



ESPE
UNIVERSIDAD DE LAS FUERZAS ARMADAS
INNOVACIÓN PARA LA EXCELENCIA

Modelos híbridos basados en datos para la predicción de carga eléctrica a corto y mediano plazo

Echeverría Simancas, Gonzalo Daniel

Departamento de Eléctrica, Electrónica y Telecomunicaciones

Carrera de Ingeniería en Electrónica, Automatización y Control

Trabajo de titulación, previo a la obtención del título de Ingeniero en Electrónica, Automatización y

Control

Dr. Gordillo Orquera, Rodolfo Xavier

22 de diciembre de 2020



Document Information

Analyzed document	1_EcheverriaGonzalo_TesisVFinal.docx (D90384841)
Submitted	12/21/2020 2:36:00 PM
Submitted by	Gordillo Orquera Rodolfo Xavier
Submitter email	rxgordillo@espe.edu.ec
Similarity	2%
Analysis address	rxgordillo.espe@analysis.arkund.com

Rodolfo X. Orquera
cc. 1001510203

Sources included in the report

SA	Universidad de las Fuerzas Armadas ESPE / Tesis_BorradorV6.pdf Document Tesis_BorradorV6.pdf (D89985209) Submitted by: rxgordillo@espe.edu.ec Receiver: rxgordillo.espe@analysis.arkund.com	 5
W	URL: http://repositorio.udec.cl/bitstream/11594/2618/3/Tesis_Herramientas_deep_learning.pdf Fetched: 12/20/2020 5:19:37 PM	 1
SA	1A_Jacinto_Herrera_Raul_Trinidad_Maestria_2019.docx Document 1A_Jacinto_Herrera_Raul_Trinidad_Maestria_2019.docx (D47626213)	 1
SA	Proyecto-Demanda con RNA (08-09-2017)Rev01.docx Document Proyecto-Demanda con RNA (08-09-2017)Rev01.docx (D30534126)	 2
W	URL: https://www.researchgate.net/publication/268214157_Metodos_bayesianos_para_la_comp ... Fetched: 12/21/2020 3:27:00 PM	 1



**DEPARTAMENTO DE ELÉCTRICA, ELECTRÓNICA Y
TELECOMUNICACIONES**

**CARRERA DE INGENIERÍA EN ELECTRÓNICA, AUTOMATIZACIÓN Y
CONTROL**

CERTIFICACIÓN

Certifico que el trabajo de titulación, “**Modelos híbridos basados en datos para la predicción de carga eléctrica a corto y mediano plazo**” fue realizado por el señor **Echeverría Simancas, Gonzalo Daniel** el cual ha sido revisado y analizado en su totalidad por la herramienta de verificación de similitud de contenido; por lo tanto cumple con los requisitos legales, teóricos, científicos, técnicos y metodológicos establecidos por la Universidad de las Fuerzas Armadas ESPE, razón por la cual me permito acreditar y autorizar para que lo sustente públicamente.

Sangolquí, 22 de diciembre de 2020

Firma:

Dr. Gordillo Orquera, Rodolfo Xavier

C. C. 1001510203



**DEPARTAMENTO DE ELÉCTRICA, ELECTRÓNICA Y
TELECOMUNICACIONES**

**CARRERA DE INGENIERÍA EN ELECTRÓNICA, AUTOMATIZACIÓN Y
CONTROL**

RESPONSABILIDAD DE AUTORÍA

Yo, **Echeverría Simancas, Gonzalo Daniel**, con cédula de ciudadanía n° 1718052614, declaro que el contenido, ideas y criterios del trabajo de titulación: **Modelos híbridos basados en datos para la predicción de carga eléctrica a corto y mediano plazo** es de mi autoría y responsabilidad, cumpliendo con los requisitos legales, teóricos, científicos, técnicos, y metodológicos establecidos por la Universidad de las Fuerzas Armadas ESPE, respetando los derechos intelectuales de terceros y referenciando las citas bibliográficas.

Sangolquí, 22 de diciembre de 2020

Firma:

.....
Echeverría Simancas, Gonzalo Daniel

C.C.: 1718052614



**DEPARTAMENTO DE ELÉCTRICA, ELECTRÓNICA Y
TELECOMUNICACIONES**

**CARRERA DE INGENIERÍA EN ELECTRÓNICA,
AUTOMATIZACIÓN Y CONTROL**

AUTORIZACIÓN DE PUBLICACIÓN

Yo **Echeverría Simancas, Gonzalo Daniel**, con cédula de ciudadanía n° 1718052614, autorizo a la Universidad de las Fuerzas Armadas ESPE publicar el trabajo de titulación: **Modelos híbridos basados en datos para la predicción de carga eléctrica a corto y mediano plazo** en el Repositorio Institucional, cuyo contenido, ideas y criterios son de mi responsabilidad.

Sangolquí, 22 de diciembre de 2020

Firma:

.....
Echeverría Simancas, Gonzalo Daniel

C.C.: 1718052614

Dedicatoria

El siguiente trabajo se lo dedico a mi familia por su paciencia.

A mi padre, por siempre escucharme y darme consejos para enfrentar cualquier desafío.

A mi madre, quien con su carisma siempre me ha motivo a seguir adelante.

A mi querida hermana Jennifer, quien siempre ha estado ahí para darme una mano.

Y por supuesto, a mi hermana Valentina, así como a mi sobrina Dalianah, quienes siempre me sacan una sonrisa con sus travesuras.

De igual forma, dedico esto a mis amigos, por esa grandiosa amistad constante e incondicional.

También, a todos los estudiantes quienes están empezando a incursionar por la investigación. Es un mundo maravilloso que trae mucho conocimiento. Les motivo a seguir explorándolo y cosechar todos los frutos que puedan de las grandes investigaciones a favor de cumplir sus objetivos.

Gonzalo Echeverría

Agradecimiento

Un especial agradecimiento a mi tutor académico, Dr. Rodolfo Gordillo, por todo su conocimiento brindado y su confianza hacia mi trabajo. Su apoyo fue un factor clave para lograr cumplir con los objetivos de esta investigación.

Mi agradecimiento hacia todos mis amigos, con quienes viví inolvidables experiencias a lo largo de la carrera. Me enseñaron que, trabajando en equipo y compartiendo conocimientos, podemos lograr grandes cosas.

Por la paciencia, consejos e incondicionalidad, quiero expresar mi gratitud a toda mi familia y personas que estuvieron a mi lado al desarrollar este proyecto de investigación. Cada palabra o deseo de superación, contó para lograr cumplir este objetivo.

También, agradecer a todos los docentes de la Universidad de las Fuerzas Armadas ESPE, quienes me han ayudado a formarme profesional y éticamente para la solución de problemas de ingeniería.

Gonzalo Echeverría

Índice de Contenidos

Hoja de resultados de la herramienta Urkund	2
Certificado del director	3
Responsabilidad de autoría.....	4
Autorización de publicación.....	5
Dedicatoria.....	6
Agradecimiento.....	7
Índice de Contenidos	8
Índice de Tablas	12
Índice de Figuras	13
Resumen	15
Abstract.....	16
Capítulo I Introducción.....	17
Antecedentes	17
Justificación e Importancia.....	19
Alcance.....	21
Objetivos	22
Objetivo General	22
Objetivos Específicos.....	22
Capítulo II Fundamentación Teórica	24
Gestión de Energía y Eficiencia Energética.....	24
Predicción de Carga Eléctrica	24

Factores que Influyen en la Predicción.....	25
Variables Exógenas a la Predicción de Carga Eléctrica	25
Clasificación de Acuerdo con Horizontes de Predicción	27
STLF	27
MTLF.....	27
LTLF	27
Modelamiento predictivo	28
Terminología asociada a modelamiento predictivo	28
Tipos de modelos de predicción.....	29
Modelos Híbridos.....	30
Técnicas de IC para Modelamiento Predictivo	31
Sistemas Basados en Lógica Difusa	32
Funciones de membresía.....	34
Redes Neuronales Artificiales.....	35
Redes Neuronales Multicapa.....	37
Algoritmos de Aprendizaje	38
Algoritmo Gradient Descent y Backpropagation	39
Algoritmo Levenberg-Marquardt.....	41
Criterio de detección temprana	42
Métodos multivariantes.....	43
Análisis de Componentes Principales (PCA)	44
Modelo de Mínimos Cuadrados Parciales Ortonormales (OPLS).....	45
Técnicas de Validación de Modelos	45
Leave One Out Cross Validation	46

	10
Métricas de desempeño para modelos de predicción	47
Error Porcentual Absoluto Medio (MAPE).....	47
Porcentaje de Desviación PBIAS	47
RMSE y MSE	48
Coeficiente de Correlación	48
Capítulo III Diseño de Modelos Híbridos	49
Estudio Previo para el Diseño de Modelos	49
Descripción de Datos de Consumo Eléctrico	49
Descripción y Análisis de Modelo: PCA con AR.....	50
Subestimación y Sobreestimación de los Modelos PCA y OPLS.....	55
Análisis de Modelo ENS: Combinación Convexa entre PCA y OPLS	56
Diseño de Modelo NCVX: Combinación no Convexa entre PCA y OPLS	58
Especificación de Entradas, Salidas y MI del Sistema Difuso	61
Especificación de Funciones de Membresía para la Entrada	61
Conjunto de Reglas de Inferencia.....	62
Especificación de Funciones de Membresía para la Salida	63
Diseño de Modelos: Combinación de PCA y Algoritmos Basados en IC	64
Diseño de Modelo PCANN.....	64
Estructura del Modelo Propuesto	64
Diseño de NN.....	66
Distribución de Datos	66
Arquitectura de NN	66
Algoritmo de Entrenamiento y Criterio de Parada	68
Características de NN para Entrenamiento	68

	11
Variables del Modelo de NN.....	69
Implementación de Modelos de NN.....	71
Diseño de Modelo PCAFUZZY	72
Especificación de Entradas, Salidas y MI del Sistema Difuso	74
Especificación de Funciones de Membresía para la Entrada	76
Conjunto de Reglas de Inferencia	77
Especificación de Funciones de Salida	77
Capítulo IV Experimentos y Análisis de Resultados	79
Resultados de diseño NCVX	79
Resultados de diseño PCANN	80
Experimentación de Número de Neuronas	81
Estudio de Orden del Modelo PCANN	82
Resultados de implementación NN	84
Resultados de diseño PCAFUZZY	87
Análisis de Resultados de Predicción en STLF	88
Análisis de Resultados de Predicción: MTLF.....	93
Resultados de predicción de NCVX.....	96
Resultados de predicción de PCANN y PCAFUZZY	96
Capítulo V Conclusiones y Recomendaciones	98
Conclusiones	98
Recomendaciones	99
Referencias Bibliográficas	101
Anexos	109

Índice de Tablas

Tabla 1 Ejemplos de tipos de modelos de predicción	30
Tabla 2 Algoritmo basado en PCA.....	50
Tabla 3 Métricas de desempeño PCA y AR	54
Tabla 4 Valores relevantes de α para indicadores MAPE [%] y PBIAS [%].....	58
Tabla 5 Intervalos de tiempo en un día con modelos de mejor desempeño	59
Tabla 6 Conjuntos de reglas para sistema difuso de NCVX	62
Tabla 7 Características de las NN para entrenamiento	69
Tabla 8 Entradas de las NN	71
Tabla 9 Características de la variable estación.....	75
Tabla 10 Especificaciones para el diseño del sistema difuso en HUF	77
Tabla 11 Número de neuronas en la capa oculta para cada NN	82
Tabla 12 Resultados de entrenamiento de NN para HUF y CEA.....	85
Tabla 13 Métricas de desempeño en las semanas $W=33,41,48$ de HUF	92
Tabla 14 Métricas de desempeño en las semanas $W=33,41,48$ de CEA	92
Tabla 15 Parámetros de desempeño en MTLF para HUF y CEA	95

Índice de Figuras

Figura 1 Ejemplo de carga vs temperatura en la ciudad de Madrid	26
Figura 2 Tipos de modelos híbridos: homogéneos y heterogéneos	31
Figura 3 Estructura básica de un sistema basado en FL	32
Figura 4 Ejemplo de funciones de membresía: triangulares, trapezoidales y suavizadas	34
Figura 5 Diagrama de una neurona.....	36
Figura 6 Tipos de funciones de activación más frecuentes en NN	37
Figura 7 Ejemplo de red neuronal multicapa	38
Figura 8 Ejemplo de GD para determinar el mínimo punto en la función E	40
Figura 9 Ejemplo de curvaturas en minimización de E	41
Figura 10 Ejemplo de caso de detección temprana	43
Figura 11 Interpretación gráfica de ejemplo de PCA.....	44
Figura 12 Ilustración de LOOCV	46
Figura 13 Formación del índice M (muestras por semana)	51
Figura 14 Varianza Acumulativa (%) y Valores Singulares para HUF.	52
Figura 15 Autovectores de 1era y 2da componente principal	53
Figura 16 Ejemplo de predicción de 1era y 2da componente principal mediante AR para HUF ..	54
Figura 17 Sobreestimación en modelo PCA y subestimación en modelo OPLS.....	55
Figura 18 Comparación de modelos PCA y OPLS, en los días 10/12/2008 y 20/03/2009.....	56
Figura 19 MAPE [%] y PBIAS [%] en función de α	57
Figura 20 Periodos de consumo en el día: 20/03/2009	59
Figura 21 Diagrama esquemático de modelo NCVX.....	60
Figura 22 Sistema difuso para NCVX	61
Figura 23 Funciones de membresía para la entrada Muestra (n)	62

Figura 24 Funciones de membresía de la salida: Alpha.....	63
Figura 25 Ilustración esquemática de la estructura del modelo PCANN	65
Figura 26 Arquitectura de NN como aproximador general de funciones.....	67
Figura 27 Primera componente principal con su tendencia lineal y promedio para HUF.	70
Figura 28 NN para la primera componente principal tal como se visualiza en Matlab Mathworks.	72
Figura 29 Diagrama esquemático de PCAFUZZY	73
Figura 30 Secciones representativas de estaciones del año para HUF.....	73
Figura 31 Estructura de sistema difuso para factor β	74
Figura 32 Funciones de membresía para la entrada Muestras [n], para HUF y CEA	76
Figura 33 Conjunto de reglas para HUF, tal como se muestran en Matlab Mathworks.....	78
Figura 34 Curvas de respuesta α vs muestra, para los MI Mamdani y Takagi-Tsugeno en NCVX	79
Figura 35 Comparación parámetros de desempeño NCVX (Mamdani y Takagi-Tsugeno) para HUF y CEA	80
Figura 36 Estudio de número de neuronas en la capa oculta vs MSE vs tiempo de entrenamiento	81
Figura 37 Experimento LOOCV para cada NN con sus componentes principales (P1, ..., P5) en HUF	83
Figura 38 Predicción de componentes principales mediante PCANN para HUF y CEA	86
Figura 39 Curva de control β vs muestras para HUF y CEA.....	87
Figura 40 <i>Parámetros de desempeño semanales en el periodo de prueba de HUF y CEA</i>	88
Figura 41 Segunda componente principal y sus cambios en el comportamiento.	90
Figura 42 Predicción de carga eléctrica para las semanas W=33,41,48 en HUF y CEA.....	91
Figura 43 <i>Gráficos de predicción de carga de modelos híbridos para MTLF en HUF y CEA</i>	93

Resumen

El consumo eléctrico ha ido aumentando en los últimos años. Se requieren estrategias para gestionar eficientemente esa energía. Los modelos de predicción de carga eléctrica aportan a ello permitiendo mejorar la planificación y operación de sistemas eléctricos. En la actualidad, modelos híbridos basados en datos, son investigados con el fin de conseguir pronósticos más precisos. El presente proyecto propone el desarrollo de tres modelos híbridos con el propósito de abordar problemas de subestimación y sobreestimación, así como para mejorar las prestaciones de desempeño. El primero, consiste en la combinación no convexa entre el método de análisis de componentes principales (PCA) con modelamiento autorregresivo (AR) y mínimos cuadrados parciales ortonormales (OPLS), a través de un sistema difuso. El segundo, combina PCA y modelos basados en redes neuronales. Finalmente, el tercero agrega una variable al segundo modelo mediante un sistema difuso. Para evaluar la hipotética mejora, se comparó las prestaciones entre los modelos híbridos y los individuales. Para este fin, se utilizaron datos de dos centros de cuidado de salud y se analizó en los horizontes de tiempo de corto y mediano plazo. Además, se consideró al error porcentual absoluto medio (MAPE), raíz del error cuadrático medio (RMSE) y al porcentaje de sesgo (PBIAS) como principales indicadores de desempeño. En general, los resultados presentaron que los indicadores disminuyeron para la mayoría de los casos. Esto evidenció que la integración de modelos permitió mejorar la precisión de la predicción, con un consecuente aumento en su complejidad computacional.

- Palabras Clave:

- **MODELOS HÍBRIDOS BASADO EN DATOS**
- **PREDICCIÓN DE CARGA ELÉCTRICA**
- **REDES NEURONALES**
- **SISTEMAS DIFUSOS**

Abstract

Electricity consumption has been increasing in recent years. Strategies are required to efficiently manage that energy. Electrical load forecasting models contribute to this by improving planning and operation of electrical systems. Currently, the hybridization of data driven models, is being researched to achieve more accurate forecasts. The present project proposes the development of three hybrid models to address problems of under and overestimation, as well as to improve performance. The first one, consists of the non-convex combination between the principal component analysis (PCA) with autoregressive modeling (AR) and orthonormal partial least squares (OPLS), using a fuzzy logic system. The second one, combines PCA and models based on neural networks. Finally, the third model incorporates a variable to the second one through a fuzzy logic system. To test the hypothetical improvement, the performance of the hybrid and individual models were compared. For this purpose, data from two health care centers were used and analyzed in short and medium term. Furthermore, mean absolute percentage error (MAPE), root mean square error (RMSE) and percentage bias (PBIAS) were the main performance indicators. In general, the results showed that these metrics decreased for most of the cases. This showed that the hybridization of models allowed to improve the precision of the forecasting, with a consequent increase in its computational complexity.

- Keywords:

- **DATA DRIVEN HYBRID MODELS**
- **LOAD FORECASTING**
- **NEURAL NETWORKS**
- **FUZZY LOGIC SYSTEMS**

Capítulo I

Introducción

En este primer capítulo se exponen los antecedentes, justificación e importancia, así como el alcance y los objetivos del presente proyecto de investigación. Estos servirán como introducción para el entendimiento de los modelos propuestos.

Antecedentes

Con frecuencia, la toma de decisiones de un ser humano se basa en la información y experiencia adquirida en el pasado. Pensamientos sobre lo que podría ocurrir en el futuro, permiten a un individuo prepararse en el presente. En resumen, se intenta predecir lo que sucederá para mejorar la toma de decisiones (Kuhn & Johnson, 2016). La predicción de carga eléctrica se basa en los mismos principios. Los datos del comportamiento de carga predichos pueden ser de gran ayuda para establecer políticas de consumo o mejorar la gestión de energía (Zhang & Hong, 2019).

En los últimos años, el volumen de datos ha aumentado notablemente. Información proveniente de experimentos, simulaciones y registros históricos, cada vez son más frecuentes. Por esta razón, la necesidad de algoritmos eficientes para su manejo ha conllevado a un progreso en la ciencia basada en datos (*data-driven science*). Esta toma importancia en el estudio de sistemas complejos, permitiendo revolucionar la manera de modelar, predecir y controlar dichos sistemas (Brunton & Kutz, 2019).

Consecuentemente, el avance de esta ciencia y las técnicas de inteligencia computacional (IC), permiten que modelos basados en datos sean efectivos para pronosticar la carga eléctrica utilizando registros históricos de consumo (Cai, Pipattanasomporn, & Rahman, 2019).

El desarrollo de los modelos requiere que se definan los horizontes de tiempo a predecir. Generalmente, pronosticar periodos cortos corresponde al objetivo de buscar una mayor eficiencia y confiabilidad en la operación de sistemas de energía (Hong, Li, & Fan, 2019). Por otra parte, los periodos medianos o largos toman importancia en planificación de mantenimiento de equipos, análisis de requerimientos futuros de capacidad, unidades de generación de energía, entre otros (Ganguly, Kalam, & Zayegh, 2017; Anwar, Sharma, Chakraborty, & Sirohia, 2018).

En la actualidad, existen diversos tipos de modelos para la predicción de carga. Entre ellos, se encuentran los híbridos, los cuales, combinan técnicas para proveer nuevos algoritmos tendientes a mejorar el desempeño del pronóstico (Hong, Li, & Fan, 2019). En los últimos años este campo de estudio continúa en desarrollo por sus beneficios en la predicción.

En el Ecuador, la previsión de carga eléctrica es un tema que ha ido tomando importancia. Específicamente, se desarrolló un proyecto para la Empresa Eléctrica Ambato Regional Centro Norte S.A. (EEASA) que se dedica a la generación, transmisión y distribución de energía eléctrica. Se implementaron técnicas de aprendizaje automático para el modelamiento con el fin de mejorar la predicción. Los resultados demostraron que se cumplieron los objetivos con el uso de estos algoritmos (Guamán, 2019).

Asimismo, en la Universidad de las Fuerzas Armadas ESPE (UFA-ESPE) se llevó a cabo un proyecto que consistió en la implementación de un sistema de monitorización de consumo eléctrico y la predicción de carga a corto plazo para un laboratorio de la universidad (Llumiquinga, 2018). Se obtuvieron resultados favorables y se recomendó el uso de variables exógenas para la predicción. Sin embargo, cabe recalcar que en ninguna de las investigaciones mencionadas se implementaron modelos híbridos.

Adicionalmente, en la UFA-ESPE en conjunto con otras instituciones, se desarrolló una investigación que analiza y predice el consumo de carga eléctrica en centros de cuidado de salud. Para ello, se utilizó el método de análisis de componentes principales (PCA, por sus siglas en inglés, *principal component analysis*) con modelamiento autorregresivo (AR) y el esquema supervisado que utiliza mínimos cuadrados parciales ortonormales (OPLS, por sus siglas en inglés, *orthonormal partial least squares*), así como la combinación convexa entre ambos métodos. Los resultados obtenidos fueron satisfactorios, no obstante, recomendaron la idea de que métodos más avanzados, como algoritmos basados en IC o híbridos, podrían ser útiles para mejorar los resultados, motivando su uso para futuras investigaciones (Gordillo-Orquera et al., 2018).

Justificación e Importancia

Los modelos de predicción de carga brindan distintos beneficios, que en general, están alineados a objetivos de incrementar la eficiencia en sistemas energéticos, permitiendo un ahorro en costos y energía (Dong, Zhang, & Hong, 2018). Un pronóstico de carga brinda ventajas tanto a generadores como a consumidores de electricidad. La información obtenida, permite la planificación estratégica entre lo que se produce y consume (Jiménez, Donado, & Quintero, 2017). Adicionalmente, se afirma que la habilidad para predecir el comportamiento y la demanda de energía será parte de la inteligencia que requerirán las redes inteligentes (*smart grids*) (De Felice & Yao, 2011). Por lo tanto, el presente proyecto toma importancia y cuenta como aporte a investigaciones relacionadas a la gestión de energía (*energy management*) y eficiencia energética.

Para el pronóstico de carga eléctrica, existen modelos tradicionales que utilizan representaciones matemáticas que se pueden volver muy complejas al existir relaciones no lineales en la información (Ganguly, Kalam, & Zayegh, 2017). Consecuentemente, algoritmos

que utilizan técnicas de IC han logrado mejorar la exactitud del pronóstico mediante aproximaciones no lineales. Sin embargo, en la actualidad, se ha conseguido mejorar más el desempeño de la predicción aplicando modelos híbridos, que permiten tomar las ventajas de varios modelos individuales (tradicionales o inteligentes) con el fin de obtener resultados más satisfactorios (Cheng, Hong, Sheng, & Huang, 2016; Li, Geng, Hong, & Zhang, 2018). La precisión de la predicción toma gran relevancia en el ámbito económico. De acuerdo con Bunn & Farmer (1985), un incremento del 1% en el error, puede generar millones de dólares en costos de operación.

La importancia de este proyecto se basa en el desarrollo de modelos híbridos que permitan mejorar las prestaciones de desempeño obtenidas por los modelos desarrollados en la investigación previa por Gordillo-Orquera et al. (2018). Se aprovechan las características obtenidas proyectándose a desarrollar modelos de baja complejidad y pretendiendo disminuir errores en la predicción, tales como subestimación y sobreestimación de datos. Esto se logrará combinando los métodos utilizados en aquella investigación con algoritmos basados en IC. Cabe recalcar que, la presente investigación plantea el desarrollo de dos algoritmos nuevos.

El incremento del consumo de energía eléctrica y el requerimiento de garantizar un suministro, han llevado a la necesidad de modelos de predicción confiables (Jiménez, Donado, & Quintero, 2017). El consumo de energía tiene un gran efecto en el progreso socioeconómico y tecnológico de un país (Khan, Mahmood, Safdar, Khan, & Khan, 2016). En el Ecuador no es la excepción. De acuerdo con CENACE (2019), en el año 2018 se registró un aumento en la demanda energética en 4.57% con respecto al año 2017, por lo cual, el uso eficiente de energía va tomando más trascendencia. Se recalca que proyecto aportará conocimiento científico orientado a la planificación y uso eficiente de energía.

Alcance

El presente proyecto de investigación se centra en el diseño de modelos combinados o híbridos basados en datos. Para lo cual, se tomarán como base los resultados obtenidos con los modelos desarrollados por Gordillo-Orquera et al. (2018), así como fuente primaria de información a la misma. Se aprovecharán las características del método no supervisado de análisis de componentes principales (PCA) y esquemas supervisados utilizando mínimos cuadrados parciales ortonormales (OPLS) con el fin de realizar combinaciones añadiendo algoritmos de inteligencia computacional.

