Jornadas de Automática

Comparativa de metodologías de análisis de identificabilidad estructural y observabilidad para modelos no lineales

Ionescu Leoca, Dorin Alexandru[a], Shams Falavarjani, Mahmoud[a], Villaverde, Alejandro F.[a,b,*]

[a] Universidade de Vigo, Departmento de Ingeniería de Sistemas y Automática, 36310 Vigo, Galicia, España
[b] CITMAga, 15782 Santiago de Compostela, Galicia, España

To cite this article: Ionescu Leoca, D. A., Shams Falavarjani, M., Villaverde, A. F. 2025. Comparison of methodologies for structural identifiability and observability analysis of nonlinear models. Jornadas de Automática, 46. https://doi.org/10.17979/ja-cea.2025.46.12168

Resumen

La identificabilidad y la observabilidad estructurales son propiedades que describen la posibilidad teórica de inferir los parámetros y el estado de un sistema dinámico a partir de las mediciones disponibles. En este trabajo consideramos tres metodologías para el análisis local de observabilidad e identificabilidad estructural de modelos no lineales. Una de ellas, la Condición de Rango de Observabilidad, se basa en conceptos de geometría diferencial y es de naturaleza simbólica. Las otras dos metodologías se basan en cálculo numérico y de trayectorias para obtener Gramianos de observabilidad empíricos o matrices de sensibilidades que permiten establecer la observabilidad e identificabilidad estructurales bajo ciertas condiciones. Los tres métodos cuentan con implementaciones en MATLAB. Mediante la aplicación de estas herramientas a un conjunto amplio de casos de estudio, determinamos sus ventajas e inconvenientes y proporcionamos recomendaciones para su uso.

Palabras clave: Modelado e identificación, Observabilidad, Identificabilidad, Gramianos, Geometría diferencial.

Comparison of methodologies for structural identifiability and observability analysis of nonlinear models

Abstract

Structural identifiability and observability are properties that describe the theoretical possibility of inferring the parameters and state of a dynamical system from the available measurements. In this work, we consider three methodologies for analyzing the local structural observability and identifiability of nonlinear models. One of them, the Observability Rank Condition, is based on differential geometry and is symbolical by nature. The other two methodologies are based on numerical and trajectory calculations to obtain empirical observability Gramians or sensitivity matrices, which can determine structural identifiability and observability under certain conditions. All three methods have implementations in MATLAB. By applying these tools to a broad set of case studies, we determine their advantages and disadvantages and provide recommendations for their use.

Keywords: Identification and modelling, Observability, Identifiability, Gramians, Differential geometry.

1. Introducción

La identificabilidad estructural y la observabilidad son propiedades que describen la posibilidad teórica de inferir, respectivamente, los parámetros y el estado de un sistema dinámico a partir del conocimiento de su entrada y su salida. La identificabilidad estructural es condición necesaria (no suficiente) para calibrar un modelo, es decir, para estimar sus parámetros. Análogamente, la observabilidad es condición necesaria para el diseño de observadores de estado y, por tanto, para la monitorización en línea del estado de un sistema.

En sistemas lineales, el análisis de observabilidad se realiza mediante el cálculo del rango de la matriz de observabilidad, la cual se obtiene fácilmente mediante la multiplicación de matrices que son independientes del estado del sistema

* Autor para correspondencia: afvillaverde@uvigo.gal

(Kalman, 1960). En sistemas no lineales, la observabilidad e identificabilidad de un sistema dependen del punto de operación y se puede determinar también a partir del rango de una matriz de observabilidad-identificabilidad, pero su cálculo es más complejo ya que la construcción de dicha matriz requiere del cálculo de derivadas de Lie (Hermann and Krener, 1977). La observabilidad y la identificabilidad estructural están estrechamente relacionadas; de hecho, la segunda propiedad puede reducirse a la primera si consideramos los parámetros de un modelo como variables de estado cuya derivada es siempre nula (Tunali and Tarn, 1987). De esta forma, es posible analizar ambas propiedades de forma conjunta.

En la última década se han desarrollado varias metodologías para llevar a cabo estos análisis, acompañadas de implementaciones públicamente disponibles. Un enfoque, ya mencionado, se basa en la construcción de una matriz de observabilidad e identificabilidad y en el cálculo de su rango. Este test es la llamada Condición de Rango de Observabilidad (ORC en sus siglas en inglés), para cuya comprobación se utilizan conceptos de geometría diferencial como las derivadas de Lie. Esta metodología se encuentra implementada en la toolbox de MATLAB STRIKE-GOLDD, presentada originalmente por Villaverde et al. (2016); una versión computacionalmente más eficiente fue desarrollada por Díaz-Seoane et al. (2023).

