# IMO II Exam 2

Ana Milena Forero Pinto

April 3rd, 2022

## 1 Question 1

Consider two groups of cells that grow for 156 hours under optimal growth conditions (i.e. they are in the proximity of a blood vessel): see xlsx datasheet cellDensity_100G. **(25%)**

a. Assume there is no cell death, and that cell proliferation is a function of space availability (it is not restricted by anything other than space). Develop a model describing the temporal evolution of the cell concentration.

b. Using maximum likelihood estimation, fit the model to the two group of cells separately.

c. Visualize the data and fitted model, summarize the inferred model parameters in a table.

### 1.1 Solution

In this case we can model the space constraint on the growth rate as a logistic function, where the carrying capacity of the cell concentration indicates the maximum number of cells per volume that the environment can sustain due to space. Therefore, the growth of each of the cell types (A and B) can be modeled with ODEs as follows:

$$\frac{da(t)}{dt} = r_a a(t)\left(1 - \frac{a(t)}{K_a}\right) \tag{1}$$

$$\frac{db(t)}{dt} = r_b b(t)\left(1 - \frac{b(t)}{K_b}\right) \tag{2}$$

Where a(t) and b(t) are the concentrations of cells, of type A and B respectively, as a function of time.
The results for the fitting of both ODEs, separately, using the first data point as the initial condition for the ODE solver and calculating the mean value -across replicates- of each cell density over time, gives:

| **cells** | **growth rate** ($r$) | **carrying capacity** ($K$) |
|---|---|---|
| GROUP A | 0.032289 | 103.716168 |
| GROUP B | 0.043785 | 98.981869 |

**Table 1:** Growth rates and carrying capacities for the fitted curved of groups A and B in figure 1.

On the other hand, the distributions of the parameters are shown in figure 2. Where the parameter values obtained by fitting each replicate with the equations (1) and (2) are listed in the table 2. Note that in the figure 2, the stars correspond to the parameter values of the fitting of the mean across replicates, listed in table 1.

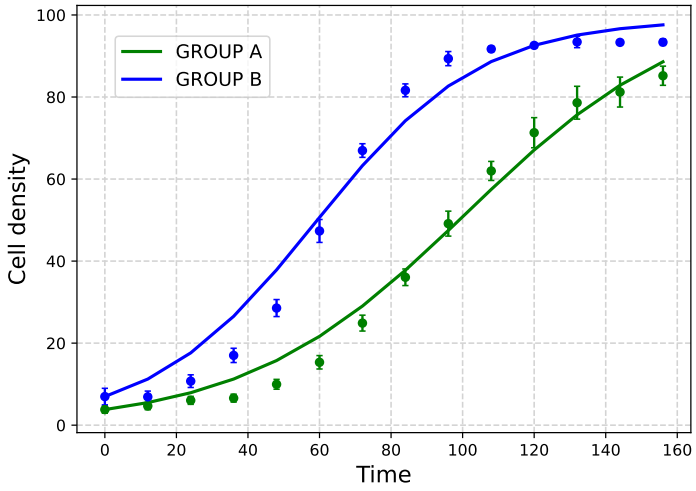



**Figure 1:** Results of the fitting by MLE - assuming the residuals to be normally distributed- for both groups of cells. Group A in green and Group B in blue. The firring was performed over the mean value -across replicates- of each cell density in the time series data. Error bars correspond to the standard error of the mean.