Los datos que se utilizarán corresponden al consumo eléctrico real de un edificio de salud y de un centro de cuidados de salud. Los registros fueron adquiridos por mediciones muestreados uniformemente a lo largo de 8 años en el Hospital Universitario de Fuenlabrada (HUF) y de 4 años en el Centro de Especialidades El Arroyo (CEA), los cuales se ubican en Madrid, España. Este tipo de lugares se caracterizan por tener un alto consumo.

Se plantean dos escenarios para el desarrollo de estos nuevos modelos híbridos. En primer lugar, se desarrolla un nuevo modelo que permita realizar una combinación no convexa entre los dos métodos mencionados previamente (PCA y OPLS), mediante el uso de un controlador basado en lógica difusa.

El desarrollo del presente modelo se lo realiza diseñando el sistema difuso, para el cual, se definen sus funciones de membresía basándose en los datos de entrada existentes, logrando así aprovechar las ventajas de ambos predictores.

En segundo lugar, se desarrolla un modelo basado en las componentes principales de la predicción, considerando solo los resultados obtenidos por el modelo no supervisado PCA (con un total de 5 componentes principales). El cual, se combina con algoritmos de inteligencia computacional, en específico, se evaluará con sistemas difusos y redes neuronales artificiales.

Las componentes principales del modelo son representadas por auto vectores y serán las entradas al nuevo modelo, además se experimentará con variables exógenas que se puedan conseguir mediante la distribución de tiempo (como las estaciones del año o días festivos) con el objetivo de conseguir un buen pronóstico. La salida será la predicción de la carga.

Ambos modelos de predicción serán evaluados con métricas de desempeño como MAPE, RMSE y PBIAS, las cuales permitirán conocer si los métodos mejoran las prestaciones de predicción. Cabe recalcar que, se evaluarán ambos modelos en términos de corto y mediano plazo.

Al tratarse de modelos basados en *Machine Learning*, la validación se la realiza utilizando en general dos técnicas: la partición de los datos en modo entrenamiento (*Training*) y prueba (*Test*), así como la selección del orden del modelo utilizando validación cruzada (LOOCV, por sus siglas en inglés, *Leave-one-out cross-validation*). Se buscará que ambos modelos sean de baja complejidad con el fin de que puedan ser de uso general para escenarios similares.

Objetivos

Objetivo General

Desarrollar algoritmos híbridos basados en datos e inteligencia computacional para la predicción de consumo eléctrico a corto y mediano plazo a fin de mejorar las prestaciones de desempeño.

Objetivos Específicos

- Diseñar un modelo que permita aprovechar las características de los métodos no supervisado y supervisado mediante una combinación no convexa, implementada por un controlador basado en lógica difusa.

- Desarrollar un segundo modelo combinando las componentes principales del método PCA y considerar variables exógenas temporales para incorporarlas a un algoritmo inteligente.
- Aplicar los algoritmos desarrollados a la predicción de datos históricos de consumo de carga a corto y mediano plazo.
- Evaluar las prestaciones obtenidas en los modelos diseñados mediante métricas de desempeño utilizadas en pronósticos de carga eléctrica.

Capítulo II

Fundamentación Teórica

En este capítulo se presentan los principios teóricos necesarios para la comprensión de los modelos desarrollados. Inicialmente, se abordan los conceptos de gestión de energía y predicción de carga eléctrica. Seguidamente, se detallan definiciones relacionadas a modelamiento predictivo, modelos híbridos, técnicas de IC para modelamiento y métodos multivariantes. Finalmente, se mencionan las métricas de desempeño utilizadas.

Gestión de Energía y Eficiencia Energética

El alto consumo de energía eléctrica ha presentado diversos desafíos para la industria energética. Diversas tecnologías se han ido involucrando, proyectando a tener sistemas integrados e inteligentes (Zhang, Shah, & G. Papageorgiou, 2013). Esto va alineado a los objetivos de la gestión de energía, cuyo fin es proveer el servicio eléctrico con el menor costo e impacto posible al medio ambiente (EL-Shimy, 2018). La gestión de energía consiste en actividades como monitoreo, análisis, control y redirección de energía para el uso eficiente de esta (Schulze et al., 2015).

Las nuevas tecnologías han surgido como estrategia para conseguir eficiencia energética. Sin embargo, el comportamiento de los consumidores es un elemento que se debe considerar. Corrigiendo hábitos de consumo, se puede aportar en gran medida al ahorro de energía (Mahapatra et al, 2018).

Predicción de Carga Eléctrica

La predicción de carga eléctrica es una técnica que permite estimar la carga por adelantado, ayudando a conocer la futura demanda energética. Esto toma gran importancia en actividades relacionadas a la gestión de sistemas energéticos (Ganguly, Kalam, & Zayegh, 2017; Anwar, Sharma, Chakraborty, & Sirohia, 2018).

Por lo tanto, conociendo un pronóstico de carga, las industrias generadoras de energía eléctrica y los otros participantes, pueden tomar decisiones, planificar, analizar y realizar mantenimientos con el fin de mejorar la eficiencia energética (Anwar, Sharma, Chakraborty, & Sirohia, 2018).

Factores que Influyen en la Predicción

El consumo de energía eléctrica depende de diversos factores. Estos permiten describir el comportamiento de carga del consumidor. En general se deben tomar en cuenta los factores de tiempo y del clima (Feinberg & Genethliou, 2005).

Los factores de tiempo incluyen época del año, día de la semana y hora del día. Por ejemplo, el consumo en la noche será mayor que el consumo en el día, debido a la necesidad de iluminación. Por otro lado, los factores climáticos involucran variables como la temperatura, humedad y sensación térmica. Que en conjunto permiten determinar el efecto de las estaciones del año (Ganguly, Kalam, & Zayegh, 2017; Gordillo-Orquera et al., 2018). Esto toma importancia en regiones donde el consumo de energía se ve afectado por estas estaciones. Por ejemplo, en invierno se esperaría un mayor consumo de energía debido al uso de sistemas de aire acondicionado.

Variables Exógenas a la Predicción de Carga Eléctrica

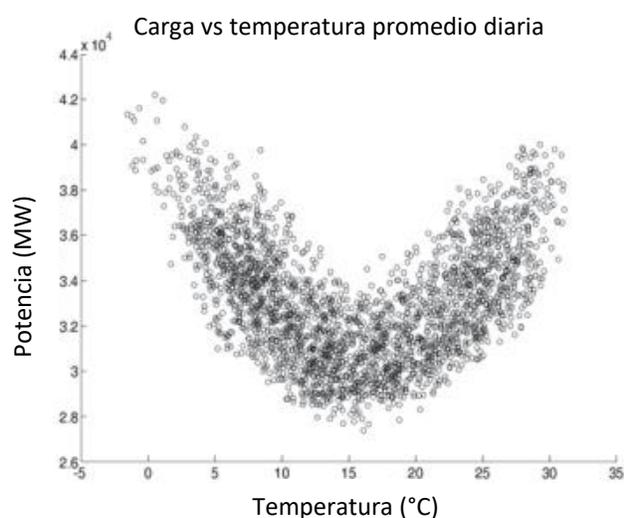
En el mismo contexto, las variables exógenas (VE) pueden ser temperatura, humedad, Producto Interno Bruto (PIB), población, así como otros parámetros que sean ajenos a la predicción (Salas, 2016). Al considerar estas variables se puede conseguir un mejor desempeño en el modelo. Por ejemplo, Taylor & Buizza (2003) determinaron que la variable temperatura, tiene una gran importancia en modelos de predicción a corto y mediano plazo. Un ejemplo de

esto, se muestra en la Figura 1, cuyos datos corresponden a la ciudad de Madrid, donde se evidencia que valores bajos y altos de temperatura causan un aumento en la carga.

De acuerdo con Salas (2016), las VE se pueden dividir climáticas, de calendario y económicas. Las climáticas coinciden con los factores climáticos mencionados anteriormente. Las de calendario coinciden con los factores de tiempo y las variables económicas suelen incluir al PIB (Producto Interno Bruto).

Figura 1

Ejemplo de carga vs temperatura en la ciudad de Madrid



Nota. Obtenido de “Empirical Comparison of Neural Network and Auto-Regressive Models in Short-Term Load Forecasting” (p. 144) por López, Sans, Valero, & Senabre, 2018, Short-Term Load Forecasting by Artificial Intelligent Technologies, MDPI.

También se suele considerar a la variable población, que consiste en datos demográficos de alguna localidad. La selección de las VE depende generalmente de los horizontes de predicción. Usualmente, para periodos cortos y medios se consideran variables históricas de carga, factores climáticos y de tiempo. Mientras que, para periodos largos, se consideran variables de población y PIB (Feilat, Al-Sha’abi, & Momani, 2017).

Clasificación de Acuerdo con Horizontes de Predicción

En términos de tiempo, la predicción de carga eléctrica puede ser dividida en tres grandes categorías: predicción de carga a corto plazo (STLF, por sus siglas en inglés, *short term load forecasting*), mediano plazo (MTLF, por sus siglas en inglés, *mid term load forecasting*) y largo plazo (LTLF, por sus siglas en inglés, *long term load forecasting*). Estos se denominan horizontes de predicción y se los selecciona con base en las necesidades del pronóstico, tales como planificación requerida, mantenimiento, operación, costos, entre otros (Gordillo-Orquera et al., 2018; Anwar, Sharma, Chakraborty, & Sirohia, 2018).

STLF

La STLF consiste en periodos de tiempo cuyos rangos pueden estar en términos de minutos, horas, días o semanas. Este tipo de predicción se suele utilizar para lograr reducir costos de operación, planificación a corto plazo y, además, permite tomar decisiones para evitar sobrecargas (Anwar, Sharma, Chakraborty, & Sirohia, 2018; Hong, Li, & Fan, 2019).

MTLF

La MTLF tiene un rango de tiempo que puede ir desde un mes a uno o dos años. Con la información obtenida se puede decidir cuándo realizar mantenimiento de equipos, planificar pruebas, acuerdo de precios y determinar el tiempo de interrupción de plantas o equipos. Los métodos de análisis utilizados para el MTLF son similares a los de STLF (Abu-Shikhah, Elkarmi, & Aloquili, 2011; Gordillo-Orquera et al., 2018).

LTLF

El LTLF comprende predicciones que van de un año en adelante, permitiendo estimar los requerimientos futuros de capacidad y planificar expansiones de equipos para determinadas áreas (Ammar, Sulaiman, Fateh, & Mohamad, 2018).

Modelamiento predictivo

El modelamiento predictivo consiste en encontrar el modelo más adecuado para predecir con precisión una salida o el comportamiento de datos (Geisser, 1993). El modelamiento se solía realizar basándose en principios básicos o empíricos. Sin embargo, la alta complejidad de los datos cambió este paradigma. En la actualidad, los modelos basados en datos han sido preferidos por su facilidad de diseño con base en registros históricos disponibles (Brunton & Kutz, 2019).

El modelamiento basado en datos utiliza técnicas estadísticas y de aprendizaje automático (ML, por sus siglas en inglés *machine learning*) logrando abarcar una gran variedad de métodos (Bishop, 2006).

Terminología asociada a modelamiento predictivo

En los modelos de predicción de datos, existe una terminología muy común y útil para el entendimiento del proceso de diseño. El significado de estos términos se resume en la siguiente lista:

- Las *muestras (samples)* son unidades de datos obtenidos a partir de registros históricos (Kuhn & Johnson, 2016).
- En referencia a lo mencionado anteriormente, los *datos de entrenamiento (training set)* son utilizados para obtener los parámetros del modelo, mientras que, los *datos de prueba y validación (test and validation set)* sirven para evaluarlo (Bourdeau et al., 2019).
- Los *predictores*, son los datos usados como entradas al modelo, para obtener la salida deseada, mientras que los datos objetivos (*targets*) son los datos reales de los registros históricos.

- Las *variables categóricas* representan valores que permiten categorizar la información. Por ejemplo, utilizar valores binarios para diferenciar los datos (1- se cumple, 0- no se cumple) (Kuhn & Johnson, 2016).
- El *preprocesamiento* de datos es un procedimiento que se realiza con el fin de mantener una misma escala en estos, lo cual, agiliza el entrenamiento de los modelos. Entre las técnicas más comunes está la normalización que consiste en adecuar los datos entre los valores 0 y 1, así como la centralización de la media, que consigue que el promedio de los datos sea 0.
- El sobreajuste (*overfitting*), es un efecto que ocurre cuando sobreentrenamos el modelo, obteniendo bajos indicadores de error para el periodo de entrenamiento, pero altos valores cuando se prueba con nuevos datos (Ying , 2019).

Tipos de modelos de predicción

De acuerdo con Cheng, Hong, Sheng, & Huang (2016), “los algoritmos se pueden clasificar en tres grandes categorías: tradicionales, modernos e híbridos o combinados”. Los tradicionales se basan en herramientas estadísticas, los modernos en inteligencia computacional (IC) y los híbridos en la combinación de modelos existentes (Hong, Li, & Fan, 2019; Cheng, Hong, Sheng, & Huang, 2016). En la Tabla 1, se muestran algunos ejemplos de los tipos de modelos mencionados.

Tabla 1

Ejemplos de tipos de modelos de predicción

Tipo de modelo	Ejemplos
Tradicionales	Modelos autorregresivos (como ARIMA, ARMAX), regresión multivariable y aproximación de suavizado exponencial.
Modernos inteligentes	Redes neuronales artificiales o perceptrón multicapa (MLP), sistemas difusos, máquinas de vector de soporte (SVM).
Híbridos	ARIMA-SVM, ARIMA-MLP, híbridos basados en redes neuronales, entre otros.

Nota. Adaptado de “A Hybrid Seasonal Mechanism with a Chaotic Cuckoo Search Algorithm with a Support Vector Regression Model for Electric Load Forecasting” (p. 24), por Dong, Zhang, & Hong, 2018, *Energies*, MDPI.

Modelos Híbridos

Los autores Feinberg & Genethliou (2005), mencionan que el desarrollo y progreso de herramientas matemáticas permiten la mejora de la precisión en las predicciones. Con base en esta idea surgieron los modelos híbridos, cuyo objetivo es aprovechar las ventajas de otros modelos.

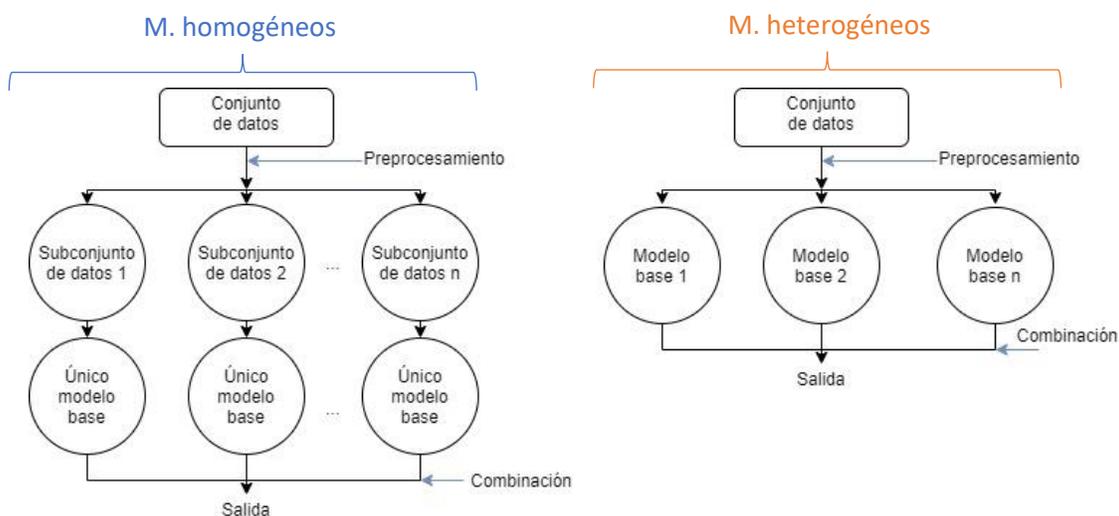
Conceptualmente los modelos combinados o híbridos (*ensembled or hybrid models*) representan la combinación de modelos existentes, denominados modelos base, que proveen nuevos algoritmos tendientes a mejorar el desempeño de la predicción (Li, Geng, Hong, & Zhang, 2018).

Muchos de estos modelos combinan métodos lineales y no lineales. Además, en los últimos años, las investigaciones se orientan en la combinación de métodos tradicionales, algoritmos de ML y técnicas de descomposición (Mohan, Soman, & Kumar, 2018).

De acuerdo con Wang & Srinivasan (2017), los modelos híbridos se clasifican generalmente en dos categorías, homogéneos y heterogéneos. En la Figura 2, se muestra cada uno de ellos. Los homogéneos, distribuyen los datos en subconjuntos manteniendo el mismo algoritmo de entrenamiento, mientras que, los heterogéneos, mantienen un solo conjunto de datos para entrenar distintos algoritmos.

Figura 2

Tipos de modelos híbridos: homogéneos y heterogéneos



Nota. Adaptado de "A review of artificial intelligence-based building energy use prediction: Contrasting the capabilities of single and ensemble prediction models" (p. 8), Wang & Srinivasan, 2017, Renewable and Sustainable Energy Reviews.

Técnicas de IC para Modelamiento Predictivo

Anteriormente, se mencionó que los modelos híbridos pueden incorporar técnicas de IC. En esta subsección se detallan los métodos de IC convenientes para el entendimiento del presente proyecto. Específicamente se dará una introducción de sistemas difusos y redes neuronales artificiales.

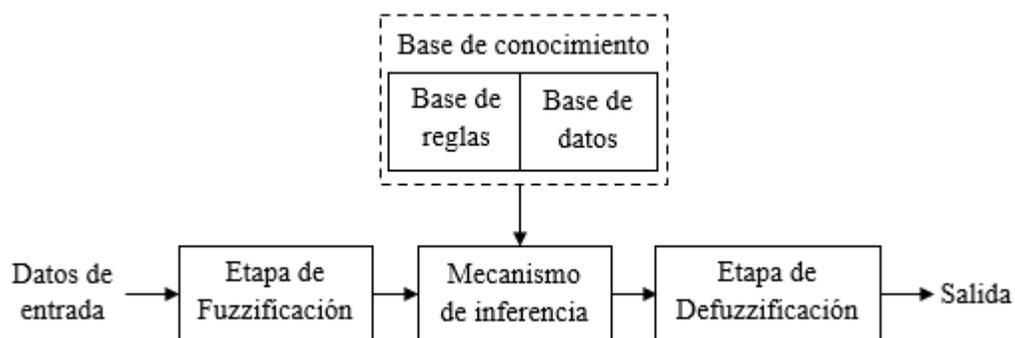
Sistemas Basados en Lógica Difusa

La lógica matemática convencional permite trabajar únicamente con dos valores de verdad, lo cual se considera una debilidad para ciertos problemas. Esto se puede solucionar mediante la lógica difusa (FL, por sus siglas en inglés, *fuzzy logic*), que permite introducir numerosos valores entre 0 y 1. Con esto, se puede definir el grado de pertenencia de un elemento a un conjunto. Cabe mencionar que, las ventajas de la FL han hecho que sea muy utilizada en la ingeniería de control y tecnologías de automatización (Ertel, 2017).

Los sistemas basados en FL o sistemas difusos suelen ser utilizados para resolver problemas asociados a sistemas complejos e inciertos, permitiendo direccionar la imprecisión de las entradas y salidas mediante conjuntos difusos (Kamal & Ibrahim, 2018; Khettab, Bensafia, Bourouba, & Azar, 2018). En general, la estructura básica de un sistema basado en FL consiste en algunas etapas, tal como se ilustra en la Figura 3.

Figura 3

Estructura básica de un sistema basado en FL



Nota. Adaptado de Estructura de un Sistema Basado en Reglas Difusas, Diciembre Sanahuja, 2017, Universidad Jaime I.

Con base en esta estructura se procederá a explicar el funcionamiento de estos sistemas. Inicialmente se deben establecer las variables de entrada y salida, así como sus universos de discurso, es decir, sus posibles valores.

La *etapa de fuzzificación* realiza el proceso de mapear los datos de entrada en valores definidos por un conjunto difuso, el cual está representado mediante funciones de membresía o pertenencia (Guo & Wong, 2013).

El *mecanismo de inferencia* (MI) o también denominado sistema de inferencia difuso, permitirá la interpretación de las reglas SI-ENTONCES (IF-THEN) para obtener un valor en la salida del sistema. Los MI comúnmente utilizados son *Mamdani* y *Takagi-Tsugeno* (Diciembre Sanahuja, 2017). El primero, es muy conocido dado a su facilidad en formalizar e interpretar, mientras que el segundo se utiliza cuando se requiere mayor eficiencia computacional y precisión. Este último se diferencia del primero dado a que su salida es una función constante o lineal que representa al problema (Jassbi, Alavi, Serra, & Ribeiro, 2007).

La *base de conocimiento* brindará información que sea representativa para resolver el problema, esta se puede obtener basándose en el conocimiento de expertos. Para lograrlo, se deben definir la *base de reglas* y la *base de datos*. La primera contiene las reglas difusas que serán evaluadas por el sistema, también se denomina conjunto de reglas de inferencia, mientras que la segunda almacena las funciones de membresía (Diciembre Sanahuja, 2017).

Para el caso de *Mamdani*, la *etapa de defuzzificación* realiza lo opuesto a la *fuzzificación*, es decir, produce la salida del sistema a partir de las funciones de membresía definido para esta. El método de Defuzzificación más conocido es el del centroide, que matemáticamente determina el centro de gravedad de la función. Para *Takagi-Tsugeno*, esta etapa consiste en evaluar las funciones constantes o lineales definidas para la salida (Guo & Wong, 2013).

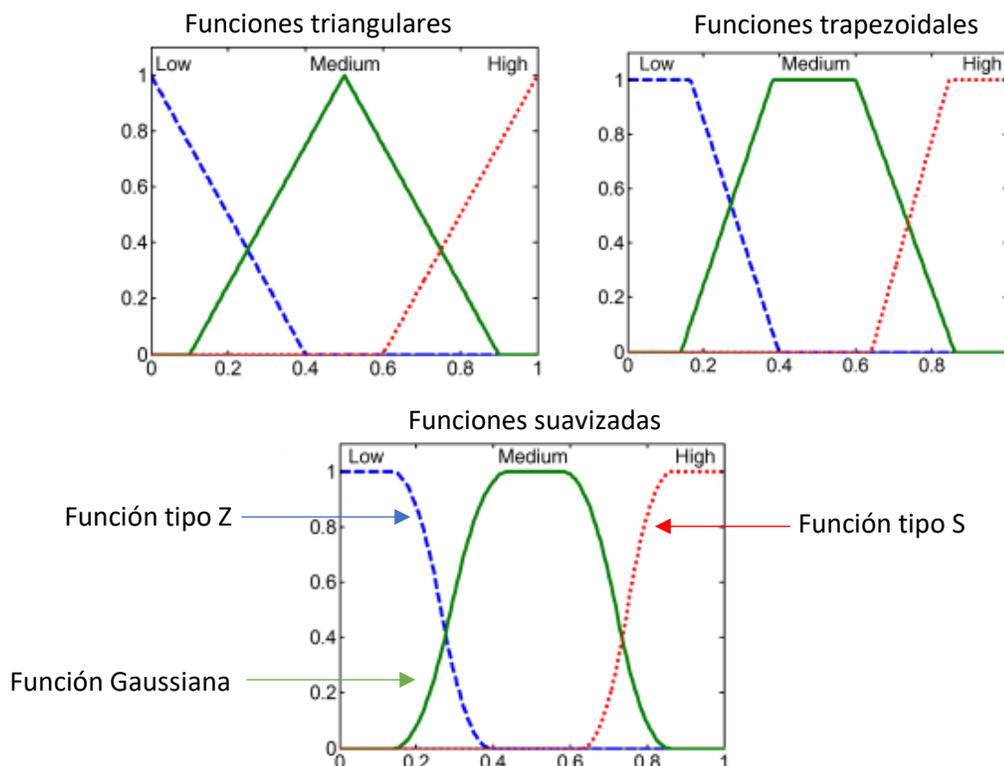
Funciones de membresía

Las funciones de membresía comúnmente utilizadas son la triangular, trapezoidal, fuzzificadores Gaussianos, función tipo S y función tipo Z. Estas permiten determinar el grado de pertenencia de las entradas hacia un conjunto difuso (Guo & Wong, 2013).

En la Figura 4, se ilustra un ejemplo, donde se ubican tres gráficos que representan las funciones de membresía triangulares, trapezoidales y suavizadas. Se puede observar claramente la diferencia entre cada una de ellas. Cabe mencionar que, para cada gráfico, se tienen tres funciones, correspondientes a un único término (*Low*, *Medium* o *High*). En particular, el gráfico de las funciones suavizadas contiene tres tipos de funciones: tipo Z, tipo S y Gaussiana.

Figura 4

Ejemplo de funciones de membresía: triangulares, trapezoidales y suavizadas



Nota. Adaptado de *Examples of common fuzzy membership functions*, de Keller, Liu, & Fogel, 2016, IEEE.

De acuerdo con Acosta (2006), los diseñadores escogen las funciones de membresía con base en su experiencia y conocimiento del problema. Las funciones más complejas, como las suavizadas, generalmente se usan para problemas de mayor complejidad.

