

# iGEM Toronto 2018 Mathematical Modelling

B. Kell, F. Vermehr, Q. Vilchez, P. Malik

October 2018

## Contents

<b>1</b>	<b>Introduction</b>	<b>3</b>
<b>2</b>	<b>A Differential Bioreactor Model</b>	<b>3</b>
2.1	Goals . . . . .	3
2.2	Mathematical Formulation . . . . .	3
2.2.1	Concepts . . . . .	3
2.2.2	The Model . . . . .	4
2.2.3	Bioreactor Assumptions . . . . .	5
2.3	Equations . . . . .	5
2.4	Numerical Solution . . . . .	6
2.5	Sensitivity Analysis . . . . .	6
2.5.1	The Effect of Dilution Rate on Bioreactor Performance . . . . .	7
2.5.2	The Effect of Initial Metal Concentration on Bioreactor Performance . . . . .	8
2.5.3	The Effect of Attachment Rate on Bioreactor Performance . . . . .	8
2.5.4	The Effect of Detachment Rate on Bioreactor Performance . . . . .	9
2.5.5	The Effect of Initial Cellular Population on Bioreactor Performance . . . . .	9
2.5.6	The Effect of Maximum Specific Growth Rate on Bioreactor Performance . . . . .	10
2.6	Hypothetical Applications . . . . .	10
<b>3</b>	<b>A Mechanistic Buoyancy Model: Maximum Carrying Capacity Determination</b>	<b>11</b>
3.1	Goal . . . . .	11
3.2	Mathematical Formulation . . . . .	11
3.2.1	Concepts . . . . .	11
3.2.2	Assumptions . . . . .	12
3.2.3	The Model . . . . .	12
3.3	Analytic Solution . . . . .	12
3.4	Determining Buoyant Force: Temporal Tracking Algorithm . . . . .	13
3.5	Discussion On Maximum Carrying Capacity . . . . .	15
<b>4</b>	<b>Growth Dynamics</b>	<b>16</b>
4.1	Objective . . . . .	16
4.2	Approach . . . . .	16
4.3	Results . . . . .	16
<b>5</b>	<b>Discussion and Future Directions</b>	<b>18</b>
<b>6</b>	<b>Appendix</b>	<b>20</b>
6.1	Variable and Parameter Definitions and Units . . . . .	20
6.2	The MATLAB Bioreactor ODE Code . . . . .	21



# 1 Introduction

Our goal in the dry lab this year was to create four different models that allow our wet lab team to characterize their results, and allow future researchers to benchmark their results creating standard measures in the field of cellular flotation. First, we created a generic differential bioreactor model that allowed our team to predict the effectiveness of our *E. coli* cells to clean waste-waters if coupled with any surface binding method. We performed a complete sensitivity analysis on this model to allow future researchers to reuse this model with completely different parameters, strains of bacteria and object of waste. Then we created an algorithm that can track cellular flotation from frame to frame, and characterize exactly how the cells float; previously, we could only tell whether they floated or not. This coupled with our ODE buoyancy model allows us to define a maximum carrying capacity for each strain. Both of these models allowed our team to benchmark their results and will allow future researchers to quantify the performance of flotation as well.

**Note:** In the very end of this paper, we included a nomenclature defining all the variables used.

## 2 A Differential Bioreactor Model

### 2.1 Goals

- Explore a possible application for our genetically engineered *E. coli* biomass that utilizes flotation.
- Develop a generic bioreactor that can be reused in many different conditions and for a variety of purposes.

### 2.2 Mathematical Formulation

#### 2.2.1 Concepts

##### Monod Equation [7]

In order to run a bioreactor, we need to understand at what rate the *E. coli* grow, and how they respond to environmental conditions. Bacterial growth can be separated into four phases:

1. The Lag Phase: Little growth is observed.
2. The Exponential Phase: Exponential growth is observed after the cells get used to their environment, and so long as the limiting substrate is still in surplus.
3. The Stationary Phase: No growth is observed, death rate and growth rate are approximately equal.
4. Death Phase: There is no more substrate, thus the population dies.

If  $N$  is the concentration of biomass population, then the exponential growth phase is given by:

$$\dot{N} = \mu N$$

where  $\mu$  is the specific growth rate, the "rate of increase of cell concentrations per unit cell concentration" ( $h^{-1}$ ) [7].  $\mu$  depends on the substrate concentration; the relation is given by the Monod equation:

$$\mu(S) = \mu_m \frac{S}{K_S + S}$$

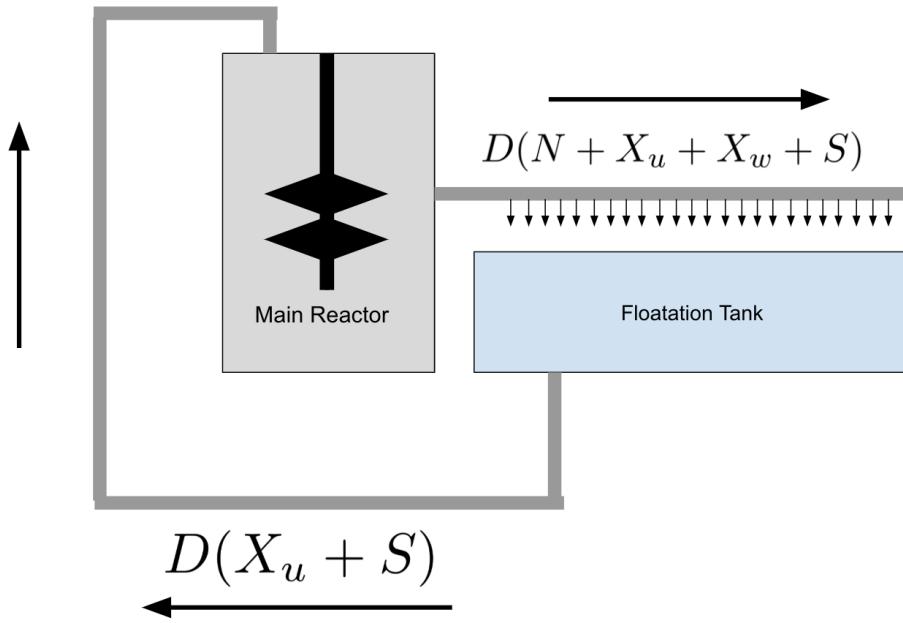
where  $S$  is the substrate concentration (in mol L<sup>-1</sup>) , and  $K_S$  is the Monod Constant (in mol L<sup>-1</sup>). As substrate is consumed, more biomass is being created at the rate of:

$$\frac{dN}{dS} = -\gamma \frac{dS}{dt}$$

where  $t$  is time.  $\gamma$  should be interpreted as “the ratio of the mass of cells formed to the mass of substrate consumed” [7].

