POLITECNICO DI TORINO
Repository ISTITUZIONALE

Two-step global sensitivity analysis of a non-local integro-differential model for Cancer-on-Chip experiments

Original
Two-step global sensitivity analysis of a non-local integro-differential model for Cancer-on-Chip experiments / Campanile, Elio; Colombi, Annachiara; Bretti, Gabriella. - In: MATHEMATICAL BIOSCIENCES. - ISSN 0025-5564. - 378:(2024), pp. 1-17. [10.1016/j.mbs.2024.109330]

Availability:
This version is available at: 11583/2994112 since: 2024-11-04T02:14:13Z

Publisher:
Elsevier

Published
DOI:10.1016/j.mbs.2024.109330

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright
Elsevier preprint/submitted version

Preprint (submitted version) of an article published in MATHEMATICAL BIOSCIENCES © 2024, http://doi.org/10.1016/j.mbs.2024.109330

(Article begins on next page)

07 May 2026

Two-step global sensitivity analysis of a non-local integro-differential model for Cancer-on-Chip experiments

Elio Campanile[a,b], Annachiara Colombi[c,*], Gabriella Bretti[d]

[a]*Fondazione the Microsoft Research, University of Trento, Centre for Computational and Systems Biology (COSBI), Piazza Manifattura 1, Rovereto, 38068, Italy,*
[b]*Department of Mathematics, University of Trento, Via Calepina, 14, Trento, 38122, Italy,*
[c]*Department of Mathematical Sciences (DISMA) Politecnico di Torino, DISMA, C.so Duca degli Abruzzi 24, Torino, 10129, Italy,*
[d]*Istituto per le Applicazioni del Calcolo, CNR, Via dei Taurini 19, Rome, 00185, Italy,*

Abstract

The present work focuses on a non-local integro-differential model reproducing Cancer-on-chip experiments where tumor cells, treated with chemotherapy drugs, secrete chemical signals stimulating the immune response. The reliability of the model in reproducing the phenomenon of interest is investigated through a global sensitivity analysis, rather than a local one, to have global information about the role of parameters, and by examining potential non-linear effects in greater detail.
Focusing on a region in the parameter space, the effect of 13 model parameters on the *in silico* outcome is investigated by considering 11 different target outputs, properly selected to monitor the spatial distribution and the dynamics of immune cells along the period of observation. In order to cope with the large number of model parameters to be investigated and the computational cost of each numerical simulation, a two-step global sensitivity analysis is performed. First, the screening Morris method is applied to rank the effect of the 13 model parameters on each target output and it emerges that all the output targets are mainly affected by the same 6 parameters. The extended Fourier Amplitude Sensitivity Test (eFAST) method is then used to quantify the role of these 6 parameters.
As a result, the proposed analysis highlights the feasibility of the considered space of parameters, and indicates that the most relevant parameters are those related to the chemical field and cell-substrate adhesion. In turn, it suggests how to possibly improve the model description as well as the calibration procedure, in order to better capture the observed phenomena and, at the same time, reduce the complexity of the simulation algorithm. On one hand, the model could be simplified by neglecting cell-cell alignment effects unless clear empirical evidences of their importance emerge. On the other hand, the best way to increase the accuracy and reliability of our model predictions would be to have experimental data/information to reduce the uncertainty of the more relevant parameters.

Keywords: discrete and continuous mathematical model, cancer-on-chip, global sensitivity analysis, Morris method, eFAST method
2020 MSC: : 92-10, 90C31, 92C50, 49K40

1. Introduction

Mathematical modeling is a powerful asset to explore biological phenomena and thus support advances, for instance, in medicine, biology and biotechnology. Among the new technologies, an increasingly important role is played by Organs-on-Chips (OoCs), see among others [20, 21, 43]. OoCs are bio-engineered microfluidic chips designed to simulate activities of an organ or an organ system, thanks to their partition in different areas, called compartments, connected to each other by micro-channels. Among OoC categories, Cancer-on-chips (CoCs) are used to observe micro-environmental factors that influence tumor cells in response to anticancer therapies [2, 12, 13, 14, 18, 19, 29]. In the last decades, laboratory experiments based on these microfluidic devices generated an extensive collection of images and data. Some mathematical-based *in silico* models have been then created to replicate these biological *in vitro* experiments in order to gain more insights on cellular behaviors [30, 31, 60, 61, 62].

In this context, the aspect of our interest is the dynamics of immune cells (ICs) in response to chemical signals secreted by tumor cells (TCs), accounting also for cell-cell and cell-substrate mechanical interactions. With this aim in mind, the first *in silico* model developed by one of the authors on CoC experiments is [33], where a mathematical model based on reaction-diffusion equations with chemotaxis inspired by Keller-Segel model [17] was proposed. Such a model describes cell death processes, effects of chemoattractants, interactions, and competition between different cell species. In [34] techniques for estimating parameters in the model [33] have been proposed. A discrete-in-continuous hybrid approach has also been formulated by the same author as a PDE reaction-diffusion partial model for the evolution of the chemicals, coupled with an ODE particle model for cell motion, see [15, 16]. Well-posedness and asymptotic behavior of solutions for this class of hybrid coupled system have been studied in [36, 37, 38]. Recently, the mean-field limit of a general class of deterministic hybrid macro-micro models was proposed in [39]. Successively, an agent-based model for the chip environment dynamics based on cellular automata approach was developed and the related sensitivity analysis of model parameters was carried out in [32, 35], respectively.

Despite the significant efforts in literature, an accurate model to describe the immune response in tumor microenvironment is still a challenge. One of the main aspects to face is the calibration of model parameters, since they may be difficult to measure with

*annachiara.colombi@polito.it

precision, or even impossible to estimate *in vitro*. This may happen either due to technical issues or when model parameters do not represent measurable physical quantities. It is further worth mentioning that mathematical models implementing complex biological phenomena with interconnected processes, possibly occurring at different time and spatial scales, may depend on a large number of parameters.

In this context, a proper sensitivity analysis (SA) of the mathematical model is crucial. It indeed evaluates the impact of variations in the parameter values on the model predictions, and possibly highlights how different model components interplay [6, 42]. SA thus provides insights about the accuracy of the model and the considered parameter settings in reproducing the phenomenon of interest, stating therefore both the robustness of the model and the reliability of the model calibration. Moreover, SA outcomes can be used to identify possible model improvements or simplifications; drive the calibration of the model in different scenarios; or even suggest experimental research priorities to reduce model uncertainty. In this perspective, the sensitivity analysis should be part of the modelling process [40].

In this respect, in several of the above-cited works by Bretti et al. on CoCs [16, 32, 34], the proposed mathematical models have been coherently provided by sensitivity analysis to shed light on the role of the parameters on *in silico* outcomes. However, they have always performed a so-called *local sensitivity analysis*, i.e., they have investigated the effect of small perturbations in the parameter values one at a time around a fixed (nominal) setting. Although easy to use and demanding limited computational resources, local sensitivity analysis methods have important limitations [65, 66]. Being based on the assumption of independence between the model parameters, and the limited perturbation of parameter values, they may lead to heavily biased results in the case of nonlinear models. In addition, the importance of parameters will be underestimated in the case of interactions between model parameters, i.e. if the effect of a parameter depends on the value of others parameters and thus the parameter effects are not additive. A different and more challenging approach is the so-called *global sensitivity analysis* (GSA) [6, 41, 42]. On one hand, methods of GSA are based on the full exploration of a given, even large or infinite, region of the space of parameters, by means of proper sampling methods. On the other hand, the effect of variations in each model parameter on the outcome is regarded globally, i.e. averaged over variations of the other parameters. In addition, GSA methods are able to deal with nonlinear effects and highlight interactions between the model parameters. In turn, GSA methods allow us to rank model parameters from the most to the least globally affecting the evolution of the system, thereby pointing out the main factors/processes.

In this regard, in the present work we perform GSA of a slight variation on recent non-local integro-differential models proposed in [15, 16]. Our aim is to gain a more complete overview of the potentials and critical issues of this modelling framework in capturing the biological COC experiments by Vacchelli et al. [14]. In particular, among

their experiments, we here focus on those performed with tumor cells that, having been treated with a chemotherapy drug, undergo apoptosis by secreting chemical signals stimulating the response of immune cells. From a mathematical point of view, in [15, 16], cells are described as discrete point-wise entities, while a continuous approach is adopted for the chemical signal. The latter evolves according to a reaction-diffusion equation provided by a non-local source term related to the position of TCs. For simplicity, the apoptotic cells are here assumed static. The dynamics of ICs are conversely given by a second-order differential equation including non-local terms implementing cell-cell and cell-chemical interactions. In the present work, inspired by [26], a Cucker and Smale-like alignment term [27] is included in the equation for ICs dynamics, as in [16].

The GSA is here applied to investigate the role of 13 model parameters, i.e. all those involved in the equations for either the ICs or the chemical field apart from cell radii whose values are taken from [24, 25]. In order to study how these 13 parameters affect ICs dynamics, our GSA is performed by focusing on a representative scenario designed to mimic the experiments in [14], avoiding excessively high computational costs. According to [15, 16], the spatial domain reproduces a small portion of the CoC device monitored in [14] and proper boundary conditions for the chemical field reproduce chemical-secreting TCs placed outside the domain. For simplicity, a randomly generated spatial distribution of the TCs inside the domain and an ICs inflow consistent with data in [14] are fixed equal for all the simulations. The range of values of the 13 model parameters investigated in our GSA are defined on the basis of the pertinent biological and mathematical literature. The choice and the characterization of the outputs to be investigated are crucial to perform a useful sensitivity analysis of a model. In this particular case, IC paths cannot be fully determined by a small number of output values. For this reason, the temporal evolution of ICs distribution and dynamics are described through 11 distinct scalar measures, calculated over 3 disjoint sub-intervals of the period of observation. Some of these measures have been inspired by the statistical paper by Agliari et al. [1].

The GSA is here carried out in two-step in order to cope with the large number of model parameters selected, the non-linearity of the model, and the computational cost of each numerical simulation. First, we apply the Elementary Effects method by Morris [4], in the improved version proposed by Campolongo et al. [3]. It belongs to the class of *screening methods*, i.e., designs conceived to efficiently treat non-linear models with tens or hundreds of parameters that return a rank of the input factors in order of importance (qualitative information), but do not quantify the effect of each parameter nor differentiate among them [6]. The ranks obtained for each output quantity, over each sub-interval, are compared by computing the Spearman's rank correlation coefficient. Interestingly, this allows us to extract a group of 6 parameters mainly responsible for the variability of all the output quantities and investigate them with a more accurate, although more expensive in terms of required number of model realizations, GSA method. Specifically,

we apply the *variance-based method* eFAST (Extended Fourier Amplitude Sensitivity Test) [10, 44, 46, 54, 55], which estimates, by means of spectral analysis, the so-called Sobol sensitivity indices/measures denoting the fraction of the output variance due to each model parameter singularly (*first order/main effects*) or in combination with other parameters (*total effects*).

As detailed in the following sections, the first step of our analysis highlights that (i) the considered region in the 13-dimensional parameter space results in reliable *in silico* scenarios; and (ii) that both cell dynamics and spatial distribution are mainly affected by parameters regulating either the evolution of the chemical field, cell chemotactic sensing/motion or their adhesion to the substrate. The second step of our analysis conversely quantifies how the 6 most relevant parameters affect the variance of the outcome quantities capturing different aspects of system evolution (e.g. cell mean speed, cell clustering around tumor cells and so on) as time goes by. In turn, these additional information identify the parameters, and related processes, included into the model (i.e. chemotaxis, adhesion to the substrate, chemical sensing) that are mainly responsible for the variability of system behavior over time. In particular, it emerges that (i) all the monitored output quantities are strongly affected by the chemical secretion rate and the diffusion coefficient, while (ii) the decay rate has substantially negligible effects; (iii) cell substrate adhesion and chemical sensing have an important effect mainly on the initial evolution of cell distribution and mean speed. On one hand, this approach therefore allows to validate the proposed model in combination with the considered region of space parameter. On the other hand, it suggests which parameter/processes mainly require careful calibration.

It is finally worth clarifying that a two-step sensitivity approach based on the application of the Morris method followed by the estimation of the Sobol indices of the most important parameters is not new in literature. For instance, it is nowadays a consolidated practice in agronomy, to analyze complex crop models (see, among others [47, 48, 49, 50, 51, 52, 53]). On the other hand, as far as we know, there are few works in literature where GSA methods have been used to investigate the behavior of mathematical models for spatial cell dynamics (see, for instance, [56, 57] for ABM models) or for other relevant biological phenomena (see e.g. [58, 59]). Regardless of cell dynamics, an example of application of GSA methods is found in [63] where Sobol's sensitivity indices are used to identify non-influential parameters that can be fixed to reduce the complexity of an ordinary differential equation model reproducing the immune response to an *S. aureus* infection in mice. However, we have not found works where the Morris method coupled with a variance-based method has been applied to study mathematical models for spatial cell dynamics or, more specifically, non-local integro-differential models coupling continuous and discrete descriptions.

The rest of the manuscript is organized as follows. Section 2 is devoted to summarize preliminary knowledge at the basis of our work and to define our two-step GSA approach.

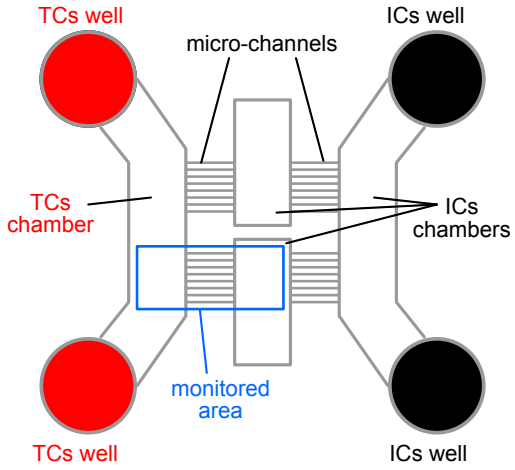

Figure 1: Microfluidic chip environment. Schematic representation of the microfluidic chip architecture.

In detail, Section 2.1 introduces the biological experiments inspiring the model in [15]. The mathematical model here analyzed is described in Section 2.2, while details on the applied numerical scheme are reported in Appendix A. In Section 2.3 the case study and the parameter space to investigate are introduced. Then, the output quantities used to characterize the system dynamics are defined. The main aspects of the standard GSA methods of Morris and eFAST, as well as our strategy to use them, are delineated in Section 2.4 and in Appendix B. Section 3 is devoted to the presentation of obtained results. First, in Section 3.1 a representative numerical simulation, performed with a fixed parameter setting, is described to help readers not familiar with the model to interpret the forthcoming sensitivity analysis. The SA performed with the subsequent application of the Morris and the eFAST methods are presented in Sections 3.2 and then discussed in Section 3.3. Conclusions and future perspectives of the work are highlighted in Section 4.