Redes Neuronales Artificiales

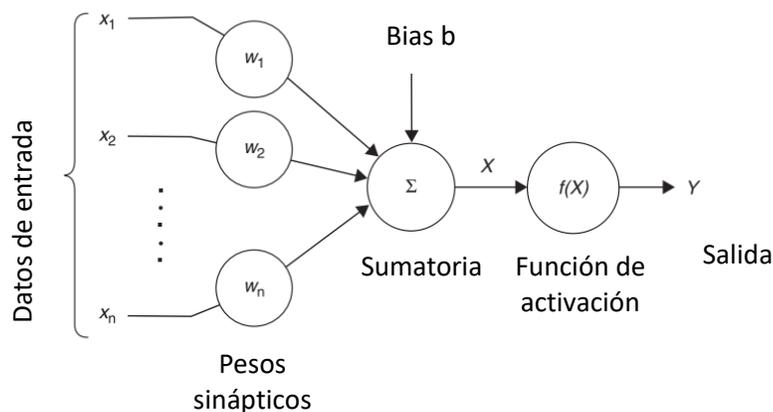
Las redes neuronales artificiales usualmente llamadas únicamente como redes neuronales (NN, por sus siglas en inglés, *neural networks*), consisten en un modelo computacional que simula el mecanismo de aprendizaje de organismos biológicos. De manera específica, se basan en las células del sistema nervioso, las neuronas, que se conectan entre sí mediante unas regiones denominadas sinapsis (C. Aggarwal, 2018).

Las NN presentan un modelo de aprendizaje computacional que permite realizar tareas relacionadas al ML. Una de sus principales características es que pueden ser utilizadas como aproximadores universales permitiendo resolver una gran variedad de problemas, tales como predicción, modelamiento y clasificación (Guo & Wong, 2013).

En este contexto, una neurona es una función matemática que puede tener x_i , para $i = 1, \dots, n$ entradas y a cada una de ellas, se le aplica un peso sináptico denotado por w_i que posteriormente pasarán a una sumatoria donde se añade el *bias* o sesgo b . El diagrama esquemático de una neurona se ilustra en la Figura 5 (C. Aggarwal, 2018; Guo & Wong, 2013).

Figura 5

Diagrama de una neurona



Nota. Adaptado de Diagrama de una neurona (p. 27), por Guo & Wong, 2013, The Hong Kong Polytechnic University.

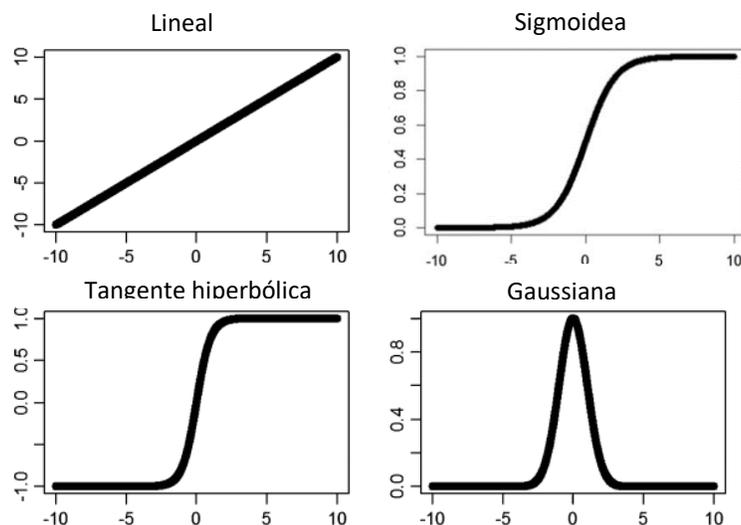
Este diagrama corresponde a un perceptrón que se considera uno de los ejemplos más sencillos de NN. Consta de una capa de entrada y una de salida considerada también como capa de cómputo, dado que contiene una función de activación. Además, al observar el diagrama, es sencillo notar que las entradas con su respectivo peso sináptico entran en una sumatoria para obtener la variable X , la cual pasa por una función de activación $f(X)$ y así obtener la salida Y (Guo & Wong, 2013). Esto también se evidencia en la ecuación (1).

$$Y = f\left(\sum_{i=1}^n x_i w_i + b\right) \quad (1)$$

Existe una gran variedad de funciones de activación, cuyo objetivo en general es limitar la sumatoria observada en la ecuación (1), por lo cual, esta función se diseña con base en un valor umbral (Ertel, 2017). Entre las diversas funciones existentes, las más comúnmente utilizadas se pueden visualizar en la Figura 6, las cuales son: función lineal (*purelin*), sigmoidea (*logsig*), tangente hiperbólica (*tansig*) y Gaussiana.

Figura 6

Tipos de funciones de activación más frecuentes en NN



Nota. Adaptado de Tipos de funciones de activación alternativas (p. 385), Dinov, 2018, Springer.

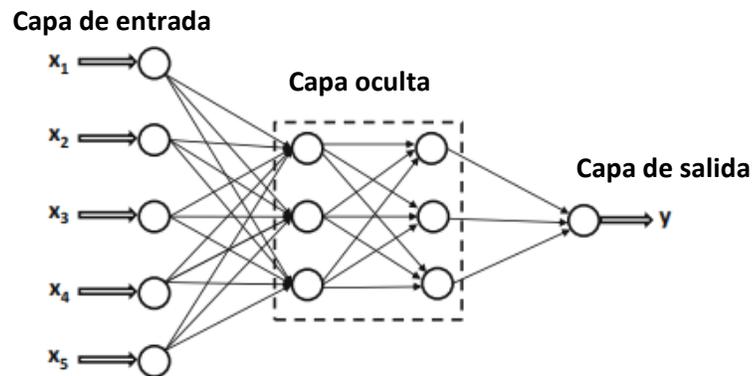
Dinov (2018) afirma que “se pueden seleccionar una de estas funciones basandose en el codominio” (p. 385), es decir, los valores que podría tomar la función ante una respectiva entrada, como por ejemplo, el caso de *logsig*, que evaluandose puede estar entre 0 y 1.

Redes Neuronales Multicapa

También denominadas perceptrón multicapa, son redes que contienen más de una capa de cómputo, es decir existen capas intermedias entre la entrada y salida denominadas capas ocultas (*hidden layers*). Debido a la arquitectura sucesiva donde cada capa “alimenta” a la siguiente en dirección desde la entrada hacia la de salida, estas redes también se denominan *Feed Forward Networks* (FFN) (C. Aggarwal, 2018). Un ejemplo específico de FFN, se puede observar en la Figura 7.

Figura 7

Ejemplo de red neuronal multicapa



Nota. Adaptado de The basic architecture of a feed-forward network with two hidden layers and a single output layer (p. 18), Aggarwal, 2018, Springer.

Este ejemplo cuenta con cinco entradas y una salida, además, se dice que es una red de dos capas, haciendo notar que la capa de entrada no cuenta, dado que no realiza ningún cómputo como en las otras. Cabe mencionar que existen diversas arquitecturas que pueden combinar más capas ocultas, diversas salidas, *bias* para cada capa, entre otros. Esto se define dependiendo de la complejidad del problema (C. Aggarwal, 2018).

Algoritmos de Aprendizaje

Para las NN, los algoritmos de aprendizaje permiten calcular los parámetros de la red, es decir, los pesos sinápticos y *bias* que la componen. Existen dos tipos de aprendizaje, el no supervisado y el supervisado. El no supervisado considera únicamente los datos de entrada para identificar patrones o características y así realizar el aprendizaje. Mientras que el supervisado, requiere de una muestra de datos de entrada con su correspondiente salida para lograr el aprendizaje. En aplicaciones de regresión, se utilizan los algoritmos supervisados (Guo & Wong, 2013).

Existen diversos algoritmos para el entrenamiento de las NN. En general, se los puede clasificar en dos grandes grupos: de primer orden y de segundo orden. Para el propósito del presente proyecto solo se mencionarán los algoritmos: *Backpropagation* y *Levenberg Marquardt*, de 1er y 2do orden, respectivamente.

Como consideraciones previas, estos métodos están orientados a mejorar el indicador llamado Error Cuadrático Medio (MSE, por sus siglas en inglés, *Mean Squared Error*), el cual, se multiplica por $\frac{1}{2}$ con el fin de agilizar cálculos posteriores. Además, el número total de muestras se denotará como N y el número de iteraciones o épocas como p .

Algoritmo Gradient Descent y Backpropagation

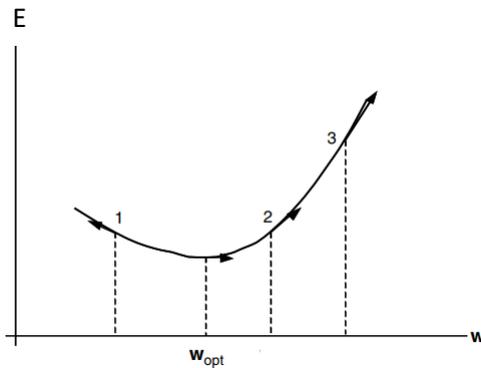
El algoritmo *backpropagation* de entrenamiento se utiliza para aplicaciones de regresión y utiliza métodos como el algoritmo Descenso de Gradiente (GD, por sus siglas en inglés, *gradient descent*) o sus variantes, para estimar los parámetros que minimizan una función. Para el caso de las NN, estima los pesos sinápticos que permitan minimizar el MSE. La ecuación (2) se utiliza por el cálculo del MSE en el algoritmo GD (Samarasinghe, 2007).

$$E = MSE_{BP} = \frac{1}{2N} \sum_{i=1}^N (y_i - \hat{y}_i)^2 \quad (2)$$

Donde y_i denotan los datos reales (*targets*) para la muestra i -ésima y \hat{y}_i los datos obtenidos en la salida de la red. Por otro lado, los pesos de la NN w_p (para la época p) y los *bias* b_p tendrán una inicialización aleatoria y se irán actualizando en el proceso de entrenamiento. El algoritmo GD, utiliza la primera derivada de la superficie del MSE en función de los pesos y *bias*. Por ejemplo, en la Figura 8 se ilustra como el algoritmo utiliza la derivada, es decir, la pendiente en un punto, para determinar el valor de w que permitan que la función decrezca, hasta llegar a un valor mínimo u óptimo.

Figura 8

Ejemplo de GD para determinar el mínimo punto en la función E



Nota. Adaptado de La pendiente a través de la superficie del error (p. 141), por Samarasinghe, 2007, Taylor & Francis Group.

Dado a que solo se utiliza la primera derivada, se considera un algoritmo de primer orden. La ecuación (3) permite describir el algoritmo GD para determinar y actualizar los pesos de una NN que minimizan la función.

$$\Delta w_p = -\alpha \frac{\partial E}{\partial w_p} \quad (3)$$

Donde α se denomina coeficiente de aprendizaje y determina el tamaño de los pasos de actualización de los parámetros hacia el punto óptimo. En general se le asigna un valor entre 0 y 1. Cabe recalcar que, una ecuación similar se aplica para los *bias de la NN*.

El algoritmo *backpropagation* está orientado a redes multicapa. Este método utiliza la derivada del indicador de error con respecto a los pesos de salida y la derivada con respecto a los pesos de la capa oculta para después concatenar el proceso mediante la regla de la cadena (Samarasinghe, 2007).

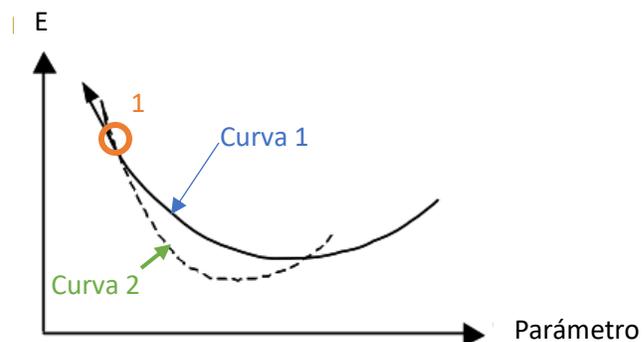
Algoritmo Levenberg-Marquardt

El algoritmo *Levenberg-Marquardt* (LMA, por sus siglas en inglés, *Levenberg-Marquardt algorithm*) permite un entrenamiento rápido utilizando técnicas de optimización numéricas. Se considera de 2do orden por usar a su beneficio la segunda derivada del indicador de error con respecto a los parámetros de la red.

Recalcando que, la primera derivada determina el cambio del error con respecto a los parámetros (pendiente), la segunda determinará el cambio de la pendiente con respecto a estos. De forma que, se puede escoger la curva que permita llegar al punto mínimo más velozmente (Samarasinghe, 2007). En la Figura 9 se ilustra un ejemplo de esta idea, comparando la curva 1 y 2, se observa que a partir del punto 1, se puede escoger cualquiera de las dos para llegar a un mínimo, sin embargo, solo una de ellas permitirá llegar más rápido.

Figura 9

Ejemplo de curvaturas en minimización de E



Nota. Adaptado de Dos superficies de error con diferente curvatura (p. 172), por Samarasinghe, 2007, Taylor & Francis Group.

Para LMA, la actualización de los pesos de la red es diferente al descrito anteriormente. La ecuación (4) permite determinar los pesos de la NN. Se observa que esta no tiene un coeficiente de aprendizaje, pero incluye un nuevo parámetro λ .

$$\Delta w_p = -\frac{d_p}{d_p^s + e^\lambda} \quad (4)$$

Para esta ecuación las derivadas son denotadas como: $d_p = \partial E / \partial w_p$, $d_p^s = \partial^2 E / \partial w_p^2$.

El término e^λ permite estabilizar el proceso evitando que el error aumente (Lv et al., 2018; Samarasinghe, 2007). Para modelos de 2do orden es común representar los cálculos en forma matricial. Por ello, se adopta el término matriz Hessiana (H) para representar las 2das derivadas, obteniendo la ecuación reescrita (5).

$$\Delta w_p = -\frac{d_p}{H + e^\lambda I} \quad (5)$$

Por otro lado, la matriz Hessiana, puede ser aproximada por (6)

$$H = J^T J \quad (6)$$

Donde J representa la matriz Jacobiana. Mediante este reemplazo, la actualización de los pesos se puede describir con la ecuación (7).

$$\Delta w_p = -(J^T J + e^\lambda I)^{-1} J^T E \quad (7)$$

Aplicando este algoritmo, se logran obtener los pesos sinápticos de la red, así como los *bias*. Para poder aplicar en MLP, se procede a emplear los principios de *backpropagation* consiguiendo así, calcular todos los parámetros de la NN (Lv et al., 2018).

Criterio de detección temprana

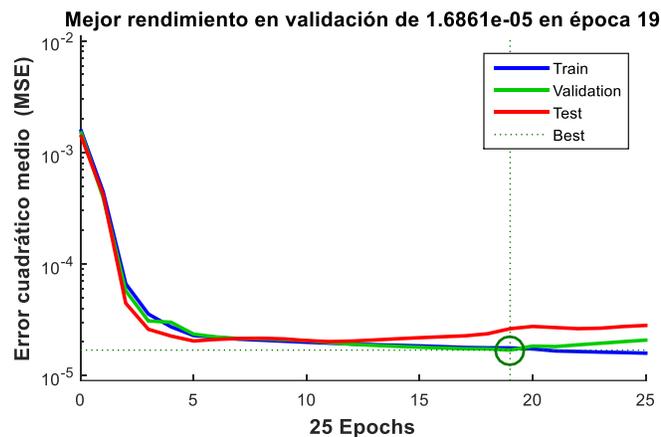
El criterio de detección temprana (ES, por sus siglas en inglés, *early stopping*) permite mejorar la generalización de un modelo. Se lo aplica en el entrenamiento de la NN y es una de las estrategias que permiten reducir el sobreajuste (Demuth & Beale, 2004).

Para aplicar este método, es necesario reservar un conjunto de datos para validación. Es decir, los datos originales se deben dividir en 3 subconjuntos: entrenamiento, validación y prueba. Los datos de validación permitirán aplicar el criterio ES.

Al inicio del entrenamiento del modelo el indicador MSE empezará a reducir. Sin embargo, este comenzará a aumentar en algún punto indicando la existencia de sobreajuste. En este punto ES permitirá detener el entrenamiento y conservará los parámetros del menor MSE obtenido para este subconjunto (Demuth & Beale, 2004).

Figura 10

Ejemplo de caso de detección temprana



En la Figura 10, se puede visualizar este comportamiento. Se grafica el MSE de los 3 subconjuntos de datos y su evolución para cada iteración (épocas). En este caso, en la época 19 se logró el mejor MSE para validación, por lo que se escogerán los parámetros de entrenamiento obtenidos en este punto. Se observa también que a partir de ahí el MSE de validación empieza a crecer indicando un posible sobreajuste, por lo que el proceso de entrenamiento solo llegó a hacerse hasta la época 25.

Métodos multivariantes

Los métodos multivariantes permiten reducir la dimensionalidad de un conjunto de datos, logrando utilizarlos más eficientemente. Para ellos, se extrae la información más relevante y se evade la más redundante (Jolliffe & Cadima, 2016). En esta subsección se mencionan dos métodos significativos para este proyecto.

Análisis de Componentes Principales (PCA)

El método de análisis de componentes principales PCA extrae información relevante de un conjunto de datos permitiendo describir su estructura interna en forma de variables ortogonales, las cuales, se denominan componentes principales. Se dice que se ha reducido la dimensionalidad cuando el número de características obtenidas es menor al de variables originales (Muñoz-Romero, 2015; Gordillo-Orquera et al., 2018).

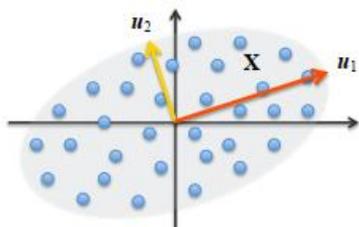
PCA es un método *no supervisado*, es decir, solo toma en cuenta los datos de entrada que se disponen. Además, se basa en autovectores y trata de encontrar las direcciones con máxima varianza en los datos. Cada una de las componentes principales debe recoger la varianza más alta posible y ser ortogonal a las anteriores (Muñoz-Romero, 2015).

Por ejemplo, el autor Muñoz-Romero (2015) presenta un caso de un conjunto bidimensional X de datos, cuyos autovectores y autovalores son: u_1, u_2 y λ_1, λ_2 respectivamente. Los autovectores indican las direcciones de la mayor varianza, mientras que los autovalores presentan su magnitud. Con base en esto se determinan las componentes principales z_1 y z_2 , tal como se ilustra en la Figura 11.

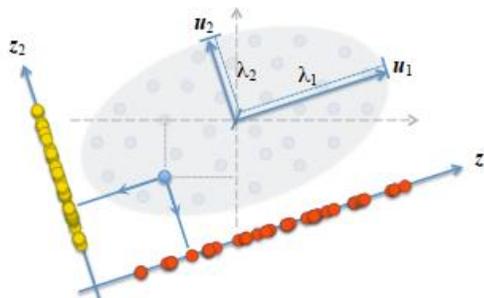
Figura 11

Interpretación gráfica de ejemplo de PCA

Direcciones de máxima varianza



Proyección de datos al espacio de características



Nota. Obtenido de Interpretación Gráfica del PCA, Muñoz-Romero, 2015, Universidad Carlos III de Madrid.

Como se observa los datos de color rojo están proyectados en la componente z_1 , mientras que los de color amarillo en z_2 . En resumen, el método consiste en obtener una matriz que mapea los datos de entrada X en el espacio de características, para posteriormente reducir la dimensionalidad y recuperar los datos, trayendo estos al espacio original. Una de las técnicas que suelen ser utilizadas al aplicar PCA, es la de descomposición por valores singulares (SVD, por sus siglas en inglés, *singular value decomposition*).

Modelo de Mínimos Cuadrados Parciales Ortonormales (OPLS)

El método de los Mínimos Cuadrados Parciales Ortonormales (OPLS) es un método supervisado que toma en cuenta los datos disponibles de entrada y salida. El método OPLS se suele utilizar para problemas de regresión o clasificación (Muñoz-Romero, 2015).

Normalmente en el procedimiento para aplicar el método se determinan las matrices de auto covarianza, covarianza cruzada y se calcula una matriz que contiene los vectores de proyección la cual permitirá determinar las características mediante para posteriormente resolver un problema de optimización. Cabe recalcar, que el objetivo de OPLS es determinar los vectores que proyecten una mejor aproximación de los datos de salida minimizando el error cuadrático medio (Muñoz-Romero, Arenas-García, & Gomez-Verdejo, 2013).

Técnicas de Validación de Modelos

Anteriormente se mencionó la validación simple, que permitía entrenar el modelo con un subconjunto de datos y realizar su evaluación con otro subconjunto. Sin embargo, existen otras técnicas, como es el caso de la validación cruzada (CV, por sus siglas en inglés, *cross validation*). Este tipo de validación consiste en dividir los datos disponibles en un número determinado de subconjuntos, para posteriormente escoger uno de ellos para la prueba del modelo y los sobrantes para el entrenamiento. Esto se lo realiza con cada uno de los subconjuntos repetidamente. Por ejemplo, se pueden dividir los datos en 5 grupos, los primeros

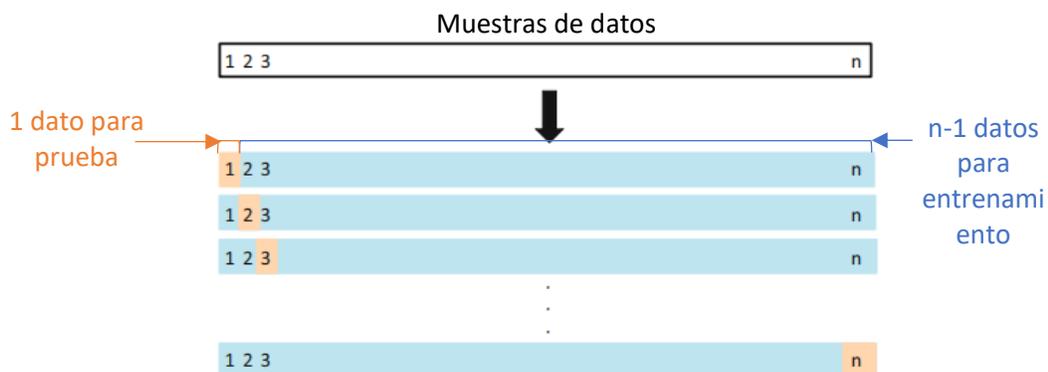
cuatro se utilizan para entrenar el modelo y el quinto para probarlo, después, se utilizan los grupos 1,2,3,5 para entrenar y el 4to para probarlo, así repetidamente hasta que todos los grupos hayan sido utilizados como prueba del modelo (Reunanen, 2003). Este método también es conocido como *k-fold cross validation* y tiene sus variaciones, de las cuales, una de ellas se menciona a continuación.

Leave One Out Cross Validation

El método de validación cruzada dejando uno fuera (LOOCV, por sus siglas en inglés, *leave one out cross validation*) consiste en separar los datos, donde se define un grupo para entrenamiento que contendrá todas las muestras excepto una, la cual, se utilizará para la prueba del modelo y este proceso se repite tantas veces como número de muestras se tengan disponibles (James, Witten, Hastie, & Tibshirani, 2017).

Figura 12

Ilustración de LOOCV



Nota. Adaptado de A schematic display of LOOCV (p. 179), por James, Witten, Hastie, & Tibshirani, 2017, Springer.

La Figura 12 presenta una ilustración donde se tienen n muestras, y se aplica LOOCV iterativamente para cada una de ellas, terminado el proceso con la muestra n . Para cada iteración se calcula el indicador MSE, para finalmente obtener el promedio de este, tal como se muestra en la ecuación (8)

$$MSE_{prom} = \frac{1}{n} \sum_{i=1}^n MSE_i \quad (8)$$

De acuerdo con James, Witten, Hastie, & Tibshirani (2017), una de las desventajas de LOOCV es su alto costo computacional, dado que se debe entrenar el modelo n veces. Cabe destacar que esta estrategia se puede usar de manera general para los modelos predictivos.

Métricas de desempeño para modelos de predicción

Existen distintas métricas orientadas a evaluar los modelos de predicción de carga eléctrica. Para el presente, se mencionarán indicadores MAPE (%), PBIAS (%) y RMSE (kW), que fueron utilizados para evaluar la predicción. Asimismo, se describe el MSE y el coeficiente de correlación que fueron otros indicadores útiles para entrenamiento de los modelos. Con el fin de entender las ecuaciones, se denotará a n como el número total de muestras, y como los datos reales de consumo y \hat{y} a los datos obtenidos de la predicción.

Error Porcentual Absoluto Medio (MAPE)

El Error Porcentual Absoluto Medio (MAPE, por sus siglas en inglés, *mean absolute percentage error*), ver ecuación (9), es un indicador muy utilizado para la medida de desempeño de modelos de predicción de consumo eléctrico (Myttenaere, Golden, Le Grand, & Rossi, 2016).

$$MAPE(\%) = \frac{100}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \quad (9)$$

Porcentaje de Desviación PBIAS

Otro indicador utilizado es el porcentaje de sesgo (PBIAS, por sus siglas en inglés *percentage of BIAS*), ver ecuación (10), el cual permite conocer si existe subestimación (valores positivos de PBIAS) o sobreestimación (valores negativos) de datos (Gordillo-Orquera et al., 2018).

$$PBIAS(\%) = 100 * \frac{\sum_{i=1}^n (y_i - \hat{y}_i)}{\sum_{i=1}^n y_i} \quad (10)$$

RMSE y MSE

La raíz del error cuadrático medio (RMSE, por sus siglas en inglés, *root mean square error*), ver ecuación (11) y el MSE (por sus siglas en inglés, *mean square error*), ver ecuación (12), son indicadores que permiten estimar la calidad del modelo. Una de las ventajas del RMSE es que tiene la misma escala de los datos (Hyndman & Koehler, 2006). Por otro lado, el MSE es muy utilizado en algoritmos de entrenamiento para modelos (James, Witten, Hastie, & Tibshirani, 2017).

$$RMSE(kW) = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (11)$$

$$MSE(kW^2) = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (12)$$

Coefficiente de Correlación

El coeficiente de correlación (R), generalmente, se utiliza en los modelos de predicción con el fin de determinar la correlación que existe entre la salida del modelo y los datos reales (*targets*). Si el valor de R es cercano a 1, correspondería a que el modelo se ajusta bien a los datos reales (Demuth & Beale, 2004).

