

A model for membrane degradation using a gelatin invadopodia assay

Giorgia Ciavolella¹² Nathalie Ferrand³ Michèle Sabbah³
 Benoît Perthame⁴ Roberto Natalini⁵

April 10, 2024

Abstract

One of the most crucial and lethal characteristics of solid tumors is represented by the increased ability of cancer cells to migrate and invade other organs during the so-called metastatic spread. This is allowed thanks to the production of matrix metalloproteinases (MMPs), enzymes capable of degrading a type of collagen abundant in the basal membrane separating the epithelial tissue from the connective one. In this work, we employ a synergistic experimental and mathematical modelling approach to explore the invasion process of tumor cells. A mathematical model composed of reaction-diffusion equations describing the evolution of the tumor cells density on a gelatin substrate, MMPs enzymes concentration and the degradation of the gelatin is proposed. This is completed with a calibration strategy. We perform a sensitivity analysis and explore a parameter estimation technique both on synthetic and experimental data in order to find the optimal parameters that describe the *in vitro* experiments. A comparison between numerical and experimental solutions ends the work.

2010 *Mathematics Subject Classification.* 35K57; 35Q92; 35R30; 65M06; 65M22

Keywords and phrases. Reaction-diffusion equations; Finite difference methods; Tumour degradation and invasion models; Parameter estimation; Sensitivity analysis

1 Introduction

Tumors are complex diseases characterised by high diversity and incidence, Sung *et al.* [33]. One of the most crucial and lethal processes is represented by the increased ability of cancer cells to migrate and invade other organs during the so-called *metastatic spread*, Dillekås *et al.* [14]. In the primary tumor, cancer cells could acquire useful mutations that allow them to penetrate

¹INRIA Bordeaux-Sud-Ouest, Institut de Mathématiques de Bordeaux, CNRS UMR 5251 & Université de Bordeaux, 351 cours de la Libération, 33405 Talence Cedex, France

²Corresponding author: giorgia.ciavolella@inria.fr, ORCID number 0000 0002 6614 3966

³Sorbonne Université Cancer Biology and Therapeutics, INSERM, CNRS, Institut Universitaire de Cancérologie, Saint-Antoine Research Center (CRSA), F-75012, Paris, France

⁴Sorbonne Université, Inria, Université de Paris, Laboratoire Jacques-Louis Lions, UMR7598, 4 place Jussieu, 75005 Paris, France

⁵Istituto per le Applicazioni del Calcolo 'M. Picone', Rome, Italy

different barriers and to disseminate themselves into secondary organs. Then, we observe the transition from an *in situ* stage to an invasive one. It is now well known that metastatic cancer cells typically move in clusters which have greater predisposition of forming metastasis than single cells, Aceto *et al.* [1], Bubba *et al.* [6], Hong *et al.* [21].

Invasion of cells through layers of extracellular matrix is a key step in tumor metastasis, inflammation, and development. The process of invasion involves several stages, including adhesion to the matrix, degradation of proximal matrix molecules, extension and traction of the cell on the newly revealed matrix, and movement of the cell body through the resulting gap in the matrix, Friedl and Wolf [16]. Each of these stages of invasion is executed by a suite of proteins, including proteases, integrins, GTPases, kinases, and cytoskeleton-interacting proteins. One of the most difficult barriers for cells to cross is the basement membrane, also composed by extracellular matrix (ECM). This kind of membrane separates the epithelial tissue from the connective one and, among its functions, we can distinguish a supportive role and an isolating one. Tumor cell invasion is associated with an enhanced capability of tumor cells to degrade ECM. Tumor cells produce several hydrolytic enzymes including matrix metalloproteinases (MMPs). This process is, for instance, observed for breast tumors, as considered in our study.

Many questions concerning invasion details remain unanswered. Over the last decade, the research interest on this process is increasing in order to highlight the main cues with the aim of controlling and treating the phenomenon (Chaplain *et al.* [7], Ciavolella *et al.* [11], Franssen *et al.* [15], Gallinato *et al.* [17], Givero *et al.* [18]).

In particular, we are interested in better describing one of the main biological phenomenon responsible for the invasion one, which is membrane degradation. In fact, *in vitro* invasion investigations (using the XCELLigence technology, Connolly and Maxwell [12], Ke *et al.* [23], Martinez-Serra *et al.* [26], Obr *et al.* [29], Turker *et al.* [34], Zaoui *et al.* [36]) are not able to give details on this process. Consequently, we build here a mathematical model which describes degradation of an ECM-like biological membrane through the production of MMPs by cancer cells. At the same time, we provide numerical simulations with a sensitivity analysis followed by a parameters estimation study. The work is completed by experimental results consisting of cells seeded in wells containing at their bottom fluorescent gelatin. We focus the attention on breast cancer, the most common malignancy among women. The Triple Negative Breast Cancer (TNBC) patients are mainly treated with combinations of chemotherapy with severe side effects and frequent recurrence of the metastasis. If non-metastatic, it has high chances of healing, but, on the contrary, advanced breast cancer with metastasis are considered incurable with the present therapies, Harbeck *et al.* [20], Waks and Winer [35].

The outline of the paper is as follows. In Sections 2 and 2.1, we introduce the mathematical model and its dimensionless form. Section 3 is divided in two subsections describing the sensitivity analysis and the parameter estimation problem. In Subsection 3.1, we briefly present the concept of sensitivity analysis and then the results on our model. The same is done in Subsection 3.2, in the case of both synthetic data and experimental ones. The subsection ends with numerical simulations which are compared with the biological experiments. Finally, in Section 4, we conclude the paper presenting also some perspectives that both improve our work and develop it. At the end of the paper, the reader can find the Supplementary Materials S5 in which we give more details concerning the numerical method behind the simulations presented.

2 Mathematical model

We build a mathematical model describing the invasive behaviour of tumor cells on a gelatin coated plate. Their movement is influenced by the gelatin, since their primary role is to degrade

it, due to MMPs enzymes. The major hypothesis is that cells degrade gelatin below them and, after, they move to degrade around. Experiments capturing this behaviour are realised using the QCM™ Gelatin Invadopodia Assay. Tumor cells are plated on a culture surface coated with a thin layer of green fluorescent gelatin and they are filmed over several days. Videos can show both cells movement and the consequent creation of holes in the gelatin.

In accordance with this kind of experimental setting, we consider a domain Ω representing a top view of a single plate. We do not include the third dimension because the thickness of the gelatin layer is not significant compared to cells size. Consequently, the movement is only realised on a surface. For simplicity, we take the rectangular two-dimensional domain Ω . For $x \in \Omega$, $t > 0$, we consider a three species system for cells density $u(x, t)$, MMPs enzymes concentration $m(t, x)$, and the damage function $d(t, x) \in [0, 1]$, related to the amount of gelatin $q(t, x) = 1 - d(t, x) \in [0, 1]$. Equations write as

$$\begin{cases} \partial_t u = \text{div}(D(d)\nabla u), & \text{in } \Omega, \\ \partial_t m = D_m \Delta m + \beta(1 - d)u - \alpha m, & \text{in } \Omega, \\ \partial_t d = \gamma m(1 - d), & \text{in } \Omega, \end{cases} \quad (1)$$

where

$$D(d) = D_L d + D_G(1 - d) = D_G + (D_L - D_G)d.$$

We impose no-flux boundary conditions on $\partial\Omega$ for u and m , *i.e.*

$$\begin{cases} D(d)\nabla u \cdot \mathbf{n} = 0, \\ D_m \nabla m \cdot \mathbf{n} = 0, \end{cases}$$

where \mathbf{n} is the outward unit normal at the boundary. We complete the system with the following initial conditions

$$\begin{cases} u(0, x) = u_0(x), \\ m(0, x) = m_0(x) = 0, \\ d(0, x) = d_0(x) = 0, \end{cases}$$

with u_0 a random function on Ω .

Equation for u . The first equation in (1) describes the evolution of the cells density. The diffusion coefficient depends on the diffusion $D_L > 0$ into the liquid and $D_G > 0$ on the gelatin, with $D_L > D_G$. We do not consider the case $D_L = D_G$, otherwise we would have a standard diffusion equation. We observe that when the gelatin is intact ($d = 0$, then $q = 1$), cells move randomly on the gelatin, whereas when it is completely destroyed ($d = 1$, then $q = 0$), cells diffuse into the liquid. The expression of $D(d)$ is not new in applications. In the literature we can find it, for example, in models describing the chemical transformation of calcium carbonate stones under the attack of sulphur dioxide, Aregba-Driollet *et al.* [2], Guarguaglini and Natalini [19].

