tumour cells, degraded linearly and diffusing in the standard Fickian way.

This model recaptures the key results of the original paper (Figure 1), namely that travelling waves exist and that an acellular gap is possible. However, the modelling approach used is phenomenological, with the nonlinear diffusion term proposed based purely on the modelling assumption that tumour cell movement should be opposed by normal cells. It does not try to derive the form of the diffusion term from a detailed study of individual cell behaviour. In the next section, we consider some models in which the transport term at the cell density level is derived by hypothesising the local cell behaviour and then taking the appropriate limits.

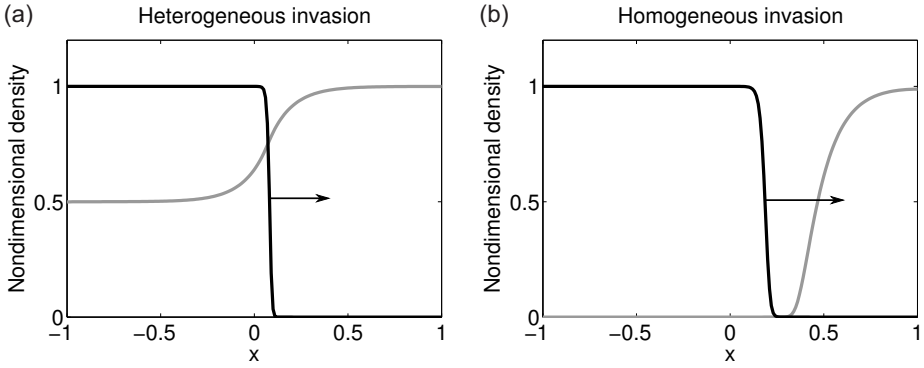


FIGURE 1. Numerical solution of model system (1)-(3) for an initially compact tumour invading the host tissue. Travelling wave solutions: (a) co-existence, (b) competitive exclusion. Note that in (b) the model predicts an acellular gap between the two cell types. Dark lines represent tumour cell density, light lines normal cell density. See [26] for full details.

3. **From the microscopic to the macroscopic.** With the recent advances in biotechnology, it is now possible to observe cell behaviours at the single cell level and therefore we need techniques for developing macroscopic (tissue level) models, such as the one in the previous section, from local, microscopic, cell-cell interactions. The seminal paper of Othmer and Stevens [31] is one of the first to do this using the idea of reinforced random walks. Using a master equation approach, they derived a series of equations based on different transition probabilities (describing how cell movement is affected by interactions at an individual level) and, on taking the appropriate limit, obtained the corresponding partial differential equation behaviour, where the local dynamics are encoded in the functional forms of the transport equations.

These ideas have been extended in [32], which considers a number of approaches for deriving models for chemosensitive cell movement which take into account the finite size of individuals ("volume filling") and cell density (quorum) sensing mechanisms. Further biologically motivated extensions include cell-to-cell adhesion and it can be shown, in this case, that the formal continuum limit of a discrete one-dimensional model of cell motility is a nonlinear diffusion equation in which the diffusivity can undergo a sign change [1].

3.1. **Off-lattice approaches.** More recently, Dyson *et al.* [10] used a master equation approach to consider a simple off-lattice individual-based model with volume exclusion. Taking the parabolic limit of the resulting conservation equation they showed how, in the macroscopic limit, a position jump process with volume exclusion could be written as a diffusion equation with a nonlinear diffusion coefficient which depended on the details of the jump probability density function. For example, if individuals moved with jump rate $\alpha$ a constant distance $d$ to the left or right, then the diffusion equation, under certain simplifying assumptions, takes the form

$$\frac{\partial n}{\partial t} = D\frac{\partial}{\partial x}\left[\left(1+4R\frac{N-1}{N}n\right)\frac{\partial n}{\partial x}\right], \tag{4}$$

where $n(x,t)$ is the average total cell density at position $x$ and time $t$, $N$ is the total cell number, $R$ is the cell radius (assumed the same for all cells) and $D = \lim_{\alpha\to 0} \alpha d^2/2$. Note that, in this case, crowding actually enhances diffusion. However, we should be careful in using this equation to model high density cell populations since as $N$ increases the limiting procedure becomes invalid and in fact we reach the jamming limit, where crowding inhibits movement.