2. Materials and methods

2.1. The Cancer-on-Chip biological experiment

Biological phenomena inspiring the present work are described in [14, 29], where detailed laboratory settings of in vitro experiments performed on CoC are reported. The immune-oncology chip designed for the experiments presents a complex geometry sketched in Fig. 1. It is composed of cylindrical wells containing, separately, cell cultures of TCs (red circles) or ICs (black circles). These wells are connected by culture chambers (white areas) and microchannels (striped gray areas). Microchannels, having a width and length of 12 μm and 500 μm, respectively, allow both the flow of the chemical and the migration of cells between the chambers. Biologists observed that this type

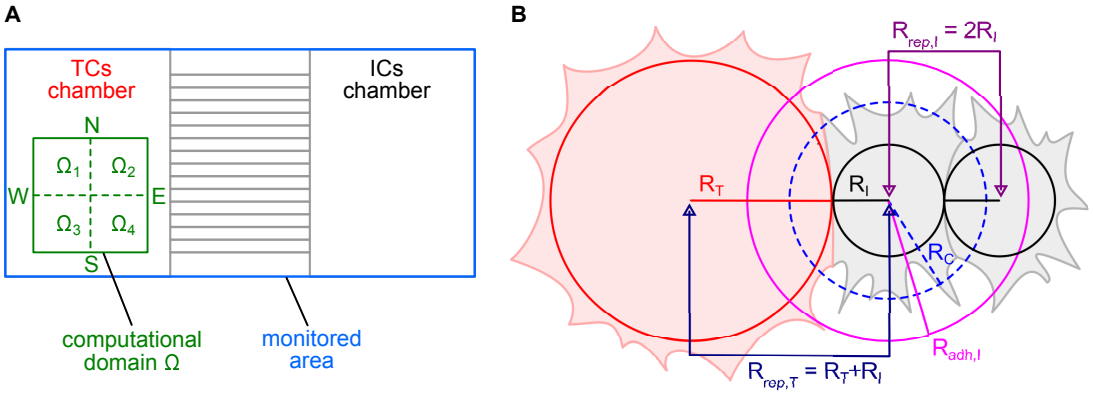

Figure 2: Panel A: Schematic representation of the chip area monitored in the experiments by Vacchelli et al. [14]. The green square area is the computational domain considered in the present work. For the sensitivity analysis, the domain is partitioned in 4 disjoint square areas: Ω_q, with $q = 1, 2, 3, 4$: i.e. Ω_1 (North-West), Ω_2 (North-East), Ω_3 (South-West) and Ω_4 (South-East). The sides of the domain are denoted by N (North), E (East), W (West), S (South). Panel B: Representation of cell morphologies and cell radii used to model cell chemotaxis and cell-cell interactions.

of configuration is able to reproduce quite realistically the physiochemical environment and the mechanical stresses acting on the living cells inside it. It is worth noting that we reported the schematic representation of the chip for the sake of completeness. However, video recordings of the laboratory experiments only focus on the monitored area depicted in blue in Fig. 1.

In particular, among the experimental settings proposed in [14], we focus on the case where TCs, having been treated with a chemotherapy drug, are dying and secrete a chemical signal that acts as an attractant for wild-type ICs. Being apoptotic, TCs are quasi-static, and thus, in addition to the cells placed in the TC wells, experimentalists also planted some cancer cells in the TCs chamber in adherence to the glass slide. The immune population consists of peripheral blood mononuclear cells (PBMCs), i.e. any peripheral blood cell having a round nucleus. It is a heterogeneous population including different cell species: monocytes, dendritic cells, and T and B lymphocytes. However, a classification of the experimental data with respect to the different IC types in the PBMC population is not provided. Note that PBMCs measure on average about 8-10 μm of diameter [25], while TCs measure about 20 μm in diameter [24].

The laboratory experiments in [14] last 48-72 h, and video frames are acquired at a constant rate of 2 min. In detail, the area monitored during the laboratory experiments is a portion of the entire chip (see Fig. 1) consisting of a rectangular area of size 1362 μm $\times$ 1702 μm, which includes part of the TC chamber and the central one, with the microchannels connecting them.

2.2. The mathematical model

The mathematical model derived from [15], which we will analyze in the next section, is here reported. This model gathers the advantages of both microscopic and macroscopic descriptions: cells are treated as discrete entities, whereas the concentration of the chemoattractant is modeled as a continuum, and with our approach we aim at describing short-range interactions between TCs and ICs. It is worth noting that the experiment in [14] involves 2D culture of immune cells in the liquid with TCs adherent to the glass slide and mainly static, while ICs are floating. The video footage of the experiment is recorded on a small portion of the chip and at fixed height, thus registering the dynamics of ICs travelling towards TCs that are in adhesion with the slide. For these reasons, the third spatial dimension in our framework is neglected since a 2D mathematical model it is well-suited to describe the bidimensional trajectories of ICs towards TCs observed in video recordings.

System evolution is then described on a bidimensional domain Ω denoting either a portion or the entire complex geometry of the chip. In particular, according to [15], we hereafter focus on a square region of size 600 μm $\times$ 600 μm in the TCs chamber, see the blue area in the Panel A of Fig. 2. This allows a detailed analysis of short-range interactions and pattern formation in the chip environment. Furthermore, it reduces the computational cost of numerical simulations.

Cells are described as point-wise particles characterized by their positions: $\mathbf{Y}_j \in \Omega$ with $j = 1, ..., N_T$ for static treated TCs; and $\mathbf{X}_i(t) \in \Omega$ with $i = 1, ..., N(t)$ for ICs. The number of ICs $N(t)$ actually present in the domain varies in time due to a continuous inflow and outflow at the boundary. Specifically, having in mind the experiments in [14], we assume that ICs enter in the domain from the top side of the domain, and definitively left the domain as soon as they reach the boundary.

The concentration of the chemical is represented through its spatial distribution $\varphi : \mathcal{I} \times \Omega \mapsto \mathbb{R}_+$, where $\mathcal{I}$ is the time interval $[0, T]$, with $T > 0$, denoting the period of observation. The temporal evolution of the system is then given by

$$\partial_t \varphi = \underbrace{D\,\Delta\varphi}_{\text{diffusion}} + \underbrace{\xi\, F_S(\mathbf{Y})}_{\text{production}} - \underbrace{\eta\,\varphi}_{\text{decay}} \quad \text{in } \Omega \tag{1}$$

$$\ddot{\mathbf{X}}_i = \underbrace{-\zeta\,\dot{\mathbf{X}}_i}_{\substack{\text{substrate}\\ \text{adhesion}}} + \underbrace{\gamma\,\mathbf{F}_C(\mathbf{X}_i, \varphi)}_{\text{chemotaxis}} + \underbrace{\mathbf{F}_T(\mathbf{X}_i, \mathbf{Y})}_{\substack{\text{IC-TC}\\ \text{interactions}}} + \underbrace{\mathbf{F}_I(\mathbf{X})}_{\substack{\text{IC-IC}\\ \text{interactions}}} + \underbrace{\beta\,\mathbf{F}_A(\mathbf{X}, \dot{\mathbf{X}})}_{\text{alignment}} \quad \text{for } i = 1, ..., N(t) \tag{2}$$

where the vectors $\mathbf{X} := [\mathbf{X}_1, ..., \mathbf{X}_{N(t)}]$ and $\mathbf{Y} := [\mathbf{Y}_1, ..., \mathbf{Y}_{N_T}]$ denote the positions of ICs and TCs respectively; $D, \xi, \eta, \zeta, \gamma, \beta$ are positive constants; and F_S, $\mathbf{F}_C$, $\mathbf{F}_T$, $\mathbf{F}_I$, $\mathbf{F}_A$ are suitable functions defined below. Specifically, in the reaction-diffusion equation (1), D is the diffusion coefficient; ξ and η are the default production and decay rate of the chemical, respectively; while the function F_S implements the production of

chemoattractant by TCs. In the second order differential equation (2), ζ is the damping coefficient due to cell adhesion to the substrate. $\mathbf{F}_C$ is the contribution in cell acceleration triggered by the amount of chemoattractant sensed by the cell and γ is said coefficient of chemotactic effect. $\mathbf{F}_T$ and $\mathbf{F}_I$ implement IC interactions with TCs and ICs, respectively; while $\mathbf{F}_A$ is related to IC mutual alignments and β is the alignment coefficient. Hereafter we report the functional forms of F_S, $\mathbf{F}_C$, $\mathbf{F}_T$, $\mathbf{F}_I$, and $\mathbf{F}_A$.

Chemoattractant secretion F_S. In Eq. (1), F_S implements the chemoattractant substance constantly secreted by TCs through their membrane. It is thus defined as

$$F_S(\mathbf{Y}) = \sum_{j=1}^{N_T} f_{\mathcal{B}(\mathbf{Y}_j, R_T)}(\mathbf{x}), \tag{3}$$

where R_T is the radius of a TC (see Panel B in Fig. 2), $\mathcal{B}(\mathbf{y}, r)$ denotes the bidimensional ball with radius r centered at $\mathbf{y}$, and $f_{\mathcal{A}}$ is the identity function of the set $\mathcal{A}$.

Chemotactic response $\mathbf{F}_C$. In Eq. (2), the chemotactic contribution $\mathbf{F}_C$ depends on the chemoattractant sensed by the $i-$th IC in its neighborhood. It is thus computed as a weighted average of the gradient $\nabla\varphi$ over a ball centered at $\mathbf{X}_i$ whose radius R_C denotes the extension of cell sensing extracellular molecules (see again Panel B in Fig. 2). Similar non-local chemotaxis terms have been already proposed in literature, see [67] for a review and references therein. Specifically, $\mathbf{F}_C$ is assumed as

$$\mathbf{F}_C(\mathbf{X}_i, \varphi) = \frac{1}{\mathcal{W}} \int_{\mathcal{B}(\mathbf{X}_i, R_C)} w_i(\boldsymbol{x})\, \chi(\varphi)\, \nabla\varphi(\boldsymbol{x}, t)\, d\mathbf{x}, \tag{4}$$

where the function $w_i : \mathbb{R}^2 \mapsto \mathbb{R}_+$ is a truncated Gaussian weight function implementing the distribution of chemical receptors. It is defined as $w_i(\boldsymbol{x}) = 2^{(1-\|\boldsymbol{x}-\mathbf{X}_i\|/R_C)^2)}$ when $\mathbf{x} \in \mathcal{B}(\mathbf{X}_i, R_C)$ and null otherwise, so that the value of $\mathcal{W} := \int_{\mathcal{B}(\mathbf{X}_i, R_C)} w_i(\boldsymbol{x}) d\mathbf{x}$ is independent on i. It only depends on the value of R_C. Conversely, the function $\chi : \mathbb{R}_+ \mapsto \mathbb{R}_+$, also known as receptor saturation function, imposes a limitation on the maximum chemoattractant local amount, hence reducing cell migration in high concentration areas. According to [28], it is given by

$$\chi(\varphi) = \frac{k_1}{(k_2 + \varphi)^2}, \tag{5}$$

where k_1 represents the cellular drift velocity, while k_2 is the receptor dissociation constant indicating how many molecules are necessary to bind the receptors.

IC-TC interactions $\mathbf{F}_T$. The function $\mathbf{F}_T$ in Eq. (2) implements a repulsion effect affecting the dynamics of the i-th IC whereas it collides with some TCs, i.e. as the distance between their centers is less than $R_{\mathrm{p},T} := R_I + R_T$ (see Panel B in Fig. 2). Assuming that the resulting repulsion term is due to the superposition of pairwise isotropic metric interactions, $\mathbf{F}_T$ is defined as

$$\mathbf{F}_T(\mathbf{X}_i, \mathbf{Y}) = -\omega_{\mathrm{rep},T} \sum_{j\,:\,\mathbf{Y}_j \in \mathcal{B}\left(\mathbf{X}_i, R_{\mathrm{rep},T}\right)} \left(\frac{1}{\|\mathbf{Y}_j - \mathbf{X}_i\|} - \frac{1}{R_{\mathrm{rep},T}}\right) \frac{\mathbf{Y}_j - \mathbf{X}_i}{\|\mathbf{Y}_j - \mathbf{X}_i\|}, \tag{6}$$

where $\omega_{\mathrm{rep},T}$ is said repulsion coefficient.

IC-IC interactions $\mathbf{F}_I$. The function $\mathbf{F}_I$ in Eq. (2) includes attraction-repulsion effects between ICs. In particular, repulsion occurs if the distance between the centers of two ICs is less than $R_{\mathrm{rep},I} := 2R_I$ (see Panel B in Fig. 2). Conversely, attraction, implementing adhesive interaction between cells via filopodia, occurs if the mutual distance is between $R_{\mathrm{rep},I}$ and $R_{\mathrm{adh},I} > R_{\mathrm{rep},I}$ (see Panel B in Fig. 2). Assuming again the superposition of pairwise isotropic metric interactions, $\mathbf{F}_I$ writes

$$\mathbf{F}_I(\mathbf{X}) = \sum_{\substack{j=1 \\ j\neq i}}^{N(t)} \mathcal{K}(\|\mathbf{X}_j - \mathbf{X}_i\|)\, \frac{\mathbf{X}_j - \mathbf{X}_i}{\|\mathbf{X}_j - \mathbf{X}_i\|}, \tag{7}$$

where, given positive constants $\omega_{\mathrm{rep},I}$ (repulsion coefficient) and $\omega_{\mathrm{adh},I}$ (adhesion coefficient), the function $\mathcal{K}$ is assumed as

$$\mathcal{K}(r) = \begin{cases} -\,\omega_{\mathrm{rep},I}\left(\dfrac{1}{r} - \dfrac{1}{R_{\mathrm{rep},I}}\right), & \text{if } r \le R_{\mathrm{rep},I}, \\ \omega_{\mathrm{adh},I}\left(r - R_{\mathrm{rep},I}\right), & \text{if } R_{\mathrm{rep},I} < r \le R_{\mathrm{adh},I}. \end{cases} \tag{8}$$

Note that for both IC-TC interactions in Eq. (6) and IC-IC interactions in Eq. (7), repulsive kernels are assumed to go as $1/r$, as proposed in literature for cell-cell interactions for instance in [26, 64].

Alignment $\mathbf{F}_A$. In Eq. (2), the function $\mathbf{F}_A$ implementing cell alignment is defined as a Cucker and Smale-like flocking term [27]:

$$\mathbf{F}_A(\mathbf{X}, \dot{\mathbf{X}}) = \frac{1}{N_A(t)} \sum_{\substack{j\,:\,j\neq i \\ \mathbf{X}_j \in \mathcal{B}(\mathbf{X}_i, R_{\mathrm{rep},I})}} \left(1 + \frac{\|\mathbf{X}_j - \mathbf{X}_i\|^2}{R_{\mathrm{rep},I}^2}\right)^{-\frac{1}{2}} \left(\dot{\mathbf{X}}_j - \dot{\mathbf{X}}_i\right), \tag{9}$$

where $R_{\mathrm{rep},I}$ has been chosen as a suitable radius of influence based on [26], and $N_A(t) := \mathrm{card}\{j = 1, \ldots, N(t) \,:\, j \neq i,\ \mathbf{X}_j(t) \in \mathcal{B}(\mathbf{X}_i, R_{\mathrm{rep},I})\}$ is the number of ICs

close enough to the i-th cell to allow alignment effects. We remark that the alignment term was first introduced in [27] for birds and applied to morphogenesis in the setting of conditional flocking, see [26]. In [16], this term has been added for the first time to the model proposed in [15] in order to better mimic the IC flocking movements observed in cancer-on-chip experiments and it is explored in depth in through different scenarios showing the effect of the alignment term coupled with the other effects, such as chemotaxis and adhesion-repulsion terms.

The initial condition is given by the position and velocity of ICs at the initial time $t = 0$ h, i.e. $\mathbf{X}(0) = \mathbf{X}_0$ and $\dot{\mathbf{X}}(0) = \mathbf{V}_0$; together with the distribution of the TCs $\mathbf{Y}$ and the initial concentration of the chemoattractant $\varphi(\boldsymbol{x}, 0) = \varphi_0(\boldsymbol{x})$.
As already anticipated, in order to reproduce the experiments described in Section 2.1, an influx of ICs over time is required. For simplicity, given the inflow rate $1/\tau$, at every instant time t' equal to a multiple of τ, a new IC is added to the system by increasing by one the amount of $N(t')$; selecting a position $\mathbf{X}_{N(t')}(t')$ over the top boundary of the domain; and defining an inward directed velocity. On the other hand, since the boundary of the domain does not represent a physical barrier, ICs are always allowed to trespass it and leave the domain. For simplicity, we assume that once they leave the domain, they will not re-enter. In mathematical terms, if the i-th IC enters the domain at time instant t_{in} and moves inside Ω until t_{out}, when it reaches the boundary, i.e., $\mathbf{X}_i(t) \in \Omega$ for any $t \in [t_{\text{in}}, t_{\text{out}})$ and $\mathbf{X}_i(t_{\text{out}}) \in \partial\Omega$, then we state that $\mathbf{X}_i(t) \notin \Omega$ for any $t > t_{\text{out}}$.
Due to the peculiarity of the chosen computational domain Ω, we further assign a non-homogeneous Robin boundary condition to reproduce the flow of the chemoattractant through the boundary:

$$D\frac{\partial \varphi}{\partial \mathbf{n}} + a\varphi = b, \text{ on } \partial\Omega, \tag{10}$$

where a is the local rate of exchange of the chemoattractant with the external environment, and b is the outflow of the chemical. In particular, labelling the four sides of Ω by S (South), E (East), N (North), W (West) (see Panel B in Fig. 2), the values assumed by a and b over them are respectively denoted as a_S, a_E, a_N, a_W and b_S, b_E, b_N, b_W.

