



Universidade do Minho
Escola de Engenharia

Ficha de Trabalho I

Students:	Alexandre Miguel Magalhães Esperança	PG45963
	André Filipe Pereira da Silva	PG45462
	António Nuno Carrilho Canatário Duarte	PG45464
	Mónica Rafaela Machado Leiras	PG45473
	Roberto Costa Bullitta	PG45474
	Vânia Miguel Bento	PG45971

Course: Mestrado em Bioinformática (MBINF)

Course Unit: Modulação de Processos Biológicos

Conteúdo

1		2
a)	2
b)	2
2		3
a)	3
b)	4
c)	5
d)	5
e)	6
f)	8
g)	9

GITHUB - https://github.com/rocobull/Modelacao_Grupo_1

1

Nesta primeira parte, pretende-se modelar os diferentes modos de crescimento da bactéria *Zymomonas mobilis*, tendo por base as seguintes reações químicas.

Crescimento oxidativo em glucose: $k_1 G \rightarrow^{\mu_1} X$

Crescimento fermentativo em sacarose: $k_2 S \rightarrow^{\mu_2} X + k_3 E$

Crescimento oxidativo em etanol: $k_4 E \rightarrow^{\mu_3} X$

Cinética da reação 1: $\mu_1 = \mu_{max1} \frac{G}{K_G + G} = 0.4 \frac{G}{0.1 + G}$

Cinética da reação 2: $\mu_2 = \mu_{max2} \frac{S}{K_S + S} = 0.2 \frac{S}{0.3 + S}$

Cinética da reação 3: $\mu_3 = \mu_{max3} \frac{E}{K_E + E} = 0.25 \frac{E}{0.4 + E}$

De acordo com as reações fornecidas, relativamente ao consumo de glucose e sacarose para produção de etanol, determinaram-se os modelos dinâmicos em dois modos de cultivo: primeiro em modo *batch* (descontínuo) e de seguida em modo *fed-batch* (semi-contínuo). Consideram-se ainda 2 correntes de alimentação para o modo *fed-batch*: uma de glucose (com um fluxo de entrada de 150 g/L) e uma de sacarose (com um fluxo de entrada de 100 g/L).

a)

Para o **modelo descontínuo**, vemos que tanto a glucose como a sacarose variam negativamente com o aumento da concentração de biomassa. A variação da biomassa aumenta com maiores concentrações da própria biomassa. Já o etanol varia tanto positivamente como negativamente com a biomassa, ou seja, neste caso o valor dessa variação vai depender da extensão das reações de consumo de sacarose (e consequente produção de etanol) e consumo de etanol. Uma vez que estamos a considerar o modo descontínuo, não existe variação de volume do sistema ($\frac{dV}{dt} = 0$).

$$\frac{d[G]}{dt} = -k_1 \mu_1 [X]$$

$$\frac{d[S]}{dt} = -k_2 \mu_2 [X]$$

$$\frac{d[E]}{dt} = -k_4 \mu_3 [X] + k_3 \mu_2 [X]$$

$$\frac{d[X]}{dt} = \mu_1 [X] + \mu_2 [X] + \mu_3 [X]$$

$$\frac{d[V]}{dt} = 0$$

b)

Para o **modelo semi-contínuo**, uma vez que temos duas correntes de alimentação, a variação do volume do sistema vai depender tanto do fluxo de entrada de glucose (FeG) como o de sacarose (FeS), pelo que temos de representar ambos nas equações.

$$\frac{d[G]}{dt} = -k_1 \mu_1 [X] + \frac{F_{eG}}{V} G_0 - [G] \left(\frac{F_{eG} + F_{eS}}{V} \right)$$

$$\frac{d[S]}{dt} = -k_2 \mu_2 [X] + \frac{F_{eS}}{V} S_0 - [S] \left(\frac{F_{eG} + F_{eS}}{V} \right)$$

$$\frac{d[E]}{dt} = -k_4 \mu_3 [X] + k_3 \mu_2 [X] - [E] \left(\frac{F_{eG} + F_{eS}}{V} \right)$$

$$\frac{d[X]}{dt} = \mu_1 [X] + \mu_2 [X] + \mu_3 [X] - [X] \left(\frac{F_{eG} + F_{eS}}{V} \right)$$

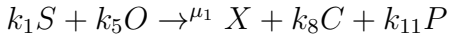
$$\frac{d[V]}{dt} = F_{eG}(t) + F_{eS}(t)$$

Desta forma, passamos a considerar as taxas de diluição em cada equação, uma vez que o volume aumenta devido à entrada de substrato a uma taxa constante. Nas equações de variação de glucose e sacarose, a diluição é mitigada pela concentração inicial de glucose e sacarose, respetivamente.

