

# SISTEMA HÍBRIDO PARA LA PREDICCIÓN DEL FUNCIONAMIENTO DE UNA CELDA DE COMBUSTIBLE BASADA EN HIDRÓGENO, EMPLEADA EN EL ALMACENAMIENTO DE ENERGÍA

Juan Aurelio Montero-Sousa, Esteban Jove, José-Luis Casteleiro-Roca, Héctor Quintián, José Luis Calvo-Rolle  
 Universidade da Coruña, Departamento de Ingeniería Industrial  
 Avda. 19 de febrero s/n, 15405, Ferrol, A Coruña, España  
 {juan.montero.sousa,esteban.jove, jose.luis.casteleiro, hector.quintian, jlcalvo}@udc.es

Héctor Aláiz-Moretón  
 Universidad de León, Depart. de Ing. Eléctrica y de Sist. y Automática, Campus de Vegazana, 24071 León, España  
 hector.moreton@unileon.es

Tomás González-Ayuso  
 CIEMAT, Departamento de Energía, Av. Complutense 40, 28040 Madrid, España  
 tomas.gonzalez@ciemat.es

## Resumen

*En la actualidad, en gran medida debido al auge del vehículo eléctrico, los sistemas de almacenamiento energético son cada vez una necesidad mayor siendo tanto las baterías eléctricas como las pilas de combustible, las dos tecnologías en mayor desarrollo en los últimos años. Sin embargo, no es suficiente sólo desarrollar sistemas de almacenamiento de energía, sino que es indispensable maximizar la eficiencia de los mismos, para garantizar el máximo aprovechamiento de la energía almacenada. Para alcanzar dicho objetivo, uno de los aspectos más relevantes es el poder predecir con suficiente exactitud y antelación tanto la generación como el consumo energético que se hará sobre el dispositivo de almacenamiento. Es por ello, que la presente investigación se centra en el desarrollo de un sistema híbrido de modelado de una celda de combustible mediante técnicas de aprendizaje no supervisado para agrupamientos, combinadas con técnicas de regresión para modelado. Finalmente, los modelos generados con conjunto de datos real proveniente de un sistema de generación y almacenamiento de energía mediante una celda de hidrógeno, son validados obteniendo resultados altamente satisfactorios.*

**Palabras clave:** Pila de hidrógeno, almacenamiento de energía, SVM, ANN, BHL.

## 1. INTRODUCCIÓN

El cuidado del medio ambiente no es sólo una moda en la actualidad, sino que es un tema de gran

interés tanto para la sociedad como para los gobiernos actuales. A pesar de que conseguir un impacto cero en el medio ambiente se antoja muy difícil o incluso imposible, aspectos como la sostenibilidad y la reducción de la contaminación medioambiental son muy importantes [22]. En este aspecto, las energías renovables juegan un papel principal para contribuir a reducir el impacto medioambiental y las emisiones contaminantes [20], sin embargo éstas tampoco tienen un impacto nulo [31].

Ya que no es posible conseguir un impacto medioambiental cero, es una obligación legal el optimizar y maximizar la eficiencia de las nuevas instalaciones que se creen [40].

El sector eléctrico es un sistema complejo debido a diversos factores [25], siendo uno de los más relevantes la necesidad de ajustar la producción de energía eléctrica a la demanda de la misma, por lo que este sistema puede fácilmente desestabilizarse [24].

Con la aparición de las energías renovables instaladas en edificios conectados a la red, se contribuyó a complicar aún más un sistema de por sí muy complejo. Es por ello que aparece la necesidad de disponer de herramientas que permitan una correcta gestión energética tanto en los puntos de generación como en los de consumo.

Por todo lo descrito hasta ahora, nace el concepto de “Smartgrid” [1], con la necesidad de ser capaz no sólo medir la energía generada y consumida, sino también tratar de predecirlas con suficiente antelación con objeto de tomar las decisiones

correctivas adecuadas, haciendo el sistema global más eficiente.

De hecho, hay zonas geográficas donde edificios con demandas energéticas no disponen de conexión a la red eléctrica, por lo que una posible alternativa puede ser el combinar sistemas de energías renovables con sistemas de almacenamiento energético [10].

Por todo lo anteriormente mencionado, los sistemas de almacenamiento energético son cada vez más una necesidad para trabajar con la energía eléctrica. En la actualidad existen diversas tecnologías para ello, incluso alguna de ellas son relativamente antiguas e ineficientes como el bombeo de agua a presas [43]. Sin embargo, debido a las necesidades actuales, en los últimos años se han desarrollado proyectos y investigaciones en dicho ámbito [9]. El vehículo eléctrico ha sido el principal motivo del gran y rápido desarrollo de sistemas de almacenamiento de energía eléctrica [41].

Las tecnologías de almacenamiento de energía en las que más se ha trabajado en los últimos años son las baterías eléctricas y las pilas de combustible [17]. Siendo el hidrógeno [7], en el caso de estas últimas, el combustible primario más usado [39].

Además de disponer de sistemas de almacenamiento energético, es necesario ser capaces de modelar dichos sistemas con el objetivo de predecir la generación energética de los sistemas de generación para tomar decisiones que optimicen la eficiencia del sistema energético en su totalidad [30].