2.3. Case study and definition of the observable quantities

Case study and parameter space. Our sensitivity analysis is performed by focusing on a representative numerical setting, denoted hereafter as a *case study*, designed on the domain Ω by fixing the period of observation $\mathcal{I}$, the initial condition, the inflow of ICs, and the Robin boundary conditions, to reproduce the experiments in [14] described in Section 2.1.

Specifically, the period of observation is fixed to $\mathcal{I} = [0, 24]$ h and it represents the interval of time $[24, 48]$ h in the experiment. Indeed, this is the period presenting the more interesting dynamics, with ICs migrating towards the tumor chamber in the left compartment. The initial condition for the chemoattractant is $\varphi_0(\boldsymbol{x}) = 0$ mol μm^{-2} for

any $\boldsymbol{x} \in \Omega$. A fixed number of TCs, given by $N_T = 35$, are randomly distributed in Ω while no ICs are initially located within the domain, i.e. $N(0) = 0$. The influx of ICs has been defined based on a qualitative observation of the video footage of laboratory experiments reported in [14], by focusing on the monitored area corresponding to the computational domain and FPR1 CC cells, as well as the related statistics reported in [2]. Specifically, ICs are assumed to progressively enter the domain from the top side of the domain over the period of time $[0, 15]$ h, corresponding to the period $[24, 39]$ h, with an inflow rate of about 1 cell every $\tau = 6$ min resulting into a total amount of 150 ICs. The positions over the boundary domain where the cells enter the domain will be randomly selected only once and fixed for all the forthcoming simulations. The velocity of the entering cells is always assumed downward directed with modulus equal to 0.03 μm s^{-1}, i.e. the mean value reported in [2].

Following preliminary results in [15], the Robin boundary conditions in Eq. (10) are set with a uniform rate of exchange over the entire boundary, i.e. $a_N = a_E = a_N = a_W = 10^5$ μm s^{-1}, and different flux intensities on each side of Ω, i.e. $b_S = 2.2$ mol μm^{-1}s^{-1}, $b_E = 1.2$ mol μm^{-1}s^{-1}, $b_N = 0$ mol μm^{-1}s^{-1}, $b_W = 1.8$ mol μm^{-1}s^{-1}. In fact, these values qualitatively well capture the experimentally observed chemical distributions, which, in reality, are affected also by the TCs placed outside the domain, i.e. in the rest of the TCs chamber and, mainly, in the bottom TCs wells.

Remark. *The effect of Robin boundary conditions in Eq.* (10) *on the overall dynamics is strong, as shown in [15, 16] and in the Supplementary Material S1. Indeed, such boundary conditions involving the values of parameters a and b, regulate the inflow/outflow of chemicals in the observed domain. Note that the numerical values here imposed allow us to qualitatively reproduce the ICs dynamics observed in experimental movies, but they do not correspond to reality. For the mentioned reasons, here we keep them fixed in order to not affect the results of the sensitivity analysis.*

Concerning the space of parameter to investigate, it is first worth noticing that the cell radii R_T and R_I can be easily measured *in vitro*. R_T and R_I are therefore hereafter fixed to 10 μm and 4 μm, respectively, accounting for the experimental measurements reported in Section 2.1. In turn, $R_{\text{rep},T}$ and $R_{\text{rep},I}$ are consistently defined equal to 14 μm and 8 μm, respectively. Our sensitivity analysis therefore deal with the 13 model parameters listed in Table 1. Notice that this will highlight the role of the different features of the mathematical model. In fact, we have: coefficients regulating the evolution of the chemical signal, i.e. D, ξ, η; the damping coefficient due to cell adhesion to the substrate ζ; parameters characterizing ICs sensing of the chemical signal, i.e. γ, R_C, k_1, k_2; and parameters linked to cell-cell interactions, i.e. $\omega_{\text{rep},T}$, $\omega_{\text{rep},I}$, $\omega_{\text{adh},I}$, $R_{\text{adh},I}$, β. Specifically, taking into account either estimates in literature and preliminary numerical simulations in [15], qualitatively reproducing phenomena observed experimentally, model parameters are assumed independent and uniformly distributed over the ranges reported in Table 1.

Par.	Description	Units	Range	Value	Ref.
D	Diffusion coefficient of the chemical	$\mu m^2 s^{-1}$	$[1.5 \cdot 10^2, 1.5 \cdot 10^3]$	900	[23]
ξ	Default production rate of the chemical	μm^{-2} mol s^{-1}	$[5 \cdot 10^{-7}, 10^{-5}]$	$1.6 \cdot 10^{-6}$	[-]
η	Decay rate of the chemical	s^{-1}	$[10^{-5}, 10^{-3}]$	10^{-4}	[22]
ζ	Damping coefficient (IC-substrate adhesion)	s^{-1}	$[10^{-3}, 5 \cdot 10^{-3}]$	$2.1 \cdot 10^{-3}$	[-]
γ	Coefficient of chemotactic effect	μm^{-1}	$[5, 10^2]$	$2 \cdot 10$	[-]
R_C	Detection radius of chemicals	μm	[5, 15]	7	[-]
k_1	Cellular drift velocity	mol s^{-1}	$[10^{-9}, 10^{-8}]$	$3.9 \cdot 10^{-9}$	[23]
k_2	Receptor dissociation constant	mol μm^{-2}	$[10^{-14}, 10^{-13}]$	$5 \cdot 10^{-14}$	[23]
$\omega_{\text{rep},T}$	Repulsion coefficient between ICs and TCs	$\mu m^2 s^{-2}$	$[10^{-3}, 10^{-2}]$	$8.5 \cdot 10^{-3}$	[-]
$\omega_{\text{rep},I}$	Repulsion coefficient between ICs	$\mu m^2 s^{-2}$	$[10^{-4}, 10^{-3}]$	$5 \cdot 10^{-4}$	[-]
$\omega_{\text{adh},I}$	Adhesion coefficient between ICs	s^{-2}	$[10^{-8}, 10^{-6}]$	10^{-7}	[-]
$R_{\text{adh},I}$	Radius of action of adhesion between ICs	μm	[9, 11]	10	[-]
β	Alignment coefficient	s^{-1}	$[5 \cdot 10^{-2}, 5]$	10^{-1}	[26]

Table 1: Model parameters investigated in the GSA. For each parameter, we report a short description; the unit of measurements; a range of values where the parameters are considered uniformly distributed used in our sensitivity analysis; a nominal value used for the simulation in Section 3.1; references used to estimate the used values. In the last column '[-]' means that the values have been estimated through preliminary numerical simulations.

Output quantities. A crucial point in performing a useful sensitivity analysis is the choice of the output quantities to be considered. In our case, on one hand, it is hard to characterize with a single scalar measure the ICs dynamics arising from Eqs. (1)-(2). On the other hand, as highlighted for instance by the several statistics reported in [1, 2], there are many relevant readouts and features of CoC biological experiments to be considered: e.g. the clusterization of ICs around TCs, the spatial distribution of ICs, as well as their speed and direction of motion. In fact, from a biological point of view, the experimental evaluation of these statistics may support the identification of how possible manipulations of the scenario affect ICs dynamics. In this perspective, our sensitivity analysis deals with multiple model outcomes (defined below) to highlight how the model parameters affect the spatial distribution, the speed, and the direction of motion of ICs, in the considered case study. In order to look at the temporal evolution of the system, the observation period $\mathcal{I}$ is divided into three sub-intervals $\mathcal{I}_h = [t_{h-1}, t_h]$, with $h = 1, 2, 3$, with the same length. The symbol $< \cdot >_h$ then introduced to denote the time average of the quantity " $\cdot$ " over the time interval $\mathcal{I}_h$, i.e,

$$< \cdot >_h := \frac{1}{(t_h - t_{h-1})} \int_{\mathcal{I}_h} \cdot \, dt. \tag{11}$$

The decision to divide the observation period into three intervals is based on preliminary studies of the experiment and numerical simulations of cell dynamics, which highlighted

three distinct cell behaviors over time. During the first time interval $\mathcal{I}_1 = [0, 8]$ h the few ICs present in the domain are usually located in the top part of the domain. During $\mathcal{I}_2 = [8, 16]$ h there still are ICs entering the domain and, depending on the cases, the other ICs might either still move downwards following the chemical field, cluster around a TC or leave the domain. During $\mathcal{I}_3 = [8, 16]$ h no more cells are entering the domain, and enough time has passed to observe the ICs' final 'destiny', i.e. whether they leave the domain or cluster around an encountered TC. These three time intervals therefore allow to monitor possible variations in time of both ICs dynamics and detect the role of the model parameters.

To monitor the spatial distribution of the ICs, the domain Ω is partitioned into 4 square disjoint subdomains Ω_q, with $q = 1, 2, 3, 4$, i.e. Ω_1 (North-West), Ω_2 (North-East), Ω_3 (South-West) and Ω_4 (South-East), as depicted in Panel A in Fig. 2. The spatial distribution of ICs during each time interval $\mathcal{I}_h$ with $h = 1, 2, 3$, is then characterized by the following quantities.

$< N >_h$**:** the average number of ICs in the domain Ω during the time period $\mathcal{I}_h$, i.e. $< N >_h :=< N(t) >_h$, being $N(t)$ the number of ICs inside the domain at the instant of time t.

$< N_q >_h$ **:** the average number of ICs located in the portion of the domain $\Omega_q, q = 1, \ldots, 4$ during the period of time $\mathcal{I}_h$, i.e. $< N_q >_h :=< N_q(t) >_h$ being $N_q(t)$ the number of ICs such that $\mathbf{X}_i(t) \in \Omega_q$ at the instant of time t.

$p_{\mathbf{out}}(t_h)$**:** the fraction of ICs that have left the domain Ω before the instant of time t_h. Specifically, it is defined as

$$p_{\text{out}}(t_h) := \frac{N_{\text{out}}(t_h)}{N_{\text{in}}(t_h)}, \tag{12}$$

where $N_{\text{in}}(t)$ and $N_{\text{out}}(t)$ are the number of ICs that respectively entered or exited from the domain Ω over the period $[t_0, t]$. Notice that if $N_{\text{in}}(t_h) = 0$ then obviously also $N_{\text{out}}(t_h)$ nullifies, and it is then consistent to set $p_{\text{out}}(t_h) = 0$.

$p_{\mathbf{cl}}(t_h)$**:** the fraction of ICs clustered around a TC at the instant time t_h, i.e.

$$p_{\text{cl}}(t_h) := \frac{N_{\text{cl}}(t_h)}{N(t_h)}, \tag{13}$$

where N_{cl} is the number of ICs sufficiently close to a TC to be considered clusterized. Specifically, we assume

$$N_{\text{cl}}(t) := \text{card}\left\{ i = 1, \ldots, N(t) \ : \ \min_{j=1,\ldots,N_T} \|\mathbf{X}_i(t) - \mathbf{Y}_j(t)\| \leq R_{\text{cl}} \right\}, \tag{14}$$

where $R_{\rm cl}$ is set equal to $R_T + 3R_I$. This value of $R_{\rm cl}$ allows to detect clusters constituted by at most two layers of ICs around a TC. Also in this case, if the denominator $N(t_h)$ nullifies, then also as also $N_{\rm cl}(t_h) = 0$, and it is thus consistent to set $p_{\rm cl}(t_h) = 0$.

Notice that the computation of the quantities $p_{\rm out}$ and $p_{\rm cl}$ is quite expensive. Hence, they are not calculated as means on $\mathcal{I}_h$, but only at the final instants of each time interval.

Concerning instead the dynamics of ICs, we hereafter distinguish between cell speed and direction of motion by writing IC velocities as $\dot{\mathbf{X}}_i(t) = v_i(t)\,(\cos\theta_i(t), \sin\theta_i(t))$, where $v_i := \|\dot{\mathbf{X}}_i\|$, while θ_i is the angle between $\dot{\mathbf{X}}_i$ and the direction $\mathbf{n} = (0, -1)$ toward the bottom side S of the domain. Cell dynamics at each time interval $\mathcal{I}_h$ with $h = 1, 2, 3$, is thus described by the following quantities.

$< \overline{v} >_h$**:** the time average over the time interval $\mathcal{I}_h$ of the mean speed $\overline{v}(t)$ of the ICs located in the domain Ω at t, i.e.

$$\overline{v}(t) := \frac{1}{N(t)} \sum_{i=1}^{N(t)} v_i(t). \tag{15}$$

$< v_{\rm std} >_h$**:** the time average over the interval time $\mathcal{I}_h$ of the standard deviation $v_{\rm std}(t)$ of the speed located in the domain Ω at t, i.e. $v^h_{\rm std} :=< v_{\rm std}(t) >_h$ with

$$v_{\rm std}(t) := \sqrt{\sum_{i=1}^{N(t)} \frac{(v_i(t) - \overline{v}(t))^2}{N(t)}}. \tag{16}$$

$C^h_{\mathbf{n}}$**:** the correlation between cell direction of motion and the fixed direction $\mathbf{n}$. Specifically, it is defined as

$$C^h_{\mathbf{n}} = \frac{1}{N(t)} \sum_{i=1}^{N(t)} < \cos(\theta_i(t)) >_h . \tag{17}$$

C^h_C**:** the correlation between the direction of motion of ICs and the gradient of the chemical cue sensed by cells. Denoting by $\phi(\mathbf{x})$ the angle between the gradient of the chemical field $\nabla\varphi$ at $\mathbf{x} \in \Omega$ and the direction $\mathbf{n}$, we define

$$C^h_C := \frac{1}{N(t)} \sum_{i=1}^{N(t)} < \cos(\phi(\mathbf{X}_i(t)) - \theta_i(t)) >_h . \tag{18}$$

Notice that if $C^h_C \approx 1$ indicates that all individuals move according to the chemical cue over the entire period $\mathcal{I}_h$.

2.4. Our GSA approach

Accounting for a large number of model parameters, the non-linearity of the model, and the considerable computational cost of a single *in silico* realization (see Appendix A for further details about the numerical scheme), the GSA is performed in two successive steps. First, the improved method of Morris is used to screen the model parameters and rank them according to their importance in affecting the model outcome. Then the more computational expensive eFAST method is used to quantify the effect of only the few most important parameters detected by the Morris method.

