

## Revisión e identificación de sistema para empresa

INGENIERÍA MATEMÁTICA  
MODELACIÓN EXPERIMENTAL

---

**Profesor:**  
Daniel Rojas  
**Año Académico:**  
2022-2

**Camilo Oberndorfer Mejía,**  
**Miguel Valencia Ochoa,**  
**Luisa Toro Villegas**

---

### 1. Introducción

En esta ocasión nos contacta una empresa pidiéndonos ayuda con la siguiente situación: el ingeniero anterior de la empresa estaba identificando el sistema y los parámetros de la función de transferencia que lo representaran de forma adecuada sin conocimiento previo sobre este. Sin embargo, tuvieron que prescindir de él durante el proceso de identificación, por esto el proceso no está terminado y nos piden ayuda para que revisemos los adelantos hechos por el ingeniero anterior, demostremos su validez o invalidez y si el proceso es válido, terminar el proceso de identificación de la estructura del sistema y aproximación de sus parámetros y explicar y analizar los resultados.

En el livescript *plantillap1.mlx* en Matlab nos describen el proceso efectuado por el anterior ingeniero y la información preliminar descrita a continuación:

- El sistema cuenta con una única entrada y una única salida.
- El sistema tiene una entrada de 1000 unidades.

Con esta información preliminar lo primero que hizo el ingeniero anterior fue observar la salida con valores de entrada variando desde 500 a 1500 como se muestra en la Figura 1.

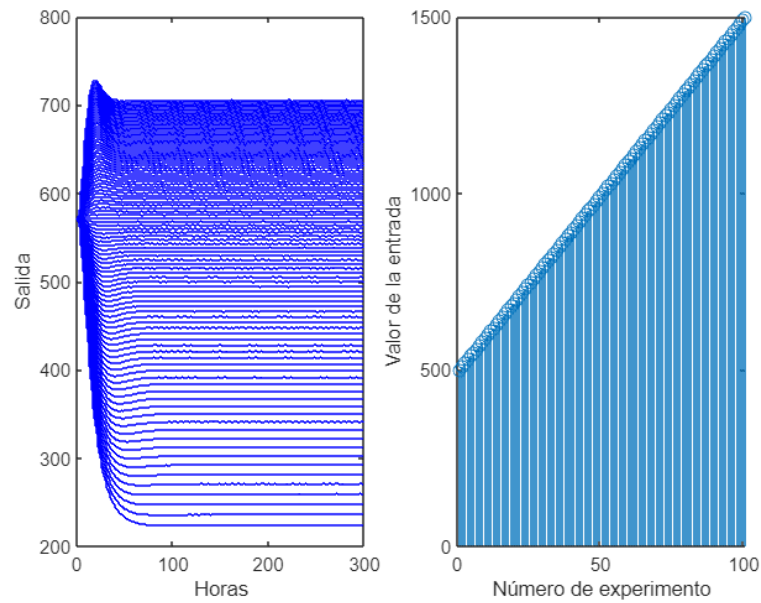


Figura 1: Salidas con variación de entradas entre 500 y 1500

De esta figura vemos que antes de la hora 100 ya presenta un comportamiento estable en general, aunque si se observa de cerca los valores presentan pequeños cambios (se podría decir que ruido). Luego hace una prueba de cambiar la entrada de 1000 a 1120 repentinamente y con la mayor velocidad de muestreo posible observa el comportamiento de la salida, Figura 2.

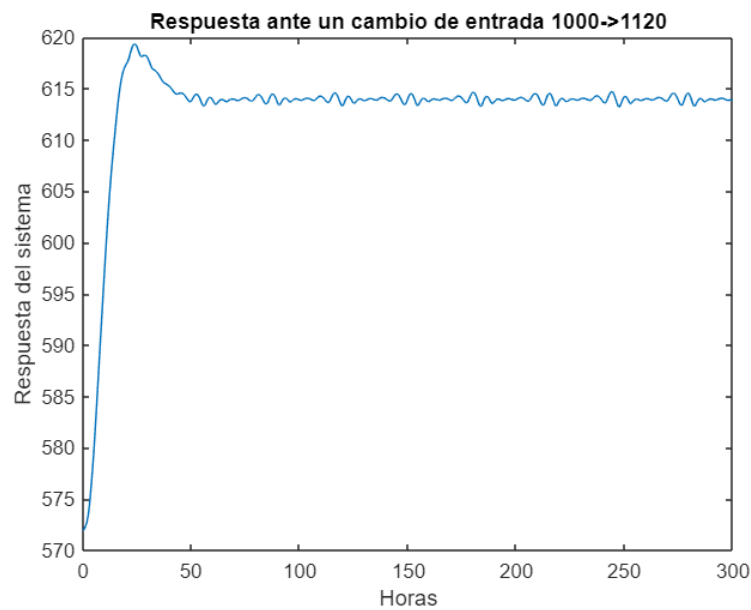


Figura 2: Salida muestreo máximo variando la entrada de 1000 a 1120

Después de esto establece un tiempo de muestreo de 1 hora y toma como entrada una señal PRBS que varía en el intervalo (900, 1120) con un periodo de una hora, lo cual se observa en la Figura 4.