En cuanto a las técnicas de modelado de procesos, si bien existen gran variedad de las mismas, una de las más usadas es la basada en modelos de Análisis de Regresión Múltiple [30], habiendo sido usado previamente con éxito en diversas aplicaciones [18, 29, 5, 12]. Sin embargo, el método presenta ciertas limitaciones en varios casos [30, 6], siendo los sistemas no lineales uno de ellos. En estos casos, los sistemas inteligentes han demostrado capacidades superiores para trabajar dichos sistemas no lineales [34, 26, 37, 42], así como diversas técnicas de “Soft Computing” [28, 35, 11, 19]. Sin embargo, incluso con sistemas inteligentes el modelado de ciertos procesos es complejo debido, entre otros, a diversos estados de funcionamiento del mismo sistema. En estos casos, es posible dividir el problema a solventar mediante técnicas de “clustering” como “k-means” para un posterior modelado más sencillo [23, 15, 4, 32, 16].

En el presente trabajo, se realiza el modelado de una pila de combustible basándose en la predicción de hidrógeno, usando para ello diversas variables medidas en un sistema real de almacenamiento de energía. Debido a las no linealidades que presen-

ta el sistema, se ha optado por usar técnicas de agrupamiento combinadas con técnicas de modelado de “Soft Computing”. Para ello, se ha usado un método para analizar la estructura interna del conjunto de datos, llamado “Beta Hebbian Learning” (BHL) [33], y obtener así el número adecuado de grupos para una posterior fase de modelado. La principal ventaja del método propuesto, es que es posible incorporar el conocimiento experto humano durante la fase de “clustering” para posteriormente generar, mediante técnicas de regresión, modelos más sencillos y precisos para cada “cluster”.

Este trabajo de investigación está organizado comenzando con esta introducción y posterior presentación del caso de estudio. A continuación se describe el sistema híbrido de modelado propuesto, así como los materiales y métodos de que se dispone. Por último, tras presentar en detalle los experimentos llevados a cabo, y los resultados obtenidos, se discuten las conclusiones y los trabajos futuros.

## 2. CASO DE ESTUDIO

El caso de estudio consiste en un sistema experimental localizado en el “Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas” (CIEMAT). El sistema consta de dos fuentes de energías renovables: eólica y solar. En la figura 1 se presenta un esquema del sistema.

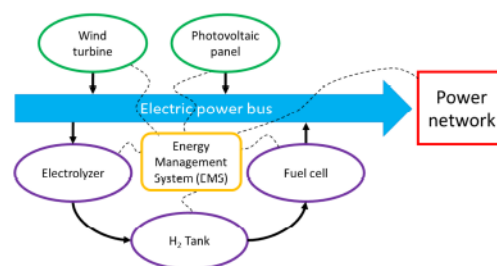


Figura 1: Sistema de gestión energético (EMS, Energy Management System)

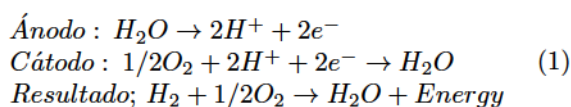
Tanto la generación eólica como la solar y la pila de combustible, son puntos de entrada al sistema (generación), siendo el equipo de generación de hidrógeno la única salida (consumo), a parte de la propia red eléctrica. El sistema de almacenamiento, está formado por un electrolizador, un tanque de hidrógeno y una pila de combustible. El propósito final es poder almacenar la energía no usada en fases de gran generación, para poder ser usada en momentos de gran consumo, en las que los sistemas primarios no sean capaces de abastecer dicha demanda.



## 2.1. SISTEMA DE ALMACENAMIENTO DE ENERGÍA

El sistema presentado en la Figura 1, consta de:

- **Electrolizador:** Produce el hidrógeno que alimenta a la pila de combustible. Para ello el equipo usa agua que va a través del desionizador donde es mezclada con hidróxido de potasio (20-30%), creando el electrolito. A continuación, aplicando una corriente eléctrica entre los electrodos negativo y positivo del electrolizador, se obtiene  $H_2$  y  $O_2$  por separado. El  $O_2$  es enviado directamente a la atmósfera y el  $H_2$  es filtrado para eliminar posibles restos de  $O_2$  mezclado. El  $H_2$  es almacenado en un tanque de hidrógeno para su uso posterior, dicho tanque cuenta con una válvula de seguridad en caso de sobrepresión.
- **Tanque de  $H_2$ :** El hidrógeno obtenido es almacenado a una presión de 10 bar. Ya que el sistema es experimental, el tamaño del tanque no es relevante, siendo necesario su correcto dimensionado en el caso de implantarse en una aplicación real. En el sistema experimental usado se precisa de un compresor para almacenar el hidrógeno a la presión de 10 bar requerida.
- **Pila de combustible:** La pila de combustible genera energía eléctrica a partir del  $H_2$  almacenado. La energía obtenida se basa en la reacción química expresada por la ecuación 1. Dicha reacción ocurre en cada celda, de modo que las celdas que se conecten en serie permiten sumar los voltajes individuales de cada una de ellas.



## 3. SISTEMA HÍBRIDO DE MODELADO

El sistema propuesto consiste en la combinación de técnicas de aprendizaje no supervisado para “clustering”, para generar grupos en el conjunto de datos disponible. A continuación, se obtienen modelos para cada “cluster” con técnicas de regresión. De este modo se genera un sistema híbrido más sencillo y preciso que el obtenido por un único modelo de todo el conjunto de datos.

El sistema híbrido consta de 3 etapas:

- Primero se realiza un preprocesado del conjunto de datos, eliminando “outliers” y medidas erróneas, procediendo entonces a la normalización del conjunto de datos.
- Segundo, BHL es aplicado para identificar la estructura interna del conjunto de datos y definir posibles “clusters” en el mismo de forma visual, incluyendo el conocimiento humano.
- Finalmente, diversas técnicas de regresión son aplicadas sobre cada uno de los grupos generados en el paso anterior, obteniéndose así el mejor modelo para cada “cluster”.