Capítulo III

Diseño de Modelos Híbridos

Inicialmente, en este capítulo se detalla un estudio específico de la investigación previa. Por consiguiente, se presentan los diseños de los modelos propuestos, denominados NCVX, PCANN y PCAFUZZY. Asimismo, se mencionarán los métodos de IC utilizados para cada caso. Cabe mencionar que, estos fueron implementados usando *Neural Network Toolbox*, *Fuzzy Logic Toolbox* y otras funcionalidades del *software Matlab Mathworks*.

Estudio Previo para el Diseño de Modelos

Para el diseño de los modelos planteados, surgió la necesidad de realizar un análisis de los datos y resultados de la investigación previa. Específicamente, se estudian los registros reales de consumo eléctrico, las características de los modelos implementados, así como los parámetros de desempeño obtenidos en la predicción.

Descripción de Datos de Consumo Eléctrico

De acuerdo con Gordillo-Orquera et al., (2018), los datos de consumo eléctrico fueron registrados cada 15 minutos. La información pertenece a los centros de cuidado de salud HUF y CEA, ubicados en la provincia de Madrid, España. El conjunto de datos de HUF fue registrado en las fechas comprendidas entre el 01 de septiembre del 2008 y el 31 de agosto del 2016 (8 años). De esta información, se utilizaron 7 años como datos de entrenamiento, es decir, desde el 01 de septiembre del 2008 hasta el 31 de agosto del 2015. Mientras que, para los datos de prueba se tomó un año, comprendido entre el 01 de septiembre del 2015 al 31 de agosto del 2016.

Por otro lado, para CEA se dispuso de registros de 4 años comprendidos entre el 01 de enero del 2013 y 31 de diciembre del 2016, de los cuales, los primeros 3 años se usaron para entrenamiento y el último para prueba. Cabe destacar que, tanto para HUF como CEA, los horizontes de tiempo corresponden a MTLF (Gordillo-Orquera et al., 2018).

Para el presente proyecto, se utiliza el mismo rango de datos para el entrenamiento y prueba de los modelos. Los horizontes de predicción para el análisis de resultados serán STLF y MTLF. Además, para el diseño de algunos modelos, se consideró inicialmente solo los datos de HUF. Mientras que los de CEA, para aquellos casos, se utilizaron con el fin de comprobar los modelos.

Descripción y Análisis de Modelo: PCA con AR

El método PCA se utilizó para reducir la dimensionalidad de los datos, mientras que, el modelamiento *autorregresivo* (AR) permitió realizar la predicción de estos. De manera estratégica, se combinaron ambos métodos, de tal forma que al obtener las componentes principales estas puedan ser modeladas con AR (Gordillo-Orquera et al., 2018). Conforme a esto, los autores desarrollaron un algoritmo el cual, se resume en la Tabla 2.

Tabla 2

Algoritmo basado en PCA

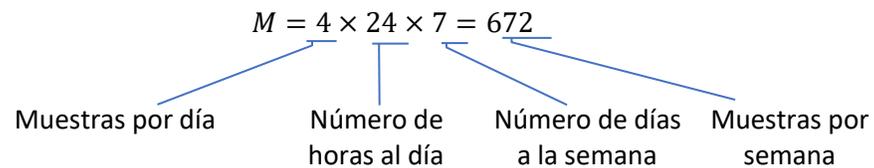
Algoritmo PCA
<p>Entrada: \mathbf{y}</p> <ol style="list-style-type: none"> 1. Redimensionamiento de \mathbf{y}, en matriz $\mathbf{X} \in \mathbb{R}^{M \times N}$. 2. Centrar en la media a \mathbf{X}, obteniendo $\tilde{\mathbf{X}}$ (Cumpliendo que $\mathbf{X} = \tilde{\mathbf{X}} + \bar{\mathbf{X}}$). 3. Implementar SVD a $\tilde{\mathbf{X}}$, de tal forma que $\tilde{\mathbf{X}} = \mathbf{U}\mathbf{S}\mathbf{V}^T$. 4. Determinar los P componentes más significativos. 5. Identificar la tendencia lineal y la media periódica. 6. Predecir los coeficientes $\widehat{\mathbf{V}}_r$. 7. Calcular la aproximación $\hat{\mathbf{X}} = \bar{\mathbf{X}} + \mathbf{U}_r \mathbf{S}_r \widehat{\mathbf{V}}_r^T$ 8. Redimensionar en forma de vector $\hat{\mathbf{y}}_{PCA} = \mathbf{vec}(\hat{\mathbf{X}})$ <p>Salida: $\hat{\mathbf{y}}_{PCA}$, \mathbf{U}</p>

Nota. Adaptado de Analyzing and Forecasting Electrical Load Consumption in Healthcare Buildings (p. 5), por Gordillo-Orquera et al., 2018, Energies, MDPI.

Al analizar el algoritmo, se observa que inicialmente, se redimensiona la variable de entrada y , la cual contiene los registros históricos de carga, esto con el fin de facilitar la aplicación de PCA. Específicamente, y se redimensiona en la matriz $X \in \mathbb{R}^{M \times N}$, donde $M = 4 \times 24 \times 7 = 672$, corresponde al número de muestras por semana y N al número de semana (Gordillo-Orquera et al., 2018). Con fines ilustrativos, en la Figura 13 se presenta la composición de M .

Figura 13

Formación del índice M (muestras por semana)

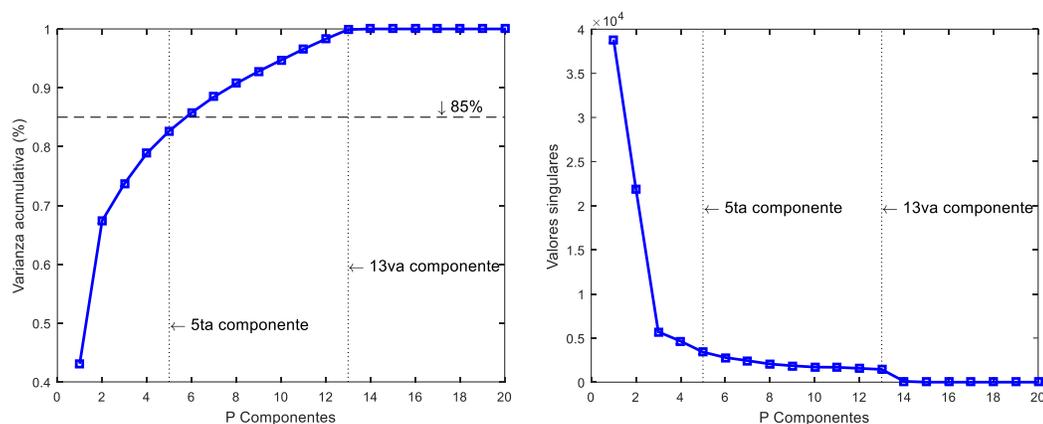


Seguidamente, se realiza un preprocesamiento de los datos, el cual, consiste en centrarlos en la media, de tal forma que $X = \tilde{X} + \bar{X}$. El valor que se utiliza a partir de este punto es \tilde{X} . Por consiguiente, se realiza la descomposición en valores singulares (SVD), obteniendo las matrices U , S y V , lo cual, equivale a la ecuación: $\tilde{X} = USV^T$.

Continuando, se determina los componentes más representativos del conjunto de datos. Es decir, se establece el orden del modelo más adecuado, buscando utilizar el mínimo número de componentes. Para ello, se analiza la varianza acumulativa en función de los componentes principales (Gordillo-Orquera, y otros, 2018).

Figura 14

Varianza Acumulativa (%) y Valores Singulares para HUF.



Nota. Adaptado de Analyzing and Forecasting Electrical Load Consumption in Healthcare Buildings (p. 9), por Gordillo-Orquera et al., 2018, Energies, MDPI.

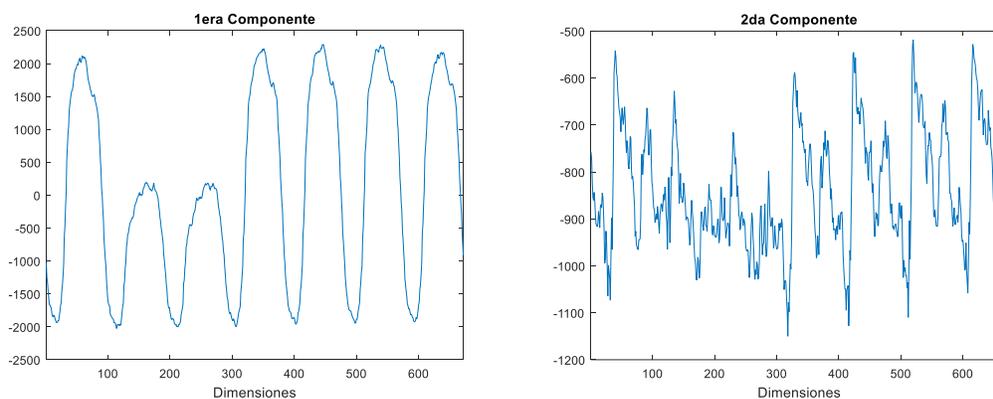
En la Figura 14, se observa el análisis de los componentes principales para HUF. Específicamente, se aplica una regla general (*rule of thumb*) utilizada comúnmente como criterio para la reducción de orden del modelo. La regla consiste en considerar las P componentes que permitan obtener una varianza acumulada mayor al 85% (Gordillo-Orquera, y otros, 2018; Zhang & Wang, 2009). Se determinó que en la 5ta componente principal ($P=5$) se cumple la regla. Por tal razón, se utilizaron las primeras 5 componentes principales para el modelo. Por otro lado, al observar los valores singulares se demuestra que al mayor número de componentes principales, estos se aproximan a cero. Además, es evidente que en la 13va componente ($P=13$), la varianza acumulada tiene un valor aproximado al 100% (Gordillo-Orquera, y otros, 2018).

Cabe mencionar que cada componente contribuye de una manera específica para formar nuevamente el comportamiento de consumo. Por ejemplo, en la Figura 15 se presentan los autovectores de la primera y segunda componente principal. Analizando los patrones, la 1era

aporta en el comportamiento semanal del consumo. Mientras que, la 2da corresponde a detalles más específicos de consumo. De la misma manera, las siguientes componentes (3ra, 4ta y 5ta) aportan características cada vez más específicas. Estas se pueden visualizar en el Anexo A.

Figura 15

Autovectores de 1era y 2da componente principal

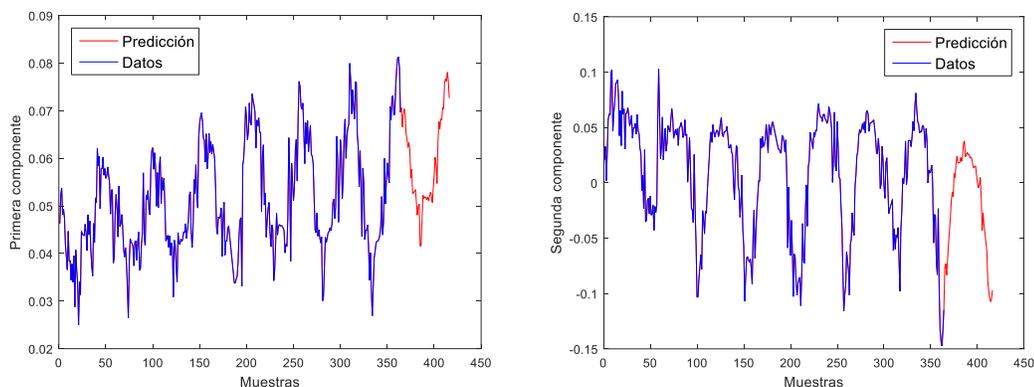


Nota. Adaptado de Analyzing and Forecasting Electrical Load Consumption in Healthcare Buildings (p. 13), por Gordillo-Orquera et al., 2018, Energies, MDPI.

Prosiguiendo el análisis, se asumió que los componentes principales tenían una periodicidad anual y tendencia lineal, dando lugar la posibilidad de aplicar modelamiento AR (Gordillo-Orquera et al., 2018). Mediante esta técnica, se procedió a predecir los valores de \hat{V}_r . La Figura 16 ejemplifica los resultados obtenidos en la predicción de la primera y segunda componente principal para HUF. Los datos de color azul corresponden a las componentes principales, mientras que los de color rojo, son la predicción obtenida con el algoritmo mencionado.

Figura 16

Ejemplo de predicción de 1era y 2da componente principal mediante AR para HUF



Nota. Adaptado de Analyzing and Forecasting Electrical Load Consumption in Healthcare Buildings (p. 13), por Gordillo-Orquera et al., 2018, Energies, MDPI.

Posteriormente se reconstruye los datos con las componentes principales predichas. Se suma la media quitada en el preprocesamiento, de tal forma que se obtiene un vector: $\hat{X} = \bar{X} + U_r S_r \hat{V}_r^T$. Finalmente, se redimensiona como vector a \hat{X} y se evalúan los parámetros de desempeño para el periodo prueba. Cabe mencionar que, el procedimiento fue similar para CEA. Los resultados obtenidos se resumen en la Tabla 3.

Tabla 3

Métricas de desempeño PCA y AR

Método	PCA	
	Datos	
	HUF	CEA
Métrica		
RMSE (kW)	28.91	13.33
MAPE (%)	5.06	16.21
PBIAS (%)	-2.87	7.58

Nota. Adaptado de Analyzing and Forecasting Electrical Load Consumption in Healthcare Buildings (p. 10), por Gordillo-Orquera et al., 2018, Energies, MDPI.

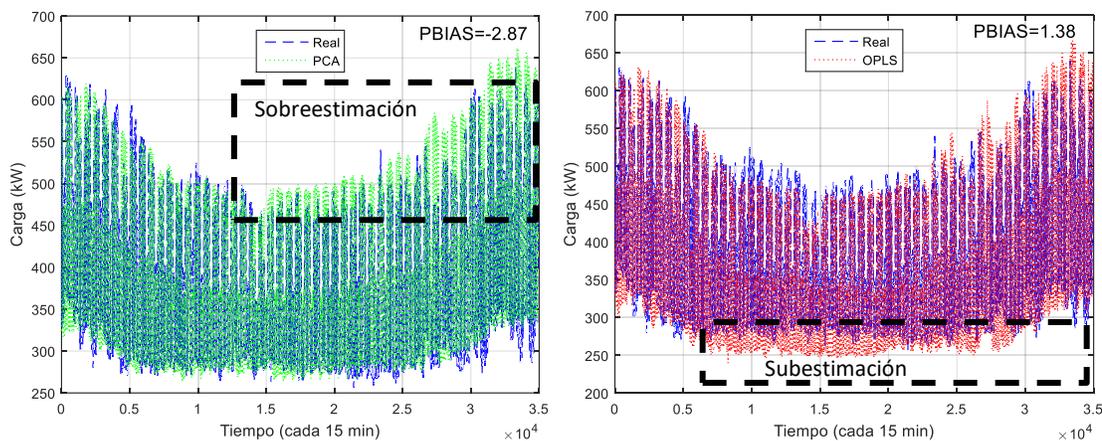
Los autores concluyeron que los resultados fueron satisfactorios para los métodos implementados. No obstante, identificaron la presencia de subestimación y sobreestimación de datos que se mencionará más detalladamente en la siguiente subsección.

Subestimación y Sobreestimación de los Modelos PCA y OPLS

Los modelos PCA con AR y OPLS exhibieron una característica de desviación evidenciada en los parámetros de desempeño. Se analizó en el periodo prueba de HUF y se observó que el modelo PCA presentó una sobreestimación de carga, mientras que, para OPLS ocurrió lo contrario, presentando una subestimación de carga, tal como se observa en la Figura 17.

Figura 17

Sobreestimación en modelo PCA y subestimación en modelo OPLS



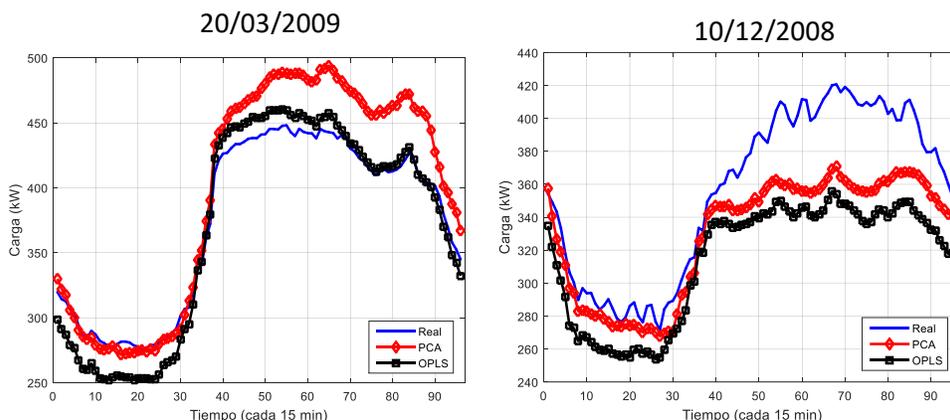
Nota. Adaptado de Analyzing and Forecasting Electrical Load Consumption in Healthcare Buildings (p. 13), por Gordillo-Orquera et al., 2018, Energies, MDPI.

Cabe mencionar que el PBIAS, indicado en la parte superior derecha de los gráficos, indica un valor negativo para PCA y uno positivo para OPLS. Esto confirma que, en efecto, para la mayoría de los datos, se tiene una sobreestimación y subestimación respectivamente. Por

otro lado, se comprueba nuevamente en periodos de tiempo más cortos. La Figura 18 ilustra el comportamiento de los predictores para 24 horas en 2 días diferentes.

Figura 18

Comparación de modelos PCA y OPLS, en los días 10/12/2008 y 20/03/2009



Nota. Adaptado de Analyzing and Forecasting Electrical Load Consumption in Healthcare Buildings, por Gordillo-Orquera et al., 2018, Energies, MDPI.

Al analizar la gráfica de la fecha 20/03/2009, se cumple que, para las primeras muestras del día, PCA realiza una mejor aproximación que OPLS, mientras que, para las siguientes, ocurre lo contrario. Además, el efecto de subestimación y sobreestimación también es evidente. Por otra parte, la gráfica de la fecha 10/12/2008, evidencia un caso particular donde se visualiza una subestimación de datos por parte de los dos modelos.

A partir de este ejemplo, se puede deducir que existen días en los que no se cumple estrictamente el rasgo de sobreestimación de PCA y subestimación de OPLS. A pesar de esto, en general, si ocurre este efecto, dado a que el parámetro PBIAS, calculado a lo largo del año, demuestra que en la mayoría de los días se cumple esta característica.

Análisis de Modelo ENS: Combinación Convexa entre PCA y OPLS

Anteriormente se mencionó que la combinación de modelos de predicción suele ser utilizada para mejorar el desempeño del pronóstico. Con este objetivo, Gordillo-Orquera (2018)

propuso una combinación convexa entre los algoritmos PCA y OPLS, para la cual, la formulación matemática se basó en la siguiente ecuación:

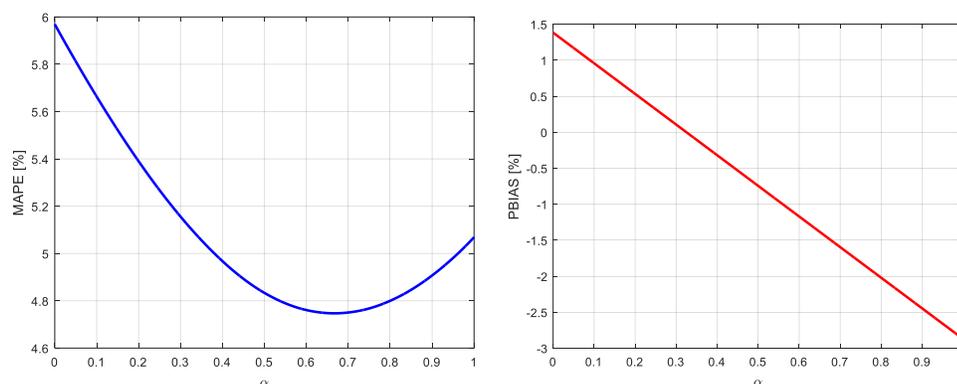
$$\hat{y}_{CVX} = \alpha \hat{y}_{PCA} + (1 - \alpha) \hat{y}_{OPLS} \quad (13)$$

Donde $\alpha \in [0,1]$ es un escalar que permite realizar combinación lineal entre la salida del predictor PCA (\hat{y}_{PCA}) y de OPLS (\hat{y}_{OPLS}). Dependiendo de los valores de α , la salida \hat{y}_{CVX} será la combinación entre ambos algoritmos. Si $\alpha = 0$, la salida corresponderá al predictor PCA, mientras que si $\alpha = 1$ será la salida del predictor OPLS. Para este caso, se asignó el valor de $\alpha = 0.5$. Este método se denominó ENS (por sus siglas en inglés, *ensembled*) y logró mejorar las prestaciones del pronóstico (Gordillo-Orquera et al., 2018).

A partir de esto, se precede a realizar un estudio del efecto del parámetro α en la predicción para el periodo de tiempo *Test*, para lo cual, se analiza la variación de los indicadores de desempeño MAPE [%] y PBIAS [%] en función de este parámetro, tal como se ilustra en la Figura 19.

Figura 19

MAPE [%] y PBIAS [%] en función de α



El rango de α únicamente puede tomar valores entre $[0,1]$, por lo cual se evaluó en ese intervalo incrementando los valores con un espaciamiento de 0.001. La información más relevante obtenida a partir de los gráficos se los resume en la Tabla 4, en la cual, se presentan

los valores de α , que permiten obtener los parámetros óptimos. La primera columna indica el MAPE óptimo y en la siguiente PBIAS más conveniente. Además, en la última columna se muestran los valores para $\alpha = 0.5$, que corresponde a ENS.

Tabla 4

Valores relevantes de α para indicadores MAPE [%] y PBIAS [%]

	$\alpha_{MAPE\acute{o}ptimo} = 0.67$	$\alpha_{PBIAS\acute{o}ptimo} = 0.32$	$\alpha = 0.50$
MAPE [%]	4.75	5.11	4.83
PBIAS [%]	-1.47	0.02	-0.74

Esto sugiere que, se podría mejorar el desempeño de la predicción escogiendo un α que permita reducir al máximo ambos indicadores. Sin embargo, esta solución no es adecuada dado que estos valores optimizan los indicadores solo para el periodo de prueba, pero no necesariamente para nuevos datos, es decir, se puede ocasionar un sobreajuste del modelo.

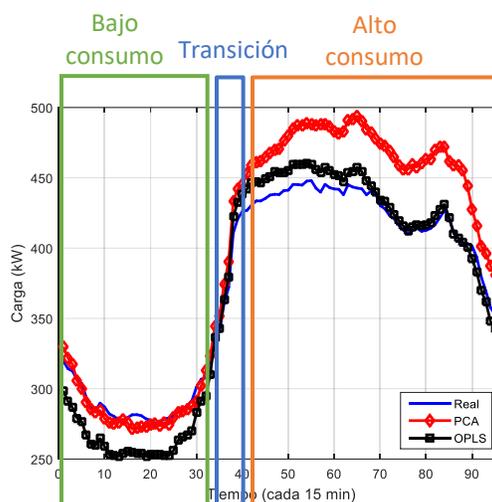
Diseño de Modelo NCVX: Combinación no Convexa entre PCA y OPLS

Se propone un nuevo modelo de combinación no convexa (NCVX) entre los algoritmos PCA y OPLS, proyectándose a disminuir el efecto de sobreestimación y subestimación. Este modelo entra en la categoría de modelos híbridos heterogéneos, por utilizar un mismo conjunto de datos, pero diferentes modelos base.

Como punto de partida, se observan los resultados obtenidos en las muestras de un día escogido aleatoriamente. Mediante observación, se determina que las primeras muestras corresponden a un bajo consumo y las que prosiguen a un alto consumo. En general, se pueden establecer tres rangos de tiempo, que representen este comportamiento: periodo de bajo consumo, transición y alto consumo. El día y los periodos mencionados se ilustran en la Figura 20.

Figura 20

Periodos de consumo en el día: 20/03/2009



Del análisis presentado en la sección Subestimación y Sobreestimación de los Modelos PCA y OPLS, se determinó que el modelo PCA realizaba una buena aproximación para las muestras del periodo de bajo consumo y a su vez una sobreestimación para las de alto consumo. Mientras que OPLS presentaba un mejor desempeño en las de alto consumo y una subestimación en las bajo consumo. Esto sugiere que, en las muestras de bajo consumo, es más conveniente utilizar los resultados de PCA y en las de alto OPLS.

Por lo tanto, se diseñó un sistema difuso que permita llevar a cabo esta idea. El objetivo es mantener el modelo PCA para las muestras de bajo consumo y el OPLS en las de alto consumo. Esto se detalla de manera más específica en la Tabla 5.

Tabla 5

Intervalos de tiempo en un día con modelos de mejor desempeño

Periodo de consumo	Rango de horas	Intervalo de muestras	Modelo de mejor aproximación
Bajo consumo	00:15 a 9:00	[1,36]	PCA
Transición	09:15 a 9:45	[37,39]	PCA o OPLS
Alto consumo	10:00 a 00:00	[40,96]	OPLS

Cabe destacar que, para el periodo de transición se estima que cualquiera de los dos métodos funcionará correctamente, por ello, no se especifica un modelo en particular para este rango. Además, se observa que los datos tienen correspondencia con la realidad, dado que en las horas de la madrugada se tiene un bajo consumo, mientras que en las del día y noche, el consumo aumenta.