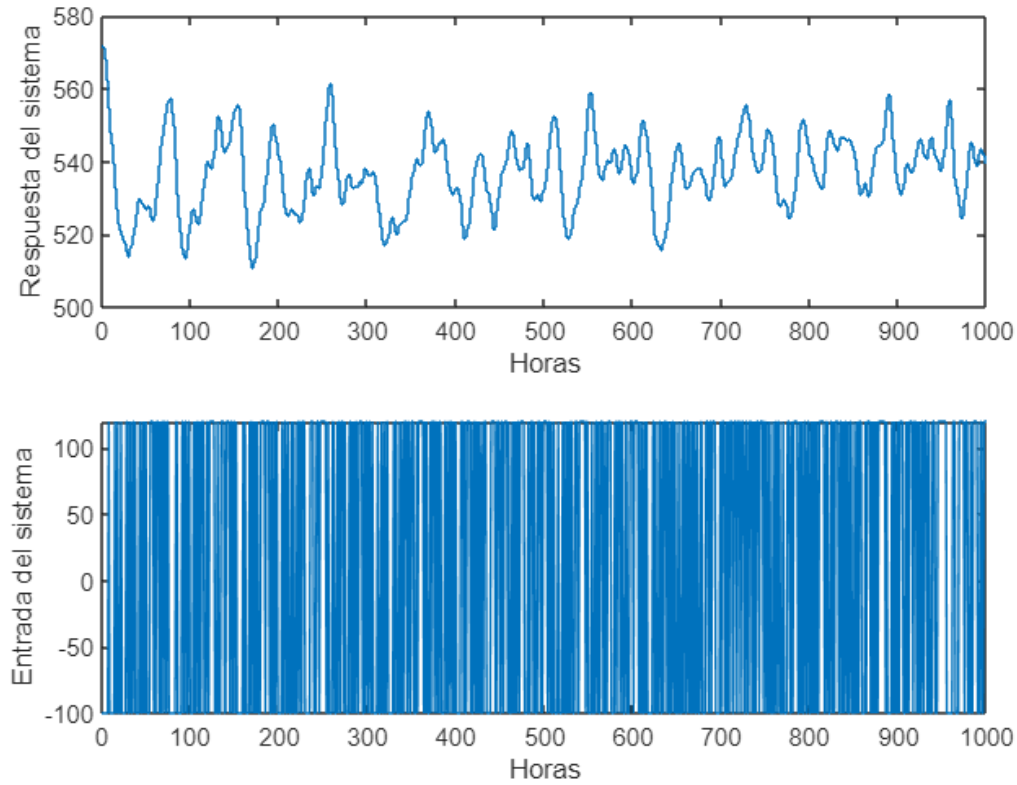


Figura 3: Respuesta a señal PRBS

La identificación de este sistema es sumamente importante para la empresa para poder establecer un controlador que permita regularlo. Esto implica que la aproximación sea lo mejor posible y que se muestre rigurosamente el proceso de revisión y aproximación, un análisis completo de los resultados obtenidos y una validación donde se comparen los valores de salida obtenidos con los originales.

## 2. Metodología

Se usará la metodología establecida en el curso de modelación experimental para identificación de sistemas no lineales.

### 2.1. Hallar el rango de linealidad

Para hallar el rango de linealidad de un sistema, donde la aproximación lineal es válida para explicar el comportamiento del sistema, el primer paso es comparar el valor al que converge las salidas dependiendo de la entrada a la que están sujetas. Luego, si los valores finales de las salidas son no lineales con respecto a los valores de entrada entonces se establece que el sistema es no lineal. Sin embargo, con datos reales el tomar solo el último valor no es lo mas confiable ya que puede estar altamente contaminado con errores por lo que se hace un promedio de los últimos 50 valores de cada salida como reemplazo.

Se establece un punto de interés para el sistema basado en información previa y se crea una línea tangente a la curva de linealidad que pase por este punto. Esta línea va a demostrar el rango en el que funciona el sistema linealizado, demostrando cuando este este demasiado desviado del sistema original, dentro de un rango de entradas llamadas rango de linealidad.

### 2.2. Establecer un periodo de muestreo

Usaremos los valores de salida de la Figura 2 para establecer un tiempo de convergencia. Tomaremos el tiempo donde la salida llega al 10 % del valor de convergencia y el tiempo donde llega al 90 %, este tiempo lo definiremos como  $T_s$  o tiempo de subida y estableceremos nuestro tiempo de muestreo  $T_m$  como cualquier valor que cumpla  $\frac{T_s}{10} \leq T_m \leq \frac{T_s}{5}$ .

### 2.3. Establecer el tipo de entrada

Escogemos una entrada PRBS (Pseudo Random Binary Sequence) que cubra nuestro rango de linealidad y que tenga un periodo igual a nuestro tiempo de muestreo para cumplir con las características de una buena entrada para identificar un sistema. Estas son un orden de excitación persistente y un factor de cresta alto. Esto se hace para excitar al sistema lo más posible y de tantas formas como se pueda.

### 2.4. Preprocesamiento de los datos

#### 2.4.1. Eliminación de outliers

Para reemplazar outliers utilizamos la función de filloutliers(movemean) de Matlab donde establecemos ventanas de tiempo donde se revisa que datos están con una separación de mas de 3 desviaciones estándar de la media de la ventana de tiempo, luego estos outliers son cambiados por el promedio entre el valor anterior y el valor siguiente. Se escogió eliminación por ventanas de tiempo ya que, al ser una serie de tiempo, es posible que los patrones de datos varíen a través de las horas de muestreo, y algunos datos que podrían ser normales en algunos intervalos, no lo sean luego.

#### 2.4.2. Eliminación de tendencias

La eliminación de tendencias se hace a través de la función de Detrend en el toolbox, que nos permite eliminar la tendencia de los datos. Esta se hace restando el dato actual del dato anterior.

$$Y(i) = Y(i) - Y(i - 1)$$

#### 2.4.3. Eliminación de medias

La eliminación de medias se hace a través de la función de Detrend en el toolbox, que nos permite restarle la media a los datos y revisar como este afecta a la entrada.

$$Y(i) = Y(i) - \mu(Y)$$

### 2.5. Algoritmo humano de identificación

La identificación consta de tres características identificables: identificación de la estructura del modelo, determinación del orden del modelo y determinación del retardo. Para las tres características usaremos análisis de la función de costos generada por la comparación entre el modelo previsto y los datos originales y el análisis residual donde vemos que la correlación entre las salidas y el ruido se mantenga menor a un 10 %. Estos análisis se harán con los resultados y gráficas obtenidas del System Identification toolbox de Matlab.