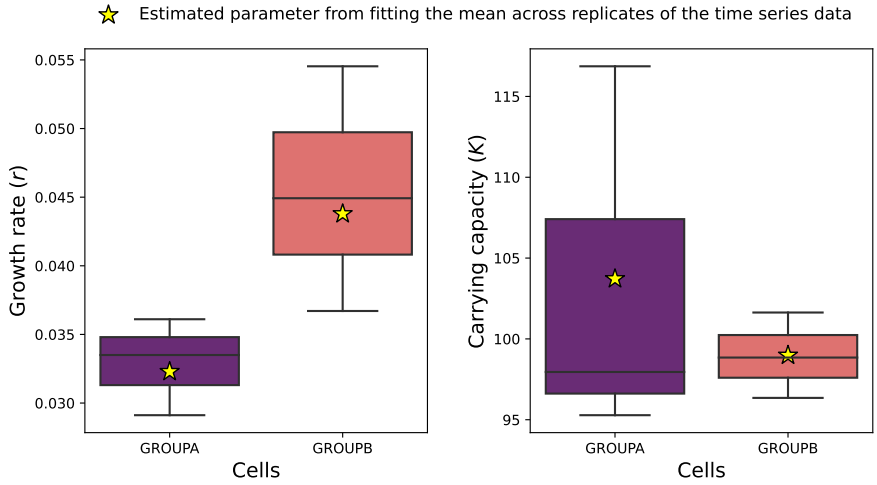



**Figure 2:** Distribution of the parameters, growth rate and carrying capacity, inferred from the fitting of each replicates for cells of groups A(purple) and B(pink). The stars are the corresponding parameter value estimated by fitting the mean across replicates of the time series data, see table 1.

| **cells** | **replicate** | **growth rate** ($r$) | **carrying capacity** ($K$) |
|---|---|---|---|
| GROUP A | 1 | 0.036106 | 97.953778 |
| GROUP A | 2 | 0.029115 | 116.864713 |
| GROUP A | 3 | 0.033499 | 95.280367 |
| GROUP B | 1 | 0.054535 | 96.350655 |
| GROUP B | 2 | 0.036707 | 101.632245 |
| GROUP B | 3 | 0.044923 | 98.846004 |

**Table 2:** Growth rates and carrying capacities from the results of the fitting of each replicate time series data. The distributions of these values per each cell group are summarized in figure 2.

# 2 Question 2

Now consider the same two groups of cells growing under sub-optimal access to nutrients (e.g. cells that are further away from blood vessels). We know: **(25%)**

a. Nutrients are initially present at high concentrations in the blood vessels (e.g. after a meal), but once consumed by the cells they are not subsequently replenished for the duration of the experiment (i.e. there is no second meal).

b. Proliferation rate is a function of nutrient access. If nutrient concentration drops below a critical level, proliferation starts to decrease.

c. Nutrient consumption is a function of proliferation rate.

d. There is no cell death.

Model these dynamics by expanding the model developed in 1) to an ODE model which explicitly describes i) Glucose dynamics (i.e. concentration and consumption by cells), and ii) dependency of cell proliferation rate on glucose concentration in addition to space availability. Provide model description/equations. List all used parameters.

## 2.1 Solution

Consider the variable $n$ as the cell density ($a$ or $b$, from 1.1). Then, the growth dynamics of the density of cells $n$, according to the conditions listed in the question, is:

$$\frac{dn(t)}{dt} = \underbrace{r_0\left(\frac{g}{\rho+g}\right)}_{r(g)}\left(1-\frac{n(t)}{K}\right) \tag{3}$$

where $r(g)$ is the proliferation rate as a function of glucose access, $g$ is the glucose concentration, and $K$ the carrying capacity. $r(g)$ is modeled using the Michaelis-Menten kinetics (a special case of the Hill function with Hill coefficient equal to 1).

On the other hand, the glucose dynamics is described by the following equation:

$$\frac{dg(t)}{dt} = -\underbrace{\alpha r}_{m(r)} n(t)g(t) \tag{4}$$

where $m(r)$ is the uptake rate per cell, and is a linear function of the proliferation rate $r(g)$, which states that cells that proliferate faster consume more glucose than cells that proliferate slowly. All the model parameters and their numerical values used in the solution of equations 3 and (4) are listed in table 3. Figure 3 shows the numerical solution when the initial values of cell density and glucose concentrations are 2.5 and 500, respectively. As can be seen in the figure, the glucose concentration decreases monotonously as indicated by equation (4) and characteristic **a.** of the question statement.