## 4. MATERIALES Y MÉTODOS

A continuación se describen los diversos algoritmos empleados en cada una de las 3 etapas en que se divide el sistema híbrido desarrollado.

### 4.1. PREPROCESADO

Como paso previo al uso de los datos, se ha aplicado una etapa de preprocesado consistente en la eliminación de “outliers” y medidas erróneas, para finalmente realizar una normalización de los datos restantes. Para ello se ha aplicado el criterio de normalización de “Maxmin Scaler” [36], que se presenta en la Ecuación (2):

$$\frac{X_i - \min(x)}{\max(x) - \min(x)} \quad (2)$$

El objetivo de este proceso de normalización es evitar una convergencia temprana en las primeras iteraciones de la fase de entrenamiento de los algoritmos de regresión empleados para el modelado [13].

### 4.2. CLUSTERING: BETA HEBBIAN LEARNING

El algoritmo BHL [33] es una red neuronal de aprendizaje no supervisado perteneciente a la familia de EPP, el cual emplea la distribución Beta en la actualización de los pesos de la regla de aprendizaje durante el entrenamiento. Con ello consigue ajustar de forma óptima la Función de Densidad de Probabilidad de los residuos con la distribución particular del conjunto de datos empleado.

En las ecuaciones (3), (4) y (5), se describe el proceso de aprendizaje durante la fase entrenamiento de la red neuronal.

$$\text{Feedforward: } y_i = \sum_{j=1}^N W_{ij}x_j, \forall i \quad (3)$$

$$\text{Feedback} : e_j = x_j - \sum_{i=1}^M W_{ij} y_i \quad (4)$$

$$\begin{aligned} \text{Actualización de pesos :} \\ \Delta W_{ij} = \eta (e_j^{\alpha-2} (1 - e_j)^{\beta-2} \\ (1 - \alpha + e_j (\alpha + \beta - 2))) y_i \end{aligned} \quad (5)$$

### 4.3. ALGORITMOS DE REGRESIÓN

Tras aplicar BHL para analizar la estructura interna del conjunto de datos y obtener los diversos “clusters”, se procede a aplicar diversas técnicas de regresión para generar modelos en cada grupo generado en la etapa anterior.

#### 4.3.1. Maquinas de Vector Soporte

Maquinas de Vector Soporte (SVMs, Support Vector Machines) [8] son una técnica de aprendizaje automático muy popular en problemas de clasificación, regresión y análisis de datos. Las SVMs funcionan construyendo hiperplanos en alta dimensionalidad con objeto de obtener separaciones lineales de los datos donde la separación de las clases es máxima. Este desarrollo de las SVMs es conocido como Clasificador de Vector Soporte, el cual puede ser extendido para solventar problemas de regresión.

Los principales parámetros de las SVRs (aplicación para regresión de las SVMs) son:

- Función de Kernel: Es la función de similitud elegida por las SVRs.
- C: Controla el margen de error de clasificación, cuanto mayor sea C, menor margen y cuando menor, mayor margen.
- Gamma: Define el grado y forma de influencia de cada muestra durante el aprendizaje.
- Grado: es el grado del polinomio del kernel.

#### 4.3.2. Extra Trees Regressor

Esta técnica es un “ensemble” de árboles de decisión basado en un meta estimador, para ello genera diversos arboles de decisión trabajando cada uno de ellos con un subconjunto del total de datos disponibles para luego, gracias a un enfoque de valores promedio, incrementar la precisión y evitar el sobre ajuste en comparación con un único árbol de decisión para todo el conjunto de datos [14, 38].

Los principales parámetros de dicha técnica son:

- Número de estimadores: el número de arboles de decisión.
- “Bootstrap”: técnica de división del conjunto de datos en subconjuntos.
- “Max depth”: máxima profundidad del árbol.
- “Min samples leaf”: número mínimo de muestras necesarios para alcanzar un nodo hoja.
- “Min samples split”: número mínimo de muestras necesarios para dividir un nodo.

#### 4.3.3. Red Neuronal Perceptrón Multicapa (MLP)

Ésta es una red neuronal de aprendizaje supervisado capaz de aprender gracias a la función:  $Fun(\cdot) : R^n \rightarrow R^0$ . En este trabajo, se ha implementado una MLP para regresión mediante “Python Scikit-Learn” [3]. Para el entrenamiento se usó el método de “back-propagation”, usando la función de activación lineal en la capa de salida. Los parámetros usados por la MLP implementada fueron:

- Coeficiente de aprendizaje inicial: Valor inicial del paso de aprendizaje, permite también controlar como varía dicho valor a medida que los pesos se van actualizando.
- Número de capas ocultas: determina cuantas capas ocultas se usarán en la red.
- Tamaño de las capas ocultas: número de neuronas en las capas ocultas.
- Función de activación: determina la función de activación (sigmoideal, lineal, etc.), usada en las neuronas de cada capa.
- “Solver”: define cómo se gestiona la optimización de los pesos, con las siguientes tres opciones. “LBFGS”, orientado a usar el algoritmo de la familia del método cuasi-Newton [27], “SGD” para usar el algoritmo de descenso de gradiente estocástico [2], y “ADAM”, orientado al algoritmo optimizador basado en gradiente estocástico [21].

## 5. CONJUNTO DE DATOS

El conjunto de datos usado en este artículo fue obtenido del sistema presentado en el caso de estudio. Dicho conjunto de datos consiste en un total de 281 muestras con 6 variables de entrada y 1 de salida, con un periodo de muestreo de 1 min:

- Variables de entrada:

- Temperatura del proceso.
  - Voltaje.
  - Corriente.
  - Potencia.
  - Presión de Ánodo.
  - Presión de Cátodo.
- Variable de salida:
    - $H_2$  generado.

