



Caracterización híbrida de un dispositivo cosechador de energía

Mikel Merino-Olagüe^{1,2}, Aitor Plaza¹, Xabier Iriarte^{1,2}, Julen Bacaicoa^{1,2}

¹ Dpto. de Ingeniería. Universidad Pública de Navarra. [mikel.merino, aitor.plaza, xabier.iriarte, julen.bacaicoa]@unavarra.es

² Institute of Smart Cities (ISC). Universidad Pública de Navarra. [mikel.merino, xabier.iriarte, julen.bacaicoa]@unavarra.es

La obtención de modelos matemáticos para sistemas mecánicos constituye una disciplina fundamental en el ámbito de la ingeniería. Estos modelos son esenciales para llevar a cabo cálculos, simulación y diseño, obteniéndose principalmente a partir de principios físicos o del ajuste de un modelo paramétrico de caja negra. Sin embargo, ambas metodologías tienen una serie de limitaciones: los modelos basados en física pueden no tener en cuenta ciertos fenómenos y los modelos de caja negra son complicados de interpretar. En este trabajo, se pretende caracterizar un dispositivo cosechador de energía cuya física es parcialmente desconocida debido a la complejidad de modelar una interacción magnética. Debido a las limitaciones de las metodologías de identificación clásicas, se ha desarrollado una nueva metodología que combina los principios físicos con el modelado caja negra. En este caso concreto, se combina la segunda ley de Newton con una red neuronal, modelando esta segunda la interacción magnética desconocida. Tras el proceso de ajuste o entrenamiento del modelo híbrido, se han obtenido los valores de los parámetros con sentido físico del dispositivo. Además, el modelo obtenido puede reorganizarse para realizar integraciones numéricas y predecir su comportamiento a futuro, puede linealizarse y permite variar los valores de los parámetros físicos. Los resultados muestran que esta metodología posibilita la caracterización del cosechador de energía de forma precisa e interpretable, permitiendo la utilización del modelo obtenido para llevar a cabo tareas de análisis, simulación y diseño.

1. Introducción

En todas las áreas de la ingeniería, el modelado se ha convertido en una herramienta fundamental para llevar a cabo cálculos, simulaciones y diseño. Los modelos más comunes están basados en principios básicos de la física que relacionan las variables de interés. Las leyes de Newton se utilizan para modelar sistemas mecánicos, las leyes de Maxwell para sistemas electromagnéticos y en el caso de sistemas que experimentan cambios de temperatura, presión y volumen junto con transferencia de energía en forma de calor, se aplican los principios de la termodinámica. La gran mayoría de estos modelos tiene como objetivo predecir el comportamiento de los sistemas que modelan con suficiente precisión. La identificación de sistemas tiene como objetivo construir modelos que puedan ajustarse para que sus predicciones coincidan con los datos obtenidos de experimentos diseñados para este propósito. Cuando se construyen modelos paramétricos cuyos parámetros se suponen conocidos, se les denomina modelos de *caja blanca*. En caso de que los valores de los parámetros no se conozcan *a priori* y se estimen, se les llama modelos de *caja gris*.

Un sistema mecánico puede ser modelado mediante ecuaciones diferenciales de segundo orden a partir de la aplicación de los principios de la física. Estas ecuaciones generalmente dependen de una serie de parámetros que tienen sentido físico (masa, inercia, fricción, amortiguamiento, rigidez, etc.) que pueden resultar complicados de medir directamente. En consecuencia, el problema de identificación de sistemas se convierte en un problema de estimación de parámetros, en el que se estiman los valores de los parámetros que mejor se ajustan a los datos experimentales de acuerdo a un criterio concreto [1,2]. La estimación de parámetros ha sido empleada para la obtención de modelos en diversas áreas de la ingeniería mecánica, tales como la robótica [3-5], la aeronáutica [6,7], los vehículos [8,9] o los aerogeneradores [10].

En determinadas circunstancias, la mera modificación de los parámetros no proporciona suficiente flexibilidad para ajustar el modelo a los datos experimentales. Esta falta de precisión del modelo puede ocurrir cuando no existen principios físicos robustos para que permitan representar adecuadamente el sistema, ya sea debido a que las ecuaciones no se ajustan a la realidad o cuando existen fenómenos no modelados o difíciles de modelar. En estos casos, es posible utilizar modelos de *caja negra*, en los cuales las ecuaciones que representan el sistema no se fundamentan en principios físicos, pero poseen una estructura suficientemente general para representar las relaciones entrada-salida a partir de datos experimentales. El modelo autorregresivo Auto-Regressive Moving Average with eXogenous inputs (ARMAX) y las redes neuronales (RN) son dos ejemplos de este tipo de modelos. Las ventajas de los modelos *caja negra*, son la flexibilidad que les permite representar cualquier relación entrada-salida y el escaso (o nulo) conocimiento del que requieren para modelar un sistema. Sin embargo, son difíciles de interpretar y no suelen ser generalizables fuera del rango de datos con el que han sido ajustados.

Recientemente, con el desarrollo de la Inteligencia Artificial (IA) y específicamente del *Machine Learning* (ML) están apareciendo nuevas técnicas de identificación de sistemas y modelado [11,12]. Estas técnicas son ejemplos de modelado *caja negra* puesto que relacionan entradas y salidas mediante un modelo que depende de una serie de parámetros que carecen de sentido físico.

Por lo tanto, resulta evidente que las técnicas de modelado e identificación de *caja gris* y de *caja negra* tienen características complementarias. Los modelos de *caja gris* provienen de principios físicos, siendo interpretables y consistentes. No obstante, en algunos casos pueden no ser capaces de modelar fenómenos complejos debido a su falta de generalidad y posibles simplificaciones. Por su parte, los modelos de *caja negra* no responden a principios físicos y pueden ser tan generales como sea necesario. No obstante, estos modelos no suelen ser interpretables y se debe tener cuidado con el sobreajuste en su proceso de entrenamiento.

En este sentido, los investigadores están tratando de combinar las ventajas de ambos tipos de modelos añadiendo principios físicos a los algoritmos de ML, desarrollando una serie de técnicas denominadas *Physics Enhanced Machine Learning* (PEML). Las referencias [13,14] ofrecen una revisión exhaustiva de los algoritmos desarrollados con un enfoque general, mientras que [15] se centra en la mecánica de sólidos y fluidos, [16] aborda aplicaciones en mecatrónica y [17] se enfoca en dinámica estructural y vibroacústica. Recientemente, Haywood-Alexander *et al.* [18] han propuesto una clasificación de las técnicas actuales de PEML en forma de espectro, siendo los modelos más claros u oscuros dependiendo de la cantidad de datos y física que requieren. Los modelos de *caja blanca* tienen una alta dependencia de la física, por lo tanto, requieren pocos datos. Por su parte, los modelos de *caja negra* necesitan una gran cantidad de datos debido a que su dependencia de la física es escasa.