The improved Morris method. The screening method of the Elementary Effects defined by Morris in [4], as well as its improved version proposed by Campolongo et al. [3], are suited to efficiently treat deterministic models with a large number d of model parameters mutually independent, and/or fairly expensive to simulate. In [4], the concept of *elementary effect* associated with the k-th model parameter on a scalar target output Q is defined as

$$\mathrm{EE}_k(\mathbf{z}) := \frac{Q(\mathbf{z} + \Delta\mathbf{e}_k) - Q(\mathbf{z})}{\Delta} \tag{19}$$

where $\mathbf{z}$ is a point in the region of interest $\mathcal{H}$ of the d-dimensional parameters space; $\mathbf{e}_k$ are the standard basis vectors in $\mathbb{R}^d$; and Δ is such that $(\mathbf{z} + \Delta\mathbf{e}_k) \in \mathcal{H}$. The basic idea of the Morris method is to characterize the distribution of EE_k, by randomly sampling different $\mathbf{z}$ from $\mathcal{H}$, (see Appendix B or [3] for further details), and then use the mean μ_k and the standard deviation σ_k of the sampled EE_k as sensitivity measures. However, as stated in [3], using the mean of the absolute value of EE_k, said μ_k^* rather than μ_k, avoids possible cancellation effects introduced by negative values. Specifically, μ_k^* is an indicator of the overall influence of the k-th parameter on the output, thus providing the ranking of model parameters by importance. Conversely, σ_k highlights if the k-th parameter has non-linear effects or interacts with other parameters. An effective way to visualize the qualitative information provided by the extended Morris method is to display model parameters in the $\mu_k^* - \sigma_k$ plane. This allows to identify three classes of model parameters, as stated in [3, 4, 6, 42]: negligible ones (low μ_k^* and σ_k, i.e. $\mu_k^*, \sigma_k \approx 0$); important parameters with large linear effects and no interactions (large μ_k^* and small σ_k, i.e. $\mu_k^* > \sigma$); important parameters with non-linear and/or interaction effects (large μ_k^* and large σ_k, i.e. $\mu_k^* \leq \sigma$). In practice, negligible parameters appear in the bottom left part of the $\mu_k^* - \sigma_k$ plane while important parameters are distributed on their right. In addition, important parameters located largely below the 45-degree line has purely linear effects without and no interactions, while those close or above the 45-degree line are affected by non-linear effects and/or interactions.

The eFAST method. The eFAST method is a variance-based GSA method, i.e. an approach suitable to study non-linear models, and based on the ANOVA-like decomposition of the target output variance [10, 44, 46, 54, 55]. Denoting by $\{Z_k\}_{k=1,\dots,d}$, the d

model parameters, if they are mutually independent, the variance of the output Q can be decomposed into contributions related to either the single parameters or combinations of them, i.e.

$$\mathrm{Var}(Q) = \sum_{k=1,\dots,d} D_k(Q) + \sum_{\substack{k,i=1,\dots,d \\ k<i}} D_{ki}(Q) + \ \dots \ + D_{12\dots d}(Q), \tag{20}$$

where $D_k(Q) := \mathrm{Var}[\mathbb{E}(Q|Z_k)]$; while $D_{ki}(Q) := \mathrm{Var}[\mathbb{E}(Q|Z_k, Z_i)] - D_k(Q) - D_i(Q)$ and so on. Then, the eFAST method returns, for each k-th model parameter, the Sobol' sensitivity indices [7] defined as

$$S_k(Q) = \frac{D_k(Q)}{\mathrm{Var}(Q)}, \quad S_k^T(Q) = S_k(Q) + \sum_{\substack{i=1,\dots d \\ i<j}} \frac{D_{ki}(Q)}{\mathrm{Var}(Q)} + \ \dots \ + \frac{D_{12\dots d}(Q)}{\mathrm{Var}(Q)} = \sum_{g\in \#k} \frac{D_g(Q)}{\mathrm{Var}(Q)}, \tag{21}$$

where $\#k$ denotes all the possible combinations from $\{1, ..., d\}$ containing k. In Eq. (21), the first-order indices $S_k(Q)$ quantify the fraction of the output variance due to changes in the k-th model parameter only, and are thereby also called *main effects indices* [54, 55]. Conversely, the so-called *total effects* $S_k^T(Q)$, introduced in [44], assess the impact of parameter k including all possible interactions with the other factors. It follows that the difference $S_k^T - S_k$ represents the fraction of the output variance due to interactions involving the k-th parameter. Practical criteria to interpret these quantities are provided in [11]: (i) the threshold 0.01 discriminates between *negligible* effects/parameters (affecting less than 1% of the output variance) and *relevant* once; (ii) Sobol indices above 0.1 identify the *key* parameters influencing more than the 10% of $\mathrm{Var}(Q)$.
The general idea at the basis of the eFAST method for computing first-order and total-effect indices is reported in Appendix B. Refer to [10] and references therein for a detailed description.

Two-step GSA. For each one of the output quantities introduced in Section 2.3, the improved Morris method is first used to rank the 13 parameters in Table 1 according to their importance. Specifically, the sampling method by Campolongo et al. [3] has been used and allows us to identify 280 points in the 13-dimensional parameter space (see Appendix B for further details about the Morris sampling). Each one of these points identifies a parameter setting used to compute, by numerically solving the model, the values of our 11 output quantities over the 3 distinct time intervals, as detailed in Section 2.3. All these values are then used to estimate the elementary effects of the 13 model parameters on each output quantity Q. The relative sensitivity measure μ_k^*, defined in [3] and recalled above, identifies for each output quantity Q a parameters' ranking of the 13 model parameter by importance (i.e. by decreasing values of μ_k^*).
The resulting rankings are then compared two-by-two by using the Spearman's Rank correlation coefficient to measure their degree of similarity. High values of this coefficient,denoting by definition a strong similarity, in fact indicate that the two compared

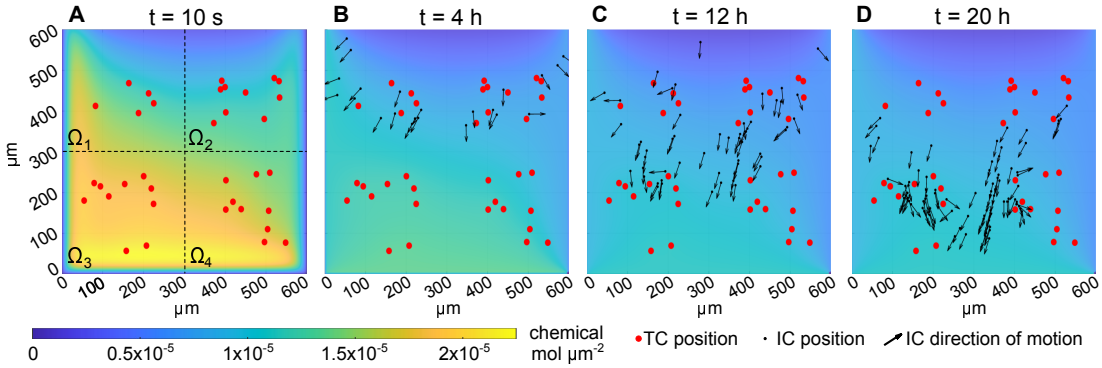

Figure 3: Representative model realization. Eqs. (1)-(2) are solved by considering the case study scenario in Section 2.3, by fixing the 13 investigated model parameters equal to the nominal parameters values in Table 1. The concentration of the chemoattractant φ, with units mol, is plotted with 2D contour lines. Red dots represent the static 35 TCs randomly distributed in the domain. Black dots denote the present positions of the moving ICs, and the black arrows their direction of motion. In Panel A, the names of the subdomains Ω_l, with $l = 1, \ldots, 4$ are reported.

output quantities are mainly affected by the same model parameters. It therefore follows that, if all the Spearman's Rank correlation coefficients are sufficiently high, it is possible to identify a single group of relevant parameters mostly determinant for all the considered model outputs. Otherwise, it could be necessary to consider different groups of relevant parameters to study in the second step of our analysis or find other strategies. Interestingly, as detailed in the next section, in our case, all the model outputs of our interest are mainly affected by the same 6 parameters.

The eFAST method is then applied to investigate a smaller space of parameters, obtained by fixing the parameters denoted as negligible by the Morris method, equal to an arbitrary value in the range used for the screening. The sampling method in [10] here results in a sampling of 3000 points in the 6-dimensional parameter space (see Appendix B). Notice that, despite the eFAST method has been here used to investigate only 6 parameters, it requires approximately one order of magnitude more than the Morris method to screen 13 parameters. Such a different computational cost is intrinsically linked to the sampling approach at the basis of each method, specifically designed to provide either qualitative (Morris) or quantitative (eFAST) information about the role of the model parameters. This underlines the computational usefulness of coupling these two methods, instead of applying eFAST to all the 13 model parameters.

Both Morris and Sobol analysis have been performed by using the SALIB Python library [8, 9].

3. Results

3.1. Representative numerical realization of the case study

This section is devoted to a representative numerical realization of the model shown in Fig. 3, that has been obtained by considering the case study scenario introduced in Section 2.3 (see, for instance, the video footage of laboratory experiments Movie S13 and Movie S14 in the Supplementary Material of [14]). The 13 investigated model parameters are here fixed equal to the nominal values in Table 1, within the ranges considered for the sensitivity analysis. The aim of this simulation is to help readers not familiar with the model to better appreciate the forthcoming sensitivity analysis.

As specified in Section 2.3, at the initial time instant $t = 0$ h, there are no ICs inside the domain, $N_T = 35$ TCs are randomly distributed as shown in Panel A (red dots), and the chemoattractant field is $\varphi_0(\boldsymbol{x}) = 0$ mol μm^{-2} for any $\boldsymbol{x} \in \Omega$. Due to the presence in the reaction-diffusion equation (1) of the non-local source term F_S defined in Eq. (3), a certain amount of chemical is progressively generated around the TCs since the first time instant. At the same time, the chemoattractant rapidly spreads across the domain driven by the diffusion and decay terms in Eq. (1), in addition to the Robin conditions in Eq. (10), with the values defined in Table 1 and Section 2.3, respectively. A feasible chemical field consistent with experimental evidence in [14], see Panel A in Fig. 3, is thus obtained before the first ICs enter the domain at $t = \tau (= 6$ min). Specifically, the emerged chemical field is characterized by higher concentrations at the bottom left of the domain by mimicking the presence of chemical secreting TCs outside Ω, especially the large number of them located in the bottom well, see Fig. 1. Such a chemical pattern then drives first ICs entered in the domain, to move downwards to possibly approach the TCs, see Panel B in Fig. 3. As time goes by, see Panels C-D in Fig. 3, the chemical field gradually becomes more homogeneous due to the diffusion and decay terms, in addition to the boundary conditions defined above. Notice that ICs are coherently mainly attracted toward the densest groups of TCs. However, during the evolution of the system, some ICs reach the boundary and leave the domain. The resulting reduction in the downward chemical gradient is consistent with IC dynamics observed in [14]. Meanwhile, the ICs continue to mostly move downwards following the chemical field, by possibly temporarily clustering around the encountered TCs.

3.2. Two-step global sensitivity analysis

3.2.1. First step: Morris sensitivity analysis

First, as described in Section 2.4, the improved Morris method is applied to investigate the hypercube in the space of parameters defined by the ranges in Table 1. Figures 4-5 show the values of the output quantities resulted by solving the model with each one of the 280 sampled parameter settings. It is remarkable that all the obtained values are consistent with the experimental observations in [1, 2, 14]. This confirms the feasibility of the space of parameters we are considering and support our SA.

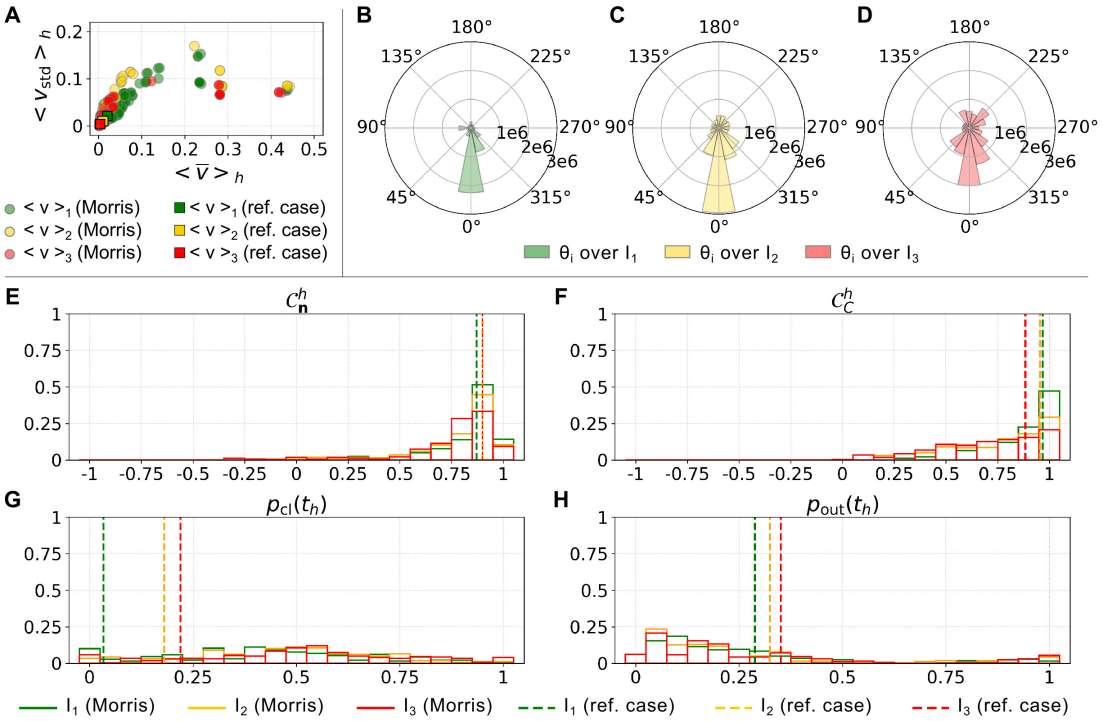

Figure 4: Dynamics of ICs in the model realizations considered for the Morris analysis. Each panel shows the values of an output quantity obtained over all the 280 parameter settings constituting the Morris sampling. In all panels, data referring to three time intervals $\mathcal{I}_1 = [0, 8]$ h, $\mathcal{I}_2 = [8, 16]$ h, and $\mathcal{I}_3 = [16, 24]$ h are respectively represented in green, yellow, and red. Panel A: $< \overline{v} >_h$ vs. $< v_{\text{std}} >_h$, see Eqs. (15)-(16). Circle markers denote the values obtained with the parameter settings identified by the Morris sampling. Square markers refer to the representative simulation in Section 3.1. Panels B-C-D: Absolute frequency of the angles θ_i denoting the direction of motion of ICs over $\mathcal{I}_1$ (B), $\mathcal{I}_2$ (C), and $\mathcal{I}_3$ (D). Panel E: Distributions of the correlation between cell direction of motion and the fixed direction $\mathbf{n}$, i.e. $C^h_{\mathbf{n}}$ in Eq. (17). Panel F: Distributions of the correlation between cell direction of motion and the gradient of the chemical field, i.e. C^h_C in Eq. (18). Panel G: Distributions of the fraction of ICs clustered around a TC at t_h, i.e. $p_{\text{cl}}(t_h)$ in Eq. (13). Panel H: Distributions of the fraction of ICs that have left the domain at t_h, i.e. $p_{\text{out}}(t_h)$ in Eq. (12). In Panels E-H, x-axis and y-axis denote the values assumed by the considered quantity and relative frequencies, respectively; and dashed lines represents the values observed in Section 3.1.