La evaluación de la ORC puede ser computacionalmente costosa para modelos complejos, especialmente si el número de parámetros y variables desconocidas supera ampliamente al de salidas. Como alternativa existen métodos basados en la integración numérica. Uno de ellos es el basado en Gramianos empíricos, que son una extensión de los Gramianos clásicos para sistemas no lineales. Su cálculo en base a trayectorias fue descrito por Lall et al. (1999), con miras a su uso para reducción de modelos. El paquete emgr desarrollado por Himpe and Ohlberger (2013); Himpe (2018, 2023) implementa esta metodología. Otra metodología existente es la propuesta por Stigter and Molenaar (2015), inicialmente para análisis de identificabilidad estructural, y extendida posteriormente para observabilidad y controlabilidad (Van Willigenburg et al., 2022). Se basa en el cálculo de una matriz de sensibilidad y en su descomposición en valores singulares. Existe una implementación como toolbox de MATLAB llamada StrucID (Stigter and Joubert, 2021).

El término *estructural* hace referencia a que las propiedades son determinadas totalmente por las ecuaciones del modelo, en contraste con la identificabilidad y observabilidad *prácticas*, en las cuales influyen también la cantidad y calidad de las medidas experimentales disponibles. Así, un parámetro identificable estructuralmente puede no serlo desde el punto de vista práctico, pero no a la inversa: las propiedades estructurales son condición necesaria pero no suficiente para sus versiones prácticas correspondientes. Por otra parte, cabe resaltar que todos los métodos mencionados analizan la identificabilidad y observabilidad estructurales de forma *local*, es decir, determinan la posibilidad de distinguir el valor de un parámetro o variable de estado de otros valores próximos, pero no necesariamente de otros valores alejados (para lo cual habría que realizar análisis *globales*). En lo sucesivo tenderemos a omitir el adjetivo "local" (y en ocasiones también "estructural"), por no recargar excesivamente el texto, sobreentendiéndose que las propiedades analizadas tienen dicho carácter.

Algunos resultados y observaciones sugieren que el método basado en ORC utiliza cálculo simbólico y en principio proporciona resultados exactos, pero puede resultar costoso computacionalmente. En contraste, los métodos basados en Gramianos y en sensibilidades son más rápidos, pero por su carácter numérico podrían dar lugar a resultados engañosos. Sin embargo, el balance deseable entre rigor y eficiencia computacional de las diferentes herramientas no ha sido evaluado de forma sistemática. Algunos resultados parciales en este sentido han sido publicados: Wang and Qi (2024) presentaron una metodología basada en diferenciación automática que compararon de forma limitada – usando solo dos modelos – con STRIKE-GOLDD y StrucID. Por otro lado, Rey Barreiro and Villaverde (2023) llevaron a cabo una amplia comparativa de herramientas de análisis de identificabilidad estructural, pero se centraron en métodos de cálculo simbólico. Así mismo, dos artículos publicados recientemente (durante la escritura del presente trabajo) comparan StrucID con: Gramianos empíricos, centrándose especialmente en la reducción de modelos Stigter and van Willigenburg (2025); y con STRIKE-GOLDD y otros métodos, centrándose en el análisis de identificabilidad estructural Heinrich et al. (2025). El presente trabajo complementa dichos resultados mediante un estudio comparativo de las tres herramientas, tanto a nivel teórico como práctico, usando para ello un conjunto de modelos, y prestando particular atención a su capacidad de producir resultados exactos. Para ello, en la Sección 2 se describen los fundamentos teóricos de cada método, en la Sección 3 se analizan los resultados de su aplicación a un conjunto de 13 modelos, y en la Sección 4 se proporcionan algunas conclusiones.

2. Metodologías

2.1. *Formalismo de modelado y definiciones*

Consideramos modelos de ecuaciones diferenciales ordinarias, de la forma:

$$M : \begin{cases} \dot{x}(t) = f(x(t), u(t), \theta), \\ y(t) = g(x(t), u(t), \theta), \\ x_0 = x(t_0, \theta) \end{cases} \tag{1}$$

donde f y g son funciones analíticas, $\theta \in \mathbb{R}^q$ es el vector de parámetros, $u(t) \in \mathbb{R}^r$ el de entradas, $x(t) \in \mathbb{R}^n$ el de variables de estado e $y(t) \in \mathbb{R}^m$ el de salidas.

Conceptualmente, un parámetro $\theta_i \in \theta$ es *estructuralmente identificable localmente* si, para casi cualquier vector $\theta^* \in \mathbb{R}^q$ (es decir, con la posible excepción de un conjunto de medida cero), existe un entorno $\mathcal{N}(\theta^*)$ tal que, para todo $\hat{\theta} \in \mathcal{N}(\theta^*)$, $y(t, \theta^*) = y(t, \hat{\theta}) \iff \theta_i^* = \hat{\theta}_i$ (Walter and Pronzato, 1997; Chis et al., 2011). Si esta relación no se cumple en ningún $\mathcal{N}(\theta^*)$, θ_i es estructuralmente no identificable. En este trabajo emplearemos el término "identificabilidad" para referirnos a identificabilidad estructural local.