### 2.6. Validación del trabajo del ingeniero anterior

Para validar el proceso efectuado por el ingeniero anterior encontraremos el rango de linealidad, el tiempo de muestreo y la selección de entrada por nuestra cuenta y veremos si los resultados obtenidos son los mismos o compatibles con los obtenidos con el anterior ingeniero.

## 3. Resultados

### 3.1. Rango de linealidad

El promedio de las ultimas 50 horas para cada salida correspondiente a las entradas entre 500 y 1500 se usó para trazar la siguiente curva de linealidad:

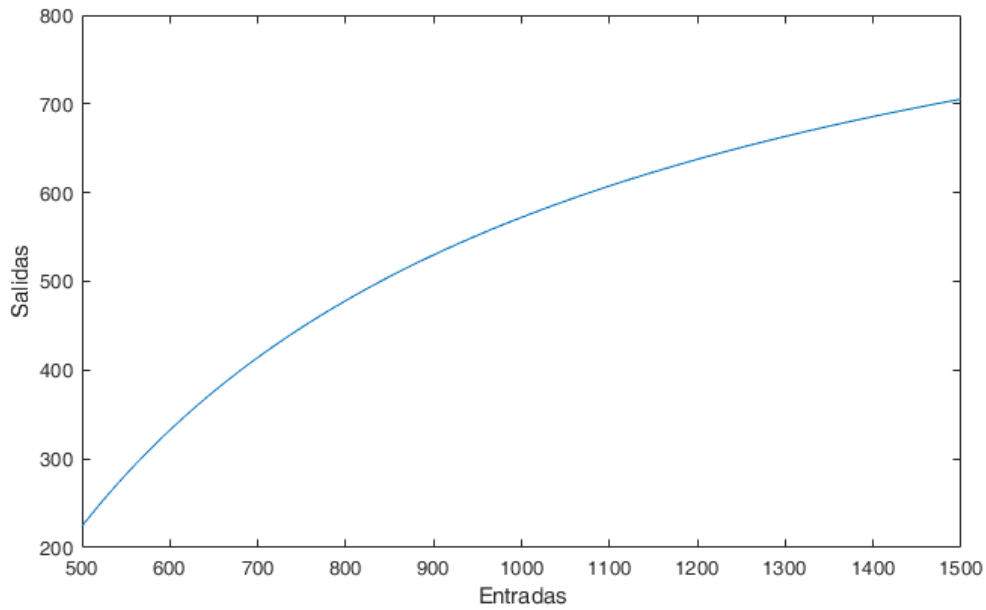


Figura 4: Curva de Linealidad

Por la Figura 4 sabemos que el sistema no es lineal, por tanto procedemos a linealizarlo en el punto  $u = 1000$  ya que este es el valor de entrada que recibe usualmente el sistema. Luego se procede a revisar si el rango de linealización que escogió el ingeniero es una buena aproximación. A través de la comparación de las respuestas de las salidas a las entradas tanto del sistema lineal como el sistema original.

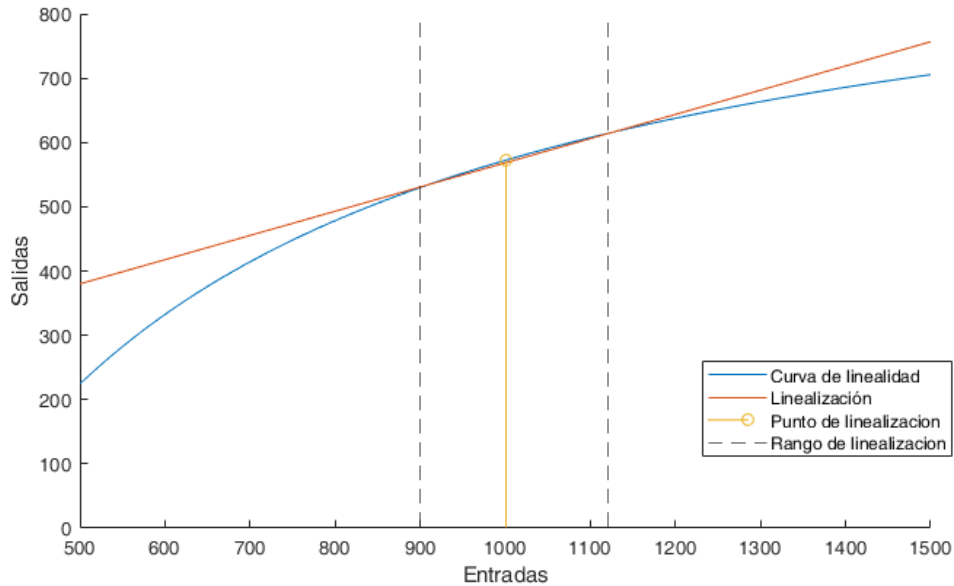


Figura 5: Rango de Linealidad

De la Figura 5 determinamos que el rango que escogió el ingeniero  $[900, 1120]$  (demarcado por las líneas punteadas verticales) ofrece una buena aproximación. Lo cual significa que hizo un buen trabajo respecto a esto. Sin embargo, se podría escoger un rango más amplio (se observa de la gráfica).