Alternatively, if it is assumed that the distance moved is normally distributed with zero mean and variance $\sigma^2$, then the diffusion equation, under the same simplifying assumptions, takes the form

$$\frac{\partial n}{\partial t} = \hat{D}\frac{\partial}{\partial x}\left[\left(1 + \frac{N-1}{N}\left(4R - 2\sigma\sqrt{\frac{2}{\pi}}\right)n\right)\frac{\partial n}{\partial x}\right], \tag{5}$$

where $\hat{D} = \lim_{\alpha\to 0} \alpha\sigma^2/2$.

3.2. **Dynamics of cell sheets.** In certain situations, coarse graining can also be carried out for the case where cells are actually moving coherently in a sheet, as is the case for epithelia. This system can be modelled in a number of ways. For example, assuming that the cells are point masses connected together by springs [27]. In the simple case of a row of cells connected by linear springs, it can be shown that, in the overdamped case, the continuum limit of the discrete model is a nonlinear diffusion equation of the form

$$\frac{\partial n}{\partial t} = \frac{\partial}{\partial x}\left[\frac{k}{\eta n^2}\frac{\partial n}{\partial x}\right], \tag{6}$$

where $n(x,t)$ is the normalised cell density at position $x$ and time $t$, $k$ is the spring constant and $\eta$ the cell viscosity [28]. The method used in [28] can be extended to the case of nonlinear springs [29].

An alternative approach assumes that cells are composed of elements which move to minimise a Hamiltonian representing an energy (the cellular Potts model [13]). Assuming that cells interact via a hard-core potential it can be shown that in the macroscopic limit this, too, reduces to a nonlinear diffusion equation but with the form

$$\frac{\partial n}{\partial t} = \frac{\partial}{\partial x}\left[C\frac{n_0^2 + n^2}{(n_o - n)^2}\frac{\partial n}{\partial x}\right], \tag{7}$$

where $n_o^{-1}$ is the average cell length and $C$ is a constant [21].

The above examples illustrate some possible forms of nonlinear diffusion coefficients that can arise from making biologically realistic hypotheses on how cells interact with each other locally. We now turn our attention to address the validity of the widely used mean-field assumption that individual cell positions are independent of each other.

4. **Correlations models.** Mean-field models generally assume that the positions of cells are independent of one another. However, this is frequently not the case in biology, where we often find high levels of cell clustering. There is the concern that using a mean-field assumption (MFA) to derive population-level partial differential equation models could lead to inaccurate results, thus necessitating the development of methods incorporating higher levels of information. This can be done by including information on the positions of pairs of cells; an approach which has been used in many areas of biology [7, 8, 15, 16, 19, 30, 34, 35], most recently in cellular

processes [3, 24, 25, 36, 37]. In this section, we discuss these models in the context of cell biology, highlighting insights from the work that has been done thus far.

We begin with the model developed by Baker and Simpson [3]. They consider a one-species volume-excluding system whereby agents undergo proliferation, migration and death on a regular $d$-dimensional lattice, beginning from uniform initial conditions. We regard sites as being occupied (state $A$) or unoccupied (state 0), with the lattice state of site $l$ being described by the variable $\sigma_l$. Adding greater degrees of spatial information into a population model can be done using $k$-point distribution functions, $\rho^{(k)}$, which describe the occupancies of $k$-tuplets of lattice sites. If lattice site occupancies are independent we expect the $k$-th distribution function to be the product of $k$ 1-point distribution functions. For example, if the lattice sites are independent, then we have

$$\rho^{(2)}(A_l, A_m) = \rho^{(1)}(A_l)\rho^{(1)}(A_m). \tag{8}$$

To relate the 2-point and 1-point distribution functions we can use correlation functions [3, 22, 23], defined as

$$F_{\lambda,\mu}(|\mathbf{l}-\mathbf{m}|) \coloneqq \frac{\rho^{(2)}(\sigma_l, \sigma_m)}{\rho^{(1)}(\sigma_l)\rho^{(1)}(\sigma_m)}, \tag{9}$$

