

IDENTIFICACIÓN DE SISTEMAS APLICANDO TÉCNICAS DE MODELADO PARAMÉTRICAS

Omar Santiago Acevedo Londoño^{1,*}, Juan Esteban Pinera Lopera^{1,*}, Sofía Raigosa Buriticá^{1,*}.

¹Facultad de ingeniería, Universidad de Antioquia, Calle 70 No. 52-21, Medellín, Colombia.

*Autores correspondientes:

omar.acevedo1@udea.edu.co, jesteban.pineda1@udea.edu.co, sofia.raigosab@udea.edu.co

1. INTRODUCCIÓN

La identificación paramétrica es una técnica fundamental en el modelado de sistemas dinámicos, que permite estimar una representación matemática del comportamiento de un sistema a partir de datos experimentales. Estos modelos son cruciales para el diseño de controladores, la simulación, la predicción y el análisis del comportamiento de sistemas dinámicos. Dentro del campo de la identificación de sistemas, los métodos paramétricos destacan por su capacidad para generar modelos con una estructura definida y parámetros específicos, lo que facilita su interpretación y aplicación, siendo correspondiente a lo conocido como sistemas de caja gris. Este proceso consiste en ajustar un modelo, caracterizado por un conjunto finito de parámetros, que minimice el error entre la salida real del sistema y la salida estimada por el modelo [1]. En el caso de sistemas discretos, un modelo genérico puede expresarse como $y(n) = f(\varphi(n), \theta) + v(n)$, donde $y(n)$ representa la salida, $\varphi(n)$ es el vector de regresores (compuesto por entradas y salidas pasadas), θ es el vector de parámetros a estimar, y $v(n)$ representa el error o ruido [2].

En este tipo de identificación, seleccionar una estructura adecuada es fundamental para representar correctamente la dinámica del sistema. Las estructuras más utilizadas se diferencian por los componentes que consideran: entradas $u(n)$, salidas pasadas $y(n - k)$, y el tratamiento del ruido o error $e(n)$. A continuación, se describen las principales estructuras:

Modelo AR (AutoRegresivo): Es el modelo más simple y considera solo la influencia de las salidas pasadas, sin tener en cuenta la entrada del sistema [3]. Se expresa como:

$$A(q)y(n) = e(n)$$

Ecuación 1. Forma polinomial modelo AR.

Esta estructura es útil cuando se desea modelar un proceso autoregresivo puro, es decir, uno que puede explicarse únicamente con su historial de salidas.

Modelo ARX (AutoRegresivo con entrada exógena): Integra tanto las salidas pasadas como las entradas del sistema, siendo ideal cuando la entrada es medible y el ruido es blanco [3]. Su forma polinomial es:

$$A(q)y(n) = B(q)u(n) + e(n)$$

Ecuación 2. Forma polinomial modelo ARX.

Es una de las estructuras más utilizadas por su simplicidad computacional y capacidad de representar muchos sistemas prácticos.

Modelo ARMA (AutoRegresivo de Media Móvil): Este modelo combina la estructura del modelo AR (autoregresivo) con la de un modelo MA (media móvil) aplicado al ruido, sin incluir la señal de entrada. Es adecuado cuando se desea representar un sistema cuya salida depende de sus valores pasados y de una perturbación que tiene una estructura temporal específica [3]. Se expresa en forma polinomial como:

$$A(q)y(n) = C(q)e(n)$$

Ecuación 3. Forma polinomial modelo ARMA.

Este modelo es útil cuando el sistema no tiene entrada conocida o controlada, pero se desea capturar tanto la dinámica de la señal como la del ruido. La inclusión del término $C(q)$ permite una mejor estimación en presencia de perturbaciones autocorrelacionadas.

Modelo ARMAX (AutoRegresivo de Media Móvil con entrada exógena): Es una extensión del modelo ARX que, además de considerar la entrada y las salidas pasadas, incluye un filtrado del ruido mediante un polinomio adicional. Esto permite capturar mejor la dinámica del sistema cuando el ruido tiene correlación temporal [3]. Su forma polinomial es:

$$A(q)y(n) = B(q)u(n) + C(q)e(n)$$

Ecuación 4. Forma polinomial modelo ARMAX.