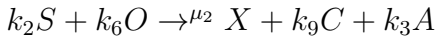
2

Nesta segunda parte, pretende-se modelar os diferentes modos de crescimento da bactéria *Escherichia coli* (estirpe BL21), tendo por base as seguintes reações químicas.

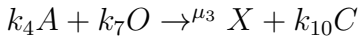
Crescimento oxidativo em glucose com produção de proteína recombinante:



Crescimento fermentativo em glucose:



Crescimento oxidativo em acetato:



Por serem gases, e portanto complicarem a criação dos modelos dinâmicos, tanto o oxigénio como o carbono não foram tidos em consideração nesta análise.

a)

De acordo com as reações fornecidas, relativamente ao crescimento oxidativo em glucose com produção de proteína recombinante, crescimento fermentativo em glucose com produção de acetato, e crescimento oxidativo em acetato, determinou-se o modelo dinâmico em modo *batch* (descontínuo).

Ao longo do relatório, passamos a tratar a glucose como "substrato", visto que é a fonte principal de alimentação da bactéria em estudo.

$$\frac{d[X]}{dt} = \mu_1[X] + \mu_2[X] + \mu_3[X]$$

$$\frac{d[S]}{dt} = -k_1\mu_1[X] - k_2\mu_2[X]$$

$$\frac{d[A]}{dt} = k_3\mu_2[X] - k_4\mu_3[X]$$

$$\frac{d[P]}{dt} = k_{11}\mu_1[X]$$

$$\frac{d[V]}{dt} = 0$$

b)

Em Python, criou-se primeiro o modelo geral em modo descontínuo para a estirpe BL21, tendo em conta as três equações de crescimento: oxidativo em glucose, fermentativo em glucose, e oxidativo em acetato.

Tendo o modelo criado, foi possível simular a variação dos parâmetros X , S , A e P ao longo do tempo, aplicando o método "ode" do *sub-package* *scipy.integrate*, indicando os valores iniciais dos parâmetros a determinar e as constantes do sistema (previamente fornecidos), e definindo um intervalo de tempo de 0 a 3 unidades de tempo (ut) na qual vão sendo determinados os valores das variáveis, de 0.1 em 0.1 ut .

Utilizando o sub-package *matplotlib.pyplot*, foi possível criar gráficos que representam a variação das concentrações das variáveis previamente referidas ao longo do tempo. Na **Figura 1**, encontra-se representado uma versão truncada do gráfico original (de 0 a 1.5 ut) para visualizar melhor os instantes iniciais do mesmo.

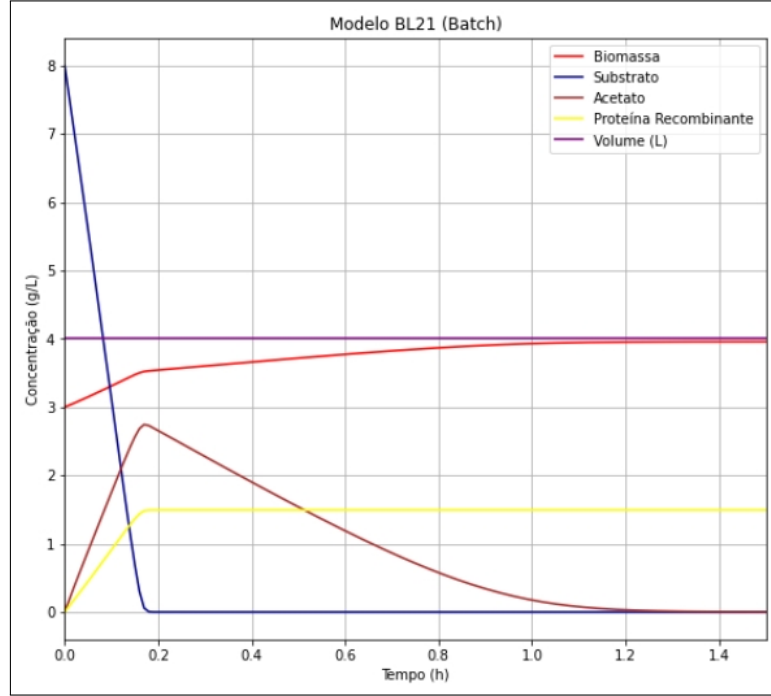


Figura 1: Modelo Batch, sendo representado a variação das concentrações de biomassa, substrato, acetato e proteína recombinante, e do volume ao longo do tempo.