where $\lambda$ and $\mu$ denote the states of sites $\mathbf{l}$ and $\mathbf{m}$, respectively. We note that if lattice sites are independent, the correlation function is equal to unity. To obtain the relevant equations, we derive the master equations for the 1- and 2-point distribution functions, respectively, in the following way:

$$\begin{aligned}\frac{d\rho^{(1)}(A_l)}{dt} =& \underbrace{P_m \sum_n \frac{\alpha_{n,l}}{z}[\rho^{(2)}(0_l, A_n) - \rho^{(2)}(A_l, 0_n)]}_{\text{Movement to/from neighbouring sites}} \\ &+ \underbrace{P_p \sum_n \frac{\alpha_{n,l}}{z}\rho^{(2)}(0_l, A_n)}_{\text{Proliferation from neighbouring sites}} - \underbrace{P_d\rho^{(1)}(A_l)}_{\text{Cell death}},\end{aligned} \tag{10}$$

and

$$\begin{aligned}\frac{d\rho^{(2)}(A_l, A_m)}{dt} =& \underbrace{P_m \sum_{n\neq l} \frac{\alpha_{n,m}}{z}[\rho^{(3)}(A_l, 0_m, A_n) - \rho^{(3)}(A_l, A_m, 0_n)]}_{\text{Movement in and out of site } m} \\ &+ \underbrace{P_m \sum_{n\neq m} \frac{\alpha_{n,l}}{z}[\rho^{(3)}(A_l, 0_m, A_n) - \rho^{(3)}(A_l, A_m, 0_n)]}_{\text{Movement in and out of site } l} \\ &+ \underbrace{P_p \left[\sum_{n\neq l} \frac{\alpha_{n,m}}{z}\rho^{(3)}(A_l, 0_m, A_n) + \sum_{n\neq m} \frac{\alpha_{n,l}}{z}\rho^{(3)}(0_l, A_m, A_n)\right]}_{\text{Proliferation from site } n} \\ &+ \underbrace{P_p \frac{\alpha_{l,m}}{z}\left[\rho^{(2)}(A_l, 0_m) + \rho^{(2)}(0_l, A_m)\right]}_{\text{Proliferation when } l \text{ and } m \text{ are nearest neighbours}} - \underbrace{2P_d\rho^{(2)}(A_l, A_m)}_{\text{death}},\end{aligned} \tag{11}$$

where $z = 2d$ ($d$ is the number of spatial dimensions of the system), and $\alpha_{n,l}$ is unity if $n$ and $l$ are nearest neighbours, otherwise it is zero. Depending on the system in

question, we could find ourselves including $k$-point functions up to the total number of lattice sites. Of course, this is not a desirable outcome as the system would rapidly become numerically intractable. Most systems choose to close at $k = 3$, expressing $\rho^{(3)}$ terms using $\rho^{(1)}$ and $\rho^{(2)}$, although some systems do go to fourth order [14]. There are a number of closures in the literature [8, 17, 18, 33], with the Kirkwood Superposition Approximation (KSA) [17, 18] being one of the most widely used. Using the KSA and simplifying, the above equations can be reduced to a system of ODEs, one for the cell density, and the rest for the system of correlation functions:

$$\frac{\mathrm{d}c_A}{\mathrm{d}t} = P_p c_A[1 - F_{A,A}(\Delta)c_A] - P_d c_A \tag{12}$$

and

$$\begin{aligned}\frac{dF_{AA}}{dt}(|\mathbf{l}-\mathbf{m}|) = {} & \frac{2P_m}{z}\sum_{n\neq l}\alpha_{n,m}[F_{AA}(|\mathbf{l}-\mathbf{n}|) - F_{AA}(|\mathbf{l}-\mathbf{m}|)] \\ & + \frac{2P_p}{1-c_A}[1 - c_A F_{AA}(|\mathbf{l}-\mathbf{m}|)]\,[1 - c_A F_{AA}(\Delta)]\left[\sum_{n\neq l}\frac{\alpha_{n,m}}{z}F_{AA}(|\mathbf{l}-\mathbf{n}|)\right] \\ & + \frac{2P_p}{c_A}\frac{\alpha_{l,m}}{z}(1 - c_A F_{AA}(\Delta)) - 2P_p(1 - F_{AA}(\Delta)c_A)F(|\mathbf{l}-\mathbf{m}|),\end{aligned} \tag{13}$$

where $\Delta$ is the lattice spacing, which is usually taken to be (at least) the diameter of a cell.