En el contexto de este trabajo, resulta de especial interés una técnica de modelado que se sitúa en la zona *clara* de del espectro comentado anteriormente. Esta técnica denominada *discrepancy modelling*, consiste en añadir un término adicional a modelos desarrollados a partir de principios físicos. El propósito de este término es compensar las discrepancias entre las predicciones del modelo basado en la física y los datos experimentales. Estas discrepancias pueden estar asociadas a pequeños efectos físicos cuya existencia es desconocida o a efectos de mayor escala cuya existencia es conocida, pero resultan difíciles de modelar. Este tipo de modelos pueden considerarse híbridos, ya que combinan partes del modelo que son de *caja gris* con partes que se modelan como una *caja negra*. Por lo tanto, estos modelos requieren conocimiento de la física, pero aceptan que ciertos efectos

físicos pueden ser desconocidos o complicados de modelar. Por ejemplo, en [19] se combinan leyes de la física con la técnica basada en datos *Sparse Identification of Nonlinear Dynamics* (SINDy) [20] para modelar el oscilador de Van der Pol y un sistema real de doble péndulo sobre un carro. Por su parte, Ebers *et al.* [21] proponen la modelización de las discrepancias mediante SINDy y tres algoritmos adicionales: *Dynamic Mode Decomposition* (DMD), *Gaussian Process Regression* (GPR) y RNs. Dichas variaciones se aplican a sistemas que pueden modelarse mediante EDOs como el oscilador de Van der Pol, el atractor de Lorenz y un sistema de masa-muelle-amortiguador. Además, también se aplican al sistema de Burgers de EDPs.

En el campo de la ingeniería mecánica, pueden encontrarse algunas aplicaciones de *discrepancy modelling*. Wohleben *et al.* [22] evaluaron 19 algoritmos de ML distintos para modelar las discrepancias entre un modelo basado en física de un péndulo doble y datos reales. Para su caso particular GPR y *bagged-trees* obtuvieron los mejores resultados. En [23] los autores van un paso más allá a la hora de aplicar *discrepancy modelling* a un mecanismo de biela manivela. No sólo modelan el muelle no lineal y un término de amortiguamiento desconocido, sino que identifican los parámetros del modelo físico (masa, inercia, etc.) de manera simultánea. Además, embeben el modelo híbrido en una Red Neuronal Recurrente (RNN) para predecir el estado del sistema en instantes de tiempo futuros. Askari *et al.* [24] utilizan la misma metodología en vehículos, identificando tanto los valores de los parámetros físicos como un modelo de neumáticos desconocido con un modelo de discrepancia embebido en una RNN.

En este trabajo se presenta una metodología híbrida para modelar el dispositivo cosechador de energía propuesto por Castellano *et al.* en [25], el cuál se puede modelarse de manera sencilla mediante leyes físicas, con la excepción de un término electromagnético que resulta complicado de modelar. Por consiguiente, la segunda ley de Newton se combina con una RN para modelar la interacción magnética desconocida entre las palas. En este caso, en vez de utilizar una RNN para predicciones futuras, se utiliza un modelo estático similar al de las *Neural ODEs*, debido a que esta metodología ofrece mejores resultados según [26]. La metodología desarrollada tiene la capacidad de añadir conocimiento físico previo al modelo y de identificar tanto los valores de los parámetros físicos como los de la RN de forma simultánea. Posteriormente, el modelo puede reorganizarse, despejando las aceleraciones de las coordenadas, para posteriormente poder realizar integraciones en el tiempo. Además, la RN puede linealizarse junto con el resto del modelo. En definitiva, el modelo identificado puede utilizarse para mejorar el diseño y analizar el dispositivo cosechador de energía.

El presente artículo se ha organizado en 5 secciones. La sección 2 describe el dispositivo cosechador de energía y la problemática de modelado que ha impulsado a aplicar la metodología desarrollada. En la sección 3 se introducen las técnicas clásicas de identificación de sistemas e identificación mediante RN. Posteriormente, se desarrolla teóricamente la metodología de modelado híbrido. La sección 4 presenta los resultados obtenidos tras la aplicación de la metodología al dispositivo cosechador de energía. Además, muestra como adaptar el modelo para realizar integraciones numéricas en el tiempo y como llevar a cabo su linealización. Por último, se presentan las conclusiones del trabajo en la sección 5.

2. Dispositivo cosechador de energía

El dispositivo cosechador de energía que se va a modelar en este trabajo, es el cosechador de ultra-baja frecuencia presentado por Castellano *et al.* en [25]. Se compone de una carcasa exterior formada por 24 bobinas que generan un campo magnético. Dentro de esta carcasa se ubican tres palas que giran de forma independientemente con respecto al eje de giro situado en el centro de la misma. Estas palas contienen una serie de imanes en configuración Halbach, de forma que concentran el campo magnético en la parte exterior de las mismas, produciendo un efecto de repulsión entre ellas. La figura 1 muestra el dispositivo real y un esquema del mismo.

El propósito de este dispositivo es transformar la energía mecánica, debida a las vibraciones que experimenta, en energía eléctrica. Su principio de funcionamiento es muy sencillo: Al moverse las palas debido a las vibraciones, se provoca una variación en el campo magnético, induciéndose una tensión en las bobinas que sería la energía eléctrica aprovechable. El dispositivo está pensando para instalarse en aerogeneradores y aprovechar las vibraciones para alimentar nodos sensores de su interior. De esta manera, se tiene un nodo sensor completamente autónomo y fácil de instalar y mantener.

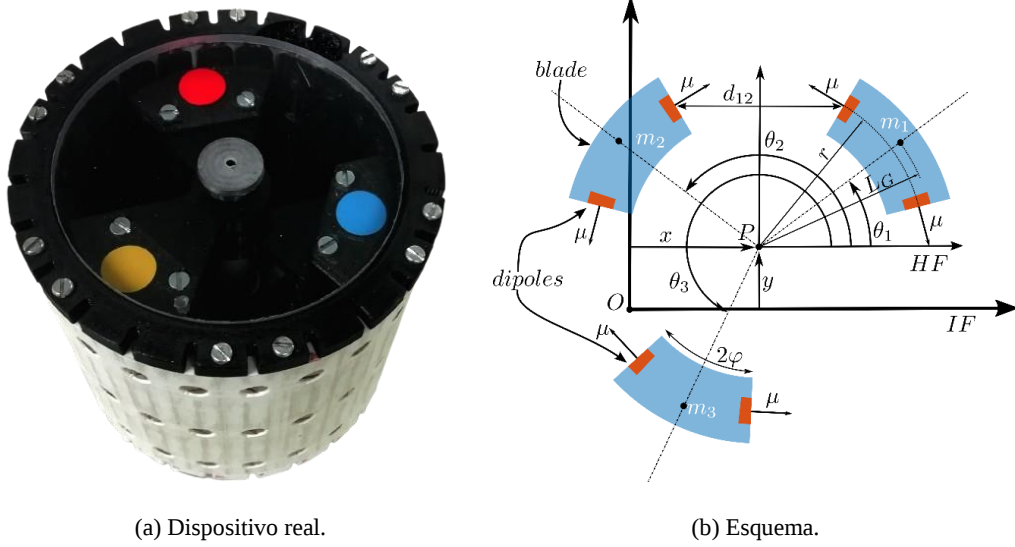


Figura 1: Dispositivo cosechador de energía.