Como podemos observar pelo gráfico (**Figura 1**), sendo que o modo de crescimento é descontínuo (não existindo entrada nem saída de substrato) o volume mantém-se constante ao longo do tempo. A biomassa aumenta até atingir o ponto máximo coincidente com o valor constante do volume do sistema ($conc = 4 \text{ g/L}$), representativo da interrupção do crescimento da biomassa. Este ponto é atingido pela escassez do substrato e do acetato. O acetato é produzido durante o crescimento fermentativo com consumo de substrato e, quando este último é totalmente consumido ($h \approx 0.2$), o acetato torna-se a fonte de alimentação. Isto é denotado pelo aumento da curva do acetato e posterior diminuição a partir do consumo total de substrato. Por último, a proteína recombinante também aumenta de acordo com o consumo de substrato, até este último ser totalmente consumido, estagnando a produção de proteína recombinante ($conc \approx 1.5 \text{ g/L}$)

c)

Considerando o modelo dinâmico em modo *fed-batch*, adicionou-se o parâmetro da diluição a cada equação do modelo em modo *batch* previamente formulado, e definindo a variação de volume como o fluxo de entrada de substrato (Fe).

$$\frac{d[X]}{dt} = \mu_1[X] + \mu_2[X] + \mu_3[X] - \frac{Fe}{V}[X]$$

$$\frac{d[S]}{dt} = -k_1\mu_1[X] - k_2\mu_2[X] + \frac{Fe}{V}(Se - S)$$

$$\frac{d[A]}{dt} = k_3\mu_2[X] - k_4\mu_3[X] - \frac{Fe}{V}[A]$$

$$\frac{d[P]}{dt} = k_{11}\mu_1[X] - \frac{Fe}{V}[P]$$

$$\frac{d[V]}{dt} = Fe$$

d)

Para simular o modelo dinâmico em modo semi-contínuo, efetuaram-se as alterações referidas na alínea anterior no modelo criado em Python. Assim, adicionou-se o caudal de entrada e a concentração inicial de substrato nos parâmetros iniciais. Estimou-se, ainda, a variação do volume (V) ao longo do tempo.