The predicted density from our correlations equations can be examined against the averaged discrete results to determine the accuracy of the model. Additionally, it can be compared with the MFA (the logistic equation in this case) to see if it provides an improvement. In their paper, Baker and Simpson show the greatly improved agreement with the averaged discrete results in both 2D and 3D. The improvement is particularly noticeable when agent death is included, as a different steady-state is observed (see Figure 2). The discrepancies observed in the mean-field case are of particular importance when a mean-field model is used for parameter estimations in a situation where we cannot assume the system is well-mixed. In [37], the authors estimate the rates of movement and proliferation for two different cell types. For one of the cell types (3T3 fibroblast cells), the MFA and correlations model provide similar estimates for the parameters in question. However, for the other cell type (MDA MB231 breast cancer cells), significantly different estimates are given by the two different modelling approaches. This is a cause for concern if we plan to use these estimated parameters for other modelling studies, as they could lead to inaccurate predictions of the system's behaviour.

The work discussed so far has demonstrated the importance of including spatial correlations when modelling biological cells. However, there are some drawbacks to this method arising from the non-uniform system of ODEs in 2D and 3D. This makes it challenging to implement in practice, especially if the correlations act at long-ranges. These issues can be circumvented, as was shown by Markham *et al.* [24], by describing the system of ODEs for the correlation functions by a PDE with associated boundary conditions. The PDE is obtained using series expansions,

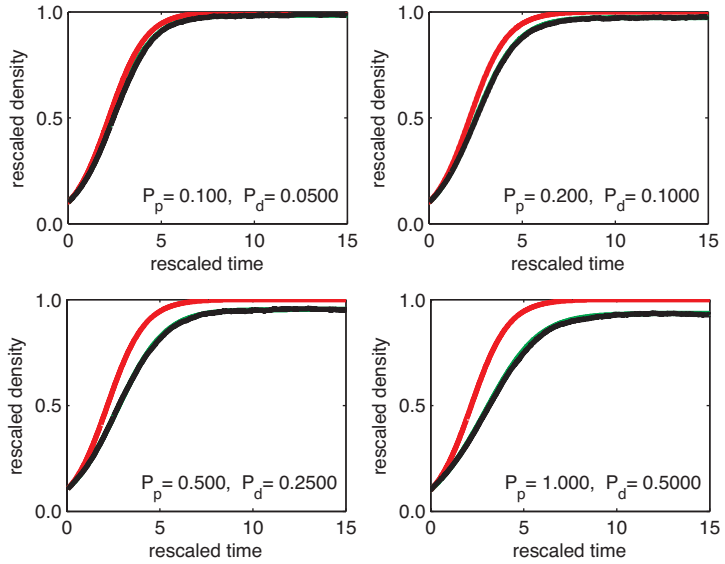


FIGURE 2. The correlations model (green/light grey) provides a good approximation to the averaged discrete results (black), in contrast to the MFA (red/dark grey). More details can be found in [3].

and has the following form in two-dimensions:

$$\frac{\partial F(s)}{\partial t} = \underbrace{\frac{P_m}{2}\left[\frac{\partial^2 F(s)}{\partial s^2} + \frac{1}{s}\frac{\partial F(s)}{\partial s}\right]}_{\text{diffusion term}} + \underbrace{2P_p c_A F(s)\left[1 - c_A F(\Delta)\right]\left[\frac{1-F(s)}{1-c_A}\right]}_{\text{reaction term}}, \quad s > \Delta, \tag{14}$$