| symbol | description | value |
|---|---|---|
| $r_0$ | Proliferation rate at optimal glucose | 0.05 |
| $K$ | Carrying capacity | 50 |
| $\rho$ | Michaelis constant: concentration at half the maximal rate $r_0$ | 0.1 |
| $\alpha$ | Constant of proportionality of the uptake rate per cell $m(r)$ | 0.1 |

**Table 3:** Summary of the model parameters and their corresponding values used in the numerical solution of equations 3 and 4.

Also, from figure 3 it can be seen that the maximum cell density is different from the carrying capacity, in fact, it is a little bit lower than the value of 50. From the code file (python notebook, attached with this document) it can be seen that when time goes to infinity, the maximum value is different from $K$.

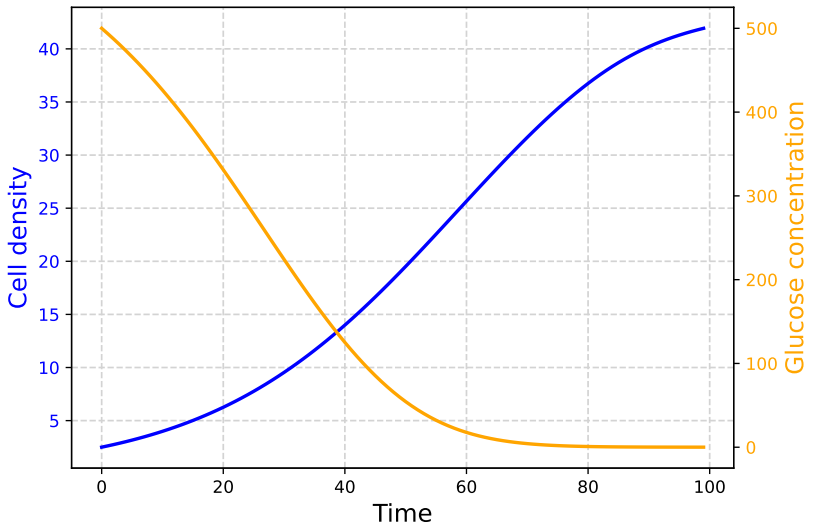



**Figure 3:** Cell density (in blue) and glucose concentration (in yellow) over time. The initial conditions for cell density and glucose concentration are 2.5 and 500, respectively. Parameters used in the calculation are listed in table 3.

# 3 Question 3

Consider cell density data recorded for the two groups of cells under depleted nutrient concentrations (20% of concentration in point 1): see xlsx datasheet cellDensity_20G. **(50%)**

a. Fit each of the two models (developed in points 1 and 2) to the data from both glucose concentrations simultaneously. Use at least two different model selection criteria to compare models 1&2 and conclude which one is superior, considering the number of model parameters.

b. Which of the inferred parameters differ between the two groups of cells? Explain how they account for differences in the observed growth behavior.

c. Which of the inferred parameters are identifiable? Would measurements of growth curves under additional nutrient concentrations improve parameter identifiability? Make recommendations for future experiments.

d. It is possible that some of the parameters for cell groups A and B are the same, whilst others are different. Perform repeated iterations of the fit from part 3a using model 2. For each iteration select a different parameter to remain fixed between groups A&B, whilst the other parameters are allowed to vary freely. Use model selection criteria to identify whether any of the resulting models are superior to the original model 2, taking into account the number of free parameters required. Use this result to comment on the most likely differences between groups A&B.