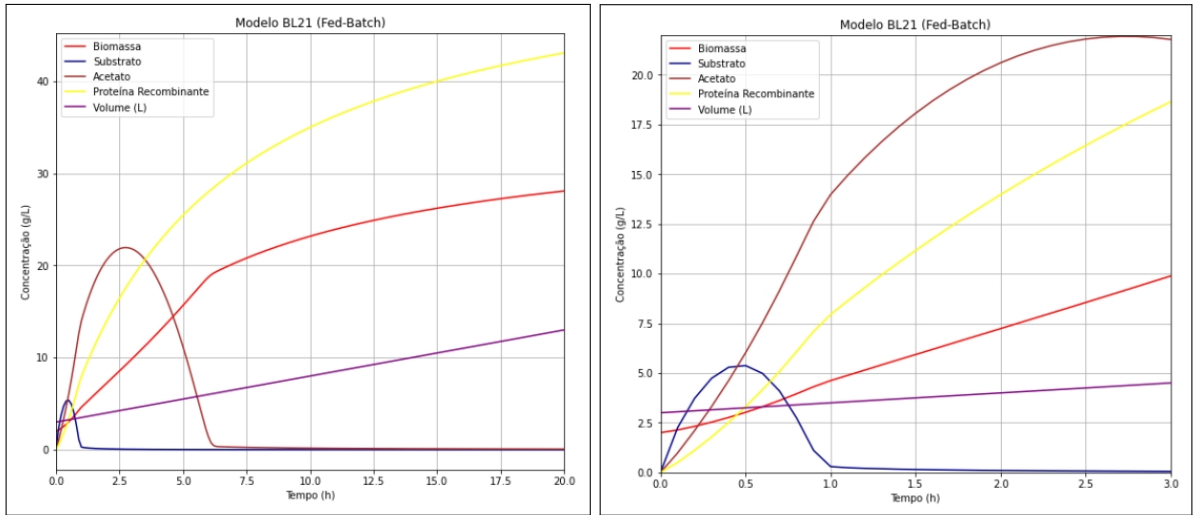


Figura 2: Modelo *fed-batch*, sendo representado a variação do volume, e das concentrações de biomassa, substrato, acetato e proteína recombinante ao longo do tempo. São apresentados dois gráficos, em que o da direita possui uma escala menor para melhor visualização do comportamento do sistema nos momentos iniciais.

Os gráficos da **Figura 2** representam um modo de crescimento semi-contínuo onde ocorre constante entrada de substrato no sistema, sem saída do mesmo, levando a um crescimento constante do volume. Inicialmente, observa-se um abrupto aumento da biomassa e, conseqüentemente, de acetato e proteína recombinante. No entanto, verifica-se uma queda na concentração de substrato ($h \approx 0.5$), pois o fluxo deste deixa de ser capaz de suportar a biomassa existente. A partir deste ponto, a biomassa recorre ao acetato como fonte de alimentação, passando este a demonstrar o mesmo comportamento ($h \approx 2.5$). Posto isto, a biomassa atingirá um equilíbrio com um constante e ligeiro aumento devido à crescente carência de substrato e acetato ao longo do tempo, conseqüentemente reduzindo a produção sistemática de proteína recombinante.

e)

Considerou-se agora o crescimento em modo *fed-batch* de uma nova estirpe **JM109**. Foram fornecidos dados experimentais onde foram medidas as concentrações de biomassa, substrato e acetato, e o volume do sistema ao longo de um intervalo de tempo equivalente a $20ut$ (medições efetuadas de 0.5 em $0.5ut$). Para verificar se o modelo criado anteriormente para a estirpe **BL21** se ajusta aos dados experimentais, criou-se um gráfico com os valores de $[X]$, $[S]$, $[A]$ e V ao longo do tempo para cada conjunto de dados, apresentado na **Figura 3**.

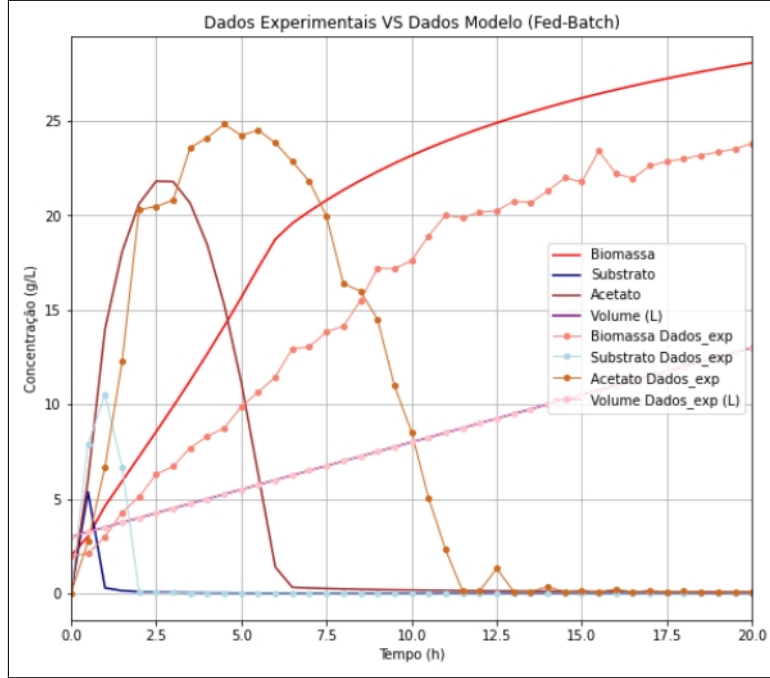


Figura 3: Representação das variações das concentrações de biomassa, substrato, acetato e do volume ao longo do tempo para o modelo BL21 em modo *fed-batch* e para os dados experimentais.