La mayor complicación que presenta el modelado de este dispositivo es la interacción magnética entre las palas, que resulta altamente no-lineal y difícil de modelar mediante leyes físicas. Por esta razón, la sección 3 se presenta una nueva metodología para modelar este dispositivo combinando leyes físicas y técnicas de inteligencia artificial.

3. Metodología

La metodología propuesta para la identificación del cosechador de energía constituye una combinación de la estimación de parámetros de los modelos de caja gris y el modelado caja negra mediante ML. De la combinación de ambas metodologías se obtiene un modelo híbrido que podría utilizarse en multitud de problemas. En este apartado, se revisan ambas metodologías aplicadas a sistemas mecánicos y se propone una nueva metodología cuya implementación detallada para el caso del cosechador de energía se explica paso a paso.

3.1. Modelado clásico de caja gris basado en física

La dinámica de un sistema mecánico genérico se describe mediante ecuaciones diferenciales de segundo orden que pueden escribirse como

$$d(q, \dot{q}, \ddot{q}, y, \phi) = 0 \quad (1)$$

donde q son las coordenadas generalizadas, $\dot{q}$ las velocidades generalizadas, $\ddot{q}$ las aceleraciones generalizadas, y las excitaciones externas y ϕ los parámetros. En la estimación de parámetros de los modelos de caja gris el objetivo es obtener los valores de ϕ que mejor describen el sistema. Las ecuaciones de los sistemas mecánicos pueden escribirse de forma lineal en los parámetros inerciales. En el caso de que otros componentes dinámicos, tales como muelles, amortiguadores y efectos de fricción también sean lineales respecto a sus parámetros, las ecuaciones dinámicas pueden escribirse como

$$K(q, \dot{q}, \ddot{q})\phi = y. \quad (2)$$

En este caso, las ecuaciones son lineales en los parámetros ϕ . Para resolver esta ecuación se evalúa la ecuación 2 en n instantes de tiempo, obteniéndose el sistema de ecuaciones

$$W(q, \dot{q}, \ddot{q})\phi = \begin{bmatrix} K(q_1, \dot{q}_1, \ddot{q}_1) \\ K(q_2, \dot{q}_2, \ddot{q}_2) \\ \dots \\ K(q_n, \dot{q}_n, \ddot{q}_n) \end{bmatrix} \phi = \begin{bmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{bmatrix} = Y \quad (3)$$

el cual es sobredeterminado. Por lo tanto, debe resolverse un problema de optimización para determinar ϕ . El procedimiento más común es la resolución del problema mediante mínimos cuadrados y la matriz pseudoinversa de Moore-Penrose, obteniéndose así la solución

$$\hat{\phi} = \arg \min_{\phi} \|Y - W\phi\| = (W^T W)^{-1} W^T Y. \quad (4)$$

En el caso de que las ecuaciones no sean lineales en los parámetros, éstas ya no pueden escribirse tal y como se ha hecho en la ecuación 2, quedando

$$\mathbf{K}(\mathbf{q}, \dot{\mathbf{q}}, \ddot{\mathbf{q}}, \boldsymbol{\phi}) = \mathbf{y} \quad (5)$$

que ha de ser resuelto mediante técnicas de optimización no lineales para obtener los valores de $\boldsymbol{\phi}$.

Esta metodología permite obtener de forma fácil modelos sencillos de interpretar y que satisfacen las leyes de la física. Además, las técnicas de optimización están altamente documentadas y testadas en la bibliografía. Sin embargo, los modelos caja gris pueden no ser lo suficientemente precisos debido a que los modelos físicos pueden no ser exactos y a la aparición de fenómenos desconocidos o complicados de modelar mediante principios físicos. Por lo tanto, en determinados casos otras metodologías pueden ser más efectivas a la hora de ajustar un modelo a datos experimentales.

3.2. Modelado caja negra mediante redes neuronales

Las redes neuronales son aproximadores universales de funciones que relacionan una serie de entradas y salidas ajustando un conjunto de parámetros (pesos y sesgos) mediante un proceso de optimización. Estas características las hacen ideales para modelar sistemas cuya dinámica es completamente desconocida. Una opción en sistemas mecánicos sería utilizar una red neuronal cuya entrada fueran $\mathbf{q}$, $\dot{\mathbf{q}}$ y $\ddot{\mathbf{q}}$, siendo la salida $\mathbf{y}$. La red neuronal ajustaría sus parámetros optimizando una función de coste hasta representar correctamente a los datos experimentales.

Las redes neuronales tienen una mayor flexibilidad que los modelos de caja gris, pudiendo modelar efectos cuya física es desconocida. Sin embargo, al tratarse de modelos de caja negra que no se basan en principios físicos, son complicados de interpretar y además pueden no respetar dichos principios. Por ejemplo, un modelo de red neuronal podría generar energía pese a no haber ninguna fuente en el sistema real. Por último, las redes neuronales pueden experimentar problemas de sobreajuste, de forma que no generalizan a otros casos que no han sido utilizados para ajustar sus parámetros.

3.3. Modelado híbrido

Las dos metodologías introducidas en las secciones previas tienen una serie de ventajas y desventajas. En esta sección se propone una nueva metodología de modelado híbrido, en la que se utiliza un modelo de caja gris al que se le añaden uno o varios términos que modelan los fenómenos físicos desconocidos o sin modelo del sistema.

Un modelo dinámico inverso de un sistema mecánico puede escribirse tal y como se muestra en la ecuación 5. La función no lineal $\mathbf{K}$ puede dividirse en dos términos

$$\mathbf{K}(\mathbf{q}, \dot{\mathbf{q}}, \ddot{\mathbf{q}}) = \mathbf{f}(\mathbf{q}, \dot{\mathbf{q}}, \ddot{\mathbf{q}}, \boldsymbol{\vartheta}) + \mathbf{g}(\mathbf{q}, \dot{\mathbf{q}}, \ddot{\mathbf{q}}) = \mathbf{y}. \quad (6)$$

El término $\mathbf{f}$ es conocido, pero depende de los parámetros $\boldsymbol{\vartheta}$ cuyos valores son desconocidos. Esta función representa la física conocida del sistema mecánico. Por su parte, $\mathbf{g}$ es una función desconocida que representa la física desconocida del sistema mecánico. Esta arquitectura es particularmente interesante cuando aparecen fenómenos físicos desconocidos como fricción, juego entre pares cinemáticos e interacciones magnéticas, eléctricas o aerodinámicas desconocidas. La modelización de esta serie de fenómenos puede realizarse mediante $\mathbf{g}$. Un modelo que integre las funciones $\mathbf{f}$ y $\mathbf{g}$ exhibirá una mayor capacidad de ajuste a los datos de entrenamiento en comparación con un modelo que contemple únicamente $\mathbf{f}$ o $\mathbf{g}$.