### 2.2.2 The Model

Figure 1: The Bioreactor Design



Call the 'Main Reactor' (M.R) and the 'Flootation Tank' (F.T). Let  $X_u$  be the concentration of E. coli (in mol L<sup>-1</sup>),  $X_w$  the concentration of metal bound to the surface of the E. coli (in mol L<sup>-1</sup>). Below we will outline the steps of the bioreactor seen in Figure 1.

1. All the substrate, waste and cellular population is homogeneously distributed between the F.T and the M.R. Initially there is no metal bound to any of the cells, i.e  $X_w = 0$ .
2. When activated, the bioreactor begins to pump solution from the M.R to the F.T at rate  $D$ , this is represented by  $D(N + X_u + X_w + S)$ .
3. In the F.T, the cells are floating on the surface, and are scraped off and removed. Thus only  $X_u$  and  $S$  remain within the solution, which is pumped back into M.R at rate  $D$ . In practice, a cellular filter alike the one used for skimming could be placed at the exit of the F.T to capture any dead cells.
4. Within M.R the unbound metal binds to the E. coli with some rate constant  $\alpha$ , and detaches with some rate  $\beta$  (equilibrium process).
5. This process operates until the metal concentration is as low as desired.

### 2.2.3 Bioreactor Assumptions

- Our cell surface engineering technique has specificity to only one substance in the waste-water (i.e. the substance desired for removal).
- We will only observe exponential bacterial growth within the reactor, i.e. the biomass has been cultured to sufficiently large optical density before being introduced into the bioreactor.
- Suspended metal particles attach to the surface of our E. coli at a rate proportional to the metal concentration and the ratio of unused area available on the surface.
- Bound metal particles detach from the surface of our E. coli back into the waste-water at a rate proportional to the metal concentration bound to the E. coli.
- Once the E. coli that are pumped out of the main chamber end up in the flotation tank, they are immediately removed – no more growth nor unbinding of metal occurs.
- We ignore all the spatial properties of pipes between the main reactor and flotation tank, as soon as part of the solution leaves the main tank, it immediately arrives in the flotation tank, and vice versa.
- The pipe leading into the flotation tank sprinkles the solution over the flotation tank so lightly that the E. coli cells do not sink into the tank – they float on the surface, allowing for immediate removal.
- The solution within the main reactor is continuously mixed; it can be considered a homogeneous solution).
- All the E. coli in the flotation tank float.
- A higher dilution rate ( $D$ ) does not affect the detachment rate.

## 2.3 Equations

The dynamics within the bioreactor are given by:

$$\dot{N} = N_i\mu(S) - N \cdot (D + \kappa) \quad (1)$$

$$\dot{S} = -\mu(S)N\gamma^{-1} \quad (2)$$

$$\dot{X}_u = \beta X_w - \alpha X_u \cdot \left( \frac{\delta N - X_w}{\delta N} \right) \quad (3)$$

$$\dot{X}_w = -\beta X_w + \alpha X_u \cdot \left( \frac{\delta N - X_w}{\delta N} \right) - DX_w \quad (4)$$

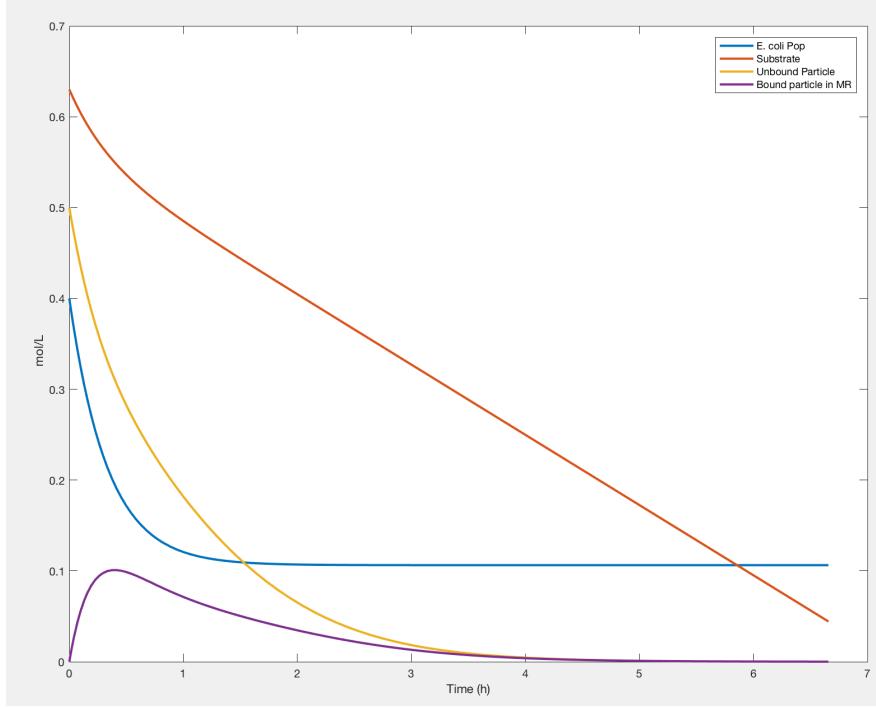
Where,  $\kappa$  is the death rate,  $\delta$  the maximum carrying capacity of the cells,  $\alpha$  the binding rate of metal to the cells and  $\beta$  the detachment rate of metal from the cells.

Equation (1) represents the change in E. coli cell concentration, equation (2) represents the change in limiting substrate concentration, equation (3) represents the change in unbound metal concentration, and equation (4) represents the change in concentration of metal bound to the surface of the E. coli.

In equation (3) we have  $\beta X_w$  [9] which represents metal unbinding from the E. coli population, it depends on  $\beta$ : the kinetic detachment rate, and  $X_w$ : the concentration of metal that is bound to the surface of the E. coli. We are also removing metal from the solution by  $\alpha X_u \cdot (\frac{\delta N - X_w}{\delta N})$ . This depends on the kinetic binding rate  $\alpha$ , the concentration of unbound metal within the solution ( $X_u$ ), and the proportion of available binding sites on the E. coli, this is represented by  $(\frac{\delta N - X_w}{\delta N})$ . Equation (4) is the negative of equation (3) but we also remove  $X_w$  at some dilution rate  $D$ .

## 2.4 Numerical Solution

Figure 2: Bioreactor Numerical Simulation



We solved the ODE system composed of equations (1) - (4) using the ODE45 solver in MATLAB (see the end of the paper of the MATLAB code), using the initial conditions  $N_i = 0.4$ ,  $S = 0.63$ ,  $X_u = 0.5$ ,  $X_w = 0$  and parameters:

- $\gamma = 1.1$
- $\kappa = 0.01$
- $D = 3$
- $\beta = 0.03$
- $\alpha = 1.5$
- $\mu_m = 0.8$
- $K_s = 2.87 \times 10^{-7}$
- $\delta = 1.5$

The above simulation was done for removing Cobalt from mining waste-water effluent. It utilizes a metal binding mechanism outlined in [2] which gave us a range of viable  $\alpha$ . We got  $K_s$  and  $\mu_m$  from [8],  $\beta$ , and  $\kappa$  from [9]. We estimated reasonable values for  $D$ ,  $\gamma$ . See §3.5 for calculating the value of  $\delta$ .