### 3.2. Periodo de muestreo

Como ya se linealizó el modelo, sabemos que para el intervalo  $[900, 1120]$  nos va a dar mayor o igual el tiempo de muestreo que variar la entrada entre  $[1000, 1120]$ . Por lo cual el resultado más conservador está dado por variar la entrada entre  $[1000, 1120]$ , de esta manera obtenemos que el 90% del cambio generado está en el

$t = 145 * (6min)$  y el 10% del cambio en el tiempo  $t = 39 * (6min)$  por lo tanto el tiempo de crecimiento  $T_r = 106 * (6min)$  y con esto establecemos que el tiempo de muestreo  $T_s$  sea un valor en el intervalo [1.06 hora - 2.12 horas] aproximadamente, entonces al ingeniero tomar un tiempo de muestreo como 1 hora, esta logrando capturar correctamente el comportamiento del sistema.

### 3.3. Escoger el tipo de entrada

Se establece una entrada PRBS que puede tomar valores en el intervalo [900, 1120] con periodo de 1 hora y se obtiene los resultados de la Figura 4.

### 3.4. Preprocesamiento de los datos

#### 3.4.1. Eliminación de outliers

Al intentar eliminar los outliers de los datos, se encuentra que no había ninguno como se observa en 6. Por tanto, estos no se van a remover. para la eliminación de estos, se eligieron ventanas móviles de 1 hora, para que estuvieran en sintonía con el tiempo de muestreo.

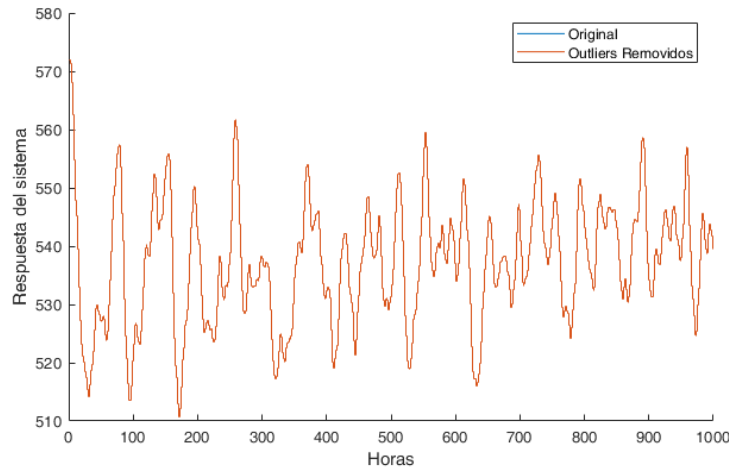


Figura 6: Eliminación de Outliers

#### 3.4.2. Remover medias

La eliminación de medias en los datos, modificaba la entrada, haciendo que esta perdiera sus características para la buena identificación del sistema 7. Por tanto se descarto este método para el preprocesamiento de los datos

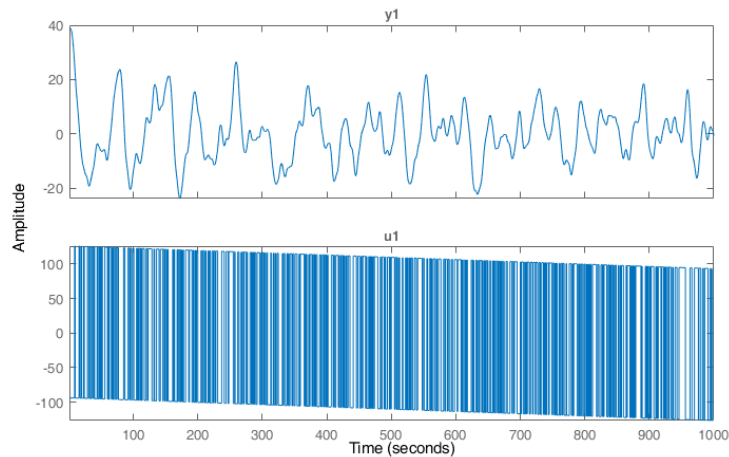


Figura 7: Eliminación de Medias

### 3.4.3. Remover tendencias

La eliminación de la tendencia de los datos, nos permite mantener a la entrada con las características para la buena identificación del sistema, a la misma vez que nos lo facilita por tanto se va a utilizar este método para estimar al sistema.

