XX Simposio CEA de Control Inteligente
25-27 de junio de 2025, Huelva

Modelado de sistemas industriales no lineales mediante SINDyc para la construcción de gemelos digitales

González-Herbón, R.[a,*], Fuertes, J.J.[a], Alonso, S.[a], González-Mateos, G.[a], Rodríguez-Ossorio, J.R.[a], Domínguez, M.[a]

[a]*Grupo de Investigación SUPPRESS, Escuela de Ingenierías, Universidad de León, Campus de Vegazana, León, 24007, España*
https://suppress.unileon.es

To cite this article: González-Herbón, R., Fuertes, J.J., Alonso, S., González-Mateos, G., Rodríguez-Ossorio, J.R., Domínguez, M. Modeling of nonlinear industrial systems using SINDyc for developing digital twins. XX Simposio CEA de Control Inteligente, Huelva (Spain), 2025.

Resumen

Este trabajo presenta una metodología basada en *Sparse Indentification of Nonlinear Dynamics with control* (SINDyc) para el desarrollo de gemelos digitales en sistemas industriales no lineales. El estudio se centra en modelar el comportamiento dinámico de una planta piloto a través de la predicción de tres variables clave del proceso: nivel, presión y caudal. El enfoque propuesto permite obtener modelos compactos, precisos e interpretables, lo que refuerza su aplicabilidad en entornos industriales reales.

Palabras clave: SINDyc, Gemelos Digitales, Modelado de Gemelos Digitales, Modelado de sistemas industriales.

Nonlinear industrial system modelling with SINDyc to develop digital twins

Abstract

This work presents a methodology based on Sparse Identification of Nonlinear Dynamics with control (SINDyc) for the development of digital twins in nonlinear industrial systems. The study focuses on modeling the dynamic behavior of a pilot plant by predicting three key process variables: level, pressure, and flow rate. The proposed approach enables the generation of compact, accurate, and interpretable models, reinforcing its applicability in real industrial environments.

Keywords: SINDyc, Digital Twin, Digital Twin Modelling, Industrial System modelling.

1. Introducción

En los últimos años la digitalización industrial ha transformado profundamente los procesos productivos, impulsada por la integración de tecnologías avanzadas como los gemelos digitales, los cuales permiten simular y optimizar sistemas físicos en tiempo real Min et al. (2019). Estas herramientas ofrecen un gran valor al representar virtualmente procesos industriales, facilitando su análisis dinámico y mejorando la toma de decisiones González-Herbón et al. (2024). En este contexto, los gemelos digitales se han consolidado como componentes esenciales dentro del paradigma de la Industria 4.0, permitiendo una optimización continua del rendimiento Tao et al. (2019).

El desarrollo de gemelos digitales requiere modelos precisos que capturen la dinámica del sistema real. Sin embargo, las condiciones adversas en entornos industriales, como no linealidades complejas introducidas por elementos como válvulas, bombas o variadores de frecuencia, junto con el ruido que proviene de los sensores o perturbaciones externas, dificultan la obtención de modelos analíticos exactos Wagg et al. (2020). Ante este desafío, los enfoques basados en datos han ganado relevancia, ya que permiten identificar modelos sin depender exclusivamente de ecuaciones físicas detalladas Sun and Ge (2021). No obstante, para que un gemelo digital opere en sincronía con el sistema físico, el modelo subyacente debe ser capaz de ejecutarse en tiempo real, garantizando una comunicación bidireccional efectiva que permita predicción, control y optimización Birk et al. (2022).

En este trabajo, se propone el uso de SINDy (*Sparse Identification of Nonlinear Dynamics*), en concreto su variante

*Autor para correspondencia: rgonzh@unileon.es

SINDyc (*SINDy with control*), para modelar un sistema industrial no lineal con el fin de desarrollar un gemelo digital. A diferencia de métodos basados en redes neuronales, SINDy identifica ecuaciones gobernantes de forma interpretable Brunton et al. (2016a). Esta característica lo hace especialmente atractivo para la implementación de gemelos digitales de entornos industriales. Además, SINDyc permite incorporar entradas de control, lo que facilita su aplicación en sistemas donde la dinámica del sistema es una dinámica forzada por las variables de control.

El resto del artículo se estructura de la siguiente manera: en la sección 2 se describe el modelo SINDy, así como su variante SINDyc. A continuación, en la sección 3 se describe tanto el sistema como las herramientas empleados, así como los experimentos y el modelado realizado con SINDyc. En la sección 4 se presentan los resultados obtenidos por el modelo para las variables del sistema estudiado. Finalmente, se exponen las conclusiones y las líneas futuras en la sección 5.