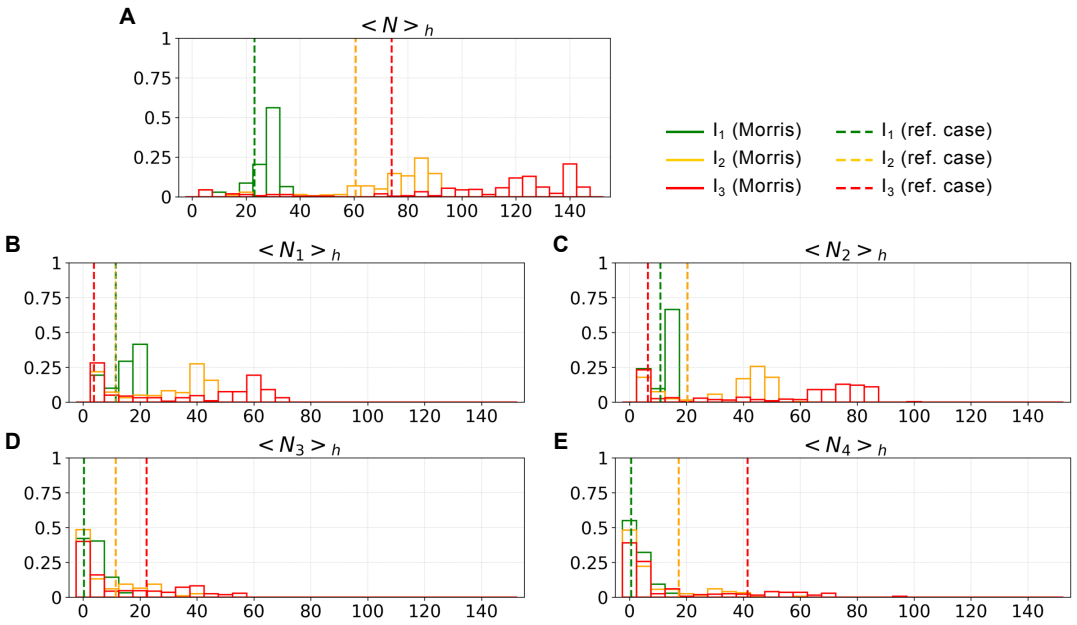

Figure 5: Number and distribution of ICs in the model realizations considered for the Morris analysis. Each panel shows the values of an output quantity obtained over all the 280 parameter settings constituting the Morris sampling. Panel A: Number of ICs in the domain $< N >_h$. Panel B-E: Number of ICs in the subdomains, i.e. in the North-West $< N_1 >_h$ (B); North-East $< N_2 >_h$ (C); South-West $< N_3 >_h$ (D); and South-East $< N_4 >_h$ (E). In all panels, data referring three time intervals $\mathcal{I}_1 = [0, 8]$ h, $\mathcal{I}_2 = [8, 16]$ h, and $\mathcal{I}_3 = [16, 24]$ h are respectively represented in green, yellow and red; x-axis and y-axis denote the values assumed by the considered quantity and relative frequencies, respectively; and dashed lines represents the values observed in the representative simulation in Section 3.1.

Referring to Panel A in Fig. 4, the ranges covered by IC speeds in the three time intervals are in fact in accordance with the values reported in Fig. 4 of [2]. All together Panels A-D in Fig. 4 show that, during the first time interval $\mathcal{I}_1$, ICs dynamics is always characterized by a faster motion directed towards the bottom side of the domain. Instead, during $\mathcal{I}_2$ and $\mathcal{I}_3$, both the mean and the standard deviation of their speed decrease, and the directions of motion of ICs are more distributed, as shown by the polar histograms in Panel C and D. This can be explained by the fact that at time $\mathcal{I}_2$ and $\mathcal{I}_3$ ICs reach their targets and have local interactions in all the directions. In addition, Panels E-F in Fig. 4 indicate that ICs over $\mathcal{I}_1$ mainly move towards the bottom side of the domain following the chemotactic field, as $C^1_{\mathbf{n}}$ and C^1_C are both mostly close to 1. Then, the alignment of ICs direction of motion to $\mathbf{n}$ and to the chemotactic field slightly decreases.

These first results suggest a system behavior similar to those observed in the reference simulation in Section 3.1 is captured for all the sampled parameter setting. In the first period $\mathcal{I}_1$, ICs dynamics are mainly driven by the chemical field, which is characterized by a higher difference in its concentration between the top and the bottom part of the domain, where the concentration is higher, see Panels A-B in Fig. 3. As time goes by, the chemoattractant diffuses inside and outside the domain reducing the spatial differences in its concentration, see Panels C-D in Fig. 3, while ICs have either left the domain or have approached the TCs so that cell-cell interactions enter into play. This explains why both the mean and the standard deviation of ICs speed decrease, while the directions of motion of ICs are more sparse.

These considerations are also supported by Panels G-H in Fig. 4. Panel G in Fig. 4 shows that the fraction of cells that reach the TCs and cluster around them slightly increases in time. Indeed, as time goes by, ICs entered in the domain progressively move downwards so that a higher portion of cells falls close to the TCs. In addition, Panel H in Fig. 4 shows that the portion of ICs that have left the domain before t_1, t_2, and t_3, is substantially the same. This suggests that p_{out} is independent on where the ICs are mainly located (i.e., in the top or bottom part of the domain) confirming that cells are allowed to trespass the boundary with no limitations.

Figure 5 provides information about the mean number and the spatial distribution of the ICs coherent with the above considerations. Panel A of Fig. 5 shows that over $\mathcal{I}_1$, in most cases, roughly 25-30 ICs are present inside the domain with slight changes among the different simulations. During $\mathcal{I}_2$ and $\mathcal{I}_3$, a greater and greater variability in the average number of ICs inside the domain is observed. This is due to the fact that, as time goes by, the Robin condition, mimicking the presence of other TCs outside Ω, may attract some of, sometimes all, the ICs outside of the domain. Concerning instead the spatial distribution of the ICs in the domain portions Ω_q, with $q = 1, ..., 4$, Panels B-C in Fig. 5 show that the time evolution of both $< N_1 >_h$ and $< N_2 >_h$, i.e. the mean number of ICs in the upper parts of the domain, are qualitatively similar to that observed over the entire domain, i.e. in Panel A of Fig. 5. Conversely, Panels D-E in Fig. 5 highlight

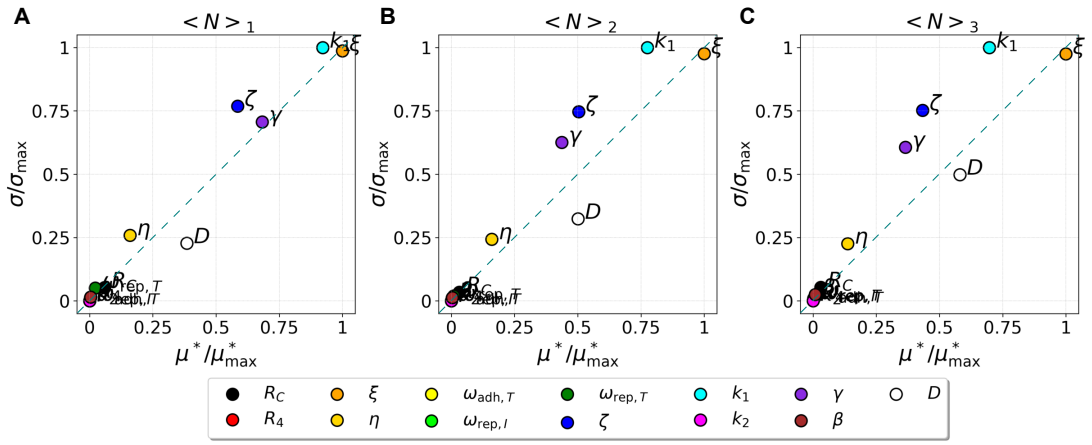

Figure 6: Distribution of the model parameters in the plane $\mu^*/\mu^*_{\max} - \sigma/\sigma_{\max}$, according to the Morris sensitivity measures relative to the observables $< N >_h$ with $h = 1, 2, 3$, given in Section 2.3, i.e. the mean amount of ICs in the domain over the time intervals $\mathcal{I}_1 = [0, 8]$ h (A), $\mathcal{I}_2 = [8, 16]$ h (B), and $\mathcal{I}_3 = [16, 24]$ h (C).

that only in a few cases all the ICs reach the bottom part of the domain, indicating that they mainly cluster around the TCs in Ω_1 and Ω_2, or leave the domain.

Once stated that the considered hypercube in the space of parameter results in admissible system outcomes, i.e. consistent with the scenarios observed in the experiments by Vacchelli et al. in [14], it is reasonable to analyze the Morris sensitivity measures. As recalled in Section 2.4, for each output quantity, the improved Morris method returns two sensitivity measures μ^*_k and σ_k for each model parameter $k = 1, ..., 13$, i.e. the mean of the absolute value and the standard deviation of their elementary effects. Recalling that these quantities provide only qualitative information about the role of the model parameters, in Fig. 6 and in Figs. S1-S4 of the Supplementary Material, the Morris measures μ^*_k and σ_k are scaled by the values $\mu^*_{\max} := \max_k(\mu^*_k)$ and $\sigma_{\max} := \max_k(\sigma_k)$, respectively.

The plots in Fig. 6 show that, despite small differences in their order, the parameters mostly affecting the mean number of ICs in the domain, over all the three time intervals, are D, ξ, η, ζ, γ, k_1, while the effect of the other model parameters results negligible. As a remark, the first three parameters are the diffusion coefficient D, the growth rate ξ, and the consumption rate η defining the evolution equation (1) for the chemical signal released by TCs. The latter are instead related to the dynamics of ICs. The damping coefficient ζ and the coefficient of chemotactic effect γ appear in the evolution equation (2). The drift velocity k_1 is involved in the receptor saturation function $\chi(\varphi)$ defined in Eq. (5) and included in Eq. (4) to implement cell chemical sensing.

Interestingly, a similar result emerges also from the plots in Figs. S1-S4 in the Supplementary Material for the other output quantities, i.e. $< \overline{v} >_h$ and $< v_{\text{std}} >_h$ (see

Suppl. Fig. S1), $< N_q >_h$ (see Suppl. Fig. S2), $p_{\text{out}}(t_h)$ and $p_{\text{cl}}(t_h)$ (see Suppl. Fig. S3), $C^h_{\mathbf{n}}$ and C^h_C (see Suppl. Fig. S4). The consistency among the rankings resulted from the Morris method by assuming the ranges of parameters in Table 1, is evaluated through the Spearman rank correlation coefficients between any pair of output quantities. All these coefficients are higher than 0.8 and thereby allow us to neglect small differences among the rankings and state that variations in all the outcomes of our interests are mainly regulated by the same 6 parameters. In other words, both the spatial distribution and the dynamics of ICs are thus mainly affected by the evolution of the chemical signal, dictated by D, ξ, and η; cell sensing of and reactivity to the chemoattractant, i.e. k_1 and γ; and cell adhesion to the substrate, i.e. ζ,.

Lastly, notice that in all panels in Fig. 6 and in Figs. S1-S4 in the Supplementary Material, the most influential parameters are mainly distributed along the bisector of the plane. This suggests that there are important non-linear and/or interaction effects, consistently with the non-linearity of the analyzed model.

3.2.2. Second step: eFAST sensitivity analysis

Following the approach stated in Section 2.4, the second step of our sensitivity analysis consists in focusing on the non-negligible model parameters detected by the Morris method, by applying the eFAST method. Having identified a small group of 6 relevant parameters mostly responsible for the variability of all the output quantities of our interest, the space of parameter to investigate here is defined by the ranges of values of D, ξ, η, ζ, γ, k_1 in Table 1. The other 7 negligible parameters can be instead equivalently fixed equal to any value in their ranges. In particular, we here set them as in the reference simulation in Section 3.1, i.e. equal to the nominal value in Table 1. For each output quantity, the eFAST method estimates the Sobol sensitivity indices S_k and S^T_k of the 6 model parameters, i.e. the main and total effects defined in Eq. (21). These are displayed in bar charts in Figs. 7-8. Specifically, in order to highlight possible effects due to model non-linearity and parameter interactions, for each time interval, the main effect S_k of each model parameter is plotted alongside the total effect bar S^T_k, see again Eq. (21). In all panels, two dotted lines denote the thresholds 0.01 (black) and 0.1 (red) defined in [11] to identify non-influential and key parameters, affecting less than 1% or more than 10% of the variance, respectively, see Section 2.4 and references therein.

Panel A in Fig. 7 first highlights that the variance of the number of ICs within the domain Ω is mainly due to the source rate of the chemical ξ and the diffusion coefficient D, over the whole observation period. Having fixed the inflow of ICs, this means that the mean number of ICs remaining in the domain mainly depends on the strength of the gradient of the chemical and how fast it spreads over the domain. In fact, the same results are obtained for the fraction of ICs $p_{\text{out}}(t_h)$ that have left the domain, see Panel B in Fig. 7. In both cases, the decay rate η of the chemical always has negligible effects. Notice that it happens for almost all the output quantities, see Figs. 7-8.

Focusing instead on each portion Ω_q, with $q = 1, ..., 4$, of the domain, see Panels

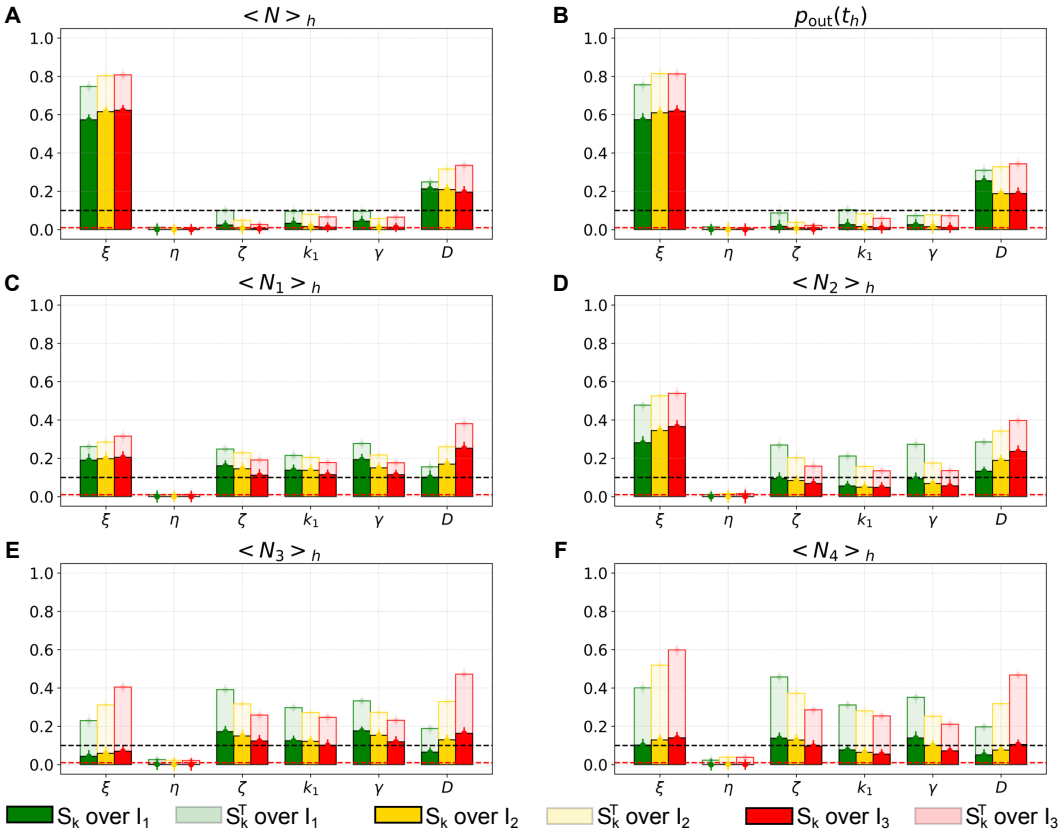

Figure 7: eFAST-estimated Sobol sensitivity indices in Eq. (21) of the output quantities $< N >_h$ (A), $p_{\text{out}}(t_h)$ (B), $< N_1 >_h$ (C), $< N_2 >_h$ (D), $< N_3 >_h$ (E), and $< N_4 >_h$ (F), with $h = 1, 2, 3$, given in Section 2.3. In all panels, filled bars represent the main effects S_k while transparent bars denote the total effects S_k^T. Data referring to the three interval times $\mathcal{I}_1$, $\mathcal{I}_2$, and $\mathcal{I}_3$ are respectively displayed in green, yellow, and red.