Una variable de estado $x_i(\tau)$ es *observable* si es posible determinarla a partir de la salida $y(t)$ y la entrada $u(t)$ en un intervalo finito $t_0 \leq \tau \leq t \leq t_f$ (Chatzis et al., 2015). La observabilidad de una variable de estado equivale a la identificabilidad de su condición inicial, si consideramos ésta como un parámetro.

Decimos que un modelo es identificable (respectivamente, observable) si todos sus parámetros (resp., variables de estado) lo son. En caso contrario, decimos que es no identificable (resp., no observable) (Wieland et al., 2021).

2.2. Gramianos empíricos (emgr)

Los Gramianos para sistemas lineales se calculan resolviendo ecuaciones matriciales. Los Gramianos empíricos son extensiones de los sistemas clásicos de Gramianos basadas en datos, y no dependen de una estructura lineal. La idea central detrás de los Gramianos empíricos es representar una media sobre los Gramianos locales cerca del punto de operación, obtenidos de simulaciones numéricas en localizaciones de una región cercana al estado estable. Una región de operabilidad se define en este contexto por una serie de perturbaciones para las entradas y el estado estable. Dentro del entorno de los Gramianos empíricos, las rotaciones están limitadas a un conjunto de matrices unidad tanto positivas como negativas. Esta limitación de la rotación reduce el conjunto de perturbaciones a direcciones y escalas. Con esta técnica, sólo entradas singulares y componentes de estado pueden ser perturbadas cada vez, lo cual, en general, es suficiente.

Gramiano empírico de observabilidad. El Gramiano *lineal* de observabilidad cuantifica el efecto de los cambios de estado del sistema en la salida del mismo, y se define cómo:

$$W_O := \int_0^\infty e^{A^T t} C^T C e^{At} dt = \int_0^\infty (e^{A^T t} C^T)(e^{A^T t} C^T)^T dt \tag{2}$$

La definición del Gramiano *empírico* de observabilidad necesita dos conjuntos no vacíos (E_x y S_x), unas trayectorias de salida ($y^{li}(t)$) para la configuración de estados iniciales ($\tilde{x}_0^{li}(t) = d_l \epsilon^i + \tilde{x}, u(t) = \hat{u}$) y unos desfases $\hat{u}, \tilde{x}, \tilde{y}^{li}$:

$$\begin{aligned} E_u &= \left\{ e^m \in \mathbb{R}^M : m = 1...M, e_i^m = \delta_{im} \right\} \\ S_u &= \left\{ c^k \in \mathbb{R} : k = 1...K, c_k \neq 0 \right\} \\ E_x &= \left\{ \epsilon^j \in \mathbb{R}^N : j = 1...N, \epsilon_i^j = \delta_{ij} \right\} \\ S_x &= \left\{ d^l \in \mathbb{R} : l = 1...L, d_l \neq 0 \right\} \end{aligned} \tag{3}$$

$$\begin{aligned} \tilde{W}_0 &= \tfrac{1}{|S_x|} \textstyle\sum_{k=1}^{|S_x|} \tfrac{1}{d_l^2} \int_0^\infty \Psi^l(t) dt \\ \Psi_{ij}^l(t) &= (y^{li}(t) - \tilde{y}^{li})^T (y^{lj}(t) - \tilde{y}^{lj}) \in \mathbb{R} \end{aligned} \tag{4}$$

Gramiano empírico de identificabilidad. Para identificación de parámetros en base a observabilidad, los parámetros se toman como estados adicionales del sistema, llevándonos así a un sistema aumentado en el que a los estados del sistema (x) se les suman los parámetros (θ), que asumimos constantes en el tiempo (por los que su campo de vectores asociado es cero):

$$\begin{gathered} \begin{pmatrix} \dot{x}(t) \\ \dot{\theta}(t) \end{pmatrix} = \begin{pmatrix} f(t, x(t), u(t), \theta(t)) \\ 0 \end{pmatrix} \\ y(t) = g(t, x(t), u(t), \theta(t)) \\ \begin{pmatrix} x(0) \\ \theta(0) \end{pmatrix} = \begin{pmatrix} x_0 \\ \theta \end{pmatrix} \end{gathered} \tag{5}$$

El Gramiano de observabilidad de este sistema aumentado será una matriz del tipo:

$$\tilde{W}_0 = \begin{pmatrix} W_O & W_M \\ W_M^T & W_P \end{pmatrix} \tag{6}$$

Con el Gramiano de observabilidad estado-espacio W_O, el de observabilidad parámetro-espacio W_I y el que junta estados y parámetros $W_M = W_M^T$. Para aislar la información de la identificabilidad, eliminamos la parte del estado-espacio.