Equation for m . The second equation is a reaction-diffusion like equation for the MMPs with diffusion coefficient $D_m > 0$, production rate $\beta > 0$ and death rate $\alpha > 0$. In particular, production of MMPs is due to cells and to the fact that they sense the gelatin below them. Moreover, we assume that MMPs diffuse locally and, consequently, we add the condition

$$\sqrt{\frac{D_m}{\alpha}} \ll 1,$$

where $\sqrt{\frac{D_m}{\alpha}}$ is homogeneous to $\sqrt{x^2}$ and corresponds to the diffusion length.

Equation for d . Finally, the equation for the damage derives from the damage mechanics, Kachanov [22]. It can be written in terms of the gelatin q as

$$\partial_t q = -\gamma m q,$$

which has an exponential decreasing solution. The damage is produced at rate γ by the MMPs m .

Remark 2.1. *The choice of a macroscopic model is dictated by the high number of tumor cells in experiments. Moreover, it is more interesting to look at the general behaviour of cells instead of the single individual. Indeed, tracking them, we could observe that over 72 hours they make local movements around the initial position and at very low velocities. No particular interactions between cells could then be taken into account. Finally, a hybrid approach could be introduced, but this would largely increase the number of parameters, adding complexity in their estimation. We deduce the multiple advantages of our simple PDE model that is able to describe degradation phenomena.*

2.1 Dimensionless model

Before analysing the model, we propose a nondimensional form. It has several advantages, as the reduction of the number of parameters and the fact that their units are unimportant, see Murray [28], Segel [32]. Upon changes of time and space variables

$$\tilde{t} = \alpha t, \quad \tilde{x} = \sqrt{\frac{\alpha}{D_m}} x,$$

and appropriate scaling for u, m , and d , namely

$$\begin{aligned} u(t, x) &= \tilde{u} \left(\alpha t, \sqrt{\frac{\alpha}{D_m}} x \right), \\ m(t, x) &= \frac{\alpha}{\gamma} \tilde{m} \left(\alpha t, \sqrt{\frac{\alpha}{D_m}} x \right), \\ d(t, x) &= \frac{D_m}{D_L - D_G} \tilde{d} \left(\alpha t, \sqrt{\frac{\alpha}{D_m}} x \right), \end{aligned}$$

we find a parametrised version, again with homogeneous Neumann boundary conditions. For simplicity in notation, eliminating the tilde, we find

$$\begin{cases} \partial_t u = \operatorname{div}((\theta + d)\nabla u), & \text{in } \Omega, \\ \partial_t m = \Delta m + k_1(1 - pd)u - m, & \text{in } \Omega, \\ \partial_t d = \frac{1}{p}m(1 - pd), & \text{in } \Omega, \end{cases} \quad (2)$$

where $\theta = \frac{D_G}{D_m}$, $p = \frac{D_m}{D_L - D_G}$, $k_1 = \frac{\gamma\beta}{\alpha^2}$. Here, $D(d) = \theta + d$. We assume all parameters positive, except p that satisfies the restriction $p \leq 1$, since we require $1 - pd \geq 0$.

3 Sensitivity analysis and parameter estimation

We provide a parameters estimation study for the dimensionless Model (2). Indeed, parameters are not accessible by direct measurements from experiments. To test the estimation technique, it is interesting to work at first with synthetic data. The idea is that we create a numerical (synthetic) solution, through our Model (2), using parameters from the literature. Then, we find

the best set of parameters that gives reasonable solutions and we can evaluate the error made in the choice of parameters, see Subsection 3.2.

Another useful study that can be done in parallel is the sensitivity analysis, see Subsection 3.1. It provides us with an instrument to examine how the choice of the parameters affects the model dynamics. This is also a key indicator in the case in which the error between one of the optimal parameters and its corresponding initial estimate is too big. In fact, the sensitivity analysis could show us that some parameters do not have important effects on solutions. Thus, a huge error on their estimation is not so important.

At the end of Subsection 3.2, we provide an example of possible experimental data and we reiterate the consolidated procedure performed with synthetic data. We evaluate the optimal parameters describing the biological experiment and we provide numerical simulations of the model solutions. Except for this experiments paragraph, in the following we integrate the theory with our model analysis considering parameters taken from Di Costanzo *et al* [13], and Braun [4],

$$D_L = 7 \times 10^{-7} \text{ cm}^2\text{s}^{-1}, \quad D_G = 10^{-7} \text{ cm}^2\text{s}^{-1}, \quad D_m = 5 \times 10^{-7} \text{ cm}^2\text{s}^{-1},$$

and from Franssen *et al* [15],

$$\alpha = 2.5 \times 10^{-6} \text{ s}^{-1}, \quad \beta = 4.9 \times 10^{-6} \text{ Ms}^{-1}.$$

Consequently, the nondimensional parameters are

$$\theta = 0.2, \quad p = 0.83, \quad k_1 = 0.78. \quad (3)$$

3.1 Sensitivity analysis

The lack of parameters availability from experiments imposes an uncertainty in the output of the model. To obtain as reliable results as possible, we have to study the influence of the parameters on the model dynamics through a *local sensitivity analysis*. Local methods are the simplest and the most common. They are based on a one-factor-at-a-time (OAT) method. It consists in perturbing one parameter at a time to see the effect on the output for each parameter of the model. Sensitivity analysis (SA) is then performed monitoring changes in the output through, for example, a derivative-based approach.

SA determines dependencies of input parameters $\mathcal{Q} = (\mathcal{Q}_1, \dots, \mathcal{Q}_n)$ and output of the model \mathcal{Y} . In the following we illustrate the procedure. The measure of the influence of the parameter \mathcal{Q}_i on the output $\mathcal{Y}(\mathcal{Q})$ is calculated by the partial derivative with respect to this parameter. Since the model parameters differ by several orders of magnitude, we introduce a normalisation with respect to the mean value. Using the finite difference approach, the sensitivity of the output with respect to the parameter \mathcal{Q}_i is obtained as

$$S = \frac{\mathcal{Y}(\mathcal{Q}_1, \dots, \mathcal{Q}_i \pm \delta, \dots, \mathcal{Q}_n) - \mathcal{Y}(\mathcal{Q}_1, \dots, \mathcal{Q}_n)}{\mathcal{Y}(\mathcal{Q}_1, \dots, \mathcal{Q}_n)} \frac{\mathcal{Q}_i}{\delta},$$

where $\delta = 0.05 \mathcal{Q}_i$ to consider a 5% perturbation of the parameters.

For our model. We report here the results obtained for our Model (2), using the dimensionless parameters from the literature in (3). We have evaluated the sensitivity of parameters $\mathcal{Q} = [\theta, p, k_1]$ taking as output \mathcal{Y} the maximum value at the final time for cells density and the total mass of enzymes and gelatin at the final time. It is not interesting to consider the cells total mass since this is a conserved quantity. The OAT method does not examine the whole range of values of the input parameters as the global SA. However, this method is easier to define

and computationally more efficient than the global sensitivity, for which different techniques are possible, see Saltelli *et al.* [31].

In Table 1, we show for each perturbed parameter $Q_i \pm \delta$, the sensitivity S_y related to the output \mathcal{Y} . Sensitivity should be less than one or around it but not much bigger in order to have low sensitivity of the model with respect to parameters. Since in our case it is at maximum around 1, we deduce that solutions have not a great influence on big changes in parameters. Of course, the bigger is the sensitivity, the bigger is the influence on solutions. Moreover, from Table 1, we deduce that the diffusion parameter does not play a remarkable role in the behaviour of enzymes m and degradation d , whereas it is the most 'critical' parameter for cells evolution. In this case, in fact, p and k_1 are of order 10^{-2} . More critical is the influence of p on degradation and k_1 on both degradation and enzymes concentration, but still sensitivity is good enough. From SA, we infer that changing the literature parameters do not greatly affect solutions behaviour. Not knowing real solutions of the model, this is a good estimate to know about.

$Q_i \pm \delta \backslash S_y$	S_{maxu}	S_{massm}	S_{massd}
$\theta + \delta$	0.6262	6.57×10^{-4}	7.05×10^{-5}
$\theta - \delta$	0.6705	7.03×10^{-4}	7.48×10^{-4}
$p + \delta$	0.0141	2.44×10^{-4}	0.9524
$p - \delta$	0.0155	2.7×10^{-6}	1.0526
$k_1 + \delta$	0.0146	0.9980	0.9982
$k_1 - \delta$	0.01469	0.9982	0.9972

Table 1: We collect sensitivity values S_y per each parameter $Q_i \pm \delta$, with $\delta = 0.05 Q_i$. S_{maxu} is the sensitivity with output the maximum of cell density at the final time (72 hours), S_{massm} is the sensitivity with output the amount of enzymes again after 72 hours and the same for S_{massd} with respect to gelatin degradation.

3.2 Inverse problem: parameter estimation with synthetic and biological data

With the knowledge that changing our parameters from the literature we do not modify too much solutions (Subsection 3.1), we can now estimate, through an inverse problem, the good parameters useful to recover reasonable solutions compared to synthetic and experimental ones.