Este conocimiento parcial de las ecuaciones, en conjunto con una red neuronal, puede ser utilizada para identificar conjuntamente la física desconocida del modelo y los valores de los parámetros con significado físico que aparecen en $\mathbf{f}$. La figura 2 muestra un esquema de la metodología de identificación híbrida propuesta en este trabajo. En este caso, la función desconocida $\mathbf{g}(\mathbf{q}, \dot{\mathbf{q}}, \ddot{\mathbf{q}})$ se ha sustituido por la red neuronal $\mathcal{N}_{\boldsymbol{\phi}}$, la cual depende de una serie de parámetros $\boldsymbol{\phi}$ sin sentido físico.

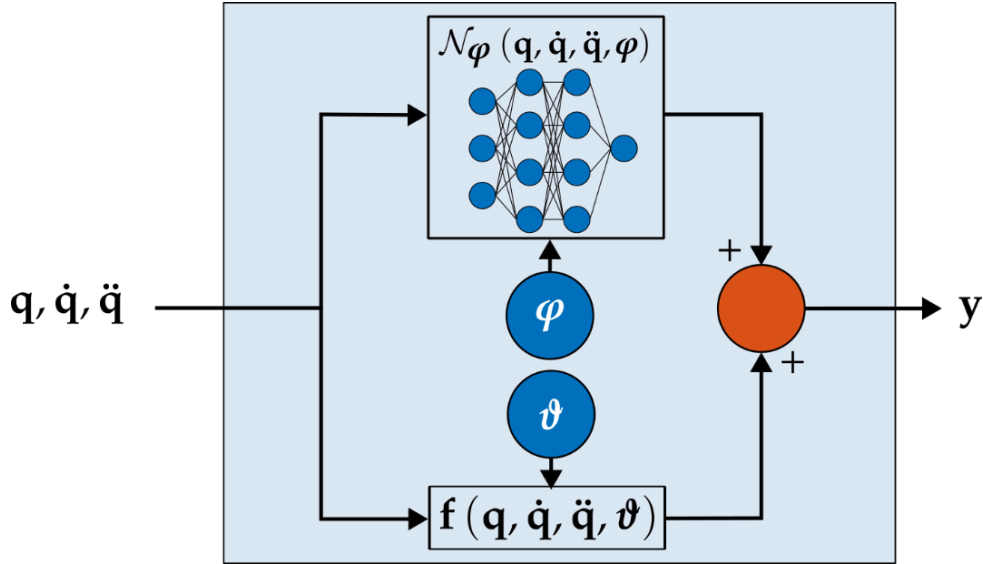


Figura 2: Estructura genérica de un modelo híbrido para la identificación de sistemas.

A esta estructura se le pueden añadir restricciones adicionales basadas en la física, de forma que se reduce el espacio de posibles soluciones, mejorándose la capacidad de generalización del modelo híbrido. Por ejemplo, si se tiene conocimiento de que la parte desconocida depende únicamente de $\mathbf{q}$, puede utilizarse una red neuronal con $\mathbf{q}$ como única entrada. Otros fenómenos como simetrías o restricciones de movimiento también podrían añadirse.

Una vez diseñado el modelo híbrido y sus restricciones, se define la función de coste

$$\mathcal{L} = \lambda_1 \|\mathbf{y} - \hat{\mathbf{y}}\| + \lambda_2 \mathbf{h} \quad (7)$$

donde $\hat{\mathbf{y}}$ son las salidas estimadas por el modelo, $\mathbf{h}$ representa los términos adicionales que podrían añadirse al modelo y λ_1 y λ_2 son dos factores de ponderación. El modelo se entrena optimizando simultáneamente los parámetros de la red neuronal ϕ y los parámetros del modelo físico θ . Una vez finalizado el proceso de entrenamiento, se lleva a cabo una validación cruzada para asegurar la generalización y viabilidad del modelo. La implementación se lleva a cabo mediante la librería PyTorch de Python, en la cual se define el modelo de forma modular. De esta forma, pueden reordenarse los distintos bloques para llevar a cabo integraciones numéricas como se mostrará en secciones posteriores.

3.4. Aplicación a un dispositivo cosechador de energía de ultra-baja frecuencia

La metodología de identificación desarrollada hasta el momento puede aplicarse a cualquier sistema mecánico. En esta sección, se procederá a la aplicación de la metodología al dispositivo cosechador de energía presentado en la sección 2, detallándose las particularidades del modelo.

La dinámica de cada una de las palas del cosechador de energía puede describirse mediante la ecuación diferencial

$$(mL_G^2 + I_{zz})\ddot{\theta}_i + (\sin^2(12\theta_i) b_g + b_0)\dot{\theta}_i + g_i(\theta_i, \theta_j, \theta_k) = \tau_i \quad (8)$$

donde aparecen cinco términos distintos: Los términos inerciales que multiplican a la aceleración angular $\ddot{\theta}_i$, los dos términos de amortiguamiento que multiplican a la velocidad $\dot{\theta}_i$, el término magnético que es desconocido y se sabe que depende de las posiciones angulares de las tres palas $\theta_i, \theta_j, \theta_k$ y el par externo τ_i . Los términos inerciales pueden agruparse mediante el parámetro base

$$I_{eq} = mL_G^2 + I_{zz} \quad (9)$$

que representa la inercia de una pala respecto al punto P de la figura 1b. El resto de parámetros son la masa de la pala m , la distancia desde el pivote hasta el centro de gravedad de la pala L_g , el momento de inercia respecto al centro de masa de la pala I_{zz} , el coeficiente de amortiguamiento eléctrico b_g y el coeficiente de amortiguamiento viscoso b_0 . Se supone que los parámetros físicos del sistema son idénticos para las tres palas y pueden organizarse en el vector

$$\boldsymbol{\theta} = \begin{Bmatrix} b_g \\ b_0 \\ I_{eq} \end{Bmatrix}. \quad (10)$$

Sustituyendo el término magnético desconocido $g_i(\theta_i, \theta_j, \theta_k)$ por una red neuronal $\mathcal{N}_{\varphi,i}(\theta, \varphi)$ que depende de los parámetros φ , puede plantearse el modelo híbrido de la figura 3 para el dispositivo cosechador de energía.