with boundary conditions

$$\begin{aligned}\frac{\mathrm{d}F(\Delta)}{\mathrm{d}t} = {} & \frac{P_m}{2}\left[2F(\sqrt{2}\Delta) + F(2\Delta) - 3F(\Delta)\right] - 2P_p F(\Delta)[1 - c_A F(\Delta)] \\ & + \frac{P_p}{2c_A}\left[1 - c_A F(\Delta)\right] + \frac{P_p}{2(1-c_A)}\left[1 - c_A F(\Delta)\right]^2\left[2F(\sqrt{2}\Delta) + F(2\Delta)\right],\end{aligned} \tag{15}$$

and

$$F(s \to \infty) = 1. \tag{16}$$

In [24], the authors demonstrate that this PDE provides excellent agreement with the system of ODEs. We note that the PDE has the form of a reaction-diffusion equation. The 'diffusion' coefficient of our equations depends only on $P_m$, and does not depend on proliferation. Correlations will decay faster for larger values of $P_m$. Intuitively, we expect higher rates of movement to break up clusters of agents, thereby agreeing with the conclusions drawn from our model. Our 'reaction' term depends on the rate of proliferation, and it can be argued it will always be non-positive [24]. This implies that increasing $P_p$ will decrease the level of correlation at that specific distance. On first inspection, this appears counterintuitive: we expect higher levels of proliferation to cause correlations to build-up. However, it is only nearest neighbour correlations ($s = \Delta$) which are positively affected by increased

proliferation, which can be observed in our ODE for the boundary conditions. Thus, the boundary at $s = \Delta$ is acting as a source of correlations, which are then dispersed by movement.

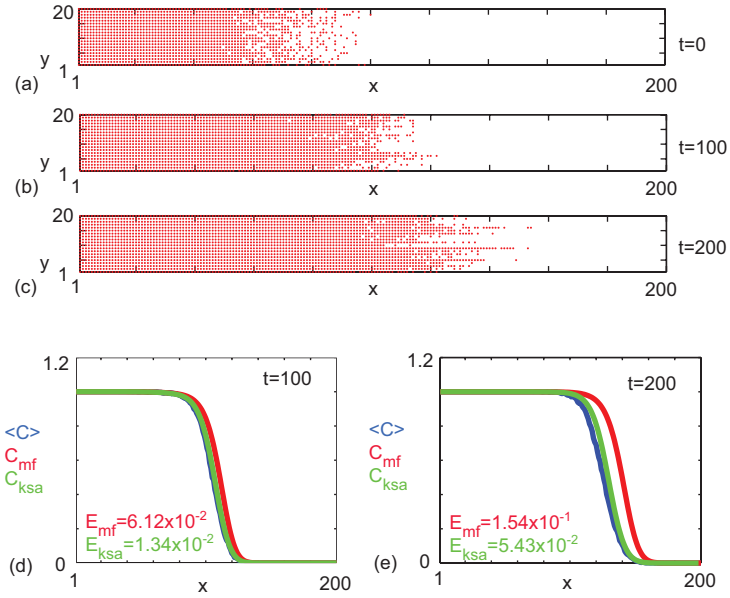


FIGURE 3. (a)-(c) show snapshots of 20 discrete results, whilst (d)-(e) show the density from the averaged discrete simulations (blue/black), mean-field (red/dark grey) and correlations models (green/light grey). We see that the invading wave travels with a faster speed in the mean-field case, whilst the correlations model provides good agreement with the averaged discrete results. More details can be found in [37].

The models discussed thus far rely on assuming translational invariance, which can be done because of the uniformity of the initial conditions. However, biologically, systems do not always start from uniform initial conditions. In 2011, Simpson and Baker extended their analysis to account for non-uniform initial conditions [36]. Compared to the mean-field model, this model provides a much better prediction of the transient behaviour in the averaged discrete case, as can be seen in Figure 3. The derivation of the correlation equations is similar to that for the case with uniform initial conditions, however we can no longer assume translational invariance. This results in a much larger system of equations as we now have a set of correlation ODEs for every lattice site. To simplify this, Ascolani *et al.* [2] develop a nearest-neighbour PDE method to examine systems with adhesion and contact interactions. The authors demonstrate the success of their method at predicting the averaged discrete results in the systems in question, despite only accounting for short-range correlation functions.

