

MODELACIÓN EXPERIMENTAL

PRACTICA 2

Nombres:

Felipe Henao Gómez
Miguel Ángel Chacón López

Fecha: 21 de octubre de 2023

Profesor: Carlos Mario Vélez

Universidad EAFIT

ESCUELA DE CIENCIAS APLICADAS E INGENIERÍA

Medellín

2023

Índice

1. Introducción	1
1.1. Resumen del significado de términos	3
1.2. Hipótesis	3
2. Métodos	4
2.1. Modelo en Simulink y verificación	4
2.2. Procedimiento de estimación de parámetros	5
2.3. Resumen de pasos	6
3. Resultados	7
3.1. Curva de Linealidad y Periodo de Muestreo	7
3.2. Datos Experimentales	8
3.3. Búsqueda Exhaustiva	9
3.4. Simulink Design Optimization Toolbox	12
3.5. Método Metaheurístico	17
4. Discusión	25
4.1. Curva de Linealidad y Periodo de Muestreo	25
4.2. Búsqueda Exhaustiva	25
4.3. Simulink Design Optimization Toolbox	25
4.4. Método Metaheurístico	26

1. Introducción

En el desarrollo de este proyecto nos basamos en el trabajo de De Pillis & Radunskaya (2003). El modelo de este trabajo representa el crecimiento de un tumor y como es su respuesta a un tratamiento, hay diversos factores que intervienen en este proceso, principalmente la severidad de la enfermedad, la aplicación del tratamiento y la fuerza del sistema inmune del paciente.

En el modelo se incluyen células inmunitarias cuyo crecimiento puede verse afectado por la presencia del tumor, las células inmunitarias pueden destruir células tumorales por medio de un proceso cinético. También se incluyen unos términos de competencia, es decir, las células normales y las tumorales compiten por los recursos disponibles, mientras que las células inmunitarias y las tumorales compiten en una relación depredador-presa. En el modelo $I(t)$ denota el número de células inmunitarias en el tiempo t , $T(t)$ el número de células tumorales o tumorales en el tiempo t , $N(t)$ el número de células normales, o huésped, en el tiempo t .

Se considera que la fuente de las células inmunitarias está fuera del sistema, por lo que es razonable suponer una tasa de afluencia constante s . Además, en ausencia de cualquier tumor, las células morirán a una tasa *per cápita* d_1 , lo que resulta en un tamaño de población a largo

plazo de s/d_1 células.

La presencia de células tumorales estimula la respuesta inmunitaria, representada por el término de crecimiento no lineal positivo de las células inmunitarias

$$\frac{\rho I(t)T(t)}{\alpha + T(t)}$$

ρ y α son constantes positivas.

La reacción de las células inmunes y tumorales puede resultar en la muerte de las células tumorales o en la inactivación de las células inmunes, representadas por los dos términos de competición

$$\frac{dI}{dt} = -c_1 I(t)T(t)$$

Tanto las células tumorales como las normales se modelan mediante una ley de crecimiento logística, con los parámetros r_i y b_i que representan las tasas de crecimiento *per cápita* y las capacidades de carga recíprocas de los dos tipos de células: $i = 1$ identifica los parámetros asociados al tumor, e $i = 2$ los asociados al tejido normal. Además, hay dos términos que representan la competencia entre las células tumorales y las del huésped.

Al juntar todos los términos mencionados anteriormente se obtiene el siguiente sistema de ecuaciones diferenciales ordinarias.

$$\begin{aligned}\dot{N} &= r_2 N(1 - b_2 N) - c_4 TN \\ \dot{T} &= r_1 T(1 - b_1 T) - c_2 IT - c_3 TN \\ \dot{I} &= s + \frac{\rho IT}{\alpha + T} - c_1 IT - d_1 I\end{aligned}$$

Se añade el efecto de la droga a este sistema. Se denota por $u(t)$ la cantidad de fármaco en la zona del tumor en el momento t . Además, se supone que el fármaco mata todos los tipos de células, pero que la tasa de destrucción diferente para cada tipo de célula, con una curva de respuesta en todos los casos dada por una exponencial

$$F(u) = a(1 - e^{-ku})$$

$F(u)$ es la muerte celular fraccionada para una cantidad dada de fármaco u , en el lugar del tumor. Dado que se desconocen los detalles de la farmacocinética, se establece $k = 1$ como en estudios preliminares. Denotamos por a_1 , a_2 y a_3 los tres coeficientes de respuesta diferentes. Se añaden estos términos al sistema de ecuaciones diferenciales anterior, así como una ecuación para $u(t)$, la cantidad de fármaco en el lugar del tumor. Ésta viene determinada por la dosis dada $v(t)$, y una tasa de decaimiento *per cápita* del fármaco una vez inyectado. Esta tasa de decaimiento incorpora todas las vías de eliminación del fármaco.

El sistema con interacción de fármacos viene dado por:

$$\begin{aligned}\dot{N} &= r_2 N(1 - b_2 N) - c_4 T N - a_3(1 - e^{-u})N \\ \dot{T} &= r_1 T(1 - b_1 T) - c_2 I T - c_3 T N - a_2(1 - e^{-u})T \\ \dot{I} &= s + \frac{\rho I T}{\alpha + T} - c_1 I T - d_1 I - a_1(1 - e^{-u})I \\ \dot{u} &= v(t) - d_2 u\end{aligned}$$

1.1. Resumen del significado de términos

- $N(t)$: El número de células normales, o huésped, en el tiempo t . Se usa $N(0) = 1$
- $T(t)$: El número de células tumorales o tumorales en el tiempo t . Se usa $T(0) = 0.25$
- $I(t)$: Denota el número de células inmunitarias en el tiempo t . Se usa $I(0) = 0.15$ y $I(0) = 0.10$
- a_i : Fracción de muerte celular, $a_3 \leq a_1 \leq a_2$ y $0 \leq a_i \leq 0.5$. Se usa $a_1 = 0.2$, $a_2 = 0.3$ y $a_3 = 0.1$
- b_1, b_2 : capacidades de barrido, $b_1^{-1} \leq b_2^{-1} = 1$. Se usa $b_1 = b_2 = 1.0$
- c_1, c_2, c_3, c_4 : Términos de competición, $c_i > 0$. Se usa $c_1 = 1.0$, $c_2 = 0.5$, $c_3 = 1.0$ y $c_4 = 1.0$
- d_1, d_2 : tasas de muerte, d_1 es la tasa de muerte *per cápita* de células inmunes, y d_2 es la tasa de muerte *per cápita* de la droga. Se usa $d_1 = 0.2$ y $d_2 = 1.0$
- r_1, r_2 : tasa por unidad de crecimiento, se normaliza de tal forma que $r_2 = 1$. En el trabajo asumimos que la población de celular tumorales crece mas rápido que la población de células normales, es decir, $r_1 > r_2$. Se usa $r_1 = 1.5$
- s : Tasa de fuente inmunitaria. $0 \leq s \leq 0.5$. Se usa $s = 0.33$
- α : Tas a de umbral inmune. Se usa $\alpha = 0.3$
- ρ : Tasa de respuesta inmune. Se usa $\rho = 0.01$, para simular un paciente con un sistema inmunológico comprometido

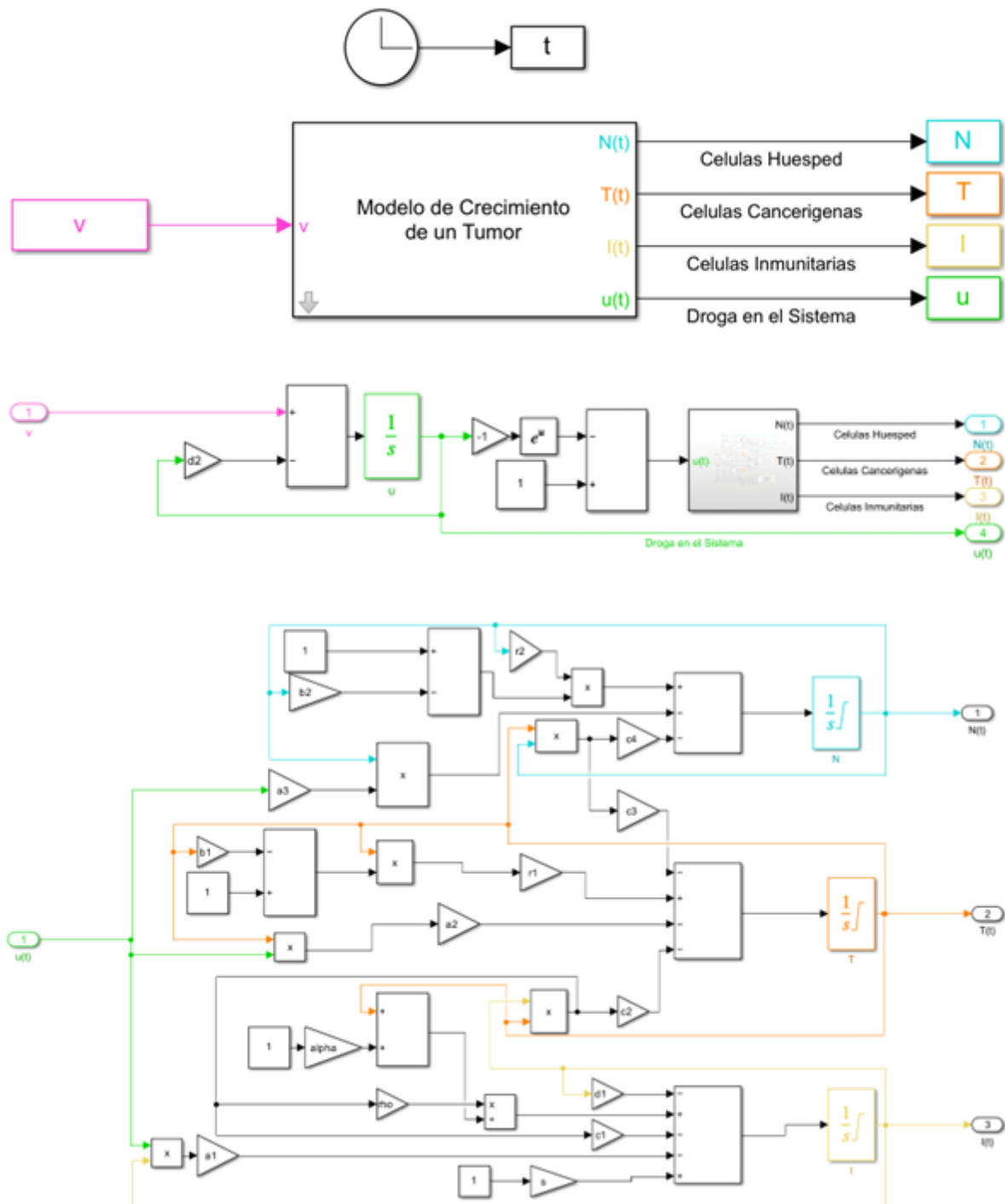
1.2. Hipótesis

1. Se espera que la función de coste con respecto a dos parámetros no sea convexa debido a la complejidad del modelo.
2. Se espera que los parámetros obtenidos usando el Simulink Identification Toolbox sean similares a aquellos obtenidos usando el método metaheurístico.