These results are incredibly promising. They show that the bioreactor can theoretically be useful at solving real problems, and that it operates within a reasonable amount of time (6.5 hours).

## 2.5 Sensitivity Analysis

In order to understand the dynamics within the bioreactor, we performed a sensitivity analysis that estimates the relative effect of a single parameter on the performance of the system. Unfortunately, it is not possible to define an explicit function relating the different parameters to each other, we

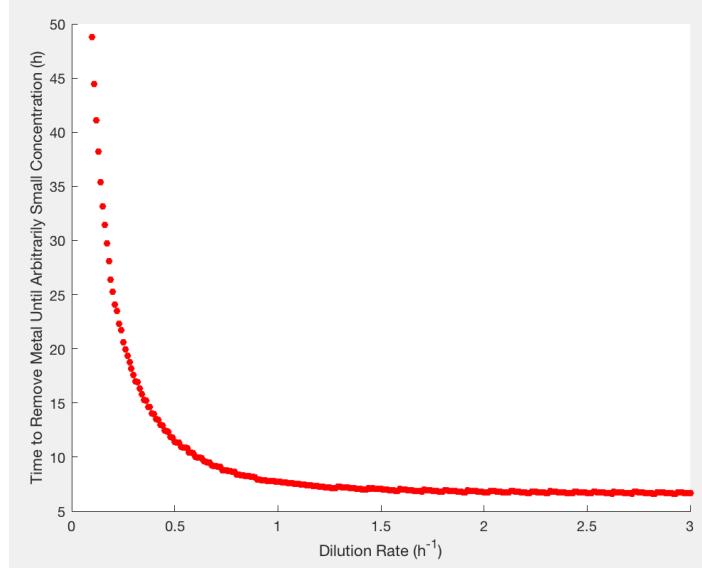
must perform a 'naive' analysis; we vary one parameter while holding all other constant, and see how this change affects the performance of the system. Performance is measured by how long it takes the bioreactor to remove the large majority of metal. We set some small  $\varepsilon = 10^{-4} \text{ mol L}^{-1}$  to define a low threshold of acceptable metal concentration. For §2.5.1 - §2.5.6, we kept all the parameters at:

- $\gamma = 1.1$
- $\alpha = 1.5$
- $\kappa = 0.01$
- $\mu_m = 0.8$
- $D = 3$
- $K_s = 2.87 \times 10^{-7}$
- $\beta = 0.03$
- $\delta = 1.5$

and only varied the parameter whose effect was measured.

### 2.5.1 The Effect of Dilution Rate on Bioreactor Performance

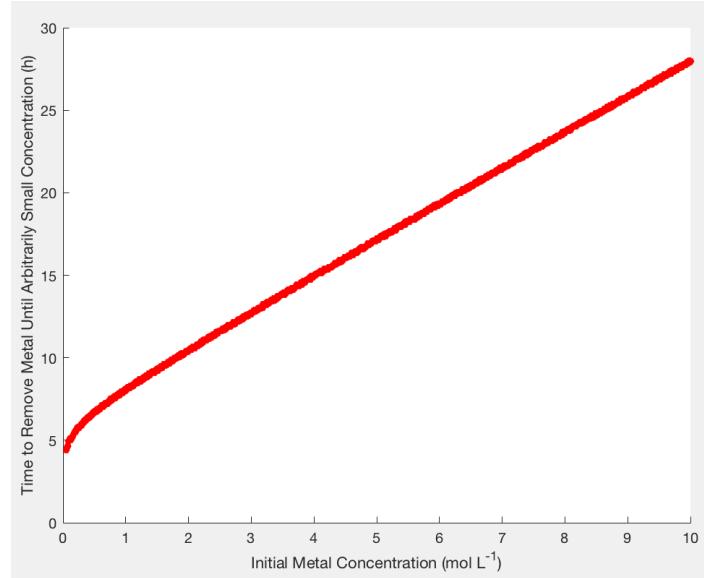
Figure 3: Dilution Rate ( $D$ ) vs. Bioreactor Performance



(See Figure 3) The higher the dilution rate, the better the performance. Above dilution rate  $D = 1$  it slowly approaches an asymptote. We should not have a  $D$  that is too large because it is energy costly. Also, a larger  $D$  leads to faster deterioration of the bioreactor.

### 2.5.2 The Effect of Initial Metal Concentration on Bioreactor Performance

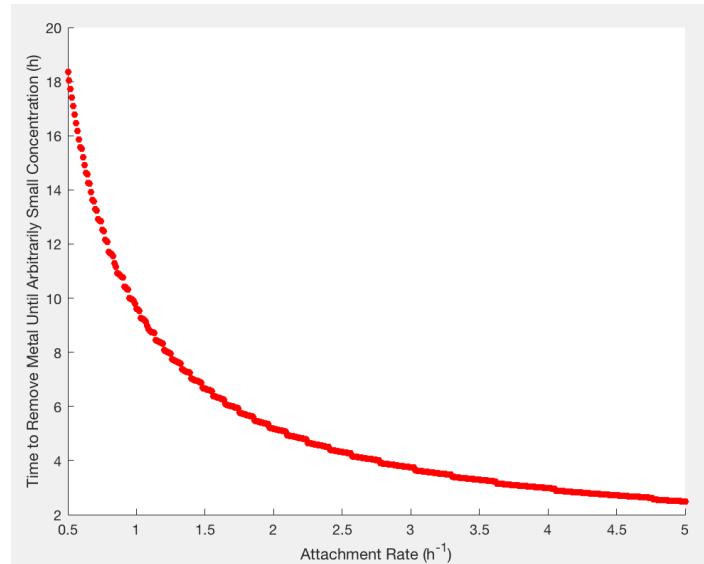
Figure 4: Initial Metal Concentration ( $X_u$ ) vs. Bioreactor Performance



(See Figure 4) The higher the initial metal concentration within the bioreactor, the longer it takes to operate.