Ao analisar o gráfico da **Figura 3**, como o sistema está em modo semi-contínuo (*fed-batch*), é possível verificar o crescimento constante do volume nos dados experimentais e nos dados previstos, tendo valores equivalentes. É possível verificar que os valores previstos do modelo da biomassa são ligeiramente superiores aos valores dos dados experimentais, enquanto as concentrações de substrato e de acetato são ligeiramente inferiores. Isto pode ser uma consequência da influência dos parâmetros no sistema.

De seguida, determinaram-se os parâmetros k_1 , μ_{max1} e K_{s1} de acordo com os dados experimentais de forma a obter uma melhor aproximação destes com os estimados pelo modelo.

Em primeiro lugar, foi necessário criar o modelo dinâmico da estirpe **JM109** que apresenta muitas semelhanças com o da estirpe **BL21**, alterando-se apenas os parâmetros de entrada, que neste caso são apenas os parâmetros desejados mencionados anteriormente, sendo que os restantes foram diretamente substituídos na função.

Para que os parâmetros desejados tenham valores que permitam aproximar o nosso modelo criado com o modelo real, de forma a obter resultados o mais semelhante possível com os experimentais, foi necessário gerar valores e testá-los de forma a obter resultados que minimizem o erro em comparação com os resultados experimentais. Para isso, criou-se uma função *estimate* que aceita os parâmetros que se desejam estimar, calcula os valores para as variáveis equivalentes aos medidos experimentalmente (através da resolução de uma ode no espaço de tempo equivalente ao da experiência), e determina o erro aplicando o método *root-sum-squared* (RSS).

Foi também criada uma classe auxiliar **Bounds** para definir os limites superiores e inferiores para os valores dos parâmetros a estimar, avaliando se são respeitados.

Desta forma, foi possível estimar os parâmetros desejados ao utilizar a função *basinhopping* do *sub-package* *scipy.optimize*, que por sua vez foi utilizado para minimizar o valor da função *estimate* ao determinar os minimizantes correspondentes aos parâmetros k_1 , μ_{max1} e k_{s1} , utilizando o método iterativo "BFGS". A função *basinhopping* verifica também a validade desses parâmetros, isto é, se estão dentro dos limites definidos utilizando um objeto da classe **Bounds**.

Ao fim de 200 iterações, obteve-se uma aproximação equivalente a 3.00 para o k_1 , 0.01 para o μ_{max1} , e

0.45 para o k_{s1} . Estes parâmetros foram utilizados para gerar os valores teóricos de $[X]$, $[S]$, $[A]$ e V , e a partir destes gerar um gráfico que representa as variações destes valores ao longo do tempo. No mesmo gráfico, representaram-se as mesmas variações obtidas pelos dados experimentais para ser possível comparar.

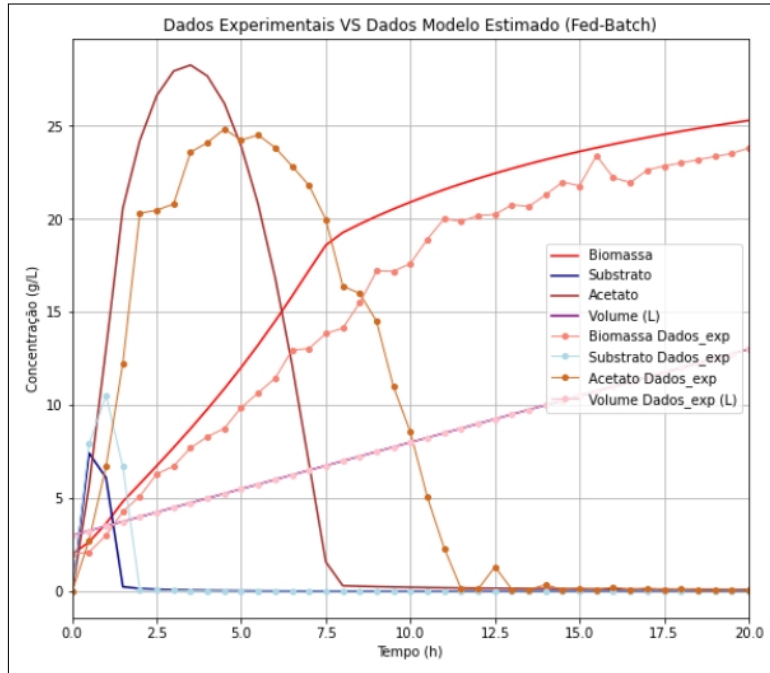


Figura 4: Representação das variações das concentrações de biomassa, substrato, acetato e de volume ao longo do tempo para o modelo JM109 em modo *fed-batch*, considerando os novos valores definidos para os parâmetros k_1 , μ_{max1} e k_{s1} , e para os dados experimentais.