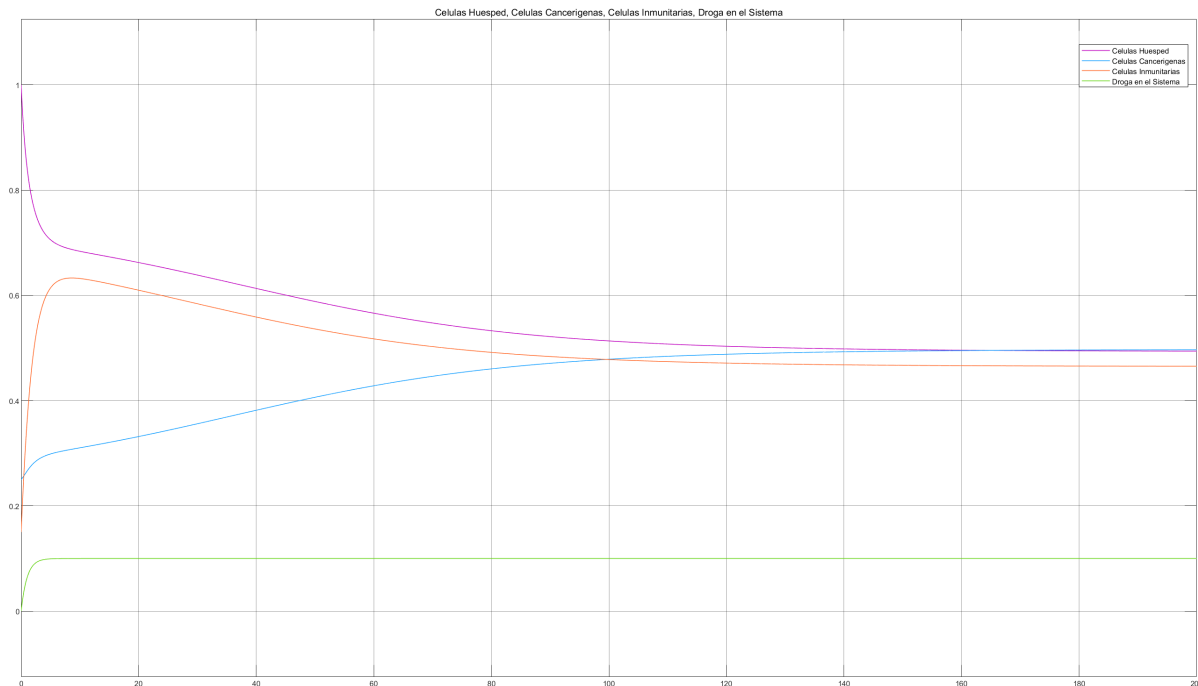
2. Métodos

2.1. Modelo en Simulink y verificación

A continuación se presenta el modelo en Simulink hecho para el modelo matemático presentado:



A modo de verificación, en primer lugar se revisó que todas las ecuaciones estén bien implementadas en el modelo y todos los signos correctos, además, se presenta la salida temporal usando las condiciones iniciales y parámetros presentados anteriormente y una entrada de 0.1:



Se puede ver que el comportamiento de las variables en el tiempo tiene sentido con el modelo, pues a medida que el tumor crece, la cantidad de células normales decrece debido a la competición por los recursos, las células inmunitarias que crecen en un principio como reacción al tumor pero luego decrecen hasta llegar a un equilibrio que también mantiene en equilibrio y muestra el comportamiento de depredador presa entre estas variables. Para la droga en el sistema se ve que esta crece hasta alcanzar el mismo valor del cual es administrado.

2.2. Procedimiento de estimación de parámetros

Para estimar los parámetros del modelo se usarán tres distintos métodos, los cuales se presentan a continuación:

■ Búsqueda Exhaustiva

El primer método que se usara sera el método de búsqueda exhaustiva, este consiste en probar distintos valores de los parámetros y quedarnos con la mejor configuración de estos. Debido a que dicho método tiene un costo de tiempo muy alto, solo se estimara 2 parámetros usando este método. Luego de obtener los parámetros estimados, se hará un gráfico 3D con todas las parejas de los valores usados de cada parámetro y su respectivo valor en la función de coste, y se hará dos gráficos 2D de los cortes transversales de cada parámetro con el mejor valor obtenido del otro parámetro.

- **Simulink Design Optimization Toolbox**

El segundo método será usando el Simulink Design Optimization Toolbox, este hace uso de métodos de optimización matemática, enfocado en modelos desarrollados en la herramienta de Simulink de MATLAB. Por medio de este método, si se estimará todos los parámetros del modelo. Luego de estimar los valores de los parámetros, por medio de la interfaz gráfica del toolbox, obtendremos los gráficos de la trayectoria de todos los parámetros y la función de coste. Haciendo uso del toolbox, también se hará los gráficos de validación para las tres distintas entradas y del análisis residual. Finalmente, usando el método del bootstrapping residual, se buscará hallar unos intervalos de confianza para los valores de los parámetros obtenidos.

- **Método Metaheurístico**

Finalmente, se desarrollará un método metaheurístico para estimar los parámetros. La ventaja de estos es que no asumen nada sobre el problema que se desea optimizar. Para esta práctica, se usará como base el algoritmo de evolución diferencial, agregando algunos cambios, dicho algoritmo se explicará más adelante. Luego de estimar los valores de los parámetros, se hará los gráficos de la trayectoria de todos los parámetros y la función de coste. También se hará los gráficos de validación para las tres distintas entradas y del análisis residual. Finalmente, para obtener los intervalos de confianza, se usará la población que resulta en el final.

2.3. Resumen de pasos

Usando el modelo desarrollado en Simulink, se obtiene lo que consideramos los datos experimentales para el desarrollo de la práctica, para esto corremos usando una entrada PRBS y agregando un ruido blanco generado por MATLAB en la salida, dicha entrada se escoge de tal forma que se salga del rango de linealidad, pero además, que no se sature el modelo. Para la validación, se hace un proceso similar pero se usa una entrada escalón y otra sinusoidal. Los valores de los parámetros usados no se usan en ninguno de los próximos pasos, solo se usan para obtener los "datos experimentales", a excepción del método de búsqueda exhaustiva ya que solo se estiman algunos de los parámetros.

En primer lugar, se usará el método de búsqueda exhaustiva para estimar dos parámetros del modelo. Primero, se escogen dos parámetros del modelo los cuales serán los estimados y los demás se dejan con el valor real, y seleccionamos los valores que se probarán. Después se corre el modelo con todas las distintas parejas de valores y se selecciona la que tenga menor error cuadrático medio. Por motivos de visualización, graficamos la función de coste con respecto a las distintas parejas de los valores de los parámetros y tomamos dos cortes transversales con respecto al mejor valor de cada uno de estos parámetros.

El segundo método usado será usando el Simulink Design Optimization Toolbox, para este se usa la interfaz gráfica. Aquí agregamos todos los datos experimentales y parámetros a estimar, dándole valores iniciales a estos y rangos en los que se encuentran. Ya el toolbox se encarga del proceso de estimación usando métodos de optimización matemática y al final hacemos gráficos de ajuste para cada entrada y análisis residual. Como algo adicional, se busca calcular los intervalos de confianza de los parámetros usando el método de bootstrapping residual.

El ultimo método que se usara es un método metaheurístico para estimar los parámetros. Para esto, primero se debe describir el método que sera usado. Después de correr dicho método, se hará gráficos de ajuste para cada entrada y análisis residual, y se buscara calcular un intervalo de confianza para los parámetros obtenidos.

Finalmente, se discutirá e interpretará los resultados obtenidos.

3. Resultados

3.1. Curva de Linealidad y Periodo de Muestreo

En primer lugar, primero vemos la curva de linealidad del modelo para evaluar sobre que valores se puede considerar la entrada, para que no se sature el modelo. La curva obtenida es la siguiente:

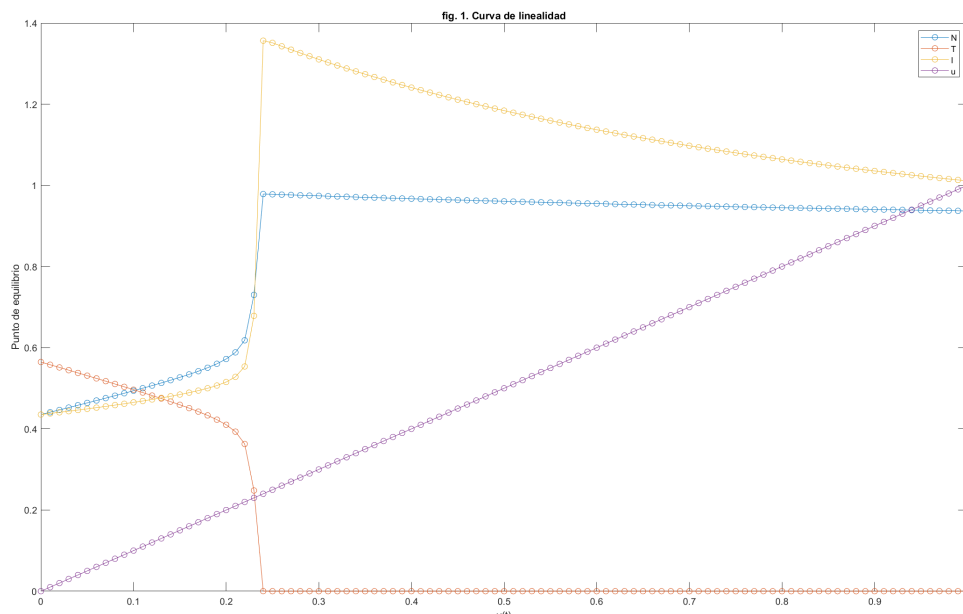


Figura 1

A partir de la curva de linealidad, podemos ver que para valores de la entrada mayores a aproximadamente 0.25, el modelo se satura, por lo que las entradas se consideran entre 0 y 0.2.