Aquí, $C(q)$ permite modelar la dinámica del ruido, lo cual mejora la precisión en escenarios donde las perturbaciones no son puramente aleatorias, sino que siguen una estructura determinada.

Modelo OE (Output Error): En este modelo, la salida se expresa como el cociente entre dos polinomios en q , representando la verdadera dinámica del sistema, y el error se agrega directamente sin modelar su dinámica [3]:

$$y(n) = \frac{B(q)}{F(q)}u(n) + e(n)$$

Ecuación 5. Forma polinomial modelo OE.

Es adecuada cuando se tiene confianza en que el ruido es blanco y se desea representar con precisión la respuesta del sistema.

Modelo Box-Jenkins (BJ): Es la estructura más general. Modela de forma separada tanto la dinámica del sistema como la del ruido, usando polinomios independientes [3]. Se expresa como:

$$y(n) = \frac{B(q)}{F(q)}u(n) + \frac{C(q)}{D(q)}e(n)$$

Ecuación 6. Forma polinomial modelo Box-Jenkins (BJ).

Es útil cuando se desea un modelo muy detallado, incluyendo la dinámica del sistema y del ruido por separado, aunque su identificación es más compleja.

Para evaluar la precisión de los modelos identificados, se utilizan métricas estadísticas que cuantifican la diferencia entre la señal observada y la señal predicha por el modelo. El Error Absoluto Medio (MAE) se calcula como:

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i|$$

Ecuación 7. Fórmula del error absoluto medio (MAE).

y proporciona una medida de la magnitud del error promedio sin considerar su dirección. Por otro lado, el Error Cuadrático Medio (MSE) penaliza los errores grandes al elevarlos al cuadrado:

$$MSE = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2$$

Ecuación 8. Fórmula del error cuadrático medio (MSE).

lo que lo convierte en una métrica más sensible a desviaciones significativas. Finalmente, la correlación de Pearson (r) se usa para evaluar la relación lineal entre las señales reales y modeladas, dada por:

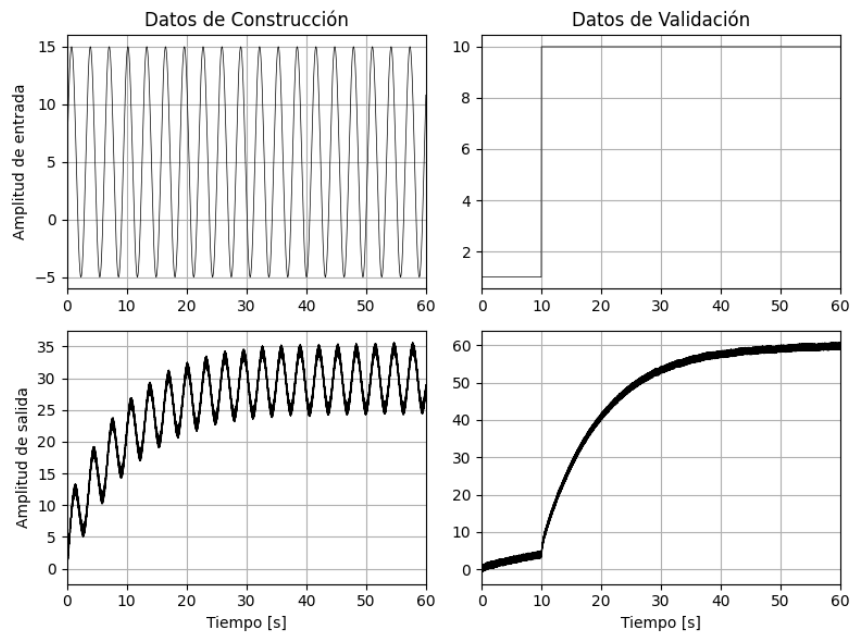
$$r = \frac{\sum_{i=1}^N (y_i - \bar{y})(\hat{y}_i - \bar{\hat{y}})}{\sqrt{\sum_{i=1}^N (y_i - \bar{y})^2 \sum_{i=1}^N (\hat{y}_i - \bar{\hat{y}})^2}}$$