La definición del Gramiano **empírico** de identificabilidad nace del complemento Schur del Gramiano de observabilidad empírico aumentado para el bloque inferior derecho:

$$W_I = W_P - W_M^T W_O^{-1} W_M \tag{7}$$

En general, es suficiente aproximar el Gramiano empírico de observabilidad por la parte inferior derecha del Gramiano aumentado de observabilidad, es decir:

$$W_I \approx W_P \tag{8}$$

2.3. Matriz de sensibilidades (StrucID)

Se trata de un método que analiza la falta de identificabilidad, observabilidad o controlabilidad a través del estudio del kernel del Gramiano en sistemas no lineales. StrucID es una aplicación en MATLAB que requiere un modelo definido en un espacio de estados como entrada, así como una ecuación de salida Stigter and Joubert (2021). El funcionamiento del algoritmo es el siguiente: Se calcula el rango de una matriz de sensibilidad con SVD, produciendo *"graph signatures"* (firmas de grafo), que son un conjunto de patrones presentes en una secuencia de grafos Balasundaram et al. (2022). Eso es útil para grafos variables en el tiempo o para buscar características comunes a un conjunto que a priori no parecería tener ninguna relación. Las firmas de grafo representan los valores singulares y vectores de kernel y proporcionan un resumen visual de la información.

StrucID se basa en una combinación de métodos numéricos y algebraicos, diferenciándose así de STRIKE-GOLDD, que solamente usa métodos simbólicos. Se calcula el kernel del Gramiano asociado, siendo este un cálculo más rápido al tratarse de un método numérico. De forma resumida, StrucID obtiene resultados numéricos en cuánto a identificabilidad y observabilidad, con lo que después continua el análisis con un cálculo simbólico. El número de cálculos necesario se reduce usando SVD en la matriz de sensibilidad de los parámetros de salida. La segunda parte del análisis comprueba los resultados numéricos y permite caracterizar la relación entre parámetros correlacionados y la reparametrización asociada del modelo.

Se ha establecido en otros trabajos Krener and Ide (2009) que existe la posibilidad de establecer una medida para el grado de observabilidad a través del conjunto de derivadas (mapa tangente) siguiente:

$$\begin{aligned} \tfrac{d\delta x(t)}{dt} &= \tfrac{\partial f}{\partial x} \delta x \\ \delta y(t) &= \tfrac{\partial h}{\partial x} \delta x \end{aligned} \tag{9}$$

en el que los Jacobianos serán evaluados en trayectorias que son soluciones de (1). Los valores singulares del mapa no lineal (1) serán definidos por los valores singulares del mapa tangente. Los valores singulares locales caracterizarán el Gramiano de observabilidad $P(x^0)$, que puede ser calculado en base a la solución fundamental $\Phi(t)$:

$$\begin{gathered} \tfrac{d\Phi(t)}{dt} = \tfrac{\partial f}{\partial x} \Phi(t) \\ \Phi(0) = I \end{gathered} \tag{10}$$

y se define el Gramiano de observabilidad como:

$$P(x^0) = \int_0^{t_f} \Phi^T(\tau)\frac{\partial h^T(\tau)}{\partial x}\frac{\partial h(\tau)}{\partial x}\Phi(\tau)d\tau \tag{11}$$

El punto de partida es el uso de las ecuaciones de sensibilidad bien conocidas. Con esto, buscaremos los valores singulares del Gramiano de observabilidad exactamente nulos, por lo que no aproximamos el Gramiano con una versión empírica del mismo. Hay que tener en cuenta que trabajamos con trayectorias conocidas, y que la elección de dicha trayectoria puede influenciar la salida del análisis.

Para encontrar un valor correcto de los valores singulares del Gramiano, es mejor aplicar una descomposición en valores singulares (SVD) a la matriz $\frac{\partial h}{\partial x}\Phi(t)$ que al Gramiano de observabilidad en sí. Esto viene de comprobar las dependencias lineales entre funciones del sistema variante en el tiempo (10) en un intervalo de tiempo $[0,t_f]$.

Por último, se ha demostrado que los valores singulares cero del Gramiano de observabilidad pueden ser usados para realizar un cálculo simbólico eficiente que permite una reparametrización del modelo (1) Joubert et al. (2020). Gracias a que los valores no identificables vienen en grupos, podemos concentrar el cálculo numérico en los parámetros no identificables para cada grupo de forma separada, lo cual ahorra enormemente recursos y se obtiene gracias al análisis SVD de la matriz de sensibilidad de salida.