2. Sparse Indentification of Nonlinear Dynamics (SINDy)

La identificación de modelos dinámicos con precisión es un paso esencial en el desarrollo de gemelos digitales, especialmente en entornos industriales donde los sistemas presentan comportamientos altamente no lineales y sujetos a perturbaciones. En este contexto, el algoritmo *Sparse Identification of Nonlinear Dynamics* (SINDy) se ha consolidado como una herramienta eficaz para descubrir ecuaciones gobernantes directamente a partir de datos, promoviendo la interpretación física de los modelos resultantes y evitando el carácter opaco de los enfoques de caja negra Wang et al. (2023b). SINDy ha demostrado su aplicabilidad en diversos contextos industriales, desde unidades de hidrotratamiento de diésel hasta reactores químicos y sistemas de manufactura complejos Wang et al. (2023a). Su capacidad para generar modelos simplificados lo hace idóneo para plataformas de simulación en tiempo real y aplicaciones de control predictivo.

El algoritmo se basa en una formulación de regresión dispersa en la que, dado un conjunto de estados temporales

$$\mathbf{X} = \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \vdots \\ \mathbf{x}_m \end{bmatrix}, \tag{1}$$

se busca una función $\mathbf{f}$ tal que:

$$\mathbf{x}_{k+1} = f(\mathbf{x}_k) \tag{2}$$

Para ello, se construye una biblioteca de funciones candidatas $\Theta(X)$, que incluye términos polinómicos, trigonométricos, la combinación de términos empíricos o principios físicos (como leyes de conservación de masa o relaciones termodinámicas):

$$\Theta(\mathbf{X}) = \left[\mathbf{1}, \mathbf{X}, \mathbf{X}^{Pd}, \sin(\mathbf{X}), \dots\right] \tag{3}$$

donde X^{Pd} representa todas las combinaciones de orden d entre las variables de estado. La tarea se reduce a resolver el problema de regresión:

$$\mathbf{X}_{k+1} = \Theta(\mathbf{X})\Xi \tag{4}$$

donde Ξ es una matriz de coeficientes dispersa, determinada mediante técnicas como LASSO o *sequential thresholded least squares* (SLS), que eliminan iterativamente los coeficientes más pequeños para mejorar la interpretabilidad y evitar el sobreentrenamiento.

2.1. Sparse Indentification of Nonlinear Dynamics with control (SINDyc)

En un entorno industrial, los sistemas dinámicos suelen estar sujetos a señales de control relativas a válvulas, motores, etc. Para capturar adecuadamente estas influencias, se ha desarrollado la variante *SINDy with control* (SINDyc), que extiende el algoritmo original para incorporar las señales de entrada al proceso Brunton et al. (2016b). En este caso, se considera un sistema de la forma:

$$\mathbf{x}_{k+1} = f(\mathbf{x}_k, \mathbf{u}_k), \tag{5}$$

donde $\mathbf{x}_k \in \mathbb{R}^n$ es el vector de estado y $\mathbf{u}_k \in \mathbb{R}^p$ representa las entradas de control en el instante k.

Para identificar este tipo de sistemas, la biblioteca de funciones se amplía para incluir no solo términos que dependen de los estados, sino también de las entradas y de combinaciones no lineales entre ambos:

$$\Theta(\mathbf{X}, \mathbf{U}) = \left[\mathbf{1},\ \mathbf{X},\ \mathbf{U},\ \mathbf{X}\cdot\mathbf{U},\ \mathbf{X}^{P2},\ \mathbf{U}^{P2},\ \dots\right], \tag{6}$$

donde $\mathbf{X}\cdot\mathbf{U}$ denota una colección de productos cruzados entre componentes de $\mathbf{X}$ y $\mathbf{U}$, tales como x_1u_1, $x_2u_1^2$, o $x_1^2u_2$, lo cual permite capturar interacciones no lineales relevantes para la dinámica del sistema.

A partir de esta biblioteca extendida, el modelo se identifica resolviendo un problema de regresión dispersa de la forma:

$$\mathbf{X}_{k+1} = \Theta(\mathbf{X}, \mathbf{U})\Xi, \tag{7}$$

donde Ξ es una matriz de coeficientes dispersa que se obtiene mediante técnicas como LASSO o *sequential thresholded least squares*.