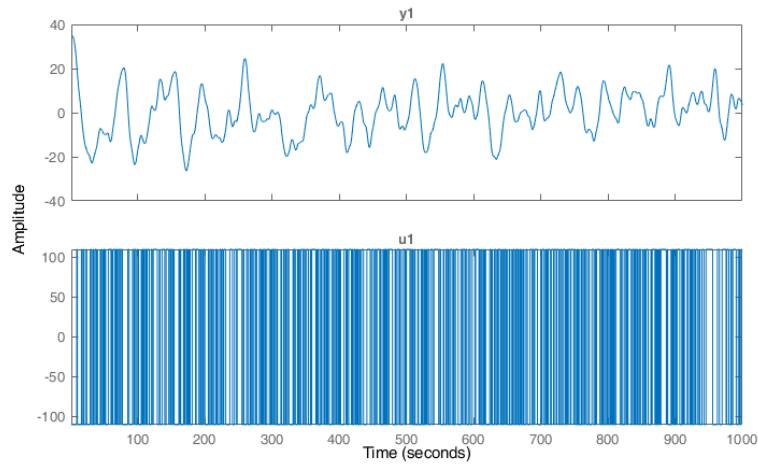


Figura 8: Eliminación de Tendencias

### 3.5. Observación de la señal de los datos

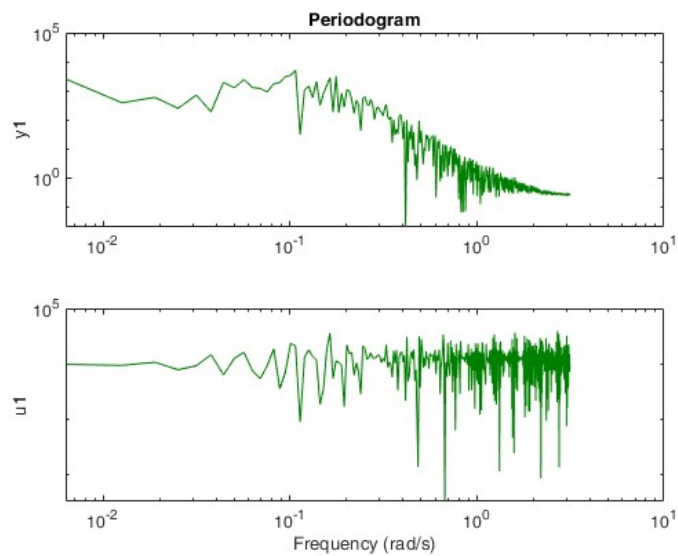


Figura 9: Periodograma

En esta figura, podemos ver la respuesta del sistema ante las frecuencias. Podemos observar también que la entrada y la salida empiezan a comportarse de forma ruidosa al mismo tiempo después de cierto instante. Lo que nos informa que el sistema no está respondiendo de forma efectiva a esas frecuencias por lo que a partir de ese punto deja de ser confiable. Por lo cual sería recomendable una señal de entrada con menos frecuencias altas.

### 3.6. Estimación del retardo del sistema

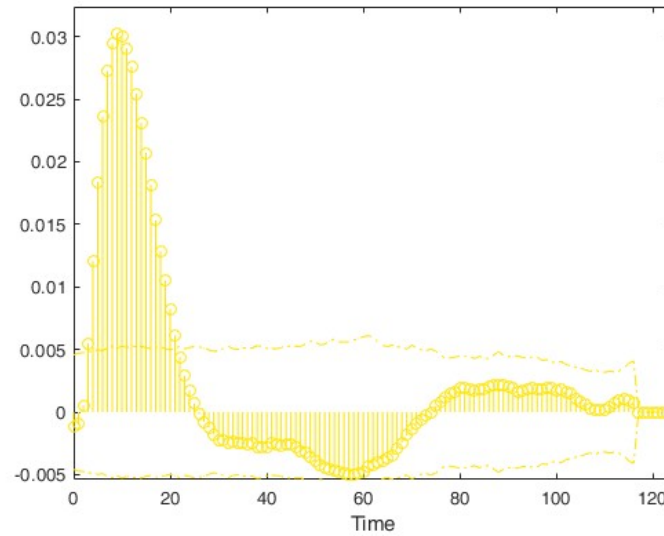


Figura 10: Respuesta al Impulso

En la imagen anterior podemos ver que al principio de esta secuencia contamos con tres ceros significativos lo cual nos indica que el retardo del sistema es de 3.

### 3.7. Algoritmo humano de identificación

Se encontró que los mejores sistemas para explicar el comportamiento de los datos para ARX, ARMAX, OE y BoxJenkins fueron:

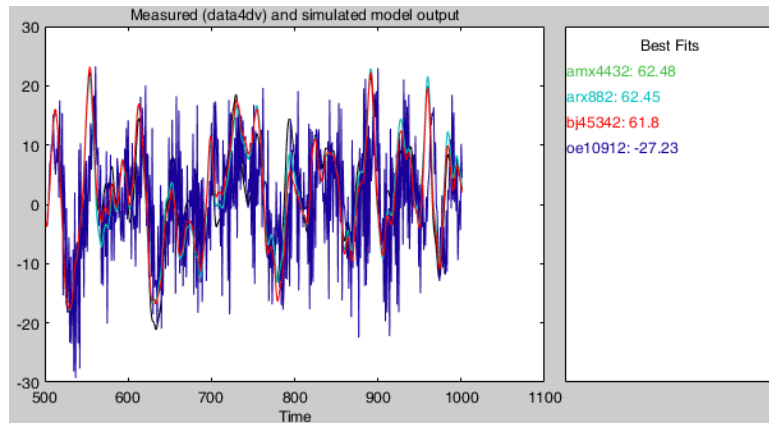


Figura 11: Sistemas mejor ajustados

#### 3.7.1. OE best fit

Dado a que todos los sistemas OE tenían un ajuste muy malo, se descartaron rápidamente, el modelo de estos que mas se ajustaba tenia  $nb = 10$ ,  $nf = 9$ ,  $nk = 12$ .



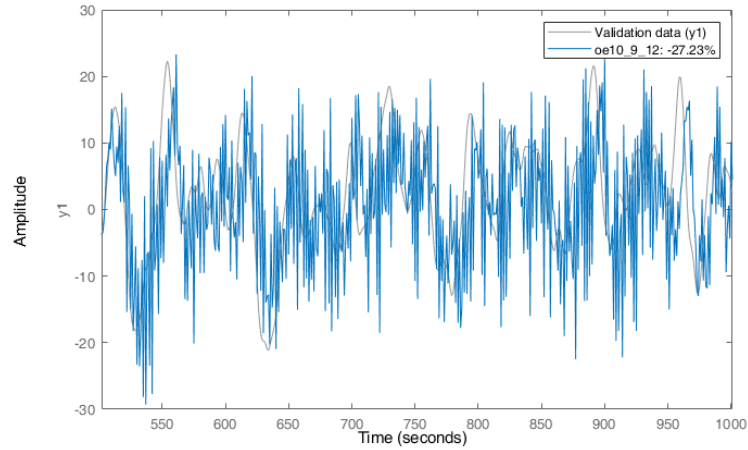


Figura 12: Ajuste del OE

### 3.7.2. ARX best fit

Para el ARX se encontró que un ARX 882, era el que mejor explicaba el comportamiento del sistema:

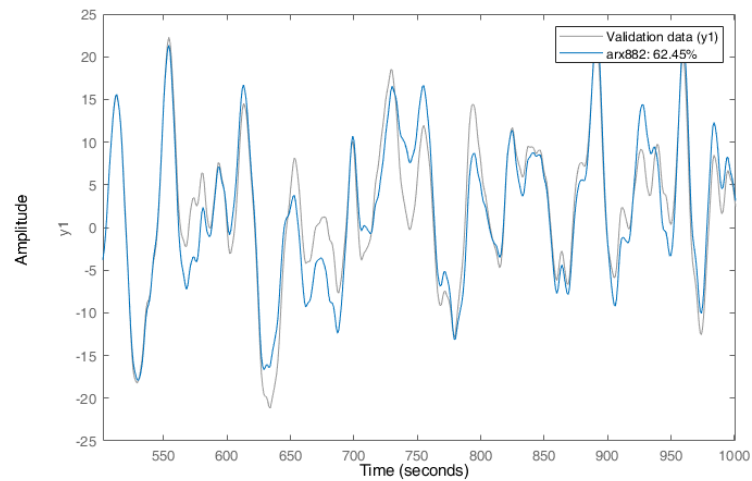


Figura 13: Ajuste del ARX882

Sin embargo este no logró presentar un ajuste tan bueno como el del ARMAX (Figura 15) o el del BoxJenkins (Figura 19) y en el análisis de residuales, este se salía de las bandas de confianza en la correlación. Además que en la correlación cruzada, este no tenía unos picos constantes fuertes. Por tanto se descartó. Además que al realizar el análisis de polos y ceros, se podía reducir la dimensionalidad y aunque se le redujera la dimensionalidad, este seguía presentando el mismo problema.

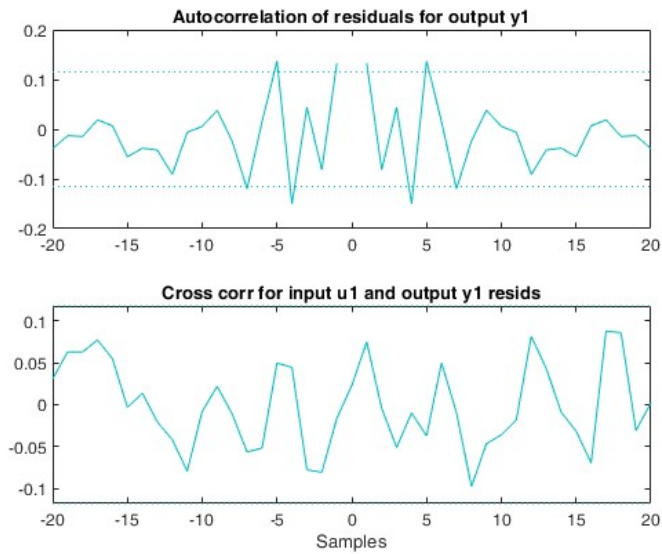


Figura 14: Análisis de residuales del ARX882

### 3.7.3. ARMAX best fit

Se encontró que el sistema que mejor ajustaba para el ARMAX es un ARMAX5452.

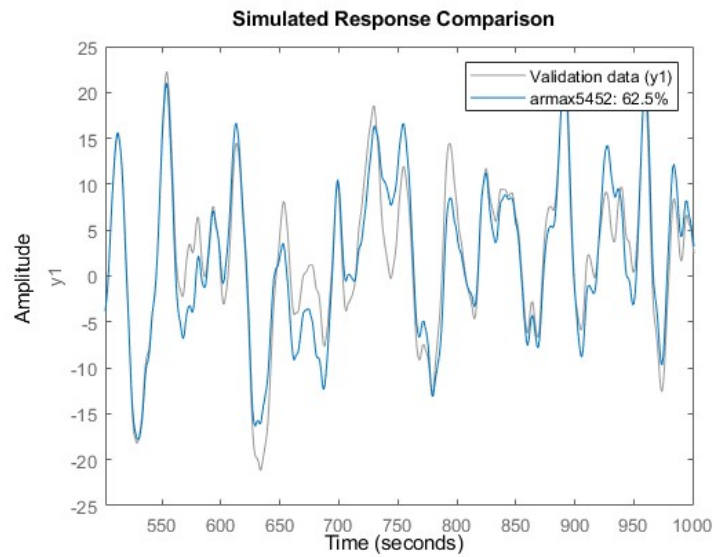


Figura 15: Ajuste del ARMAX5452

Y aunque habían otros sistemas de ARMAX que le daban una mejor explicación a los datos, los análisis de residuales de estos solían salirse de las bandas de confianza.

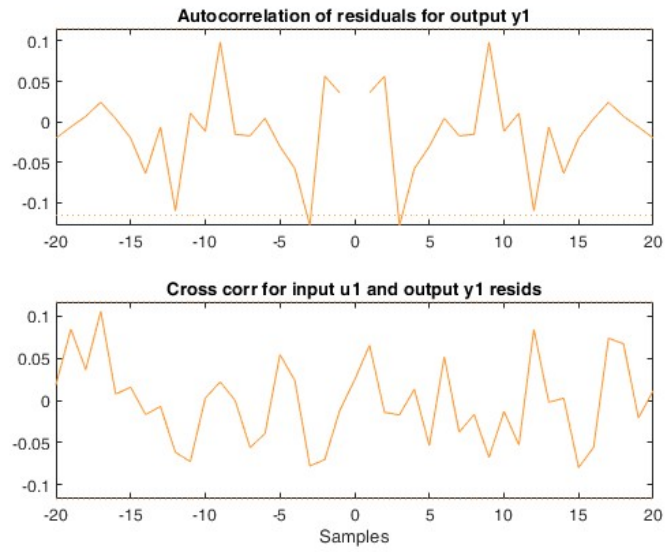


Figura 16: Análisis de residuales del ARMAX5452

Se procede a realizar el análisis de polos y ceros de este modelo y observamos que se le puede reducir la dimensionalidad.