Many biological systems involve multiple species, thus it is necessary to extend these correlation methods for them to be more widely usable. Markham *et al.* [25] tackled this problem using the simplified partial differential equation correlation approach. For a two species problem, the density results are especially striking,

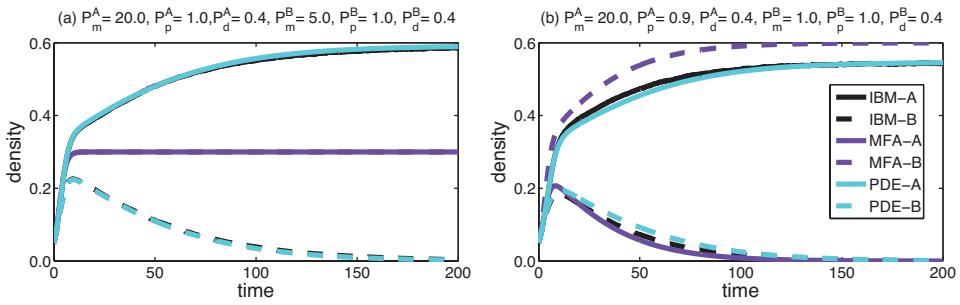


FIGURE 4. For two species systems, the MFA (purple/dark grey) sometimes predicts very different behaviour from what is seen in the averaged discrete case (black). However, the correlations model (cyan/light grey) is in good agreement with the averaged discrete results. Species $A$ is represented with solid lines, whilst species $B$ is in dashed. In (a), the averaged discrete simulations predict that species $B$ will go extinct, which the correlations model also predicts. However, the MFA predicts that the two species will co-exist with the same steady-state. In (b), the correlations model, in agreement with the averaged discrete simulations, predicts that species $B$ will go extinct. In contrast, the MFA shows species $A$ going to extinction. More details can be found in [25].

as we sometimes find the MFA predicting the opposite of what is observed in the averaged discrete case, whilst the correlations model provides a good approximation (Figure 4).

Thus we see that including spatial correlations is necessary for many biological systems, and that we must be cautious when using the MFA in situations which may not be well-mixed. The existing methods for incorporating correlations are numerically tractable, and applicable in a number of situations. Of course, there is still much work to be done, especially in the case for non-uniform initial conditions.

5. **Conclusion.** The past few years have witnessed tremendous advances in biology, allowing us to now visualise details of cell interactions in ways we could not before. This presents an exciting challenge for mathematicians, as we now must incorporate these observations into our modelling. In turn, this leads to new types of models and therefore new problems in analysis and numerical computation. In this very brief review, we have presented some recent examples in which these new generations of models have been developed and used. While the focus in Section 2 and Section 3 was on developing more realistic models of cells moving in a crowded environment, or taking into account volume exclusion, Section 4 critiqued the assumption from which many mean-field models are derived, namely that there is no spatial correlation between individual agents. Results demonstrate the care that must be taken in deriving accurate, biologically realistic models in the context of cell biology.

There is a long history of the study of spatial correlations and moment dynamics in the context of ecology (see, for example, [4, 5, 6]) in which moment equations of integro-differential form, which are analytically tractable in some cases, have been derived for plant populations. There, it has been shown that second-order

moment equations can capture spatio-temporal dynamics to a high degree of accuracy, so that third spatial moments can be neglected. In some cases translational invariance is also assumed, but this is not necessary, depending on the form of the initial conditions [20]. There have been detailed comparisons made between mean field approaches, patch models, reaction-diffusion equations, and interacting particle systems (see, for example, [9]). There are also cases in which individual-based models can be represented by partial differential equations for the pair correlation functions (see, for example, [39]).

In the aforementioned ecological examples, a distribution kernel for offspring that has quite a wide range is assumed, and then density dependent death is incorporated. In this paper, we have shown how these ideas extend to cell biology by using a different approach which incorporates volume exclusion and is off-lattice, which is perhaps the biggest difference with previous work. Analysis of the resulting models demonstrate the care that must be taken in deriving accurate, biologically realistic models in the context of cell biology.