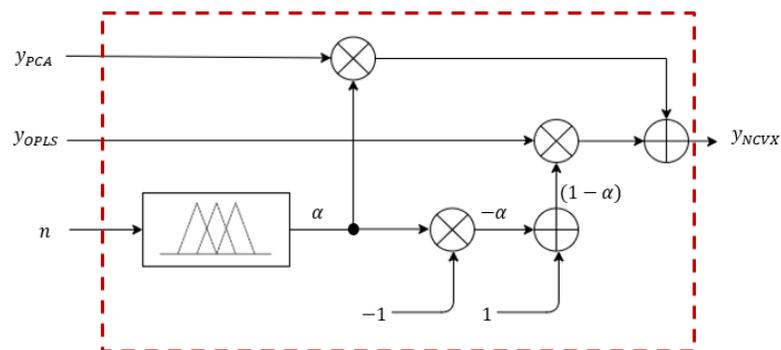
En comparación con el método ENS, la principal diferencia es que ahora α no será una constante, sino un valor cuyo comportamiento dependerá del periodo de consumo y, por ende, del número de muestra diaria n . La ecuación (14), representa la salida del modelo NCVX para el periodo de un día.

$$y_{NCVX}[n] = \alpha[n]y_{PCA}[n] + (1 - \alpha[n])y_{OPLS}[n] \quad (14)$$

Esta ecuación dicta que la predicción y_{NCVX} dependerá de α , y_{PCA} y y_{OPLS} , los cuales, a su vez, dependen de la muestra de tiempo n . Otra forma de describir el algoritmo es mediante un diagrama esquemático que se presenta en la Figura 21.

Figura 21

Diagrama esquemático de modelo NCVX



Como se observa, el parámetro α proviene de un sistema difuso, cuyo diseño se explicará después de la presente sección. Además, cabe mencionar, que el presente algoritmo

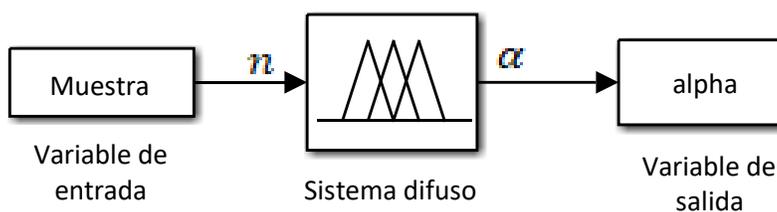
puede ser adaptado a periodos de tiempo más largo de un día, repitiendo los valores de α para el número de días requeridos y después aplicando la ecuación (14).

Especificación de Entradas, Salidas y MI del Sistema Difuso

Se procedió diseñar un sistema difuso de una entrada y una salida. La variable que ingresa es la muestra de tiempo del día n , mientras que, la salida corresponderá al parámetro α , tal como se ve en la Figura 22.

Figura 22

Sistema difuso para NCVX



La variable n , que corresponde a un día entero, está definida en el rango $[1,96]$. Por otro lado, el parámetro $\alpha \in [0,1]$ y funciona de manera similar al combinador convexo de ENS, en el cual, para $\alpha = 0$ la salida representa al predictor PCA, mientras que, para $\alpha = 1$ representa a OPLS.

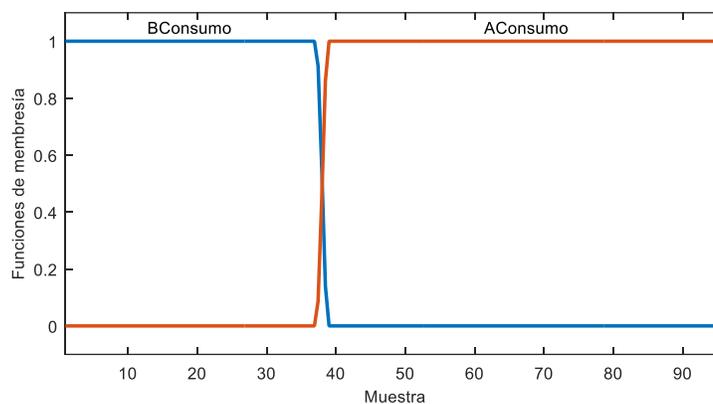
Con fines comparativos, se optó por probar el sistema difuso con los MI *Mamdani* y *Takagi-Tsugeno*. El modelo definitivo, se seleccionará basándose en los resultados obtenidos en la predicción.

Especificación de Funciones de Membresía para la Entrada

Se procede a desarrollar la etapa de *fuzzificación* para la variable de entrada. Teniendo en consideración los intervalos de muestras de la Tabla 5, se definieron dos funciones de pertenencia, las cuales se ilustran en la Figura 23.

Figura 23

Funciones de membresía para la entrada Muestra (n)



BConsumo representa las muestras de bajo consumo y consiste en una función tipo Z, la cual, permite una transición suave entre 1 y 0. Por otro lado, *AConsumo* representa las de alto consumo, con una función tipo S que de manera análoga permite la transición suave de 0 a 1.

Conjunto de Reglas de Inferencia

El conjunto de reglas de inferencia se desarrolló con base en las funciones de membresía de entrada y los valores de α que permitan alternar entre el modelo PCA y OPLS. Las reglas definidas se pueden visualizar en la Tabla 6.

Tabla 6

Conjuntos de reglas para sistema difuso de NCVX

	Salida	α
Entrada		
BConsumo		PCA ($\alpha \approx 1$)
Transición		PCA o OPLS ($\alpha \in [0,1]$)
AConsumo		OPLS ($\alpha \approx 0$)

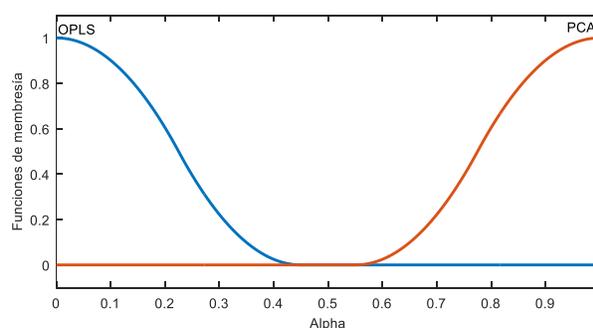
Se recalca que, para el presente modelo en el intervalo de tiempo *Transición*, el valor de α no será de mayor importancia, dado a que cualquiera de los dos métodos presenta buenos resultados en ese periodo.

Especificación de Funciones de Membresía para la Salida

Con base en el conjunto de reglas, se procedió a diseñar una etapa de *defuzzificación*, la cual, en específico permitirá implementar el modelo *Mamdani*. Se definieron dos funciones de pertenencia para la salida, permitiendo así definir los posibles valores del parámetro α . La Figura 24 ilustra las funciones mencionadas.

Figura 24

Funciones de membresía de la salida: Alpha



Se utilizó una función tipo Z para los valores que representan OPLS y una del tipo S para los de PCA. Al visualizar, se percibe que la transición entre las funciones es muy amplia, esto es para evitar que coincidan y así diferenciar claramente que los valores bajos corresponden a OPLS y los altos a PCA.

Cabe mencionar que, para *Takagi-Tsugeno* esta etapa es diferente y se deben definir funciones constantes o lineales para la salida. En el presente caso, las más adecuadas serían del tipo constantes que contengan los valores de α , es decir: $f_1 = 1$ y $f_2 = 0$.

Diseño de Modelos: Combinación de PCA y Algoritmos Basados en IC

En esta sección se presentan dos modelos denominados PCANN y PCAFUZZY. El primero incluye NN, mientras que, el segundo incluye las mismas y un sistema difuso. Se explican los diseños para cada caso.

Diseño de Modelo PCANN

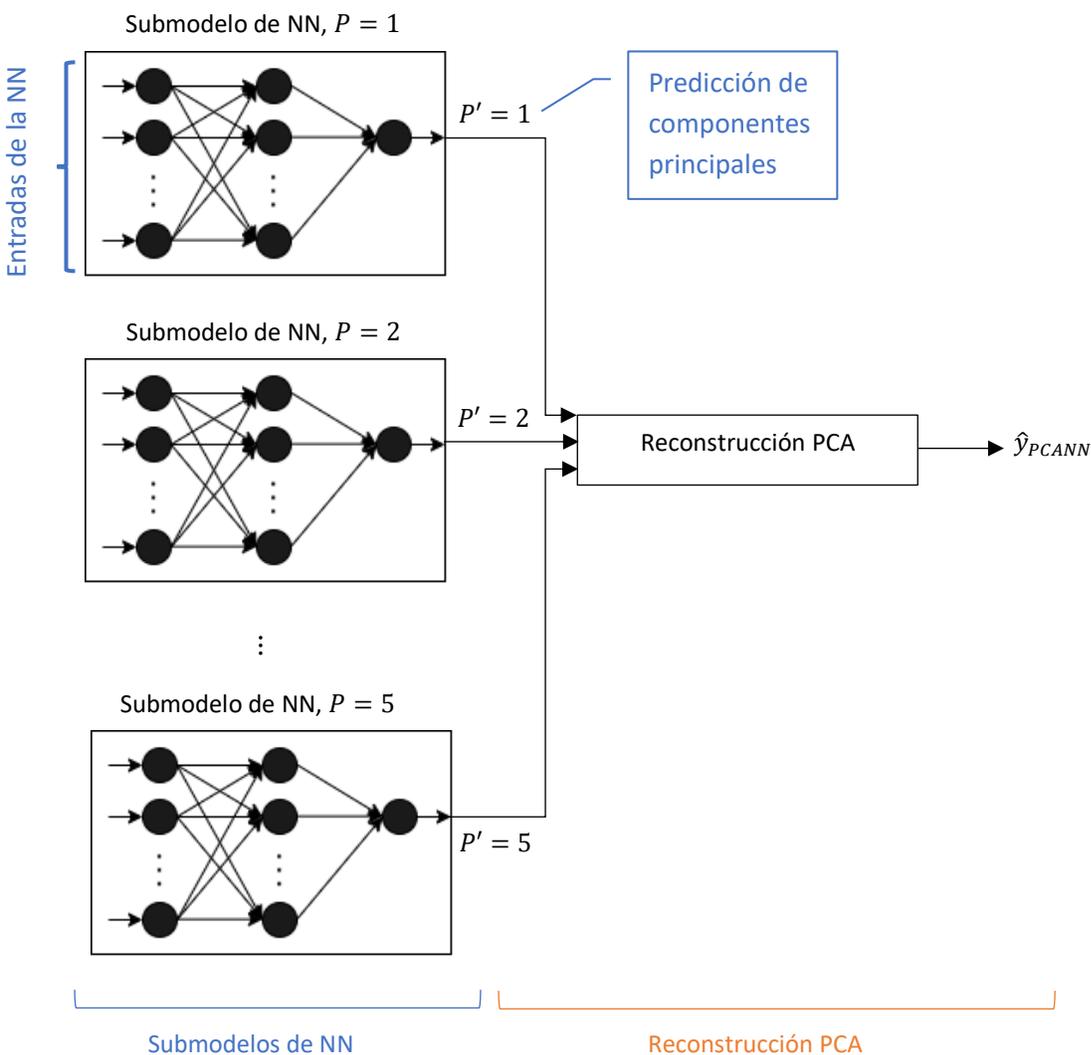
Se propone un modelo que combina el método PCA y el modelamiento por NN (PCANN). Este consiste en una metodología alternativa basada en el modelo PCA y AR (explicado anteriormente), el cual, reemplaza el modelamiento AR por uno basado en NN, manteniendo el PCA. Específicamente, en el algoritmo mostrado en la Tabla 2, se modifica el sexto paso (predicción de coeficientes \hat{V}_r). Por consiguiente, con estas consideraciones se procede a establecer una estructura del modelo.

Estructura del Modelo Propuesto

Se define una estructura para realizar el modelamiento predictivo de cada componente principal. Los datos de entrenamiento de los modelos serán los valores de las componentes principales de PCA ($P = 1, \dots, 5$). Por cada componente principal se crea un submodelo de NN, dando como resultado 5 NN, esto ocurre tanto para HUF como para CEA. Así, con los submodelos desarrollados, se predicen los nuevos valores de las componentes principales para el periodo de prueba, con los cuales, se puede reconstruir los datos originales de carga para este mismo periodo. El esquema de la estructura del modelo se ilustra en la Figura 25.

Figura 25

Ilustración esquemática de la estructura del modelo PCANN



Considerando el esquema mencionado, se recalca que el bloque *Reconstrucción PCA*, corresponde a los dos últimos pasos del algoritmo en la Tabla 2, es decir, el proceso de obtener la reconstrucción SVD, sumarle la media y vectorizar la matriz obtenida para así conseguir la salida, que en este caso será \hat{y}_{PCANN} . En definitiva, el problema de este modelo radica en el

diseño de los submodelos de NN. Además, cabe destacar que PCANN tiende a ser un modelo híbrido homogéneo, por utilizar diferentes subconjuntos de datos, es decir, las componentes principales, para entrenar submodelos del mismo tipo.

Diseño de NN

La presente sección explica el diseño de los submodelos de NN mencionados anteriormente. El procedimiento consistió en los siguientes pasos: definición de la distribución de datos, selección del modelo, arquitectura, algoritmo de entrenamiento y el criterio de parada, así como los de indicadores de desempeño y variables del modelo.

Distribución de Datos

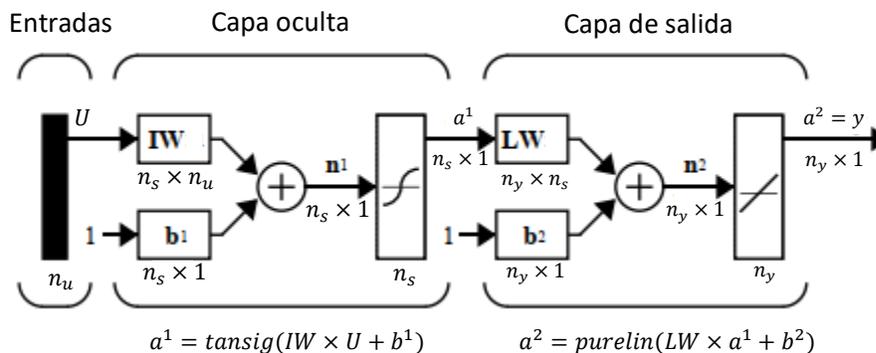
Normalmente, el *toolbox* de *Matlab* recomienda la partición de los datos en los siguientes subconjuntos: *training set*, *validation set* y *test set*, con los porcentajes de 70%, 15% y 15% respectivamente. De acuerdo con Martí (2009), el porcentaje de cada subconjunto depende de la cantidad de datos disponibles. En el presente diseño, se mantienen los porcentajes recomendados por el *Software*, dado que los datos son numerosos tanto para HUF como para CEA. Además, los datos para cada subconjunto serán seleccionados aleatoriamente.

Arquitectura de NN

Anteriormente, se mencionó que las NN funcionan como aproximadores generales de funciones. Por esta razón, una de sus aplicaciones más comunes es el modelamiento. Demuth & Beale (2004) mencionan que una de las arquitecturas usualmente usadas para este propósito, es la que se ilustra en la Figura 26, la cual será utilizada para el presente diseño.

Figura 26

Arquitectura de NN como aproximador general de funciones



Nota. Adaptado de Aproximador General de Funciones (p.5-7), por Demuth & Beale, 2004, The MathWorks, Inc.

Donde n_u es el número de entradas, n_s el número de neuronas en la capa oculta, n_y el número de salidas, U las entradas de la red, y la salida, IW la matriz de pesos en la capa oculta, LW la matriz de pesos en la capa de salida, $b^{1,2}$ los *bias* para cada capa, $a^{1,2}$ las salidas de la capa oculta y salida, respectivamente.

Específicamente, la red cuenta con una arquitectura multicapa del tipo FFN, la cual consta de una capa oculta y una de salida. Cabe destacar que, se utilizan funciones de transferencia sigmoideas del tipo *tansig* para la capa oculta y una lineal *purelin* para la capa de salida. La razón de esto se justifica en el hecho de que las funciones sigmoideas capturan la relación lineal y no lineal entre las entradas y salidas. Por otro lado, en la capa salida se pueden llegar a requerir valores fuera de un rango determinado, por ello, esta debe ser lineal (Demuth & Beale, 2004).

Otro parámetro para mencionar es el número de neuronas en la capa oculta, lo cual se analiza de manera empírica al momento de implementar las NN. Cabe mencionar, que la habilidad de una NN para comprender el comportamiento no lineal depende del número de

neuronas en la capa oculta (López, Sans, Valero, & Senabre, 2018). Sin embargo, también se afirma que el aumento de estas incrementará el costo computacional.

Algoritmo de Entrenamiento y Criterio de Parada

Para el presente caso, se escoge el algoritmo *Levenberg-Marquardt*. La razón de esta selección es por la alta rapidez de entrenamiento que este presenta. Además, al ser NN de tamaño relativamente pequeño, el uso de este algoritmo es adecuado. Con respecto a los parámetros requeridos por este método, se utilizan los que vienen por defecto en *Matlab*.

Como criterio de parada, se utiliza ES, con el fin de contrarrestar el sobreajuste de datos y lograr generalizar los modelos. El indicador seleccionado para evaluar este criterio es el MSE.

Como se mencionó anteriormente, el subconjunto de validación se utiliza para detener el entrenamiento cuando el MSE empieza a ascender con cada iteración. *Matlab* recomienda 6 iteraciones como número máximo permitido para detener el entrenamiento. Conviene poner en énfasis que los pesos sinápticos y *bias* escogidos serán los que corresponden a la iteración donde se presente el menor MSE para validación.

Características de NN para Entrenamiento

En resumen, en Tabla 7 se presentan las características finales de las NN a implementar. Los 5 submodelos mencionados mantendrán estas características. Cabe mencionar que, se probó diferentes números de neuronas en la capa oculta, en específico 7-15-20.

Tabla 7*Características de las NN para entrenamiento*

Características
Distribución de datos: 70% (entrenamiento), 15% (validación), 15% (prueba)
Número de capas: 1 capa oculta (mediante <i>tansig</i>) y 1 capa de salida (mediante <i>purelin</i>)
Número de neuronas en la capa oculta: 7-15-25
Algoritmo de entrenamiento: <i>Levenberg-Marquardt</i> (parámetros por defecto de <i>Matlab</i>)
Criterio de parada: <i>Early Stopping</i> (6 iteraciones)
Indicadores de error: MSE

Nota. El número de neuronas en la capa oculta seleccionado para cada submodelo se lo determina en la sección Experimentación de Número de Neuronas.

Variables del Modelo de NN

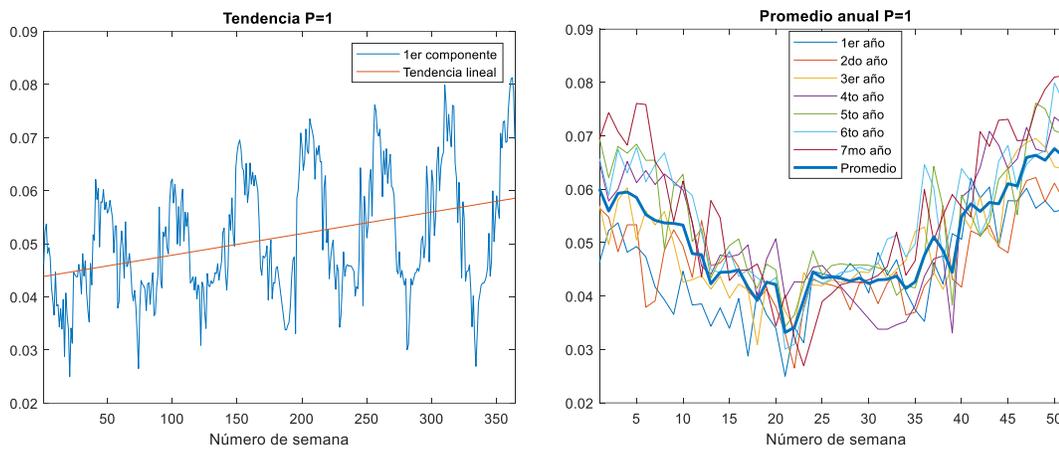
Las entradas de cada NN fueron escogidas de acuerdo con la aproximación general de una serie de tiempo, la cual, consiste en tres componentes: tendencia (*trend*), estacionalidad (*seasonality*) y estacionariedad (*stationary*) (Brockwell & Davis, 2016).

Con base en lo mencionado, se estudiaron variables que capturen los componentes de tendencia y estacionalidad, mientras que la estacionariedad se pretende capturar con las componentes principales menos significativas obtenidos por PCA. Por consiguiente, se seleccionaron las siguientes entradas:

- La *tendencia*, que permite ver la evolución lineal de la serie temporal. Esto se evidencia claramente con la primera componente principal (P=1), ver gráfico izquierdo de la Figura 27, donde se observa un incremento de los valores al pasar los años.

Figura 27

Primera componente principal con su tendencia lineal y promedio para HUF.



Nota. Adaptado de “Analyzing and Forecasting Electrical Load Consumption in Healthcare Buildings” (p. 13), por R. Gordillo-Orquera et al., 2018, *Energies*, 11 (493).

- Promedio, que corresponde al promedio anual de los datos de las componentes principales y permite determinar la estacionalidad de la serie temporal. Es decir, captura la forma de la señal anualmente, tal como se observa en la parte derecha de la Figura 27. En el Anexo B, se presenta la tendencia y promedio para todas las componentes principales.
- Temperatura promedio, la cual es una variable exógena e indica la temperatura media mensual/anual registrada en el periodo comprendido entre 1981-2010. Los datos climatológicos se obtuvieron de la Agencia Estatal Meteorológica de España (AEMet, s.f.).
- Número de semana en la muestra, la cual consiste en una variable que indica la semana en el registro de datos. Por ejemplo, al tener 8 años de registros ($8 \text{ años} = 52 \times 8 \text{ semanas} = 416 \text{ semanas}$), esta variable sera un vector comprendido por los valores de 1, ..., 416.
- Número de semana al año, que corresponde a una variable categórica y permite enumerar la semanas anualmente (52 al año). De esta manera se categoriza y diferencia el consumo en las distintas semanas, por ejemplo, en las de verano donde existe mayor consumo.

Con el fin de generalizar los modelos, las cinco variables mencionadas, serán las entradas a cada NN, mientras que las salidas a cada una, serán las componentes principales. Algunas de las variables, requieren de un preprocesamiento antes de ingresar a la red. Por ello, se han normalizado los datos que así lo ameritan, es decir, T° promedio, N. de semana en la muestra y N. semana al año. De manera más detallada, en la Tabla 8, se puede visualizar las entradas con sus detalles, escogidas para el entrenamiento de las NN.

Tabla 8

Entradas de las NN

Entrada	Detalle	Valor
u_1	Número de semana en la muestra	$\frac{Ns_1}{n_{sm}}, Ns_1 = 1, \dots, n_{sm}$, donde $n_{sm} = 416$ (para HUF) y $n_{sm} = 208$ (para CEA)
u_2	Número de semana al año	$\frac{Ns_2}{52}, Ns_2 = 1, \dots, 52$
u_3	Tendencia	Ecuación de la recta con formato: $y = mx + b$
u_4	Temperatura promedio	Registros de T° normalizados entre $[0, 1]$
u_5	Promedio	Datos promediados anualmente

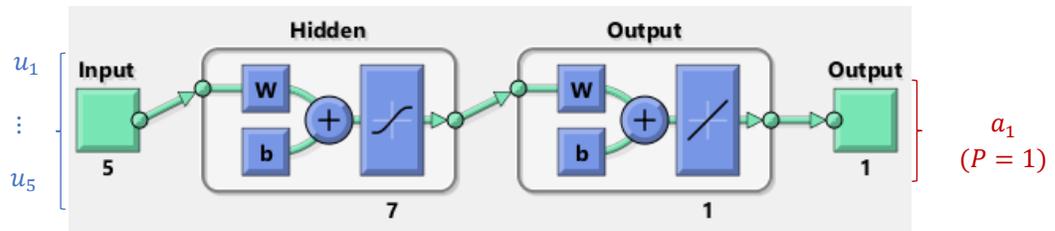
En resumen, se implementan 5 NN, con las entradas ilustradas en la Tabla 8, es decir u_1, \dots, u_5 , por otro lado, las salidas serán $a_1 (P = 1), \dots, a_5 (P = 5)$, para cada NN respectivamente.

Implementación de Modelos de NN

La implementación del diseño se resume en el ejemplo presentado en la Figura 28, específicamente, del primer submodelo de NN, es decir, el correspondiente a la primera componente principal. Como se observa, se obtuvieron en cuenta las 5 entradas y la salida (P=1), así como 7 neuronas en la capa de oculta, para este caso.

Figura 28

NN para la primera componente principal tal como se visualiza en Matlab Mathworks.



Las otras NN, tienen una estructura muy similar al presente ejemplo, el único cambio es en el número de neuronas de la capa oculta, que para cada caso es diferente, dependiendo de la complejidad de la serie de datos.