En la pila de combustible se controló la corriente de salida; de este modo, la corriente de salida fue aumentando en un ciclo de carga de 0 A a 65 A con incrementos de 5 A en cada paso, y posteriormente disminuyendo de nuevo hasta 0 A con el mismo decremento. La duración de cada paso fue de aproximadamente de 10 min. En la figura 2 se muestra el proceso de descrito previamente.

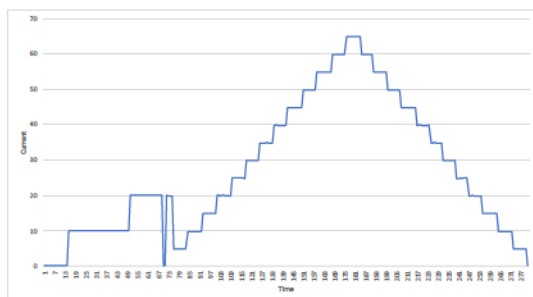


Figura 2: Salida de corriente de la celda de combustible

## 6. EXPERIMENTOS Y RESULTADOS

En esta sección se describen los experimentos llevados a cabo así como los resultados obtenidos en cada uno por los diferentes algoritmos implementados.

### 6.1. EXPERIMENTOS

El primero esta enfocado a analizar la estructura interna del conjunto de datos y determinar la existencia de posibles “clusters”. Para ello, después del preprocesado de los datos, se ha procedido a aplicar el algoritmo BHL, donde los valores de los parámetros de configuración del mismo han sido escogidos de forma experimental a través de un proceso iterativo de prueba y error, ya que dicho ajuste es muy dependiente de la propia estructura del conjunto de datos.

En un segundo experimento, se aplicaron las técnicas de regresión presentadas en la sección 4.3 a cada uno de los “clusters” identificados por BHL en el experimento anterior. Para ello el conjunto

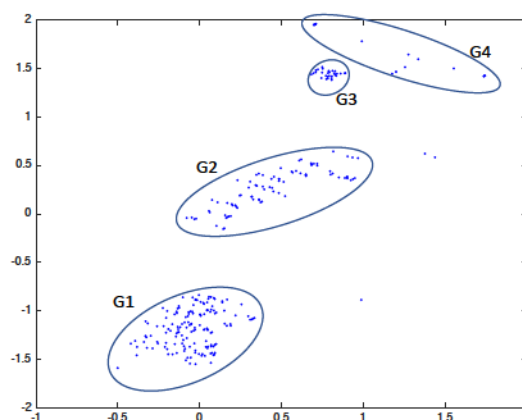


Figura 3: Proyección obtenida por BHL

de datos fue dividido en 2 partes, usando un 80 % de los datos para entrenamiento y el 20 % restante para validación del mejor modelo en términos de errores, de entre los generados para cada grupo.

Durante la fase de entrenamiento y test de los algoritmos empleados, se aplicó la técnica de validación cruzada, usando un total de 10 “folds”, mostrando finalmente los resultados obtenidos en términos de *Mean Absolute Error (MAE)*.

### 6.2. RESULTADOS

#### 6.2.1. Técnicas de “clustering”

En la figura 3, se presenta la mejor proyección obtenida por el algoritmo BHL, usando los parámetros indicados en la tabla 1.

En esta figura se pueden apreciar claramente 4 “clusters” en forma compacta y alejados unos de otros. Estos grupos se corresponden con los diferentes estados de funcionamiento del proceso y por tanto, presentan comportamientos muy diversos, por lo que son necesarios distintos modelos para cada uno de los estados.

BHL	iters=1000; lrate=0,05; $\alpha = 3, 2$ ; $\beta = 4, 3$
-----	--

Cuadro 1: Parámetros usados por BHL

#### 6.2.2. Resultados de los algoritmos de regresión

En las tablas 2, 3 y 4, se muestran las diversas combinaciones de parámetros usadas para cada uno de los algoritmos de regresión para la obtención del mejor modelo en cada caso.

En la tabla 5, se muestran los algoritmos de regresión elegidos para cada uno de los “clusters”,



Número de estimadores:	[50; 100; 200; 400; 600; 700; 800; 900]
Bootstrap:	[False; True]
Máxima profundidad:	[5; 10; 20; 40; 50; 60; 70]
Mínimo número de muestras nodo hoja:	[2; 3; 5; 7]
Mínimo número de muestras para división:	[2-6]

Cuadro 2: Combinaciones de parámetros para el algoritmo “Extra Trees Regressor”

Kernel	Lineal	Polinomial	RBF
Tol	0,1; 0,01; 0,001; 0,0001	0,1; 0,01; 0,001; 0,0001	0,1; 0,01; 0,001; 0,0001
Gamma	-	1; 0,1; 0,01; 0,001; 0,0001	1; 0,1; 0,01; 0,001; 0,0001
Degree	-	2; 3; 4; 5; 6; 7	-
C	1; 10; 100; 1000; 10000; 100000	1; 10; 100; 1000; 10000; 100000	1; 10; 100; 1000; 10000; 100000

Cuadro 3: Combinaciones de parámetros para el algoritmo SVR

Early Stopping:	[False; True]
Tamaño de las capas ocultas:	[5-20]
Nesterov momentum:	[False; True]
Solver:	[ <i>lbfgs, sgd, adam</i> ]
Coef. de aprendizaje inicial:	[0,9; 0,5; 0,01; 0,15]
Func. de activación:	[ <i>relu, tanh</i> ]
Warm start:	[True, False]
Batch size	[5; 10; 15; 20]

Cuadro 4: Combinaciones de parámetros para el algoritmo MLP para regresión

así como la mejor combinación de los parámetros presentados en las tablas anteriores, que han dado lugar al mejor modelo.