## 3.1 Solution

The parameters found from fitting the two models oped in questions 1.1 and 2.1 to the data from both glucose concentrations -100 and 20- are in tables 4 and 5 respectively. The temporal behavior of each of the cell types for the different initial glucose concentrations -100 and 20- can be seen in figures 4a and 4b. I calculated the Akaike Information and Bayesian information Criteria to conclude which of the two models is superior considering the number of parameters. The values of these measures for both models are shown in figure 4c. From this, the lower the value of these two criteria, the better the model is. Therefore, model 1 is the best one, with a value of AIC equal to 134.966598 as compared to the 138.184307 of model 2.

| symbol | description | value |
|---|---|---|
| $r_1$ | Proliferation rate of cells of type A at glucose 100 | 0.03190 |
| $k_1$ | Carrying capacity of cells of type A at glucose 100 | 103.95822 |
| $r_2$ | Proliferation rate of cells of type B at glucose 100 | 0.04338 |
| $k_2$ | Carrying capacity of cells of type B at glucose 100 | 98.32693 |
| $r_3$ | Proliferation rate of cells of type A at glucose 20 | 0.02639 |
| $k_3$ | Carrying capacity of cells of type A at glucose 20 | 67.49603 |
| $r_4$ | Proliferation rate of cells of type B at glucose 20 | 0.03916 |
| $k_4$ | Carrying capacity of cells of type B at glucose 20 | 55.04203 |

**Table 4:** Summary of the parameters of the model 1 (equations 1 and 2) for both glucose concentrations fitted simultaneously.

| symbol | description | value |
|---|---|---|
| $r_{01}$ | Proliferation rate at optimal glucose of cells of type A | 0.03982180 |
| $k_1$ | Carrying capacity of cells of type A at $g_0 = 100$ | 99.2038096 |
| $\rho_1$ | Michaelis constant of cells of type A at $g_0 = 100$ | 13.3384210 |
| $\alpha_1$ | uptake rate per cell constant of cells of type A $g_0 = 100$ | $1.0e-04$ |
| $r_{02}$ | Proliferation rate at optimal glucose of cells of type B | 0.04364732 |
| $k_2$ | Carrying capacity of cells of type B at $g_0 = 100$ | 110.0 |
| $\rho_2$ | Michaelis constant of cells of type B at $g_0 = 100$ | $1.0001e-07$ |
| $\alpha_2$ | uptake rate per cell constant of cells of type B $g_0 = 100$ | 0.11850842 |
| $k_3$ | Carrying capacity of cells of type A at $g_0 = 20$ | 65.8480636 |
| $\rho_3$ | Michaelis constant of cells of type A at $g_0 = 20$ | 8.59566612 |
| $\alpha_3$ | uptake rate per cell constant of cells of type A $g_0 = 20$ | $1.0000e-04$ |
| $k_4$ | Carrying capacity of cells of type B at $g_0 = 20$ | 109.915965 |
| $\rho_4$ | Michaelis constant of cells of type B at $g_0 = 20$ | 1.50150097 |
| $\alpha_4$ | uptake rate per cell constant of cells of type B $g_0 = 20$ | 0.07117544 |

**Table 5:** Summary of the parameters of the model 2 (equations 3 and 4) for both glucose concentrations fitted simultaneously. $g_0$ is the initial glucose concentration.

From both tables 4 and 5, none of the parameters were the same. This can be due too the minimum and maximum values I set for each parameter during the fit, because I could see some of them achieved those limit values. On the other side, from the raw data it is possible that some of the parameters are the same for the different cell types, for instance the carrying capacity (at certain initial glucose concentration) and the proliferation rate at optimal glucose $r_0$ (due to the behavior of the data at initial times). In the case of the carrying capacity, I would suggest to measure the concentration of cells for longer times in order to address the issue. Currently, this data accounts for almost 7 days, so measure cell density for additional days no further than two weeks could give better estimates of this parameter.

From my point of view, I think that the parameters related to the variation of glucose concentration are unidentifiable, since choices of different $r(g)$ and $dg/dt$ dependencies might fit the data similarly. This issue could be addressed by doing measurements of growth curves under additional nutrient concentrations and/or measure glucose concentration of the cell culture media for each specific cell type at different times in an independent experiment, or at least at the end of the same experiment reported here, in order to not alter the system too much that can affect the dynamics. If none of the experiments can be done, another alternative from the theoretical point of view would be to try different functions $r(g)$ and $dg/dt$.