Para la solución paramétrica $y_\theta(t)$, se usarán derivadas complejas para un cálculo preciso de los Jacobianos $\frac{\partial f}{\partial x}$, $\frac{\partial h}{\partial x}$, $\frac{\partial f}{\partial \theta}$ y $\frac{\partial h}{\partial \theta}$

$$\begin{aligned} \frac{dx_\theta(t)}{dt} &= \frac{\partial f}{\partial x}x_\theta + \frac{\partial f}{\partial \theta} \\ y_\theta(t) &= \frac{\partial h}{\partial x}x_\theta + \frac{\partial h}{\partial \theta} \end{aligned} \tag{12}$$

En esta ecuación, hemos representado $\frac{dx}{d\theta}$ como $x_\theta(t)$.

2.4. Condición del Rango de Observabilidad (STRIKE-GOLDD)

Como hemos dicho, la observabilidad puede ser estudiada mediante el rango de la matriz de observabilidad, que representa un relación entre la salida del modelo (y) y los estados (x). Si esta matriz tiene rango completo, el modelo será observable. Crearemos esta matriz a través de derivadas parciales de las salidas con respecto a los estados. Siguiendo esta definición, la matriz de observabilidad se construye como:

$$O(x(t)) = \begin{pmatrix} \frac{\partial}{\partial x}y(t) \\ \frac{\partial}{\partial x}\dot{y} \\ \frac{\partial}{\partial x}\ddot{y} \\ . \\ . \\ . \\ \frac{\partial}{\partial x}y^{n-1} \end{pmatrix} \tag{13}$$

Ahora, si el sistema M dado por (1) cumple que el rango de la matriz $O(x_0)$ sea igual a n, entonces dicho sistema será observable en las proximidades de x_0. Si esto se cumple, diremos que se ha cumplido la condición de observabilidad.

Con el objetivo de estudiar la identificabilidad estructural, consideraremos los parámetros p_i como estados adicionales cuyas derivadas son cero, transformado el vector de estados ($\tilde{x} = \tilde{x}(x, p)$), creando así una matriz de observabilidad-identificabilidad de la forma:

$$OI(x(t)) = \begin{pmatrix} \frac{\partial}{\partial \tilde{x}}y \\ \frac{\partial}{\partial \tilde{x}}\dot{y} \\ \frac{\partial}{\partial \tilde{x}}\ddot{y} \\ . \\ . \\ . \\ \frac{\partial}{\partial \tilde{x}}y^{n+q-1} \end{pmatrix} \tag{14}$$

Dónde vemos que la única diferencia real es que ahora tenemos q ecuaciones adicionales, que son precisamente el número de parámetros p que hemos introducido. Ahora, el rango que debe cumplir la matriz será $n + q$ para cumplir la condición de observabilidad-identificabilidad.

A partir de (14) podemos descubrir también cuales de los parámetros son estructuralmente identificables. Cada una de las columnas contiene las derivadas parciales con respecto a uno de los nuevos estados en concreto. Si eliminamos una de las columnas y el rango de la matriz OI no cambia, estaremos ante un parámetro no identificable.

Para sistemas no lineales, OI se puede calcular a través de derivadas de Lie, que se definen como

$$\begin{aligned} L_f g(\tilde{x}) &= \frac{\partial g(\tilde{x})}{\partial \tilde{x}}f(\tilde{x}, u) \\ L_f^2 g(\tilde{x}) &= \frac{\partial L_f g(\tilde{x})}{\partial \tilde{x}}f(\tilde{x}, u) \\ &. \\ &. \\ &. \\ L_f^i g(\tilde{x}) &= \frac{\partial L_f^{i-1} g(\tilde{x})}{\partial \tilde{x}}f(\tilde{x}, u) \end{aligned} \tag{15}$$

Si todas las entradas del vector de entradas son constantes, las derivadas de Lie se corresponden con las de la salida ($y^i(t) = L_f^i g(\tilde{x})$). No obstante, para entradas variantes en el tiempo debemos extender la definición de derivada de Lie:

$$L_f g(\tilde{x}) = \frac{\partial g(\tilde{x})}{\partial \tilde{x}}f(\tilde{x}, u) + \sum_{j=0}^{j=\infty} \frac{\partial g(\tilde{x})}{\partial u^j}u^{j+1} \tag{16}$$

donde u^j y u^{j+1} representan las derivadas j-ésima y $(j + 1)$-ésima de la entrada. A continuación se definen las derivadas de mayor orden de la derivada extendida de Lie, teniendo en cuenta que el sumatorio del segundo término contendrá derivadas de la entrada hasta orden $(i-2)$, por lo cual el sumatorio estará truncado:

$$L_f^i g(\tilde{x}) = \frac{\partial L_f^{i-1} g(\tilde{x})}{\partial \tilde{x}}f(\tilde{x}, u) + \sum_{j=0}^{j=i-2} \frac{\partial L_f^{i-1} g(\tilde{x})}{\partial u^j}u^{j+1} \tag{17}$$