Ecuación 9. Fórmula de la correlación de Pearson

donde \bar{y} y $\bar{\hat{y}}$ son los promedios de las señales real y estimada respectivamente. Esta métrica varía entre -1 y 1, indicando el grado de ajuste lineal entre ambas señales. En conjunto, estas herramientas permiten no solo construir modelos ajustados a los datos, sino también evaluar su fidelidad en la representación del sistema original.

2. RESULTADOS Y ANÁLISIS

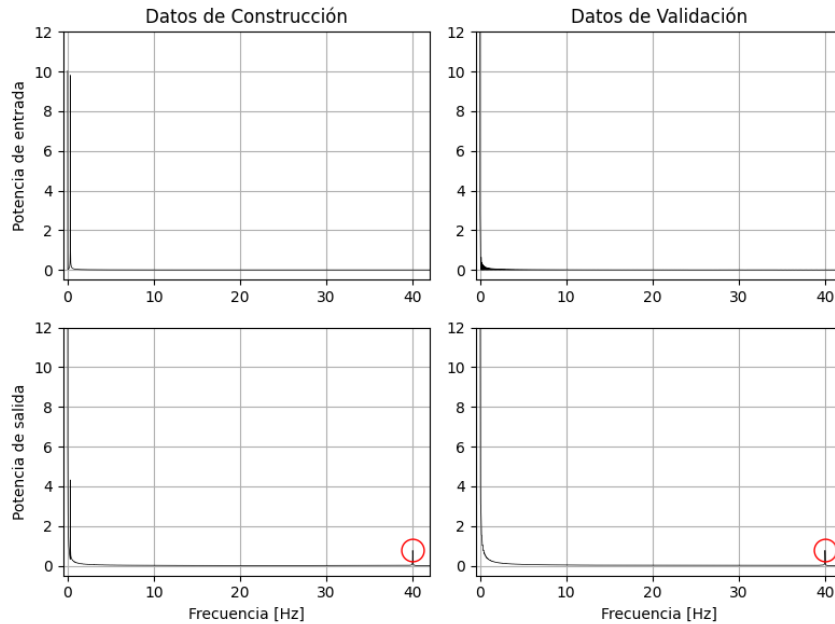
El proceso de identificación del sistema comenzó con una caracterización inicial de los datos, tanto en el dominio temporal como en el frecuencial. En la primera etapa, se graficaron las señales de entrada y salida correspondientes a los conjuntos de construcción y validación (ver Gráfica 1). En el conjunto de construcción, se observó una entrada compuesta por una señal periódica que varía con el tiempo, mientras que la salida refleja una respuesta oscilatoria creciente, que parece alcanzar una meseta hacia los 50 segundos. Esta oscilación sugiere que el sistema no es puramente estático, sino que posee una dinámica resonante o una respuesta armónica inducida por la entrada. En el conjunto de validación, la señal de entrada cambia radicalmente: se trata de un escalón que introduce un cambio abrupto a los 10 segundos. La salida, por su parte, muestra un crecimiento exponencial típico de un sistema de primer orden sin sobresalto significativo. En ambas salidas, puede observarse un leve engrosamiento de las curvas, lo que sugiere la presencia de ruido de alta frecuencia que afecta el comportamiento observado.



Gráfica 1. Caracterización visual de los datos.