A convenient reformulation of the inverse problem is to write it as a minimization problem.

Definition 3.1. Let $F(Q) = \mathcal{Y}$ be the forward problem with $Q = (Q_1, \dots, Q_n)$ the parameters and \mathcal{Y} the numerical solution found with the finite difference scheme. The inverse problem is defined as a minimisation problem of the form

$$Q_{opt} := \underset{Q \in \mathbb{R}^n}{\operatorname{argmin}} \|F(Q) - \mathcal{Y}_{exp}\|_{L^2}^2, \quad (4)$$

where Q_{opt} are the estimated parameters and \mathcal{Y}_{exp} the experimental (or synthetic) data.

Unfortunately, inverse problems are usually ill-posed. Then, initial small perturbations can lead to large ones in the results. In other words, small errors between the solution of the forward problem and the experimental data can lead to arbitrary large errors between the given

parameters and the estimated ones. Hence the necessity of a regularisation method to compute a stable approximation of the minimiser, Braun [4, 5]. Thus, we extend the minimisation in (4) with a Tikhonov regularisation as

$$\mathcal{Q}_{opt} := \underset{\mathcal{Q} \in \mathbb{R}^n}{\operatorname{argmin}} (\|F(\mathcal{Q}) - \mathcal{Y}_{exp}\|_{L^2}^2 + \lambda \|\mathcal{Q} - \mathcal{Q}_0\|_{L^2}^2),$$

where $\lambda > 0$ is the regularisation parameter and the *a priori* estimate $\mathcal{Q}_0 \in \mathbb{R}^n$ represents an *a priori* knowledge about the solution.

From gelatin invadopodia assay, it is possible to extract information mainly concerning tumor cells, since it could be difficult to quantify the gelatin and the holes created by degradation. One of the reasons is that cells have their own fluorescence and this could alter the quantification of gelatin. Up to our knowledge, it is very difficult to have data on enzymes. Thus, we take $\mathcal{Y} = u$ and $\mathcal{Y}_{exp} = u_{exp}$. Ignoring \mathcal{Q}_0 due to lack of *a priori* information on parameters, we infer that

$$\mathcal{Q}_{opt} := \underset{\mathcal{Q} \in \mathbb{R}^n}{\operatorname{argmin}} (\|F(\mathcal{Q}) - u_{exp}\|_{L^2}^2 + \lambda \|\mathcal{Q}\|_{L^2}^2). \quad (5)$$

Thus, if an unknown noise is included in the data, there might be different solutions that minimise $\|F(\mathcal{Q}) - u_{exp}\|_2^2$, but among these solutions only the one which minimises $\|\mathcal{Q}\|_2^2$ is selected. In our case, the main interest will be in minimising the error among solutions. Thus, in the following we will choose very small values of λ such that the norm of \mathcal{Q} will not play a significant role.

Parameter estimation with synthetic data for our model. Before working on biological data, it is commonly used to test the estimation on the synthetic data. Therefore, we develop the procedure used to estimate parameters for Model (2) taking the parameters from the literature. Choosing the latter as in (3)

$$\mathcal{Q}_{lit} = [\theta = 0.2, \quad p = 0.83, \quad k_1 = 0.78],$$

we derive numerical solutions u, m and $q = 1 - d$. Synthetic solutions are created from the numerical solutions of the model with the literature parameters. We choose two different synthetic data: at first, we use as u_{exp} the exact numerical solution, and in a second step we consider a perturbation of it. Indeed, in both cases, the minimisation of the functional in (5) starts with $F(\mathcal{Q})$, the numerical solution obtained with the perturbed literature parameters, and thus we aim to recover the best parameters such that $F(\mathcal{Q})$ is close to u_{exp} . So at first with u_{exp} the numerical solution of our Model (2), it should be easier to recover them, while adding a small perturbation, the idea is to reproduce an error that we also encounter with experimental data.

Thus, we perturb parameters with $n = 0.1, 0.2$ or 0.4 Gaussian noise, namely the perturbed parameters \mathcal{Q}_{pert} are composed by

$$\theta_{pert} = \theta + \omega_\theta n \theta, \quad p_{pert} = p + \omega_p n p, \quad k_{1_{pert}} = k_1 + \omega_{k_1} n k_1,$$

with ω_θ and ω_{k_1} random numbers in $[-1, 1]$, whereas ω_p is again a random number in $[-1, 1]$ in the case $n = 0.1, 0.2$ and $[-1, 0]$ if $n = 0.4$. Indeed, we recall that $p \leq 1$, so that $1 - pd \geq 0$. Then, we look for the best parameters that solve the inverse Problem (5) in an interval $I_{lit} = [\mathcal{Q}_{lit} - 0.5\mathcal{Q}_{lit}, \mathcal{Q}_{lit} + 0.5\mathcal{Q}_{lit}]$ and starting with the perturbed parameters above. So, we evaluate the parameters that minimise the error $E(\mathcal{Q})$ between our synthetic data and the new solution \hat{u} obtained with one of the parameters \mathcal{Q} in the research interval I_{lit} plus the Tikhonov regularisation with $\lambda = 10^{-6}$, namely

$$E(\mathcal{Q}) = \|\hat{u}(t) - u_{exp}(t)\|_2^2 + \lambda \|\mathcal{Q}\|_2^2, \quad \text{for } t = 48 \text{ hours.}$$

The choice of the time $t = 48$ hours is dictated by the dissipative behaviour of diffusion problems. Indeed, we can either loose information on parameters looking at the final time or not having them at all considering the initial time (initial data do not depend on parameters). Finally, we evaluate the relative error between the estimated parameters \mathcal{Q}_{opt} and the initial ones, *i.e.*

$$E_\theta = \frac{\|\theta - \theta_{opt}\|_2}{\|\theta\|_2}, \quad E_p = \frac{\|p - p_{opt}\|_2}{\|p\|_2}, \quad E_{k_1} = \frac{\|k_1 - k_{1_{opt}}\|_2}{\|k_1\|_2}.$$

We also estimate the relative error between the synthetic data u_{exp} and the solution with the estimated parameters, that is defined as

$$E_{\mathcal{Q}_{opt}}(t) = \frac{\|u_{exp}(t) - u_{\mathcal{Q}_{opt}}(t)\|_2}{\|u_{exp}(t)\|_2}, \quad \text{for } t = 0, 24, 48, 72 \text{ hours.} \quad (6)$$

Synthetic data: numerical solution

At first, we choose as synthetic data the numerical solution u . This means that the numerical solution of our model, found with the parameters from the literature, simulates the experimental data available. Thus, to find the optimal parameters with (5), we substitute u_{exp} with the numerical solution obtained with parameters from the literature and $F(\mathcal{Q})$ with the numerical solution obtained with the set of optimal parameters found at each iteration, starting with the perturbed parameters.

Example 3.1. We consider a 10% perturbation on parameters $\mathcal{Q}_{lit} = [\theta, p, k_1] = [0.2, 0.83, 0.78]$, randomly extracting $\omega_\theta = -0.48$, $\omega_p = -0.49$ and $\omega_{k_1} = 0.51$. We deduce

$$\mathcal{Q}_{pert} = [0.19, \quad 0.871, \quad 0.82].$$

Solving the inverse problem in (5), we infer that

$$\mathcal{Q}_{opt} = [0.199, \quad 0.415, \quad 0.39].$$

The errors on parameters are $E_\theta = 10^{-6}$, $E_p = 0.5$, $E_{k_1} = 0.5$, whereas

$$E_{\mathcal{Q}_{opt}} < 8 \times 10^{-6}.$$

Unfortunately, errors on p and k_1 are high. However, as expected from the sensitivity analysis in Subsection 3.1, this does not impact solutions. In Figure 1 left, we show the error $E_{\mathcal{Q}_{opt}}$ between the data u and the solution with estimated parameters. The same is provided also for the gelatin q . Both are below 8×10^{-6} indicating a good approximation of solutions. This confirms that a property of our model is stability.

Remark 3.1. The same results are obtained considering higher perturbations on parameters.

We conclude the example with a final observation. The relative error between solutions is of the same order as $\lambda = 10^{-6}$, and also lower. This implies that the minimisation does not consider anymore the norm of solutions, but only the norm of parameters \mathcal{Q} , see Equation (5). This is why we end up with a big error on parameters: the algorithm is trying to find very small parameters. Then, considering a lower $\lambda = 10^{-12}$, we could force the minimiser to decrease only the norms of the solutions. The results are in the following. We take the same data as before and, solving the inverse problem, we obtain

$$\mathcal{Q}_{opt} = [0.2, \quad 0.809, \quad 0.761].$$

The errors on parameters are significantly smaller, *i.e.*

$$E_\theta = 10^{-8}, \quad E_p = 0.02, \quad E_{k_1} = 0.02.$$

Again, we have a good estimate of solutions as before, see Figure 1 right.