### 2.5.3 The Effect of Attachment Rate on Bioreactor Performance

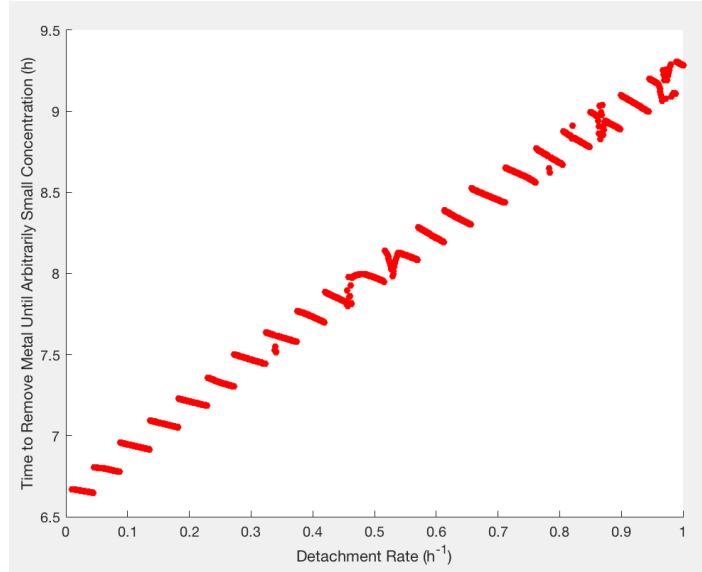
Figure 5: Attachment Rate ( $\alpha$ ) vs. Bioreactor Performance



(See Figure 5) The higher the attachment rate ( $\alpha$ ), the better the performance of the system. It does approach an asymptote quite quickly though. This is because of the maximum carrying capacity, at some point, there is just no more available space on the E. coli surface to bind more metal, regardless of the attachment rate.

#### 2.5.4 The Effect of Detachment Rate on Bioreactor Performance

Figure 6: Detachment Rate ( $\beta$ ) vs. Bioreactor Performance

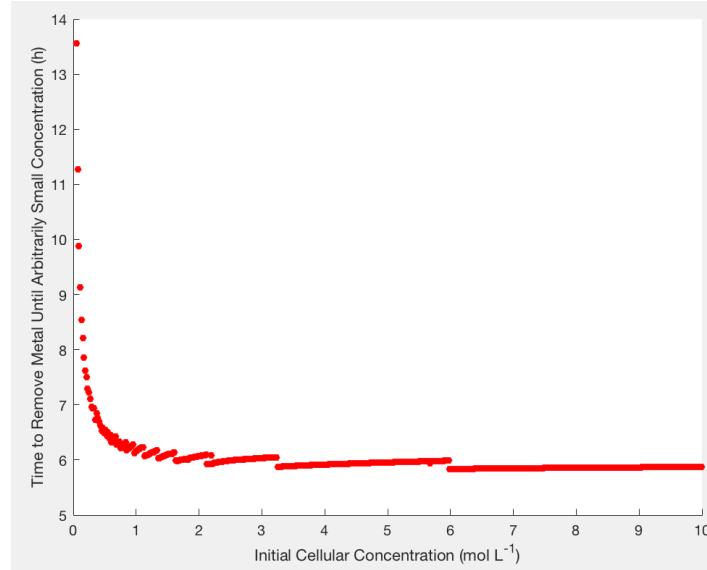


(See Figure 6) The lower the detachment rate ( $\beta$ ), the better the performance of the system.

#### 2.5.5 The Effect of Initial Cellular Population on Bioreactor Performance

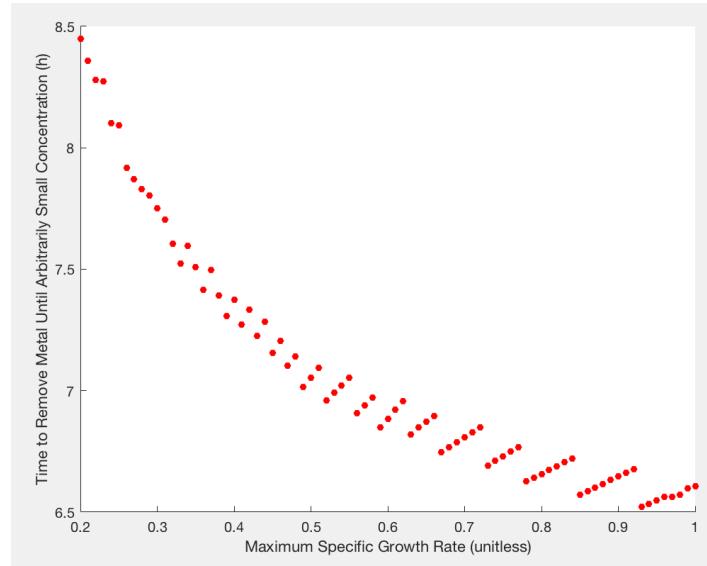
(See Figure 7) The higher the initial E. coli population ( $N_i$ ), the better the system performs. After some  $N_i \approx 1$ , it approaches an asymptote.

Figure 7: Initial Cellular Population ( $N_i$ ) vs. Bioreactor Performance



#### 2.5.6 The Effect of Maximum Specific Growth Rate on Bioreactor Performance

Figure 8: Maximum Specific Growth Rate ( $\mu_m$ ) vs. Bioreactor Performance



(See Figure 8) The higher the maximum specific growth rate, the better the performance of the system.

## 2.6 Hypothetical Applications

The numerical simulations above were all based on parameters values specific to cobalt removal from mining waste-water. If the E. coli are coupled with a binding mechanism that is not specific to cobalt,

then this generic bioreactor can be used for a large variety of applications (one simply has to change  $\alpha$  and  $\beta$ ). One very promising application is the removal of pharmaceuticals, such as penicillin, from municipal waste-water. The reactor is viable for many different binding methods, strains of E. coli, and initial concentrations of waste, as displayed in §2.5. This reactor is independent of volume, all the parameters are relative to each other, this system would perform the same with 1 L of effluent as with  $1 \times 10^{10}$  L of effluent. Evidently, the size of the chambers would need to be scaled appropriately.

### 3 A Mechanistic Buoyancy Model: Maximum Carrying Capacity Determination

#### 3.1 Goal

The buoyancy model has 2 main goals:

1. Determine the mean buoyant force experienced by the genetically engineered E. coli biomass as a result of gas vesicle formation from ARG1 over-expression. This effectively determines a mechanical upper bound for carrying capacity of the biomass.
2. Characterize and quantify flotation observations from wet lab experimentation. It is not well known what the role of some of the secondary GVPs in the ARG1 construct is for gas vesicle formation, and being able to quantify flotation facilities comparison of different combinations of secondary GVPs for optimization of a gene construct specifically engineered for biomass flotation.