Para verificar esta hipótesis, se realizó una caracterización frecuencial a través del análisis del espectro de potencia de las señales, mediante la Transformada Rápida de Fourier (FFT). La Gráfica 2 muestra los resultados de este análisis, donde se pueden comparar las componentes espectrales de entrada y salida, tanto para construcción como para validación. Las señales de entrada están dominadas por componentes de baja frecuencia, coherentes con su naturaleza periódica o escalonada. Sin embargo, en ambas salidas se detecta un componente adicional alrededor de los 40 Hz, marcado en las gráficas mediante círculos rojos. Este pico, aunque de baja amplitud, confirma la presencia de ruido de alta frecuencia que no está presente en las entradas, y por tanto debe atribuirse a características propias del sistema o del entorno de medición. Esta observación es

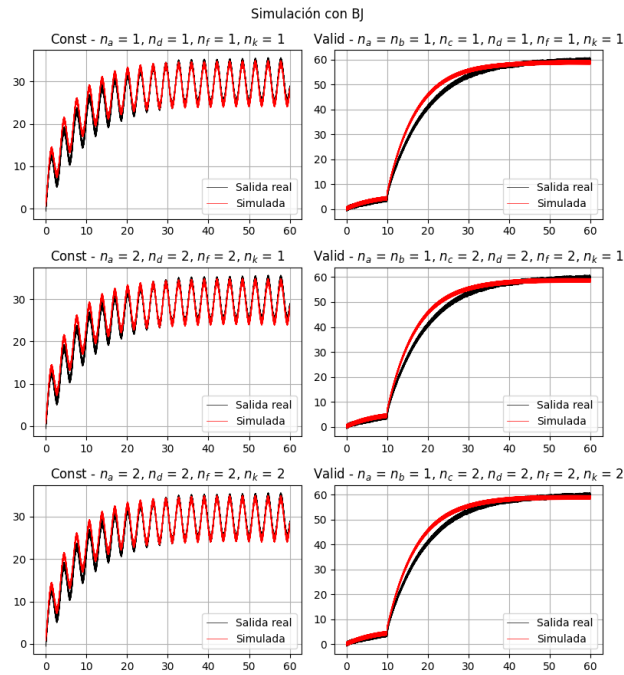
clave para la elección de estructuras de modelos que no solo consideren entradas y salidas, sino que también permitan representar la dinámica del ruido, como ARMAX, OE y Box-Jenkins



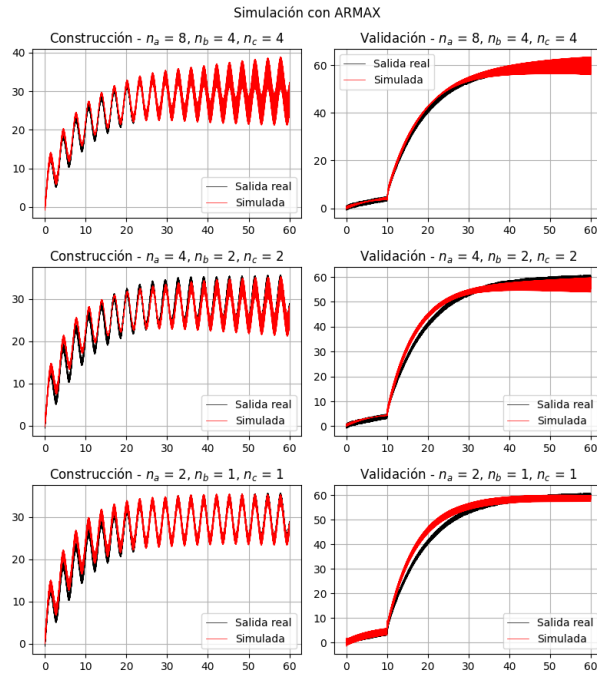
Gráfica 2. Caracterización frecuencial de los datos.