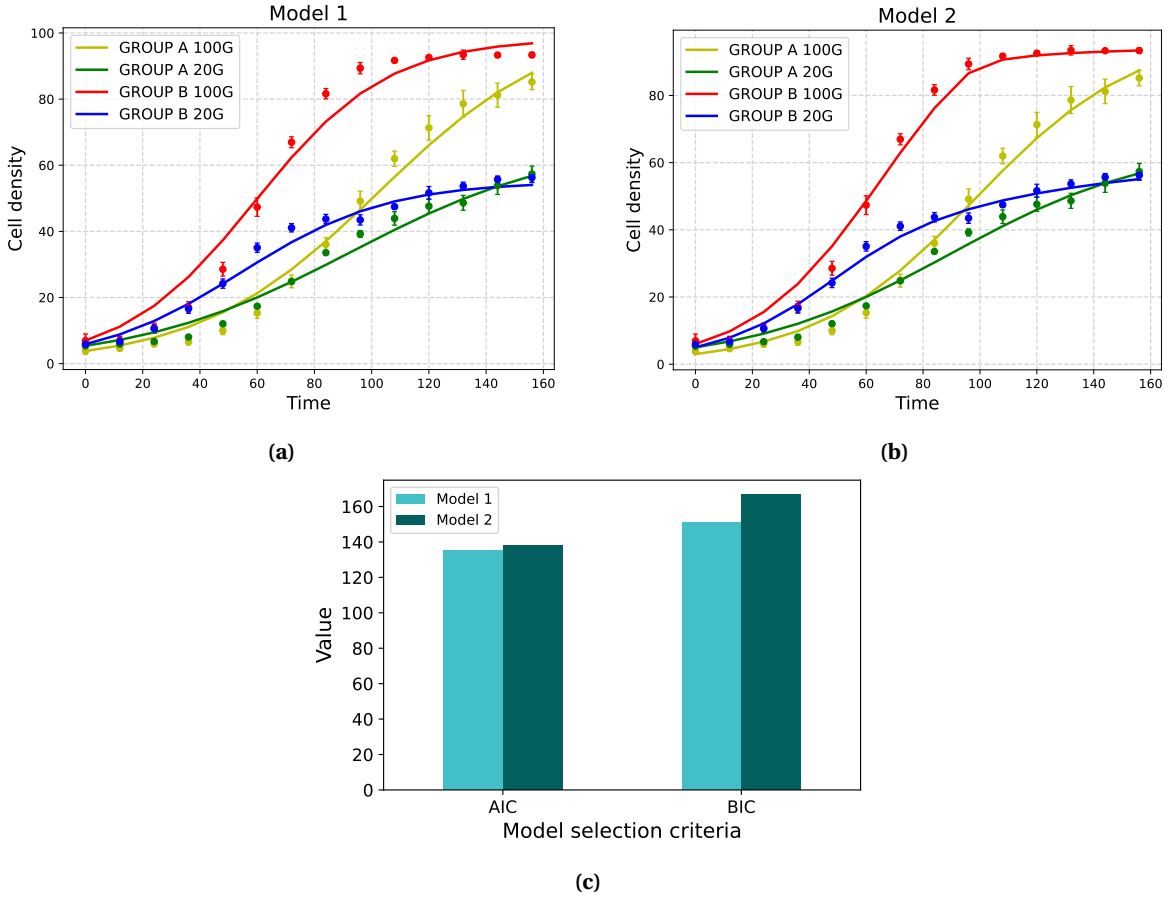
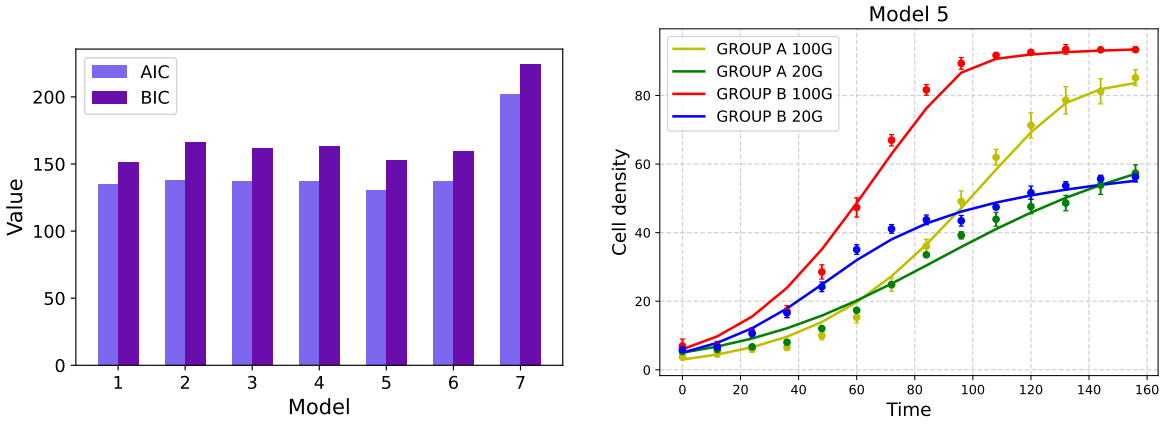