CLUSTER 1:	ETR; MAE: 1,0451; Número de estimadores: 600; Bootstrap: False; Profundidad máxima: 20; Mínimo número de muestras nodo hoja: 2; Mínimo número de muestras para división: 2
CLUSTER 2:	ETR; MAE: 1,2814; Número de estimadores: 50; Bootstrap: False; Profundidad máxima: 20; Mínimo número de muestras nodo hoja: 2; Mínimo número de muestras para división: 5
CLUSTER 3:	SVR; MAE: 1,2184; Kernel: RBF; Tol: 0,1; Gamma: 0,01; C: 1000
CLUSTER 4:	MLP; MAE: 0,0103; Early Stopping: True; Tamaño de las capas ocultas: 19; Nesterovs momentum: True; Solver: LBFSGS; Coef. de aprendizaje inicial: 0,9; Func. de activación: tanh; Warm start: False; Batch size: 5

Cuadro 5: Algoritmos elegidos y sus parámetros asociados para los mejores modelos en cada “cluster”

En las tablas 6, 7 y 8 se muestran los resultados obtenidos para los mejores modelos generados en cada “cluster”.

Cluster	1	2	3	4
LMLS	0,4771	0,6266	0,6857	0,0003
MSE	1,5253	2,9704	2,2844	0,0006
MAPE	0,0600	0,1231	0,2293	Inf
MAE	1,0451	1,2814	1,4021	0,0164
SMAPE	0,0589	0,1219	0,2037	1,0574
MASE	1,0451	1,2814	1,4021	0,0164

Cuadro 6: Medidas de error de “Extra Tree Regression”

En las figuras 4, 5 y 6, se muestran comparativas entre las salidas reales (en azul) y las predichas por los modelos (en rojo). El eje “Y” representa el valor de Hidrógeno, mientras que el eje “X” representa cada una de las muestras del conjunto de datos de validación (20%) de cada “cluster”.

Finalmente en la tabla 9, se muestran los resultados en términos de MAE para un único modelo global (sin fase de clustering), para cada uno de



Figura 4: Salida real vs. predicción de “Extra Tree Regressor”