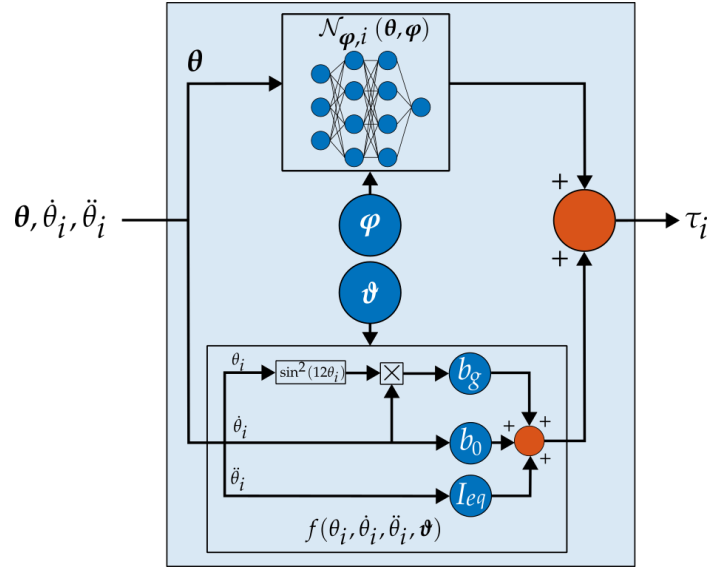


Figura 3: Modelo híbrido para la identificación del cosechador de energía.

Como puede verse, el modelo de identificación estima el par τ_i que experimenta una única pala. Para ello, utiliza las posiciones angulares de las tres palas θ , debido a que el término magnético depende de las posiciones relativas entre las palas, la velocidad angular de la pala i -ésima $\dot{\theta}_i$ y la aceleración angular de la pala i -ésima $\ddot{\theta}_i$. La simetría del sistema permite determinar el par para cada una de las palas permutando cíclicamente las entradas de la red neuronal: el par para la pala i se calcula utilizando como entrada de la red neuronal el vector $\{\theta_i, \theta_j, \theta_k\}^T$, para la pala j se utiliza el vector $\{\theta_j, \theta_k, \theta_i\}^T$ y para la pala k el vector $\{\theta_k, \theta_i, \theta_j\}^T$. La posibilidad de permutar las entradas hace que el modelo tenga un menor tamaño y pueda ajustarse fácilmente a un mayor número de palas. Nótese que la permutación puede realizarse debido a que los parámetros ϑ se suponen iguales para las tres palas.

Por otro lado, se han añadido restricciones adicionales al modelo. Los parámetros ϑ deben tener sentido físico, restringiéndose su valor a ser positivo en este caso. Además, la potencia generada por el dispositivo viene dada por

$$p_g = \left(\sum_{i=1}^3 \sin^2(12\theta_i) \dot{\theta}_i^2 \right) b_g \quad (11)$$

y puede medirse en el dispositivo real a partir de la tensión de salida de las bobinas de la carcasa. Por lo tanto, es posible comparar la potencia medida con la potencia teórica. El error cometido al estimar la potencia en todos los instantes de una época del proceso de entrenamiento puede añadirse a la función de coste, quedando

$$\mathcal{L} = \lambda_1 \|\mathbf{T} - \hat{\mathbf{T}}\| + \lambda_2 \|\mathbf{p}_g - \hat{\mathbf{p}}_g\| \quad (12)$$

donde λ_1 y λ_2 son dos factores de ponderación, $\mathbf{T} = \{\tau_i, \tau_j, \tau_k\}^T$ es el vector de pares externos medidos, $\hat{\mathbf{T}} = \{\hat{\tau}_i, \hat{\tau}_j, \hat{\tau}_k\}^T$ es el vector de pares externos estimados, $\mathbf{p}_g$ es la potencia medida y $\hat{\mathbf{p}}_g$ es la potencia estimada mediante la ecuación 12.

4. Experimentos y resultados

El procedimiento experimental utilizado es análogo al presentado por Plaza *et al.* en [27]. Los datos utilizados y su descripción pueden encontrarse en [28].

4.1. Entrenamiento y validación del modelo

Para llevar a cabo el entrenamiento es necesario medir las entradas y salidas que posteriormente se utilizarán en el modelo híbrido. Las posiciones, velocidades y aceleraciones angulares de las palas se obtienen a partir del procesamiento de vídeo. El par externo se calcula mediante

$$\tau_i = mL_G(\ddot{x} \sin \theta_i - \ddot{y} \cos \theta_i) \quad (13)$$

donde $\ddot{x}$ se mide mediante el acelerómetro montado en el cosechador de energía y m y L_G se miden físicamente. Nótese que $\ddot{y}$ se considera nulo puesto que los experimentos se llevan a cabo en una plataforma que se traslada únicamente en dirección x .

Se han llevado a cabo un total de 12 experimentos, realizando 3 experimentos para cada una de las 4 frecuencias de excitación (500 mHz, 750 mHz, 1000 mHz, 1200 mHz). Los 3 experimentos a 1000 mHz se han empleado para validar el modelo, mientras que los 9 experimentos restantes a 500 mHz, 750 mHz y 1200 mHz se han utilizado para llevar a cabo el entrenamiento. Los mejores resultados se han obtenido utilizando una ponderación de $\lambda_1 = \lambda_2 = 1$ y el método de optimizando ADAM con un factor de aprendizaje de 10^{-3} . La red neuronal empleada consta de 2 capas ocultas con 20 neuronas cada una, utilizando la función de activación *Tanh*. Se ha seleccionado esta arquitectura tras un proceso iterativo con diferentes configuraciones.

El modelo ha sido entrenado durante 120 épocas, con 200 instantes de tiempo en cada una, obteniéndose resultados muy satisfactorios. Nótese que en cada una de las épocas el modelo es evaluado en 3 ocasiones, permutando en cada una de ellas las componentes de θ según se ha explicado anteriormente. De esta manera, se calcula cada una de las componentes de $\hat{T}$ y posteriormente el valor de la función de coste.

La figura 4 muestra la comparación de las series temporales entre la predicción de par del modelo híbrido y el par medido de los datos de validación.