2.2. Limitaciones y desafíos

A pesar de sus ventajas, el uso de SINDy y SINDyc en entornos industriales reales conlleva ciertos desafíos que deben considerarse al diseñar modelos digitales robustos:

- **Dependencia de conocimiento previo:** La efectividad de la biblioteca de funciones $\Theta(\cdot)$ depende en gran medida de la inclusión de términos relevantes y físicamente significativos. Por ejemplo, en sistemas termoquímicos, omitir expresiones como $e^{-E_a/RT}$ puede deteriorar gravemente la precisión del modelo si no se consideran los mecanismos cinéticos involucrados Wang et al. (2023b).
- **Maldición de la dimensionalidad:** El número de términos candidatos en la biblioteca crece exponencialmente con el número de variables del sistema (n) y el grado de los polinomios incluidos (d), lo que genera una complejidad de orden $O(n^d)$. Esto aumenta significativamente el coste computacional y el riesgo de sobreentrenamiento, especialmente en aplicaciones industriales donde pueden coexistir decenas de variables de entrada y salida.

3. Metodología

3.1. Sistema industrial no lineal

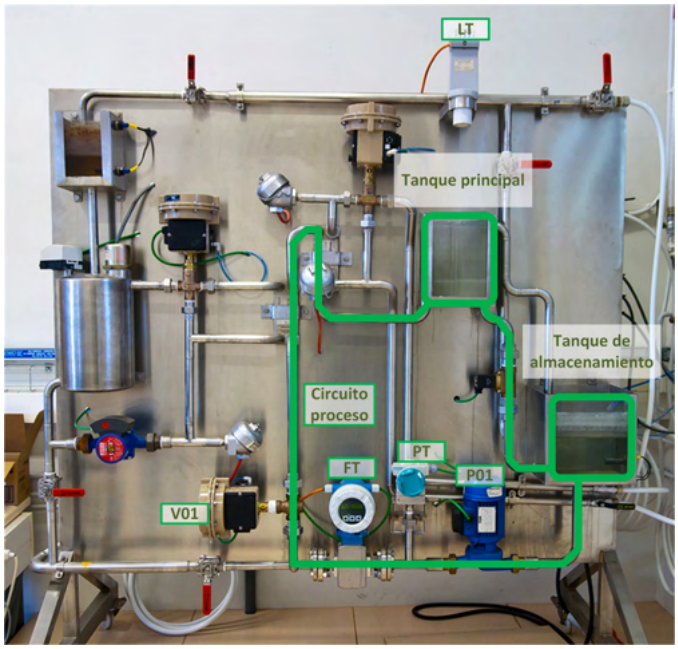

Figura 1: Sistema industrial no lineal

El sistema utilizado corresponde a una planta piloto que emula un proceso industrial en un entorno controlado, pero representativo de aplicaciones reales, como se muestra en la Figura 1. Este sistema está compuesto por tres circuitos, de los cuales se ha trabajado únicamente con el circuito principal o de proceso, representado en verde.

Este circuito consiste en la circulación de un fluido entre dos tanques situados a distinta altura (tanque principal y tanque de almacenamiento). Las variables que se han considerado en este estudio son el nivel del tanque superior, la presión del circuito y el caudal del fluido, las cuales han sido medidas mediante los sensores LT, PT y FT, respectivamente. El fluido es impulsado por una bomba centrífuga (P01) accionada mediante un variador de frecuencia. Adicionalmente, una válvula neumática (V01) regula el caudal del circuito.

3.2. Modelado mediante SINDyc

El objetivo del modelado es predecir las variables dinámicas del proceso: nivel, presión y caudal. Estas salidas están influenciadas por dos señales de entrada forzada: el estado de operación de la bomba (P01) y la apertura de la válvula de control (V01).

Para abordar esta tarea, se ha utilizado el algoritmo *Sparse Identification of Nonlinear Dynamics with Control* (SINDyc) en su versión discreta. La formulación empleada sigue el enfoque de regresión dispersa, donde se busca una función no lineal que describa la dinámica del sistema a partir de los datos temporales como se muestra en la ecuación 5.

La biblioteca de funciones $\Theta(\mathbf{X}, \mathbf{U})$ se ha construido utilizando términos polinomiales hasta segundo orden, lo que permite capturar efectos no lineales como las pérdidas de carga cuadráticas generadas por la válvula y la bomba. Además, se ha incorporado un término de raíz cuadrada, motivado por las relaciones de tipo Bernoulli para los fluidos. De este modo, se consideran interacciones tanto lineales como no lineales entre las variables de estado y las entradas.