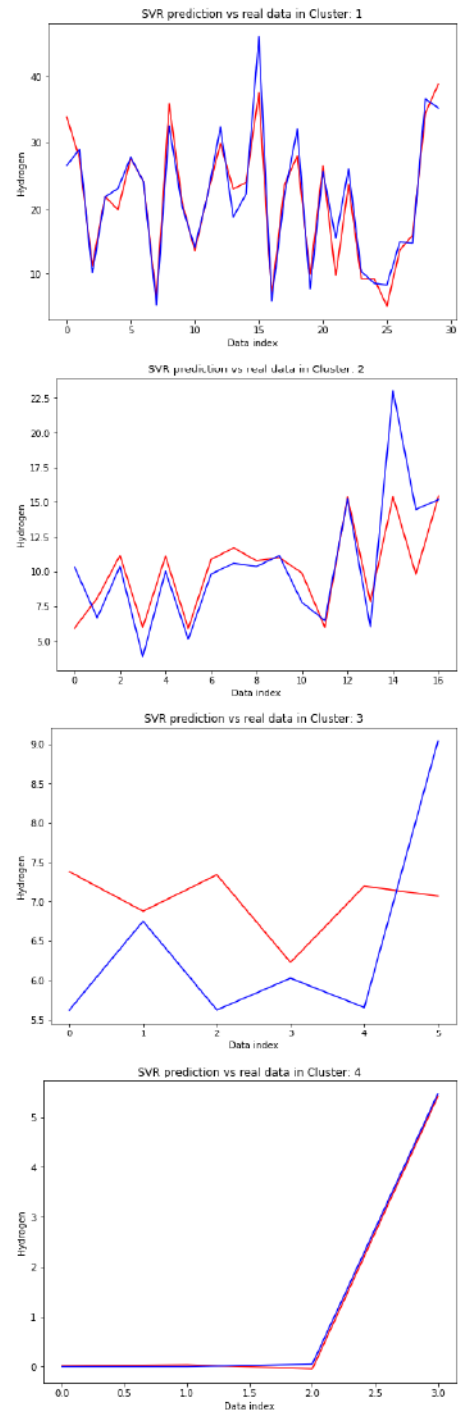


Figura 5: Salida Real vs. predicción de SVR

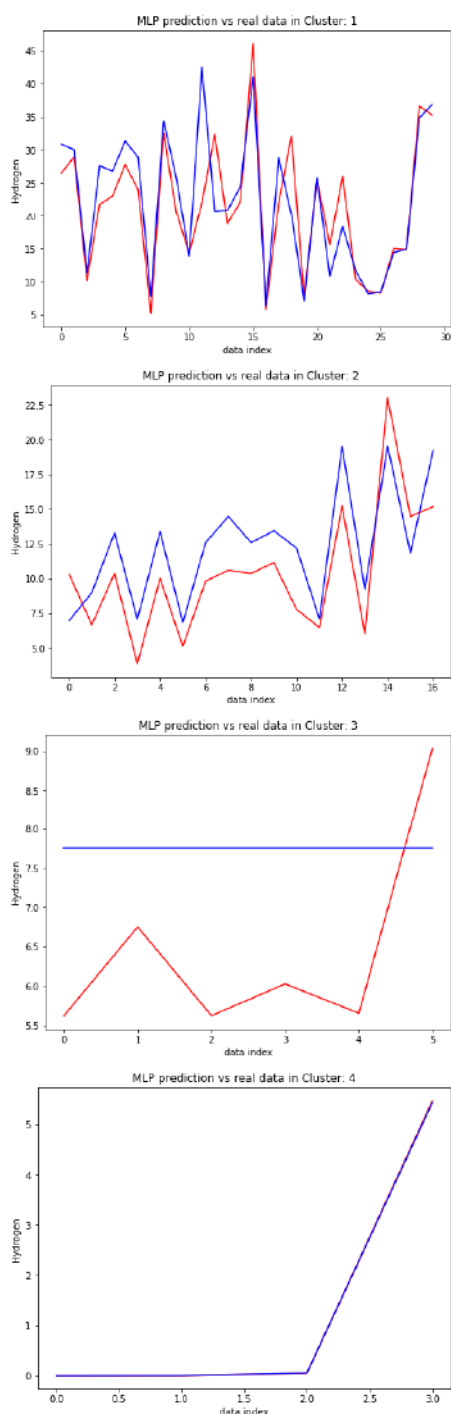


Figura 6: Salida real vs. predicción de MLP

Cluster	1	2	3	4
LMLS	1,1093	0,8485	0,6212	0,0018
MSE	9,2785	6,9840	2,0567	0,0036
MAPE	0,0640	0,0273	0,0281	Inf
MAE	2,2364	1,7867	1,2184	0,0522
SMAPE	0,1221	0,1833	0,1784	1,5023
MASE	2,2364	1,7867	1,2184	0,0522

Cuadro 7: Medidas de error de SVR

Cluster	1	2	3	4
LMLS	1,6354	1,6509	0,9118	0,0001
MSE	33,6201	9,7880	3,2027	0,0002
MAPE	0,1768	0,3291	0,2851	Inf
MAE	3,8177	2,9828	1,7331	0,0103
SMAPE	0,1672	0,2839	0,2492	1,0237
MASE	3,8177	2,9828	1,7331	0,0103

Cuadro 8: Medidas de error de MLP para regresión

los algoritmos usados (ETR, MLP y SVR) y para la mejor combinación de los modelos generados con fase de “clustering”. En el caso de la combinación de modelos, la medición del MAE se ha obtenido de dos formas distintas:

- Media Simple (MS): Cada “cluster” tiene el mismo peso en el cálculo final del error, por lo que el MAE final ha sido calculado como la suma de los MAE de cada “cluster” (C1, C2, C3, C4), entre el número total de grupos.
- Media ponderada (MP): Cada “cluster” tiene un peso en cálculo final del error. Dicho peso es proporcional al tamaño del grupo, C1 (53,7%), C2 (30%), C3 (10,5%) y C4 (5,8%), por lo que el valor final del MAE es calculado como la suma ponderada de los MAE de cada “cluster”.

Model	MAE
MODELOS GLOBAL ETR	2,195
MODELO GLOBAL SVR	2,203
MODELO GLOBAL MLP	2,467
MODELO LOCAL (MS)	0,888
MODELO LOCAL (MP)	1,074

Cuadro 9: Comparación de los modelos globales y locales en términos de MAE

En base al MAE obtenido (siendo esta la medida de error más representativa de entre las empleadas) en las tablas 6, 7, 8 y 9, se puede concluir que la combinación de diferentes modelos mejora de forma significativa los resultados obtenidos usando un sólo modelo, en concreto la mejor combinación de modelos corresponde con el uso de ETR



para los “clusters” 1 (MAE: 1,0451) y 2 (MAE: 1,2814), SVR para el “cluster” 3 (MAE: 1,2184) y MLP par el “cluster” 4 (MAE: 0,0103). En el caso del “cluster” 4, el pequeño tamaño del mismo es uno de los motivos del bajo valor de MAE obtenido.

Finalmente, el uso de un sólo modelo para el conjunto completo de datos obtiene resultados mucho peores independientemente de la técnica usada (ETR, SVR y MLP), por lo que la combinación de modelos de cada “cluster” se presenta como una alternativa que ofrece mejores resultados.

## 7. CONCLUSIONES Y TRABAJOS FUTUROS

En este artículo, se ha desarrollado un sistema híbrido de modelado consistente en la combinación de técnicas de “clustering” y modelado basado en regresión en un conjunto de datos real para la predicción de la generación de  $H_2$  del hidrolizador de una pila de combustible. Obteniéndose como resultado que dicho modelo híbrido mejora los modelos tradicionales de regresión.

En vista de los resultados obtenidos, es posible extraer las siguientes conclusiones: primero, el modelo híbrido propuesto ha superado el problema propuesto en la investigación. Segundo, se ha demostrando que la combinación de técnicas de “clustering” junto con técnicas de regresión mejoran los modelos de predicción tradicionales consistentes únicamente en una fase de modelado por regresión. También cabe destacar que en el conjunto de datos usado se han obtenido errores significativamente menores en un “cluster” respecto a los demás, principalmente motivado por el reducido tamaño del mismo, lo que facilita el generar un modelo mucho más sencillo y preciso.

Finalmente, se propone como trabajos futuros la aplicación de otras técnicas de modelado combinadas con técnicas de “clustering”, con objeto de generalizar la aplicabilidad del modelo híbrido para obtener modelos más sencillos y precisos.