#### 3.2 Mathematical Formulation

##### 3.2.1 Concepts

- Newtonian mechanics:  $\vec{F} = m\vec{a} \equiv m\frac{d\vec{v}}{t}$ 
  - $\vec{F}$   $\equiv$  vector sum of forces on body
  - $m$   $\equiv$  mass of body
  - $\vec{a}$   $\equiv$  acceleration as a function of time
  - $\vec{v}$   $\equiv$  velocity as a function of time
- Stokes-Einstein Drag:  $\vec{F}_D = -6\pi\eta R\vec{v}$ 
  - $\vec{F}_D$   $\equiv$  drag force (note it opposes direction of velocity)
  - $\eta$   $\equiv$  viscosity of fluid medium
  - $R$   $\equiv$  (approximate) radius of body in spherical approximation.
- Integrating factor method to solve first order linear ODE:

$$\frac{dy}{dt} + f(t)y = g(t) \quad (5)$$

Let  $\mu(t) = e^{\beta t}$ . Multiply both sides by  $\mu(t)$ .

Chain rule  $\implies \frac{d}{dt}(y(t)\mu(t)) = h(t)\mu(t)$

Integrate, divide by  $\mu(t)$   $\implies y(t)$ .

### 3.2.2 Assumptions

- The biomass separates into clumps that can be approximated by spheres.
- Buoyant force is constant.
- Cell motility and in-plane motion (motion perpendicular to the vertical axis) is negligible. This effectively reduces our system to one dimension.

### 3.2.3 The Model

Stokes-Einstein, gravitational force near the surface of the earth  $F_g = mg$  ( $g \approx 9.81 \text{ ms}^{-2}$ ), Newton's second law lets us write:

$$F_{net} = F_g + F_B + F_D \implies m \frac{dv}{dt} = -mg + F_B - 6\pi\eta Rv \quad (6)$$

Rearrange...

$$\frac{dv}{dt} - \frac{6\pi\eta R}{m}v = \left(\frac{F_B}{m} - g\right) \quad (7)$$

Simplify notation with  $\alpha := \frac{6\pi\eta R}{m}$  and  $\beta := \left(\frac{F_B}{m} - g\right)$  we have:

$$\frac{dv}{dt} - \alpha v = \beta \quad (8)$$

Observe the model is now in the form of equation 5.

## 3.3 Analytic Solution

Now to get a solution we can directly apply the integrating factor method as described in the Concepts sections.

The integrating factor is:

$$\mu(t) = e^{\alpha t} \quad (9)$$

Multiplying both sides we have:

$$\frac{d}{dt}(ve^{\alpha t}) = \beta e^{\alpha t} \quad (10)$$

Now we integrate, divide by integrating factor, and impose an initial condition of  $v(t = 0) = 0 \text{ ms}^{-1}$  (not moving initially):

$$v(t) = \frac{\beta}{\alpha} \left(1 - e^{-\alpha t}\right) \equiv \left(\frac{F_B - mg}{6\pi\eta R}\right) \left(1 - e^{-\frac{6\pi\eta R}{m}t}\right) \quad (11)$$

This is a closed form time-dependent solution for velocity.

Now we integrate one more time and impose an initial condition of  $z(t = 0) = 0 \text{ m}$ :

$$z(t) = \left(\frac{F_B - mg}{g\pi\eta R}\right) \left(t + \frac{m}{6\pi\eta R} \left(e^{\frac{6\pi\eta R}{m}t} - 1\right)\right) \quad (12)$$

Finally, we have a closed form time-dependent solution for vertical displacement.

**NB:**  $m$ ,  $g$ ,  $\eta$ ,  $R$  are considered to be known, empirical constants.  $F_B$  is left as a parameter that can be determined in a curve fit regression to a time-series of vertical displacement data for a floating biomass.

### 3.4 Determining Buoyant Force: Temporal Tracking Algorithm

Now that we have an expression for vertical displacement as a function of time which depends on the buoyant force as a parameter, we want to determine the magnitude of this buoyant force. In principle, this is not a difficult task as it can be solved using the built-in least-squares optimization curve-fit functions, in say, MATLAB. The difficulty lies in acquiring experimental data to fit to. Our proposed solution to this is to acquire images from a stationary point of view at evenly spaced, small time intervals and perform image segmentation and analysis techniques to track the vertical position of floating biomass frame-to-frame. In reality, the biomass will be clumped into many clusters, so clustering and labeling algorithms are employed. The advantage of using a visual tracking algorithm to observe vertical displacement is that researchers can exactly characterize flotation, benchmark their results, and calculate how different variables affect flotation.

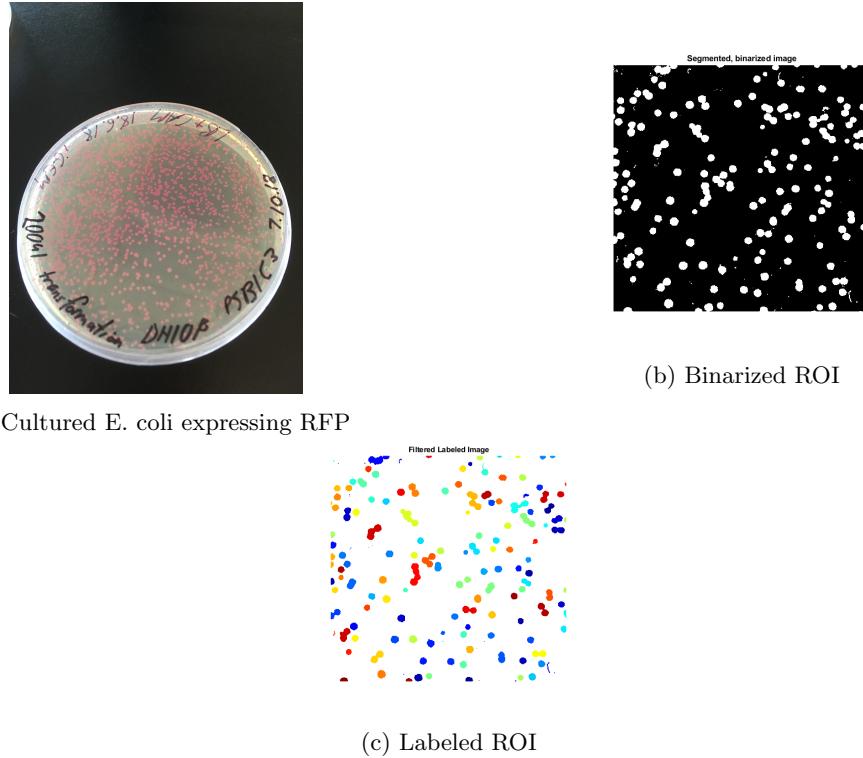


Figure 9: Example of image segmentation and clustering

Below is pseudo-code for a stochastic temporal tracking algorithm that is intended to maintain consistent cluster labeling from frame-to-frame and account for clusters combining and splitting. The time interval between frames should be chosen to be sufficiently small such that the probability of two binding/un-binding events occurring between any two frames can be assumed to be zero.

- Load directory with images.
- For  $i$  up to number of frames (i.e. time steps):
  - Read pixel data from image file.
  - Manually set crop margins to ROI encapsulating the region of flotation. (apply same cropping margins in subsequent frames programatically).
  - Convert from RGB to grayscale, perform thresholding using Otsu's method.