**Figure 4:** Results of the fitting for (a) model 1 and (b) model 2. The parameters corresponding to these results are listed in tables 4 and 5 respectively. Panel (c) shows the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC) values for each of the two models.

In order to study if some of the parameters for cell groups A and B are the same, I used different models described in table 6, in which one of the parameters in model 2 remain fixed between groups A and B, and across initial glucose concentrations. The model selection metrics (AIC and BIC) for each of the models are in figure 5a, and the values for the parameters corresponding to each of those can be found in the attached jupyter notebook. The figure shows that given the number of parameters, model 1 (AIC = 134.97, BIC = 151.17) and 5 (AIC = 130.70, BIC = 152.98) are the best ones. The fitted curves for model 5 are shown in figure 5b, which reflects how the carrying capacity for all of the cell types is the same (see how well is the fit at large times). Maybe another function of $r(g)$ will help to better fit for data at early times, and after testing another function, it will be worth to try to combine model 5 and 4 (the next good model for the AIC criterion), i.e. with $k$ and $r_0$ constant.

| model | description | number of parameters |
|---|---|---|
| 1 | not explicit dependency on glucose. See table 4 | 8 |
| 2 | explicit dependency on glucose. See table 5 | 14 |
| 3 | model 2 with $k_i$ $(i = 3,4)$ constant across glucose but different for cell types $(i = 1,2)$ | 12 |
| 4 | model 2 with $r_0$ constant | 13 |
| 5 | model 2 with $k_i$ $(i = 1,2,3,4)$ constant, same between groups and glucose | 11 |
| 6 | model 2 with $\rho_i$ $i = 1,2,3,4$ constant, same between groups and glucose | 11 |
| 7 | model 2 with $\alpha_i$ $i = 1,2,3,4$ constant, same between groups and glucose | 11 |

**Table 6:** Summary of the different models used to fit the data.




**(a)** **(b)**

**Figure 5:** (a) Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) values for all of the models tested. (b) Results of the fitting for model 5 (the best of model 2 variations), with parameters: $r_{01} = 0.03424307, k_1 = 109.999997, \alpha_1 = 0.10559276, \rho_1 = 4.4826e-05, r_{02} = 0.04366568, \alpha_2 = 0.11840447, \rho_2 = 1.0000e-07, \alpha_3 = 0.03109165, \rho_3 = 5.21654507, \alpha_4 = 0.07117358, \rho_4 = 1.49982305$.