El modelo ha sido implementado utilizando la librería *PySINDy* de Silva et al. (2020); Kaptanoglu et al. (2022) y para la optimización de los hiperparámetros se ha empleado el *GridSampler* de *Optuna* Akiba et al. (2019) aprovechando el bajo tiempo de entrenamiento que tiene el modelo SINDyc. El algoritmo de regresión empleado para resolver la ecuación 7 ha sido *Sequentially Thresholded Least Squares* (*STLSQ*), cuya configuración se ha basado en dos hiperparámetros: el *threshold* y el *alpha*. Los valores explorados para estos parámetros se detallan en la Tabla 1.

Parámetro de optimización	Intervalo de valores
`threshold`	0.001 – 0.1
`alpha` (regularización L2)	0.001 – 0.05

Tabla 1: Rangos de hiperparámetros explorados para STLSQ

3.3. Generación y tratamiento de datos

Los datos empleados para el entrenamiento del modelo se han generado mediante un experimento de identificación de sistemas en lazo cerrado. En este experimento, se utilizó un controlador proporcional para regular el nivel del tanque, aplicando cambios secuenciales tanto en la consigna de nivel como en la posición de la válvula de control.

La recolección de datos se ha realizado con un tiempo de muestreo de 125 milisegundos utilizando una aplicación en Python que interactúa directamente con el PLC del sistema, para garantizar una captura de datos fiable y ajustada en tiempo. La duración total del experimento fue de 6 horas. Los datos se han remuestreado a intervalos de 1 segundo y, para garantizar una representación uniforme de todas las variables, también fueron normalizados en un rango de 0 a 1 utilizando un *MinMaxScaler*. Posteriormente, los datos fueron divididos en tres conjuntos: 60 % para entrenamiento, 20 % para validación y 20 % para test.

4. Resultados experimentales

Variable	R^2	RMSE
Nivel	0.961	0.053
Presión	0.987	0.027
Caudal	0.979	0.028

Tabla 2: Resultados de predicción para el conjunto de test

Mediante el proceso de optimización de hiperparámetros descrito en la anterior sección, se identificaron los valores óptimos para el algoritmo STLSQ, que son: `threshold` = 0.009 y `alpha` = 0.011. Estos parámetros se han seleccionado para proporcionar un equilibrio adecuado entre la complejidad de las ecuaciones del modelo (evitar expresiones con demasiados términos) y su capacidad de predicción.

Los resultados obtenidos en el conjunto de test se resumen en la Tabla 2. Se observa que el modelo alcanza valores elevados de R^2 y bajos valores de RMSE para las tres variables

Variable	Ecuación
Nivel	$\mathrm{L}[k+1] = -0{,}049\sqrt{\lvert\mathrm{Q}[k]\rvert} + 0{,}992\mathrm{L}[k] + 0{,}083\mathrm{Q}[k] + 0{,}027\mathrm{P}[k]^2 - 0{,}071\mathrm{P}[k]\mathrm{B}[k] - 0{,}041\mathrm{Q}[k]^2 + 0{,}045\mathrm{Q}[k]\mathrm{B}[k] + 0{,}033\mathrm{B}[k]^2$
Presión	$\mathrm{P}[k+1] = 0{,}038\sqrt{\lvert\mathrm{L}[k]\rvert} + 0{,}085\sqrt{\lvert\mathrm{P}[k]\rvert} + 0{,}065\sqrt{\lvert\mathrm{B}[k]\rvert} - 0{,}010 - 0{,}049\mathrm{L}[k] + 0{,}679\mathrm{P}[k] - 0{,}095\mathrm{Q}[k] - 0{,}342\mathrm{B}[k] + 0{,}030\mathrm{V}[k] + 0{,}018\mathrm{L}[k]^2 + 0{,}336\mathrm{P}[k]^2 + 0{,}171\mathrm{P}[k]\mathrm{Q}[k] - 1{,}161\mathrm{P}[k]\mathrm{B}[k] - 0{,}041\mathrm{Q}[k]^2 - 0{,}094\mathrm{Q}[k]\mathrm{B}[k] - 0{,}021\mathrm{Q}[k]\mathrm{V}[k] + 1{,}351\mathrm{B}[k]^2 + 0{,}021\mathrm{B}[k]\mathrm{V}[k] - 0{,}020\mathrm{V}[k]^2$
Caudal	$\mathrm{Q}[k+1] = 0{,}087\sqrt{\lvert\mathrm{P}[k]\rvert} - 0{,}172\sqrt{\lvert\mathrm{Q}[k]\rvert} - 0{,}034\sqrt{\lvert\mathrm{B}[k]\rvert} + 0{,}035\sqrt{\lvert\mathrm{V}[k]\rvert} + 0{,}024 - 0{,}019\mathrm{L}[k] - 0{,}049\mathrm{P}[k] + 1{,}004\mathrm{Q}[k] + 0{,}074\mathrm{V}[k] - 0{,}068\mathrm{L}[k]\mathrm{P}[k] + 0{,}063\mathrm{L}[k]\mathrm{B}[k] + 0{,}063\mathrm{P}[k]^2 + 0{,}108\mathrm{P}[k]\mathrm{Q}[k] - 0{,}374\mathrm{P}[k]\mathrm{B}[k] + 0{,}168\mathrm{P}[k]\mathrm{V}[k] - 0{,}238\mathrm{Q}[k]^2 + 0{,}188\mathrm{Q}[k]\mathrm{B}[k] - 0{,}181\mathrm{Q}[k]\mathrm{V}[k] + 0{,}323\mathrm{B}[k]^2 + 0{,}055\mathrm{B}[k]\mathrm{V}[k] - 0{,}054\mathrm{V}[k]^2$