Luego, calculamos el tiempo de muestreo para tener una idea de cada cuanto tomar los datos, esto lo hacemos utilizando el tiempo de crecimiento de la variable T, la cual sera la principal para el estudio y la identificación, así, obtenemos la siguiente gráfica para el tiempo de crecimiento

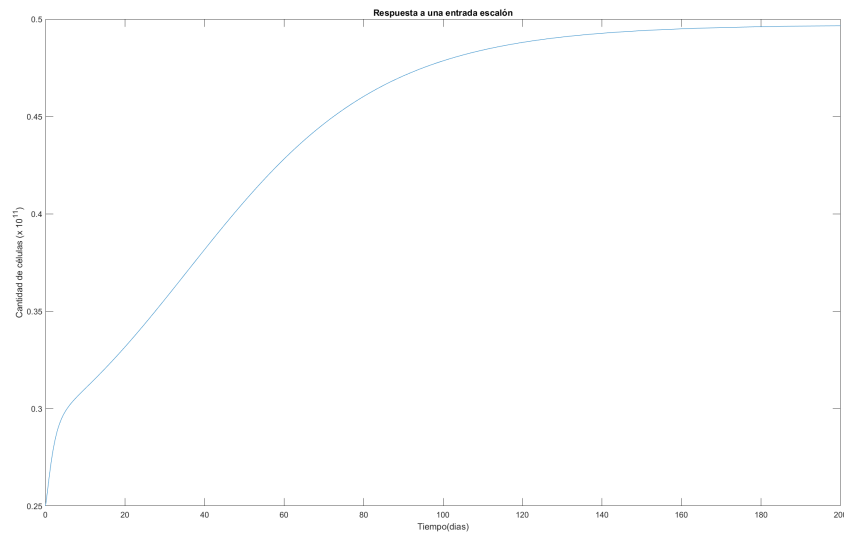


Figura 2

De esta, obtenemos que el periodo de muestreo T_s es de 9 días. Aunque no es n , se usará este periodo de muestreo para la toma de datos, pero para la estimación de parámetros este no tiene que ser fijo.

3.2. Datos Experimentales

Para la obtención de los datos experimentales, utilizamos el mismo modelo en Simulink presentado con las condiciones iniciales en el punto de operación seleccionado, tomando como entrada una PRBS que se salga del rango de linealidad pero que no sature el modelo esto es, de 0 a 0.2, y se agrega un pequeño ruido a la salida, el modelo correspondiente a esto se presenta en la siguiente figura:

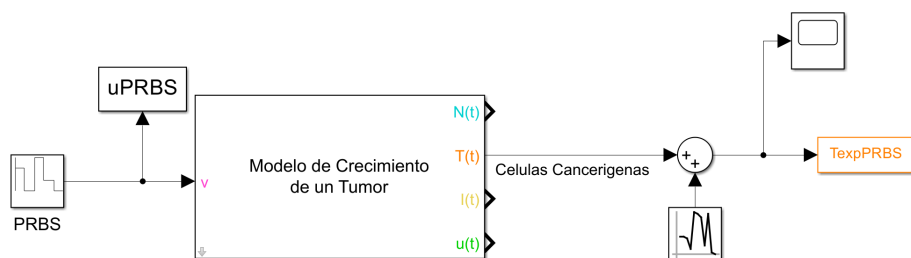


Figura 3: Modelo con una entrada PRBS

Similarmente, se hace un modelo usando una entrada escalón y sinusoidal para la validación, las siguientes figuras muestran dichos modelos:

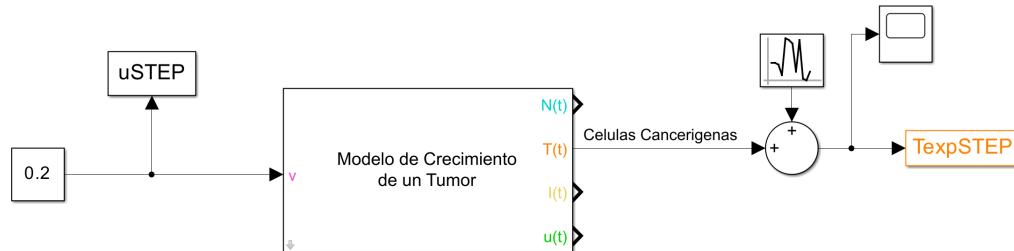


Figura 4: Modelo con una entrada escalón de 0.2 en $t = 0$

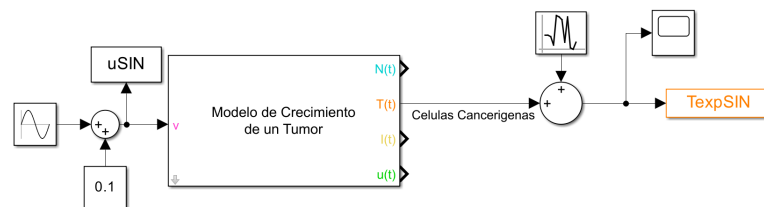


Figura 5: Modelo con una entrada sinusoidal de amplitud 0.1, centrada en 0.1 y frecuencia 0.01 rad/sec

Para la entrada sinusoidal, se tomo una frecuencia baja debido a que el modelo se demora en reaccionar al cambio en la entrada, lo que se puede ver en el tiempo de muestreo, el cual es alto.

3.3. Búsqueda Exhaustiva

Para el método de búsqueda exhaustiva, se selecciono los parámetros a_1 y s , los rangos en los que se tomo estos parámetros son $0 \leq s \leq 0.5$ $a_3 \leq a_1 \leq a_2$ y este segundo corresponde a $0.1 \leq a_1 \leq 0.3$, y se tomo como tamaño de paso para ambos de 0.005, esto es, se toman todos los valores en esos rangos con una diferencia de 0.005.

A partir de esto, usando el método de búsqueda exhaustiva con los datos experimentales con la entrada PRBS, se obtuvo que los mejores valores para los parámetros son: $a_1 = 0.2$ y $s = 0.33$.

A continuación, se muestra un gráfico 3D de la función de coste en función de los valores de los parámetros

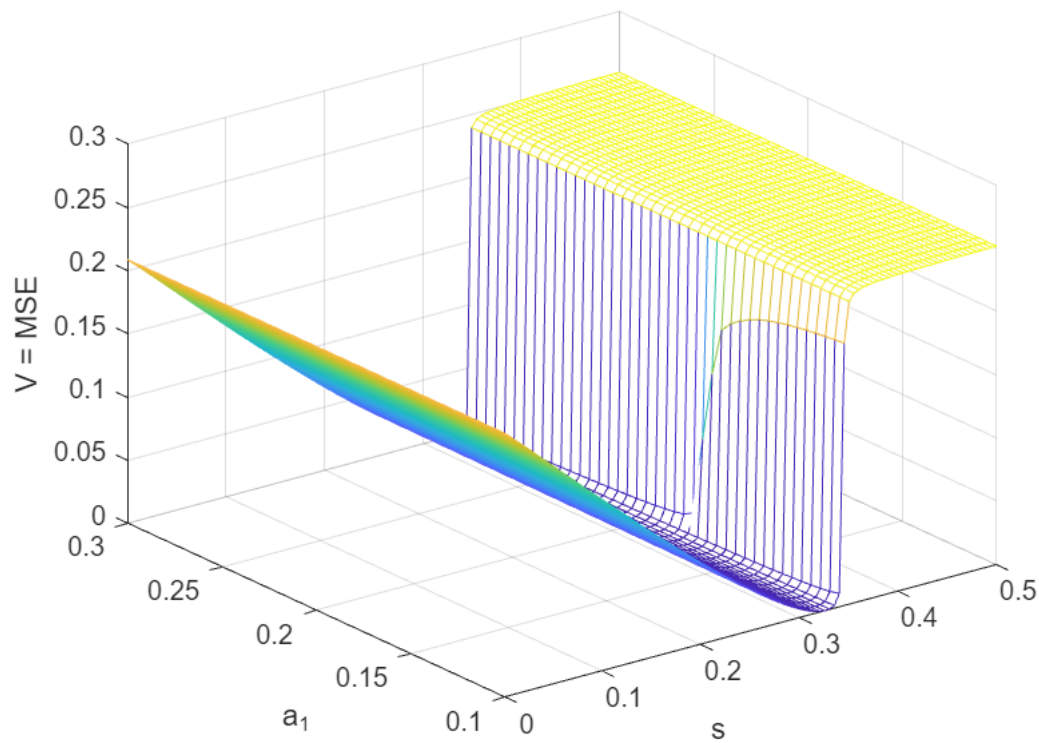


Figura 6: Función de coste en función de los valores de los parámetros

Y las siguientes gráficas muestran los cortes transversales con respecto al mejor valor de cada parámetro