C-F in Fig. 7, the eFAST analysis reveals that, apart from η, all the other five relevant model parameters have a key role on the variance of $< N_q >_h$, in terms of either main or total effects. In particular, differences among the Sobol indices values shown in Panels C-F of Fig. 7 suggest that the number and distribution of TCs characterizing each Ω_q (see Panel A in Fig. 3), as well as their distance from the source of ICs have an important role on the local cell dynamics. Comparing Panels C-F with Panel A, it also emerges that parameters ζ, γ, and k_1, involved in Eq. (2) for ICs dynamics, have more important effects on the variance of $< N_q >_h$, i.e. the local distribution of ICs, rather than on $< N >_h$, i.e. the mean number of ICs within the entire domain. In addition, notice that in Panels C-F of Fig. 7 both the main and the total effects of ζ, γ and k_1 on $< N_q >_h$ decrease in time, while those of ξ and D increase and become the most relevant parameters over the last time interval $\mathcal{I}_3$ (see red bars). It is further remarkable that non linear/interaction effects have a higher influence on the mean amount of ICs in the bottom of the domain, i.e. $< N_3 >_h$ and $< N_4 >_h$ (see Panels E-F), rather than the mean amount of ICs in the top of the domain, i.e. $< N_1 >_h$ and $< N_2 >_h$ (see Panels C-D). Such a difference, can be related to the fact that regions Ω_3 and Ω_4 are more distant from the inflow ICs boundary. For instance, according to the considered parameter setting, it may happen that a very small portion or even none of the entered ICs reach the bottom part of the domain. The analysis of different ICs influx conditions may seed light on this aspect, however this is out of the purpose of the present work.

Concerning instead the observables introduced to monitor cell dynamics, due to the strong connection between cell displacement and speed, the time evolution of the main effects S_k observed in Fig. 7 appears also for the mean speed of ICs $< \overline{v} >_h$, see Panel A in Fig. 8. Over the first time interval $\mathcal{I}_1$, γ, k_1 and ζ have key main effects on the mean speeds (see green filled bars). As time goes by, these main effects lose importance, until during the last time interval $\mathcal{I}_3$, ξ and D have the more relevant main effects on $< \overline{v} >_h$ (see red filled bars). However, all the considered model parameters, apart from η, always have relevant total effects (see the transparent bars). In fact, the model non-linearity here results in important discrepancies between the total S_k^T and main effects S_k affecting the mean ICs speed, which moreover considerably increase in time. Notice that these variations in time of the main and interaction effects lead to a remarkable increase in time of the total effects S_k^T of the model parameters ξ and D, denoting that the influence of the underling chemical field on cell speed considerably changes in time. At the beginning, ξ appears among the key parameters since the outgoing flux (i.e. the fixed Robin's boundary conditions) can generate a strong downward chemical gradient provided that the source coefficient ξ is sufficiently high. As time goes by, the value of ξ and D become more crucial with respect to the others, as the evolution of the chemical gradient is dictated by how fast the chemical is released and spreads over all of the domain. In addition, the damping coefficient ζ always appears among the key parameters regulating the mean cell speed. This is coherent both from a mathematical point of view, as it is the

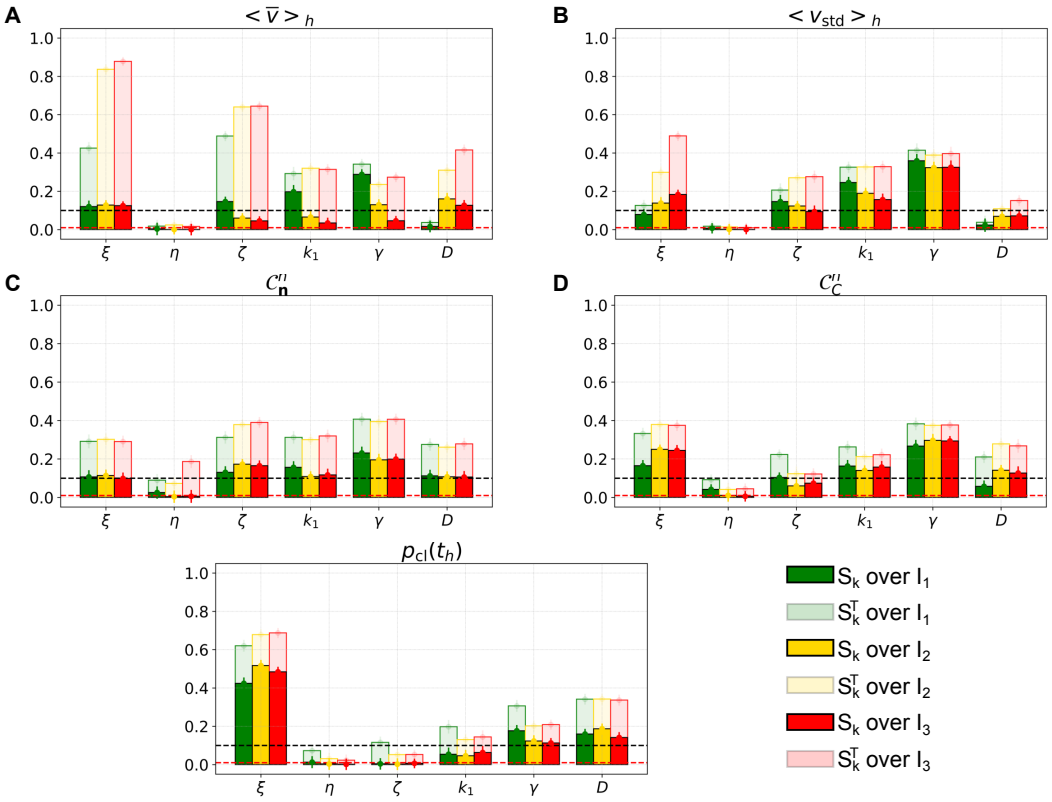

Figure 8: eFAST-estimated Sobol sensitivity indices in Eq. (21) of the output quantities $< \bar{v} >_h$ in Eq. (15) (A), $< v_{\text{std}} >_h$ in Eq. (16) (B), $C_{\mathbf{n}}^h$ in Eq. (17) (C), C_C^h in Eq. (18) (D), and $p_{\text{cl}}(t_h)$ in Eq. (13) (E), with $h = 1, 2, 3$. In all panels, filled bars represent the main effects S_k while transparent bars denote the total effects S_k^T. Data referring to the three interval times $\mathcal{I}_1$, $\mathcal{I}_2$, and $\mathcal{I}_3$ are respectively displayed in green, yellow, and red.

damping coefficient in Eq. (2), and from a biological point of view, since a higher/lower cell-adhesion to the homogeneous substrate have a constant in time effect on cell speed.

Concerning instead the standard deviation $< v_{\text{std}} >_h$ of ICs speed, see Panel B of Fig. 8, notice that also in this case the main effects of ζ, γ and k_1 decrease in time while ξ and D increase. Moreover, there is a remarkable increase in the total effect of ξ and D. Interestingly, all the relevant parameters have a higher total effect on $< v_{\text{std}} >_h$ rather than on $< \overline{v} >_h$, i.e. variations in the model parameters affect more the mean cell behavior than discrepancies among the individual behaviors.

Dealing with ICs direction of motion, Panels C-D in Fig. 8 show the effects of model parameters on $C^h_{\mathbf{n}}$ and C^h_C. On one hand, in Panel C, $C^h_{\mathbf{n}}$, i.e. cell motion alignment to the downward direction $\mathbf{n}$, is always strongly affected by all the parameters (apart from η) with important contributions due to the non-linearity of the model and interactions among the parameters. Indeed, cell trajectories are quite tortuous since they result from the interplay between the influence of the chemical field pattern and cell-cell interactions. In Panel D, also the alignment of cell motion to the gradient of the chemoattractant mainly depends on the values of ξ, γ and k_1, D and ζ, i.e. all parameters apart form η. However, in this case, important differences among total and main effects appears only for ξ and D regulating the chemical field evolution. These aspects are consistent and confirm the previous considerations about the relation between cell dynamics and chemical field.

Lastly, Panel E in Fig. 8 shows that $p_{\text{cl}}(t_h)$, i.e. cell clustering around TCs is mainly regulated by the rate of chemical secretion ξ, the diffusion coefficient D and the coefficients γ and k_1 in the chemotactic term. In other words, this is coherently mainly related by the evolution of the chemoattractant field and cell ability to locally detect the chemical gradient.

3.3. Discussion

Summing up, our study first indicates that the proposed model, by assuming parameter values within the ranges in Table 1, results in feasible scenarios qualitatively in accordance with the experimental results reported in [2, 14].

The improved Morris method then highlights that among the investigated 13 parameters listed in Table 1 only 6 of them are mainly responsible for the variability in both spatial distribution and dynamics of the ICs: the chemoattractant diffusion coefficient D; the chemical source and decay rates ξ and η; the chemotactic coefficient γ; the drift velocity k_1; and the damping coefficient ζ. Notice that parameters related to cell-cell interactions thus result negligible with respect to the influence of the chemical field and the substrate adhesion. Indeed, consistently in CoC experiments in [14], ICs do not move compactly but rather individually explore the domain by only accidentally and temporarily interacting with the other cells, as shown in Fig. 3. From a modeling point of view, the lack of parameters related to the cell-cell alignment among the most influential ones denotes that this term, at least as it is, is not fundamental in reproducing CoC experiments. This suggests that the model could be simplified by removing this

term unless specific empirical observations strongly justify its presence and possibly guide its re-calibration.

Focusing on the 6 non-negligible parameters, the eFAST method further states that: (i) the secretion rate ξ and the diffusion coefficient D of the chemoattractant strongly affect all the monitored output quantities (more than 10% of their variances over all the period or at least over $\mathcal{I}_2$ and $\mathcal{I}_3$); (ii) the decay rate η results substantially always negligible; (iii) cell adhesion to the substrate ζ and cell sensitivity to the chemical, i.e. γ and k_1, have constant relevant effects on the standard deviation of cell speed and cell directions of motion, while they strongly affect only the initial evolution of cell distribution and mean speed, i.e. only over $\mathcal{I}_1$; lastly (iv) the coefficient γ is a key parameter also for cell clustering.

In [32] a local sensitivity analysis on an agent-based CoC in-silico model was performed by varying the parameters one factor-at-time (OAT) in a range of ±20%. From the analysis carried out in [32], it turns out that the model there proposed is most sensitive to a parameter regulating the threshold value of chemical density necessary to start the migration of ICs towards the left chamber of the chip. Despite the different approach with respect to the modeling here presented, the study in [32] is substantially in accordance with our findings, since the most influential parameters result to be those linked to the chemical gradient.

However, due to the strong non-linearity of the model, it is not possible to deeply understand and study the system behavior when multiple parameters vary by changing one factor-at-time. In the present work, since all parameters are varied simultaneously (some of them in a wider range) and analyzed with GSA approach, we were able to detect both nonlinear and interaction effects (total effects) in addition to the individual role of the corresponding modeled phenomena. On one hand, the results here obtained by GSA mostly confirm the insights obtained by local sensitivity analysis. As an example, we had the confirmation that the alignment term changes very little the outcome of the experiment (very low sensitivity) as it was hypothesized in the previous work [16]. On the other hand, it is worth noting that here we get conclusions that with simpler approach could not be deduced, thanks to the observation of total effect of model parameters. See, for example, the extraordinary effect that ξ and ζ have on the velocities in terms of total effect but not in terms of main effect, as shown in Fig. 8.

For instance, focusing on the damping parameter ζ, we further investigated system evolution in the case of an almost negligible inertia. Specifically, considering the case study used in Section 3.2.1, the Morris method has been performed assuming ζ uniformly distributed within the range $[10^{-16}, 2 \cdot 10^{-16}]$, while the other 12 parameters are uniformly distributed within the ranges in Table 1. The obtained results show that during $\mathcal{I}_1$ and $\mathcal{I}_2$, the ICs move towards the bottom side of the domain faster than in the control case so that in $\mathcal{I}_3$ almost all the ICs have unrealistically left the domain. This suggests that, in order to reproduce IC dynamics over a CoC by neglecting inertial terms,

i.e., with a first-order ODE instead of Eq. (2), it is first necessary to properly revise the calibration of all the other model parameters, for instance, by reducing cell chemical sensing ξ and reactivity γ.

The Morris sensitivity measures can be used as indications about which parameters may be first considered in the re-calibration of the model, being the most affecting ones. However, this approach requires further investigations and analysis to be properly formalized.

Moreover, having observed in Fig. 7 that the distribution and number of TCs have a relevant effect, a deep investigation of the role of the initial condition is necessary. In this perspective, we have studied whether and how TCs configuration influences the evolution of the reference simulation in Section 3.1. Specifically, 1000 different model realizations have been obtained by assuming the same settings used in Section 3.1, apart from the number and distribution of TCs. These have been randomly generated for each model realization assuming that the number of TCs within each portion Ω_q, with $q = 1, ..., 4$, of the domain is uniformly distributed between 2 and 12. Interestingly, over all the three-time intervals, the obtained output quantities result more condensed around the values of the simulation reported in Section 3.1, than in Figs. 4-5 resulting from the 280 realizations identified with the Morris method. This suggests that the influence of the initial distribution of TCs is somehow less impacting than the choice of the model parameters.

4. Conclusions and future perspectives

The sensitivity analysis carried out in this work shows the effect of almost all model parameters on the system behavior and highlights that the model is able to describe the crucial role of tumor microenvironment on IC dynamics. To our knowledge, a GSA producing Sobol indices identifying the role of parameters in CoC *in silico* realizations represents an original contribution to the literature. We also believe that a rigorous GSA including the initial distribution of TCs and/or the ICs inflow among the investigated parameters, requiring the identification of a proper and sufficient samples, deserves to be investigated in future works.

The indications provided by this type of analysis may drive a fine tuning of the most influential parameters against available experimental data, in order to improve the *in silico* model in reproducing and forecasting cell dynamics observed in CoC biological experiments. To this aim, a large amount of information and data extracted from laboratory tests is necessary. A first attempt in this direction was made in [15] where ICs velocity field was computed numerically and compared with an interpolated synthetic dataset.

In order to get a higher adherence of the model to CoC biological experiment it will be necessary to: i) collect data from multiple realizations of the same biological experiment; ii) extract more experimental measures, such as:

- the localization of ICs in the microfluidic chip and in particular around TCs across time in order to reconstruct the density in the monitored domain;
- the identification of different cell species involved in the ICs population;
- the death rate of TCs;
- the number of interactions between the two cell populations, in order to evaluate the killing activity of ICs;
- the chemoattractant concentration in the microenvironment, possibly detected at different space and time points.

From a mathematical point of view, in order to set up a calibration procedure against real data, it is recommended to extend the computational domain and avoid the artificial Robin condition used here for numerical convenience. For instance, a multiscale model can be obtained by combining our model with the macroscopic approach in [33], in order to describe the overall dynamics in the entire CoC geometry with a special focus on the tumor microenvironment represented by the TC chamber. Having more specific data about chemical gradients will allow us to set proper initial and boundary conditions, and possibly consider space-dependent coefficients (e.g. for chemical diffusion or cell-substrate adhesion). In addition, a further step to improve the *in silico* model of the CoC experiment is to include the death of TCs as either spontaneous (as a consequence of drug administration) and/or induced by the activity of ICs.

Acknowledgements. G. B. and E. C. are members of the Gruppo Nazionale Calcolo Scientifico, Istituto Nazionale di Alta Matematica (GNCS-INdAM). A. C. is member of the Gruppo Nazionale di Fisica Matematica, Istituto Nazionale di Alta Matematica (GNFM-INdAM).

Funding. This work was supported by the Italian Ministry of University and Research [Project code PNC0000001, CUP B53C22006100001 - project "D3-4H—Digital Driven Diagnostics, prognostics and therapeutics for sustainable Health care" under the National Plan for Complementary Investments to the NRRP, Spoke 3 and 4]; the European Union - Next Generation EU and the Italian Ministry of University and Research [Project code P2022KHFNB, CUP E53D23017990001 - Research Project Prin2022 PNRR of National Relevance]; the Istituto Nazionale di Alta Matematica INdAM [CUP E53C23001670001 - "INdAM–GNCS Project"].