Tabla 3: Ecuaciones identificadas para las variables del proceso

analizadas. Particularmente, la presión y el caudal presentan las mejores predicciones, con R^2 superiores al 0.97 y errores cuadráticos medios muy bajos. En el caso del nivel, aunque el desempeño también es bueno ($R^2 = 0{,}961$), existe un margen de mejora que podría abordarse refinando la biblioteca de funciones.

Las ecuaciones identificadas por el algoritmo SINDyc para predecir el nivel del tanque, la presión y el caudal del circuito de proceso, pueden verse en la tabla 3.

Del análisis de la ecuación del nivel (L[k+1]), se pueden extraer varios aspectos destacables. El coeficiente 0.992 asociado al término L[k] sugiere una alta inercia en la evolución del nivel, lo cual es coherente con la física del sistema. Además, la presencia de términos cuadráticos permiten representar las diferentes relaciones cuadráticas presentes en el sistema. Cabe destacar la ausencia de la válvula en la ecuación, lo cual podría deberse a estar de forma implícita en el control del caudal, término que tiene bastante influencia. La Figura 2 muestra la comparación entre la predicción del nivel y los valores reales. En líneas generales, el modelo reproduce adecuadamente la dinámica del nivel, aunque se observan ciertas zonas donde el error es ligeramente mayor, posiblemente atribuibles a fenómenos no modelados o a dinámicas de más alta frecuencia.

La ecuación correspondiente a la presión (P[k+1]) muestra que la dinámica de la presión está fuertemente dominada por el estado de la bomba, evidenciado por el coeficiente del término B[k]^2, que presenta el mayor valor absoluto dentro de la ecuación. Esto indica que el comportamiento de la presión está principalmente determinado por la operación de la bomba, con influencias adicionales menores de otras variables como la válvula de control. La Figura 4 muestra la comparación entre la presión real y la predicha por el modelo, junto con las acciones de control de la bomba y la válvula de proceso. Se observa que la predicción sigue muy de cerca el comportamiento de la presión real, resultado coherente con los valores elevados de R^2 y bajos de RMSE obtenidos. Asimismo, se aprecia como la dinámica de la bomba y la presión son muy similares, corroborando que la presión responde principalmente a la acción de la bomba, con pequeñas variaciones introducidas por la válvula.

La ecuación correspondiente al caudal (Q[k+1]) muestra que, al igual que la presión, el caudal está fuertemente influenciado por el estado de la bomba, como se refleja en los coeficientes de los términos donde aparece esta variable. Sin embargo, a diferencia de la presión, el caudal también exhibe una mayor dependencia de la válvula de control, lo que introduce variaciones adicionales en su comportamiento. La Figura 3 muestra la comparación entre el caudal real y el predicho. Se aprecia que el caudal sigue de manera general la dinámica de la bomba, aunque las acciones de la válvula introducen fluctuaciones más pronunciadas respecto al caso de la presión. La predicción captura adecuadamente estas dinámicas, lo cual es coherente con los buenos resultados obtenidos en las métricas de validación (R^2 y RMSE).