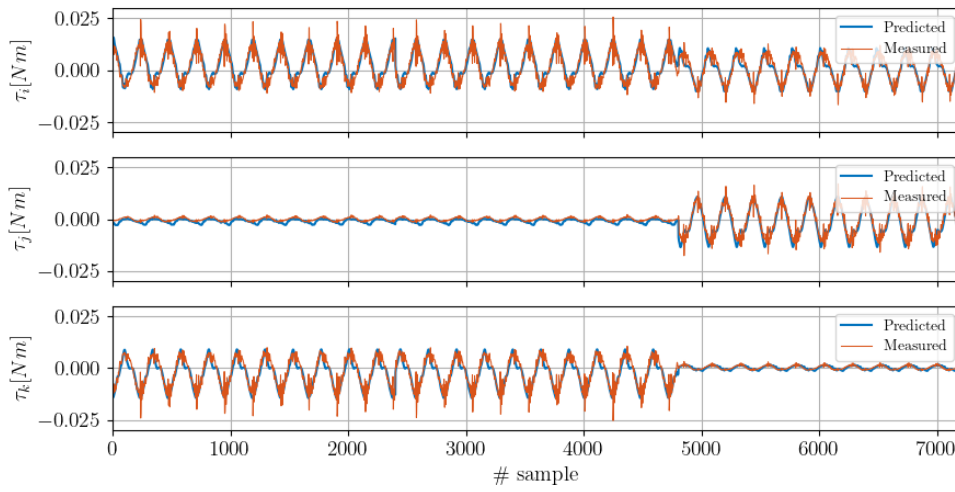


Figura 4: Comparación de par entre la predicción del modelo identificado y los datos experimentales.

La figura 5 por su parte, muestra la comparación entre la potencia predicha y medida.

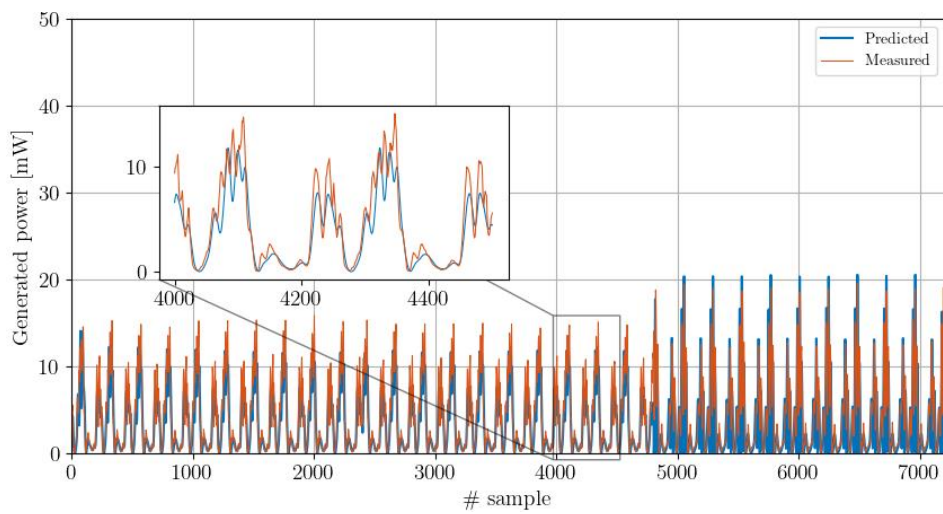


Figura 5: Comparación de potencia entre la predicción del modelo teórico y los datos experimentales.

A la vista de los resultados de ambas figuras, está claro que el modelo generaliza correctamente y no depende de la frecuencia. Asimismo, los resultados favorables en la estimación de la potencia se corresponden con una correcta estimación del parámetro b_g , debido a que es el único parámetro del que depende el modelo de potencia.

4.2. Integración numérica

Como se ha mencionado en la sección 3.4, los distintos bloques del modelo híbrido pueden reorganizarse de forma que permitan realizar integraciones numéricas. El modelo híbrido puede reorganizarse tal y como se muestra en la figura 7, de forma que se calcula la aceleración

$$\ddot{\theta}_i = \frac{1}{I_{eq}} [\tau_i - (\sin^2(12\theta_i)b_g + b_0)\dot{\theta}_i - \mathcal{N}_{\varphi,i}(\theta, \hat{\varphi})]. \quad (13)$$

De esta forma, puede calcularse la aceleración en cada instante de tiempo y utilizar integradores numéricos para predecir el comportamiento del cosechador y su producción de energía a lo largo del tiempo. Nótese que la red neuronal utiliza los parámetros $\hat{\varphi}$ estimados durante el entrenamiento. En el caso de que se desee simular el sistema original pueden utilizarse los parámetros físicos $\hat{\vartheta}$, aunque también existe la posibilidad de utilizar otros valores para dichos parámetros. De esta manera, se abre la posibilidad de analizar la influencia de la variación de estos parámetros en el dispositivo.

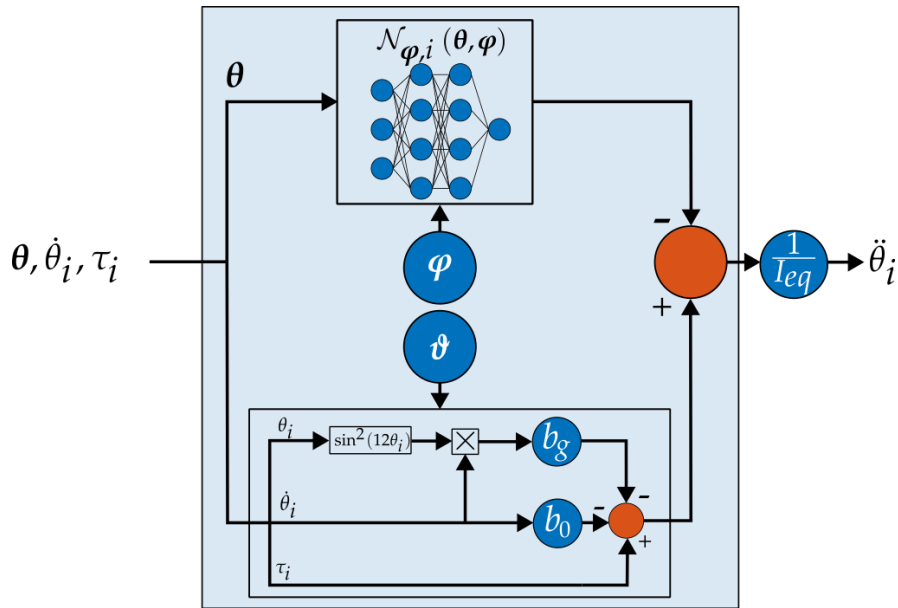
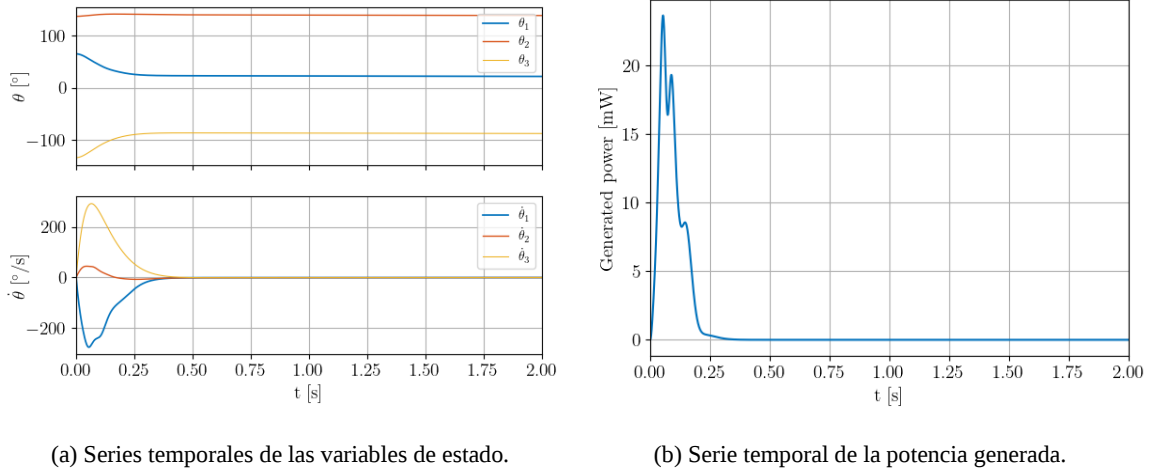