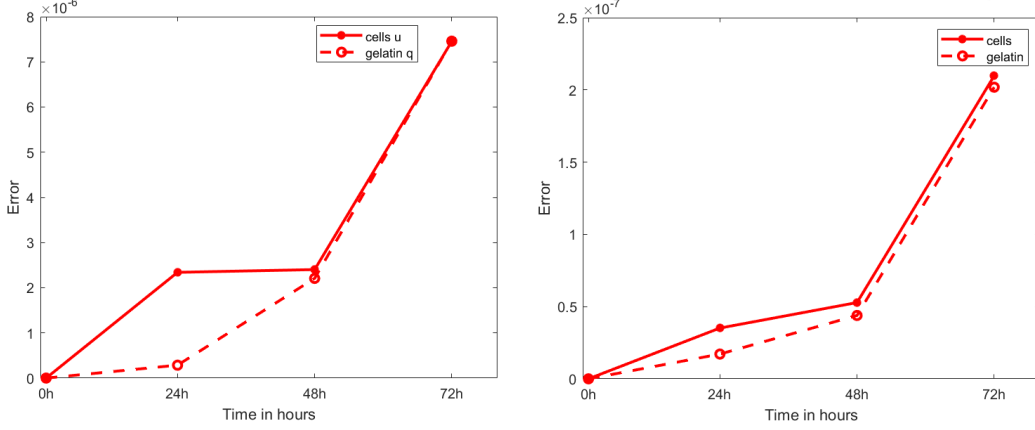


Figure 1: On the x -axis, the time in days until 3 days, which represents the duration of the experiments. On the y -axis, the error $E_{\mathcal{Q}_{opt}}$ in (6), both for cells u and gelatin q (dash line). The graphs correspond to the case in which the synthetic data are the numerical solutions, but the regularisation parameter λ changes. Indeed, on the left, $\lambda = 10^{-6}$, whereas on the right $\lambda = 10^{-12}$.

Synthetic data: perturbed numerical solution

Experimental (non synthetic) data have always some unknown noise due for instance to instrumentation. This is why in the following we consider a more realistic setting in which the synthetic data are a perturbation of the numerical solution of Model (2). We perturb with 5% Gaussian noise the numerical solution, *i.e.* $u_{pert} = u + \omega_u 0.05 u$, where ω_u is a random matrix with entries in $[-1, 1]$. This Gaussian noise is with zero mean, thus perturbed solutions are on average the exact numerical ones.

Example 3.2. We consider a 10% perturbation on parameters $\mathcal{Q}_{lit} = [\theta, p, k_1] = [0.2, 0.83, 0.78]$. Randomly extracting $\omega_\theta = -0.11$, $\omega_p = -0.15$ and $\omega_{k_1} = -0.83$, we deduce

$$\mathcal{Q}_{pert} = [0.198, \quad 0.817, \quad 0.715].$$

Solving the inverse problem in (5), we get

$$\mathcal{Q}_{opt} = [0.199, \quad 0.415, \quad 0.539].$$

This infers an error on parameters such that $E_\theta = 0.0007$, $E_p = 0.5$, $E_{k_1} = 0.3$, whereas

$$E_{\mathcal{Q}_{opt}} < 3 \times 10^{-3}.$$

As before, errors on p and k_1 are high. In Figure 2 left, we show the errors $E_{\mathcal{Q}_{opt}}$, which are for both u and q below the initial 5% error between the synthetic data and the numerical solution.

Remark 3.2. *We do not get very satisfactory results in terms of parameter estimation. However, the most important characteristic is that solutions obtained with the estimated parameters \mathcal{Q}_{opt} do not greatly differ from the experimental data (the perturbed solutions in this synthetic case). In contrast, we remark that the functional to be minimised in (5) does not include either q or m . Indeed, having more information also on gelatin and enzymes concentration could allow to include them in the functional, thus bringing us good estimations on the other parameters. In particular, the parameter p is present both in the equation for enzymes concentration m and gelatin degradation d , then we should need both of them to obtain a good estimate on p . The parameter k_1 is only in the equation for the enzymes and, indeed, we can approximate it having experimental data on enzymes.*

Remark 3.3. *A different situation promises to be the case with higher perturbation on parameters. Indeed, we notice a better convergence rate for k_1 with respect to Example 3.2, probably due to a better formulation of the minimisation problem which well represents the final minima. Otherwise, we always have good final errors on solutions.*

So if we look for the optimal parameters describing the synthetic experimental data (chosen as the perturbed numerical solution u_{pert}), we deduce:

- a good approximation of θ with an error around 0.001;
- a maximal error on p ;
- a varying error on k_1 , which is around 0.04 or 0.3 in the examples presented.

Moreover, we have good error values $E_{\mathcal{Q}_{opt}}$ between the experimental synthetic solution and the one found with the optimal parameters. Indeed, for both cells density and gelatin concentration we obtain $E_{\mathcal{Q}_{opt}} \sim 3\%$.

In order to obtain accurate optimal parameters, we should need either more experimental data, to be included in the functional definition, or eventually a model reduction. Concerning $E(\mathcal{Q})$, we could modify the value of λ , maybe a lower one, as in Example 3.1 where the error between solutions is of the same order as λ . Regarding a model reduction, even if relatively simple, System (2) could still be too complex to describe experiments. A detailed analysis of it could highlight particular behaviours that we currently ignore. We could also add an *a priori* knowledge on parameters such as $\mathcal{Q}_0 = \mathcal{Q}_{lit}$. This last addition would be less realistic but, at least, we force the minimisation to consider parameters not too far away from the *a priori* ones. An example is shown just below.

Example 3.3. *Let $\lambda = 10^{-3}$, $\mathcal{Q}_0 = \mathcal{Q}_{lit}$. We consider a 10% perturbation on parameters $\mathcal{Q}_{lit} = [\theta, p, k_1] = [0.2, 0.83, 0.78]$, randomly extracting $\omega_\theta = 0.27$, $\omega_p = -0.12$ and $\omega_{k_1} = -0.81$. Solving the inverse problem in (5), we get*

$$\mathcal{Q}_{opt} = [0.2, \quad 0.806, \quad 0.803].$$

This infers an error on parameters $E_\theta = 0.003$, $E_p = 0.03$, $E_{k_1} = 0.03$, whereas

$$E_{\mathcal{Q}_{opt}} < 3 \times 10^{-3}.$$

In Figure 2 right, we show the errors $E_{\mathcal{Q}_{opt}}$, which are for both u and q below the initial error of 5%.

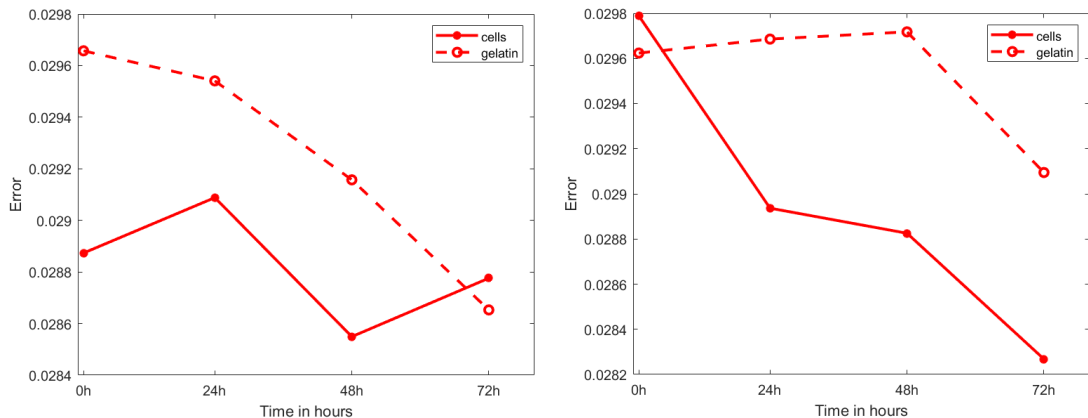


Figure 2: On the x -axis, the time in days until 3 days, which represents the duration of the experiments. On the y -axis, the error $E_{Q_{opt}}$ in (6), both for cells u and gelatin q (dash line). The graphs correspond to the case in which the synthetic data are the perturbed numerical solution, but the a priori information of parameters changes. On the left, we lack in information, whereas on the right $Q_0 = Q_{lit}$. There are no important variations in the solutions error. Instead a difference can be seen in the estimated parameters.

Parameter estimation with biological data for our model. We briefly present the experimental setting. The study is on the breast carcinoma cell line called MDA-MB-231 with a triple negative breast cancer (TNBC) phenotype. As introduced before, cells are placed on a thin layer of green fluorescent gelatin. Gelatin means to mimic basal membrane which has a thickness of 10 to 300 nm which is smaller than the size of a cell, *i.e.* 10 μ m. Fluorescence is useful to distinguish areas in which gelatin has been consumed. Such assays have also revealed that invasive cells extend small localised protrusions, called *invadopodia*, where MMP enzymes are localised and from which membrane degradation starts. Pioneered by Wen-Thien Chen in the 1980's ([8–10]), visualisation of invadopodia ECM degradation by fluorescent gelatin has emerged as the most prevalent technique for evaluating cellular invasive potential, Artym *et al.* [3], Martin *et al.* [25].