5. Conclusiones

En este trabajo se ha presentado un enfoque preliminar basado en *Sparse Identification of nonlinear Dynamics with control (SINDyc)* para la construcción de un gemelo digital de un sistema industrial no lineal. A través de la optimización de la biblioteca de funciones y la adecuada selección de hiperparámetros, se ha logrado obtener modelos parsimoniosos capaces de predecir con alta precisión las principales variables del proceso: nivel, presión y caudal.

Los resultados experimentales demuestran que la metodología propuesta, aunque preliminar, permite capturar de manera efectiva la dinámica del sistema, alcanzando valores elevados de R^2 y bajos errores de predicción (RMSE). Particularmente, las variables de presión y caudal muestran un ajuste excelente entre los valores reales y los predichos, mientras que la predicción del nivel, aunque satisfactoria, podría beneficiarse de futuras mejoras en la selección de funciones dentro de la biblioteca.

La utilización de SINDyc ha permitido no solo obtener modelos de alta precisión, sino también modelos interpretables desde un punto de vista físico, lo cual resulta esencial para su integración en aplicaciones industriales de monitorización, control y optimización.

Como líneas futuras de trabajo, se plantea explorar bibliotecas de funciones enriquecidas que incorporen términos físicos específicos del proceso, así como investigar la implementación de estrategias de identificación adaptativa que permitan actualizar el modelo en tiempo real frente a cambios en las condiciones de operación.

Agradecimientos

Este trabajo es parte del proyecto de investigación PID2020-117890RB-I00 financiado por MCIN/AEI/10.13039/501100011033.

Referencias

Akiba, T., Sano, S., Yanase, T., Ohta, T., Koyama, M., 2019. Optuna: A next-generation hyperparameter optimization framework. In: Proceedings of the