Teniendo en cuenta esta caracterización, se evaluaron cuatro estructuras paramétricas: Box-Jenkins (BJ), ARMAX, OE y ARX. Estas estructuras fueron seleccionadas estratégicamente para cubrir distintos niveles de complejidad y capacidad de modelado. Box-Jenkins fue elegida por ser la más completa, permitiendo modelar tanto la dinámica del sistema como la del ruido de manera diferenciada mediante los polinomios $B(q)$, $F(q)$, $C(q)$ y $D(q)$. ARMAX permite una modelación conjunta de entrada, salida y ruido, aunque compartiendo algunos polinomios. OE, por su parte, se centra exclusivamente en modelar la dinámica del sistema, asumiendo que el ruido es blanco, lo que la hace computacionalmente eficiente. Finalmente, ARX representa el caso más simple, considerando solo entrada y salida, y asumiendo ruido blanco no estructurado.

Las simulaciones con cada estructura se presentan en las Gráficas 3 a 6. En la Gráfica 3, correspondiente a los modelos BJ, se observa un ajuste visualmente excelente en las tres configuraciones evaluadas. Las salidas simuladas (en rojo) siguen muy de cerca a las reales (en negro), tanto en construcción como en validación, incluso en presencia de oscilaciones o rampas. El modelo con $n_a = n_b = 1$ y retardo $n_k = 2$ muestra una ligera mejora en la fase de validación, lo que indica una buena generalización. En la Gráfica 4, correspondiente a ARMAX, el modelo de mayor orden sufre problemas numéricos severos. Sin embargo, el modelo con menor orden (modelo 3) muestra un ajuste aceptable, con una respuesta suavizada y estable.



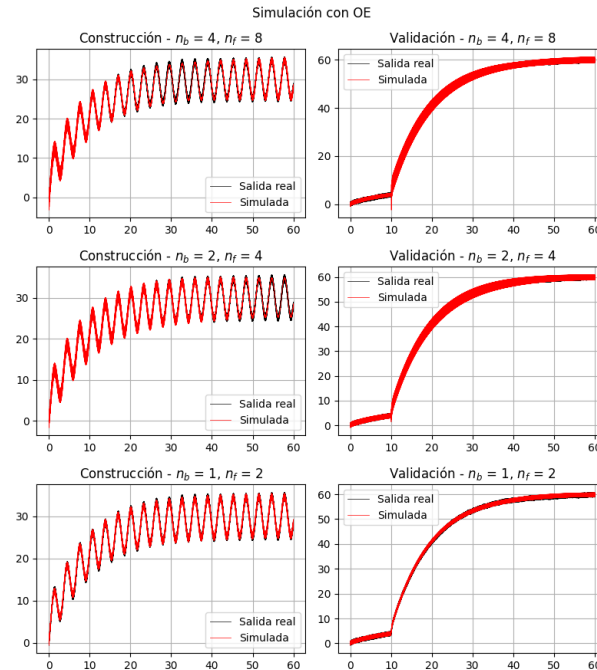
Gráfica 3. Evaluación de los datos con el modelo Box-Jenkins.



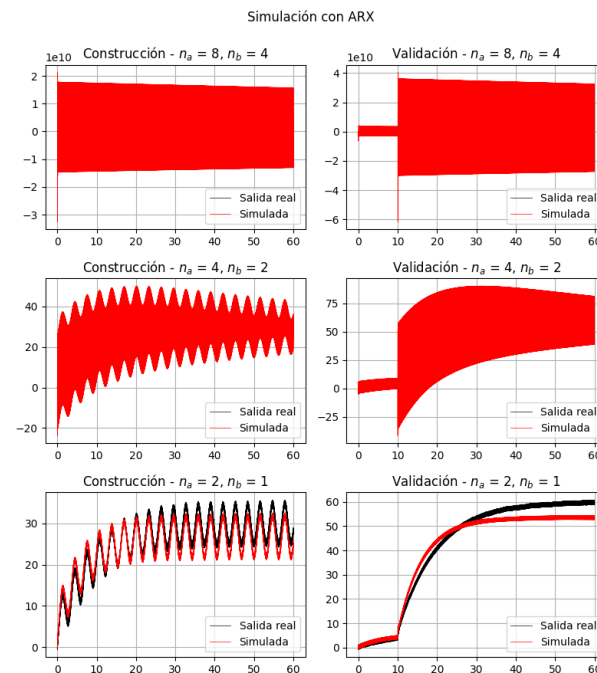
Gráfica 4. Evaluación de los datos con el modelo ARMAX.