Como podemos verificar no gráfico da **Figura 4**, e considerando os valores experimentais representados, foi possível aproximar o nosso modelo estimado aos valores dos dados experimentais, isto é, verifica-se um aumento dos valores da concentração de substrato e acetato, e diminuição dos valores da concentração de biomassa. Esta aproximação foi alcançada com alterações nos parâmetros k_1 , μ_{max1} e k_{s1} .

f)

Para determinar as sensibilidades das equações correspondentes à variação de biomassa ($[X]$) e de substrato $[S]$ em relação aos parâmetros k_1 , μ_{max1} e k_{s1} , foi preciso derivar cada equação em função de cada um destes parâmetros e determinar os respectivos resultados substituindo as incógnitas determinadas ao longo do tempo. Para determinar as equações das derivadas parciais, utilizou-se a função *isdiff* do *package SymPy*, e para substituir as incógnitas das respectivas equações, utilizou-se a função *lambdify* do mesmo *package*. Tendo obtido os valores, foi possível criar os gráficos das sensibilidades.

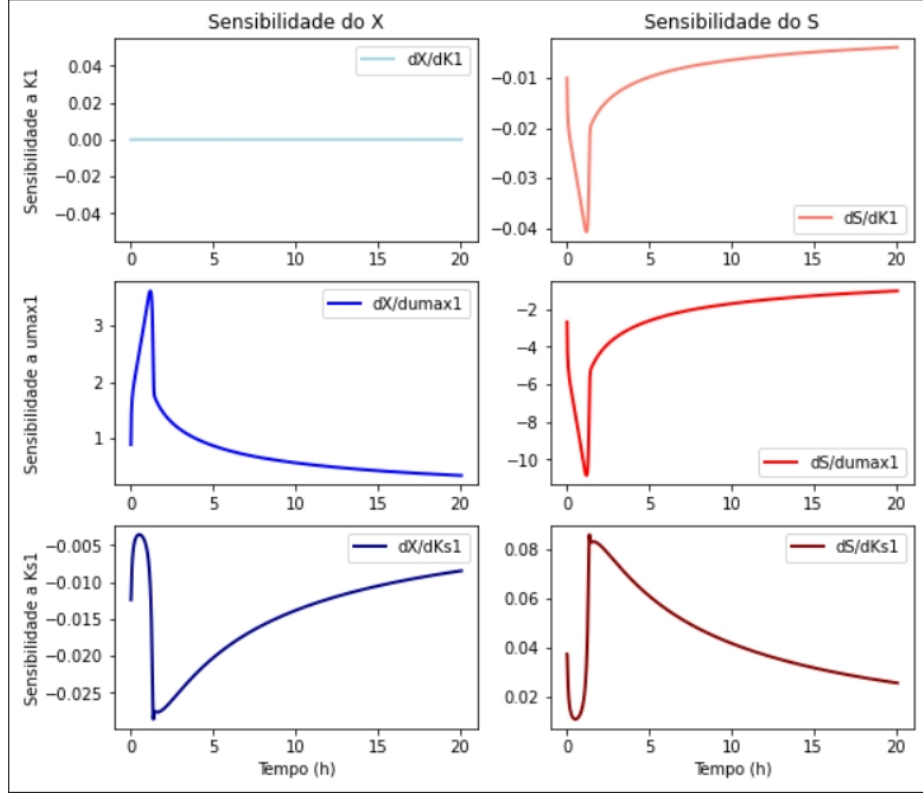


Figura 5: Gráficos que representam as sensibilidades da biomassa (Coluna 1) e do substrato (Coluna 2) em relação ao k_1 (Linha 1), μ_{max1} (Linha 2) e k_{s1} (Linha 3)