**Acknowledgments.** D.C.M. acknowledges the Oxford University Press for support through the Clarendon Fund, as well as Keble College,Oxford, for support through the Sloane-Robinson award.

## REFERENCES

[1] K. Anguige and C. Schmeiser, A one-dimensional model of cell diffusion and aggregation, incorporating volume filling and cell-to-cell adhesion, *J. Math. Biol.*, **58** (2009), 395–427.
[2] G. Ascolani, M. Badoual and C. Deroulers, Exclusion processes: Short-range correlations induced by adhesion and contact interactions, *Phys. Rev. E*, **87** (2013), 012702.
[3] R. E. Baker and M. J. Simpson, Correcting mean-field approximations for birth-death-movement processes, *Phys. Rev. E.*, **82** (2010), 041905, 12pp.
[4] B. Bolker and S. W. Pacala, Using moment equations to understand stochastically driven spatial pattern formation in ecological systems, *Theor. Pop. Biol.*, **52** (1997), 179–197.
[5] B. M. Bolker and S. W. Pacala, Spatial moment equations for plant competition: Understanding spatial strategies and the advantages of short dispersal, *Theor. Pop. Biol.*, **153** (1999), 575–602.
[6] B. M. Bolker, S. W. Pacala and S. A. Levin, Moment methods for ecological processes in continuous space, In *The Geometry of Ecological Interactions: Simplifying Spatial Complexity*, eds. U. Dieckmann, R. Law and J. A. J. Metz, chapter 20, 338–411.
[7] B. M. Bolker, S. W. Pacala and C. Neuhauser, Spatial dynamics in model plant communities: What do we really know?, *Am. Nat.*, **162** (2003), 135–148.
[8] U. Dieckmann and R. Law, Relaxation projections and the method of moments, *Cambridge University Press*, **21** (2000), 412–457.
[9] R. Durrett and S. Levin, The importance of being discrete (and spatial), *Theor. Pop. Biol.*, **46** (1994), 363–394.
[10] L. Dyson, P. K. Maini and R. E. Baker, Macroscopic limits of individual-based models for motile cell populations with volume exclusion, *Phys. Rev. E*, **86** (2012), 031903.
[11] A. Fasano, M. A. Herrero and M. R. Rodrigo, Slow and fast invasion waves in a model of acid-mediated tumour growth, *Math. Biosci.*, **220** (2009), 45–56.
[12] R. A. Gatenby and E. T. Gawlinski, A reaction-diffusion model of cancer invasion, *Cancer Res.*, **56** (1996), 5745–5753.
[13] F. Graner and J. A. Glazier, Simulation of biological cell sorting using a two-dimensional extended Potts model, *Phys. Rev. Lett.*, **69** (1992), 2013–2016.
[14] S. T. Johnston, M. J. Simpson and R. E. Baker, Mean-field descriptions of collective migration with strong adhesion, *Phys. Rev. E*, **85** (2012), 051922.
[15] M. J. Keeling, Correlation equations for endemic diseases: Externally imposed and internally generated heterogeneity, *Proc. R. Soc. Lond. B*, **266** (1999), 953–960.
[16] M. J. Keeling, D. A. Rand and A. J. Morris, Correlation models for childhood epidemics, *Proc. R. Soc. Lond. B*, **264** (1997), 1149–1156.