La Gráfica 5, que representa los resultados de la estructura OE, muestra un desempeño destacable. Incluso con modelos de orden muy bajo (como la configuración 3 con $n_b = 1, n_f = 2$), se logra un ajuste muy cercano al real tanto en construcción como en validación. Este comportamiento confirma que el sistema puede modelarse correctamente sin necesidad de

incorporar un modelo explícito del ruido, siempre que este no domine la dinámica. Finalmente, en la Gráfica 6 se presentan los resultados de los modelos ARX. Como era de esperarse, esta estructura fue la más limitada. El modelo de orden más alto generó resultados completamente erráticos e inestables. El modelo de orden medio tuvo un ajuste visual muy pobre, y solo el modelo de orden más bajo (modelo 3) logró acercarse levemente a la señal real, aunque con errores notables.



Gráfica 5. Evaluación de los datos con el modelo OE.



Gráfica 6. Evaluación de los datos con el modelo ARX.

Para cuantificar estos comportamientos, se calcularon tres métricas clásicas de ajuste: el error cuadrático medio (MSE), que penaliza grandes errores y es útil para reflejar el ajuste global; el error absoluto medio (MAE), menos sensible a valores extremos y más representativo del error promedio; y la correlación de Pearson, que mide la similitud en la forma de las señales, independientemente de su escala. Estas métricas se calcularon tanto para el conjunto de construcción como para el de validación, y posteriormente se combinaron ponderadamente para seleccionar el modelo óptimo según cada criterio.

La Tabla 1 resume los resultados obtenidos. Se puede observar que la estructura OE – Configuración 3 (con $n_b = 1, n_f = 2$) obtuvo el mejor desempeño promedio en dos de los tres criterios evaluados: fue el modelo con menor MSE combinado (1.11) y menor MAE combinado (0.91). Además, logró una correlación de 0.9935, muy cercana al mejor valor registrado. El modelo OE también destaca por su simplicidad estructural, lo cual reduce el riesgo de sobreajuste y facilita la implementación computacional.

Modelo	Configuración	MSE Promedio	MAE Promedio	Correlación Promedio
Box-Jenkins (BJ)	$n_b = 1$	4.62	1.58	0.9867
	$n_c = 1$			
	$n_d = 1$			
	$n_f = 1$			
	$n_k = 1$			
ARMAX	$n_b = 1$	4.39	1.52	0.9875
	$n_c = 2$			
	$n_d = 2$			
	$n_f = 1$			
	$n_k = 1$			
ARMAX	$n_b = 1$	4.24	1.48	0.9883
	$n_c = 2$			
	$n_d = 2$			
	$n_f = 1$			
	$n_k = 2$			
ARMAX	$n_a = 8$	3.31	1.52	0.9781
	$n_b = 4$			
	$n_c = 4$			
	$n_k = 1$			
	$n_a = 4$	4.93	1.73	0.9821
ARMAX	$n_b = 2$			
	$n_c = 2$			
	$n_k = 1$			
	$n_a = 2$	4.01	1.43	0.9899
ARMAX	$n_b = 1$			
	$n_c = 1$			
	$n_k = 1$			
	$n_b = 4$	1.42	0.92	0.9943
Output error (OE)	$n_f = 8$			
	$n_k = 1$			
	$n_b = 2$	3.04	1.50	0.9860
	$n_f = 4$			
	$n_k = 1$			
Output error (OE)	$n_b = 1$	1.11	0.91	0.9935
	$n_f = 2$			
	$n_k = 1$			

ARX	$n_a = 8$	3.17×10^{20}	1.49×10^{14}	-0.0622
	$n_b = 4$			
	$n_a = 4$	361.76	15.38	0.5321
	$n_b = 2$			
	$n_a = 2$	11.64	2.93	0.9733
	$n_b = 1$			

Tabla 1. Comparación de métricas entre modelos.