Appendix

Appendix A. Numerical Approximation of the model

We consider a square domain $\Omega = [x_L, x_R] \times [y_B, y_T]$ with $x_L = y_B = 0\ \mu m$ and $x_R = y_T = 600\ \mu m$, then it has horizontal and vertical size $L_x = L_y = 600 \mu m$. The

computational domain Ω is depicted with dotted blue line in Panel B of Fig. 1. We introduce a discretization on L_x in $N_x - 1$ subintervals of length $\Delta x = L_x/(N_x - 1)$ and a discretization on L_y in $M_y - 1$ subintervals of length $\Delta y = L_y/(M_y - 1)$. Time and spatial steps are chosen, respectively, as $\Delta t = 10s$, i.e. representing $1/12$ of the video footage timeframe (of 2 minutes) and $\Delta x = \Delta y = 5\mu m$. A Cartesian grid Ω_Δ is created consisting of grid points (x_n, y_m), where $x_n = x_L + n\Delta x$, for $n = 0, ..., N_x - 1$ and $y_m = y_B + m\Delta y$, for $m = 0, ..., M_y - 1$, and for the time interval $[0, T]$, the k-th temporal step t_k is $t_k = k\Delta t$, for $k = 0, ..., N_{\Delta t}$.

The simulations are performed over a number of frames equal to 8640, corresponding to a final time $T = 86400s$ (24h) of observations.

Besides, since experimentally it was observed that ICs leave the domain Ω, we added a ghost grid to Ω_Δ in order to manage the entrance and exit of cells and avoid numerical instabilities. The ghost grid, where the cells lie after having left the main domain, is defined as Ω_{Δ^*}. The extension of the nunerical grid is obtained by discretizing the domain $[x_L - L_{x^*}, x_R + L_{x^*}] \times [y_B - L_{y^*}, y_T + L_{y^*}]$ with the same discretization step values $\Delta x, \Delta y$ already introduced for Ω_Δ. In our tests we assume $L_{x^*} = L_{y^*} = 100\mu m$ and the equations are solved on Ω_{Δ^*}.

Appendix A.1. Discretization of the PDE in Eq. (1)

The parabolic Eq. (1) is composed of the diffusion term, the source term, and the stiff degradation term $-\eta\varphi$. Then, the classical exponential transformation is applied: $\varphi(\mathbf{x}, t) = e^{-\eta t} u(\mathbf{x}, t)$, which leads to the diffusion equation with source for $u(\mathbf{x}, t)$:

$$\partial_t u = D\Delta u + e^{\eta t}\xi \sum_{j=1}^{N_T} \chi_{\mathbf{B}(\mathbf{Y}_j, R_T)}. \tag{A.1}$$

For this equation we apply a central difference scheme in space, i.e. the 5-point stencil for the Laplacian, and the parabolic Crank–Nicolson scheme in time.
Denoting with $u^k_{n,m}$ the approximation of u at the grid point (x_n, y_m), for any time instant t_k the numerical scheme can be written as:

$$\begin{aligned}\frac{u^{k+1}_{n,m} - u^k_{n,m}}{\Delta t} = \frac{D}{2}\left(\partial_x^2 u^{k+1} + \partial_y^2 u^{k+1}\right) + \frac{D}{2}\left(\partial_x^2 u^k + \partial_y^2 u^k\right) \\ + \frac{1}{2}e^{\eta(k+1)\Delta t}\xi \sum_{j=1}^{N_T} \chi_{\mathbf{B}(\mathbf{Y}_j^{k+1}, R_T)} + \frac{1}{2}e^{\eta k\Delta t}\xi \sum_{j=1}^{N_T} \chi_{\mathbf{B}(\mathbf{Y}_j^{k}, R_T)},\end{aligned}$$

where $\partial_x^2 u^k$ (and analogously $\partial_y^2 u^k$) is defined as the central difference:

$$\partial_x^2 u^k = \frac{u^k_{n-1,m} - 2u^k_{n,m} + u^k_{n+1,m}}{\Delta x^2}.$$

Appendix A.1.1. Boundary conditions

Using the exponential transformation above, the associated boundary conditions rewrite as $D\dfrac{\partial u}{\partial \mathbf{n}} + au = e^{\eta t} b$, that we rewrite as $\dfrac{\partial u}{\partial \mathbf{n}} + pu = q(t)$, with $p = \dfrac{a}{D}$ and $q(t) = e^{\eta t}\dfrac{b}{D}$ eventually different on each side of Ω. For the discretization of the boundary conditions we use a central finite difference scheme. At the bottom and top boundaries, i.e. $y = y_B$ and $y = y_T$, we have:

$$\partial_y u^k + p_l u^k - q_l^k = \frac{u^k_{n,m+1} - u^k_{n,m-1}}{2\Delta y} + r_l u^k_{n,m} - h_l, \tag{A.2}$$

with $l = S, N$, and we proceed analogously for the vertical conditions, with the signs of r_l and h_l depending on the incoming/outgoing flow.

Appendix A.2. Discretization of the ODE

As in [15], the equation of motion (2) is reduced to a first order system with

$$\begin{aligned}\dot{\mathbf{V}}_i = \frac{\gamma}{\mathcal{W}} \int_{\mathbf{B}(\mathbf{X}_i,\mathbf{R}_0)} \chi(\varphi(\mathbf{x},t))\nabla\varphi(\mathbf{x},t) w_i(\mathbf{x}) d\mathbf{x} + \sum_{j:\mathbf{Y}_j \in \mathbf{B}(\mathbf{X}_i,\mathbf{R}_1)\setminus\{\mathbf{X}_i\}} K(\mathbf{Y}_j - \mathbf{X}_i) \\ + \sum_{j:\mathbf{X}_j \in \mathbf{B}(\mathbf{X}_i,\mathbf{R}_4)\setminus\{\mathbf{X}_i\}} K(\mathbf{X}_j - \mathbf{X}_i) + \frac{\beta}{N_i} \sum_{j:\mathbf{X}_j \in B(\mathbf{X}_i,R_3)\setminus\{\mathbf{X}_i\}} \frac{(\mathbf{V}_j - \mathbf{V}_i)}{\left(1 + \frac{\|\mathbf{X}_j - \mathbf{X}_i\|^2}{R_3^2}\right)^{\alpha}} - \zeta \mathbf{V}_i,\end{aligned} \tag{A.3}$$

$$\dot{\mathbf{X}}_i = \mathbf{V}_i, \tag{A.4}$$

for $i = 1, \ldots, N_I$. Eq. (A.3) is discretized with a one step IMEX method, putting in implicit the term containing $\mathbf{V}_i$ and in explicit the other addends, and equation (A.4) is solved with forward Euler method. The two-dimensional integral in Eq. (A.3) can be computed by a 2D quadrature formula, which due to the truncated Gaussian weight function $w_i(\mathbf{x})$ in Eq. (3), is reduced to a sum of the discretized integrand functions on the grid points belonging to the ball $\mathbf{B}(\mathbf{X}_i, \mathbf{R}_0)$. The two-dimensional integral in $\mathcal{W}$ in Eq. (3) is approximated by $\widetilde{\mathcal{W}} := \sum_{n,m \text{ s.t.}(x_n,x_m)\in\mathbf{B}(\mathbf{X}_i^k,\mathbf{R}_0)} (w_i)^{(k)}_{n,m}$. The gradients of Eq. (A.3) are approximated with first order differences $\nabla_{n,m}\varphi^k \approx \left(\dfrac{\varphi^k_{n+1,m} - \varphi^k_{n,m}}{\Delta x}, \dfrac{\varphi^k_{n,m+1} - \varphi^k_{n,m}}{\Delta y}\right)$, then Eq. (A.3)-(A.4) are discretized as follows:

$$\frac{\mathbf{V}_i^{k+1} - \mathbf{V}_i^k}{\Delta t} = \frac{\gamma}{\widetilde{\mathcal{W}}} \sum_{n,m \text{ s.t.}(x_n,x_m)\in\mathbf{B}(\mathbf{X}_i^k,\mathbf{R}_0)} \chi(\varphi^k)(\nabla_{n,m}\varphi^k)(w_i)^{(k)}_{n,m}$$

$$+ \sum_{j:\mathbf{X}_j^k \in \mathbf{B}(\mathbf{X}_i^k, \mathbf{R}_1)\setminus\{\mathbf{X}_i^k\}} K(\mathbf{X}_j^k - \mathbf{X}_i^k) + \sum_{j:\mathbf{X}_j^k \in \mathbf{B}(\mathbf{Y}_i^k, \mathbf{R}_4)\setminus\{\mathbf{Y}_i^k\}} K(\mathbf{Y}_j^k - \mathbf{X}_i^k)$$

$$+ \frac{\beta}{N_i} \sum_{j:\mathbf{X}_j^k \in B(\mathbf{X}_i^k, R_3)\setminus\{\mathbf{X}_i^k\}} \frac{(\mathbf{V}_j^k - \mathbf{V}_i^{k+1})}{\left(1 + \frac{\|\mathbf{X}_j^k - \mathbf{X}_i^k\|^2}{R_3^2}\right)^\alpha} - \zeta \mathbf{V}_i^{k+1},$$

$$\frac{\mathbf{X}_i^{k+1} - \mathbf{X}_i^k}{\Delta t} = \mathbf{V}_i^{k+1}.$$

Remark. *On a laptop equipped with an Intel Core i7-1060NG7 processor and 16 GB RAM each run of the numerical code, by considering the case study scenarion, given the time and spatial discretizations here, takes around* 15 min.

Appendix B. Additional details about GSA methods

Appendix B.1. The sampling strategy by Campolongo et al.

For readers' convenience we here summarize the sampling strategy at the basis of the improved version of the Morris method. It is first worth to notice the approach is designed for d model parameters assumed independent and uniformly distributed over $[0, 1]$, so that the region of interest in the space of parameter is the d-dimensional unit hypercube. However, it is straightforward that different parameter distributions can be easily reduced to this case through proper transformations. The region of interest is sampled following the optimized strategy defined in [3]. The d-dimensional unit hypercube is first discretized into a d-dimensional p-level grid, and a large number $M \approx 500 - 1000$ of Morris trajectories is generated. Each Morris trajectory starts from a randomly selected point of the d-dimensional p-level grid, and is obtained by performing d steps, one for each dimension, by randomly defining both the direction of motion (random choice without re-entry) and the length of each step (which is always a random multiple of the grid size, i.e. $1/p$). This way, the difference between two successive points in a Morris trajectory is equal to $\Delta \mathbf{e}_k$ where Δ is a multiple of the grid size $1/p$. Then, the 'best' $r \ll M$ Morris trajectories that maximise the coverage of the d-dimensional unit hypercube are selected and actually used to rank the model parameters. The coverage of the hypercube is estimated by evaluating the distance D_{ml} between each couple of trajectories m and l, defined in [3] as the sum of the Euclidean distances between all couples of their points. The r best Morris trajectories that maximize D_{ml} thus identify a sample of $r(d+1)$ points $\mathbf{z}$ in the space of parameters. The model is then solved $r(d+1)$ times to compute the corresponding values of the target output $Q(\mathbf{z})$ and estimate the elementary effects in Eq. (19). Notice that, by construction, the EE_k can be evaluated by considering only pair of successive points in a Morris trajectory. This means that each Morris trajectory gives k elementary effect values, one for each parameter, so that the sensitivity measures μ_k^* and σ_k are evaluated on r different values

of EE_k. Interestingly, the method is able to capture the same qualitative information provided by the variance-based method [6] (regarded as good practice in sensitivity analysis) without requiring large r, nor an extremely small grid size $1/p$. In this respect, referring to [3], we here always set $p = 4$, $M = 500$ and $r = 20$, and thus consider a sample of 280 points in the 13-dimensional space of parameters.

Appendix B.2. Computing the first-order and total-effect indices with eFAST

The extended Fourier Amplitude Sensitivity Test (eFAST) method proposed by Saltelli et al. [10], is an extended version of the FAST method by Cukier et al. [46] able to estimate also the total effects of the model parameters, rather than just the main effects. The basic idea of these approaches is that periodic oscillations of relevant parameters propagate to the model output more than periodic oscillations of non-influential parameters. Let us consider the model as a function linking the parameters $\mathbf{Z} = (Z_1, ..., Z_d)$ to the value of the output Q, i.e. $Q = f(\mathbf{Z})$. Assuming the model parameters are independent random variables, the Sobol' indices in Eq. (21) are estimated by using the monodimensional Fourier decomposition along a so-called *search curve* defined by the set of parametric equations

$$Z_k(s) = G_k(\sin \omega_k s), \quad \text{with } k = 1, ..., d. \tag{B.1}$$

In Eq. (B.1), the transformation functions G_k depend on the probability distribution of each parameter. The frequencies ω_k are d distinct positive integer frequencies chosen to satisfy several criteria given in [46, 10, 45]. It follows that varying s, i.e. moving along the search curve, all the model parameters change simultaneously, but each Z_k oscillates periodically at the corresponding frequency ω_k. As demonstrated in [45], with integer frequencies, the search curve yields a closed path entirely traveled as $s \in (-\pi, \pi)$. In turn, the output function f, as a function of s, is such that $f(s) = f(s + 2\pi)$ and can be Fourier analyzed. It is worth to stress that assuming a distinct frequency for each parameter allows to distinguish their effects during the Fourier analysis. Let us denote the Fourier coefficients of the output function by

$$A_j = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(s) \cos(js)\, ds \qquad B_j = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(s) \sin(js)\, ds \tag{B.2}$$

with $j \in \mathbb{Z}$; and the spectrum of the Fourier series expansion by $\Lambda_j = A_j^2 + B_j^2$ with $j \in \mathbb{Z}$. Noticing that Λ_0 estimate the expected mean of the output, while the second momentum is given by the sum of all harmonics, the total variance of the output writes

$$\mathrm{Var}(Q) = 2 \sum_{j=1}^{+\infty} \Lambda_j. \tag{B.3}$$

To estimate the main effect indices $S_k(Q)$ in Eq. (21), the variance of the output arising from the uncertainty of the k-th parameter is given by the spectrum of the frequency ω_k and its higher harmonics, see [46], i.e.

$$D_k(Q) = 2 \sum_{p=1}^{+\infty} \Lambda_{p\omega_k}. \tag{B.4}$$

The total-order sensitivity indices S_k^T in Eq (21) is estimated by the difference $1 - S_{-k}(Q)$, where $S_{-k}(Q)$ is the summed sensitivity index of all parameters except k, obtained by using their identification frequencies in Eq. (B.4), see [10].

From a computational point of view, to evaluate the above integrals, a sample of points over the search curve is selected. Following the improved sampling approach described in [5], the analysis in Section 3.2.2 is performed by considering a sample of 3000 different points into the 6-dimensional space of parameter.

References

[1] Agliari, E., Biselli, E., De Ninno, A., et al., *Cancer-driven dynamics of immune cells in a microfluidic environment*, Scientific Reports 4, 6639 (2014). https://doi.org/10.1038/srep06639

[2] Biselli, E., Agliari, E., Barra, A. et al., *Organs on chip approach: a tool to evaluate cancer-immune cells interactions* Scientific Reports 7, 12737 (2017). https://doi.org/10.1038/s41598-017-13070-3

[3] Campolongo, F., Cariboni, J., Saltelli, A., *An effective screening design for sensitivity analysis of large models*, Environmental Modelling & Software, 22(10):1509-1518 (2007). https://doi.org/10.1016/j.envsoft.2006.10.004

[4] Morris, M. D., *Factorial sampling plans for preliminary computational experiments*, Technometrics, 33(2):161-174 (1991). https://doi.org/10.2307/1269043

[5] Saltelli, A., *Making best use of model evaluations to compute sensitivity indices*, Computer Physics Communications, 145(2):280–297 (2002). https://doi.org/10.1016/S0010-4655(02)00280-1

[6] Saltelli, A., Tarantola, S., Campolongo, F., Ratto, M., *Sensitivity analysis in practice: a guide to assessing scientific models*, Wiley Online Library, (2004). https://doi.org/10.1002/0470870958

[7] Sobol, I. M., *Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates*, Mathematics and Computers in Simulation, 55(1-3):271-280 (2001). https://doi.org/10.1016/S0378-4754(00)00270-6

[8] Iwanaga, T., Usher, W., Herman, J., *Toward SALib 2.0: Advancing the accessibility and interpretability of global sensitivity analyses*, Socio-Environmental Systems Modelling, 4:18155, (2022). https://doi.org/10.18174/sesmo.18155

[9] Herman, J. and Usher, W., *SALib: An open-source Python library for Sensitivity Analysis*, The Journal of Open Source Software, 2(9):97 (2017). https://doi.org/10.21105/joss.00097

[10] Saltelli, A., Tarantola, S., Chan, K.P.-S. *A Quantitative Model-Independent Method for Global Sensitivity Analysis of Model Output*, Technometrics, 41(1):39-56 (1999). https://doi.org/10.1080/00401706.1999.10485594

[11] Liu, D., Li, L., Rostami-Hodjegan, A. et al., *Considerations and Caveats when Applying Global Sensitivity Analysis Methods to Physiologically Based Pharmacokinetic Models*, The AAPS Journal, 22(93) (2020). https://doi.org/10.1208/s12248-020-00480-x

[12] Parlato, S., Grisanti, G., Sinibaldi, G., et al., *Tumor-on-a-chip platforms to study cancer–immune system crosstalk in the era of immunotherapy*, Lab Chip, 21(2):234-253 (2021).