Experimental data consist of videos over 72 hours on a portion of the plate, measuring $2530 \times 2530 \mu\text{m}^2$. They focus either on cells movement or on gelatin degradation. An example of the results is given in Figure 4. At the beginning, gelatin is intact (top image right) and every day the consumption increases and this corresponds to the darker regions. On the left column, the configuration of tumor cells starting from the initial one on the top till the third day (on the bottom).

In possession of the biological data, we can apply the parameter estimation tested before considering u_{exp} as the biological experimental data, instead of the synthetic ones. We have decomposed the experimental video in 18 images (one every four hours). Using the Matlab tool *impixel*, it is possible to manually identify cells position and to have a discrete representation of them. Thus, we need to transform particles into densities in order to obtain a macroscopic view of cells, that we name u_{exp} . This procedure consists of centring a Gaussian kernel on each cell, see Braun [4, 5]. Then, discretising u_{exp} , we can compare it with our numerical solution through the inverse problem written in (5), that having reliable data only on cells, can be rewritten as

$$\theta_{opt} := \underset{\theta}{\operatorname{argmin}} (\|F(\theta)(t) - u_{exp}(t)\|_{L^2}^2 + \lambda \|\theta\|_{L^2}^2),$$

with t corresponding to the values at which we have evaluated u_{exp} , *i.e.* every 4 hours in the

interval $[0, 72]$ hours. We highlight that we are not looking for optimal parameters p and k_1 , since they do not significantly influence the behaviour of cells. Indeed, sensitivity analysis with experimental data shows very similar results to the ones in Table 1, thus p and k_1 play very little role on cells dynamics. Consequently, they are going to be fixed. Furthermore, we realised that parameters from the literature were not appropriate to describe experimental data. The first problem is in the description of cells movement that is clear only with a higher θ . Moreover, k_1 is too low to allow enzymes production and the consequent gelatin degradation. Thus, we modify them as in the following. We take p and k_1 fixed to $p = 3 \times 10^{-2}$ and $k_1 = 780$, whereas we set as $[10, 150]$ the minimisation interval for θ . We test several initial random values for $\theta \in [30, 55]$. Setting the regularisation parameter $\lambda = 10^{-9}$, in Figure 3 we show the minimisation results.

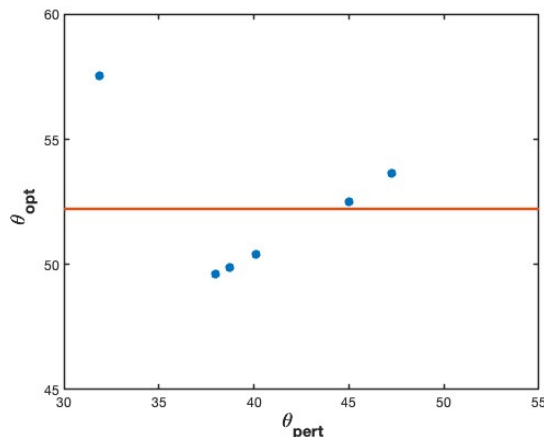


Figure 3: The optimal parameter θ found varying the initial perturbed one in the interval $[30, 55]$. We have selected 6 representative values and we can observe that the optimal θ is well confined in a small interval. Taking the mean value, we have that $\theta_{\text{opt}} = 52.23$.

The relative error in (6) between the experimental solution and the numerical one with estimated parameter θ for fixed p and k_1 is always around 0.4, which is quite good when dealing with biological data. In this case, as for the functional, we consider Expression (6) for the experimental times considered, *i.e.* every 4 hours in the interval $[0, 72]$ hours.

In Figure 5, we show an example of the numerical solution for the cells density (on the left) compared with the experimental one represented through the particle to density transformation (on the right). We have considered the numerical solution u for Model (2) with the mean optimal parameter $\theta = 52.23$, and with $p = 3 \times 10^{-2}$ and $k_1 = 780$. In Figure 6, we also propose a comparison between the numerical solution of the gelatin degradation (on the left) and the experimental data (on the right) at the final time. In this case, the fitting is much different, since apparently not all cells have produced the enzymes that degrade the gelatin and it was not possible to quantify gelatin consumption from experiments.

4 Conclusions and perspectives

In this work, we set up a preliminary mathematical model to describe gelatin degradation as a result of the action of MMPs enzymes produced by tumor cells. This process is strictly related to cancer cells invasion through a thin membrane. The model presented describes the evolution in time of tumor cells, MMPs enzymes and gelatin degradation through a macroscopic view.

The model is combined with a sensitivity analysis and a parameter estimation. Sensitivity analysis highlights how perturbation on parameters does not affect solutions of the PDE system. Indeed, it confirms the stability of our model. Then, we have performed parameter estimation both on synthetic data and on biological ones. Despite the lack of a complete biological knowledge, we were able to build an appropriate description of degradation experiments. As it can be observed in Figure 5, a very attractive perspective is to test our model looking at cells in the entire experimental domain and on different set of data. Indeed, we could both better capture the diffusive behaviour given by our model and have more information on enzymes activity and degradation, the two main poorer information that influenced the research of the complete optimal set of parameters. A deeper study on how enzymes are produced by tumor cells could be helpful, since in the experimental data proposed only a small amount of holes has appeared, see Figure 6.

More complex models considering the action of other cells involved in the invasion process, such as fibroblasts [24] could be very interesting to analyse. Moreover, as stressed during the paper, the gelatin is a very thin layer. A two-dimensional approach is then required. However, this work could be a first step in studying not only basal membrane degradation, but also ECM one. ECM is in fact a thicker layer in which cells move thanks to degradation. Furthermore, the understanding of the degradation process is crucial in modelling invasion. Indeed, the estimated parameters helps to quantify membrane permeability inside the Kedem-Katchalsky membrane conditions. Finally, a more mathematical analysis on Model (2) can reveal interesting behaviour of solutions.

Acknowledgments

G.C. has received fundings from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 740623). The work of G.C. was also partially supported by GNAMPA-INdAM. The authors are grateful to Elishan Christian Braun for fruitful discussions, concerning the numerical implementation of the parameter estimation method. Finally, authors are very grateful to reviewers for their appropriate and constructive remarks.

Statements and Declarations

Competing Interests The authors declare that they have no competing interests.

Data availability

All data supporting the findings of this study are available within the paper and its Supplementary Information.