Introduciendo esto en la matriz OI, obtenemos:

$$OI(x(t)) = \begin{pmatrix} \frac{\partial}{\partial \tilde{x}}g(\tilde{x}) \\ \frac{\partial}{\partial \tilde{x}}(L_f g(\tilde{x})) \\ \frac{\partial}{\partial \tilde{x}}(L_f^2 g(\tilde{x})) \\ . \\ . \\ . \\ \frac{\partial}{\partial \tilde{x}}(L_f^{n+q-1} g(\tilde{x})) \end{pmatrix} \tag{18}$$

Las derivadas extendidas de Lie pueden ser usadas para caracterizar la entrada necesaria para la identificabilidad estructural del sistema, al determinar cuales de las derivadas deben ser distintas de cero. La clave está en que, si la $k-$ésima derivada es igual a cero, las derivadas de orden mayor

$((k+1),(k+2),...)$ también serán cero. Así, para una *OI* de rango completo con $\dot{u}=0$, se obtiene identificabilidad estructural con una entrada constante, mientras que si se tiene rango completo con $[\dot{u}\neq 0, \ddot{u}=0]$, se requiere como mínimo una entrada en forma de rampa.

3. Resultados

En esta sección realizamos un benchmarking de las metodologías, aplicándolas a un conjunto de 13 modelos sacados de la literatura de bioprocesos, biología de sistemas, y biología sintética. Por falta de espacio no se proporcionan las ecuaciones de dichos modelos; se pueden encontrar en las publicaciones originales (citadas en el texto), así como en las implementaciones incluidas en la carpeta 'models' de la toolbox STRIKE-GOLDD (`https://github.com/afvillaverde/strike-goldd/tree/master/STRIKE-GOLDD/models`).

3.1. *Modelo lineal de dos compartimentos*

Comenzamos por estudiar el modelo de dos compartimentos analizado por Villaverde et al. (2019), quienes concluyeron que es estructuralmente identificable si la entrada tiene al menos una derivada distinta de cero, $\dot{u}(t)\neq 0$. Para estudiar este caso especificamos una entrada tipo rampa. Las tres herramientas, STRIKE-GOLDD, StrucID y los Gramianos empíricos coinciden en que el modelo es observable e identificable.

3.2. *Modelo de un bioproceso*

Este modelo se encuentra definido en (Villaverde et al., 2019). Al igual que en el anterior, si la entrada es constante el sistema es no observable, siendo los parámetros no identificables p_4, p_5, p_6 y p_7, mientras que si usamos entradas tipo rampa para u_3 y u_4, el sistema se vuelve observable e identificable. Tanto STRIKE-GOLDD como StrucID proporcionan el resultado correcto. En cuanto a los Gramianos empíricos, en este caso el resultado que se obtiene se puede interpretar como que el sistema es identificable y observable; sin embargo, el análisis de identificabilidad es complejo puesto que los resultados no se muestran de forma clara, y queda a interpretación del usuario la decisión del rango de diferencia máxima entre dos valores singulares para la determinación de la identificabilidad de las variables.

3.3. *Ruta de señalización celular NF-κB*

Para el modelo NF-κB (Lipniacki et al., 2004) los resultados coinciden de nuevo para STRIKE-GOLDD y StrucID: el sistema es no observable, con estados no identificables x_8 y x_{15} y parámetros no identificables $k_2, c_{1c}, c_{2c}, c_{3c}$ y c_4. En cuanto a los Gramianos empíricos, dan el resultado de que el sistema es observable, siendo el análisis de identificabilidad complejo una vez más, como ya explicamos para el caso anterior, aunque mayoritariamente estando de acuerdo con los resultados obtenidos en los otros modelos.

3.4. *Ruta de señalización celular JAK-STAT*

Para este modelo, propuesto por Raia et al. (2011), todas las herramientas coinciden en que estamos ante un modelo no observable, y en que los siguientes parámetros no son identificables: $\theta_{11}, \theta_{15}, \theta_{19}, \theta_{21}$ y θ_{22}. Los Gramianos empíricos reportan, adicionalmente, que los siguientes parámetros tampoco son identificables: $\theta_4, \theta_{14}, \theta_{16}, \theta_{19}$ y, de forma menos clara, θ_3 y θ_{12}. Esta discrepancia podría interpretarse como una indicación de que estos últimos parámetros son identificables desde el punto de vista estructural, pero pobremente identificables desde el punto de vista práctico-numérico; sin embargo, sería interesante confirmar esta hipótesis con algún método alternativo.

3.5. *Modelos de biología sintética*

A continuación, pasamos a estudiar un conjunto de modelos de biología sintética que aparecen en (Pardo et al., 2025). Se trata de un total de 9 modelos no lineales.