- Binarize image based off of threshold, segmentation complete.
- Cluster and label binary image.
- Centroid clusters. Store labels  $j = 1, 2, 3, \dots, N$ , positions  $(x_i^{(j)}, y_i^{(j)})$ , and approximate radius for each of the  $N$  objects identified in segmentation. Note that the radius is updated at each step to account for changes in cluster morphology affecting magnitude of drag force at each time step.
- if  $i > 1$ :
  - \* For all  $N$  labels  $j$  in frame  $i$  find position  $(x_{i-1}^{(j')}, y_{i-1}^{(j')})$  with  $y_i \geq y_{i-1}$  s.t. the distance  $d((x_{i-1}^{(j')}, y_{i-1}^{(j')}), (x_i^{(j)}, y_i^{(j)}))$  is minimized.
  - \* The label  $j$  in the  $i^{\text{th}}$  frame is considered to be a child of the label  $j'$  in the  $(i+1)^{\text{th}}$  frame.
  - \* End if-statement.
- End for-loop.
- Define function handle for solution to ODE model for buoyancy.
- Convert cluster tracking data from pixels to spacial units (based on pixel size of image), curve-fit to each set of cluster branches.
- Take average of curve-parameter determination of  $F_B$  for each branch, analyze distribution, significance of fit, variance of mean.

But wait! We passively assumed that the number of clusters  $N'$  in frame  $i - 1$  was greater than the number of clusters  $N$  in frame  $i$ . Not to worry, small modification...

1.  $N = N'$ :
  - If this is the case, the logic holds and we simply have a child label  $j'$  that is mapped to parent label  $j$ .
2.  $N > N'$ :
  - Physically, this corresponds to a cluster splitting.
  - How do we handle this? No modification of the pseudo-code algorithm is needed. There will simply be two objects with labels  $j_1, j_2$ , respectively, to which some  $j'$  is mapped to. A *branch* in the tracking tree.
3.  $N < N'$ 
  - Non-trivial case.
  - Switch search: For all  $N'$  labels  $j'$  in frame  $i + 1$  find position  $(x_i^{(j)}, y_i^{(j)})$  with  $y_i \geq y_{i-1}$  s.t. the distance  $d((x_{i-1}^{(j')}, y_{i-1}^{(j')}), (x_i^{(j)}, y_i^{(j)}))$  is minimized.
  - There will exist a label  $j$  in the  $i^{\text{th}}$  frame such that two labels  $j'_1, j'_2$  in the  $(i+1)^{\text{th}}$  frame map to it. In other words, two clusters combine.

Note: The position tracking effectively forms a directed acyclic graph (DAG)

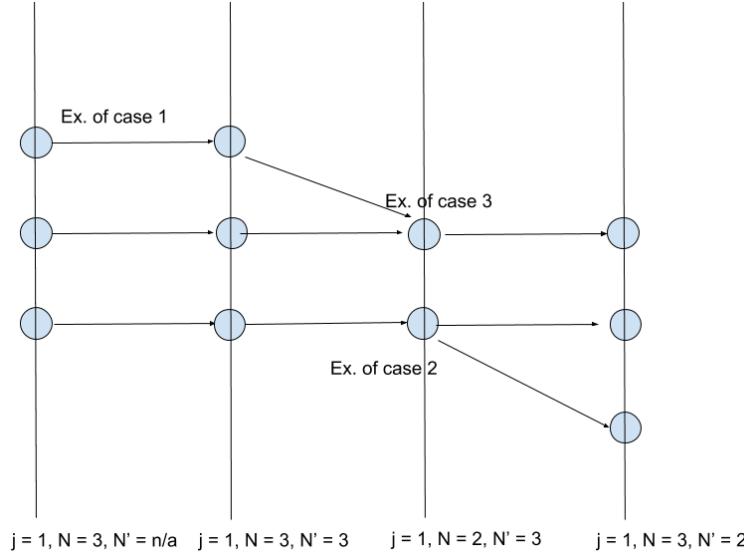


Figure 10: A few potential layers of a simple example with 3 initial clusters (nodes) to illustrate the 3 cases.

### 3.5 Discussion On Maximum Carrying Capacity

For this discussion, consider the buoyant force  $F_B$  to be a known, constant value.

In order for the biomass to float the net force on it must be greater than zero. By equation 6 we have the following constraint on buoyant force:

$$F_B \geq mg + 6\pi\eta Rv(t), \forall t \quad (13)$$

The right side of the above inequality has strictly positive monotonicity with respect to  $v$ .  $v$  is a strictly increasing function so its max on a closed, bounded interval is at the upper bound. Let  $v_{\max} = v(t_f)$ , where  $t_f$  is the solution to  $z(t_f) = h$ ,  $h$  being the height required from start point to top of container,  $z$  being as given in equation 12. The inequality in equation 13 is thus equivalent to the following:

$$F_B \geq m_{E. coli}g + 6\pi\eta R_{E. coli}v_{\max} \quad (14)$$

Taking  $m$  to be the approximate mass of an E. coli cell (without bound particle),  $R_{E. coli}$  to be the approximate radius of an E. coli cell in a spherical approximation, this constraint on buoyant force influences goals in the wet lab by setting a minimum requirement for observed buoyant force in modifying ARG1 for optimized flotation performance.

Once ARG1 is optimized for flotation and an estimate for  $F_B$  is obtained, the maximal carrying capacity can be obtained as follows:

$$m \leq \frac{F_B - 6\pi\eta R_{E. coli}v_{\max}}{g} \quad (15)$$

Here  $m = m_{E. coli} + m_{\text{bound}}$  (sum of cell mass with mass of bound particle). Solving for  $m_{\text{bound}}$ :

$$m_{\text{bound}} \leq \frac{F_B - 6\pi\eta R_{E. coli}v_{\max}}{g} - m_{E. coli} \quad (16)$$

Diving by the molecular or atomic mass, denote this  $m_{particle}$  of the substance to be sequestered we have the following value for maximal carrying capacity  $\delta$  in the bioreactor model:

$$\delta = \frac{1}{m_{particle}} \left[ \frac{F_B - 6\pi\eta R_{E. coli} v_{max}}{g} - m_{E.coli} \right] \quad (17)$$

However, in the bioreactor model, the effluent is pumped from the main reactor in such a way that cells will not sink, so  $v_{max}$  can be set to 0 and we see:

$$\delta \approx \frac{1}{m_{particle}} \left[ \frac{F_B}{g} - m_{E.coli} \right] \quad (18)$$

See §2.3. This could influence the choice of expression levels in future cell surface engineering techniques. Too many surface receptors could result in accumulation of too much particle mass, eliminating ability for flotation.