### English summary

### HYBRID MODEL IMPLEMENTATION FOR THE PERFORMANCE PREDICTION OF A FUEL CELL BASED ON HYDROGEN, EMPLOYED IN ENERGY STORAGE

#### Abstract

*Currently, largely due to the rise of the*

*electric vehicle, energy storage systems are becoming a greater need, being both electric batteries and fuel cells, the two most developed technologies in recent years. However, it is not enough just to develop energy storage system, but it is essential to maximize the efficiency of them, in order to take the maximum advantage of the stored energy. To reach this goal, one of the most relevant aspects is to predict with enough accuracy and in advance both the generation and consumption of energy that will be made on the storage device. For this reason, the present research focuses on the development of a hybrid system for modeling a fuel cell using unsupervised learning techniques for clustering combined with regression techniques for modeling. Finally, the models generated on a real dataset, coming from an experimental real generation and storage system of energy by means of a hydrogen cell, are validated obtaining highly satisfactory results.*

**Keywords:** Fuel cell, Energy storage, SVM, ANN, BHL.

### Referencias

- [1] Massoud Amin. Smart grid. *Public Utilities Fortnightly*, 2015.
- [2] Léon Bottou. Large-scale machine learning with stochastic gradient descent. In *Proceedings of COMPSTAT'2010*, pages 177–186. Springer, 2010.
- [3] Lars Buitinck, Gilles Louppe, Mathieu Blondel, Fabian Pedregosa, Andreas Mueller, Olivier Grisel, Vlad Niculae, Peter Prettenhofer, Alexandre Gramfort, Jaques Grobler, Robert Layton, Jake VanderPlas, Arnaud Joly, Brian Holt, and Gaël Varoquaux. API design for machine learning software: experiences from the scikit-learn project. In *ECML PKDD Workshop: Languages for Data Mining and Machine Learning*, pages 108–122, 2013.
- [4] José-Luis Casteleiro-Roca, José Luis Calvo-Rolle, Juan Albino Méndez Pérez, Nieves Roqueñí Gutiérrez, and Francisco Javier de Cos Juez. Hybrid intelligent system to perform fault detection on bis sensor during surgeries. *Sensors*, 17(1):179, 2017.

- [5] José Luis Casteleiro-Roca, Juan Albino Méndez Pérez, Andrés José Piñón-Pazos, José Luis Calvo-Rolle, and Emilio Corchado. Modeling the electromyogram (emg) of patients undergoing anesthesia during surgery. In *10th International Conference on Soft Computing Models in Industrial and Environmental Applications*, pages 273–283. Springer International Publishing, 2015.
- [6] Youngjun Cho and Hazim B. Awbi. A study of the effect of heat source location in a ventilated room using multiple regression analysis. *Building and Environment*, 42(5):2072–2082, 2007.
- [7] Andrew Dicks and David Anthony James Rand. *Fuel cell systems explained*. Wiley Online Library, 2018.
- [8] Kaibo Duan, S Sathiya Keerthi, and Aun Neow Poo. Evaluation of simple performance measures for tuning svm hyperparameters. *Neurocomputing*, 51:41–59, 2003.
- [9] Bruce Dunn, Haresh Kamath, and Jean-Marie Tarascon. Electrical energy storage for the grid: a battery of choices. *Science*, 334(6058):928–935, 2011.
- [10] Luis Alfonso Fernandez-Serantes, Juan Aurelio Montero-Sousa, Jose Luis Casteleiro-Roca, Xosé Manuel Vilar-Martinez, and Jose Luis Calvo-Rolle. Gestión de almacenamiento energético para instalaciones de generación-distribución. *DYNA Ingeniería e Industria*, 92(2):140–141, 2017.
- [11] Oscar Fontenla-Romero and Jose Luis Calvo-Rolle. Artificial intelligence in engineering: past, present and future. *DYNA*, 93(4):350–352, 2018.
- [12] Ramon Ferreiro Garcia, JosÃ© Luis Calvo Rolle, Javier Perez Castelo, and Manuel Romero Gomez. On the monitoring task of solar thermal fluid transfer systems using nn based models and rule based techniques. *Engineering Applications of Artificial Intelligence*, 27:129 – 136, 2014.
- [13] A. Géron. *Hands-On Machine Learning with Scikit-Learn and TensorFlow: Concepts, Tools, and Techniques for Building Intelligent Systems*. O’Reilly Media, 2017.
- [14] Pierre Geurts, Damien Ernst, and Louis Wehenkel. Extremely randomized trees. *Machine learning*, 63(1):3–42, 2006.
- [15] M H. Ghaseminezhad and A. Karami. A novel self-organizing map (SOM) neural network for discrete groups of data clustering. *Appl. Soft Comput.*, 11(4):3771–3778, June 2011.
- [16] Jose M Gonzalez-Cava, José Antonio Rebozo, José Luis Casteleiro-Roca, José Luis Calvo-Rolle, and Juan Albino Méndez Pérez. A novel fuzzy algorithm to introduce new variables in the drug supply decision-making process in medicine. *Complexity*, 2018, 2018.
- [17] Peter J Hall and Euan J Bain. Energy-storage technologies and electricity generation. *Energy policy*, 36(12):4352–4355, 2008.
- [18] RunZhi Jin, KyuMan Cho, ChangTaek Hyun, and MyungJin Son. Mra-based revised CBR model for cost prediction in the early stage of construction projects. *Expert Systems with Applications*, 39(5):5214–5222, 2012.
- [19] Esteban Jove, Jose M. Gonzalez-Cava, José Luis Casteleiro-Roca, Juan Albino Méndez Pérez, José Luis Calvo-Rolle, and Francisco Javier de Cos Juez. An intelligent model to predict ani in patients undergoing general anesthesia. In Hilde Pérez García, Javier Alfonso-Cendón, Lidia Sánchez González, Héctor Quintián, and Emilio Corchado, editors, *International Joint Conference SOCO’17-CISIS’17-ICEUTE’17 León, Spain, September 6–8, 2017, Proceedings*, pages 492–501. Springer International Publishing, 2018.
- [20] Hirushie Karunathilake, Kasun Hewage, Walter Mérida, and Rehan Sadiq. Renewable energy selection for net-zero energy communities: Life cycle based decision making under uncertainty. *Renewable energy*, 130:558–573, 2019.
- [21] Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *CoRR*, abs/1412.6980, 2014.
- [22] Tomohiro Kuwae and Masakazu Hori. Global environmental issues. *Blue Carbon in Shallow Coastal Ecosystems: Carbon Dynamics, Policy, and Implementation*, 2019.
- [23] D. Martínez-Rego, O. Fontenla-Romero, and A. Alonso-Betanzos. Efficiency of local models ensembles for time series prediction. *Expert Syst. Appl.*, 38(6):6884–6894, June 2011.
- [24] Juan Aurelio Montero-Sousa, Jose Luis Casteleiro-Roca, and Jose Luis Calvo-Rolle. The electricity sector since its inception until