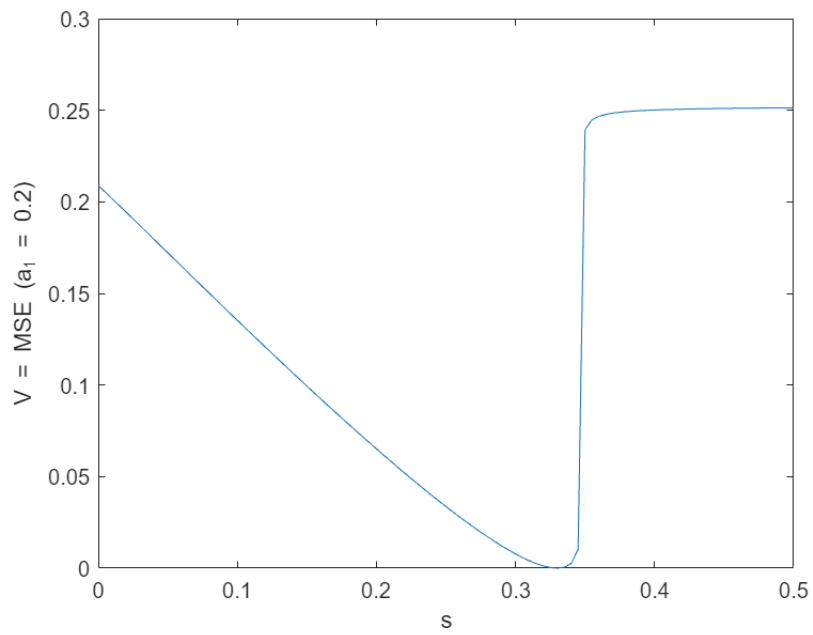


Figura 7: Corte transversal con $a_1 = 0.2$

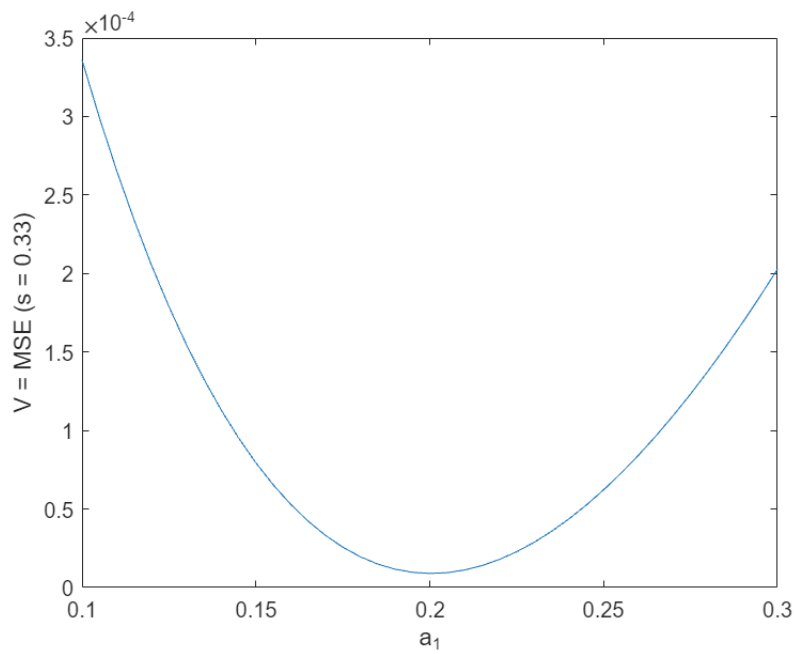


Figura 8: Corte transversal con $s = 0.33$

3.4. Simulink Design Optimization Toolbox

Ahora se uso el Simulink Design Optimization Toolbox para estimar todos los parámetros del modelo menos a b_2 . Escogimos no estimar a b_2 ya que $b_2 = 1$ por las condiciones presentadas anteriormente. Lo mismo ocurre con r_2 , ya que se normaliza para que $r_2 = 1$, pero se decidió tomar este parámetro como un parámetro de control, esto es, si el modelo estima dicho parámetro muy cerca a 1, nos da mayor indicación de que los parámetros estimados son correctos, para esto también se usa el hecho de que $a_3 \leq a_1 \leq a_2$ y $r_2 < r_1$.

Para el método, se tomaron los siguientes valores iniciales y rangos de cada parámetro:

- a_1 : inicial 0.23, rango $[0,0.5]$
- a_2 : inicial 0.26, rango $[0,0.5]$
- a_3 : inicial 0.17, rango $[0,0.5]$
- b_1 : inicial 0.9, rango $[0,1.2]$
- c_1 : inicial 0.9, rango $[0,1.2]$
- c_2 : inicial 0.6, rango $[0,1.2]$
- c_3 : inicial 0.9, rango $[0,1.2]$
- c_4 : inicial 0.9, rango $[0,1.2]$
- d_1 : inicial 0.3, rango $[0,1]$
- d_2 : inicial 0.9, rango $[0,1.1]$
- r_1 : inicial 1.3, rango $[0,2]$
- r_2 : inicial 1.0, rango $[0,2]$
- s : inicial 0.28, rango $[0,0.5]$
- α : inicial 0.2, rango $[0,1]$
- ρ : inicial 0.05, rango $[0,0.2]$

Para estimar los parámetros se tuvo en cuenta los datos experimentales con la entrada PRBS y la entrada escalón. El siguiente gráfico muestra la trayectoria de los parámetros:

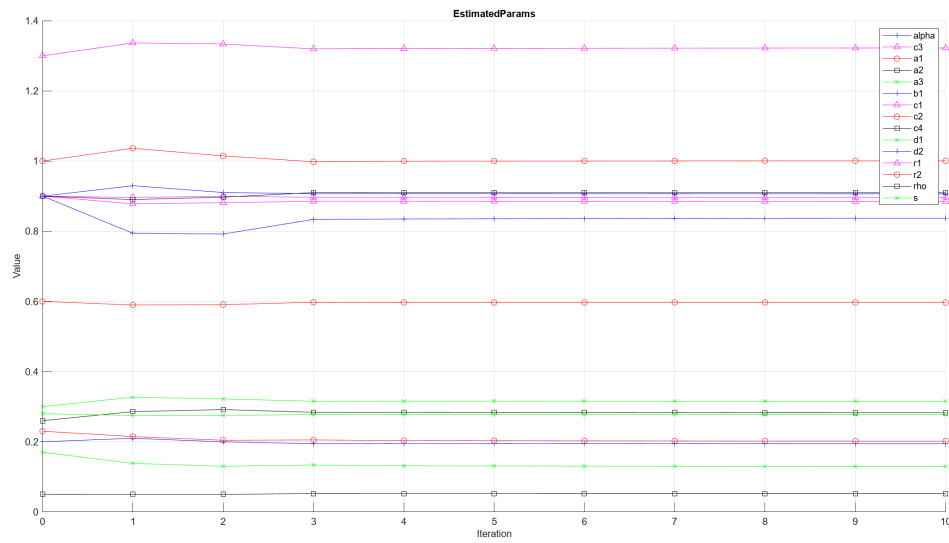


Figura 9: Trayectoria de los parámetros

y la trayectoria de las funciones de coste se muestra a continuación:

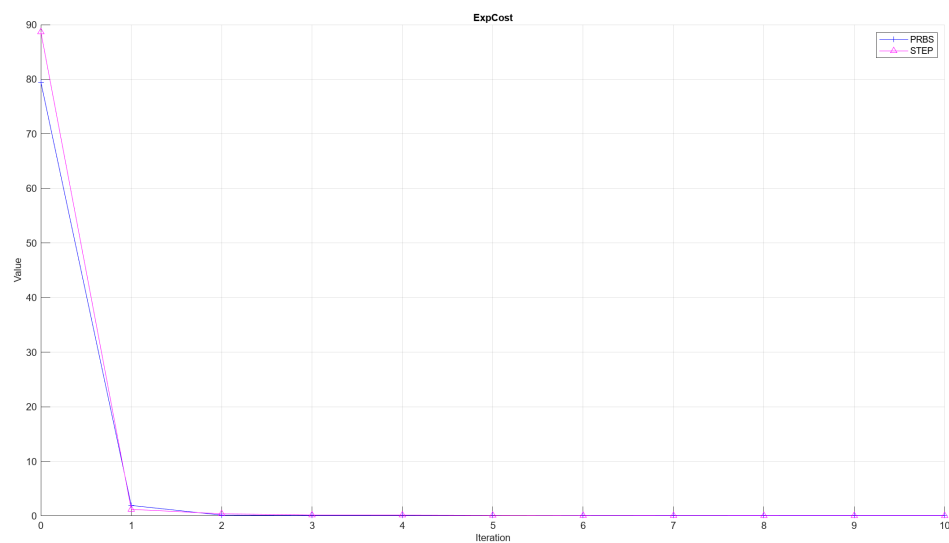


Figura 10: Trayectoria de las funciones de coste

Las siguientes tres figuras muestran los gráficos de validación con las entradas PRBS, escalón y seno, respectivamente,