Diseño de Modelo PCAFUZZY

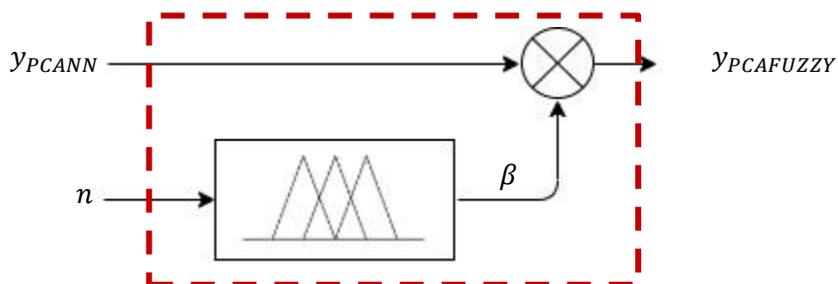
Se propone un modelo, denominado PCAFUZZY que surge a partir de los resultados de PCANN. Con el propósito de mejorar los parámetros de desempeño, se propone una modificación mediante la añadidura de un factor multiplicativo, obtenido a partir de un sistema difuso, dando así lugar al modelo PCAFUZZY. Esta estrategia se basa en la metodología propuesta por Ganguly, Kalam, & Zayegh (2017), en la cual, se realizó un procedimiento similar para incorporar variables de predicción, mediante una previa observación de los resultados en el periodo de prueba. En síntesis, la salida del presente modelo se puede representar mediante la ecuación (15).

$$y_{PCAFUZZY}[n] = \beta[n] \cdot y_{PCANN}[n] \quad (15)$$

Las variables $y_{PCAFUZZY}$, y_{PCANN} , β , n , corresponden a la salida del presente modelo PCAFUZZY, la salida de PCANN, el factor de corrección y el número de muestra al año, respectivamente. De manera gráfica, el esquemático de PCAFUZZY se ilustra en la Figura 29.

Figura 29

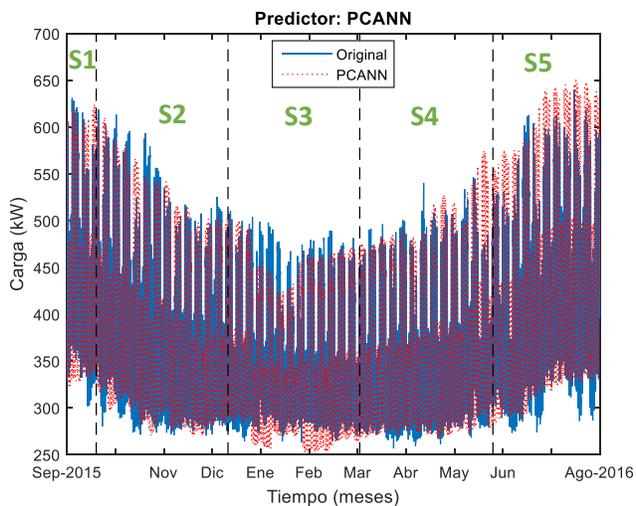
Diagrama esquemático de PCAFUZZY



El diseño del sistema difuso se lo realizó con base en las estaciones del año. Por lo tanto, se incorporará esta variable para mejorar la predicción, considerando que esta permitirá categorizar los datos en función de las estaciones. Inicialmente, se observaron los resultados del previo modelo, con el fin de determinar los valores requeridos para el factor de corrección. Para mayor facilidad de observación, se establecieron los rangos de datos correspondientes a las estaciones del año, tal como se ilustra en la Figura 30, para el caso de HUF.

Figura 30

Secciones representativas de estaciones del año para HUF.



En este gráfico se pueden observar 5 secciones (S1, S2, S3, S4 y S5) donde cada una está relacionada a una estación del año. Las secciones S1 y S5, corresponden al verano, mientras que S2, S3 y S4 conciernen a otoño, invierno y primavera respectivamente.

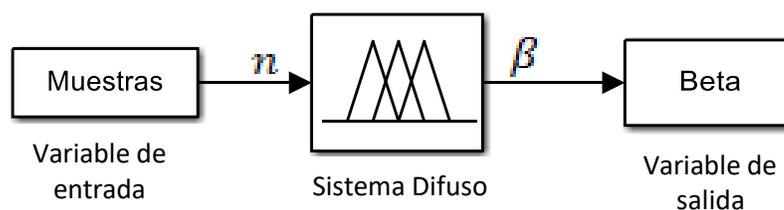
Al visualizar detalladamente, en la sección S5 (correspondiente al verano) se presenta una sobreestimación por parte del predictor PCANN. En definitiva, el factor de corrección propuesto se diseñará con el fin de disminuir este efecto. Con esto se pretenderá mejorar las prestaciones de desempeño. Por lo tanto, se presentará el diseño de un sistema difuso, el cual, se explica en la siguiente subsección. Cabe mencionar, que el diseño se realizó observando únicamente los resultados en HUF. Sin embargo, se probó también para CEA con el fin de evaluar el modelo.

Especificación de Entradas, Salidas y MI del Sistema Difuso

Se desarrolló un sistema basado en lógica difusa, cuya entrada corresponde a un vector que representa las muestras para el periodo de un año y como salida el factor de corrección β , tal como se ve en la Figura 31.

Figura 31

Estructura de sistema difuso para factor β



La variable de entrada n será definida por las muestras de tiempo a lo largo de un año, por lo cual, $n \in [1,34944]$. Mientras que, el parámetro β tendrá como objetivo mantener la señal original cuando $\beta = 1$ y reducir la sobreestimación cuando $\beta = 0.9$.

Con respecto al MI del sistema, se seleccionó *Takagi-Sugeno*, dado a su alta precisión. Para este caso, se utiliza un conjunto de funciones de salida de tipo constantes, debido a su facilidad de definición.

Las estaciones del año tendrán una importante repercusión en los datos, dado que dependen de la temperatura. En la ciudad de Madrid se presentan las cuatro estaciones del año: invierno, primavera, verano y otoño. En la Tabla 9, se detalla más información de lo mencionado.

Tabla 9

Características de la variable estación.

Estación	Fechas aproximadas	Características
Invierno	21 diciembre- 20 marzo	Temperatura promedio mínima de 1.2°C y una temperatura promedio máxima de 16.8°C.
Primavera	21 de marzo- 20 de junio	Temperatura promedio mínima de 6.9°C y máxima de 29.3°C.
Verano	21 de junio- 20 de septiembre	Temperatura promedio mínima de 14.6°C y máxima de 33.2°C.
Otoño	21 de septiembre- 20 de diciembre	Temperatura promedio mínima de 2.4°C y máxima de 20.6°C.

Nota. Adaptado de Valores climatológicos normales Getafe [Tabla], por AEMet, s.f., Aemet (www.aemet.es).

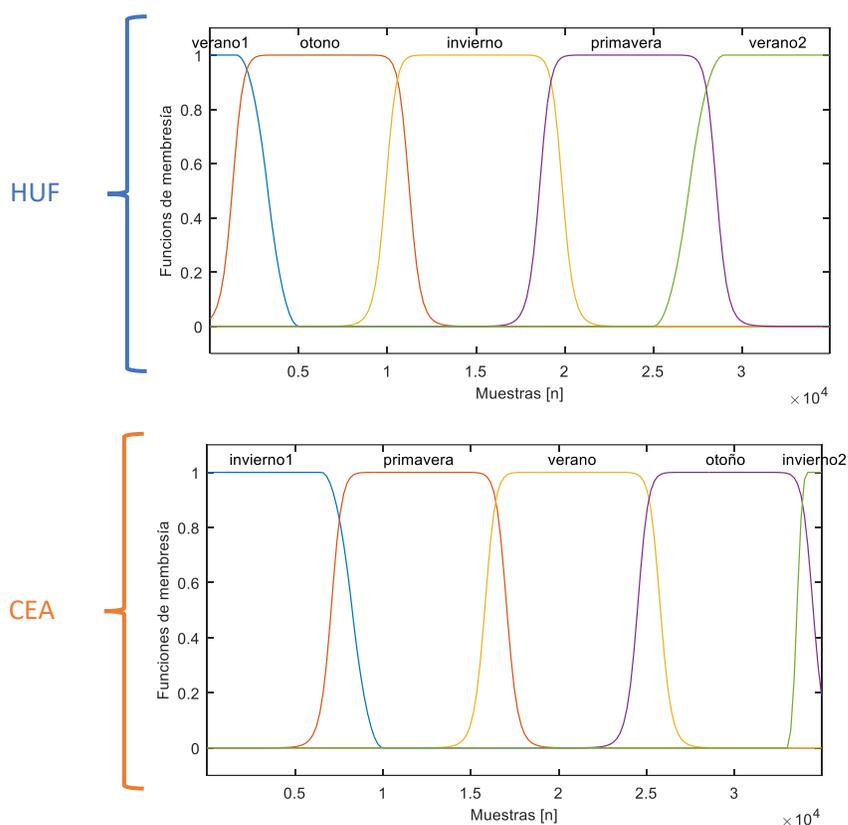
Mediante observación, se determina que existe una correspondencia entre las estaciones y los datos de consumo presentados en la Figura 30. En el periodo de verano, al aumentar la temperatura, el modelo prevé un aumento en el consumo. No obstante, para las otras estaciones no se observa mayores subestimaciones o sobreestimaciones que deban ser considerados.

Especificación de Funciones de Membresía para la Entrada

Se llevo a cabo la etapa de *fuzzificación* para la entrada de las muestras de tiempo de acuerdo con las estaciones del año. Se establecieron las funciones de membresía considerando las fechas ilustradas en la Tabla 9, dando como resultado las funciones ilustradas de la Figura 32.

Figura 32

Funciones de membresía para la entrada Muestras [n], para HUF y CEA



Para todas las funciones de membresía, se utilizó el tipo de función Gaussiana, la cual permite transiciones suaves entre 0 y 1, dependiendo de la estación. Con base en lo mencionado se procedió a establecer las reglas de inferencia.

Conjunto de Reglas de Inferencia

Inicialmente, se especifican los valores ideales del factor β para cada estación del año. Esto se determina mediante la observación de la Figura 30. Como resultado se obtienen las especificaciones ilustradas en la Tabla 10.

Tabla 10

Especificaciones para el diseño del sistema difuso en HUF

Tiempo	Valores de β	Predicción más aproximada
Verano1 [1,1920]	$(\beta \approx 1)$	\hat{Y}_{PCANN} [1,1920]
Otoño [1921,10560]	$(\beta \approx 1)$	\hat{Y}_{PCANN} [1921,10560]
Invierno [10561,19200]	$(\beta \approx 1)$	\hat{Y}_{PCANN} [10561,19200]
Primavera [19201,27936]	$(\beta \approx 1)$	\hat{Y}_{PCANN} [19201,27936]
Verano2 [27937,34944]	$(\beta \approx 0.9)$	$0.9 \times \hat{Y}_{PCANN}$ [27937,34944]

Cabe recalcar, que en el intervalo de verano1 se mantiene un valor de $\beta = 1$, dado a que en este periodo de tiempo no se obtuvo una sobreestimación considerable. Para CEA, las especificaciones son muy similares, donde el periodo de *verano* tendrá el parámetro $\beta = 0.9$. Con base en esto se procede a establecer las salidas del sistema difuso, para terminar de definir las reglas.

Especificación de Funciones de Salida

Dado a la baja complejidad del problema, se establecieron directamente las funciones constantes tomando los valores ideales de β . Como resultado, las reglas de inferencia quedaron definidas de acuerdo con la Figura 33.

Figura 33

Conjunto de reglas para HUF, tal como se muestran en Matlab Mathworks

1. If (timel is verano1) then (beta is f1) (1)
2. If (timel is otono) then (beta is f2) (1)
3. If (timel is invierno) then (beta is f3) (1)
4. If (timel is primavera) then (beta is f4) (1)
5. If (timel is verano2) then (beta is f5) (1)

Donde $f_1 = f_2 = f_3 = f_4 = 1$, mientras que $f_5 = 0.9$. Estas funciones representan al factor β . Con estos valores se pretende generalizar el modelo. La estrategia consiste en evitar especificar factores de corrección que corresponden a cada estación, logrando así impedir un sobreajuste. Cabe mencionar que, para CEA las reglas fueron muy similares, solo cambia el orden de las funciones y se especifica que la correspondiente a verano contendrá el valor de 0.9.

Capítulo IV

Experimentos y Análisis de Resultados

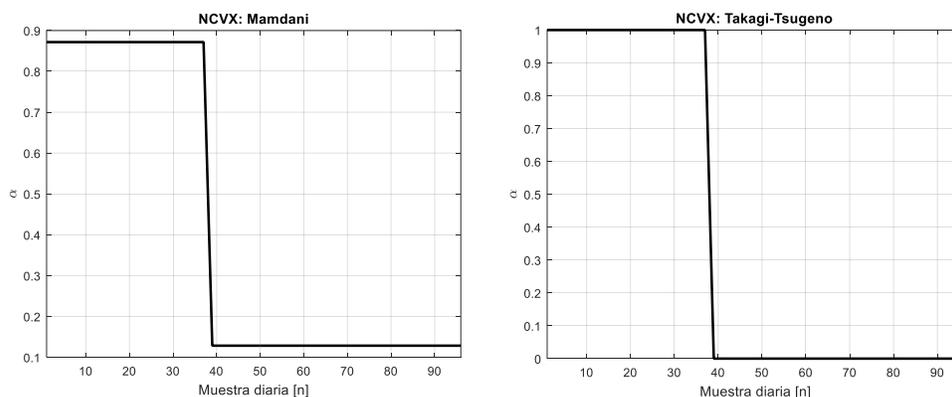
Este capítulo presenta los resultados de los tres modelos implementados: NCVX, PCANN y PCAFUZZY. Los datos obtenidos son analizados en los horizontes de tiempo: STLF y MTLF, utilizando los registros históricos disponibles de HUF y CEA. Inicialmente, se exponen los resultados del diseño de los modelos. Seguidamente, se analizan los parámetros de desempeño alcanzados y se los compara con los presentados por Gordillo-Orquera et al. (2018). Para cada caso, se exhibe un análisis e interpretación correspondiente.

Resultados de diseño NCVX

Se analizan los resultados del sistema difuso implementado para NCVX. Retomando, lo mencionado en la sección Especificación de Entradas, Salidas y MI del Sistema Difuso, se experimenta con los MI de *Mamdani* y *Takagi-Tsugeno*. Para ello, se obtienen las curvas de respuesta (entrada vs salida), presentadas en la Figura 34.

Figura 34

Curvas de respuesta α vs muestra, para los MI Mamdani y Takagi-Tsugeno en NCVX

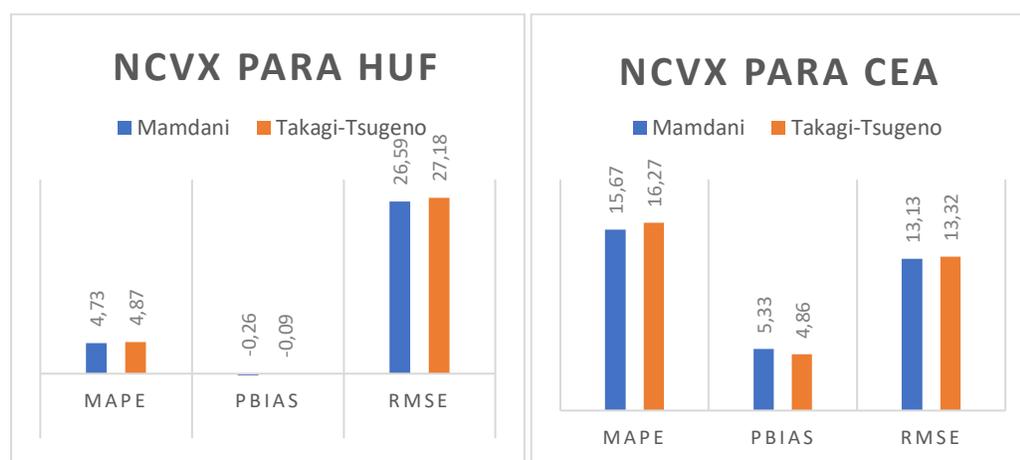


Se observa que para ambos casos el comportamiento es adecuado, dado que, en las primeras muestras [1, ...,39], el parámetro toma valores altos, representando a PCA, y posteriormente toma valores bajos, representando a OPLS.

En comparativa, el MI *Takagi-Tsugeno* demuestra ser más preciso que *Mamdani*, específicamente en los valores superiores e inferiores. Sin embargo, para completar el análisis y seleccionar el MI más adecuado, se procede a comparar brevemente los resultados previos de predicción en MTLF para cada caso, lo cual, se muestra en la Figura 35.

Figura 35

Comparación parámetros de desempeño NCVX (Mamdani y Takagi-Tsugeno) para HUF y CEA



En ambos casos, se visualiza que se obtuvo un PBIAS aceptable, lo cual, indica que se redujo el efecto de sobreestimación y subestimación. De manera específica, se observa que el MAPE y RMSE de *Mamdani* son inferiores a los de *Takagi-Tsugeno*. Sin embargo, para el PBIAS ocurre lo contrario. No obstante, este parámetro es comparable para ambos casos. Dado a que *Mamdani*, presenta en general resultados más ventajosos que *Takagi-Tsugeno*, se decidió optar por utilizar este modelo para el presente proyecto. Se recalca que, a partir de este punto, todo el análisis del modelo NCVX corresponde al sistema difuso diseñado con *Mamdani*.

Resultados de diseño PCANN

En esta sección, se presenta la experimentación para definir el número de neuronas en la capa oculta de cada NN. Además, se estudia la importancia de cada una de las entradas

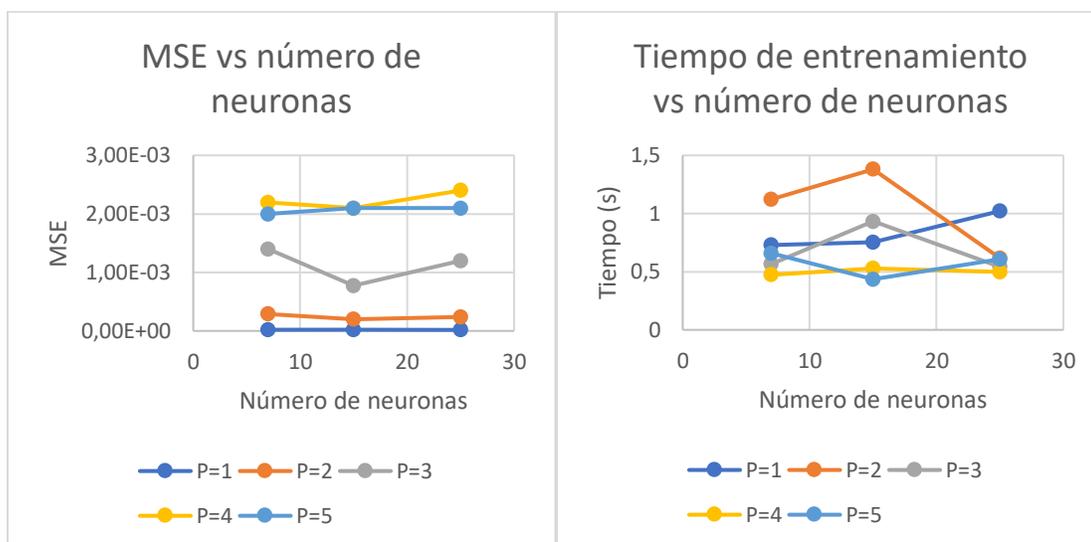
mediante una metodología LOOCV. Concluyendo con los resultados de implementación de cada modelo de NN y sus pruebas en el periodo de prueba.

Experimentación de Número de Neuronas

Se pretende estudiar la relación que existe entre el número neuronas (en la capa oculta) con el indicador de error MSE, así como, con el tiempo de entrenamiento. El MSE permitirá evaluar la precisión del modelo, mientras que el tiempo de entrenamiento representará el costo computacional. Para ello, se entrenan las NN con 7, 15 y 20 con los datos de HUF. Bajo los principios estudiados, se prevé que, al aumentar el número de neuronas, el MSE disminuirá, pero el tiempo de entrenamiento aumentará. En la Figura 36 se presenta el estudio mencionado.

Figura 36

Estudio de número de neuronas en la capa oculta vs MSE vs tiempo de entrenamiento



Al observar los gráficos, se nota que para P=1, se cumple el comportamiento previsto, sin embargo, para las otras componentes principales, se observa mucha variación a lo predicho. Por ejemplo, en P=3 el MSE disminuye con 15 neuronas, pero vuelve a aumentar con 25, mientras que, el tiempo aumenta para las 15 y disminuye para las 25. Estas variaciones se deben

a que, en el entrenamiento, el número de épocas varían de acuerdo con el criterio de parada, siendo diferente para cada prueba desarrollada. Esto afecto tanto en el MSE como en el tiempo de entrenamiento. Idealmente, para este estudio el número de épocas debe ser constante en todas las pruebas.

Sin embargo, se llegó a una conclusión de este experimento, en todos los casos se cumple que al disminuir el MSE, el tiempo de entrenamiento aumenta, independientemente del número de neuronas.

Por lo tanto, se escoge el número de neuronas con base en la complejidad de la serie de datos. Mientras más compleja sea la serie de datos, más neuronas se incorporan en el modelo, pretendiendo un equilibrio entre el costo computacional y precisión. En la Tabla 11, se muestran los números de neuronas seleccionadas para cada NN.

Tabla 11

Número de neuronas en la capa oculta para cada NN

NN	Número de neuronas en la capa oculta
Primera componente principal (P=1)	7
Segunda componente principal (P=2)	7
Tercera componente principal (P=3)	15
Cuarta componente principal (P=4)	15
Quinta componente principal (P=5)	25

Como se observa, en P=1,2, se implementan 7 neuronas dado a que estas son las menos complejas. Posteriormente, se implementan 15 neuronas para mejorar la precisión (en P=3,4) y finalmente 25 para la serie más compleja (P=5). Cabe mencionar que, esto se aplicó para HUF y CEA.

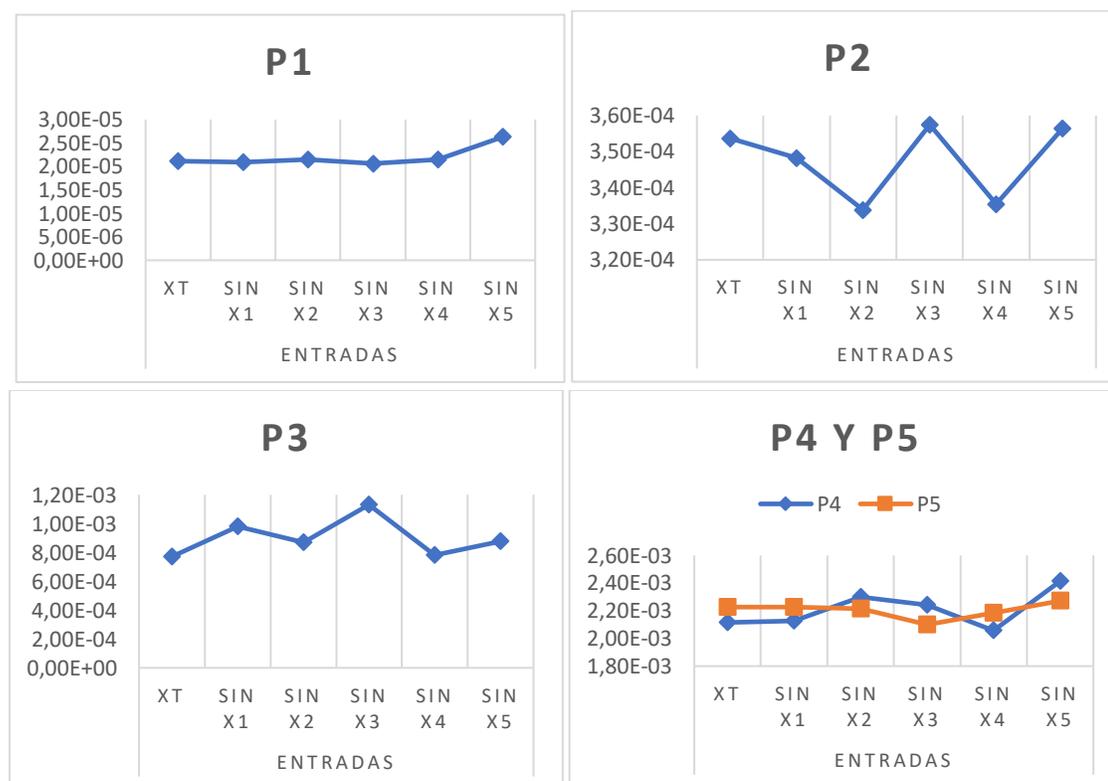
Estudio de Orden del Modelo PCANN

Se aplica LOOCV para evaluar las entradas del modelo se las NN y ver su importancia en la predicción. El proceso de entrenamiento se repitió 7 veces, obteniendo el MSE en cada

prueba y después se determinó el promedio correspondiente. Con fines de comparación, primero se realizó la prueba del modelo con todas las entradas, para posteriormente probarlo excluyendo una por una. Este proceso se llevó a cabo cinco veces (dado que son 5 entradas). En la Figura 37 se presenta el estudio mencionado para cada NN correspondiente a cada componente principal. El primer dato de los gráficos (XT) corresponde al modelo con todas sus entradas, mientras que, los datos de SIN X1 hasta SIN X5, hace referencia a la exclusión de cada una de las entradas, en el orden de la Tabla 8.

Figura 37

Experimento LOOCV para cada NN con sus componentes principales (P1, ..., P5) en HUF



El gráfico de la componente P1, evidencia que la mayor variación del MSE corresponde a la ausencia de la entrada X5 (promedio de carga anual). En general, todas las componentes presentan un aumento del MSE cuando se excluye esta entrada (en comparación con XT). Esto

sugiere que esta entrada, toma mucha importancia para el modelamiento. Por otro lado, en la componente P2, se observa que mediante la ausencia en las entrada X2 y X4 (número de semana al año y temperatura promedio) se obtiene un menor MSE. Esto indicaría que estas entradas no son muy importantes para P2. Sin embargo, el orden de este error es mínimo, por lo que excluir estas entradas no variaría mucho los resultados.

Se infiere que todas las entradas cumplen con cierta importancia para el entrenamiento del modelo. Su variación dependen de las series de datos de cada componente principal. En general, al mantener todas las entradas discutidas (XT) se obtendrá un modelo adecuado para la aplicación.