Tres de ellos (*BioSD-I*, *BioSD-II* y *BioSD-III*) son sistemas que permiten obtener derivadas de señales. Del segundo de ellos se estudian otras dos versiones, que se diferencian de la primera por usar cinética tipo Michaelis-Menten (*BioSD-II-MM-simple*, *BioSD-II-MM-complex*). Otros dos modelos describen un mecanismo de realimentación "dicótoma": *Dichotomous Feedback (I)* y *Dichotomous Feedback (k_{ap})*. Finalmente, los dos últimos usan realimentación mediada por moléculas de pequeño ARN (*sRNA-tuned autorepressor*, *closed-loop sRNA*).

Dado que los resultados de los análisis dependen de qué variables se puedan medir – es decir, de la definición de la función de salida $y(t) = g(x(t), p)$ en la ecuación (1) – y que en este tipo de aplicaciones suelen existir varias posibilidades al respecto, consideramos al igual que en Pardo et al. (2025) todas las combinaciones de salidas posibles, hasta un total de 35 configuraciones para los 9 modelos base.

Para algunas de las configuraciones resultantes encontramos las primeras discrepancias entre STRIKE-GOLDD y StrucID, que reportan diferentes conjuntos de parámetros no identificables en varios casos. En concreto, para todos los modelos mencionados, con la excepción del *closed-loop sRNA*, hay siempre alguna configuración de medidas para la que StrucID clasifica como identificables *menos* parámetros que STRIKE-GOLDD. Lo contrario – es decir, que StrucID clasifica como identificables *más* parámetros que STRIKE-GOLDD – es mucho más raro, y solo ocurre en dos modelos (*BioSDIII* y *Dichotomous Feedback (I)*), en cada caso para una configuración de medidas y afectando a un solo parámetro. En estos últimos casos, las discrepancias se producen en modelos que tienen varios parámetros identificables y varios no identificables; StrucID detecta correlaciones entre varios parámetros y clasifica como identificable uno de ellos, lo que no siempre es correcto. Notablemente, ambas herramientas coinciden plenamente en cuanto a la observabilidad de variables de estado para los modelos considerados.

En el caso de los Gramianos empíricos, ofrecen unos resultados con semejanzas a StrucID, pero con algunas discrepancias adicionales. Estas diferencias afectan también al modelo *closed-loop sRNA*, para el cual emgr clasifica erróneamente más parámetros como no identificables que las otras herramientas.

4. Conclusiones

A lo largo de este trabajo hemos hecho una comparativa de las herramientas disponibles hoy en día para el análisis de identificabilidad estructural y observabilidad de modelos matemáticos complejos y dinámicos. Análisis previos (Rey Barreiro and Villaverde, 2023) indican que los resultados proporcionados por STRIKE-GOLDD son muy fiables, coincidiendo sistemáticamente con los de otras herramientas de análisis simbólico. Por ello, el objetivo principal de el presente trabajo era analizar si las herramientas de cálculo numérico (emgr y StrucID) podían darnos unos resultados comparables en exactitud a las de cálculo simbólico, lo que permitiría reducir los tiempos y costes computacionales del cálculo sin merma de la fiabilidad de la información proporcionada sobre el sistema. Esta ventaja computacional de StrucID ha sido recientemente descrita en Heinrich et al. (2025); por ello, en el presente trabajo nos hemos centrado en el aspecto de su exactitud.

Nuestros análisis reflejan que emgr presenta limitaciones importantes, produciendo resultados que, aunque rápidos, no son precisos ni fáciles de interpretar. En contraste, StrucID proporciona un grado de exactitud en los resultados que coincide en más de un 90 % con STRIKE-GOLDD, para los modelos considerados. Las discrepancias encontradas siguen un patrón, dado que los parámetros que StrucID considera identificables también lo son en la mayoría de los casos para STRIKE-GOLDD, con algunas excepciones (discrepancia tipo 1). Por otro lado, el que StrucID clasifique como no-identificable un parámetro que STRIKE-GOLDD considera identificable (discrepancia tipo 2) es más común. Es decir, StrucID proporciona una condición de identificabilidad con un alto grado de fiabilidad, pero no con absoluta certeza. Si bien la discrepancia tipo 2 podría indicar que StrucID está informando sobre falta de identificabilidad práctica en lugar de estructural, la discrepancia tipo 1 no tiene esta justificación, lo que sugiere la existencia de problemas numéricos en el funcionamiento de esta herramienta. El esclarecer su causa exacta queda como trabajo futuro.

Agradecimientos

Esta publicación es parte del proyecto de I+D+i PID2023-146275NB-C21, financiado por MICIU/AEI/10.13039/501100011033/ y FEDER/UE, y de la ayuda CNS2023-144886 financiada por MICIU/AEI /10.13039/501100011033 y por la Unión Europea NextGenerationEU/PRTR.

Referencias

Balasundaram, B., Borrero, J. S., Pan, H., February 2022. Graph signatures: Identification and optimization. European Journal of Operational Research 296 (3), 764–775.

Chatzis, M. N., Chatzi, E. N., Smyth, A. W., 2015. On the observability and identifiability of nonlinear structural and mechanical systems. Structural Control and Health Monitoring 22 (3), 574–593.