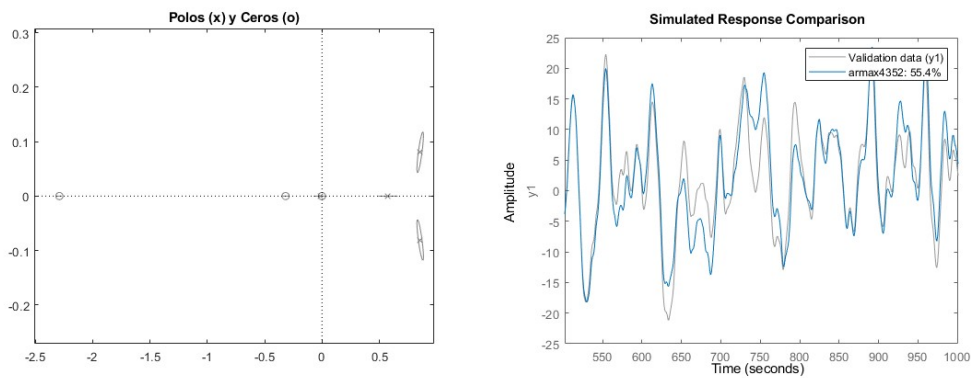


Figura 17: Polos y Ceros, y Ajuste del ARMAX4352

Al reducir la dimensionalidad del modelo revisamos su ajuste al modelo original y como todavía tiene polos cancelables, procedemos a cancelarlos.

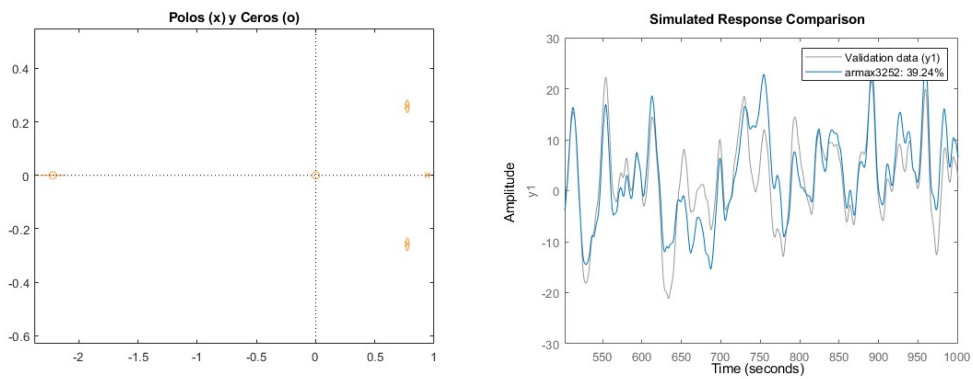


Figura 18: Polos y Ceros, y Ajuste del ARMAX3252

Al volver a revisar si esta versión del modelo tiene algún otro polo o cero cancelable, notamos que ya no hay mas.

Sin embargo, el ajuste de este modelo es muy bajo, indicando que posiblemente el sistema no es un ARMAX.

#### 3.7.4. BoxJenkins best fit

Se encontró que el modelo que mas se ajustaba entre estos fue el boxJenkins45342.

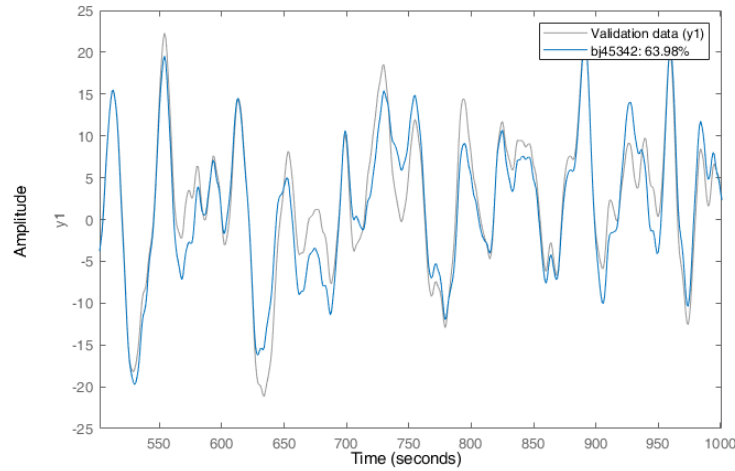


Figura 19: Ajuste del boxJenkins45342

El análisis de residuales de este también indican que este potencialmente es el sistema. Y aunque habían otros sistemas que también tenían un buen análisis de residuales con un ajuste similar entre estos, este era el mas simple y por tanto se escogió.

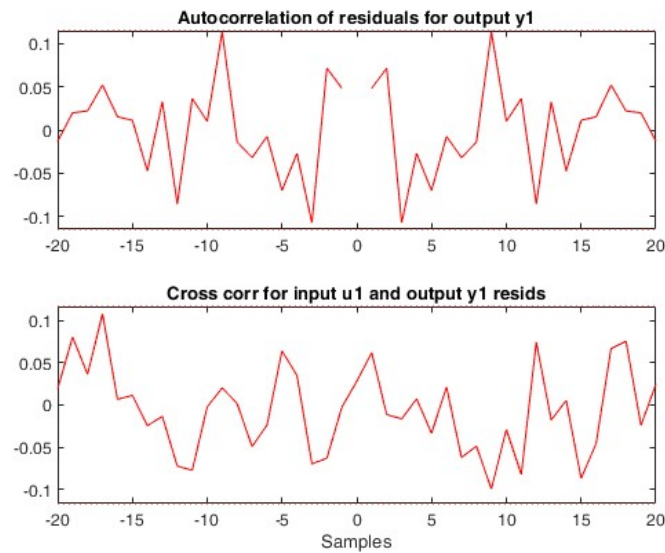


Figura 20: Análisis de residuales del boxJenkins45342

Luego procedemos a hacer el análisis de polos y ceros y notamos que hay un polo cancelable.