Cabe destacar que, dado que el modelo PCAFUZZY es una variación del modelo PCANN, el análisis LOOCV presentado también es válido para el mismo. Además, con motivos de generalización, se mantienen el mismo tipo de entradas para los datos de HUF y CEA.

Resultados de implementación NN

Para el presente análisis, por conveniencia, se identifica a cada NN por su salida, es decir, por su componente principal. En la Tabla 12 se presentan los resultados de entrenamiento para cada NN en HUF y CEA.

Tabla 12

Resultados de entrenamiento de NN para HUF y CEA

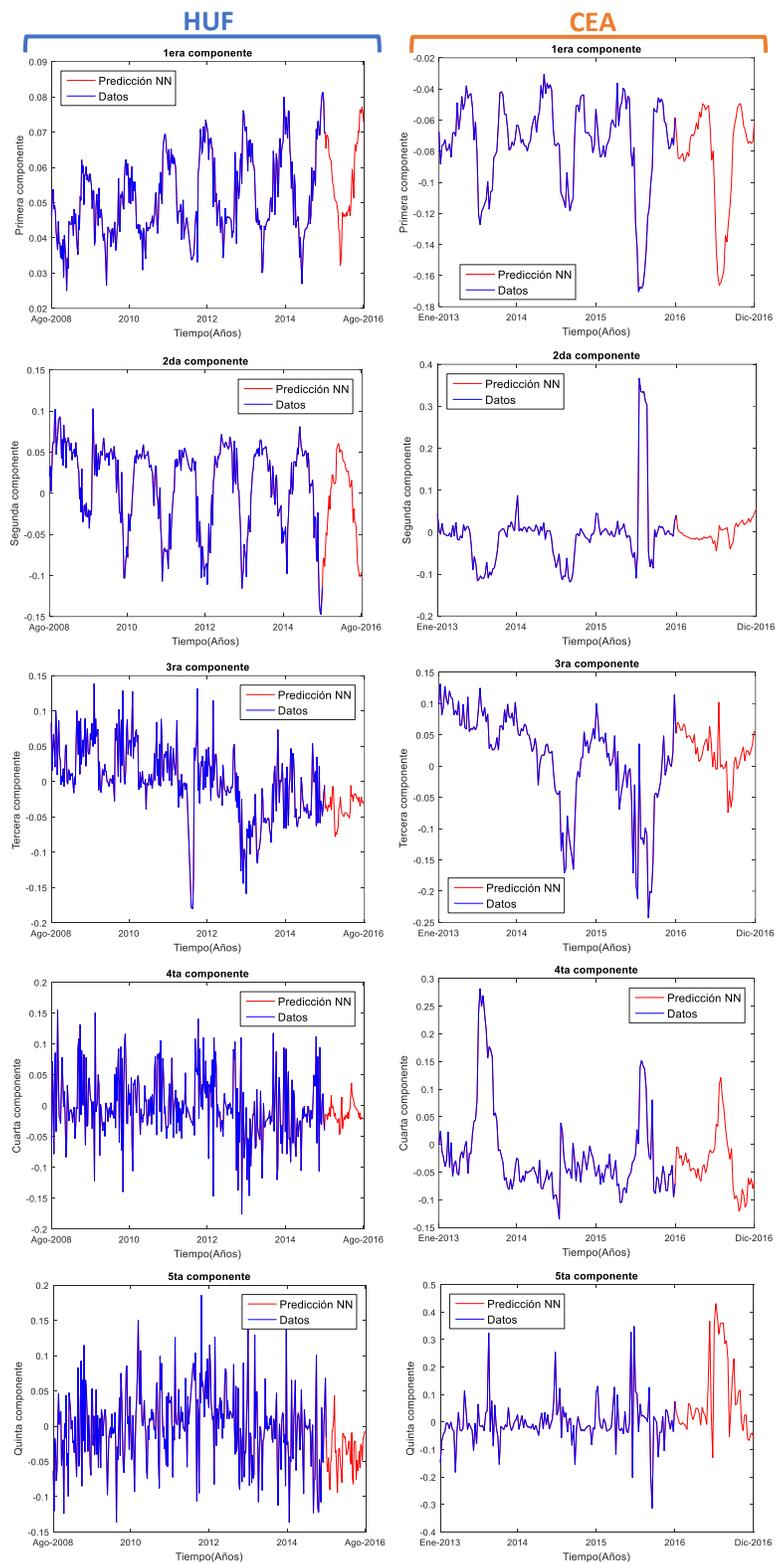
HUF				
NN	# neuronas	# épocas	R	MSE
P=1	7	6	0.918	2.19×10^{-5}
P=2	7	3	0.934	3.39×10^{-4}
P=3	15	3	0.733	1.30×10^{-3}
P=4	15	3	0.470	2.1×10^{-3}
P=5	25	5	0.531	2.00×10^{-3}
CEA				
NN	# neuronas	# épocas	R	MSE
P=1	7	9	0.977	3.68×10^{-5}
P=2	7	14	0.959	5.05×10^{-4}
P=3	15	5	0.957	5.32×10^{-4}
P=4	15	13	0.977	2.73×10^{-4}
P=5	25	1	0.186	2.00×10^{-3}

Se observa que, en las últimas NN, los valores de R son menores, representando una baja correlación entre los datos de prueba y los modelados. Esto se debe a que estas componentes son más difíciles de modelar, dado a que tienen un comportamiento poco predecible. En el Anexo C, se presentan las gráficas de R correspondientes para cada caso.

En la Figura 38, se ilustran las predicciones obtenidas para cada componente principal mediante los modelos de NN. La columna de gráficos izquierda corresponde a los datos de HUF y la derecha a los de CEA. También se debe considerar que los datos de color azul son de entrenamiento, mientras que los de color rojo son los predichos.

Figura 38

Predicción de componentes principales mediante PCANN para HUF y CEA



En estos gráficos, es más evidente lo mencionado anteriormente, donde las complejidades de los datos en las últimas componentes dificultan el modelamiento. En el caso de HUF, se observan que las primeras 3 componentes principales tienen un comportamiento periódico evidente, mientras que las últimas 2 aumentan su complejidad y no se observa un patrón tan notorio.

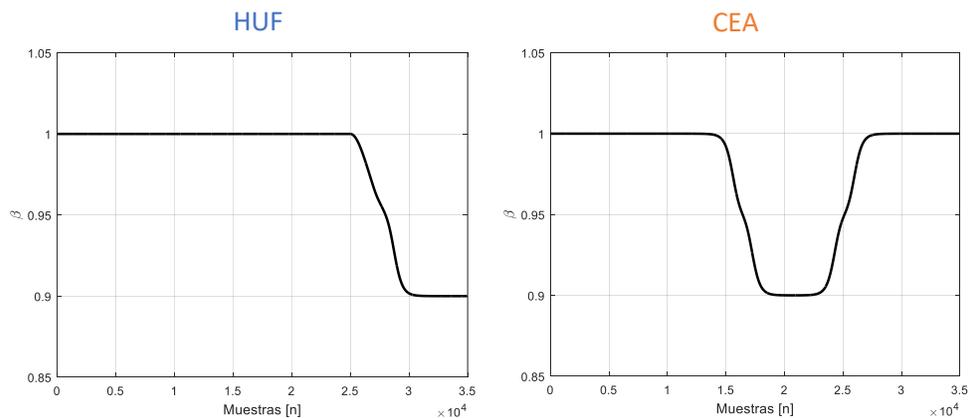
Por otro lado, para CEA, las primeras 4 componentes presentan la estacionalidad y la última exhibe la mayor complejidad. Además, también es evidente unos picos en el tercer año de las componentes: 1era, 2da y 3ra. Esto corresponde a un comportamiento de consumo de carga no habitual en esta localidad.

Resultados de diseño PCAFUZZY

Para el sistema difuso de PCAFUZZY, se analizó su curva de respuesta, la cual, se enseña en la Figura 39 con gráficas para HUF y CEA. Estas se diferencian por el hecho de que los datos anuales están organizados de diferente manera. Por lo tanto, en HUF, el periodo de verano se ubica al final, mientras que, en CEA, se ubica aproximadamente en la mitad.

Figura 39

Curva de control β vs muestras para HUF y CEA



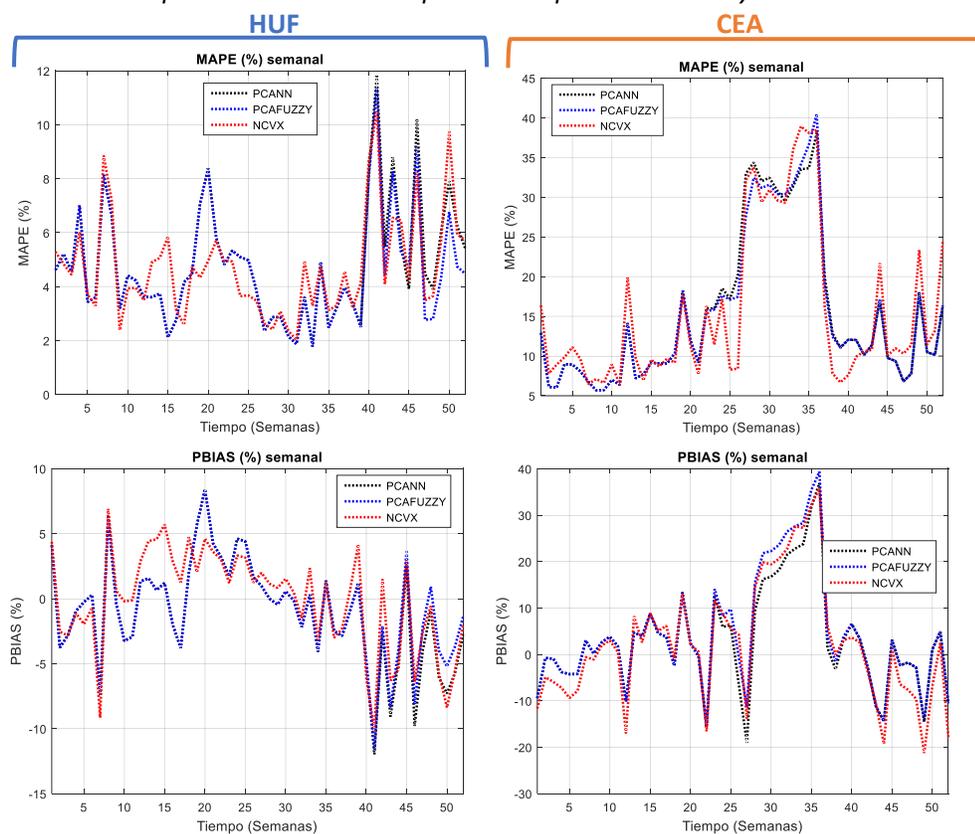
Se observa un comportamiento preciso para el factor β . Esta curva fue el resultado para el periodo de un año, dando la posibilidad de analizar los resultados en periodos de tiempo de STLF y MTLF.

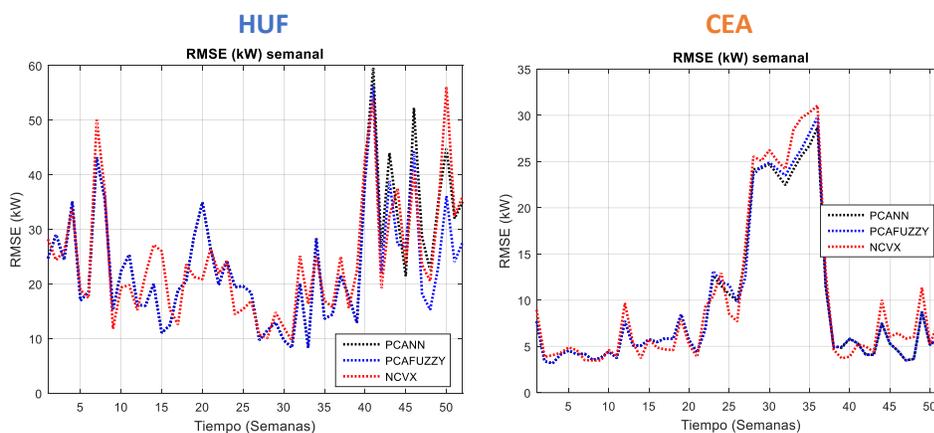
Análisis de Resultados de Predicción en STLF

En este análisis se escogió un horizonte de tiempo de una semana. Para generalizar el estudio, se determina la variación semanal de los parámetros de desempeño en el año de los datos de prueba, es decir, se calcula el MAPE, PBIAS y RMSE para las 52 semanas, denotándose a W como el número de semana del año. En la Figura 40, se muestra los gráficos correspondientes a lo mencionado, donde los del lado izquierdo son de HUF y los del derecho de CEA. Cada gráfico corresponde a MAPE, PBIAS y RMSE, respectivamente.

Figura 40

Parámetros de desempeño semanales en el periodo de prueba de HUF y CEA





A partir de estos gráficos, se reflejan algunos detalles observados. Primeramente, en aquellos correspondientes a HUF, se visualiza que para las primeras semanas los modelos PCANN y PCAFUZZY coinciden en sus valores. Esto tiene correspondencia con la descripción del método PCAFUZZY donde se detalla que $\beta = 1$ para estos datos del año. En CEA ocurre lo mismo, solo que en otros rangos de tiempo.

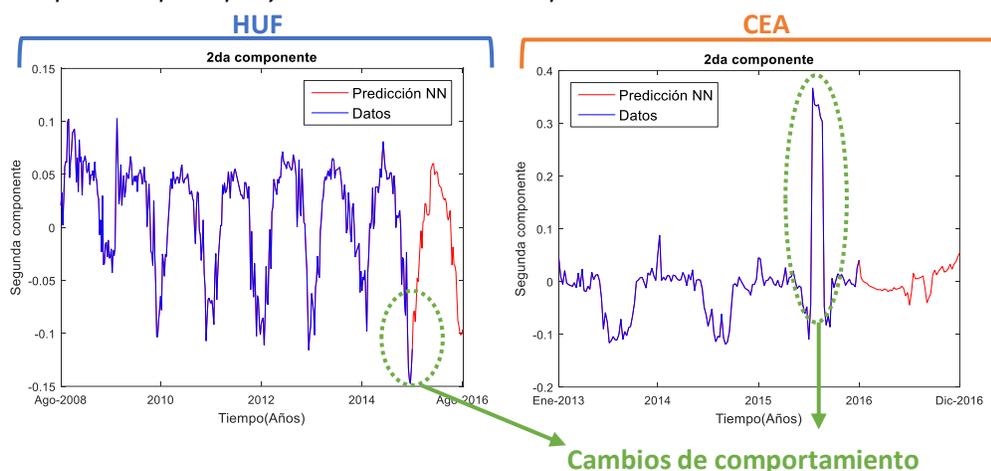
Además, en HUF a partir de la semana $W = 41$ se presentaron los mayores valores de RMSE y MAPE, así como altas variaciones de PBIAS. Este detalle coincide en los tres modelos. Estudiando la causa de esta particularidad, se determinó que al ser modelos creados a partir de PCA, los altos errores en la aproximación presentados por este método se acarrean a los modelos estudiados para el presente proyecto.

Cabe destacar que, en ambos conjuntos de datos se observa un periodo de tiempo donde las prestaciones tienen valores altos. En HUF este periodo se encuentra aproximadamente entre $W = 41, \dots, 52$ y en CEA entre $W = 25, \dots, 39$. Estos rangos corresponden al verano. Para este periodo de tiempo, se esperaría que el consumo de carga aumente, por el uso de artefactos como aire acondicionado. Sin embargo, el aumento de consumo no es tan alto como el método PCA lo estima, dando lugar a incrementos en los parámetros de desempeño. Para un análisis más detallado, en la Figura 41 se ilustra la 2da

componente de PCA y su predicción con PCANN, así como los cambios abruptos notados en el comportamiento. El gráfico del lado izquierdo es de HUF y el otro de CEA.

Figura 41

Segunda componente principal y sus cambios en el comportamiento.



Nota. Gráficos replicados de resultados de PCANN en Resultados de implementación NN.

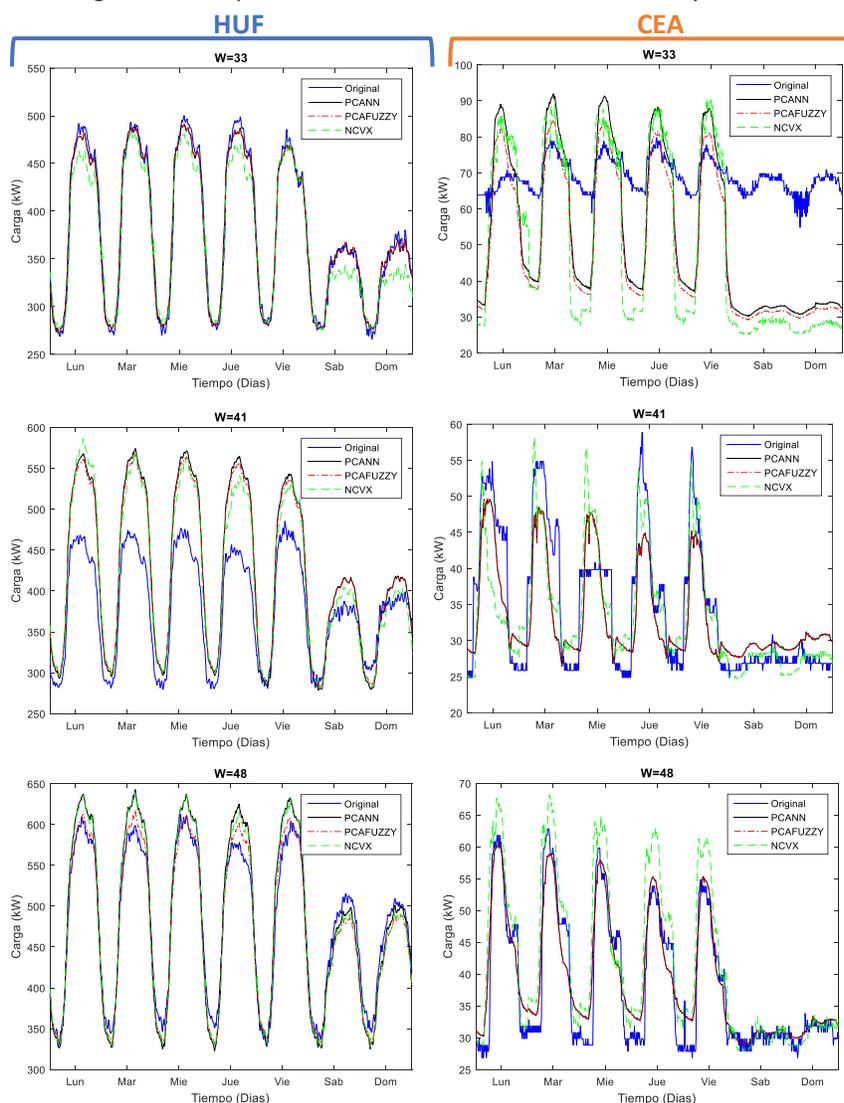
Para esta figura, en HUF, al finalizar los datos de color azul, se observa un pico inferior, el cual, ya no se replica por los datos de color rojo. Asimismo, en CEA ocurrió algo similar, se presentó un cambio súbito en el último año de entrenamiento que ya no se evidencia en los datos predichos. Esto cambios de comportamiento representan un consumo inusual que se genera a partir de los últimos datos de entrenamiento. Al ocurrir únicamente en esos casos, los modelos no son capaces de aprender lo suficiente de estos cambios y como consecuencia, presentan mayores errores para los periodos de tiempo mencionados.

Indagando en las posibles causas de los cambios de comportamiento. Las altas temperaturas presentadas en el verano de HUF como efecto del progresivo cambio climático presente mundialmente pudieron haber sido la causa de ese cambio. Por otro lado, el personal de CEA menciona la implementación de nuevas políticas de uso de aire acondicionado para ese año, causando aquel cambio súbito.

A continuación, se presenta un análisis más específico para STLF. Se escogieron 3 semanas individualmente dentro del periodo de prueba de HUF y CEA. Para HUF, se seleccionó una semana de verano ($W = 48$, 01/08/2016 – 07/08/2016), primavera ($W = 33$, 18/04/2016 – 24/04/2016) y otra correspondiente al mayor MAPE obtenido ($W = 41$, 13/06/2016 – 19/06/2016). Por otro lado, para CEA se presentan las semanas ($W = 48$, 14/11/2016 – 20/11/2016), la correspondiente a verano ($W = 33$, 01/08/2016 – 07/08/2016) y ($W = 41$, 26/09/2016 – 02/10/2016). La Figura 42, ilustra los gráficos de las semanas mencionadas.

Figura 42

Predicción de carga eléctrica para las semanas $W=33,41,48$ en HUF y CEA



Es evidente que, para los datos de HUF, las semanas $W = 33, 48$ tuvieron una buena aproximación a los datos originales. Mientras que, en la semana $W = 41$ se presentó el mayor error para los tres modelos, siendo esto causado por el cambio de comportamiento mencionado. Por otro lado, para CEA, las semanas $W = 41, 48$ fueron muy aceptables. Sin embargo, la semana $W = 33$, que pertenece a verano, presentó la menor fidelidad a los datos originales, por la razón expuesta anteriormente. Para complementar el análisis, se muestran los parámetros de desempeño en la Tabla 13 y Tabla 14.

Tabla 13

Métricas de desempeño en las semanas $W=33,41,48$ de HUF

Métrica	MAPE (%)			PBIAS (%)			RMSE (kW)			
	N° semana	33	41	48	33	41	48	33	41	48
Método										
NCVX		3.28	10.50	3.65	2.33	-10.18	-0.59	16.32	53.96	20.53
PCANN		1.78	11.81	3.91	0.31	-11.96	-0.95	8.38	59.62	22.51
PCAFUZZY		1.78	11.39	2.80	0.31	-11.62	0.93	8.38	56.30	15.29

Al observarlos, se constata lo dicho anteriormente para HUF. Destacando a $W = 41$ como la semana que presentó mayores valores de las prestaciones y donde se visualiza una sobreestimación considerable (ver PBIAS) en comparación de las otras semanas.

Tabla 14

Métricas de desempeño en las semanas $W=33,41,48$ de CEA

Métrica	MAPE (%)			PBIAS (%)			RMSE (kW)			
	N° semana	33	41	48	33	41	48	33	41	48
Método										
NCVX		36.01	9.92	11.37	27.57	2.59	-9.60	28.43	5.21	6.03
PCANN		31.37	12.01	7.79	22.78	3.38	-2.83	24.16	5.33	3.65

N° semana	MAPE (%)			PBIAS (%)			RMSE (kW)		
	33	41	48	33	41	48	33	41	48
Método									
PCAFUZZY	31.65	12.01	7.79	27.61	3.40	-2.83	25.00	5.33	3.65

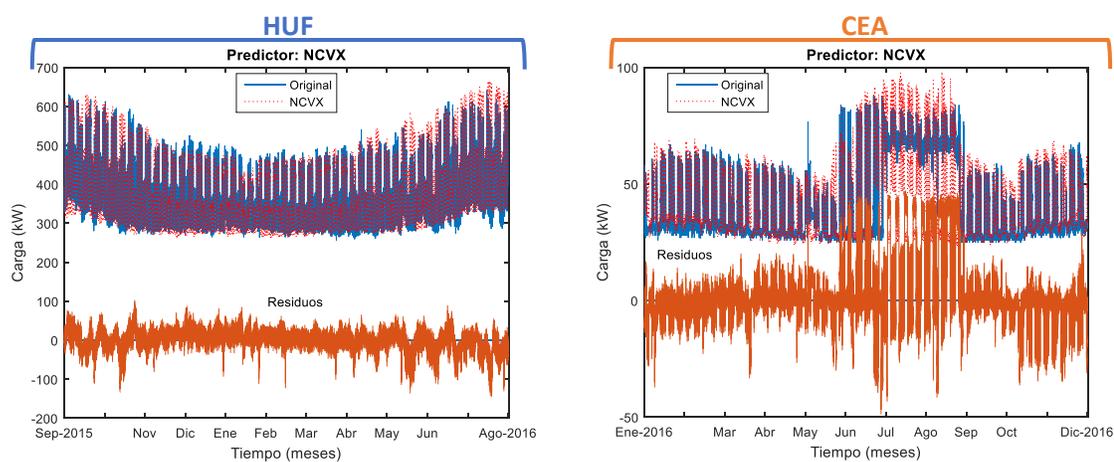
Por otro lado, los parámetros correspondientes a CEA (ver Tabla 14), también tienen concordancia con lo dicho anteriormente. Es evidente notar que, estos valores son más altos que los obtenidos para HUF. Esto se debe a varias razones, como, por ejemplo, el menor tiempo de entrenamiento usado para CEA (3 años) versus el usado en HUF (7 años) que afecta en la precisión del modelo.