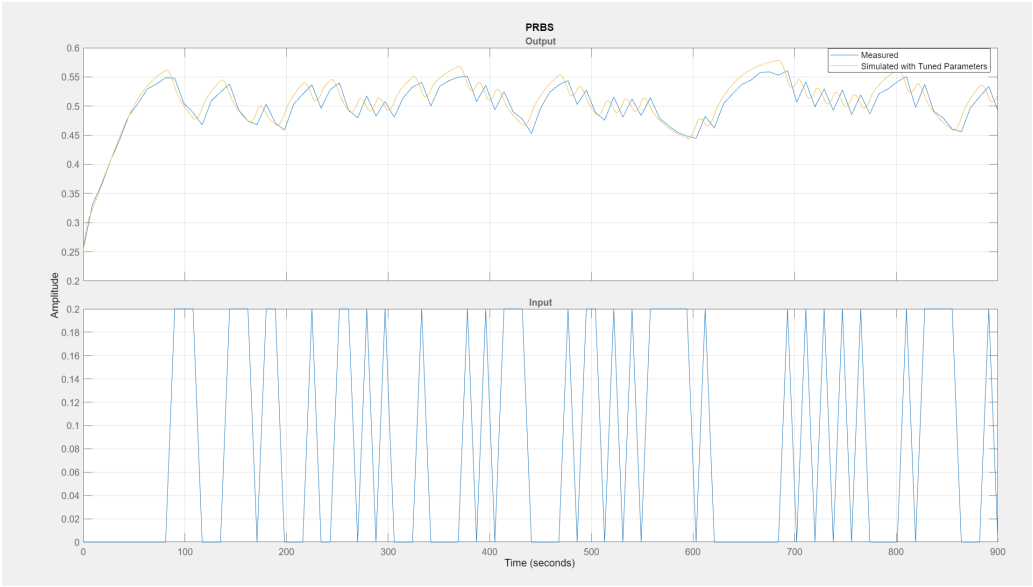


Figura 11: Gráfico de ajuste con entrada PRBS

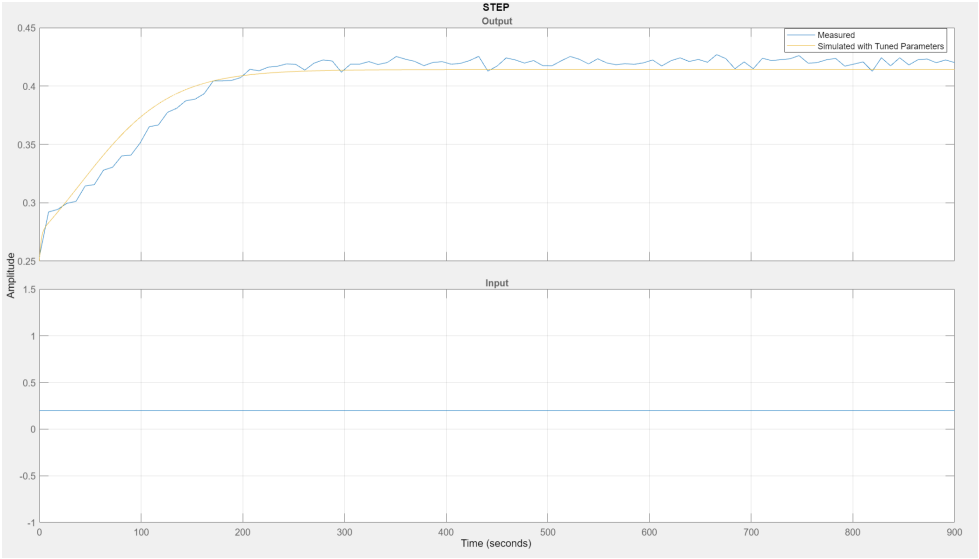


Figura 12: Gráfico de ajuste con entrada escalón

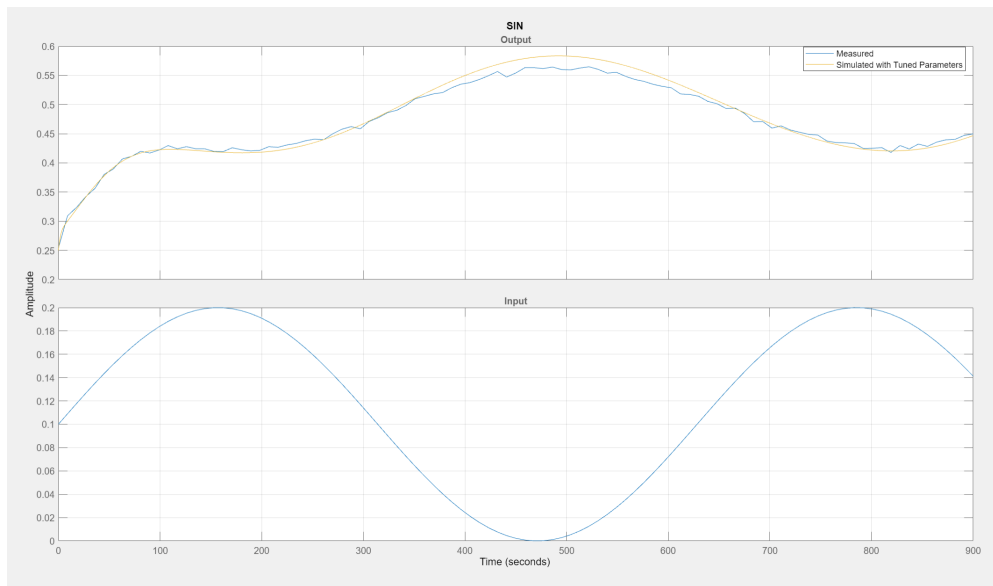


Figura 13: Gráfico de ajuste con entrada seno

y el análisis residual con la entrada PRBS

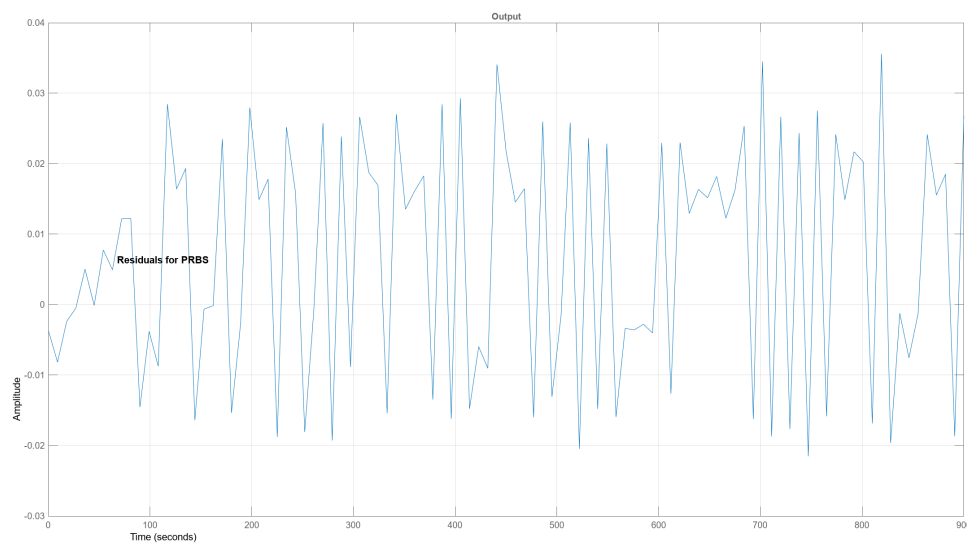


Figura 14: Gráfico del análisis residual con entrada PRBS

La interfaz gráfica del Simulink Design Optimization Toolbox se muestra a continuación:

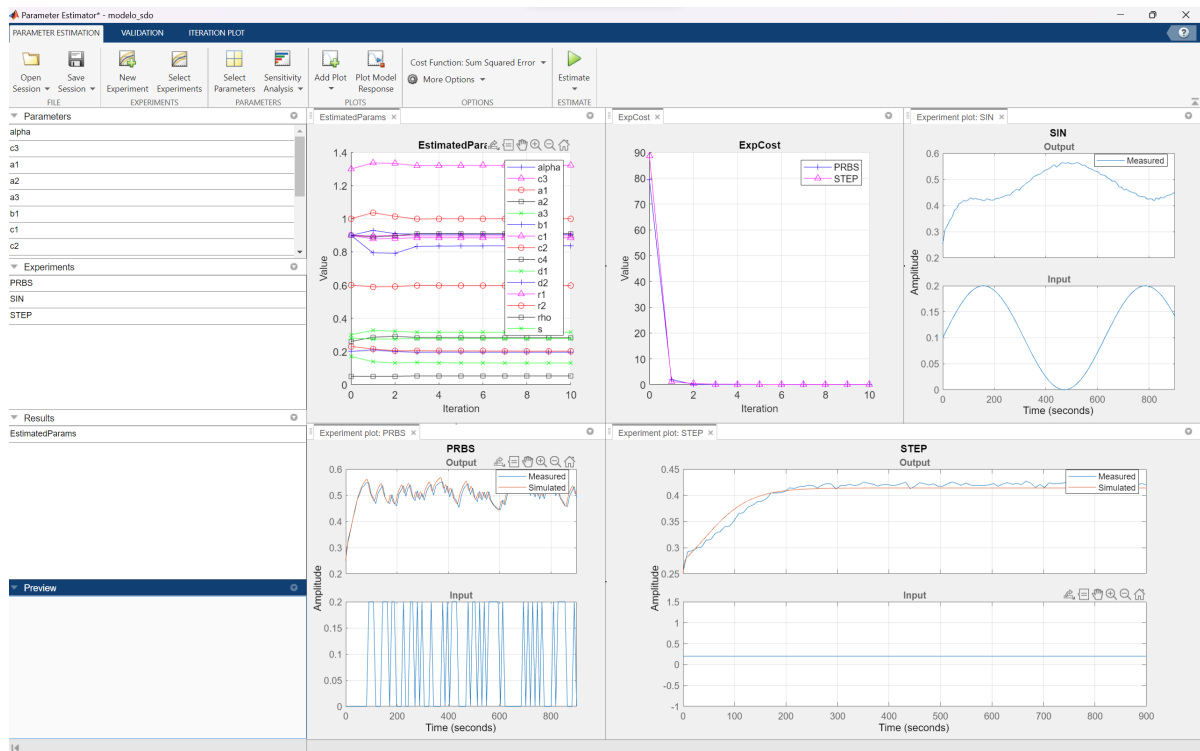


Figura 15: Interfaz gráfica del Simulink Design Optimization Toolbox

Para calcular los intervalos de confianza se usó el método del bootstrapping residual. Primero, se revisó que los residuos si fueran ruido blanco, y se obtuvo que tienen una media próxima a 0 (-0.0035) y el siguiente gráfico de autocorrelación de estos:

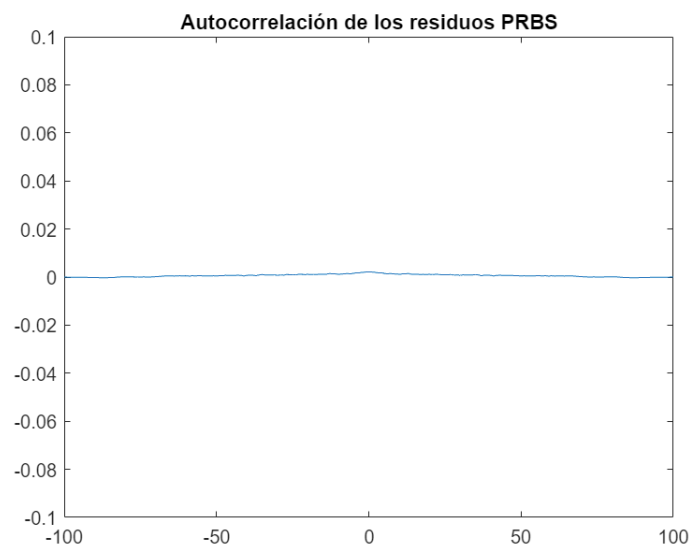


Figura 16: Gráfico de correlación de los residuos

Por tanto el ruido parece ser blanco, y entonces se puede hacer bootstrapping residual. Para esto, se hace un remuestreo con remplazo de los residuos obtenidos y se estima de nuevo pero ahora la salida a estimar sera la salida con los parámetros estimados mas el remuestreo del ruido, y esto se repite 100 veces y a partir de los resultados obtenidos se calculan los intervalos de confianza, la siguiente tabla muestra los valores de los parámetros estimados junto con sus intervalos de confianza y error relativo:

	Estimados	Intervalo de confianza	Error relativo(%)
alpha	0.19439	0.0014	0.72
c3	0.8847	0.0081	0.92
a1	0.20195	0.036	18
a2	0.28279	0.13	46
a3	0.12976	0.12	92
b1	0.90663	5e-05	0.0055
c1	0.89657	0.015	1.7
c2	0.59684	0.023	3.9
c4	0.90992	0.0068	0.75
d1	0.31548	0.0018	0.57
d2	0.8372	0.11	13
r1	1.3221	0.0074	0.56
r2	1.0005	0.0054	0.54
rho	0.052278	0.0028	5.4
s	0.27736	0.0053	1.9

Figura 17: Valores de los parámetros estimados y sus intervalos de confianza

3.5. Método Metaheurístico

Para la estimación por medio del método metaheurístico, se uso como base el algoritmo de evolución diferencial como se presenta en Wikipedia (2023). A continuación se presenta una descripción del algoritmo usado:

- Se escoge los parámetros NP, CR y F los cuales corresponden al tamaño de la población, la probabilidad de cruce y el peso diferencial.
- Se inicializan todos los individuos \mathbf{x} aleatoriamente en el espacio de búsqueda.
- Se repite lo siguiente hasta una condición de parada, para este caso se uso un numero de iteraciones definido previamente.
- Se repite lo siguiente para cada individuo \mathbf{x} :
 - Se escogen tres individuos aleatorios \mathbf{a} , \mathbf{b} y \mathbf{c} todos distintos entre ellos y distintos de \mathbf{x} .
 - Se escoge un numero R entre 1 y la dimensión del problema n .
 - Se genera un nuevo individuo \mathbf{y} como sigue:
 - Para cada $i = 1, \dots, n$ se genera un numero uniforme r .

- Si $r < CR$ o $i = R$, se hace $y_i = a_i + F * (b_i - c_i)$, si no, $y_i = x_i$. El índice R se reemplaza con seguridad.
- En caso que y_i se salga del espacio de búsqueda, este se reemplaza con un numero aleatorio dentro de este rango.
- Si el el coste de \mathbf{y} es menor al coste de \mathbf{x} , se reemplaza el individuo \mathbf{x} por \mathbf{y}
- Al final, se devuelve el mejor individuo de la población.

De esta forma, para la estimación se uso NP=50, CR=0.8 y F=0.6, dichos parámetros se escogieron de múltiples pruebas que se hizo, y además, un límite de 400 iteraciones. Adicionalmente, se uso como salida a estimar la obtenida con la entrada PRBS y como función de coste la MSE. Para evitar errores, se uso la función "try catch" de MATLAB, la cual intenta hacer algo y en caso de que salga un error se hace lo que este en la parte del catchz además, en el modelo en Simulink se agrego un if else.^{el} cual para la simulación en caso de que el valor de la variable T sea muy grande (mayor a 200) con el fin de evitar errores.

Para este método se tomaron los siguientes rangos para los parámetros:

- a_1 : inicial 0.23, rango [0,0.5]
- a_2 : inicial 0.26, rango [0,0.5]
- a_3 : inicial 0.17, rango [0,0.5]
- b_1 : inicial 0.9, rango [0,1]
- c_1 : inicial 0.9, rango [0,1.5]
- c_2 : inicial 0.6, rango [0,2]
- c_3 : inicial 0.9, rango [0,2]
- c_4 : inicial 0.9, rango [0,2]
- d_1 : inicial 0.3, rango [0,2]
- d_2 : inicial 0.9, rango [0,2]
- r_1 : inicial 1.3, rango [0,2]
- r_2 : inicial 1.0, rango [0,2]
- s : inicial 0.28, rango [0,2]
- α : inicial 0.2, rango [0,0.5]
- ρ : inicial 0.05, rango [0,0.3]

La siguiente gráfica muestra la trayectoria de los parámetros:

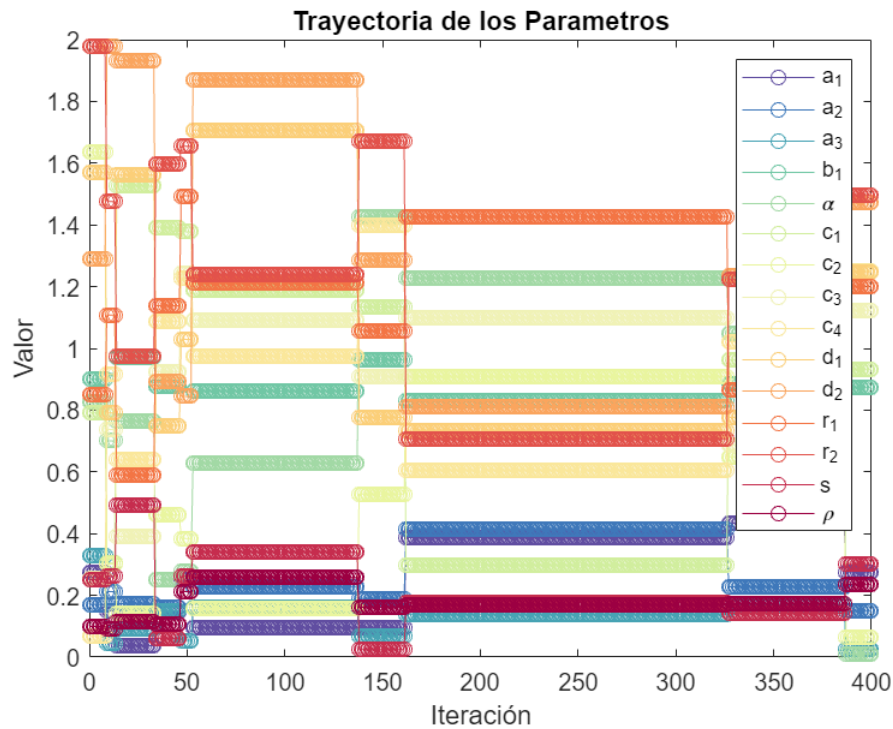


Figura 18: Trayectoria de los parámetros

la trayectoria de la función de coste:

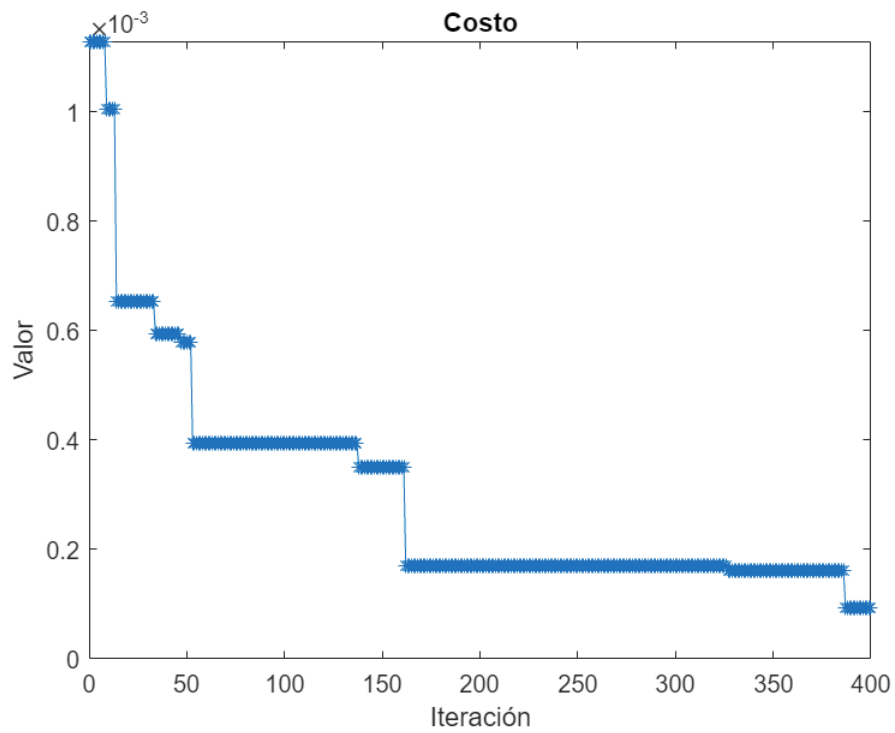


Figura 19: Trayectoria de la función de coste

Las siguientes tres figuras muestran los gráficos de validación con las entradas PRBS, escalón y seno, respectivamente,