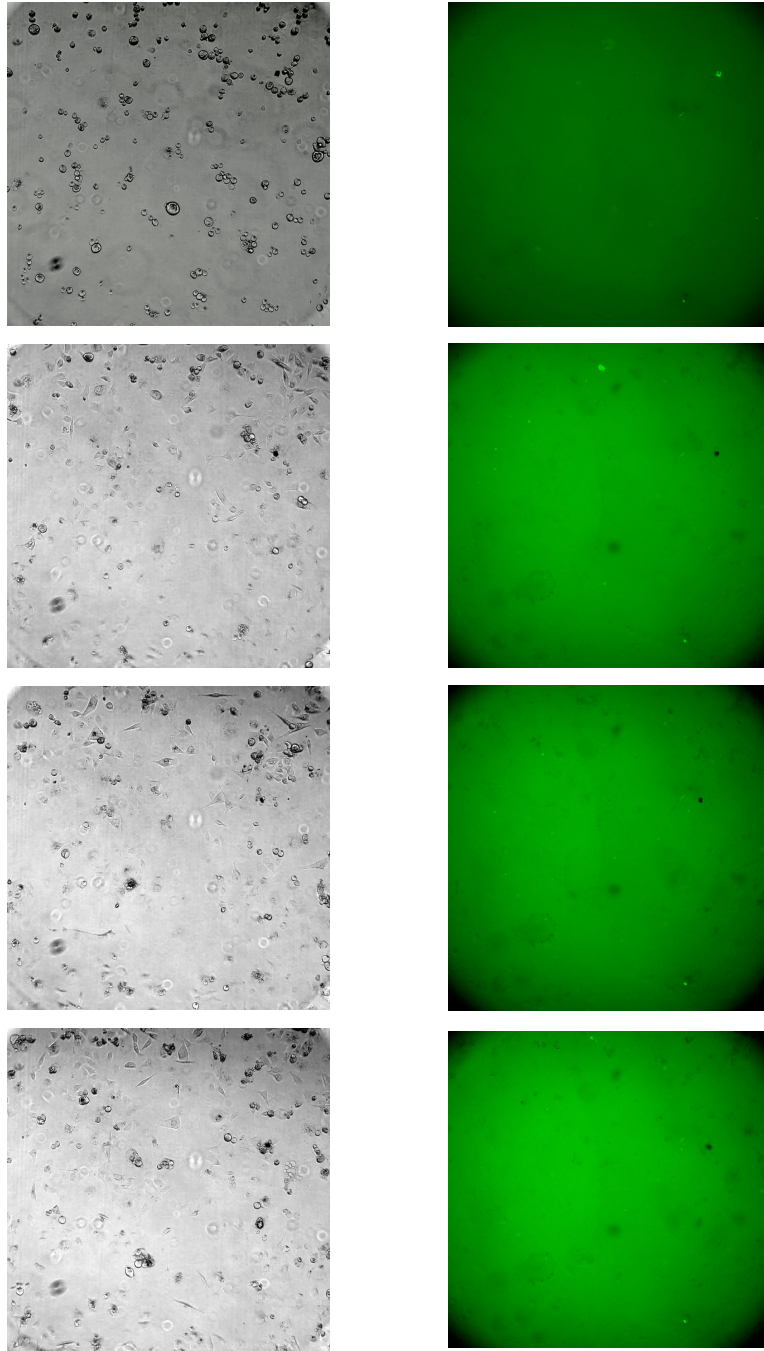


Figure 4: *Example of experimental data. On the left, cells position from the initial time to 72 hours, pictured every day, on the right the corresponding gelatin consumption. Gelatin is in green, whereas darker regions indicate its degradation.*

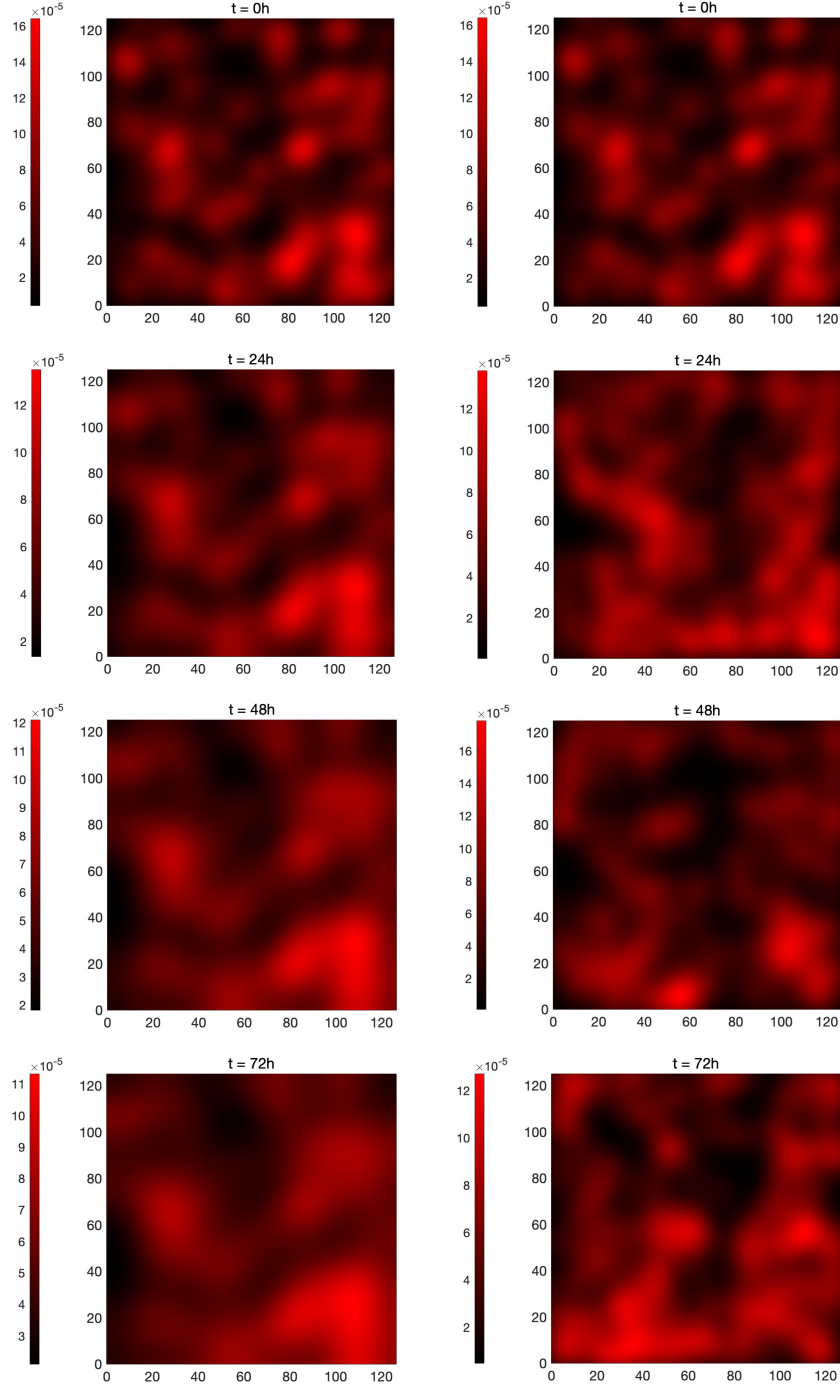


Figure 5: Comparison between our numerical cell density with optimal parameter θ (on the left) and the experimental data after transformation into densities (on the right). The domain dimension is in the corresponding dimensionless coordinates of the experimental one. In the first row, the initial data that is the same for both. The second row represents the density after one day, the third one after two days, whereas the last one after three days. We can observe some similarities between the two columns even if the match is not complete, but the error is below 0.4.

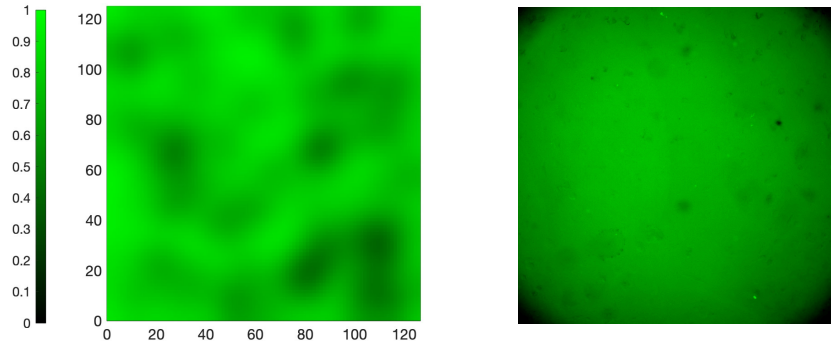


Figure 6: Gelatin degradation after 72 hours. On the left the results obtained through our simulations, and on the right the experimental data.

References

- [1] N. Aceto, A. Bardia, D. T. Miyamoto, M. C. Donaldson, B. S. Wittner, J. A. Spencer, M. Yu, A. Pely, A. Engstrom, H. Zhu, et al. Circulating tumor cell clusters are oligoclonal precursors of breast cancer metastasis. *Cell*, 158(5):1110–1122, 2014. URL <https://doi.org/10.1016/j.cell.2014.07.013>.
- [2] D. Aregba-Driollet, F. Diele, and R. Natalini. A mathematical model for the sulphur dioxide aggression to calcium carbonate stones: Numerical approximation and asymptotic analysis. *SIAM Journal on Applied Mathematics*, 64(5):1636–1667, 2004. URL <https://doi.org/10.1137/S003613990342829X>.
- [3] V. V. Artym, Y. Zhang, F. Seillier-Moiseiwitsch, K. M. Yamada, and S. C. Mueller. Dynamic interactions of cortactin and membrane type 1 matrix metalloproteinase at invadopodia: defining the stages of invadopodia formation and function. *Cancer Res.*, 66(6):3034–3043, 2006. URL <https://doi.org/10.1158/0008-5472.CAN-05-2177>.
- [4] E. C. Braun. *Organs-On-Chips: mathematical modelling and parameter estimation*. PhD thesis, Università degli Studi di Roma Tre, 2021. URL <http://www.matfis.uniroma3.it/Allegati/Dottorato/TESI/ecbraun/Thesis%20Braun%20Revision.pdf>.
- [5] E. C. Braun, G. Bretti, and R. Natalini. Parameter estimation techniques for a chemotaxis model inspired by cancer-on-chip (coc) experiments. *Int. J. Non Linear Mech.*, 140:103895, 2022. URL <https://doi.org/10.1016/j.ijnonlinmec.2021.103895>.
- [6] F. Bubba, C. Pouchol, N. Ferrand, G. Vidal, L. Almeida, B. Perthame, and M. Sabbah. A chemotaxis-based explanation of spheroid formation in 3d cultures of breast cancer cells. *J. Theor. Biol.*, 479:73–80, 2019. URL <https://doi.org/10.1016/j.jtbi.2019.07.002>.
- [7] M. A. J. Chaplain, C. Giverso, T. Lorenzi, and L. Preziosi. Derivation and application of effective interface conditions for continuum mechanical models of cell invasion through thin membranes. *SIAM J. Appl. Math.*, 79(5):2011–2031, 2019. ISSN 0036-1399. doi: 10.1137/19M124263X. URL <https://doi.org/10.1137/19M124263X>.
- [8] W.-T. Chen and S. Singer. Fibronectin is not present in the focal adhesions formed between normal cultured fibroblasts and their substrata. *Proc. Natl. Acad. Sci.*, 77(12):7318–7322, 1980. URL <https://doi.org/10.1073/pnas.77.12.7318>.