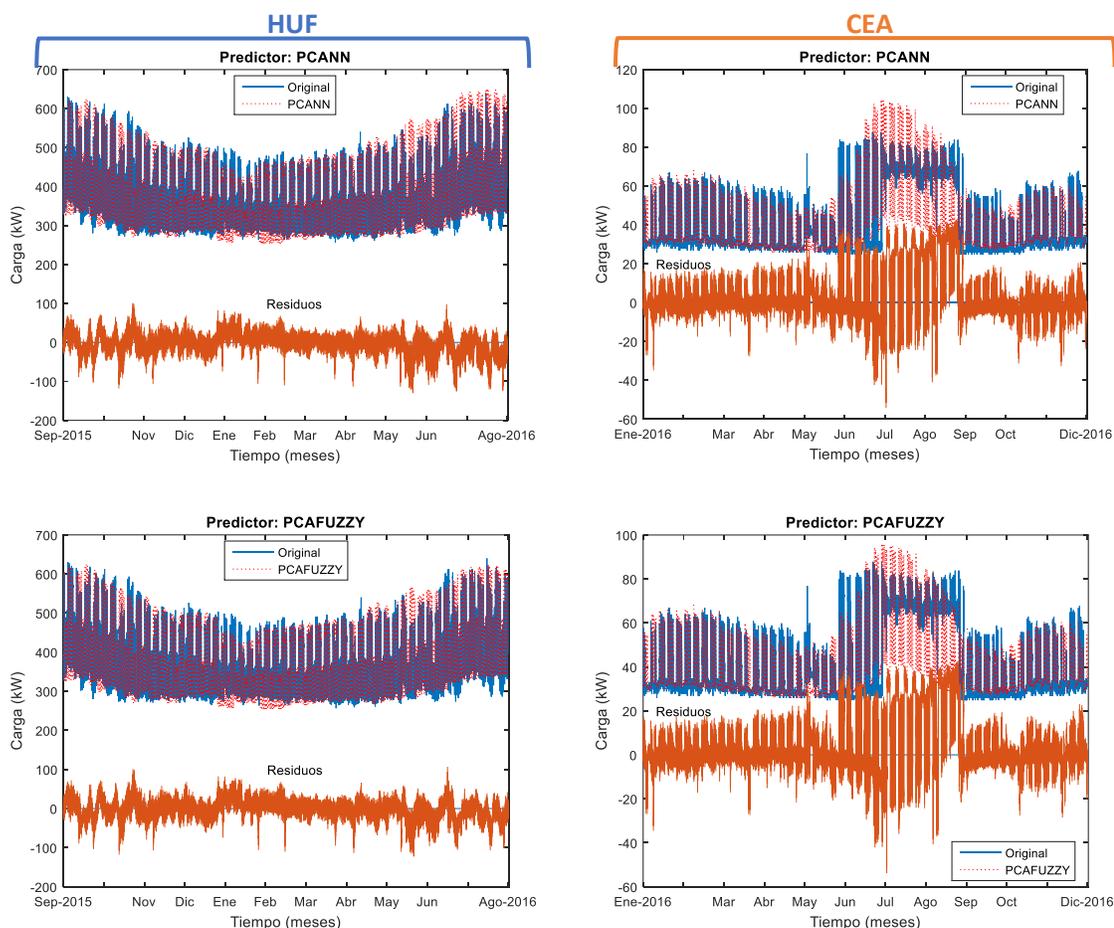
Análisis de Resultados de Predicción: MTLF

Se analizarán los resultados en el horizonte de tiempo MTLF, en el periodo de prueba de ambos conjuntos de datos. Es decir, un año para HUF y uno para CEA, cada uno con su respectivo orden de meses. En la Figura 43, se observa la comparación entre los datos de carga originales y los predichos por los modelos NCVX, PCANN y PCAFUZZY. Los gráficos de la izquierda corresponden a HUF, mientras que los de la derecha a CEA.

Figura 43

Gráficos de predicción de carga de modelos híbridos para MTLF en HUF y CEA





Se puede observar que los modelos siguen el patrón del comportamiento de carga, haciendo referencia a su correcta implementación. Cabe destacar que, se evidencia de mejor manera que, tanto en HUF como en CEA, los periodos de verano (aproximadamente entre junio y septiembre) son los que presentan mayores variaciones en la predicción. Además, se puede visualizar una hipotética mejora por parte de PCAFUZZY con respecto a PCANN, en este periodo, donde se observa como la sobreestimación disminuye.

Con el fin de demostrar la efectividad de los modelos implementados, se procedió a comparar los parámetros de desempeño obtenidos en la presente investigación, con los calculados por Gordillo-Orquera et al. (2018). En la Tabla 15, se presentan los parámetros obtenidos para HUF y CEA.

Tabla 15*Parámetros de desempeño en MTLF para HUF y CEA*

Métrica	HUF			CEA		
	MAPE (%)	PBIAS (%)	RMSE (kW)	MAPE (%)	PBIAS (%)	RMSE (kW)
Método						
PCA	5.06	-2.87	28.91	16.21	7.58	13.33
OPLS	5.96	1.38	29.53	15.14	5.81	13.19
ENS	4.83	-0.74	26.77	14.63	6.69	12.99
NCVX	4.73	-0.26	26.59	15.67	5.33	13.13
PCANN	4.83	-1.07	26.79	15.40	5.55	12.08
PCAFUZZY	4.62	-0.72	24.91	15.25	7.39	12.29

Nota. Los algoritmos originales se encuentran propuestos por Gordillo-Orquera et al. (2018) y son PCA, OPLS y ENS.

Se comparará el modelo híbrido NCVX con PCA, OPLS y ENS, teniendo en cuenta que es una mejora de este último. Además, se enfatiza en comparar PCANN y PCAFUZZY con PCA, por la misma razón.

En general, para HUF, todas las métricas mejoraron con respecto a los modelos individuales (PCA y OPLS), presentando un MAPE inferior al 5% y siendo comparables con los resultados de ENS. Mientras que, el PBIAS, demostró que existe una sobreestimación baja para los tres casos. Además, el parámetro RMSE también resultó disminuir para los tres modelos. Por otra parte, para CEA la situación fue ligeramente distinta. Considerando el MAPE, se presentaron resultados comparables. Mientras que, el PBIAS disminuyó para la mayoría de las comparaciones. Prosiguiendo el análisis, se explicará en detalle primero los resultados de NCVX y después los de PCANN y PCAFUZZY, considerando ambos conjuntos de datos para cada caso.

Resultados de predicción de NCVX

El modelo NCVX en HUF, presenta una disminución del MAPE de 0.33%, 1.23% y 0.1%, con respecto a PCA, OPLS y ENS, correspondientemente. En el mismo contexto, el PBIAS se redujo en 2.61%, 1.64% y 0.48%, con respecto a los originales. EL RMSE también resulto ser menor en todos los casos. Cabe mencionar que, retomando el análisis de la sección de Análisis de Modelo ENS: Combinación Convexa, el MAPE óptimo posible con ENS era de $\text{MAPE}(\%) = 4.75\%$ (con un $\alpha = 0.67$), el cual, es superior al MAPE obtenido con NCVX ($\text{MAPE}(\%) = 4.73\%$). En definitiva, el modelo NCVX logró mejorar el desempeño de la predicción en HUF, reemplazando el parámetro constante α por uno obtenido a partir de un sistema difuso.

Para el caso de los datos de CEA, el modelo NCVX presentó que el MAPE disminuyó 0.54% con respecto a PCA, pero aumentó 0.53% y 1.04% con respecto a OPLS y ENS. Sin embargo, el PBIAS fue el menor obtenido en contraste con los otros modelos, presentando una subestimación reducida. Mientras que el RMSE tuvo valores comparables con los modelos originales. En resumen, el indicador MAPE sugiere que, para este caso ENS es más conveniente que NCVX, pero el PBIAS indica lo contrario. Para resolver este dilema, nos basamos en la iniciativa de NCVX, la cual, consiste en abordar la problemática de subestimación y sobreestimación de datos. Por lo tanto, se puede deducir que, en efecto, está cumpliendo su objetivo al presentar valores de PBIAS menores que ENS.

Resultados de predicción de PCANN y PCAFUZZY

Los modelos PCANN y PCAFUZZY, con los datos de HUF, presentaron una reducción del MAPE de 0.23% y 0.44%, con respecto a PCA. Además, el PBIAS también mejoró en ambos casos, reduciéndose 1.8% y 2.15%, en el mismo contexto. El RMSE, presentó un comportamiento similar. Por tal motivo, se puede inferir que estos métodos mejoraron las prestaciones de predicción con respecto a PCA. Donde, PCAFUZZY presentó mejores resultados que PCANN.

Para CEA, estos modelos también presentaron mejoras. El MAPE se redujo 0.81% y 0.96%, de PCANN y PCAFUZZY con respecto a PCA, mientras que el PBIAS disminuyó 2.03% y 0.19%, en el mismo orden. El RMSE también se redujo para ambos casos. Cabe mencionar que, para estos datos, PCANN presentó una mayor reducción de PBIAS con respecto a PCAFUZZY. A pesar de que en PCAFUZZY presenta un menor MAPE, la disminución del PBIAS es más considerable. Por lo tanto, en este conjunto de datos, los resultados de PCANN son más convenientes.

En definitiva, para los dos conjuntos de datos, los modelos implementados por PCA y NN, así como los que incluyen un sistema difuso, dieron como resultado una mejora en los parámetros de desempeño, en comparación con los de PCA y AR, cumpliendo los objetivos del proyecto.

Capítulo V

Conclusiones y Recomendaciones

En este capítulo se presentan las conclusiones obtenidas del desarrollo del este proyecto. Además, se mencionan algunas recomendaciones que motivan y orientan a futuros trabajos de investigación.

Conclusiones

Se desarrollaron tres modelos híbridos basados en datos NCVX, PCANN y PCAFUZZY, que integraron técnicas de IC y modelos base como PCA y OPLS. Sus resultados fueron comparados con sus respectivos modelos base y en casi su totalidad mostraron mejoras para ambos conjuntos de datos, es decir, HUF y CEA. Tanto en STLF como en MTLF, los resultados fueron satisfactorios. Por lo tanto, se dedujo que los modelos híbridos presentaron prestaciones de desempeño favorables, permitiendo obtener una predicción más precisa que sus respectivos modelos base. Cabe destacar, que la implementación de los modelos híbridos implica un aumento en la complejidad computacional.

En particular, el diseño del modelo NCVX se hizo a través de un sistema difuso. Este realizó la combinación no convexa entre PCA y OPLS, permitiendo abordar el problema de subestimación y sobreestimación presentado por estos modelos individualmente. Los resultados demostraron una disminución del PBIAS en HUF y CEA. De este modo, se comprueba que el modelo permitió reducir este parámetro y, por ende, el problema mencionado.

Por otra parte, se desarrolló el modelo PCANN que combinó los resultados de PCA con técnicas de modelamiento por NN. Además, se consideró a una variable exógena de temperatura como entrada de predicción. Asimismo, se implementó el modelo PCAFUZZY que añadió una variable representativa a las estaciones del año, al modelo anterior. Ambas implementaciones mostraron mejoras en las prestaciones de desempeño en comparación con

el método PCA original. En relación con lo expuesto, se deduce que la integración de técnicas de IC al método PCA permitió obtener mejores resultados en la predicción.

Específicamente, los modelos se entrenaron y evaluaron con los registros históricos disponibles de HUF y CEA. Los resultados se analizaron en semanas para STLF y un año para MTLF. En STLF, los indicadores de error por semana variaban en relación con periodos de tiempo en el año, registrando mayores valores en verano. Mientras que, para MTLF los indicadores y las gráficas demostraron que el modelo se ajustaba correctamente de los datos. Se determinó que la variación en verano mencionada se debió a la complejidad de los datos en el entrenamiento. Sin embargo, en general se presentaron resultados muy satisfactorios por parte de los modelos en ambos periodos. En definitiva, los modelos exhibieron resultados aceptables para STLF y MTLF con la consideración de un aumento de las prestaciones para los periodos de verano.

Finalmente, cabe destacar que los modelos fueron evaluados con los indicadores MAPE, PBIAS y RMSE. El MAPE y RMSE disminuyeron para la mayoría de los casos, mientras que, el PBIAS disminuyó para todos. Con base en esto, se concluyó que en general los modelos híbridos presentaron mejoras en las prestaciones de desempeño, dando énfasis en la ventajosa reducción del PBIAS.

Recomendaciones

Dado a la complejidad de datos que se manifestó en verano, se sugiere analizar los comportamientos no habituales de consumo presentados por artefactos como aire acondicionado. Mediante un registro histórico de activaciones o datos que midan su consumo, se podría estudiar cómo su actividad afecta en la predicción.

Analizar los modelos utilizando una mayor cantidad de datos disponibles de los lugares HUF y CEA, con el fin de mejorar el entrenamiento y aprendizaje sobre nuevos cambios de

comportamiento de los consumidores. Consecuentemente, se podría conseguir mejores prestaciones de desempeño en la predicción.

Con fines de comprobar más la versatilidad de los modelos, se aconseja experimentar los presentes métodos con diferentes datos de lugares similares a los presentados. De esta manera, se podrá indagar en estrategias para mejorar la generalización de los algoritmos.

Referencias Bibliográficas

- Acosta, H. (01 de Octubre de 2006). Tesis Doctoral. *Diseño de controladores dedicados a la lógica difusa*. Madrid, Madrid, España: Universidad Autónoma de Madrid.
- AEMet. (s.f. de s.f. de s.f.). *Valores climatológicos normales. Getafe*. Obtenido de AEMet: <https://www.aemet.es/es/serviciosclimaticos/datosclimatologicos/valoresclimatologicos?l=3200&k=28>
- Ammar, N., Sulaiman, M., Fateh, A., & Mohamad, A. (2018). Long-term load forecasting of power systems using artificial neural network and ANFIS. *ARPN Journal of Engineering and Applied Sciences*, 828-834.
- Anwar, T., Sharma, B., Chakraborty, K., & Sirohia, H. (2018). Introduction to Load Forecasting. *International Journal of Pure and Applied Mathematics*, 1527-1537.
- Bishop, C. (2006). *Pattern recognition and machine learning*. New York: Springer.
- Bourdeau, M., Zhai, X., Nefzaoui, E., Guo, X., & Chatellier, P. (2019). Modeling and forecasting building energy consumption: A review of datadriven techniques. *Sustainable Cities and Society*, 1-27. doi:<https://doi.org/10.1016/j.scs.2019.101533>
- Brockwell, P., & Davis, R. (2016). *Introduction to Time Series and Forecasting*. AG Switzerland: Springer.
- Brunton, S., & Kutz, J. (2019). *Data-Driven Science and Engineering: Machine Learning, Dynamical Systems, and Control*. Cambridge: Cambridge University Press.
- Bunn, D., & Farmer, E. (1985). *Comparative models for electrical load forecasting*. New York: John Wiley and Sons.
- C. Aggarwal, C. (2018). *Neural Networks and Deep Learning*. Cham: Springer Nature Switzerland AG.

- Cai, M., Pipattanasomporn, M., & Rahman, S. (2019). Day-ahead building-level load forecasts using deep learning vs. traditional time-series techniques. *Applied Energy*, 1078-1088.
- CENACE. (2019). *Informe anual 2018*. Quito: CENACE.
- Cheng, Y., Hong, W., Sheng, W., & Huang, N. (2016). Electric load forecasting based on a least squares support vector machine with fuzzy time series and global harmony search algorithm. *Energies*, 1-13.
- De Felice, M., & Yao, X. (2011). Short-Term Load Forecasting with Neural Network Ensembles: A Comparative Study. *IEEE Computational Intelligence Magazine*, 47-56.
- Demuth, H., & Beale, M. (2004). *Neural Network Toolbox User's Guide*. Natick, MA: The MathWorks, Inc.
- Diciembre Sanahuja, S. (2 de Octubre de 2017). Sistemas de Control con Lógica Difusa: Métodos de Mamdani y de Takagi-Sugeno-Kang (TSK). *Estancia en Practicas y Proyecto Final de Grado*. Castellon de la Plana, Comunidad Valenciana, España: Universitat Jaume I.
- Dinov, I. (2018). *Data Science and Predictive Analytics*. Michigan: Springer.
- Dong, J.-r., Zheng, C.-y., Kan, G.-y., Zhao, M., Wen, J., & Yu, J. (2015). Applying the ensemble artificial neural network-based hybrid data-driven model to daily total load forecasting. *Neural Computing and Applications*, 603-611. doi:10.1007/s00521-014-1727-5

- Dong, Y., Zhang, Z., & Hong, W.-C. (2018). A Hybrid Seasonal Mechanism with a Chaotic Cuckoo Search Algorithm with a Support Vector Regression Model for Electric Load Forecasting. *Energies*, 1009.
- EL-Shimy, M. (2018). *Fundamentals of Energy Management and Energy Managers - A technical report and a short course*. Egypt: Ain Shams University.
- Ertel, W. (2017). *Introduction to Artificial Intelligence*. London: Springer. doi:10.1007/978-3-319-58487-4
- Feilat, E., Al-Sha'abi, D., & Momani, M. (2017). Long-Term Load Forecasting Using Neural Network Approach for Jordan's Power System. *Engineering Press*, 43-50. doi:10.28964/EngPress-1-108
- Feinberg, E., & Genethliou, D. (2005). Load Forecasting. En J. Chow, F. Wu, & J. Momoh, *Applied Mathematics for Restructured Electric Power Systems* (págs. 269-285). Boston, MA: Springer.
- Ganguly, P., Kalam, A., & Zayegh, A. (2017). Short Term Load Forecasting using Fuzzy Logic. *International Conference on Research in Education and Science (ICRES)* (págs. 355-361). Ephesus-Kusadas: ICRES.
- Geisser, S. (1993). *Predictive Inference: An Introduction*. London: Chapman and Hal.
- Gordillo-Orquera, R., Lopez-Ramos, L. M., Muñoz-Romero, S., Marques, A., Rojo-Álvarez, J., Iglesias-Casarrubios, P., & Arcos-Avilés, D. (2018). Analyzing and Forecasting Electrical Load Consumption in Healthcare Buildings. *Energies*, 493.
- Guamán, G. (1 de Julio de 2019). Trabajo de Titulación previo a la Obtención del Título de Ingeniero Eléctrico. *Mejora del Pronóstico de Carga a Corto Plazo en el Sistema de Distribución de la Empresa Eléctrica Ambato Región al Centro Norte*

- S.A. (EEASA) *Basada en Criterios de BIG DATA*. Quito, Pichincha, Ecuador: Escuela Politécnica Nacional.
- Guo, C., & Wong, W. (2013). Fundamentals of artificial intelligence techniques for apparel management applications. En C. X. Guo, W. K. Wong, & S. Leung, *Optimizing Decision Making in the Apparel Supply Chain Using Artificial Intelligence (AI)* (págs. 13-40). N.A.: Woodhead. doi:<https://doi.org/10.1533/9780857097842.13>
- Hong, W.-C., Li, M.-W., & Fan, G.-F. (2019). *Short-Term Load Forecasting by Artificial Intelligent Technologies*. Basel: MDPI.
- Hyndman, R., & Koehler, A. (2006). Another look at measures of forecast accuracy. *International Journal of Forecasting* 22, 679 – 688. doi:10.1016/j.ijforecast.2006.03.001
- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2017). *An Introduction to Statistical Learning*. New York: Springer.
- Jassbi, J., Alavi, S., Serra, P., & Ribeiro, R. (2007). Transformation of a Mamdani FIS to First Order Sugeno FIS. *IEEE International Conference on Fuzzy Systems, Imperial College* (págs. 23-26). London: IEEE. doi:10.1109/FUZZY.2007.4295331
- Jiménez, J., Donado, K., & Quintero, C. (2017). A Methodology for Short-Term Load Forecasting. *IEEE LATIN AMERICA TRANSACTIONS*, 400-407. doi:10.1109/TLA.2017.7867168
- Jolliffe, I., & Cadima, J. (2016). Principal component analysis: A review and recent developments. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, n. p.

- Kamal, N. A., & Ibrahim, A. M. (2018). Conventional, Intelligent, and Fractional-Order Control Method for Maximum Power Point Tracking of a Photovoltaic System: A Review. En A. T. Azar, A. G. Radwan, & S. Vaidyanathan, *Fractional Order Systems: Optimization, Control, Circuit Realizations and Applications* (págs. 603-671). Cairo: Elsevier Inc. doi:<https://doi.org/10.1016/B978-0-12-816152-4.00020-0>
- Keller, J., Liu, D., & Fogel, D. (2016). *Fundamentals of computational intelligence*. Canada: The Institute of Electrical and Electronics Engineers, Inc.
- Khan, A., Mahmood, A., Safdar, A., Khan, Z., & Khan, N. (2016). Load Forecasting and Dynamic Pricing based Energy Management in Smart Grid-A Review. *Renewable and Sustainable Energy Reviews*, 1311-1322.
- Khettab, K., Bensafia, Y., Bourouba, B., & Azar, A. T. (2018). Enhanced Fractional Order Indirect Fuzzy Adaptive Synchronization of Uncertain Fractional Chaotic Systems Based on the Variable Structure Control: Robust H_∞ Design Approach. En A. T. Azar, A. G. Radwan, & S. Vaidyanathan, *Mathematical Techniques of Fractional Order Systems* (págs. 597-624). Cairo: Elsevier. doi:<https://doi.org/10.1016/B978-0-12-813592-1.00020-9>
- Kuhn, M., & Johnson, K. (2016). *Applied Predictive Modeling*. New York: Springer.
- Li, M.-W., Geng, J., Hong, W.-C., & Zhang, Y. (2018). Hybridizing Chaotic and Quantum Mechanisms and Fruit Fly Optimization Algorithm with Least Squares Support Vector Regression Model in Electric Load Forecasting,. *Energies*, 2226.
- Llumiquinga, L. (27 de 11 de 2018). Trabajo de titulación previo a la obtención del título de Ingeniero en Electrónica, Automatización y Control. *Sistema de monitorización*

- de consumo eléctrico no invasivo de los equipos del laboratorio de circuitos electrónicos para el desarrollo de un algoritmo de predicción de carga a corto plazo.* Sangolquí, Pichincha, Ecuador: Universidad de las Fuerzas Armadas ESPE.
- López, M., Sans, C., Valero, S., & Senabre, C. (2018). Empirical Comparison of Neural Network and Auto-Regressive Models in Short-Term Load Forecasting. *Short-Term Load Forecasting. Energies* 2018, 139-157.
- Lv, C., Xing, Y., Zhang, J., Na, X., Li, Y., Liu, T., . . . Wang, F.-Y. (2018). Levenberg–Marquardt Backpropagation Training of Multilayer Neural Networks for State Estimation of a Safety-Critical Cyber-Physical System. *IEEE Transactions on Industrial Informatics*, 3436 - 3446. doi:10.1109/TII.2017.2777460
- Mahapatra, K., Alm, R., Hallgren, R., Bischoff, L., Tuglu, N., Kuai, L., . . . Umoru, I. (2018). A behavioral change-based approach to energy efficiency in a manufacturing plant. *Energy Efficiency*, 1103–1116.
- Martí, P. (1 de Febrero de 2009). Tesis Doctoral. *Aplicación de redes neuronales artificiales para predicción de variables en ingeniería del riego: evapotranspiración de referencia y pérdidas de carga localizadas en emisores integrados.* Valencia, Valencia, España: Universidad Politécnica de Valencia.
- Mohan, N., Soman, K., & Kumar, S. S. (2018). A data-driven strategy for short-term electric load forecasting using dynamic. *Applied Energy*, 229-244. doi:https://doi.org/10.1016/j.apenergy.2018.09.190
- Muñoz-Romero, S. (18 de 05 de 2015). *Análisis Multivariante: Soluciones eficientes e interpretables.* Madrid: Universidad Carlos III de Madrid.

- Muñoz-Romero, S., Arenas-García, J., & Gomez-Verdejo, V. (2013). Iterative orthonormalized partial least squares with sparsity constraints. *2013 IEEE International Conference on Acoustics, Speech and Signal Processing* (págs. 3387-3391). Vancouver: IEEE. doi:10.1109/ICASSP.2013.6638286
- Myttenaere, A., Golden, B., Le Grand, B., & Rossi, F. (2016). Mean Absolute Percentage Error for regression models. *Neurocomputing*, 38-48. doi:10.1016/j.neucom.2015.12.114
- Reunanen, J. (2003). Overfitting in Making Comparisons Between Variable Selection Methods. *Journal of Machine Learning Research* 3, 1371-1382.
- Salas, J. S. (2016). *Trabajo Fin de Master*. Valladolid: Universidad de Valladolid.
- Samarasinghe, S. (2007). *Neural Networks for Applied Sciences and Engineering*. New York: Taylor and Francis Group .
- Schulze, M., Nehler, H., Ottosson , M., & Thollander , P. (2015). Energy management in industry – a systematic review of previous findings and an integrative conceptual framework. *Journal of Cleaner Production*, 3692-3708.
- Wang, Z., & Srinivasan, R. S. (2017). A review of artificial intelligence based building energy use prediction: Contrasting the capabilities of single and ensemble prediction models. *Renewable and Sustainable Energy Reviews*, 796-808. doi:https://doi.org/10.1016/j.rser.2016.10.079
- Ying , X. (2019). An Overview of Overfitting and its Solutions. *Journal of Physics: Conference Series*, 1-6. doi:10.1088/1742-6596/1168/2/022022

- Zhang, D., Shah, N., & G. Papageorgiou, L. (2013). Efficient energy consumption and operation management in a smart building with microgrid. *Energy Conversion and Management*, 209-222. doi:<https://doi.org/10.1016/j.enconman.2013.04.038>
- Zhang, S., & Wang, D. (2009). Medium and Long-Term Load Forecasting Based on PCA and BP Neural Network Method. *2009 International Conference on Energy and Environment Technology* (págs. 389-391). Guilin: IEEE. doi:10.1109/ICEET.2009.559
- Zhang, Z., & Hong, W. (2019). Electric load forecasting by complete ensemble empirical mode decomposition adaptive noise and support vector regression with quantum-based dragonfly algorithm. *Nonlinear Dynamics*, 1107-1136. doi:10.1007/s11071-019-05252-7

Anexos

Anexo A. AUTOVECTORES SEMANALES DE COMPONENTES PRINCIPALES DE HUF Y CEA

Anexo B. TENDENCIA Y PROMEDIO DE COMPONENTES PRINCIPALES PARA HUF Y CEA

Anexo C. GRÁFICOS DE R PARA CADA SUBMODELO EN HUF Y CEA

Anexo D. EXPERIMENTOS

Anexo E. CÓDIGOS MATLAB