[17] J. G. Kirkwood, Statistical mechanics of fluid mixtures, *J. Chem. Phys.*, **3** (1935), 300–314.
[18] J. G. Kirkwood and E. M. Boggs, The radial distribution function in liquids, *J. Chem. Phys.*, **10** (1942), 394–403.
[19] R. Law, D. J. Murrell and U. Dieckmann, Population growth in space and time: Spatial logistic equations, *Ecology*, **84** (2003), 252–262.
[20] M. A. Lewis and S. Pacala, Modeling and analysis of stochastic invasion processes, *Theor. Pop. Biol.*, **41** (2000), 387–429.
[21] P. M. Lushnikov, N. Chen and M. Alber, Macroscopic dynamics of biological cells interacting via chemotaxis and direct contact, *Phys. Rev. E*, **78** (2008), 061904.
[22] J. Mai, V. N. Kuzovkov and W. von Niessen, A theoretical stochastic model for the $a+1/2b_2 \to 0$ reaction, *J. Chem. Phys.*, **98** (1993), 10017–10025.
[23] J. Mai, V. N. Kuzovkov and W. von Niessen, A general stochastic model for the description of surface reaction systems, *Physica A*, **203** (1994), 298–315.
[24] D. C. Markham, M. J. Simpson and R. E. Baker, Simplified method for including spatial correlations in mean-field approximations, *Phys. Rev. E*, **87** (2013), 062702.
[25] D. C. Markham, M. J. Simpson, P. K. Maini, E. A. Gaffney and R. E. Baker, Incorporating spatial correlations into multispecies mean-field models, *Phys. Rev. E*, **88** (2013), 052713.
[26] J. B. McGillen, E. A. Gaffney, N. K. Martin and P. K. Maini, A general reaction–diffusion model of acidity in cancer invasion, *J. Math. Biol.*, **68** (2014), 1199–1224.
[27] F. A. Meineke, C. S. Potten and M. Loeffler, Cell migration and organization in the intestinal crypt using a lattice-free model, *Cell Prolif.*, **34** (2001), 253–266.
[28] P. J. Murray, C. M. Edwards, M. J. Tindall and P. K. Maini, From a discrete to a continuum model of cell dynamics in one dimension, *Phys. Rev. E*, **80** (2009), 031912.
[29] P. J. Murray, C. M. Edwards, M. J. Tindall and P. K. Maini, Classifying general nonlinear force laws in cell-based models via the continuum limit, *Phys. Rev. E*, **85** (2012), 021921.
[30] D. J. Murrell, U. Dieckmann and R. Law, On moment closures for population dynamics in continuous space, *J. Theor. Biol.*, **229** (2004), 421–432.
[31] H. G. Othmer and A. Stevens, Aggregation, blowup, and collapse: The ABC's of taxis in reinforced random walks, *SIAM J. Appl. Math.*, **57** (1997), 1044–1081.
[32] K. J. Painter and T. Hillen, Volume-filling and quorum-sensing in models for chemosensitive movement, *Canadian Appl. Math. Quarterly*, **10** (2002), 501–543.
[33] M. Raghib, N. A. Hill and U. Dieckmann, A multiscale maximum entropy moment closure for locally regulated space-time point process models of population dynamics, *J. Math. Biol.*, **62** (2011), 605–653.
[34] K. J. Sharkey, Deterministic epidemic models on contact networks: Correlations and unbiological terms, *Theor. Pop. Biol.*, **79** (2011), 115–129.
[35] K. J. Sharkey, C. Fernandez, K. L. Morgan, E. Peeler, M. Thrush, J. F. Turnbull and G. B. Bowers, Pair-level approximations to the spatio-temporal dynamics of epidemics on asymmetric contact networks, *J. Math. Biol.*, **53** (2006), 61–85.
[36] M. J. Simpson and R. E. Baker, Corrected mean-field models for spatially dependent advection-diffusion-reaction phenomena, *Phys. Rev. E*, **83** (2011), 051922.
[37] M. J. Simpson, B. J. Binder, P. Haridas, B. K. Wood, K. K. Treloar, D. L. S. McElwain and R. E. Baker, Experimental and modelling investigation of monolayer development with clustering, *Bull. Math. Biol.*, **75** (2013), 871–889.
[38] O. Warburg and F. Dickens, The metabolism of tumors, *Am. J. Med. Sci.*, **182** (1931), 123.
[39] W. R. Young, A. J. Roberts and G. Stuhne, Reproductive pair correlations and the clustering of organisms, *Nature*, **412** (2001), 328–331.

Received February 2014; revised May 2014.

*E-mail address*: markham@maths.ox.ac.uk
*E-mail address*: baker@maths.ox.ac.uk
*E-mail address*: maini@maths.ox.ac.uk