Chis, O.-T., Banga, J. R., Balsa-Canto, E., 2011. Structural identifiability of systems biology models: a critical comparison of methods. PloS one 6 (11), e27755.

Díaz-Seoane, S., Rey Barreiro, X., Villaverde, A. F., 2023. STRIKE-GOLDD 4.0: user-friendly, efficient analysis of structural identifiability and observability. Bioinformatics 39 (1), btac748.

Heinrich, M., Rosenblatt, M., Wieland, F.-G., Stigter, H., Timmer, J., 2025. On structural and practical identifiability: Current status and update of results. Current Opinion in Systems Biology, 100546.

Hermann, R., Krener, A., 1977. Nonlinear controllability and observability. IEEE Transactions on Automatic Control 22 (5), 728–740.

Himpe, C., 2018. emgr—the empirical gramian framework. Algorithms 11 (7), 91.

Himpe, C., 2023. emgr–empirical gramian framework version 5.99. ACM Transactions on Mathematical Software 49 (3), 1–8.

Himpe, C., Ohlberger, M., 2013. A unified software framework for empirical gramians. Journal of Mathematics 2013 (1), 365909.

Joubert, D., Stigter, J., Molenaar, J., 2020. An efficient procedure to assist in the re-parametrization of structurally unidentifiable models. Mathematical Biosciens 323, 108328.

Kalman, R. E., 1960. Contributions to the theory of optimal control. Boletín de la Sociedad Matemática Mexicana 5 (2), 102–119.

Krener, A. J., Ide, K., December 2009. Measures of unobservability. Proceedings of the 48th IEEE Conference on Decision and Control (CDC) held jointly with 2009 28th Chinese Control Conferece, 6401–6406.

Lall, S., Marsden, J. E., Glavaški, S., 1999. Empirical model reduction of controlled nonlinear systems. IFAC Proceedings Volumes 32 (2), 2598–2603.

Lipniacki, T., Paszek, P., Brasier, A. R., Luxon, B., Kimmel, M., 2004. Mathematical model of nf-κb regulatory module. Journal of theoretical biology 228 (2), 195–215.

Pardo, A., Díaz-Seoane, S., Ionescu, D. A., Papachristodoulou, A., Villaverde, A. F., 2025. Analysing control-theoretic properties of nonlinear synthetic biology circuits. In: 14th IFAC Symposium on Dynamics and Control of Process Systems, including Biosystems (DYCOPS 2025). Accepted, preprint available at arXiv:2411.05450.

Raia, V., Schilling, M., Böhm, M., Hahn, B., Kowarsch, A., Raue, A., Sticht, C., Bohl, S., Saile, M., Möller, P., et al., 2011. Dynamic mathematical modeling of il13-induced signaling in hodgkin and primary mediastinal b-cell lymphoma allows prediction of therapeutic targets. Cancer research 71 (3), 693–704.

Rey Barreiro, X., Villaverde, A. F., 2023. Benchmarking tools for a priori identifiability analysis. Bioinformatics 39 (2), btad065.

Stigter, J., Joubert, D., 2021. Computing measures of identifiability, observability, and controllability for a dynamic system model with the strucid app. IFAC-PapersOnline 54 (7), 138–143.

Stigter, J. D., Molenaar, J., 2015. A fast algorithm to assess local structural identifiability. Automatica 58, 118–124.

Stigter, J. H., van Willigenburg, L. G., 2025. Model reduction of complex dynamical systems: A sensitivity based approach. IFAC-PapersOnLine 59 (1), 457–462.

Tunali, E., Tarn, T.-J., 1987. New results for identifiability of nonlinear systems. IEEE Transactions on Automatic Control 32 (2), 146–154.

Van Willigenburg, L., Stigter, J., Molenaar, J., 2022. Sensitivity matrices as keys to local structural system properties of large-scale nonlinear systems. Nonlinear Dynamics 107 (3), 2599–2618.

Villaverde, A. F., Barreiro, A., Papachristodoulou, A., 2016. Structural identifiability of dynamic systems biology models. PLoS computational biology 12 (10), e1005153.

Villaverde, A. F., Evans, N. D., Chappell, M. J., Banga, J. R., April 2019. Input-dependent structural identifiability of nonlinear systems. IEEE Control Systems Letters 3 (2), 272–277.

Walter, E., Pronzato, L., 1997. Identification of parametric models from experimental data. Springer.

Wang, L., Qi, J., 2024. Efficient structural parameter identifiability analysis for generator dynamic models. In: 2024 56th North American Power Symposium (NAPS). IEEE, pp. 1–6.

Wieland, F.-G., Hauber, A. L., Rosenblatt, M., Tönsing, C., Timmer, J., 2021. On structural and practical identifiability. Current Opinion in Systems Biology 25, 60–69.