Figura 7: Modelo híbrido del cosechador de energía reorganizado para integración numérica.

La figura 8a muestra las series temporales de posición y velocidad obtenidas a partir de una simulación numérica. En dicha simulación, el cosechador parte de una posición cercana a la de equilibrio, sin experimentar la aplicación de cargas externas. Para los parámetros con sentido físico, se han utilizado los valores estimados $\hat{b}_g$, $\hat{b}_0$, $\hat{I}_{eq}$.



(a) Series temporales de las variables de estado.

(b) Serie temporal de la potencia generada.

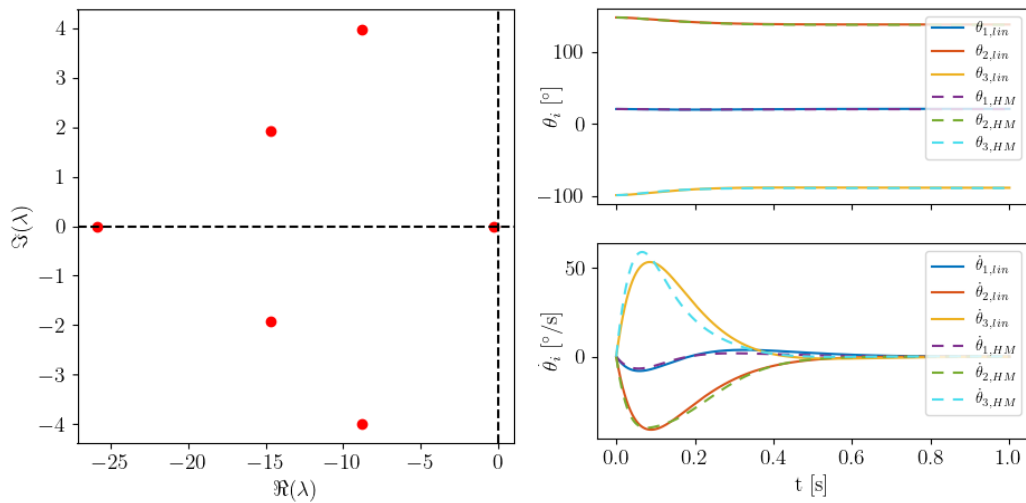
Figura 8: Integración numérica del modelo sin carga externa.

Como era de esperar, las palas tienden a retornar a su posición de equilibrio rápidamente, lo que resulta en una reducción de las velocidades hasta ser prácticamente nulas en menos de 0.5 segundos. Por su parte, la figura 8b muestra la variación de energía generada por el dispositivo. Se observa un primer pico al acelerarse las palas, el cual se atenúa rápidamente conforme éstas disminuyen su velocidad. Nuevamente, era el comportamiento esperado puesto que tal y como puede verse en la ecuación 11, el modelo de potencia va multiplicado por la velocidad al cuadrado. A la vista de los resultados, el modelo se asemeja a la realidad fielmente.

4.3. Linealización

El modelo híbrido puede linealizarse en torno a un punto de equilibrio gracias a la diferenciación automática implementada en la librería PyTorch. De este modo, para linealizar el modelo completo se linealiza $\mathbf{f}$ siguiendo el procedimiento habitual en ecuaciones diferenciales. Posteriormente se linealiza la red neuronal calculando la derivada parcial de la salida respecto de la entrada y se evalúa en el punto de equilibrio calculado previamente. Este punto de equilibrio cumple que tanto las velocidades como las aceleraciones angulares son nulas, el par externo es nulo y las palas no experimentan fuerza magnética.

El modelo linealizado puede validarse mediante la comparación de la simulación del modelo híbrido original y la del modelo híbrido linealizado en un punto cercano al de equilibrio. La figura 9a muestra la comparación de dicha simulación, en la que se perturba la posición inicial de las palas trasladando las palas j y k 10° y -10° respecto a su posición de equilibrio. Nótese que se han tomado las velocidades iniciales como nulas y que no hay excitación externa.



(a) Dispositivo real.

(b) Esquema.

Figura 9: Dispositivo cosechador de energía.

Como puede verse, las series temporales son muy similares, de forma que el modelo linealizado es válido en un entorno cercano al punto de equilibrio y pueden analizarse sus valores propios. Estos valores propios se muestran en la figura 9b. El valor propio puramente real cercano al eje imaginario domina la dinámica del sistema de forma que el sistema es sobreamortiguado. La simulación 9a muestra claramente este comportamiento, amortiguándose rápidamente el movimiento de las palas y no sobreoscilando ninguna de ellas en su respuesta.

5. Conclusiones

Este trabajo presenta una nueva metodología para la identificación y modelización de sistemas mecánicos que poseen física parcialmente desconocida o en los que aparecen fenómenos desconocidos. La metodología se ha aplicado a un dispositivo cosechador de energía, en el que la interacción magnética entre las palas es desconocida. Se han combinado las ecuaciones dinámicas del sistema con una red neuronal que modela la parte desconocida, de forma que se ha obtenido un modelo híbrido que depende de un conjunto de parámetros en el que se combinan los que tienen sentido físico con los que no. Durante el proceso de entrenamiento se ha obtenido el valor óptimo de estos parámetros.

El modelo obtenido depende de 3 parámetros con sentido físico cuyos valores pueden modificarse sin afectar a la validez del modelo. Este aspecto es particularmente interesante puesto que estos parámetros se pueden modificar convenientemente para modificar el diseño del cosechador de energía. Además, los distintos bloques del modelo pueden reorganizarse fácilmente de forma que se pueden realizar simulaciones con integradores numéricos conocidos. Asimismo, la posibilidad de linealizar el modelo híbrido es particularmente interesante para hacer un análisis de valores propios e implementar técnicas de control lineal.