[13] Mattei, F., Andreone, S., Mencattini, A., Businaro, L., et al., *Oncoimmunology meets Organs-on-chip*, Front Mol Biosci, 8 (2021). https://doi.org/10.3389/fmolb.2021.627454

[14] Vacchelli, E., Ma, Y., Baracco, E. E., Sistigu, A. et al., *Chemotherapy-induced antitumor immunity requires formyl peptide receptor 1*, Science, 350(6263):972-978 (2015). https://doi.org/10.1126/science.aad0779

[15] Bretti, G., De Ninno, A., Natalini, R., Peri, D., Roselli, N., *Estimation Algorithm for a Hybrid PDE–ODE Model Inspired by Immunocompetent Cancer-on-Chip Experiment*, Axioms, 10(4) (2021). https://doi.org/10.3390/axioms10040243

[16] Bretti, G., Campanile, E., Menci, M., Natalini, R., *A Scenario-Based Study on Hybrid PDE-ODE Model for Cancer-on-Chip Experiment*, In: d'Onofrio, A., Fasano, A., Papa, F., Sinisgalli, C. (eds) Problems in Mathematical Biophysics. SEMA SIMAI Springer Series, vol 38. Springer, Cham. https://doi.org/10.1007/978-3-031-60773-8_3

[17] Evelyn, F. K., Lee, A. S., *Initiation of Slime Mold Aggregation Viewed as an Instability*, Journal of Theoretical Biology, 26(3):399-415 (1970). https://doi.org/10.1016/0022-5193(70)90092-5

[18] Maulana, T. I., Kromidas, E., Wallstabe, L., Cipriano, M. et al., *Immunocompetent cancer-on-chip models to assess immuno-oncology therapy*, Advanced Drug Delivery Reviews (2021). https://doi.org/10.1016/j.addr.2021.03.015

[19] Businaro, L., De Ninno, A., Schiavoni, G. et al., *Cross talk between cancer and immune cells: exploring complex dynamics in a microfluidic environment*, Lab Chip, 13(2) (2013). https://doi.org/10.1039/c2lc40887b

[20] Low, L. L., Mummery, C., Berridge, B. R., Austin, C. P., Tagle, D. A., *Organs-on-chips: into the next decade*, Nature Reviews Drug Discovery, 20:345-361 (2020). https://doi.org/10.1038/s41573-020-0079-3

[21] Gori, M., Simonelli, M. C., Giannitelli, S. M., Businaro, L. et al., *Investigating Nonalcoholic Fatty Liver Disease in a Liver-on-a-Chip Microfluidic Device*, PLOS ONE, 11:1-15 (2016). https://doi.org/10.1371/journal.pone.0159729

[22] Curk, T., Marenduzzo, D., Dobnikar, J., *Chemotactic Sensing towards Ambient and Secreted Attractant Drives Collective Behaviour of E. coli*, PLOS ONE, (2013). https://doi.org/10.1371/journal.pone.0074878

[23] Murray, J. D., *Mathematical Biology II: spatial models and biomedical applications* Springer New York, (2003). https://doi.org/10.1007/b98869

[24] Boulter E., Grall, D., Cagnol, S., Van Obberghen-Schilling E. et al., *Regulation of cell-matrix adhesion dynamics and Rac-1 by integrin linked kinase*, The FASEB Journal, 20(9):1489-1491 (2006). https://doi.org/10.1096/fj.05-4579fje

[25] Wiśniewski, J. R., Hein, M. Y., Cox, J., Mann, M., *A "Proteomic Ruler" for Protein Copy Number and Concentration Estimation without Spike-in Standards*, Mol Cell Proteomics, 13(12):3497-3506 (2014). https://doi.org/10.1074/MCP.M113.037309

[26] Di Costanzo, E., Natalini, R., Preziosi, L., *A hybrid mathematical model for self-organizing cell migration in the zebrafish lateral line*, Journal of Mathematical Biology, 71:171-214 (2015). https://doi.org/10.1007/s00285-014-0812-9

[27] Cucker, F., Smale, S., *Emergent Behavior in Flocks*, IEEE Transactions on Automatic Control, 52(5):852-862 (2007). https://doi.org/10.1109/TAC.2007.895842

[28] Lapidus, I. R., Schiller, R., *Model for the chemotactic response of a bacterial population*, Biophysical Journal, 16(7):779-789 (1976). https://doi.org/10.1016/S0006-3495(76)85728-1

[29] De Ninno, A., Bertani, F. R., Gerardino, A., Schiavoni, G. et al., *Microfluidic co-culture models for dissecting the immune response in in vitro tumor microenvironments*, Journal of Visualized Experiments, 30 (2021). https://doi.org/10.3791/61895

[30] Kirschner, D., Panetta, J. C., *Modeling immunotherapy of the tumor-immune interaction*, Journal of Mathematical Biology, 37:235-252 (1998). https://doi.org/10.1007/s002850050127

[31] Wei, S., Lee, L., Seager, R. J., Litvak, F. et al., *Integrated in silico and 3D in vitro model of macrophage migration in response to physical and chemical factors in the tumor microenvironment*, Integrative Biology (Camb), 12(4):90-108 (2020). https://doi.org/10.1093/intbio/zyaa007

[32] Pompa, M., Torre, D., Bretti, G., De Gaetano, A., *Sensitivity Analysis of a 2D Stochastic Agent-Based and PDE Diffusion Model for Cancer-on-Chip Experiments*, Axioms, 12(10) (2023). https://doi.org/10.3390/axioms12100930

[33] Braun, E. C., Bretti, G., Natalini, R., *Mass-Preserving Approximation of a Chemotaxis Multi-Domain Transmission Model for Microfluidic Chips*, Mathematics, 9(6) (2021). https://doi.org/10.3390/math9060688

[34] Braun, E. C., Bretti, G., Natalini, R., *Parameter estimation techniques for a chemotaxis model inspired by Cancer-on-Chip (COC) experiments*, International Journal of Non-Linear Mechanics, 140:103895 (2022). https://doi.org/10.1016/j.ijnonlinmec.2021.103895

[35] Bretti, G., De Gaetano, A., *An Agent-Based Interpretation of Leukocyte Chemotaxis in Cancer-on-Chip Experiments*, Mathematics, 10(8):1338 (2022). https://doi.org/10.3390/math10081338

[36] Di Costanzo, E., Menci, M., Messina E., Natalini, R., Vecchio, A., *A hybrid model of collective motion of discrete particles under alignment and continuum chemotaxis*, Discrete and Continuous Dynamical Systems - B, 25(1) (2020). https://doi.org/10.3934/dcdsb.2019189

[37] Menci, M., Papi, M., *Existence of solutions for hybrid systems of differential equations under exogenous information with discontinuous source term*, Nonlinear Analysis, 221:112885 (2022). https://doi.org/10.1016/j.na.2022.112885

[38] Menci, M., Papi, M., Porzio, M. M., Smarrazzo, F., *On a coupled hybrid system of nonlinear differential equations with a nonlocal concentration*, Journal of Differential Equations, 361:288-338 (2023). https://doi.org/10.1016/j.jde.2023.02.044

[39] Menci, M., Natalini, R., Paul, T., title=Microscopic, kinetic and hydrodynamic hybrid models of collective motions with chemotaxis: a numerical study, ArXiv (2023). https://api.semanticscholar.org/CorpusID:259224724

[40] Saltelli, A., Tarantola, S., Campolongo, F., *Sensitivity Analysis as an Ingredient of Modeling*, Statistical Science, 15(4):377–395 (2000). https://www.jstor.org/stable/2676831

[41] Iooss, B., Lemaître, P., (2015) *A Review on Global Sensitivity Analysis Methods*. In: Dellino, G., Meloni, C. (eds) Uncertainty Management in Simulation-Optimization of Complex Systems. Operations Research/Computer Science Interfaces Series, vol 59. Springer, Boston, MA. https://doi.org/10.1007/978-1-4899-7547-8_5

[42] Saltelli, A., *Global sensitivity analysis. The primer*, John Wiley & Sons, Ltd (2008). https://doi.org/10.1002/9780470725184.fmatter

[43] Zarrintaj, P., Saeb, M. R., Stadler, F. J., Yazdi, M. K. et al., *Human Organs-on-Chips: A Review of the State-of-the-Art, Current Prospects, and Future Challenges*, Advanced Biology, 6(1):2000526 (2021). https://doi.org/10.1002/adbi.202000526

[44] Homma, T., Saltelli, A., *Importance measures in global sensitivity analysis of nonlinear models*, Reliability Engineering & System Safety, 52(1):1-17 (1996). https://doi.org/10.1016/0951-8320(96)00002-6

[45] Cukier, R. I., Fortuin, C. M., Shuler, K. E., Petschek, A. G., Schaibly, J. H., *Study of the sensitivity of coupled reaction systems to uncertainties in rate coefficients. I Theory*, The Journal of Chemical Physics, 59:3873-3878 (1973). https://doi.org/10.1063/1.1680571.

[46] Cukier, R. I., Levine, H. B., Shuler, K. E., *Nonlinear sensitivity analysis of multiparameter model systems*, Journal of Computational Physics, 26(1):1-42 (1978). https://doi.org/10.1016/0021-9991(78)90097-9

[47] Colombi, A., Bancheri, M., Acutis , M., Basile, A., Botta, M., Perego, A., *A sound understanding of a cropping system model with the global sensitivity analysis*, Environmental Modelling & Software, 173:105932 (2024). https://doi.org/10.1016/j.envsoft.2023.105932

[48] Bregaglio, S., Giustarini, L., Suarez, E., Mongiano, G., De Gregorio, T., *Analysing the behaviour of a hazelnut simulation model across growing environments via sensitivity analysis and automatic calibration*, Agricultural Systems, 181:102794 (2020). https://doi.org/10.1016/j.agsy.2020.102794

[49] Specka, X., Nendel, C., Wieland, R., *Analysing the parameter sensitivity of the agro-ecosystem model MONICA for different crops*, European Journal of Agronomy, 71:73-87 (2015). https://doi.org/10.1016/j.eja.2015.08.004

[50] Confalonieri, R., Bregaglio, S., Cappelli, G., Francone, C. et al., *Wheat modeling in Morocco unexpectedly reveals predominance of photosynthesis versus leaf area expansion plant traits*, Agronomy for Sustainable Development, 33:393-403 (2013). https://doi.org/10.1007/s13593-012-0104-y

[51] DeJonge, K.C., Ascough , J.C., Ahmadi, M., Andales, A. A., Arabi, M., *Global sensitivity and uncertainty analysis of a dynamic agroecosystem model under different irrigation treatments*, Ecological Modelling, 231:113-125 (2012). https://doi.org/10.1016/j.ecolmodel.2012.01.024

[52] Silvestro, P. C., Pignatti, S., Yang, H., Yang, G. et al., *Sensitivity analysis of the Aquacrop and SAFYE crop models for the assessment of water limited winter wheat yield in regional scale applications*, PLOS ONE, 12(11):1-30 (2017) https://doi.org/10.1371/journal.pone.0187485

[53] Xiang Z., Bailey, R. T., Kisekka, I., *Using DSSAT-MODFLOW to determine the controls of groundwater storage and crop yield in groundwater-based irrigated regions*, Journal of Hydrology, 612:128161 (2022). https://doi.org/10.1016/j.jhydrol.2022.128161

[54] Sobol, I. M., *Sensitivity estimates for nonlinear mathematical models*, Matematicheskoe Modelirovanie, 2:112–118 (1990), in Russian, translated in English in Sobol' (1993).

[55] Sobol, I. M., *Sensitivity analysis for non-linear mathematical models*, Mathematical Modelling and Computational Experiment, 1:407-414 (1993), English translation of Russian original paper Sobol' (1990).

[56] Martinson, W. D., McLennan, R., Teddy, J. M., McKinney, M. C. et al., *Dynamic fibronectin assembly and remodeling by leader neural crest cells prevents jamming in collective cell migration*, eLife, 12:e83792 (2023). https://doi.org/10.7554/eLife.83792

[57] Dela, A., Shtylla, B., de Pillis, L., *Multi-method global sensitivity analysis of mathematical models*, Journal of Theoretical Biology, 546:111159 (2022). https://doi.org/10.1016/j.jtbi.2022.111159

[58] Marino, S., Hogue, I. B., Ray, C. J., Kirschner, D. E., *A methodology for performing global uncertainty and sensitivity analysis in systems biology*, Journal

of Theoretical Biology, 254(1):178-196 (2008). https://doi.org/10.1016/j.jtbi.2008.04.011

[59] A. van der Vegt, S. A., Polonchuk, L., Wang, K., Waters, S. L., Baker, R. E., *Mathematical modelling of autoimmune myocarditis and the effects of immune checkpoint inhibitors*, Journal of Theoretical Biology, 537:111002 (2022). https://doi.org/10.1016/j.jtbi.2021.111002

[60] Checcoli, A., Pol, J. G., Naldi, A., Noel, V. et al., *Dynamical Boolean Modeling of Immunogenic Cell Death*, Front Physiol., 11:590479 (2020). https://doi.org/10.3389/fphys.2020.590479

[61] Lewin, T.D., Avignon, B., Tovaglieri, A., Cabon, L. et al., *An* in silico *Model of T Cell Infiltration Dynamics Based on an Advanced in vitro System to Enhance Preclinical Decision Making in Cancer Immunotherapy*, Front Pharmacol., 13:837261 (2022). https://doi.org/10.3389/fphar.2022.837261

[62] Yang, T.D., Park, J.-S., Choi, Y., Choi, W., Ko, T.-W., Lee, K.J., *Zigzag Turning Preference of Freely Crawling Cells*, PLOS ONE, 6:e20255 (2011). https://doi.org/10.1371/journal.pone.0020255

[63] Jarrett, A. M., Liu, Y., Cogan, N. G., Hussaini, M. Y., *Global sensitivity analysis used to interpret biological experimental results*, Journal of Mathematical Biology, 71:151–170 (2015). https://doi.org/10.1007/s00285-014-0818-3

[64] Carrillo, J. A., Colombi, A., Scianna, M., *Adhesion and volume constraints via nonlocal interactions determine cell organisation and migration profiles*, Journal of Theoretical Biology, 445:75-91 (2018). https://doi.org/10.1016/j.jtbi.2018.02.022

[65] Saltelli, A., Annoni, P., *How to avoid a perfunctory sensitivity analysis*, Environmental Modelling & Software, 25(12):1508-1517 (2010). https://doi.org/10.1016/j.envsoft.2010.04.012

[66] Li, D., Jiang, P., Hu, C., Yan, T., *Comparison of local and global sensitivity analysis methods and application to thermal hydraulic phenomena*, Progress in Nuclear Energy, 158:104612 (2023). https://doi.org/10.1016/j.pnucene.2023.104612

[67] Chen, L., Painter, K., Surulescu, C., Zhigun, A., *Mathematical models for cell migration: a non-local perspective*, Philosophical Transactions of the Royal Society B: Biological Sciences, 375:20190379, (2020). https://doi.org/10.1098/rstb.2019.0379