Através da análise da Figura 5, relativa às sensibilidades da biomassa (X) e do substrato (S), verificou-se que o parâmetro k_1 não tem qualquer influência sobre a variável X ao longo do tempo, no entanto, tem influência sobre a variável S, de forma negativa, isto é, diminuição deste parâmetro irá provocar um aumento na concentração de substrato. No que diz respeito ao parâmetro μ_{max1} , para a variável X, verifica-se uma sensibilidade positiva, o que significa que a diminuição deste parâmetro irá provocar uma diminuição na produção de biomassa, enquanto para a variável S, a sensibilidade é negativa, pelo que a diminuição deste parâmetro irá aumentar a concentração de substrato ao longo do tempo. Finalmente, para o parâmetro k_{s1} verificou-se que este apresenta para a variável X, uma influência negativa na sensibilidade, pelo que o aumento neste parâmetro irá provocar uma diminuição na produção de biomassa, e em relação à variável S, observa-se o oposto: uma sensibilidade positiva, ou seja, um aumento do valor de k_{s1} traduz-se numa menor quantidade de substrato.

Assim, agora é possível perceber como a alteração destes parâmetros levou à aproximação do modelo inicial aos valores experimentais. O modelo inicial previa concentrações de biomassa superiores enquanto as concentrações de substrato seriam inferiores em relação aos dados experimentais (Figura 3). Verificamos que as alterações destes parâmetros foi suficiente para aproximar o modelo aos dados experimentais, uma vez que a redução dos primeiros dois parâmetros (k_1 e μ_{max1}) leva a uma redução na produção de biomassa, exclusivamente por parte do μ_{max1} , e ao aumento do substrato ao longo do tempo, enquanto o aumento do parâmetro k_{s1} causa uma menor produção da biomassa e menor quantidade de substrato. No geral, estas alterações dos parâmetros no modelo levaram a uma redução da produção da biomassa e um aumento da concentração de substrato ao longo do tempo, que se verificou na Figura 4.

g)

Por fim, vamos considerar o sistema em quimiostato, determinar os estados estacionários, e avaliar as suas respectivas estabilidades em 2 condições: quando a biomassa tem concentração igual a 0 e quando tem concentração superior a 0. Para isso, determinaram-se os pontos estacionários num modelo de quimiostato criado, através de aproximações iniciais dos parâmetros e aplicando a função *fsolve* do *package scipy.optimize*. Neste caso, usaram-se 2 aproximações iniciais de acordo com as duas condições mencionadas anteriormente, obtendo respetivamente os pontos **[0, 300, 0, 0]** e **[35.88, 9.06×10^{-03} , 0.14, 55.96]**. De seguida, criou-se uma função para determinar o jacobiano de acordo com os símbolos das incógnitas e as equações fornecidas como parâmetros, sendo possível determinar as derivadas parciais de cada reação (usando as funcionalidades do *sympy*).

Para deduzir o comportamento de ambos os estados estacionários, obtivemos os respetivos determinantes e traços. Procedeu-se com a substituição dos valores dos pontos determinados na jacobiana definida e, recorrendo ao *package numpy*, utilizou-se o *linalg.det* para o cálculo do determinante e o *trace* para o traço. Desta forma, obteve-se os seguintes resultados:

Tabela 1: Valor do Determinante (D), Traço (τ) e Discriminante (Δ) de cada estado estacionário.

X	D	τ	Δ
$= 0$	-0.0005	0.5661	0.3185
> 0	9.1173×10^{-16}	-0.2500	0.0625

Através da análise da **Tabela 1**, inferimos que o primeiro ponto comporta-se como um **ponto sela (instável)**, já que o determinante é negativo. Assim, nas suas imediações o sistema tem um comportamento de "fuga", ou seja, qualquer perturbação no sistema (por exemplo, a adição de biomassa) leva a um afastamento do estado do sistema deste ponto estacionário.

Por outro lado, o segundo comporta-se como um **nodo estável**, pois o determinante e o Δ são positivos e o traço negativo. Este representa o ponto de equilíbrio do sistema, onde qualquer perturbação não terá qualquer influência no mesmo, pois ao longo do tempo o sistema tenderá novamente para este ponto. Neste caso, a biomassa possui um valor constante de **35.88**, o substrato de **9.06×10^{-03}** , o acetato de **0.14** e a proteína recombinante de **55.96**.