## 4 Growth Dynamics

### 4.1 Objective

Our objective in this paper is to describe the growth dynamics of BL21 E. coli strain in lysogeny broth (LB). We have obtained two sets of observations from our wet lab team:

1. 25 observations with  $\Delta t = 30\text{mins}$
2. 40 observations with  $\Delta t = 30\text{mins}$

### 4.2 Approach

To model all the observations, we will use the Gompertz function[10], which can be expressed in the following way:

$$f(t) = ae^{-be^{-ct}}$$

with:

$$b, c > 0$$

The reason for using this model is that it best describes the lag phase, the exponential phase and stationary phase.

### 4.3 Results

The observations were given to us in  $OD_{600}$ . We know that for bacterial cell cultures  $OD_{600}$  of  $1.0 = 8.0 \times 10^8 \text{cells/ml}$  [6].

- For the first set of observations we fit the curve to our data.

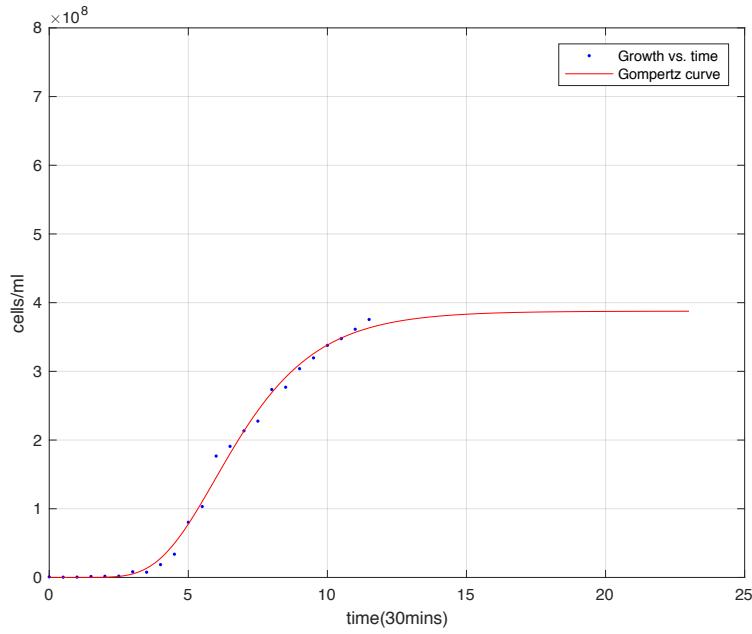


Figure 11: Fitted data

with the following parameters:

General model :

$ans(x) = a * \exp(-b * \exp(-c * x))$

Coefficients (with 95% confidence bounds) :

$a = 3.876 \text{e}+08 \quad (2.782 \text{e}+08, 4.97 \text{e}+08)$

$b = 18.97 \quad (-22.7, 60.64)$

$c = 0.4938 \quad (0.1035, 0.884)$

- For the second set of observations, we fit the same model. However, we obtain different parameters, since we enter the stationary phase after the exponential phase, which we did not observe in the first data set. But we also notice that they fall into the 95% confidence interval for the parameters of the first observations. Hence we can assume, with some certainty, that the parameters we got for this curve fitting process would also fit for the previous one.

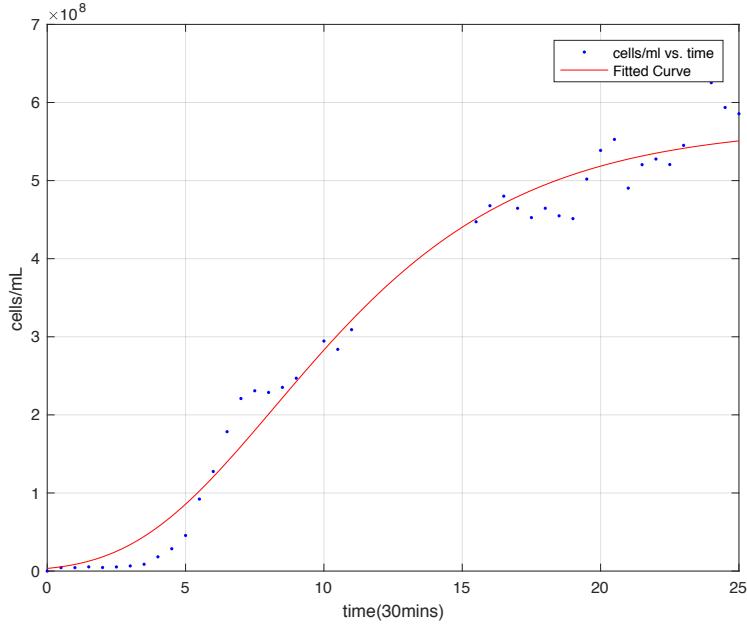


Figure 12: Fitted data

The parameters are as follows:

General model:

```
ans(x) = a*exp(-b*exp(-c*x))
Coefficients (with 95% confidence bounds):
  a = 5.707e+08  (5.355e+08, 6.059e+08)
  b = 5.127      (3.785, 6.469)
  c = 0.1989     (0.1603, 0.2376)
```

## 5 Discussion and Future Directions

Based on the results of our differential bioreactor model we postulate that a bioreactor of this design could perform at appropriately small time-scales with a sufficiently optimized flotation construct in a bioremediation context. This bodes well for future laboratory endeavours where the bioreactor schema along with an engineered cell-line optimized for flotation from gas vesicle formation could be tested in a small scale laboratory model of the system to test its empirical performance. This is useful for model validation, and proof of concept.

An improvement to the bioreactor could be seen in the form of incorporation of an experimental growth dynamics model as demonstrated in the Growth Dynamics section with the Gompertz equation. The Monod equation was chosen initially for this model because it is scalable and its parameters are known for BL21; however, it only describes the exponential phase of growth and it is not guaranteed that the biomass will remain in exponential phase during the bioreactor process. Another shortcoming is that it is a single-substrate model and in practice, this would not be the case as the likely nutrient source would be from carbohydrates, lipids, and organics present in the waste-water

effluent. For future application-based analysis using the bioreactor model, Goempertz coefficients could be determined for a biomass of industrially relevant size for the volume demands dictated by the industry requiring bioremediation using a similar experimentation and analysis technique as described in the Growth Dynamics section.

To further enrich the analysis of the behaviour of the bioreactor model, in addition to the existing sensitivity analysis, identification and characterization of stable points resulting from different combinations of parameter values could be performed.

It would also be informative to apply the stochastic temporal tracking algorithm to real flotation data (images), once obtained, to estimate the buoyant force for ARG1 and compare different modifications of ARG1 to determine an optimal gene combination for flotation.

A next major step is coupling cellular flotation assays with cell surface engineering for application-based testing for different bioremediation tasks. Much research has been done relating to selecting peptide sequences for metal binding [4][5]. Existing techniques could be coupled with a signal transduction pathway to induce expression of ARG1, or modified versions of it optimized for flotation.