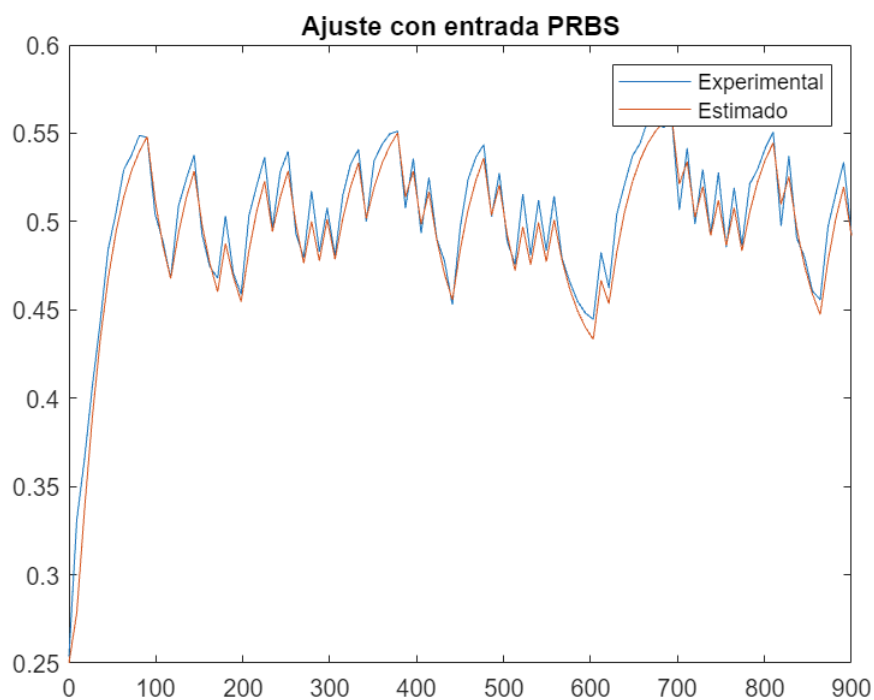


Figura 20: Gráfico de ajuste con entrada PRBS

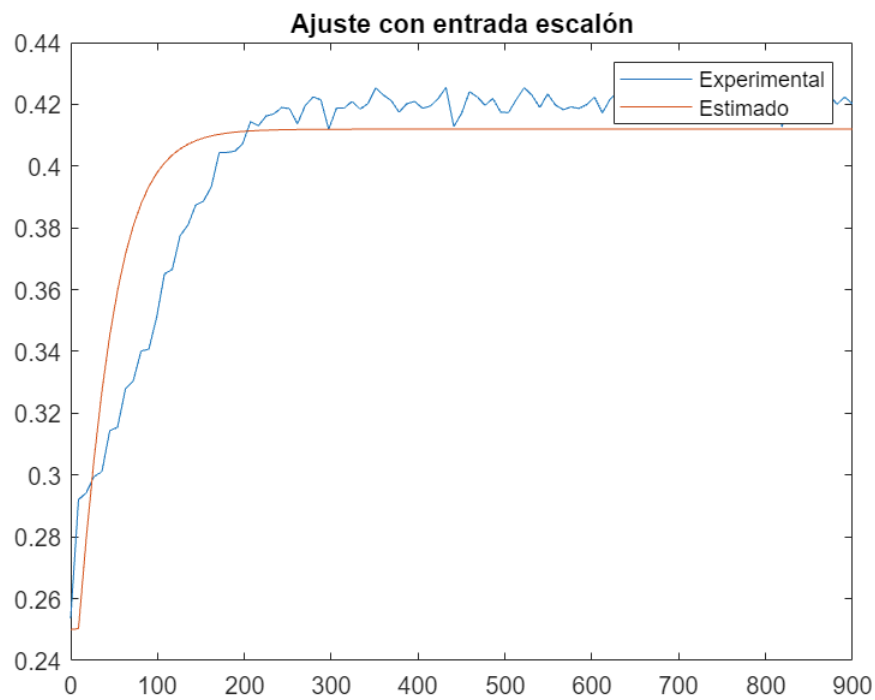


Figura 21: Gráfico de ajuste con entrada escalón

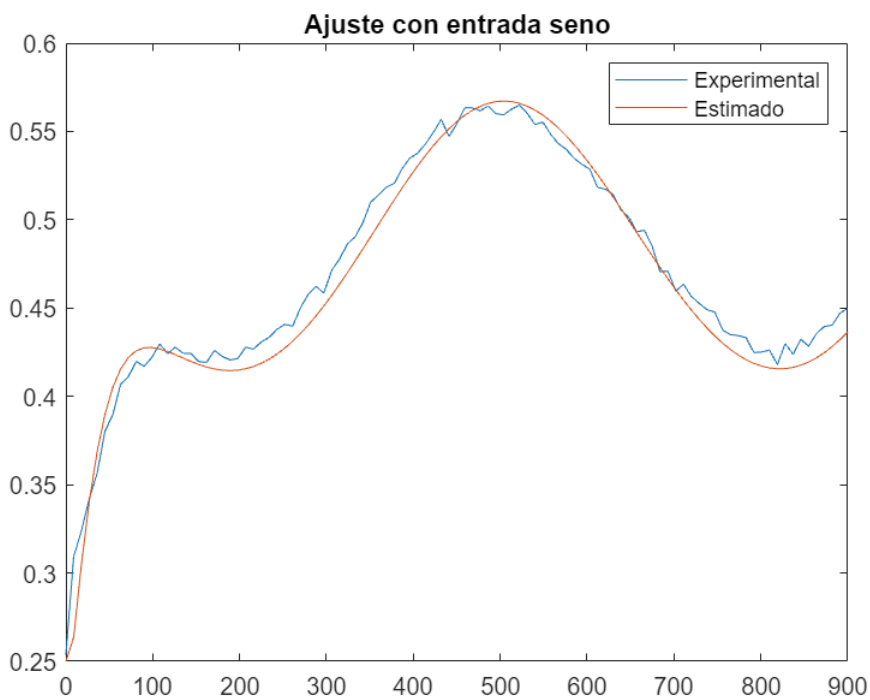


Figura 22: Gráfico de ajuste con entrada seno

y el análisis residual con la entrada PRBS

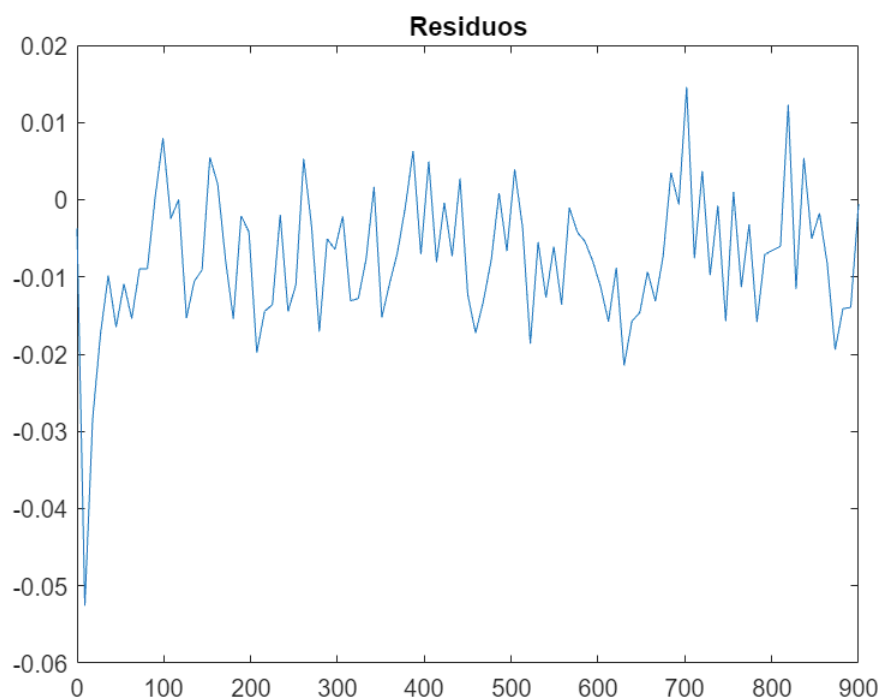


Figura 23: Gráfico de los residuos con entrada PRBS

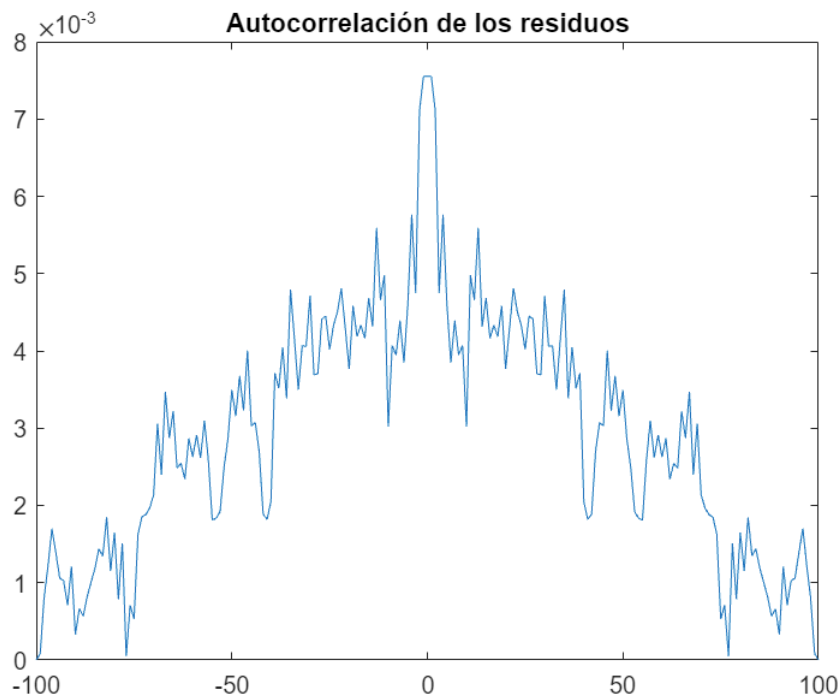


Figura 24: Gráfico de la autocorrelación de los residuos con entrada PRBS

Finalmente, usando la población en la ultima iteración, se calcularon los intervalos de confianza de los parámetros, la siguiente figura muestra estos intervalos y los valores estimados de los parámetros:

	Estimados	Intervalo de confianza	Error relativo(%)
a1	0.27346	0.042	15
a2	0.15077	0.087	58
a3	0.024692	0.11	450
b1	0.87461	0.006	0.69
alpha	0.010179	0.69	6800
c1	0.93306	0.17	18
c2	0.063631	0.51	800
c3	1.1236	0.059	5.3
c4	1.2507	0.15	12
d1	1.2456	0.17	14
d2	1.4743	0.33	22
r1	1.2012	0.099	8.2
r2	1.4979	0.022	1.5
rho	0.30272	0.1	33
s	0.23602	0.072	31

Figura 25: Valores de los parámetros estimados y sus intervalos de confianza

4. Discusión

4.1. Curva de Linealidad y Periodo de Muestreo

Podemos observar en la figura 1, la cual corresponde a la curva de linealidad, que el comportamiento de la cantidad de droga en el sistema (u) es lineal. Además, podemos ver que para entradas mayores a aproximadamente 0.25, el modelo se satura, pues la cantidad de células tumorales se va a 0.