- the second world war. *DYNA*, 92(1):43–47, 2017.
- [25] Juan-Aurelio Montero-Sousa, Jose-Luis Casteleiro-Roca, and Jose-Luis Calvo-Rolle. Evolution of the electricity sector after the 2nd world war. *DYNA*, 92(3):280–284, 2017.
- [26] Aitor Moreno-Fernandez-de Leceta, Jose Manuel Lopez-Guede, Leire Ezquerro Insagurbe, Nora Ruiz de Arbulo, and Manuel Graña. A novel methodology for clinical semantic annotations assessment. *Logic Journal of the IGPL*, 26(6):569–580, 2018.
- [27] Philipp Moritz, Robert Nishihara, and Michael Jordan. A linearly-convergent stochastic l-bfgs algorithm. In *Artificial Intelligence and Statistics*, pages 249–258, 2016.
- [28] J. Nieves-Acedo, I. Santos-Grueiro, and P. Garcia-Bringas. Enhancing the prediction stage of a model predictive control systems through meta-classifiers. *Dyna*, 88(3):290–298, 2013.
- [29] J Osborn, Dani Guzman, FJ de Cos Juez, AG Basden, TJ Morris, E Gendron, Timothy Butterley, RM Myers, A Guesalaga, F Sanchez Lasheras, et al. Open-loop tomography with artificial neural networks on canary: on-sky results. *Monthly Notices of the Royal Astronomical Society*, 441(3):2508–2514, 2014.
- [30] Cameron W Potter, Allison Archambault, and Kenneth Westrick. Building a smarter smart grid through better renewable energy information. In *Power Systems Conference and Exposition, 2009. PSCE'09. IEEE/PES*, pages 1–5. IEEE, 2009.
- [31] Ravi Prakash, Inder Krishnan Bhat, et al. Energy, economics and environmental impacts of renewable energy systems. *Renewable and sustainable energy reviews*, 13(9):2716–2721, 2009.
- [32] Héctor Quintián, Jose Luis Calvo-Rolle, and Emilio Corchado. A hybrid regression system based on local models for solar energy prediction. *Informatica*, 25(2):265–282, 2014.
- [33] Héctor Quintián and Emilio Corchado. Beta hebbian learning as a new method for exploratory projection pursuit. *Int. J. Neural Syst.*, 27(6):1–16, 2017.
- [34] J A Rincon, V Julian, C Carrascosa, A Costa, and P Novais. Detecting emotions through non-invasive wearables. *Logic Journal of the IGPL*, 26(6):605–617, 2018.
- [35] A Suárez Sánchez, P Riesgo Fernández, F Sánchez Lasheras, Francisco Javier de Cos Juez, and PJ García Nieto. Prediction of work-related accidents according to working conditions using support vector machines. *Applied Mathematics and Computation*, 218(7):3539–3552, 2011.
- [36] Scikit-learn. Min max scaler, 2018.
- [37] F Segovia, J M Górriz, J Ramírez, F J Martínez-Murcia, and M García-Pérez. Using deep neural networks along with dimensionality reduction techniques to assist the diagnosis of neurodegenerative disorders. *Logic Journal of the IGPL*, 26(6):618–628, 2018.
- [38] DA Sunter, PM Berkeley, and DM Kammen. City-integrated photovoltaics sustainably satisfy urban transportation energy needs. *WIT Transactions on Ecology and the Environment*, 204:559–567, 2016.
- [39] Haimin Tao, Jorge L Duarte, and Marcel AM Hendrix. Line-interactive ups using a fuel cell as the primary source. *IEEE Transactions on Industrial Electronics*, 55(8):3012–3021, 2008.
- [40] Max Wei, Shana Patadia, and Daniel M Kammen. Putting renewables and energy efficiency to work: How many jobs can the clean energy industry generate in the us? *Energy policy*, 38(2):919–931, 2010.
- [41] Michael Hereward Westbrook. *The Electric Car: Development and future of battery, hybrid and fuel-cell cars*. IET Digital Library, 2001.
- [42] Szymon Wojciechowski. A comparison of classification strategies in rule-based classifiers. *Logic Journal of the IGPL*, 26(1):29–46, 2018.
- [43] Chi-Jen Yang and Robert B Jackson. Opportunities and barriers to pumped-hydro energy storage in the united states. *Renewable and Sustainable Energy Reviews*, 15(1):839–844, 2011.



© 2019 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution CC BY-NC-SA 4.0 license (<https://creativecommons.org/licenses/by-nc-sa/4.0/deed.es>).