- [9] W.-T. Chen, K. Olden, B. A. Bernard, and F. F. Chu. Expression of transformation-associated protease (s) that degrade fibronectin at cell contact sites. *J. Cell. Biol.*, 98(4): 1546–1555, 1984. URL <https://doi.org/10.1083/jcb.98.4.1546>.
- [10] W.-T. Chen, J.-M. Chen, S. J. Parsons, and J. T. Parsons. Local degradation of fibronectin at sites of expression of the transforming gene product pp60src. *Nature*, 316(6024):156–158, 1985. URL <https://doi.org/10.1038/316156a0>.
- [11] G. Ciavolella, N. David, and A. Poulain. Effective interface conditions for a porous medium type problem. *to appear in Interface Free Bound*, 2023. URL <https://arxiv.org/abs/2105.02063>.
- [12] L. Connolly and P. Maxwell. Image analysis of transwell assays in the assessment of invasion by malignant cell lines. *Br. J. Biomed. Sci.*, 59(1):11–14, 2002. URL <https://doi.org/10.1080/09674845.2002.11783627>.
- [13] E. Di Costanzo, V. Ingangi, C. Angelini, M. F. Carfora, M. V. Carriero, and R. Natalini. A macroscopic mathematical model for cell migration assays using a real-time cell analysis. *PLoS One*, 11(9):e0162553, 2016. URL <https://doi.org/10.1371/journal.pone.0162553>.
- [14] H. Dillekås, M. S. Rogers, and O. Straume. Are 90% of deaths from cancer caused by metastases? *Cancer med.*, 8(12):5574–5576, 2019. URL <https://doi.org/10.1002/cam4.2474>.
- [15] L. C. Franssen, T. Lorenzi, A. E. Burgess, and M. A. Chaplain. A mathematical framework for modelling the metastatic spread of cancer. *Bull. Math. Biol.*, 81(6):1965–2010, 2019. URL <https://doi.org/10.1007/s11538-019-00597-x>.
- [16] P. Friedl and K. Wolf. Plasticity of cell migration: a multiscale tuning model. *J. Cell Biol.*, 188(1):11–19, 2010. URL <https://doi.org/10.1083/jcb.200909003>.
- [17] O. Gallinato, T. Colin, O. Saut, and C. Poignard. Tumor growth model of ductal carcinoma: from in situ phase to stroma invasion. *J. Theoret. Biol.*, 429:253–266, 2017. URL <https://doi.org/10.1016/j.jtbi.2017.06.022>.
- [18] C. Giverso, T. Lorenzi, and L. Preziosi. Effective interface conditions for continuum mechanical models describing the invasion of multiple cell populations through thin membranes. *Appl. Math. Lett.*, 125:107708, 2022. URL <https://doi.org/10.1016/j.aml.2021.107708>.
- [19] F. Guarguaglini and R. Natalini. Fast reaction limit and large time behavior of solutions to a nonlinear model of sulphation phenomena. *Communications in Partial Differential Equations*, 32(2):163–189, 2007. URL <https://doi.org/10.1080/03605300500361438>.
- [20] N. Harbeck, F. Penault-Llorca, J. Cortes, M. Gnant, N. Houssami, P. Poortmans, K. Ruddy, J. Tsang, and F. Cardoso. Breast cancer. *Nat. Rev. Dis. Primers*, 5(1):1–31, 2019. URL <https://doi.org/10.1038/s41572-019-0111-2>.
- [21] Y. Hong, F. Fang, and Q. Zhang. Circulating tumor cell clusters: What we know and what we expect. *Int. J. Oncol.*, 49(6):2206–2216, 2016. URL <https://doi.org/10.3892/ijo.2016.3747>.

- [22] L. Kachanov. *Introduction to continuum damage mechanics*, volume 10. Springer Science & Business Media, 1986. URL <https://doi.org/10.1007/978-94-017-1957-5>.
- [23] N. Ke, X. Wang, X. Xu, and Y. A. Abassi. The xcelligence system for real-time and label-free monitoring of cell viability. In *Mammalian cell viability*, pages 33–43. Springer, 2011. URL https://doi.org/10.1007/978-1-61779-108-6_6.
- [24] N. Malaquin, C. Vercamer, F. Bouali, S. Martien, E. Deruy, N. Wernert, M. Chwastyniak, F. Pinet, C. Abbadie, and A. Pourtier. Senescent fibroblasts enhance early skin carcinogenic events via a paracrine mmp-par-1 axis. *PloS One*, 8(5):e63607, 2013. URL <https://doi.org/10.1371/journal.pone.0063607>.
- [25] K. H. Martin, K. E. Hayes, E. L. Walk, A. G. Ammer, S. M. Markwell, and S. A. Weed. Quantitative measurement of invadopodia-mediated extracellular matrix proteolysis in single and multicellular contexts. *J. Vis. Exp.*, (66):e4119, 2012. URL <https://dx.doi.org/10.3791/4119>.
- [26] J. Martinez-Serra, A. Gutierrez, S. Muñoz-Capó, M. Navarro-Palou, T. Ros, J. C. Amat, B. Lopez, T. F. Marcus, L. Fueyo, A. G. Suquia, et al. xcelligence system for real-time label-free monitoring of growth and viability of cell lines from hematological malignancies. *OncoTargets Ther.*, 7:985–994, 2014. URL <https://doi.org/10.2147/OTT.S62887>.
- [27] K. W. Morton and D. F. Mayers. *Numerical solution of partial differential equations: an introduction*. Cambridge university press, 2005. URL <https://doi.org/10.1017/CB09780511812248>.
- [28] J. D. Murray. *Mathematical Biology I. An Introduction*. Springer, 2002. URL <https://doi.org/10.1007/b98868>.
- [29] A. Obr, P. Röselová, D. Grebeňová, and K. Kuželová. Real-time monitoring of hematopoietic cell interaction with fibronectin fragment: The effect of histone deacetylase inhibitors. *Cell Adh. Migr.*, 7(3):275–282, 2013. URL <https://doi.org/10.4161/cam.24531>.
- [30] A. Quarteroni, R. Sacco, and F. Saleri. *Numerical mathematics*. Springer Science & Business Media, 2010. URL <https://doi.org/10.1007/b98885>.
- [31] A. Saltelli, M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, M. Saisana, and S. Tarantola. *Global sensitivity analysis. The Primer*. John Wiley & Sons, 2008. URL <https://doi.org/10.1002/9780470725184>.
- [32] L. A. Segel. Simplification and scaling. *SIAM review*, 14(4):547–571, 1972. URL <https://doi.org/10.1137/1014099>.
- [33] H. Sung, J. Ferlay, R. L. Siegel, M. Laversanne, I. Soerjomataram, A. Jemal, and F. Bray. Global cancer statistics 2020: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.*, 71(3):209–249, 2021. URL <https://doi.org/10.3322/caac.21660>.
- [34] L. Türker Şener, G. Albeniz, B. Dinç, and I. Albeniz. icelligence real-time cell analysis system for examining the cytotoxicity of drugs to cancer cell lines. *Exp. Ther. Med.*, 14(3):1866–1870, 2017. URL <https://doi.org/10.3892/etm.2017.4781>.
- [35] A. G. Waks and E. P. Winer. Breast cancer treatment: a review. *Jama*, 321(3):288–300, 2019. URL <https://doi.org/10.1001/jama.2018.19323>.

- [36] M. Zaoui, M. Morel, N. Ferrand, S. Fellahi, J.-P. Bastard, A. Lamazière, A. K. Larsen, V. Béréziat, M. Atlan, and M. Sabbah. Breast-associated adipocytes secretome induce fatty acid uptake and invasiveness in breast cancer cells via cd36 independently of body mass index, menopausal status and mammary density. *Cancers*, 11(12):2012, 2019. URL <https://doi.org/10.3390/cancers11122012>.

Supplementary Materials

S5 Discretisation

We illustrate the two-dimension numerical method, Morton and Mayers [S27], Quarteroni *et al.* [S30]. We present the discretisation of the two-dimensional System (2) on the domain $\Omega = [a, b] \times [c, d]$.