## 6 Appendix

### 6.1 Variable and Parameter Definitions and Units

#### Nomenclature

$\alpha$  Attachment Rate ( $\text{h}^{-1}$ )

$\beta$  Detachment Rate ( $\text{h}^{-1}$ )

$\delta$  Maximum Carrying Capacity Constant ( $\text{h}^{-1}$ )

$\eta$  Dynamic viscosity of waste-water ( $\text{kg ms}^{-1}$ )

$\gamma$  Yield Constant (unitless)

$\kappa$  E. coli Death Rate ( $\text{h}^{-1}$ )

$\mu_m$  Maximum Specific Growth Rate ( $\text{h}^{-1}$ )

$D = \frac{F}{V}$  Dilution Rate ( $\text{h}^{-1}$ )

$F$  Flow Rate ( $\text{Lh}^{-1}$ )

$K_s$  Monod Constant ( $\text{mol L}^{-1}$ )

$m_{\text{E. coli}}$  Aproximate mass of E. coli cell

$m_{\text{particle}}$  Atomic or molecular mass of substance to be sequestered ( $\text{g mol}^{-1}$ )

$N_i$  Initial Cellular Concentration ( $\text{mol L}^{-1}$ )

$R_{\text{E. coli}}$  Approximate radius of E. coli cell in spherical approximation (m)

$S_i$  Initial Limiting Substrate Concentration ( $\text{mol L}^{-1}$ )

$V$  Total Volume (L)

$v_{\max}$  Maximum achievable velocity based on vertical displacement specification.

$X_{u_i}$  Initial Metal Concentration within Waste-water ( $\text{mol L}^{-1}$ )

$X_w$  Concentration of Bound Metal ( $\text{mol L}^{-1}$ )

## 6.2 The MATLAB Bioreactor ODE Code

```

gamma = 1.1; % yield constant (?)
kappa = 0.01; % death rate h^-1
D = 3; % dilution/flow rate h^-1
beta = 0.03; % detachment rate of particle from biomass surface h^-1
alpha = 1.5; % attachment rate h^-1
mu_m = 0.8; % maximum specific growth rate (monod equation) unitless
K_s = 2.87*10^(-7); % monod constont mol L^-1
delta = 1.5; % maximum carrying capacity
N_i = 0.4; % initial cellular concentration L^-1
S_i = 0.63; % initial substrate concentration L^-1
Xu_i = 0.5; % initial unbound metal concentration L^-1
Xw_i = 0; % initial bound metal concentration L^-1
epsilon = 0.0001; % Acceptable metal concentration threshold

f = @(t,x) [ N_i * ((mu_m * x(2)) / (K_s + x(2))) - x(1) * (D + kappa);
              -((mu_m * x(2)) / (K_s + x(2))) * x(1) / gamma;
              beta * x(4) - alpha * x(3) * (delta * x(1) - x(4))/(delta * x(1));
              - (beta * x(4) - alpha * x(3) * (delta * x(1) - x(4))/(delta * x(1))) - D*x(4) ];

[t,xa] = ode45(f,[0 50],[N_i S_i Xu_i Xw_i]);
fprintf('ode solved\n')

ind = find(xa(:,3) <= epsilon);
xap = xa(1:ind,:);
tp = t(1:ind);

figure;
plot(tp,xap(:,1)); % Cell concetration mol L^-1
hold on
plot(tp,xap(:,2)); % Substrate concentration
hold on
plot(tp,xap(:,3)) % Unbound metal concentration mol L^-1
hold on
plot(tp,xap(:,4)) % Bound metal concentration mol L^-1
hold on
set(findall(gca, 'Type', 'Line'), 'LineWidth',2);

legend('E.coli.Pop', 'Substrate', 'Unbound_Particle', 'Bound_particle_in_MR')
xlabel('Time_(h)'), ylabel('mol/L')

```

## 7 References

- [1] Oladeji, S. O., & Saeed, M. D. (2015). *Assessment of cobalt levels in wastewater, soil and vegetable samples grown along Kubanni stream channels in Zaria, Kaduna State, Nigeria*. African Journal of Environmental Science and Technology, 9(10), 765-772. doi:10.5897/ajest2015.1969
- [2] Whittaker MM, Mizuno K, Bächinger HP, Whittaker JW. *Kinetic Analysis of the Metal Binding Mechanism of Escherichia coli Manganese Superoxide Dismutase*. Biophysical Journal. 2006;90(2):598-607. doi:10.1529/biophysj.105.071308.
- [3] *Monod substrate affinity constant (K<sub>s</sub>) for strain ML30 of E. coli growing on glucose as the only source of carbon and energy*. (n.d.). Retrieved October 13, 2018 from <http://bionumbers.hms.harvard.edu/bionumber.aspx?s=n&v=1&id=111051>.
- [4] N., Cetinel, S., Omar, S. I., Tuszyński, J. A., & Montemagno, C. (2017). *A computational method for selecting short peptide sequences for inorganic material binding*. Proteins: Structure, Function, and Bioinformatics, 85(11), 2024-2035. doi:10.1002/prot.25356
- [5] Li, Pengsong & Tao, Huchun. (2013). *Cell surface engineering of microorganisms towards adsorption of heavy metals*. Critical reviews in microbiology. 41. 10.3109/1040841X.2013.813898.
- [6] *Bacterial cell number (OD600)* (n.d.). Retrieved October 13, 2018 from <http://www.labtools.us/bacterial-cell-number-od600/>
- [7] Bengt Carlsson. (2009). *An introduction to modeling of bioreactors*. Dept of Systems and Control, Information Technology Uppsala University. Retrieved October 13, 2018 from <https://www.it.uu.se/edu/course/homepage/modynsyst/vt11/Lecture/DynSystBior2009.pdf>.
- [8] *Monod substrate affinity constant (K<sub>s</sub>) for strain ML30 of E. coli growing on glucose as the only source of carbon and energy*. (n.d.). Retrieved October 13, 2018
- [9] Chen, L., & Chai, L. (2005). *Mathematical Model and Mechanisms for Biofilm Wastewater Treatment Systems*. World Journal of Microbiology and Biotechnology, 21(8-9), 1455-1460. doi:10.1007/s11274-005-6565-2
- [10] Parolini, N, & Carcano, S. (2009). *A MODEL FOR CELL GROWTH IN BATCH BIOREACTORS*. Faculty of Systems Engineering, Polytechnic University of Milan. Retrieved October 13, 2018 from [https://www.politesi.polimi.it/bitstream/10589/2082/1/2010\\_07\\_Carcano.pdf](https://www.politesi.polimi.it/bitstream/10589/2082/1/2010_07_Carcano.pdf).