El periodo de muestreo obtenido, con un valor de 9 días, indica que el comportamiento de nuestro modelo es gradual, esto es congruente con la dinámica que ocurre en situaciones de crecimiento tumoral en la vida real. En este contexto, no es necesario realizar un seguimiento constante de las poblaciones celulares en intervalos muy cortos, sino más bien en periodos equivalentes de alrededor de una o dos semanas.

4.2. Búsqueda Exhaustiva

A partir de las figuras 6, 7 y 8 podemos ver que la función de coste con respecto a los parámetros estimados parece ser una función convexa, esto es, tiene un único mínimo que corresponde a los valores de los parámetros obtenidos. Esto anterior desmiente la primera hipótesis establecida. Adicionalmente, podemos ver que cuando $s > 0.35$ aproximadamente, la función de coste parece ser igual sin importar el valor de s y a_1 , esto puede ser debido a que el modelo se satura para dichos valores de s , pues hay una alta fuente de células inmunes lo que resulta en que el tumor se muera, esto es, que la población de células cancerígenas se vuelva 0.

4.3. Simulink Design Optimization Toolbox

Con respecto a la figura 9, podemos ver que los parámetros no parecen cambiar mucho después de algunas iteraciones, pero esto puede ser debido a que cambian poco y no se nota por la escala de la gráfica. De la figura 10 podemos ver que para los parámetros iniciales la función de coste de estos es muy alto pero a medida que se hacen mas iteraciones este disminuye considerablemente llegando a un poco menos de 0.1 para ambas salidas. Algo que cabe notar fue que el algoritmo se detuvo debido a que el tamaño del paso era menor a 10^{-5} , esto es, el cambio en los parámetros era menor a este valor. De esto podemos decir que es bastante probable que el algoritmo halla encontrado un mínimo local pero no tenemos certeza en esto.

Con respecto a las figuras 11, 12 y 13 podemos ver que el ajuste de los parámetros obtenidos parece ser relativamente bueno aunque no llega a ser perfecto. Podemos evidenciar que hay cierta separación entre la salida estimada y la experimental, que no es mucha pero no es muy deseable especialmente en el área de medicina donde cualquier error puede llegar a tener consecuencias graves sobre la salud del paciente. Además, se puede notar que la salida estimada tiene un comportamiento similar a la experimental pero con cierto adelantamiento para el caso de la figura 11. Para 12 se puede ver que la salida estimada parece estar por debajo de la experimental lo cual no sería muy bueno ya que significa que estamos estimando que hay menos células cancerígenas de las que hay realmente lo que puede llevar a un diagnostico erróneo. Para 13,

podemos ver que en la mayor parte se estima muy bien la salida, sin embargo, en una franja de tiempo hay un error considerable en el que estamos estimando más de lo que hay realmente.

Del análisis residual, mostrado en la figura 14, podemos ver que el ruido parece tener una media de 0, sin embargo, llega a tener valores de 0.03 lo que es un valor alto para el modelo, pues son $0.03 * 10^8$ células. Esto significa, en la práctica, que los parámetros obtenidos pueden ser incorrectos o si son correctos entonces puede haber una muy mala medición de estos valores.

La figura 17 nos muestra los valores de los parámetros estimados e intervalos de confianza de estos. Primero, podemos ver que el valor de r_2 , nuestro parámetro de control, es muy cercano a 1, lo que nos da indicaciones que los parámetros estimados pueden ser correctos, además, se cumple que $a_3 \leq a_1 \leq a_2$ y $r_1 > r_2$.

Con respecto a los intervalos de confianza, podemos decir que en mayor parte, los intervalos son pequeños y el error bajo lo que indica una mayor confiabilidad. Sin embargo algunos parámetros como a_2 y a_3 presentan un alto error y amplio intervalo de confianza, lo que en un principio se podría decir que no confiable la estimación, pero se debería realizar un análisis de sensibilidad e incertidumbre para evaluar si estos parámetros son sensibles, pues si no lo son, tiene sentido que el intervalo de confianza sea amplio ya que no tiene un gran efecto en la salida del modelo. Bajo el contexto del modelo, a_2 y a_3 representan la tasa de muerte por la droga de las células cancerígenas y normales, respectivamente, por lo que una baja certeza de estos parámetros puede llevar a problemas con el paciente, y dado el caso en que dichos parámetros no sean sensibles, esto indicaría que no tiene sentido investigar en nuevos medicamentos para incrementar la tasa de muerte celular sobre el tumor debido a estos medicamentos, lo cual no tiene mucho sentido. Por tanto, no se debe confiar totalmente en los parámetros obtenidos, lo cual se puede ver un poco en el ajuste, ya que este no es del todo bueno.

4.4. Método Metaheurístico

Al utilizar este método encontramos en la figura 18 que después de aproximadamente la iteración 160 los valores de los parámetros cambian muy poco, una posible causa es como se menciono anteriormente la escala del gráfico o también porque no genera un individuo mejor luego los parámetros no cambian, debido a la aleatoriedad del algoritmo. En la función de coste, que se ve en la figura 19, observamos como a medida que aumenta el numero de iteraciones el valor de esta función va decreciendo, también podemos notar que en este proceso en algunas iteraciones como por ejemplo en desde alrededor de la 50 hasta la iteración 200 , o desde alrededor de la 160 hasta mas o menos la iteración 380, no parece cambiar mucho el valor de la función de coste, esto puede deberse a que la solución queda atrapada en una especie de mínimo local sin embargo debido a la naturaleza aleatoria que se encuentra presente en el algoritmo evolutivo que utilizamos esta solución logra salir de estos mínimos locales y seguir mejorando en su desempeño. Este bajo cambio también puede ser debido a que simplemente no encuentra una mejor solución.

Con respecto a la figura 20 podemos ver que estima de forma muy parecida a los datos experimentales pero con cierto desplazamiento hacia abajo, el cual no es mucho pero como se explico anteriormente, este puede ser significativo debido a la naturaleza del modelo que se esta estimando. Podríamos decir que los parámetros estimados usando el metaheurístico estiman

mejor la salida que los obtenidos con el Simulink Design Optimization Toolbox (SDO). De la figura 21, podemos evidenciar que los parámetros obtenidos estiman la salida de una forma similar a los obtenidos con el SDO, pero se puede ver que hay un error considerable en esta estimación. De la figura 22, podemos ver que no estima totalmente bien la salida pero aun así esta bastante cerca y tiene un comportamiento muy similar, pero podemos decir que para este caso los parámetros obtenidos con el SDO estiman de una mejor forma esta salida.

De los residuos, presentados en la figura 23, podemos ver que en un principio estos son negativos lo que indica que la salida estimada esta un poco por debajo de la salida real, pero luego se puede ver que estos tienden a estar alrededor del cero, y se puede ver que, a excepción de los primeros valores, estos residuos son menores que aquellos obtenidos con el SDO, que indica una mejor estimación. De la gráfica 24, podemos decir que los residuos prácticamente no tienen auto-correlación entre ellos.

La figura 17 nos muestra los valores de los parámetros estimados e intervalos de confianza de estos. Lo primero que se debe notar es que nuestro parámetro de control r_2 no esta lo suficientemente cerca para considerar que sea 1, por tanto podemos decir que los parámetros obtenidos, aunque en cierto modo puedan mostrar un buen ajuste, no son correctos. Además, se evidencia que no se cumple que $a_3 \leq a_1 \leq a_2$ y $r_1 > r_2$, lo que nos indica con mas razón que dichos parámetros son incorrectos. Con respecto a los intervalos de confianza, aunque no tendría sentido calcularlos debido a que se tiene resultados erróneos, podemos ver que algunos de los parámetros tienen un error muy alto, lo que indica que hay una muy alta variabilidad en la población y que el resultado obtenido no es muy confiable. Para aquellos en los que error relativo es bajo, podríamos decir que para esos parámetros el algoritmo ya había convergido de cierto modo a un valor.

A partir de esto, lo que nos permite ver es que fuera del espacio de búsqueda puede haber configuraciones de los parámetros que en cierto modo pueden estimar bien la salida. Sin embargo, estos errores pueden deberse a varios factores, como por ejemplo que no se le dio el suficiente tiempo al algoritmo debido a la cantidad de parámetros que se tiene y el tiempo que se demora en evaluarlos (simular) o simplemente dicho algoritmo no es muy bueno para resolver este tipo de problemas. Otra causa puede ser a que dependiendo de la configuración de los parámetros, el modelo se puede saturar, lo que resulta en un valor grande en la función de coste, por tanto hace mas difícil la búsqueda, aparte de la cantidad de parámetros.

Finalmente, podemos decir que los resultados obtenidos son bastante diferentes a aquellos con el SDO, lo que desmiente nuestra segunda hipótesis.

Para concluir, consideramos que se debe hacer un trabajo mas a fondo para evaluar si los parámetros obtenidos si son viables, y de forma mas profunda el método metaheurístico para evaluar que fue lo que llevo a dichos errores.

Referencias

De Pillis, Lisette G, & Radunskaya, Ami. 2003. The dynamics of an optimally controlled tumor model: A case study. *Mathematical and computer modelling*, **37**(11), 1221–1244.

Wikipedia. 2023. Differential Evolution. [En línea; Accedido 7 de octubre de 2023].