Concerning space discretisation, we consider a mesh over Ω such that $\Delta x = \frac{b-a}{N_x+1} = \Delta y$. Thus, we divide $[a, b]$ into $N_x \in \mathbb{N}$ intervals such that a (respectively, b) corresponds to $j = 1$ ($j = N_x + 1$), and $[c, d]$ into $N_y \in \mathbb{N}$ intervals such that c (respectively, d) corresponds to $i = 1$ ($i = N_y + 1$). The mesh is formed by the intervals

$$J_j = \left(x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}}\right), \quad j = 1, \dots, N_x + 1, \quad I_i = \left(y_{i-\frac{1}{2}}, y_{i+\frac{1}{2}}\right), \quad i = 1, \dots, N_y + 1.$$

The intervals are centred in $x_j = j\Delta x$, $j = 1, \dots, N_x + 1$ and $y_i = i\Delta y$, $i = 1, \dots, N_y + 1$. Moreover, we add ghost points to build the extremal intervals centred at the boundaries for $x = a, b$ and $y = c, d$. At a given time, the spatial discretisation of $u(t, x, y)$ (the same for functions m and d), interpreted in the finite volume sense, is of the form

$$u_{i,j}(t) \approx \frac{1}{\Delta x} \frac{1}{\Delta y} \int_{J_j} \int_{I_i} u(t, x, y) dx dy, \quad \text{for } j = 1, \dots, N_x + 1, i = 1, \dots, N_y + 1.$$

For the time discretisation, we consider a time step Δt , and set $t^n = n\Delta t$, with $n \in \mathbb{N}$. The discrete approximation of $u(t, x, y)$ (or of functions m or d), for $n \in \mathbb{N}$, $j = 1, \dots, N_x + 1$, $i = 1, \dots, N_y + 1$ is now

$$u_{i,j}^n \approx \frac{1}{\Delta x} \frac{1}{\Delta y} \int_{J_j} \int_{I_i} u(t^n, x, y) dx dy.$$

Equation for d. We solve $\partial_t d = \frac{1}{p} m \left(1 - \frac{1}{p} d\right)$. We discretise the solution $d = \frac{1}{p} \left(1 - e^{\int_0^t m ds}\right)$, $\forall x, y \in \Omega$, using the trapezoidal rule for the integral

$$\int_0^t m ds \approx \frac{\Delta t}{2} \left[m_{i,j}^n + 2 \sum_{k=1}^{n-1} m_{i,j}^k \right],$$

for $j = 1, \dots, N_x + 1$, $i = 1, \dots, N_y + 1$. We remember that initially $d_{i,j}^0 = 0$, $\forall i = 1, \dots, N_y + 1, j = 1, \dots, N_x + 1$, since the gelatin is not damaged.

Discretisation for D(d). We analyse separately the diffusion coefficient for cell density u . We recall the expression in (2), namely $D(d) = \theta + d$ in the parametrised version. Its discretisation is easy, but we have to deal with ghost points that appear in the discretisation for u . We derived boundary conditions which preserve the mass in our system. Namely, we obtain

$$\begin{aligned} D(d)_{i,0}^n &= D(d)_{i,2}^n, & D(d)_{i,N_x+2}^n &= D(d)_{i,N_x}^n, \\ D(d)_{0,j}^n &= D(d)_{2,j}^n, & D(d)_{N_y+2,j}^n &= D(d)_{N_y,j}^n. \end{aligned} \tag{S1}$$

Equation for u. To achieve the time discretisation, we adopt an Euler method, and we write $\frac{du}{dt}(t)$ as $\frac{u^{n+1} - u^n}{\Delta t}$. To obtain space discretisation, we recall the following explicit one-dimensional method

$$\partial_x(a(x)\partial_x u) \approx \frac{(a_{j+1} + a_j)(u_{j+1} - u_j) - (a_{j-1} + a_j)(u_j - u_{j-1})}{2\Delta x^2}.$$

In the vector analysis, we can use this one-dimensional approximation on each axis derivative. Then, full discretisation of System (2) reads for $n \in \mathbb{N}, j = 1, \dots, N_x + 1, i = 1, \dots, N_y + 1$ as

$$\begin{aligned} u_{i,j}^{n+1} - u_{i,j}^n = & \mu[(D_{i-1,j}^n + D_{i,j}^n)u_{i-1,j}^n + (D_{i+1,j}^n + D_{i,j}^n)u_{i+1,j}^n + \\ & -(D_{i-1,j}^n + D_{i+1,j}^n + 4D_{i,j}^n + D_{i,j-1}^n + D_{i,j+1}^n)u_{i,j}^n + (D_{i,j-1}^n + D_{i,j}^n)u_{i,j-1}^n + \\ & + (D_{i,j+1}^n + D_{i,j}^n)u_{i,j+1}^n] + k_2 u_{i,j}^n \left(1 - \frac{u_{i,j}^n}{\alpha_3}\right), \end{aligned} \quad (\text{S2})$$

with $\mu_u = \frac{\Delta t}{2\Delta x^2}$. Finally, we deduce the system

$$\begin{aligned} u_{i,j}^{n+1} = & [1 - \mu(D_{i-1,j}^n + D_{i+1,j}^n + 4D_{i,j}^n + D_{i,j-1}^n + D_{i,j+1}^n)]u_{i,j}^n + \\ & + \mu(D_{i-1,j}^n + D_{i,j}^n)u_{i-1,j}^n + \mu(D_{i+1,j}^n + D_{i,j}^n)u_{i+1,j}^n + \mu(D_{i,j-1}^n + D_{i,j}^n)u_{i,j-1}^n + \\ & + \mu(D_{i,j+1}^n + D_{i,j}^n)u_{i,j+1}^n + k_2 u_{i,j}^n \left(1 - \frac{u_{i,j}^n}{\alpha_3}\right), \end{aligned} \quad (\text{S3})$$

with second order discretisation of the Neumann boundary conditions in $x = a, b$, and $y = c, d$ as

$$u_{i,0}^n = u_{i,2}^n, \quad u_{i,N_x+2}^n = u_{i,N_x}^n, \quad u_{0,j}^n = u_{2,j}^n, \quad u_{N_y+2,j}^n = u_{N_y,j}^n. \quad (\text{S4})$$

Previous conditions give the relations of the extremal ghost points. Substituting ghost points Relations (S1), (S4) into Equation (S3), we obtain the discretised equation for u of System (2) defined on the spatial grid.

Equation for m . Instead of the equation for m , in order to avoid stiffness problems due to the presence of the term $-m$, we discretise the equation for $w = e^t m$ which is of the form $\partial_t w = \Delta w + k_1 u(1 - pd)e^t$. At the end, the discretised density for the enzymes m can be derived from the numerical solution w , as $m_{i,j}^n = e^{-t^n} w_{i,j}^n$. We infer that

$$w_{i,j}^{n+1} = (1 - 4\mu_m)w_{i,j}^n + \mu_m(w_{i-1,j}^n + w_{i+1,j}^n + w_{i,j-1}^n + w_{i,j+1}^n) + k_1 u_{i,j}^n(1 - pd_{i,j}^n)e^{n\Delta t}, \quad (\text{S5})$$

with $\mu_m = \frac{\Delta t}{\Delta x^2}$. Boundary conditions are the same as in (S4).

Calling the vector solutions at time t^n as

$$U^n = (u_{i,1}^n, \dots, u_{i,N_x+1}^n)^T, \quad W^n = (w_{i,1}^n, \dots, w_{i,N_x+1}^n)^T,$$

and the reaction vectors as

$$R^n = k_2 U^n \left(1 - \frac{U^n}{\alpha_3}\right), \quad S^n = e^{\alpha n \Delta t} \beta U^n,$$

we can write the discretised systems in a matrix form as $U^{n+1} = AU^n + \Delta t R^n$ coupled with $W^{n+1} = BW^n + \Delta t S^n$. Coefficients of the matrices A and B can be found substituting ghost points in the discretised equations. In particular, B is the standard matrix related to the heat equation.

It is important to guarantee the positiveness of the coefficients in the previous Equations (S3), (S5), in order to preserve positiveness and stability. This brings conditions on the time interval Δt which has to be such that

$$\Delta t < \min(dt_1, dt_2), \quad \text{where} \quad dt_1 = \frac{\Delta x^2}{4(\theta + 1)}, \quad dt_2 = \frac{\Delta x^2}{4}, \quad (\text{S6})$$

since $dt_1 = \min\left(\frac{2\Delta x^2}{8\theta + (d_{i-1,j}^n + d_{i+1,j}^n + 4d_{i,j}^n + d_{i,j-1}^n + d_{i,j+1}^n)}\right) = \left(\frac{2\Delta x^2}{8\theta + \max(d_{i-1,j}^n + d_{i+1,j}^n + 4d_{i,j}^n + d_{i,j-1}^n + d_{i,j+1}^n)}\right)$, and the maximum value for d is 1.