Al reducir la dimensionalidad del modelo el ajuste no disminuyó significativamente, y ya no hay mas polos ni ceros. Por tanto el modelo que se opina que se debería considerar como el modelo del sistema es un Box Jenkins con  $nb = 3$ ,  $nc = 5$ ,  $nd = 3$ ,  $nf = 3$ ,  $nk = 2$ .

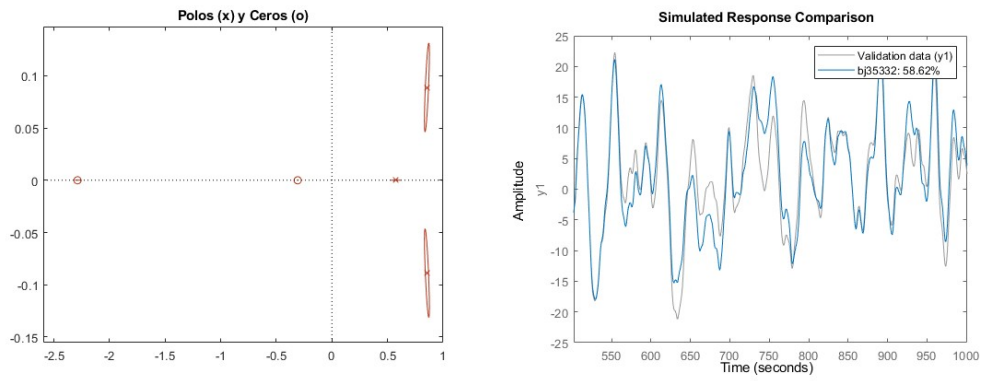
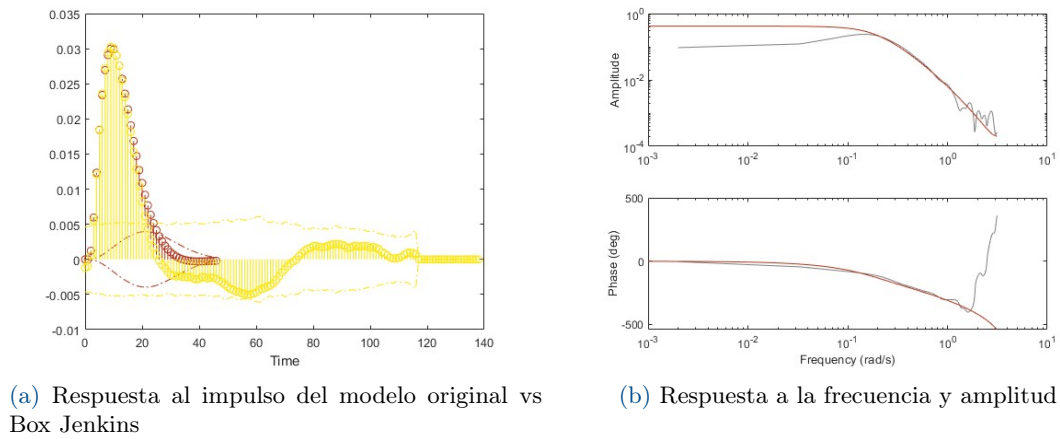


Figura 21: Polos y Ceros, y Ajuste del Box Jenkins

### 3.7.5. Comprobación del Sistema



(a) Respuesta al impulso del modelo original vs Box Jenkins

(b) Respuesta a la frecuencia y amplitud

Figura 22: Respuesta del sistema

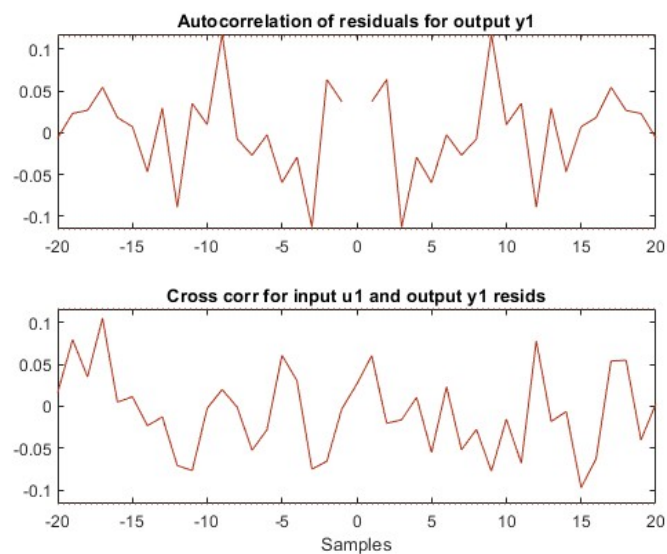


Figura 23: Análisis de residuales del Box Jenkins

Utilizamos finalmente la estructura BJ35332 ya que se mantiene dentro del rango aceptable de correlación del ruido con la salida en el análisis residual y su función de costo con respecto al sistema original no empeora considerablemente con la cancelación de polos y ceros. Como se puede ver en las imágenes, esta estructura de Box Jenkins se comporta de manera muy similar al sistema original con respecto a una entrada tipo impulso y a la respuesta a la frecuencia, sugerimos utilizar esta estructura para representar la planta en el rango de linealidad entre 900 y 1120. Sin embargo, se sugiere intentar tomar mediciones con un menor error si es posible para volver a hacer la identificación del sistema de manera mas precisa.