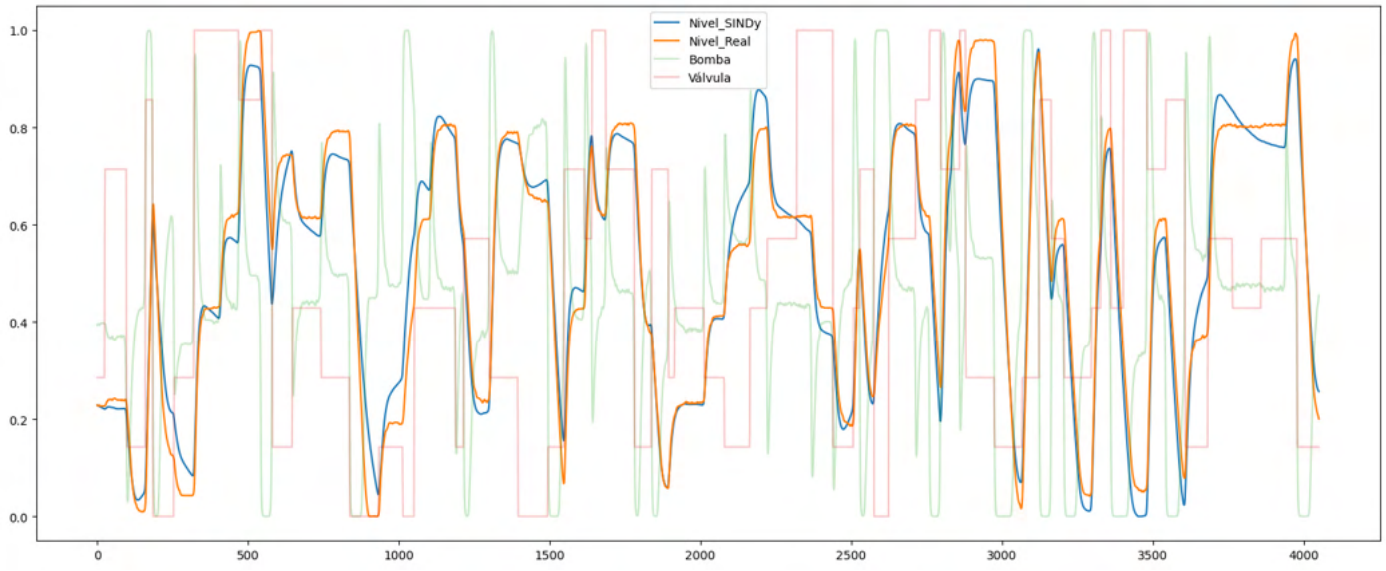

Figura 2: Resultado del modelo SINDyc para la variable nivel

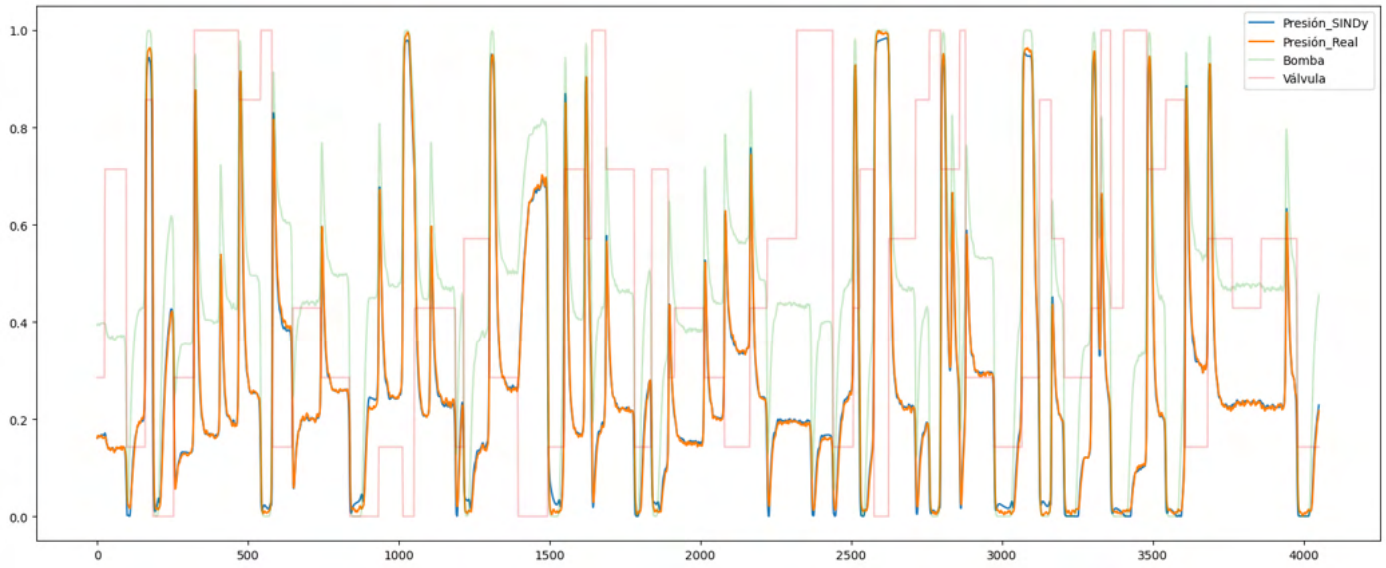

Figura 3: Resultado del modelo SINDyc para la variable presión

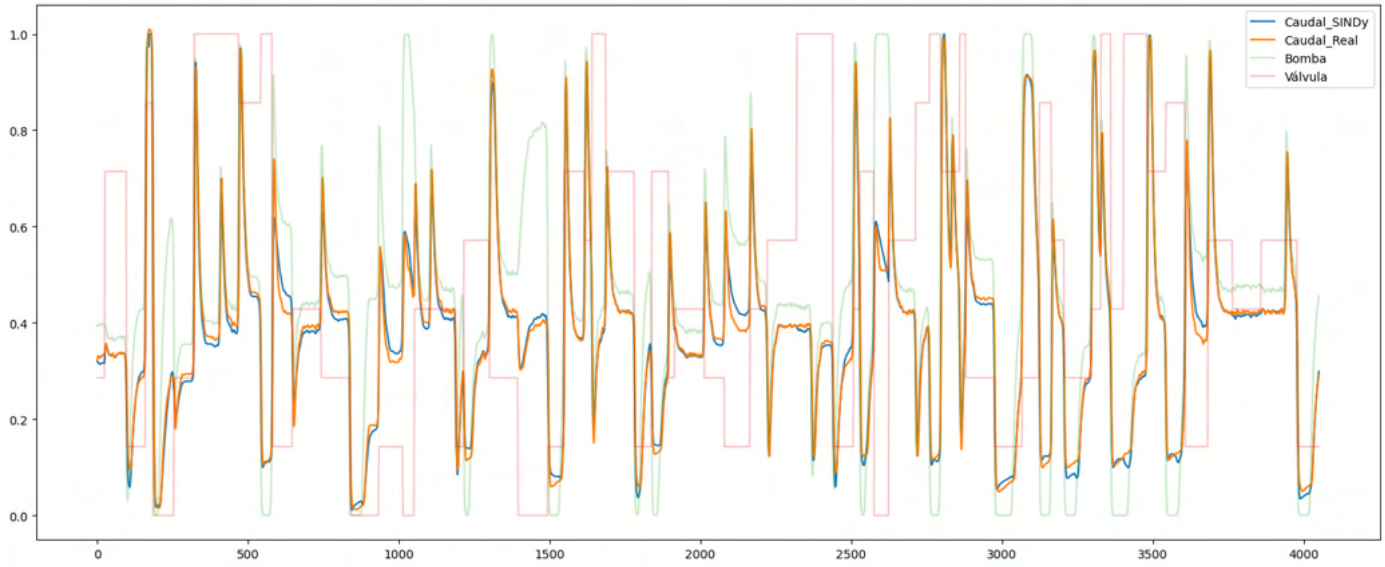

Figura 4: Resultado del modelo SINDyc para la variable caudal

25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.
URL: https://doi.org/10.1145/3292500.3330701
DOI: 10.1145/3292500.3330701

Birk, W., Hostettler, R., Razi, M., Atta, K., Tammia, R., 2022. Automatic generation and updating of process industrial digital twins for estimation and control - a review. Frontiers in Control Engineering.
DOI: 10.3389/fcteg.2022.954858

Brunton, S. L., Proctor, J. L., Kutz, J. N., 2016a. Discovering governing equations from data by sparse identification of nonlinear dynamical systems. Proceedings of the national academy of sciences 113 (15), 3932–3937.
DOI: 10.1073/pnas.1517384113

Brunton, S. L., Proctor, J. L., Kutz, J. N., 2016b. Sparse identification of nonlinear dynamics with control (sindyc). IFAC-PapersOnLine 49 (18), 710–715.
DOI: 10.1016/j.ifacol.2016.10.249

de Silva, B., Champion, K., Quade, M., Loiseau, J.-C., Kutz, J., Brunton, S., 2020. Pysindy: A python package for the sparse identification of nonlinear dynamical systems from data. Journal of Open Source Software 5 (49), 2104.
URL: https://doi.org/10.21105/joss.02104
DOI: 10.21105/joss.02104

González-Herbón, R., González-Mateos, G., Rodríguez-Ossorio, J., Domínguez, M., Castro, S. A., Fuertes-Martínez, J. J., 2024. An approach to develop digital twins in industry. Sensors.
DOI: 10.3390/s24030998