Por otro lado, se han introducido múltiples restricciones físicas adicionales en el modelo, tales como el uso de una única red permutando sus entradas, el hecho de que la red neuronal dependa exclusivamente de las posiciones angulares y el restringir los parámetros físicos a valores positivos. Durante el proceso de entrenamiento, se ha optimizado una función de coste con un término adicional que utiliza las medidas de potencia. Esta serie de restricciones han permitido restringir el espacio de búsqueda de la solución y obtener un modelo que generaliza a casos distintos a los de entrenamiento.

En resumen, se ha diseñado una metodología que permite modelar, analizar, simular y optimizar un dispositivo con física parcialmente conocida como es el cosechador de energía.

6. Referencias

- [1] R. C. Aster, B. Borchers, C. H. Thurber, *Parameter estimation and inverse problems*, 2005.
- [2] E. Walter, L. Pronzato, *Identification of Parametric Models from Experimental Data*, Springer-Verlag, New York, 1997.
- [3] W. Khalil, E. Dombre, *Modeling identification and control of robots*, CRC Press, 2002.
- [4] J. Hollerbach, W. Khalil, M. Gautier, *Model identification*, Springer handbook of robotics (2016) 113–138.
- [5] N. Farhat, V. Mata, Page, F. Valero, *Identification of dynamic parameters of a 3-DOF RPS parallel manipulator*, Mechanism and Machine Theory 43 (2008) 1–17.
- [6] E. A. Morelli, V. Klein, *Aircraft system identification: theory and practice*, volume 2, Sunflyte Enterprises Williamsburg, VA, 2016.
- [7] R. V. Jategaonkar, *Flight vehicle system identification: a time domain methodology*, American Institute of Aeronautics and Astronautics, 2006.
- [8] M. Doumiani, A. Charara, A. Victorino, D. Lechner, *Vehicle dynamics estimation using Kalman filtering: experimental validation*, John Wiley & Sons, 2012.
- [9] D. Alberer, H. Hjalmarsson, L. Del Re, *System identification for automotive systems: Opportunities and challenges*, Identification for Automotive Systems (2012) 1–10.
- [10] N. Bao, Q. Chen, T. Jiang, *Modelling and identification of a wind turbine system*, Wind Engineering 20 (1996) 203–218.
- [11] T. A. Tutunji, *Parametric system identification using neural networks*, Applied Soft Computing 47 (2016) 251–261.
- [12] P. Rajendra, V. Brahmajirao, *Modeling of dynamical systems through deep learning*, Biophys Rev 12 (2020) 1311–1320.
- [13] S. Cuomo, V. S. Di Cola, F. Giampaolo, G. Rozza, M. Raissi, F. Piccialli, *Scientific Machine Learning Through PhysicsInformed Neural Networks: Where we are and Whats Next*, J Sci Comput 92 (2022) 88.

- [14] Z. K. Lawal, H. Yassin, D. T. C. Lai, A. Che Idris, Physics-Informed Neural Network (PINN) Evolution and Beyond: A Systematic Literature Review and Bibliometric Analysis, *Big Data and Cognitive Computing* 6 (2022) 140.
- [15] S. A. Faroughi, N. M. Pawar, C. Fernandes, M. Raissi, S. Das, N. K. Kalantari, S. Kourosh Mahjour, Physics-Guided, Physics-Informed, and Physics-Encoded Neural Networks and Operators in Scientific Computing: Fluid and Solid Mechanics, *Journal of Computing and Information Science in Engineering* 24 (2024) 040802.
- [16] S. Ayankoso, P. Olejnik, Time-Series Machine Learning Techniques for Modeling and Identification of Mechatronic Systems with Friction: A Review and Real Application, *Electronics* 12 (2023) 3669.
- [17] B. Zapparoli Cunha, C. Droz, A.-M. Zine, S. Foulard, M. Ichchou, A review of machine learning methods applied to structural dynamics and vibroacoustic, *Mechanical Systems and Signal Processing* 200 (2023) 110535.
- [18] M. Haywood-Alexander, W. Liu, K. Bacsá, Z. Lai, E. Chatzi, Discussing the Spectrum of Physics-Enhanced Machine Learning; a Survey on Structural Mechanics Applications, 2024.
- [19] K. Kaheman, E. Kaiser, B. Strom, J. N. Kutz, S. L. Brunton, Learning Discrepancy Models From Experimental Data, 2019.
- [20] S. L. Brunton, J. L. Proctor, J. N. Kutz, Discovering governing equations from data by sparse identification of nonlinear dynamical systems, *Proceedings of the national academy of sciences* 113 (2016) 3932–3937.
- [21] M. R. Ebers, K. M. Steele, J. N. Kutz, Discrepancy Modeling Framework: Learning missing physics, modeling systematic residuals, and disambiguating between deterministic and random effects, 2023.
- [22] M. Wohlleben, A. Bender, S. Peitz, W. Sextro, Development of a Hybrid Modeling Methodology for Oscillating Systems with Friction, in: G. Nicosia, V. Ojha, E. La Malfa, G. La Malfa, G. Jansen, P. M. Pardalos, G. Giuffrida, R. Umeton (Eds.), *Machine Learning, Optimization, and Data Science*, Springer International Publishing, Cham, 2022, pp. 101–115.
- [23] W. De Groote, E. Kikken, E. Hostens, S. Van Hoecke, G. Crevecoeur, Neural Network Augmented Physics Models for Systems With Partially Unknown Dynamics: Application to Slider–Crank Mechanism, *IEEE/ASME Trans. Mechatron.* 27 (2022) 103–114.
- [24] E. Askari, D. Gorgoretti, G. Crevecoeur, Hybrid modeling of multibody vehicles with partially known physics: discovering complex behaviors of tires, *Multibody Syst Dyn* (2024).
- [25] C. Castellano-Aldave, A. Carlosena, X. Iriarte, A. Plaza, Ultra-low frequency multidirectional harvester for wind turbines, *Applied Energy* 334 (2023) 120715.
- [26] R. T. Q. Chen, Y. Rubanova, J. Bettencourt, D. Duvenaud, Neural Ordinary Differential Equations, 2019.
- [27] A. Plaza, X. Iriarte, C. Castellano-Aldave, A. Carlosena, Comprehensive characterisation of a low-frequency-vibration energy harvester, *Sensors* 24 (2024).
- [28] J. Bacaicoa, M. Hualde-Otamendi, M. Merino-Olage, A. Plaza, X. Iriarte, C. Castellano-Aldave, A. Carlosena, Dataset for the identification of a ultra-low frequency multidirectional energy harvester for wind turbines, *Data in Brief* 57 (2024) 111126.