En función de los resultados obtenidos, se seleccionó como modelo final la estructura OE – Configuración 3, por combinar alto rendimiento cuantitativo, simplicidad estructural y estabilidad visual en ambas fases del análisis. Esta elección se justifica además por la caracterización espectral previa, que indicó la presencia de ruido acotado y de alta frecuencia, pero no dominante, lo cual permite que un modelo sin dinámica de ruido, como OE, sea suficiente para representar el sistema de forma precisa. La combinación de análisis gráfico, frecuencial, cuantitativo y estructural permitió llegar a una decisión bien fundamentada sobre el modelo óptimo.

Para el modelo seleccionado, los polinomios identificados fueron:

- $B(q) = 0.02384 q^{-1} - 0.022343 q^{-2}$
- $F(q) = 1 - 1.9336 q^{-1} - 0.9337 q^{-2}$

Reemplazando estos polinomios en la expresión general del modelo OE, se obtiene la ecuación en el dominio Z:

$$Y(z) = \frac{0.02384 z^{-1} - 0.022343 z^{-2}}{1 - 1.9336 z^{-1} - 0.9337 z^{-2}} \cdot U(z)$$

Ecuación 10. Función de transferencia en el dominio Z

Esta expresión representa la función de transferencia discreta del sistema identificado, y describe cómo la entrada $U(z)$ es transformada en la salida $Y(z)$ a través de una relación causal y lineal en el dominio Z. El término $e(n)$ no se incluye en la ecuación de la función de transferencia ya que se asume como ruido blanco no estructurado, tratado como residuo del ajuste.

Esta forma es fundamental para análisis en frecuencia, simulación, implementación en controladores digitales, y validación de modelos identificados mediante simulación directa o por filtros recursivos.

3. CONCLUSIONES

A partir del proceso de identificación realizado, se concluye que la caracterización previa del sistema, tanto en el dominio temporal como en el frecuencial, es una etapa crítica para orientar la selección de estructuras adecuadas. La presencia de componentes armónicas y ruido de alta frecuencia detectados en la salida del sistema justificó el uso de modelos capaces de representar dinámicas complejas y perturbaciones estructuradas. Esta estrategia permitió anticipar que modelos como ARX, que no consideran la dinámica del ruido, tendrían limitaciones importantes en su capacidad de generalización.

En segundo lugar, el análisis comparativo entre estructuras evidenció que una mayor complejidad en el modelo no necesariamente garantiza un mejor ajuste. Si bien las estructuras como Box-Jenkins y ARMAX demostraron capacidad

para seguir la señal de salida, su implementación fue más costosa y, en algunos casos, numéricamente inestable. En contraste, la estructura OE logró igualar e incluso superar su desempeño con una configuración mínima, mostrando que, ante ruido blanco o poco estructurado, una representación simple pero bien parametrizada puede ser más eficaz y eficiente.

Finalmente, el uso combinado de métricas cuantitativas como MSE, MAE y correlación de Pearson permitió una evaluación integral del ajuste, capturando tanto el error promedio como la fidelidad en la forma de la señal. Esta evaluación múltiple fue determinante para seleccionar el modelo final de forma objetiva, más allá de la inspección visual. En conjunto, estos resultados validan la importancia de un enfoque sistemático y multifactorial en la identificación de sistemas, donde la decisión sobre el mejor modelo surge del equilibrio entre precisión, estabilidad y simplicidad estructural.

4. REFERENCIAS

- [1] L. Ljung, *System Identification: Theory for the User*, 2nd ed. Upper Saddle River, NJ, USA: Prentice-Hall PTR, 1999.
- [2] G. E. P. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung, *Time Series Analysis: Forecasting and Control*, 5th ed. Hoboken, NJ, USA: John Wiley & Sons, 2015.
- [3] T. Söderström and P. Stoica, *System identification*. London, UK: Prentice Hall, 1989.