Kaptanoglu, A. A., de Silva, B. M., Fasel, U., Kaheman, K., Goldschmidt, A. J., Callaham, J., Delahunt, C. B., Nicolaou, Z. G., Champion, K., Loiseau, J.-C., Kutz, J. N., Brunton, S. L., 2022. Pysindy: A comprehensive python package for robust sparse system identification. Journal of Open Source Software 7 (69), 3994.
URL: https://doi.org/10.21105/joss.03994
DOI: 10.21105/joss.03994

Min, Q., Lu, Y., Liu, Z., Su, C., Wang, B., 2019. Machine learning based digital twin framework for production optimization in petrochemical industry. International Journal of Information Management.
DOI: 10.1016/J.IJINFOMGT.2019.05.020

Sun, Q., Ge, Z., 2021. A survey on deep learning for data-driven soft sensors. IEEE Transactions on Industrial Informatics.
DOI: 10.1109/TII.2021.3053128

Tao, F., Zhang, H., Liu, A., Nee, A. Y. C., 2019. Digital twin in industry: State-of-the-art. IEEE Transactions on Industrial Informatics.
DOI: 10.1109/TII.2018.2873186

Wagg, D., Worden, K., Barthorpe, R., Gardner, P., 2020. Digital twins: State-of-the-art and future directions for modeling and simulation in engineering dynamics applications. ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part B: Mechanical Engineering.
DOI: 10.1115/1.4046739

Wang, J., Moreira, J., Cao, Y., Gopaluni, R. B., 2023a. Neural network and sparse identification of nonlinear dynamics integrated algorithm for digital twin identification. In: IFAC-PapersOnLine. Vol. 56. Elsevier, pp. 6921–6926.
DOI: 10.1016/j.ifacol.2023.10.503

Wang, J., Moreira, J., Cao, Y., Gopaluni, R. B., 2023b. Simultaneous digital twin identification and signal-noise decomposition through modified generalized sparse identification of nonlinear dynamics. Computers and Chemical Engineering 177, 108294.
DOI: 10.1016/j.compchemeng.2023.108294